A KNOWLEDGE ACQUISITION SYSTEM FOR PRICE CHANGE RULES

A Thesis

Submitted to the Faculty of Graduate Studies and Research

In Partial Fulfilment of the Requirements

For the Degree of

Master of Science

In

Computer Science

University of Regina

By

Moslema Jahan

Regina, Saskatchewan

January, 2016

UNIVERSITY OF REGINA

FACULTY OF GRADUATE STUDIES AND RESEARCH

SUPERVISORY AND EXAMINING COMMITTEE

Moslema Jahan, candidate for the degree of Master of Science in Computer Science**,** has presented a thesis titled, ***A Knowledge Acquisition System For Price Change Rules,*** in an oral examination held on December 2, 2015. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner:	*Dr. Vlado Keselj, Dalhousie University
Supervisor:	Dr. Howard J. Hamilton, Department of Computer Science
Committee Member:	Dr. Robert Hilderman, Department of Computer Science
Committee Member:	Dr. Malek Mouhoub, Department of Computer Science
Chair of Defense:	Dr. Andrew Cameron, Department of Biology

*Via SKYPE

ABSTRACT

Knowledge acquisition is the process of extracting and organizing knowledge from one source and storing it in some other location such as a knowledge base. Our research developed a new approach to knowledge acquisition concerning motor fuel pricing and implemented it in the Knowledge Acquisition System for Price ChangE Rules (KASPER) software system. Store managers want to understand the pricing strategies at competing stores or brands. The main goal of our research is to provide decision rules with high predictive accuracy on unseen data that may explain why a store or brand made a price change in a specific category. These decision rules should relate prices at one store to those at other stores or brands in the same city.

Our approach is able to generate directional and categorical price change rules. The approach can use brand-based or distance-based store-to-store relations or use brand-to-brand relations. KASPER was applied to data from four cities to generate decision rules from these relations. We tested the decision rules on unseen data and found that most decision rules had high predictive accuracy in cases where the price changes tend to fluctuate more. Our approach was more effective in the two cities where price changes of varied sizes occur than in the two cities where price changes are of consistent, small sizes. We found that high variability of price changes allows the system to match corresponding behaviours more effectively.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Dr. Howard J. Hamilton for his excellent guidance and patience, and for providing me with an excellent atmosphere for doing research. I also would like to thank the University of Regina Faculty of Graduate Studies and Research, and the Natural Science and Engineering Research Council of Canada, via a Discovery Grant and a Collaborative Research and Development administrated by Dr. Hamilton, for financial support that helped me satisfy my daily basic needs and stay focused on my research. I would like to thank Dr. Hilderman for his valuable comments on my research and especially his suggestions of related research. I would also like to thank Dr. Malek Mouhoub for serving on my supervisory committee. He was more than generous with his expertise and precious time.

I would like to thank one of my favourite colleagues, Mondelle Simeon, who supported me with effective discussions whenever I faced problems in my research. One of my experienced colleagues, Dr. Richard Dosselmann, also aided me with the LaTeX formatting language.

POST DEFENCE ACKNOWLEDGEMENT

I would like to gratefully acknowledge my external examinar, Dr. Vlado Keselj, for his valuable comments and special attention to my thesis. Thanks to Dr. Andrew Cameron for chairing my thesis defence.

DEDICATION

I would like to dedicate this thesis to my father Moklesur Rahman, M.D., and my mother Ruma Begum. I would like to thank them for their love and support throughout my life. My younger brothers, relatives, and friends deserve my wholehearted thanks as well. They have always given me personal support while I was completing my degree. Finally, thanks to my husband for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Algorithms

List of Acronyms

AAA American Automobile Association

AC accuracy

CAT category

DB database

DC double-component

DIR direction

DOR diagnostic odds ratio

DRS decision rule selector

E error rate

EODP end of day price

EODPC end of day price change

F F-measure

FDR false discovery rate

FN false negative

FNR false negative rate

FP false positive

FPR false positive rate

G G-mean

KASPER Knowledge Acquisition System for Price ChangE Rules

KB	knowledge base
LB	lower bound
LR+	positive likelihood ratio
LR-	negative likelihood ratio
Max	maximum
MED	median
Min	minimum
M-TAR	Momentum-Threshold Autoregressive
OPIS	Oil Price Information Service
P	precision
PC	price change
Q1	first quartile
Q2	second quartile
Q3	third quartile
RG	rule generator
RTP	real-time price
RT	rule tester
RV	rule validator
SC	single-component
STDEV	standard deviation

TAR Threshold Autoregressive

TN true negative

TNR true negative rate

TP true positive

TPR true positive rate

UB upper bound

UI user interface

Chapter 1

INTRODUCTION

Finding an effective pricing strategy is a key part of running a successful retail business. A variety of pricing strategies can be used by a business when selling a product or service. The price of a product or service may be set to maximize profit, to increase the sales of other services or products, to increase market share, or to achieve any other goal set by the seller. Obtaining the maximum profit is the most common target when an organization sets its prices. A higher price for a product or service may not cause a higher profit because of lost potential sales. The price may be set higher or lower depending on the needs and behaviours of customers in a particular area. For example, if local consumers are highly price conscious, then having relatively low prices may be a prerequisite to attracting customers.

To set a reasonable price for a product or service, a retailer may want to develop a pricing strategy. In some cases, the process of setting retail prices at a store or a brand is strongly influenced by the prices of identical or similar products in other stores or brands. One pricing strategy is to observe price changes made at other stores or brands for identical or similar products or services and adjust prices according to *price change rules*, i.e., rules that specify how to change the price at a store or brand when the price changes at another store or brand. Stores or brands that follow rules for setting prices do not necessarily follow the same rules. Thus, when attempting to form a price setting strategy for product at a store or brand; it may be of strong interest to automatically derive price change rules that are consistent with the price changes observed at other stores or brands.

This thesis focuses on automatically generating and validating price change rules that explain changes in retail prices for commodities. We concentrate on commodities that are purchased regularly in small amounts by consumers such that the commodities are not amenable to online commerce and they frequently change in price. For example, prices of motor fuel may change once a week or several times a day. To deduce the price change rules used for setting such prices, our research captures relationships between frequent changes in motor fuel prices at various stores and brands. After receiving knowledge about the hypothesized price change rules of a store or a brand, people can make more informed decisions when creating pricing strategies.

1.1 Motivation and Applications

Price change rules are of interest to store managers. Store managers are motivated to know about price changes at nearby stores as well as at other stores or brands. A price change rule with high predictive accuracy could be used as an explanation of why a change occurred at another store. These decision rules could contribute to the design of a pricing strategy.

Two types of users who may benefit from our approach are as follows:

(1) *Specific store managers*: A manager of a specific store may apply our approach to learn about the price change rules used at nearby stores or at stores of the same brand. This information may potentially be used to set prices at the manager's store.

(2) *Specific brand managers*: A brand manager may apply our approach to learn about the price change rules used by other brands.

Table 1.1: Price change rules for set of stores.

	PC categories					
		1	...	j	...	z
Stores	s_1	r_{11}	...	r_{1j}	...	r_{1z}
	⋮	⋮	⋮	⋮	⋮	⋮
	s_i	r_{i1}	...	r_{ij}	...	r_{iz}
	⋮	⋮	⋮	⋮	⋮	⋮
	s_n	r_{n1}	...	r_{nj}	...	r_{nz}

1.2 Research Problem

Our research is intended to design a knowledge acquisition system that generates meaningful price change rules as well as to evaluate the effectiveness of this system at producing price change rules with high predictive accuracy. In this context, *knowledge acquisition* is the process of extracting and organizing knowledge from one source and storing it in some other location such as a knowledge base [14].

A large data set containing reports of times, locations, and prices of motor fuel is available as input to generate price change rules. A system was desired such that the user gives commands by selecting parameters and then according to the commands and input data values, knowledge is extracted and price change rules are generated.

A major concern of our research is to provide decision rules that relate price changes at one store or brand to price changes at other stores or brands. It is also desired that these decision rules have high predictive accuracy on unseen data. The target of this approach is to generate rules with precision more than or equal to 60% and accuracy more than or equal to 80%. Table 1.1 shows an example of the desired output. This table shows possible price change rules for set of stores $S = \{s_1, s_2, \ldots, s_n\}$, where the number of price change categories is z. The set of rules for store i is represented by $R = \{r_{i1}, r_{i2}, \ldots, r_{in}\}$. Rule r_{ij} from Table 1.1 gives a possible explanation, in the form of a price change rule, of why a store s_i made a change in price in category j. Thus, the goal of our research is to generate a decision rule for a store in each of the price change categories. In this formulation, the term “store” could be replaced by “brand” throughout.

This research problem is distinct from predicting prices. The goal is to provide a high quality, comprehensible rule for every brand or every store rather than an arbitrarily complex prediction rule. In previous research we evaluated the effectiveness of two rule based classification algorithms and six decision tree based classification algorithms on a subset of the data used in the experiments described in this thesis [21]. Although obtaining accurate predictions is part of our goal it is not the single focus. We use a wider definition of quality, based on seventeen measures as described in Section 2.5.

1.3 Contributions of Our Research

Our research developed an approach for generating and validating decision rules concerning motor fuel price changes and implemented it as the Knowledge Acquisition System for Price ChangE Rules (KASPER). Our approach generates rules about price changes for a key store or brand in relation to other stores or brands with some measures that automate judging whether or not a rule is useful. From hundreds or thousands of preliminary generated price change rules, this system selects some rules as decision rules, i.e., if the end of day price of some other store or brand is in a price change (PC) category then the end of day price of the key store or brand will be in a particular PC category. Overall, this system generates a highly accurate and precise decision rule for each PC category for a key store or brand. According to such decision rules, if other stores or brands make a change in a specific PC category, then the key store or brand will also make a change in the same PC category. To determine the effectiveness of our approach, we also tested every decision rule on unseen data and measured its quality.

Contributions of this research are given below:

- Our system generates and validates thousands of price change rules while previous research on pricing knowledge made and tested a small number of researcher formulated hypotheses [34].
- Our system can generate and test more rules than a person.
- A novel method of assigning a utility score based on seventeen existing measures was designed to select decision rules.
- KASPER provides the first implementation of an integrated, automated approach for generating price change rules for motor fuel prices.

1.4 Thesis Overview

The approach of this thesis is to first provide background material and then describe the design, implementation, and evaluation of the KASPER system. The contents of the remaining chapters are summarized below.

Chapter 2 presents a review of previous work related to our research and explains the background topics required for developing a methodology to generate price change rules from store-to-store and brand-to-brand relationships. First, the factors relevant to pricing strategies of motor fuel are described. Then, we define a knowledge acquisition system and describe the architecture of a typical one. Details are given about statistical terminology for analyzing results, techniques for measuring rule quality, and techniques for choosing good decision rules. Finally, previous research about motor fuel pricing knowledge is discussed.

Chapter 3 provides an overview of our approach and describes the KASPER. This approach derives end of day price change (EODPC) for multiple stores from a historical data set and generates decision rules for price changes. This chapter describes pricing terminology, rule formats, and the specific methodology encompassed in the system.

The experimental setup and the results from applying the method to a large commercial data set are discussed in Chapter 4. First, this chapter describes the experimental environment, the historical data set, the price change categories, the sample generated rules with quality measures, as well as the behaviour of the average EODPC for four cities. Second, this chapter analyzes the overview of results for brand-based and distance-based store-to-store rules and brand-to-brand rules. Third, a detailed explanation of accuracy, precision, true positive rate, and F-measure are given for brand-based store-to-store rules. Fourth, the median and mean number of competitors are presented for distance-based store-to-store rules for several radiuses. Fifth, the heuristic rank-based variation of KASPER on stores for two brands of two cities are described. According to our evaluation, KASPER is effective at generating decision rules with high predictive accuracy for two out of four major urban areas that were analyzed. Finally, our research is compared with other previous research.

Chapter 5 summarizes the thesis work and draws conclusions. Possible related future work is also discussed.

Two appendices are included in the thesis. Appendix A gives information on general product pricing. Appendix B provides detailed experimental results referenced in Chapter 4.

Chapter 2

BACKGROUND AND RELATED WORK

This chapter discusses background information and previous research related to this thesis. The first section describes some factors affecting motor fuel price changes. The second section describes the steps of the knowledge acquisition process and its application to this research. The third section covers the format of the classification rules that are used to represent knowledge about price changes. The fourth section discusses statistical terminology that is used for analyzing experimental results. The fifth section explains rule quality measures. The sixth section describes techniques for choosing decision rules for the system. Finally, the last section reviews similar research on motor fuel pricing. Some general information on product pricing is given in Appendix A.

2.1 Factors Affecting Motor Fuel Prices

Changes in motor fuel prices depend on a variety of factors. Some factors cause prices to move slowly and others cause prices to move quickly. Some factors have a global effect on prices, while others apply only in certain areas or at certain times.

For example, *tax* is an important factor for determining motor fuel prices. If the tax increases or decreases, it affects all stores in an area. This factor does not change frequently and leads to an almost equivalent change for every store in an area.

Similarly, *distance* from supplier is a factor that affects prices. If a store is farther from its supplier than another store, then its cost of obtaining motor fuel tends to be higher than the other store. This factor is ordinarily fixed for a particular store unless it changes its supplier or its location.

Production cost, refining cost, distribution cost, cost of brand affiliation, and *business cost* (such as rent of store and salaries of employees) are relatively constant for a certain store but may differ from store to store. So, a store may change motor fuel prices after switching suppliers or brands, after moving its location, or in response to changes in the other costs mentioned. However, such changes occur relatively rarely.

Volume of sales is another factor that influences product prices. If one store can sell a product in high volume, then it has a reduced overhead per litre. Such a store may be able to afford to sell the product at a lower price.

The season of the year (*seasonal effect*) influences prices at all stores. For example, motor fuel prices are generally higher in summer than winter. Less volatile grades of fuel must be used in summer than winter, and such grades are more expensive. Also, more people travel on vacations in summer, so demand for motor fuel increases [7].

Other factors, such as *crude oil price change*, also have a global effect on all stores. This type of change affects the wholesale price paid by a store. The *wholesale price change* has an effect on all stores. According to a naive view, if the wholesale price increases then the retail price should also increase [3]. However, this naive view does not take into account hidden factors, such as the business strategy of the store. For example, if any store has a stock of motor fuel that was obtained at a low wholesale cost then this store can continue to sell fuel at the same price even though the wholesale price has increased.

Nonetheless, if one examines the price of a particular motor fuel product, one will observe that the price at a store changes more frequently than can be explained by the above factors. For example, the price may change several times a week or even several times a day, even though the input costs have not changed significantly. Clearly, stores use a pricing strategy that depends on more than the input costs.

Competitors and *brand pricing strategies* are two factors that have an impact on the frequent price changes observed at many stores. Different brands follow different strategies when setting retail product prices. The effect of competitors on prices is recognized by noticing that in many cases all nearby stores have exactly the same price for a product. Some brands focus on providing motor fuel at the lowest price, others focus on the customer experience, and still others focus on attracting customers to their store and selling them other products with high profit margins.

Based on the above analysis, we hypothesized that in order to understand patterns in price changes we should study three relationships: the relationship between price changes at a store and those at its potential competitors, the relationship between price changes at a store and those at other stores with the same brand, and the relationship between price changes for one brand and those for other brands.

2.2 Knowledge Acquisition

Recall that knowledge acquisition is the process of extracting and organizing knowledge from one source and storing it in some other location such as a knowledge base [14]. In

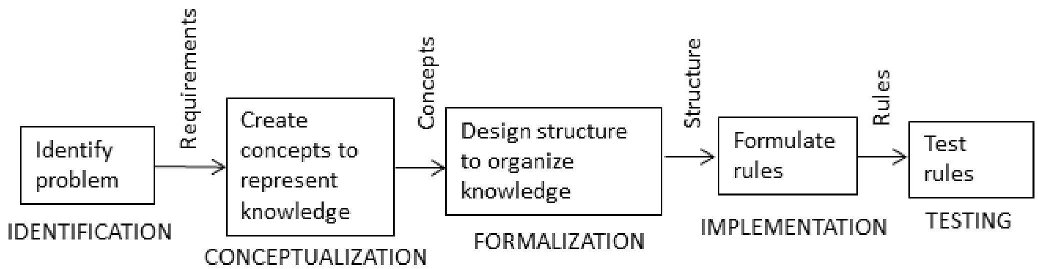


Figure 2.1: Steps of the knowledge acquisition process [5].

this thesis, we consider the case where the structured knowledge is a collection of rules. The steps of the knowledge acquisition process are shown in Figure 2.1 [5].

The *identification* step identifies the knowledge acquisition problem that we are going to solve. The *conceptualization* step creates the basic concepts required to solve the problem. The *formalization* step designs a structure to organize the knowledge and a method (algorithm) to obtain this type of knowledge. As mentioned, the knowledge is structured as a collection of rules. The *implementation* step implements the method defined in the formalization step. The *testing* step tests the rules produced by the implementation.

Knowledge acquisition has been applied to a diverse set of problems. A knowledge acquisition approach was used to develop diagnosis-specialized knowledge models [6]. A knowledge base, encoded with the NKRL (Narrative Knowledge Representation Language) formalism, was used in the management of gas/oil facilities to support tasks such as noticing gas leakage and taking steps to activate systems [39]. Several knowledge-acquisition strategies were applied to the market data of 144 internationalizing Swedish firms [1]. Knowledge acquisition was used in a supply chain partnership [15]. Knowledge acquisition was also applied to 385 manufacturer-supplier exchanges in China [40]. We did not find any previous application of a knowledge acquisition system to motor fuel prices.

2.3 Knowledge Representation

A *knowledge base* is a system for storing knowledge. As mentioned in the previous section, our research uses rules to represent knowledge. In particular, we use classification rules. *Classification* is a technique that develops a model from a data set of objects with class labels (training data set) that can classify objects into separate classes according to their properties [17]. A classification model can be represented as a set of IF-THEN rules, a decision tree, a neural net, etc.

For our research, we use conjunctive IF-THEN rules to represent price change rules. There are two major parts in an IF-THEN rule: a condition, given in the IF part, and a conclusion, given in the THEN part. We express IF-THEN rules in the following form:

IF a certain condition happens (is TRUE),

THEN a certain conclusion may occur.

For example, consider rule R1: “IF it is winter THEN people will use winter tires.” In a conjunctive IF-THEN rule, the condition part may be specified by one term or several terms joined by ANDs. For reasons explained in Chapter 3, we refer to these rules as single-component rules and double-component rules, respectively.

Consider the following group of rules [28]:

Rule1, R1: **IF** a **THEN** g1

Rule2, R2: **IF** a **and** b **THEN** g2

Rule3, R3: **IF** c **THEN** g3

Rule4, R4: **IF** d **and** f **THEN** g4

Among this group, rules R1 and R3 are single-component rules and rules R2 and R4 are double-component rules.

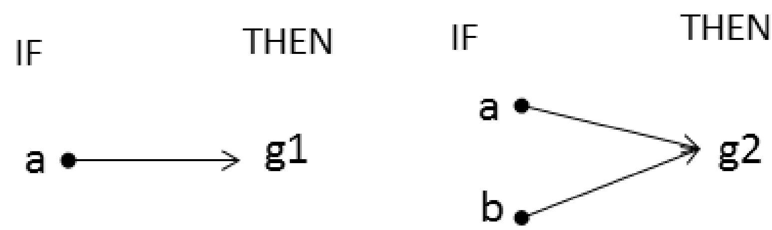


Figure 2.2: Mapping IF-THEN rule, condition to conclusion adapted from [28].

A graphical representation of the mapping between the conditions and conclusions for R1 and R2 is shown in Figure 2.2.

2.4 Statistical Terminology

In this subsection, we review common statistical terminology with examples.

Population and sample [37]: Any set can be considered to be a *population* and a subset of a population is a *sample*. In Figure 2.3, P is the population and S is the sample. For

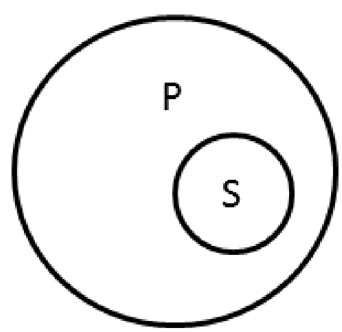


Figure 2.3: Population and sample.

example, the set P of all price reports for a certain product can be treated as a population and the subset $P \subseteq S$ of these price reports in five particular days can be treated as a sample.

Definition 5.3 [4, 28]: Let Ω be the sample space of a certain experiment and let P be a probability function on Ω. For each h, $e \subseteq \Omega$ with $P(e) > 0$, the *conditional probability* of h given e, denoted by $P(h|e)$, is defined as

$$P(h|e) = \frac{P(h \cap e)}{P(e)} \tag{2.1}$$

A conditional probability $P(h|e)$ is often called a *posterior probability.*

Three kinds of means are described below. The examples use the following set of observations:

$X = \{80.24, 77.47, 83.84, 77.67, 76.02, 75.82, 79.37, 75.82, 80.17, 75.61\}$

Definition 2.5.1:

The *arithmetic mean* [31, 32, 37] is the average of all observations in a sample. Consider a set of observations: $\{x_1, x_2, x_3, \ldots, x_n\}$. The arithmetic mean $\bar{x}$ of these observations is calculated as:

$$\bar{x} = \frac{\sum x_i}{n} \tag{2.2}$$

where n is the number of observations.

Example 2.5.1:

The arithmetic mean for set X calculated using Equation 2.2 is $\bar{x} = \frac{782.03}{10} = 78.2$.

The arithmetic mean is best used in situations where no extreme outliers exist in the set of observations and individual observations are not dependent upon each other.

Definition 2.5.2:

The *harmonic mean* [31] is the number of observations divided by sum of reciprocals of each observation. Consider a set of observations $\{x_1, x_2, x_3, \ldots, x_n\}$.

The harmonic mean of these observations is calculated as:

$$\bar{H} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots + \frac{1}{x_n}} \tag{2.3}$$

Example 2.5.2:

The harmonic mean for set X calculated using Equation 2.3 is $\bar{H} = 78.12$.

The harmonic mean is well suited to providing a true picture of the average where extreme outliers exist. Unlike the arithmetic mean, it attaches little importance to the highest and lowest outliers.

Definition 2.5.3:

The *geometric mean* [31] is the average of the observations as indicated by taking the square root of their product. Consider a set of observations $\{x_1, x_2, x_3, \ldots, x_n\}$.

The geometric mean of these observations is calculated as:

$$\text{G-mean} = \sqrt{x_1 x_2 x_3 \ldots x_n} \tag{2.4}$$

Example 2.5.3:

The geometric mean for set X is calculated using Equation 2.4 as G-mean = 78.16.

This type of mean is normally used where observations are inter-related, such as returns on investment or interest rates.

Definition 2.5.4:

The *standard deviation (STDEV)* [32, 37] provides information about how much a set of values fluctuates from their arithmetic mean. Consider a set of observations $\{x_1, x_2, x_3, \ldots, x_n\}$. The standard deviation of these observations is calculated as:

$$STDEV = \sqrt{\frac{\sum x_i - \bar{x}}{n - 1}} \tag{2.5}$$

Example 2.5.4:

The standard deviation for set X calculated using Equation 2.5 is $STDEV = 2.68$

Definition 2.5.5:

The *median (MED)* [32,37] is the value of the central components in a sorted list of values. Consider a set of observations $\{x_1, x_2, x_3, \ldots, x_n\}$. The median of these observations is calculated as:

$$MED = \begin{cases} x_{\lfloor n/2 \rfloor} \ + \ 1, \text{ if } n \text{ is odd} \\ \frac{x_{\lfloor n/2 \rfloor} \ + \ x_{\lfloor n/2 \rfloor + 1}}{2} \text{ if } n \text{ is even} \end{cases} \quad (2.6)$$

Example 2.5.5:

Consider X sorted where the values of X have been sorted in ascending order. Then, X sorted $= \{75.61, 75.82, 75.82, 76.02, 77.47, 77.67, 79.37, 80.17, 80.24, 83.84\}$.

The median for set X calculated using Equation 2.6 is $MED = \frac{77.47+77.67}{2} = 77.57$.

Definition 2.5.6:

A *percentile* [32, 37] is a calculated value that tells us about how observations are spread between minimum and maximum values. A statement that the i^{th} position of the observations is in the p^{th} percentile indicates that the value of the i^{th} position of the observations is greater than or equal to at least p percent of the items.

The following steps can be used to calculate the p^{th} percentile:

1. Sort the data set in ascending order.

2. Compute i, which indicates the position of the p^{th} percentile value,

$i = (p/100)n$ [36].

3. If i is an integer, where n is the number of observations, the p^{th} percentile is the average of the values in positions i and $(i + 1)$.

4. Otherwise, the p^{th} percentile is the value in the i^{th} position.

Example 2.5.6:

Thus, to determine the 50^{th} percentile for set X sorted, we set $p = 50$ and $n = 10$. We calculate $i = (50/100) \times 10 \ = \ 5$. Since i is an integer, the 50^{th} percentile is $\frac{77.47+77.67}{2} = 77.57$.

Definition 2.5.7:

A *quartile* [32, 37] is the value of a specific percentile. The first quartile (Q1), is the 25^{th} percentile, the second quartile (Q2), is the 50^{th} percentile, which is the median, and the third quartile (Q3), is the 75^{th} percentile.

Example 2.5.7:

The first and third quartiles of X sorted are:

Q1 = 25^{th} percentile = 75.87 and

Q3 = 75^{th} percentile = 79.97, respectively.

Definition 2.5.8:

According to a general definition, an *outlier* is "an observation that lies outside the overall pattern of a distribution" [32]. Here, we use a more specific definition, where an *outlier* is "a point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile" [35]. The following steps can be used to detect an outlier from quartile values [32, 37]:

1. Compute Q1 and Q3.

2. Compute the interquartile range (IQR) as the difference between $Q3$ and $Q1$ ($IQR = Q3 - Q1$).

3. Calculate the lower bound, LB = $Q1 - 1.5 \times IQR$.

4. Calculate the upper bound, UB = $Q3 + 1.5 \times IQR$.

5. Check each value x_i in the set of observations,

$$Outlier(x_i) = \begin{cases} TRUE, \text{ if } x_i < LB \text{ or } x_i > UB \\ FALSE, \text{ otherwise} \end{cases} \tag{2.7}$$

Example 2.5.8:

$X = \{80.24, 77.47, 83.84, 77.67, 65.02, 75.82, 79.37, 87.63, 80.17, 75.61\}$. For detecting outliers in X, the required parameters are shown in Table 2.1a and whether or not each observation is an outlier is shown in Table 2.1b. The two outliers in these observations are shown in boldface; one is less than the lower bound (LB) and the other is greater than the upper bound (UB).

Definition 2.5.9:

Boxplot and whiskers is a graphical presentation of a data set showing the quartile, me-

Table 2.1: Required parameters and checking for outliers.

(a) Parameters required for checking outliers.

Parameters	**Meaning**	**value**
Q1	1^{st} quartile	76.23
Q3	3^{rd} quartile	80.22
IQR	Inter quartile range	3.99
LB	Lower bound	70.25
UB	Upper bound	86.21

(b) Checking for the existence of outliers.

Observations (ob)	$x_i < LB$	$x_i > UB$	Outlier (x_i)
80.24	FALSE	FALSE	NONE
77.47	FALSE	FALSE	NONE
83.84	FALSE	FALSE	NONE
77.67	FALSE	FALSE	NONE
65.02	**TRUE**	FALSE	YES
75.82	FALSE	FALSE	NONE
79.37	FALSE	FALSE	NONE
87.63	FALSE	**TRUE**	YES
80.17	FALSE	FALSE	NONE
75.61	FALSE	FALSE	NONE

dian, minimum (Min), and maximum (Max) values. It is used to promote a clear understanding of how the data are distributed in the data set [32, 37]. With this technique, a box is drawn from the first quartile to the third quartile, and whiskers (vertical lines with horizontal strokes) are drawn from the minimum value to the first quartile, and from the third quartile to the maximum. This presentation is also helpful for noticing outliers and the relative position of the median.

Example 2.5.9:

To enable plotting values by drawing boxplot and whiskers, we first calculate the values of the parameters, as shown in Table 2.2a, and then compute the differences between values, as shown in Table 2.2b. The boxplot and whiskers for the calculated values of Table 2.2b is shown in Figure 2.4. The Y-axis shows the values of *Min*, *Q1*, *MED*, *Q3*, and *Max*.

Table 2.2: Values used for boxplot and whiskers example.

(a) Values of parameters.

Parameters	Meaning	Value
Min	Minimum	75.61
$Q1$	1^{st} quartile	75.87
MED	Median	77.57
$Q3$	3^{rd} quartile	79.97
Max	Maximum	83.84

(b) Differences between parameters.

Terms	Difference	Value
$Q1_d$	$Q1 - Min$	0.26
MED_d	$MED - Q1$	1.70
$Q3_d$	$Q3 - MED$	2.40
Max_d	$Max - Q3$	3.87

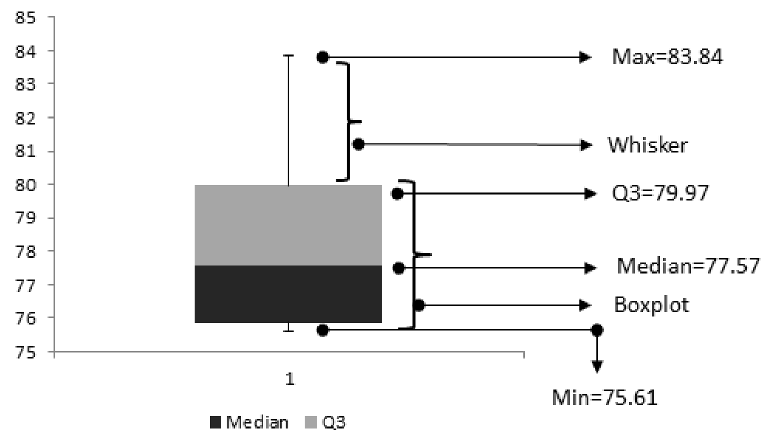


Figure 2.4: Boxplot and whiskers.

2.5 Rule Quality Measures

Seventeen rule quality measures are described in this section.

A *confusion matrix* [24] represents information about the actual and predicted classification of a classifier. This matrix indicates the predictive quality of the classifier when applied to instances. There are four basic measures:

- *true positive (TP)* is the total number of **positive** instances that are **correctly** classified.
- *false positive (FP)* is the total number of **positive** instances that are **incorrectly** classified.
- *true negative (TN)* is the total number of **negative** instances that are **correctly**

Table 2.3: Confusion matrix for basic measures.

		Actual	
		Positive	**Negative**
Prediction	**Positive**	TP	FP
	Negative	FN	TN

classified.

- *false negative (FN)* is the total number of **negative** instances that are **incorrectly** classified.

The confusion matrix of basic measures is shown in Table 2.3. Different measures can be determined from the relationships between the predicted and actual classifications. The following terms can be defined from the confusion matrix:

- The *accuracy (AC)* is the proportion of the total number of true values to the total number of all values, as shown in Equation 2.8.

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.8}$$

- The *precision (P)* is the proportion of the number of true positive values to the total number of true positive and false positive values, as shown in Equation 2.9.

$$P = \frac{TP}{TP + FP} \tag{2.9}$$

- The *true positive rate (TPR)* is the proportion of the number of true positive values to the total number of true positive and false negative values, as shown in Equation 2.10. This term is also known as *recall*, *hit rate*, or *sensitivity*.

$$TPR = \frac{TP}{TP + FN} \tag{2.10}$$

- The *true negative rate (TNR)* is the proportion of the number of true negative values to the total number of true negative and false positive values, as shown in Equation 2.11. This term is also known as *specificity*.

$$TNR = \frac{TN}{TN + FP} \tag{2.11}$$

- The *false positive rate (FPR)* is the proportion of the number of false positive values to the total number of false positive and true negative values, as shown in Equation 2.12. This term is also known as *type-I error*.

$$FPR = \frac{FP}{FP + TN} \tag{2.12}$$

- The *false negative rate (FNR)* is the proportion of the number of false negative values to the total number of false negative and true positive values, as shown in Equation 2.13. This term is also known as *type-II error*.

$$FNR = \frac{FN}{FN + TP} \tag{2.13}$$

- The *false discovery rate (FDR)* is the proportion of the number of false positive values to the total number of true positive and false positive values, as shown in Equation 2.14.

$$FDR = \frac{FP}{TP + FP} \tag{2.14}$$

- The *error rate (E)* is the proportion of the total number of false values to the total number of all values, as shown in Equation 2.15. This term is the same as $(1 - Accuracy)$.

$$E = \frac{FP + FN}{TP + TN + FP + FN} \tag{2.15}$$

There are other measures that are illuminating for some cases when trying to evaluate a rule appropriately. We will motivate the discussion of these measures by presenting an

Table 2.4: Confusion matrix for two scenarios.

(a) Confusion matrix for scenario I.

		Actual	
		Positive	**Negative**
Prediction	**Positive**	TP = 0	FP = 0
	Negative	FN = 10	TN = 9,990

(b) Confusion matrix for scenario II.

		Actual	
		Positive	**Negative**
Prediction	**Positive**	TP = 8	FP = 32
	Negative	FN = 2	TN = 9,960

example where a rule has high accuracy but nonetheless, is not of high quality. We will also show that the true positive rate is not sufficient to measure the quality of a rule.

Example 2.5.10:

Consider a classification rule with two classes, increase and not increase.

"IF end of day price change on a day d of a store A increases

THEN end of day price change on day d of key store K will increase."

Two possible scenarios are shown in Table 2.4. In both cases, the total number of instances is 10,000. In scenario I, TP = 0, FP = 0, FN = 10, and TN = 9,990. The accuracy for scenario I is calculated using Equation 2.8 as $AC = \frac{0+9990}{10000} = 99.9\%$, which seems good. However, if we calculate the true positive rate using Equation 2.10, we obtain $TPR = \frac{0}{0+10}$, which is 0, meaning that there is no positive instance that is classified correctly. So, accuracy is not a sufficient criterion to judge the quality of a rule.

Consider scenario II in Table 2.4b where TP = 8, FP = 32, FN = 2, and TN = 9,960. Here the true positive rate is $TPR = \frac{8}{8+2} = 80\%$, but the number of false positives is 32, which is large compared to the number of true positives. So, we can also calculate the precision. From Equation 2.9, the precision is $P = \frac{8}{8+32} = 20\%$. Although the value for the true positive rate is high, the value for precision is not; therefore, we conclude that this rule is not trustworthy. Thus, the true positive rate is not a sufficient indicator of the quality of a rule.

- Another measure is the *F-measure (F)* [38], which is the harmonic mean of precision and recall. It is calculated using Equation 2.16:

$$\text{F-measure} = 2 \times \frac{P \times TPR}{P + TPR} \tag{2.16}$$

- The *G-mean (G)* [25] is the geometric mean of precision and recall, as shown in Equation 2.17:

$$\text{G-mean} = \sqrt{P \times TPR} \tag{2.17}$$

None of the previous measures provide information about the likelihood of using a rule when making predictions (testing) on unseen data. Likelihood ratios [10, 29] give information about the likelihood of an increase or a decrease in the chance of using a rule. There are two types of likelihood ratios: positive likelihood ratio and negative likelihood ratio.

- The *positive likelihood ratio (LR+)* is the ratio between TPR and FPR. The positive likelihood ratio is defined as:

$$LR+ = \frac{TPR}{FPR} \tag{2.18}$$

In our application, the positive likelihood ratio of a rule is the ratio between (a) the pre-test probability that the rule predicts that a key store or brand will make a change in a PC category and it actually does so, and (b) the pre-test probability that the rule predicts that a key store or brand will make a change in a PC category but it does not actually do so.

- The *negative likelihood ratio (LR-)* is the ratio between FNR and TNR. The negative likelihood ratio is defined as:

$$LR- = \frac{FNR}{TNR} \tag{2.19}$$

Table 2.5: Interpretation for likelihood ratio values in several ranges (adapted from [10]).

LR	**Interpretation**
>10	Large and often conclusive increase in the likelihood of a key store or brand making a change in a PC category.
5 - 10	Moderate increase in the likelihood of a key store or brand making a change in a PC category.
2 - 5	Small increase in the likelihood of a key store or brand making a change in a PC category.
1 - 2	Minimal increase in the likelihood of a key store or brand making a change in a PC category.
1	No change in the likelihood of a key store or brand making a change in a PC category.
0.5 - 1.0	Minimal decrease in the likelihood of a key store or brand making a change in a PC category.
0.2 - 0.5	Small decrease in the likelihood of a key store or brand making a change in a PC category.
0.1 - 0.2	Moderate decrease in the likelihood of a key store or brand making a change in a PC category.
< 0.1	Large decrease in the likelihood of a key store or brand making a change in a PC category.

In our application, the negative likelihood ratio is the ratio between (a) the pre-test probability that the rule predicts that a key store or brand will not make a change in a PC category and it does not actually do so, and (b) the pre-test probability that the rule predicts a key store or brand will not make a change in a PC category and it actually does so.

An interpretation of likelihood ratio values in several ranges is shown in Table 2.5 (adapted from [10]) . Higher values of LR+ indicate higher probabilities that the key store or brand will actually make a specific change when the rule predicts that change. Lower values of LR- indicate lower probabilities that a key store or brand will make a specific change when the does not predict that change.

- The *diagnostic odds ratio (DOR)* [29] shows information about how LR+ behaves with LR-. It is defined as follows:

$$DOR = \frac{LR+}{LR-} \tag{2.20}$$

2.6 Decision Rule Selection

There is no single accepted way of measuring the quality of a rule [30]. Different algorithms use different techniques (utility functions) for selecting good decision rules, depending on the problem definition and solution procedure. A *utility function* is a function that rates an object by its utility value. A utility function that depends on a model is called a *model based utility function* [16].

Table 2.6, provides a guideline for the interpretation of low (indicated by MIN) and high (indicated by MAX) values for seventeen quality measures. The first column describes the measures, the second shows how to increase quality according to this measure, and the third shows which type of values (maximum ones or minimum ones) indicate high quality according to the measure. Definition of these measures are given in the previous section.

Table 2.6: Interpretation of min and max value for seventeen quality measures.

Measure	High Quality	Interpretation
TP	MAX	A higher value of TP indicates a higher frequency of correct classification of positive instances.
FP	MIN	A lower value of FP indicates a lower frequency of incorrect classification of positive instances.
FN	MIN	A lower value of FN indicates a lower frequency of incorrect classification of negative instances.
TN	MAX	A higher value of TN indicates a higher frequency of correct classification of negative instance.
AC	MAX	A higher value of AC indicates a higher frequency of correct classification in comparison to correct and incorrect classification of instances.
E	MIN	A lower value of E, which is $(1 - AC)$ indicates a lower frequency of incorrect classification in comparison to correct and incorrect classification of instances.
P	MAX	A higher value of P indicates a higher frequency of correct classification of positive instances in comparison to correct and incorrect classification of positive instances.
FDR	MIN	A lower value of FDR indicates a lower frequency of incorrect classification of positive instances in comparison to correct and incorrect classification of positive instances.
TPR	MAX	A higher value of TPR indicates a higher frequency of correct classification of positive instances in comparison to correct classification of positive instances and incorrect classification of negative instances.
TNR	MAX	A higher value of TNR indicates a higher frequency of correct classification of negative instances in comparison to correct classification of negative instances and incorrect classification of positive instances.
FPR	MIN	A lower value of FPR indicates a lower frequency of incorrect classification of positive instances in comparison to correct classification of negative instances and incorrect classification of positive instances.
FNR	MIN	A lower value of FNR indicates a lower frequency of incorrect classification of negative instances in comparison to correct classification of positive instances and incorrect classification of negative instances.
F	MAX	A higher value of F indicates a higher value of harmonic mean value for TPR and P.
G	MAX	A higher value of G indicates a higher value of geometric mean value for TPR and P.
LR+	MAX	A higher value of LR+ indicates a higher probability for TPR.
LR-	MIN	A lower value of LR- indicates a higher probability for TNR.
DOR	MAX	A higher value of DOR indicates a higher probability for LR+.

2.7 Motor Fuel Pricing

Our research is related to retail motor fuel prices. This section discusses previous research related to this topic. First, this section describes the three findings concerning the behaviour of prices in the U.S. Second, this section examines research that shows prices rise more rapidly than they fall. Third, this section describes experiments concerning price uniformity in the retail motor fuel market. Finally, this section explains price leadership and coordination among motor fuel stores.

Hosken et al. reported three findings concerning the behaviour of retail motor fuel prices [19]. Data was collected from three sources. The first source was a three year data set of weekly motor fuel prices based on fleet card transactions from 272 motor fuel stores located in the Northern Virginia suburbs of Washington, DC. These data were collected from the Oil Price Information Service (OPIS). The second source was annual surveys of roughly 600 stores, which collected the address, attributes (e.g., whether a convenience store is present and the number of pumps), and the vertical relationship between the store and its supplier. The third source was census information on neighbourhood characteristics (measured at the zip-code level), which include median household income, population, population density, and commuting time. Hosken et al. used the prices at the pump for regular (87-octane) motor fuel (including tax) as the retail price, the average "branded rack" price in a week as the wholesale price, and the retail price less the branded rack price and taxes as the *margin*. Hosken et al. claim that the primary sources of retail price variation results from (a) a store changing its price in response to a change in the wholesale price and (b) a store changing its price relative to other stores.

After conducting experiments on the collected data set, Hosken et al. stated three findings, given below:

- *Finding 1: The retail margins vary substantially over time.*

- *Finding 2: Stores do not follow simple pricing rules.*

- *Finding 3: Many stores change their pricing strategy over time.*

Finding 1 resulted from an examination of the 25^{th}, 50^{th}, and 75^{th} percentiles of the distribution of the weekly margins from 1997 through 1999. Hosken et al. observed that the retail margin shifts significantly over time.

Finding 2 resulted from examining *retail price dispersion* as the deviation of the price of a store from the region's mean price at a point in time. "Price dispersion occurs when different sellers offer different prices for the same good in a given market" [18]. They analyzed retail price dispersion by examining the residuals from the following regression:

$$p_{it} = \sum_{t} \gamma_t(WeekIndicator_{it}) + e_{it} \tag{2.21}$$

where $p_{i,t}$ represents store i's motor fuel price in week t, and γ_t represents the coefficients corresponding to the weekly indicators. Hosken et al. used "convenience store", "provides repair service", "outdated format", and "self serve only" as indicator variables. They estimated p_{it} values for the variables on the right hand side of Equation 2.21 using data for each store and time period. The residual is the difference between the actual price and the predicted price.

Hosken et al. found 56% of prices are within 2.5 cents per gallon of the region's mean and 71% of prices are within 3.5 cents per gallon [19]. Hosken et al. also found that 3.5% of prices are more than 10 cents per gallon from the mean [19]. They also found that the distribution of the residuals is not normal. If the residuals were normal, they would have expected that between 47% and 62% of prices would be within 2.5 and 3.5 cents per gallon of the mean, and 1.2% of prices would be more than 10 cents per gallon from the mean. So, the overall results supported the rejection of the null hypothesis that the residuals have a normal distribution.

They also analyzed a store's price changes by defining the store's relative price in week t to be the residual from Equation 2.21, i.e., the difference between store i's price in week t and the mean price of all stores in week t. The authors rounded the residual to the nearest cent and constructed a Markov transition matrix, where the elements of the matrix showed the probability of being y cents per gallon above (or below) the mean in period t, conditional on being x cents per gallon above (or below) the mean in period $t-1$.

To examine the role of heterogeneity in characterizing retail motor fuel pricing, they controlled for both time effects and time-variant-store effects using Equation 2.22:

$$p_{it} = \sum_i \theta_i(Store\ Indicator_{it}) + \sum_t \gamma_t(Week\ Indicator_{it}) + u_{it} \tag{2.22}$$

where θ_i is a coefficient representing the store-specific fixed effects at store i. That is, θ_i is store i's mean relative price. θ_i is zero for a store with the mean price.

For finding 3, they hypothesized that stores change their relative prices over time. To examine this hypothesis, they used a slightly modified version of Equation 2.22 and allowed the store effects to vary by calendar year ($q = 1997, 1998, 1999$):

$$p_{it} = \sum_t \gamma_t(Week\ Indicator_{it}) + \sum_{i,q} \theta_i^q(Store\ Indicators_{it})(Year_{it}) + w_{it} \tag{2.23}$$

They determined that θ_i^{1997}, θ_i^{1998}, and θ_i^{1999} were all unequal and gave the conclusion as finding 3.

Al-Gudhea et al. examined the idea that retail motor fuel prices rise more rapidly than they fall [3], using threshold and momentum models of co-integration developed by Enders et al. [12, 13]. This study examined the behaviour of the response times of daily retail motor fuel prices to upstream and downstream price changes at different stages of the distribution chain [23]. “Upstream” and “downstream” are common business terms used in reference to the supply chain in the oil and motor fuel industry. *Upstream* refers to the raw material extraction or production elements of the supply chain [20]. *Downstream* refers to firms closer to the end users or consumers [20]. They investigated pairwise dissimilar adjustments between (a) the crude oil price and the retail motor fuel price, (b) the crude oil price and the spot motor fuel price, (c) the spot motor fuel price and the wholesale motor fuel price, and (d) the wholesale motor fuel price and the retail motor fuel price. They used the West Texas Intermediate spot price as their crude oil price and the average of New York, Gulf Coast, and Los Angeles conventional regular motor fuel spot prices as the spot price of motor fuel. Prices were collected on a daily basis from the U.S. Department of Energy for the period from December 1998 to January 2004. Daily wholesale and retail prices were collected from the OPIS.

Al-Gudhea et al. studied the behavior of downstream price responses to changes in the upstream prices of motor fuel. They considered the following long run relationship between the upstream and downstream prices of motor fuel:

$$y_t = \beta_0 + \beta_1 x_t + \mu_t \tag{2.24}$$

where y_t is the downstream price, β_0 and β_1 are constants, x_t is the upstream price, and μ_t is a stationary random variable that represents the deviation from the long run equilibrium, if any. For dissimilar adjustments to the model [13], the deviation from the long-run equilibrium μ_t in Equation 2.24 behaves as a Threshold Autoregressive (TAR) process:

$$\Delta\mu_t = I_t \rho_1 \mu_{t-1} + (1 - I_t)\rho_2 \mu_{t-1} + \sum_{i=1}^{p} \beta_i \Delta\mu_{t-i} + \epsilon_t \tag{2.25}$$

where I_t is the Heaviside indicator such that:

$$I_t = \begin{cases} 1 \ \text{ if } \mu_{t-1} \geq \tau \\ 0 \text{ if } \mu_{t-1} < \tau \end{cases} \tag{2.26}$$

and τ is the value of a threshold. If the Heaviside indicator uses the amount of change in μ_t (i.e., $\Delta\mu_t$) instead of the level of μ_t, then Equation 2.26 becomes the following:

$$I_t = \begin{cases} 1 \ \text{ if } \Delta\mu_{t-1} \geq \tau \\ 0 \text{ if } \Delta\mu_{t-1} < \tau \end{cases} \tag{2.27}$$

where $\Delta\mu_{t-1} = \mu_t - \mu_{t-1}$

Equation 2.27 is relevant whenever the series exhibits more "momentum" in one direction than the other [12, 13]. This is called the Momentum-Threshold Autoregressive (M-TAR) model.

Al-Gudhea et al. showed the importance of the size of the oil price shocks in determining the outcome of the ultimate motor fuel price response. For large shocks, the response of downstream prices and upstream prices seem similar for all pairwise relations except for pairwise relations at the retail level. The dissimilar responses in the pairwise relation at the retail level were more evident with small shocks than with large shocks. This paper considered \$1 shocks as large; the average daily shock is 60 times smaller.

Eckert and West experimented with price uniformity in the retail motor fuel market [11]. *Price uniformity* occurs when the prices of the same type of product are the same everywhere in a market. This paper considered two alternative types of pricing in a retail motor fuel market. The first type of pricing is tacitly collusive pricing at the brand level; the second type is non-collusive pricing in a spatial market. A competitive market model in the retail motor fuel market in Canada was adapted from the following assumptions [11]:

(1) Consumers are able to move and check motor fuel prices charged at different stores in the same geographic market at low or zero cost.

(2) Retail stores post prices so that rival motor fuel stores can check each other's prices at low or zero cost.

(3) Individual retail motor fuel stores can set their prices by themselves.

Another explanation for price uniformity that is consistent with both spatial and product differentiation, is that certain firms use price uniformity to support tacit collusion and to coordinate their behaviour. It seems that major brands have more control over tacitly collusive prices than fringe firms.

To examine price uniformity, this paper used store-specific daily retail motor fuel prices for the period from March 1 to August 31, 2000 for the Vancouver, BC metropolitan area. The retail prices used were reported by consumers to the website http://www.gastips.com. Each price report consisted of the price charged, the store location, the store brand, and the time and date. Motor fuel store addresses and characteristics were obtained from Kent Marketing Limited year 2000 outlet facility reports. The sample consisted of 426 stores with 6651 unique price reports over 80 days, where 35 stores were missing and 391 stores were observed. Each store had an average over the 80 days. This paper used the econometric model to describe the equilibrium pricing pattern, then tested this pattern on the competitive and tacitly collusive pricing. For this purpose, the sample was divided into two groups: prices below the mode price and prices at or above the mode price.

The econometric model is:

$$I_{it} = \begin{cases} 1 \text{ if } X_{it}\beta + \epsilon_{it} > 0 \\ 0 \text{ otherwise} \end{cases} \tag{2.28}$$

where ϵ_{it} is normally distributed with a mean of zero and a variance of one, and where $\text{Cov}(\epsilon_{it}, \epsilon_{js}) = 0$ for $i \neq j$ or $t \neq s$ or both. I_{it} is an indicator variable that equals one if store i sets the price to the mode price or above on day t, and zero otherwise. The X variables were considered to capture brand effects, spatial and product characteristics, market structure, and time series variation. Eckert and West estimated the coefficients in the econometric model and measured significance at the 1 percent, 5 percent, and 10 percent levels. They also measured the probability of matching prices with the mode price on the average day. It was assumed in this paper that a store is either a major brand store with supplier control over price, a major brand store with dealer control over price, an ARCO or Tempo store, or another major brand store. After analyzing the data, they concluded that the competitive market model is rejected as the explanation for motor fuel store pricing in Vancouver. The results are more consistent with tacitly collusive pricing behaviour in this market.

Lewis showed price leadership and coordination among motor fuel stores in the Midwestern United States where prices were highly cyclical [27]. He investigated whether a price cycles existed or not in particular cities. A *price cycle* is the cyclical fluctuations or periodic jump of prices. According to this paper, "the cycles consist of a period of aggressive price undercutting followed by a very rapid and universal relenting of prices back to more profitable levels [27]". This study also focused on the pricing behaviour in some markets where cycles exist and some where they do not. This paper claims that *"Midwestern U.S. retail prices often fall at an average of a cent per day or more for a week or two and jump 10 to 20 cents in one day before starting to fall again"*. For experimentation purposes, this research collected data from three different sources. The first data set contained daily average retail prices from 280 cities nationwide from October 2004 to July 2010, which were used to determine cities where retail price cycles occur. These prices come from the reported prices on the American Automobile Association (AAA) website which were based on information from a store price survey collected by the OPIS. The second set of data contained store-specific retail motor fuel prices collected by OPIS for 165 cities from July 2008 to July 2010. These prices were for regular grade (87-octane) motor fuel. The brand of motor fuel sold, the business name of the store, and its street address are included with the data set. These data were used to determine the timing and coordination of price movements in cycling cities. The final data set was collected from the OPIS daily store level prices. OPIS store level data gives information of all stores in a city. Therefore, this data can be used for estimating market share. To understand the exact timing of price movements during restoration, the author collects pump prices for every Speedway store at 3 hour intervals throughout the day from August 2008 to July 2010.

From this study, Lewis found that a particular retail chain in each city acts as a price leader initiating each price restoration. Price restoration happens when the leader signals the new price level to competitors by simultaneously jumping prices at all its stores to a single price, and competitors follow quickly with a large majority of stations jumping to the exact same price. From the structure of the retail motor fuel market, this paper claims that *"final retail pricing decisions are now largely made by dealer operators or store owners rather than by the branded supplier (i.e. refining company)."*

Lewis analyzed the data set and showed empirical results in the following ways:

(1) *Where a price cycle exists*

Lewis examined the average retail prices of 280 cities and identified cyclical pricing behaviour by applying the statistical indicator proposed by Lewis [26] and also used by Doyle et al. [9]. The median daily change in the city's average retail price is a good metric for determining the presence of Edgeworth price cycles [26]. For cycling markets, daily prices fall in small increments and occasionally show large increases. Thus, the median of daily price changes are distinctly negative. On the other hand, price changes in non-cycling markets respond primarily to cost fluctuations and tend to have a median daily price change very close to zero. Similar to [26], Lewis examined median daily change in the average retail price and considered a median price change below -0.2 cents per gallon as a strong indicator of cycling pricing. From the experiment the author found that 46 samples showed strong signs of cycling pricing ($median\ \delta p < -0.2$), out of the 280 cities. From the second data set the author found that 52 cities out of 165 in the OPIS sample showed cycling behaviour. These 52 cities consisted of the 46 cities with cycling markets and 6 additional small cities that were not included in the AAA sample but which clearly showed cycling behaviour based on the median daily price changes. From the market share, the author found that Speedway and Quik Trip have more market share in a cycling pricing market than a non-cycling. These were the independent retailers and stores of these retailers were trying to dominate price changes in the cycling market. "Cycles are more likely in cities with more independent stores" was found in [26] and Lewis, in another paper [9], argues that motor fuel stores with convenience stores have more impact on motor fuel prices because they can make profit from customers in-store purchases. They showed that cycles occur more frequently in a market where stores with convenience stores have more market share.

(2) *Coordinating price restorations*

From the analysis of the daily distribution of prices during rapid price jumps in cyclic and non-cyclic cities, Lewis concluded that the cycling markets were unique, not only in the frequency of price jumps but in the way that competitors coordinate their price increases. Lewis identified 3288 cyclical price restoration events over the 2 year period in 52 cities that showed cyclical behaviour. The median of a cycle was 8 days, and the 25^{th} and 75^{th} percentiles of cycle length were 6 days and 14 days, respectively. The following topics are related to coordinating price restorations:

(a) *Leaders of price restorations*

To determine the leader for price restorations, Lewis estimated the coefficients and standard error for every independent dealer and branded dealer from the comparison of the probability that a store raises its prices on the first day of a cyclical restoration to the probability that a store in the same city that was not part of one of the identified retail chains did so. Experimental results showed that the coefficient of the Speedway and Quik Trip stores was significantly larger than almost every other retailer. These two were independent retail chains. This paper defined a *price jump* as an increase in price of at least 5 cents over one or two days. Therefore, if a store restores its price to a new peak that is less than 5 cents above its old price, the change will not be identified as a cyclical price jump.

(b)*Price coordination and signalling*

From the previous analysis, Speedway and Quik Trip had more market share than other stores, since these two retailers had the ability to coordinate, both the prices during a price restoration and the timing of the price jump. For this purpose, Lewis used a third data set, which consisted of the price reports of all Speedway stores collected in 3 hours intervals throughout the day from August 2008 to July 2010. Lewis calculated the median price and 5^{th}, 25^{th}, 75^{th}, and 95^{th} percentiles of Speedway's prices in the city during each day following a cyclical price restoration for each of the 38 cities where Speedway held at least 5% market share. Percentiles were then averaged across restorations within a city and across cities using a weighted average based on the number of stores in the city. The results showed that Speedway stores unify their prices within each city on the day of price restoration. Most Speedway store's prices jumped to the exact same level, and the 5^{th}

percentile prices were less than two cents below the median on average.

(c)*Retailer price aggressiveness during the undercutting phase*

To compare pricing behaviour of cycles occurring at different times and in different cities, Lewis relied on the adjusted price, which is the store's price relative to the mean price that the store jumped to on the first day of the most recent restoration. From the results of a regression of each store's adjusted price during the first day after a cyclical restoration and prices from the sixth day after a restoration, the paper confirmed that most retailers tend to jump their prices fairly closely to the new median restoration price on the first day of the cycle. However, a few firms consistently undercut the new citywide price. This study showed that the Speedway and Quik Trip brands do not appear to be particularly aggressive in undercutting competitors.

Chapter 3

THE KASPER SYSTEM

This chapter describes the Knowledge Acquisition System for Price ChangE Rules (KASPER). Section 3.1 describes pricing terminology. Section 3.2 gives an overview of the modules of KASPER Section 3.3 introduces notation and format of rules. Section 3.4 explains two strategies for choosing groups of relevant stores or brands for making rules. Section 3.5 provides information about rule quality measures and a utility function. Section 3.6 explans our method for generating price change rules. Finally, Section 3.7 describes a rank-based method for generating price change rules.

3.1 Pricing Terminology

The following terms are employed throughout the remainder of this thesis:

- A *directional rule* is a rule that indicates the direction (increase or decrease) of an expected price change.
- A *categorical rule* is a rule that indicates an interval-based category of an expected price change.
- A *key store* is a store for which rules are generated and validated to identify other stores that are relevant to price changes for the key store.

- A *key brand* is a brand for which rules are generated and validated to identify other brands that are relevant to price changes for that brand.

- A *single-store rule* is an IF-THEN rule stating that if a price change in a specific category occurs at another store on then a price change in a specific category can be expected at the key store on the same day.

- A *single-brand rule* is defined analogously to a single-store rule, with brand replacing store.

- A *single-component (SC) rule* is a single-store rule or a single-brand rule.

- A *double-store rule* is an IF-THEN rule stating that if price changes in two categories (the same or different) occur for two different stores on the same day then a price change in a specific category can be expected at the key store on the same day.

- A *double-brand rule* is defined analogously to a double-store rule, with brand replacing store.

- A *double-component (DC) rule* is a double-store rule or a double-brand rule.

- The *utility function* is a function that computes a score representing the quality of a rule based on seventeen measures. It is formally defined in Section 3.5.

- The *independent measures* are the measures that are calculated directly from the raw values from data set.

- The *first dependent measures* are calculated from the independent measures.

- The *second dependent measures* are calculated from the first dependent measures.

- The *third dependent measures* are calculated from the second dependent measures.

- The *coverage* is the fraction of possible combinations of stores and price change categories for which we have decision rules.

- The *end of day price (EODP)* is the last price reported on a specific date. The EODP on day d is denoted $EODP_d$. If there are no prices reported on a day, the EODP is considered to be missing. In the example shown in Figure 3.1, the EODP on 1 January 2015 is 3.05.

- The *real-time price (RTP)* is the price reported at a specific time on a specific date. The RTP reported at time t on day d is denoted $RTP_{d,t}$. In the example shown in Figure 3.1, the RTP at 8:00 AM on 1 January 2015 is 3.10 , the RTP at 1:00 PM is 3.10, and so on.

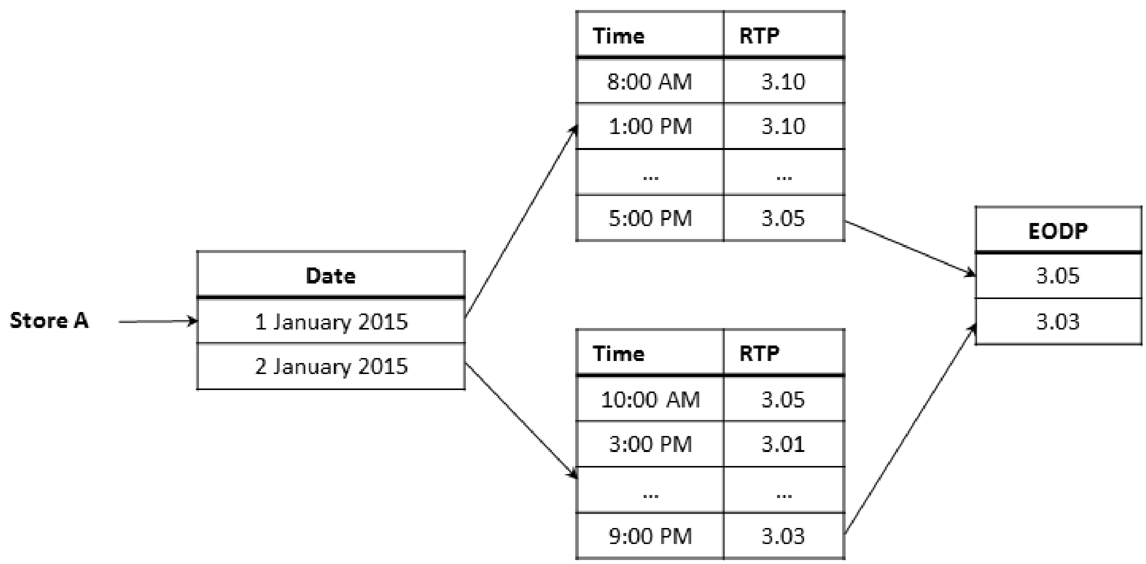


Figure 3.1: Computing end of day prices.

- The *end of day price change EODPC* is defined in Equation 3.1 as the change in price between the EODP for a specific date and the one from the immediately preceding date.

$$EODPC_d = EODP_d - EODP_{d-1} \tag{3.1}$$

The EODPC on day d is denoted $EODPC_d$. If either price is missing, the EODPC is null. The EODPC can be defined for an individual store or a brand.

Table 3.1: End of day prices and price changes.

(a) EODP.

Date	Price
2 January 2015	3.34
3 January 2015	3.39
5 January 2015	3.35
7 January 2015	3.35
8 January 2015	3.36
9 January 2015	3.36
11 January 2015	3.37
12 January 2015	3.35
13 January 2015	3.33

(b) EODPC.

Date	Price change
3 January 2015	0.05
5 January 2015	null
7 January 2015	null
8 January 2015	0.01
9 January 2015	0.00
11 January 2015	null
12 January 2015	−0.02
13 January 2015	−0.02

Table 3.2: Price changes for directional and categorical rules.

(a) Directions of price change ($z = 2$).

Direction	Meaning	Price change
		City1, City2, City3, and City4
NC	No change	$= 0$
INC	Increase	> 0
DEC	Decrease	< 0

(b) Categories of price change ($z = 6$).

Category	Price Change (PC)	
	City1, City2, and City4	City3
CAT0	$= 0$	$= 0$
CAT1	> 0 and ≤ 0.05	> 0 and ≤ 0.02
CAT2	> 0.05 and ≤ 0.10	> 0.02 and ≤ 0.04
CAT3	> 0.10	> 0.04
CAT4	< 0 and ≥ -0.05	< 0 and ≥ -0.02
CAT5	< -0.05 and ≥ -0.10	< -0.02 and ≥ -0.04
CAT6	< -0.10	< -0.04

From Table 3.1a, the EODP for two consecutive days is as follows:

Date: 2 January 2015, Price: 3.34

Date: 3 January 2015, Price: 3.39

$EODPC_d = 3.39 - 3.34 = 0.05$ for $d =$ 3 January 2015, as shown in Table 3.1b.

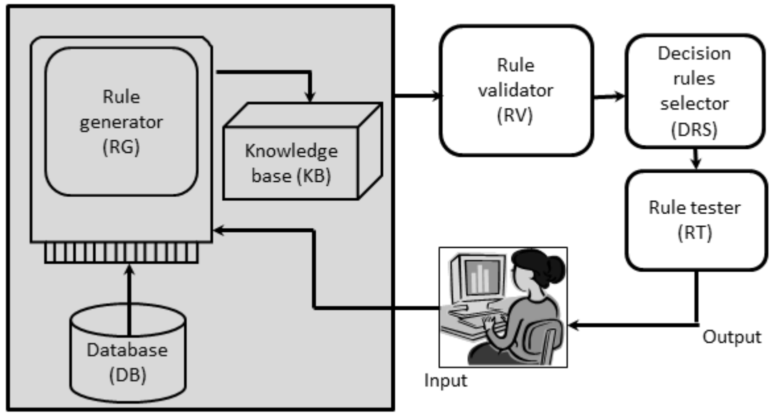


Figure 3.2: Overview of KASPER

- A *price change (PC) category* is a positive integer in the range 1 to z, where z is a small positive integer, representing the range of a price change. We consider two possibilities, 2 and 6, for z in this thesis. The categories are defined in Table 3.2. For $z = 2$, we use the symbols INC and DEC to represent the possible values 1 and 2, respectively, as shown in Table 3.2a. For, $z = 6$, we use the symbols CAT1, CAT2, CAT3, CAT4, CAT6 to represent the possible values 1 to 6, as shown in Table 3.2b. We refer to the possible categories as directions. We use the directions for making directional rules and the 6 categories for making categorical rules. NC and CAT0 both represent a case where no price change occurred. In our approach, we do not make rules for such cases.

3.2 Overview of KASPER

KASPER allows the user to learn price change rules relevant to a store or a brand. The overall approach is to generate hundreds or thousands of rules, evaluate them on data, and then select a group of the best ones as decision rules.

There are seven modules in the implementation of KASPER, as shown in Figure 3.2. The *database (DB)* contains the training data set that will be used for rule generation.

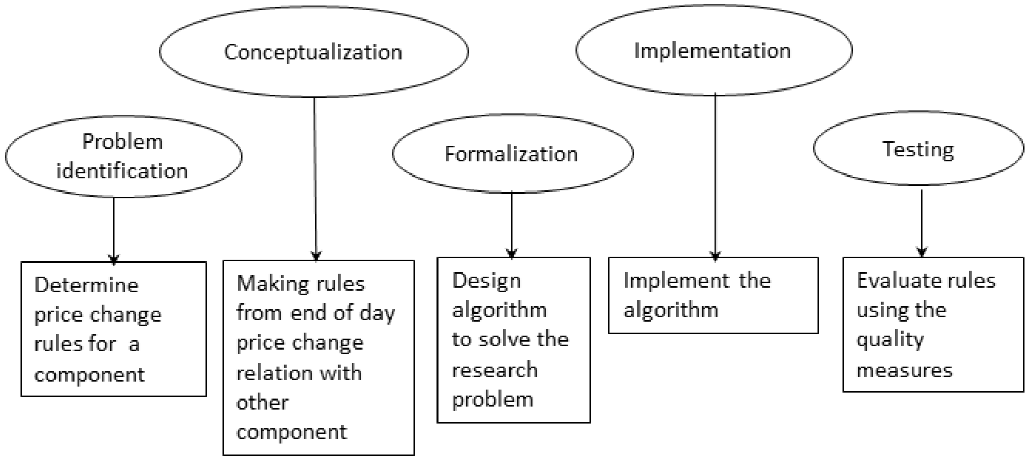


Figure 3.3: Mapping KA steps to our approach.

The *rule generator (RG)* produces rules that are consistent with the training data set. The *knowledge base (KB)* is an SQL database that contains single and double-component rules produced by the rule generator. The *rule validator (RV)* measures the quality of each rule on the validation data set. The *decision rule selector (DRS)* selects the best single and double component rule using utility function. Then one of them is selected as potential decision rule according to the higher score. If the accuracy of a potential decision rule is greater than or equal to a threshold then this rule is considered to be a decision rule. The *rule tester (RT)* tests the performance of each decision rule on unseen data set. The *user interface (UI)* interacts with a human user to obtain parameter settings and data. The parameters specify the city, the type of component (brand or store), the key component, the other relevant components, directional or categorical price change, threshold radius of distance for getting price change rules with respect to potential competitors. Then one window of UI will show the generated price change rules corresponding to the specified parameters.

The preliminary goal of KASPER is to acquire knowledge about price changes. The basic steps of knowledge acquisition, as previously shown in Figure 2.1, can be mapped to our approach, as shown in Figure 3.3. The *identification* step defines the problem statement, which is to determine price change rules for a key component from the relationship between price changes at this component and those at other components. The *conceptualization* step provides an idea to solve our research problem by making rules based on end of day price change relation with other components. The *formalization* step designs the algorithm for making decision rules. The *implementation* step implements the designed algorithm. The *testing* step is similar to the testing phase for KASPER.

3.3 Rule Format

This section introduces the notation used to specify rules and the formats of single and double-component rules with examples. Let S be a set of stores, $S = \{s_1, s_2, \ldots, s_{||S||}\}$, where $s_1, s_2, \ldots, s_{||S||}$ are individual stores. Let B be a set of brands, $B = \{b_1, b_2, \ldots, b_{||B||}\}$, where $b_1, b_2, \ldots, b_{||B||}$ are individual brands. We use $C = \{c_1, c_2, \ldots, c_n\}$, called the set of components, to denote either a set of stores S or a set of brands B. Let direction (DIR) = {INC, DEC} be a set of two possible price change categories. Let category (CAT) = {CAT1, CAT2, CAT3, CAT4, CAT5, CAT6} be a set of 6 possible price change categories.

KASPER generates EODPC rules in the following formats using meta variables $\mathbf{C_i}, \mathbf{C_k} \in C$ and $\mathbf{X}, \mathbf{Y} \in DIR$ or $\mathbf{X}, \mathbf{Y} \in CAT$:

Single-component rule:

IF $EODPC_d$ of component $\mathbf{C_i}$ is in category $\mathbf{X}$ THEN $EODPC_d$ of component $\mathbf{C_k}$ will be in category $\mathbf{Y}$.

Example: IF $EODPC_d$ of component c_1 is in category CAT1 THEN $EODPC_d$ of component c_k will be in category CAT1.

Double-component rule:

IF $EODPC_d$ of component $\mathbf{C_i}$ is in category $\mathbf{X}$ and $EODPC_d$ of component $\mathbf{C_j}$ is in category $\mathbf{Z}$ THEN $EODPC_d$ of component $\mathbf{C_k}$ will be in category $\mathbf{Y}$.

Example: IF $EODPC_d$ of component c_1 is in category CAT1 and $EODPC_d$ of component c_2 is in category CAT1 THEN $EODPC_d$ of component c_k will be in category CAT1.

The generalization of the format to triple-component rules is straightforward. However, in practice, we did not find any such rules with sufficient support on the data sets we investigated.

3.4 Relevant Stores for Generating Rules

We consider two approaches to determine a set of relevant stores for generating rules. One approach is based on distance and the other is based on brands. Let D_k be the set of relevant stores for key store k. The stores in D_k are used as potential competitors for key store k.

Definition: A *distance-based relevant store* of a key store is any other store within a specified distance (the *distance threshold*) of the key store.

The distance d between any two nearby points, point1 and point2, on the earth's surface is calculated according to the Haversine formula [33], shown in Equation 3.2:

$$d = 2 \times r \times \sin^{-1}\sqrt{\sin^2(\frac{\phi_1 - \phi_2}{2}) + \cos(\phi_1)\cos(\phi_2)\sin^2(\frac{\lambda_1 - \lambda_2}{2})} \qquad (3.2)$$

where d is the distance, ϕ_1 and ϕ_2 are the latitudes of point1 and point2, λ_1 and λ_2 are the longitudes of point1 and point2, and r is the radius of the earth (6373 km).

We use Algorithm 1 to find D_k, the distance between a key store k and some other store s using Equation 3.2.

Algorithm 1: Determine distance-based relevant stores.

Input: key store, $k\langle id, longitude, latitude\rangle$, distance threshold, set of other stores, S
Output: set of distance-based relevant stores D_k
$D_k \leftarrow \phi$
for ***each*** $s \in S$ **do**
    calculate distance between store s and key store k using Equation 3.2
    if *calculated distance* $\leq$ *distance threshold*
    then
        add store s to D_k

Definition: A *brand-based relevant store* for a key store is any other store in the same city that has the same brand as the key store.

Example: Suppose brand b has 10 stores $s_1, s_2, \ldots, s_{10}$ in a city. If we consider store s_2 to be the key store, then the brand-based relevant stores are $s_1, s_3, \ldots, s_{10}$. In this case $D_k = \{s_1, s_3, \ldots, s_{10}\}$.

3.5 Rule Quality Measures and Utility Function

Qulaity measures:

The quality of each rule is determined by the seventeen quality measures, as shown in Table 3.3. All these measures are defined in Section 2.5.

Table 3.3: Seventeen measures for rule validation.

Independent	**1st dependent**	**2nd dependent**	**3rd dependent**
TP	TPR	F	DOR
FP	TNR	G	
FN	FPR	LR+	
TN	FNR	LR-	
	FDR		
	AC		
	P		
	E		

Table 3.4: Parameters for the utility function.

Independent	**1st dependent**	**2nd dependent**	**3rd dependent**
isMax(TP)	isMax(TPR)	isMax(F)	isMax(DOR)
isMax(-FP)	isMax(TNR)	isMax(G)	
isMax(-FN)	isMax(-FPR)	isMax(LR+)	
isMax(TN)	isMax(-FNR)	isMax(-(LR-))	
	isMax(-FDR)		
	isMax(AC)		
	isMax(P)		
	isMax(-E)		

Utility function:

We use a utility function based on the sum of the "isMax" values of the seventeen rule quality measures or their negations, as shown in Table 3.4, to calculate the score of a rule r among a group of rules R for the set of seventeen measures M. Let M be the set of seventeen measures. The utility function is defined as:

$$Score(M, r, R) \;=\; \sum_{m \in M} isMax(m, r, R) \tag{3.3}$$

$$isMax(m, r, R) = \begin{cases} 1, \text{ if } \forall r' \in R, \; m(r) \geq m(r') \\ 0, \text{ otherwise} \end{cases} \tag{3.4}$$

Where r and R are clear from contexts in which we write $isMax(m)$.

Example

Suppose a set of rules **R** consists of three rules **R1**, **R2**, and **R3** and their TP values are 797, 804, and 750, respectively. So, $isMax$(TP, **R1**, **R**) = 0 for **R1**, $isMax$(TP, **R1**, **R**) = 1 for **R2**, and $isMax$(TP, **R1**, **R**) = 0 for **R3**. The same methodology can be applied to all seventeen measures, as shown in Table 3.5.

From Table 3.5, we can see that **R2** gets a score of 10, which is the highest score among the three rules. Since rule **R2** has the maximum value for TP, it will be selected as the best rule among the set of rules.

Table 3.5: Calculation of a score using the utility function.

Measures	**R1**	**R2**	**R3**	Parameters of utility function	**R1**	**R2**	**R3**
TP	797	804	750	isMax(TP)	0	1	0
FP	10	11	11	isMax(-FP)	1	0	0
FN	98	91	95	isMax(-FN)	0	1	0
TN	186	185	235	isMax(TN)	0	0	1
P	98.76	98.65	98.55	isMax(P)	1	0	0
FDR	0.01	0.01	0.14	isMax(-FDR)	1	1	0
TPR	89.05	89.83	88.76	isMax(TPR)	0	1	0
FPR	5.1	5.61	4.47	isMax(-FPR)	0	0	1
F	93.65	94.04	93.73	isMax(F)	0	1	0
G	93.78	94.14	93.87	isMax(G)	0	1	0
AC	90.1	90.65	90.28	isMax(AC)	0	1	0
E	9.9	9.35	9.71	isMax(-E)	0	1	0
TNR	94.9	94.39	95.52	isMax(TNR)	0	0	1
FNR	10.95	10.17	11.24	isMax(-FNR)	0	1	0
LR+	17.45	16.01	22.25	isMax(LR+)	0	0	1
LR-	0.12	0.11	0.11	isMax(-(LR-))	0	1	1
DOR	151.27	148.59	202.27	isMax(DOR)	0	0	1
				Score	3	10	6

3.6 Procedure for Generating Price Change Rules

Here we describe the procedure that KASPER follows to generate store-to-store and brand-to-brand decision rules for price changes. It is summarized in Figure 3.4. According to our approach, KASPER builds initial profiles and then uses them to make component profiles. All these profiles are prepared before generating any rules. Profile construction is described in Section 3.6.1

In the training phase, single-component rules are generated from the training data set. If the frequency of a rule is higher than a threshold, then the rule is added to the set of preliminary conflicting rules. A *conflict* occurs whenever we get multiple distinct values for the PC category for key component and a single value for the PC category of the other component. A *conflict* is resolved by applying the utility function and selecting the rule with the highest score. The result is the unconflicting SC rules. For every PC

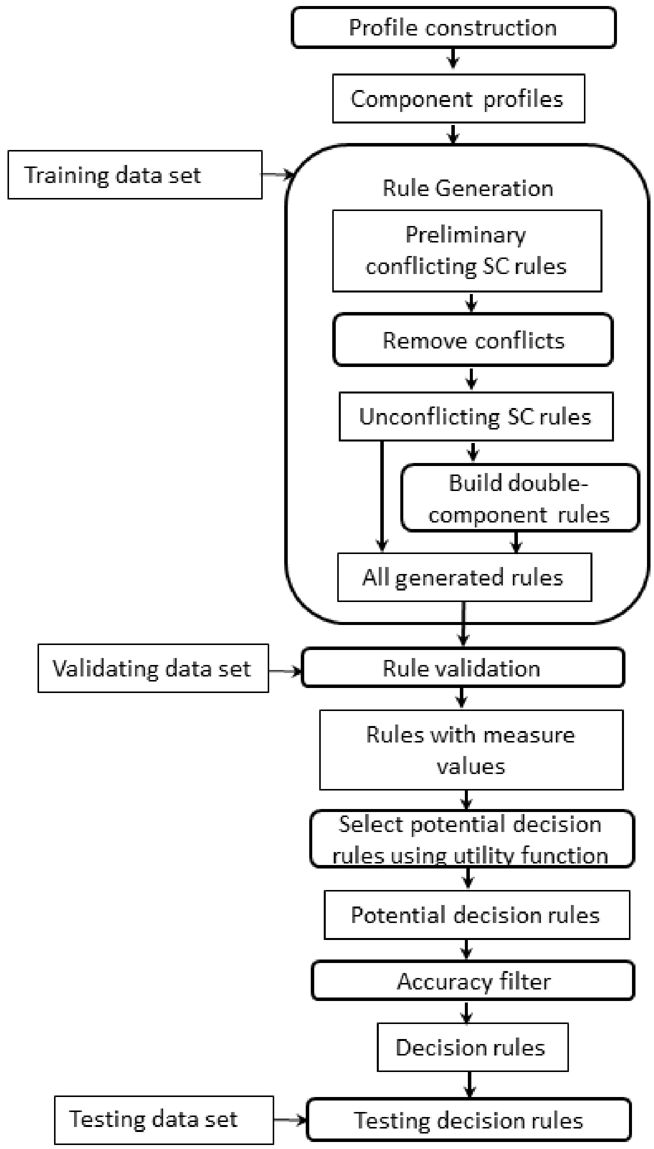


Figure 3.4: Procedure for generating and testing price change rules with KASPER

category for the key component, a group of unconflicting SC rules is used as the basis for building DC rules.

In the validation phase, quality measures are used to determine the quality of each single and double-component rule. After the rules are applied to the validation data set, we get rules with measure values. Then the best single rule and the best double-component rule is chosen according to the utility scores for the rules. Finally, according to the higher utility score among the best SC and DC rules, a single or double-component rule is selected as the potential decision rule. If the scores are equal, the single-component rule is chosen. If the accuracy of the potential decision rule is greater than or equal to a threshold, then the potential decision rule is considered to be a decision rule.

To allow us to assess the effectiveness of our method, we test the decision rules on unseen data. For simplicity, testing is also performed by the KASPER software by applying the decision rules for the key component on the testing data set. For completeness, the testing phase is also shown in Figure 3.4. It would not be used as part of generating decision rules in a production environment.

3.6.1 Profile Construction

For generating store-to-store rules or brand-to-brand rules, we first build store profiles and brand profiles, respectively. A *profile* is a set of data that characterizes an entity. Thus, a *store profile* is a collection of data that characterizes a store. A *brand profile* is defined analogously.

Algorithm 2: Construct initial profiles.

1 `Input: Set of price reports in a database with attributes` $\langle storeId, dateTime, price \rangle$

2 `Output: Initial profile with attributes for every store` $s \in S$ `for every day` $d \in D$ `as a tuple with attributes` $\langle storeId, date, List\langle time, price \rangle \rangle$

3 **for** ***each*** $s \in S$ **do**

4   **for** ***each*** $d \in D$ **do**

5   `select all` $\langle time, price \rangle$ `pairs for store` s `on a day` d `from the database and save them in a hashmap in the form` $\langle storeId, date, List\langle time, price \rangle \rangle$.

Initial profile

For each store, we construct and maintain an *initial profile*, which is the basis for building both store and brand profiles. The data set that is used for our research describes prices at a set of stores with brands, cities, and other attributes. In order to develop the store and brand profiles, we need to access an individual tuple from a large set of data in a database. For efficiency, we select the necessary tuples and attributes from the database and make an initial profile for each store using Algorithm 2. The initial profile for a store $s \in S$ is a $\langle storeId, date, List\langle time, price \rangle \rangle$ tuple. Let D be the set of all days in the data set. For each day $d \in D$ for each store $s \in S$, the necessary tuples with required attributes are selected from the database and stored in a hashmap, as shown in step 5 of Algorithm 2. We construct the store and brand profiles from the initial profile. Generating rules is faster from profiles than from the database because it avoids calculating the same values repeatedly.

Store profiles

A *store profile* is a description for a particular store $s \in S$ for a particular $d \in D$ that shows the time and price for every report, augmented with the EODPC value, which are included for convenience. The EODPC value is calculated using steps 6 to 9 in Algorithm 3. The value of EODPC is used for generating store-to-store rules. If two consecutive days have a record of a price change between them, then we calculate EODPC based on

Algorithm 3: Construct a store profile.

Input: Set of initial profiles in the format $\langle s, date, List\langle time, price\rangle\rangle$ for every store $s \in S$.
Output: Tuples in the format $\langle s, date, List\langle time, price, EODPC\rangle\rangle$
for ***each*** $s \in S$ **do**
  Initialize list L to empty
  for ***each*** *time value* t *in* $List\langle time, price\rangle\rangle$ *on a unique day* d **do**
    if $(EODP_{d-1} \neq null)$ **then**
      $EODPC_d = EODP_d - EODP_{d-1}$
    else
      $EODPC_d = null$
    $add\langle time, price, EODPC_d\rangle$ to list L
  $add\langle storeId, d, L\rangle$ to the hashmap

the difference in price; otherwise the value of EODPC is null.

Brand profiles

As a step towards making brand-to-brand price change rules, a brand profile is constructed using Algorithm 4. A brand profile consists of a brand id, a date, and an EODPC. The EODP for a brand is calculated in step 5 in Algorithm 4 as the average EODP of all stores in the same brand that have reported prices on that specific date. The assumption is that the stores that report are representatives of all stores in the brand, which may not be true for particular days. Then EODPC is calculated in step 7 from the difference between EODP on a specific day and the previous day.

3.6.2 Rule Generation

As explained in Section 3.3, we use the term "component" to refer to either a store or a brand. Here we describe the process of generating single and double-component rules; this process can be applied to generate either store-to-store or brand-to-brand rules. We first define a component-to-component EODPC relation using price change categories. A rule generated from such a relation describes the correspondence between changes in

Algorithm 4: Construct a brand profile.

Input: set of tuples in the format $\langle s, date, List\langle time, price\rangle\rangle$ for store $s \in S$, set of brands B, set of days D

Output: For every brand $b \in B$, tuples in the format $\langle brandId, date, EODPC\rangle$

for ***each*** $b \in B$ **do**

  Let $\{s_1, s_2, \ldots, s_n\}$ be the subset of stores in S with brand b **for** ***each*** $d \in D$ **do**

    $EODP_d(B) = Avg\{EODP_d(s_1), EODP_d(s_2), \ldots, EODP_d(s_n)\}$

    if $(EODP_{d-1}(B) \neq null)$ **then**

      $EODPC_d(B) = EODP_d(B) - EODP_{d-1}(B)$

    else

      $EODPC_d(B) = null$

    $add\langle b, d, EODPC_d(B)\rangle$ to the hashmap

EODP at the key component in comparison to changes at other components. Table 3.6 shows the joint frequency function f of price change categories between another component and a key component. For example, $f_{(u,v)}$ shows the joint frequency if the EODPC of the other component is in category u and EODPC of the key component is in category v.

Characteristics:

All PC categories for a component are disjoint from each other. Thus, there is no chance of the occurrence of two categories of price changes for a component on a day, as shown in Figure 3.5. Here a black circle indicates CAT1 for component c_1 and a grey circle indicates a specific PC category for component c_k.

If the other component changes its price in a category then KASPER generates rules

Table 3.6: Joint frequency function f for price change categories.

		PC category of key component		
		$v = 1$	$\ldots$	$v = m$
PC category of other component	$u = 1$	f[CAT1][CAT1]	$\ldots$	f[CAT1][CATm]
	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$u = m$	f[CATm][CAT1]	$\ldots$	f[CATm][CATm]

for all categories of price changes for the key component and calculates the joint frequency in every case.

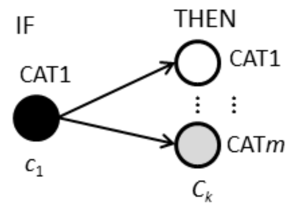


Figure 3.5: Price change category for other component to key component.

Figure 3.5 shows that when PC category for a component, c_1 is in CAT1, then the price change for the key component c_k will be in one of the price change categories CAT1 to CATm, but there is no possibility of more than one PC category occurring for c_1 or c_k on the same day.

Since PC categories are distinct, any useful conjunctive rule with two clauses must refer to two distinct stores or brands. For example, a condition such as "If $PC_{cat}(s1)$ = CAT2 and $PC_{cat}(s2)$ = CAT2" is potentially useful, but one such as "If $PC_{cat}(s1)$ = CAT2 and $PC_{cat}(s1)$ = CAT3" is never useful.

Single-component rule generation:

Single-component rules are generated using Algorithm 5 and stored in the database. For a key component c_k and every other component $c \in C - \{c_k\}$, KASPER calculates the joint frequency of the PC categories of the other component and the key component using step 12. Finally, KASPER generates rules using steps 16 to 18 if the joint frequency is greater than or equal to a threshold.

Filtering:

After generating single-component rules, we filter these rules for possible conflicts.

Algorithm 5: Generate single-component rules

Input: key component c_k, set of components C, set of tuples in hashmap (training data set), set of unique days in the hashmap D

Output: For every component $c_i \in C - \{c_k\}$, set of single component rules $R_{k,i}$

for ***each*** $c_i \in C - \{c_k\}$ **do**

    $R_{k,i} = \phi$ initialize frequency table f to zeros

    for ***each*** $d \in D$ **do**

        $p_i = Retrieve(hashmap, c_i, d)$

        if $(p_i \neq null)$ **then**

            $p_k = Retrieve(hashmap, c_k, d)$

            if $(p_k \neq null)$ **then**

                $u = Classify(p_i)$

                $v = Classify(p_k)$

                $f[u][v]++$

    for ***each*** $u = 1$ *to* m **do**

        for ***each*** $v = 1$ *to* m **do**

            if $f[u][v] \geq threshold$ **then**

                generate single-component rule r in the format

                IF $EODPC_d$ of other component c_i is u THEN $EODPC_d$ of key component c_k is v.

                add single-component rule r to $R_{k,i}$.

A *conflict* occurs if there are multiple rules for the same key component with the same condition but different conclusions.

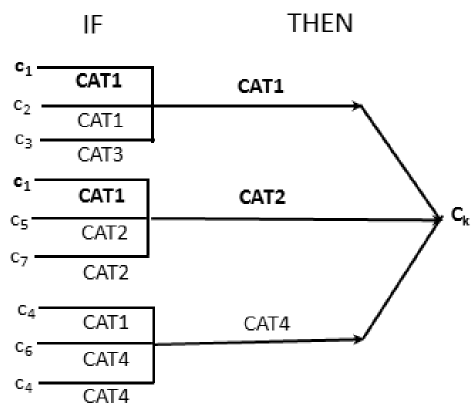


Figure 3.6: Tree structure for conflicting single-component rules.

Figure 3.6 shows an example of a conflict. Three groups of rules exist for three distinct price change categories for c_k, namely CAT1, CAT2, and CAT4. Details about the rules are described below:

For key component c_k, $EODPC_d$ category CAT1 PC category of the other components are:

(a) $EODPC_d$ of other component c_1 is in category CAT1

(b) $EODPC_d$ of other component c_2 is in category CAT1

(c) $EODPC_d$ of other component c_3 is in category CAT3

If the category is CAT2, the second group of responsible PC category of other components are:

(a) $EODPC_d$ of other component c_1 is in category CAT1

(b) $EODPC_d$ of other component c_5 is in category CAT2

(c) $EODPC_d$ of other component c_7 is in category CAT2

$EODPC_d$ category CAT4, the third group of responsible PC category of other components are:

(a) $EODPC_d$ of other component c_4 is in category CAT1

(b) $EODPC_d$ of other component c_6 is in category CAT4

(c) $EODPC_d$ of other component c_4 is in category CAT4

Conflict: In the first and second groups, we can see two cases where the condition is "IF $EODPC_d$ of component c_1 is in category CAT1". Then $EODPC_d$ of key component, c_k can be in category CAT1 or CAT2.

We select the best rule among the conflicting rules using the utility function and resolve ties arbitrarily.

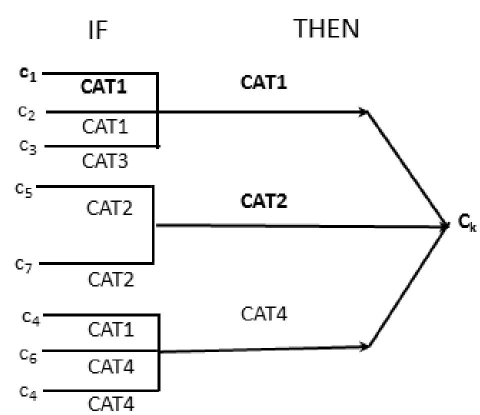


Figure 3.7: Tree structure for unconflicting single-component rules.

After removing any conflicts, we obtain the scenario shown in Figure 3.7, where the first and third groups consist of three rules and the second group consists of two rules.

Building double-component rules

A double-component rule is formed by combining two single-component rules for two separate other components. Sometimes rules can be generated for the same other component for two PC categories. For a particular key component, if there are multiple SC rules with the same other component in their conditions, we choose the SC rule with the higher score on the validate data. In Figure 3.7, two rules for component c_4, one for CAT1 and the other for CAT4. From the third group, we build a DC rule from the c_4 and c_6 rules. One of two rules generated for c_4 is selected by the utility function. If SC rules are generated for n components, then the number of DC rules will be $\frac{n(n-1)}{2}$.

The DC rules made from the first group of SC rules Figure 3.7 are shown below:

(a) IF $EODPC_d$ of c_1 is in CAT1 and $EODPC_d$ of c_2 is in CAT1 THEN $EODPC_d$ of c_k will be in CAT1

(b) IF $EODPC_d$ of c_1 is in CAT1 and $EODPC_d$ of c_3 is in CAT1 THEN $EODPC_d$ of c_k will be in CAT1

(c) IF $EODPC_d$ of c_2 is in CAT1 and $EODPC_d$ of c_3 is in CAT3 THEN $EODPC_d$ of c_k will be in CAT1

The only DC rule made from the second group of SC rules is shown below:

(a) IF $EODPC_d$ of c_5 is in CAT2 and $EODPC_d$ of c_7 is in CAT2 THEN $EODPC_d$ of c_k will be in CAT2

The only DC rule made from the third group of SC rules is shown below:

(a) IF $EODPC_d$ of c_4 is in CAT4 and $EODPC_d$ of c_6 is in CAT4 THEN $EODPC_d$ of c_k will be in CAT4

3.6.3 Rule Validation

KASPER employs seventeen measures to assess the quality of a rule on the validation data set. To measure the quality of a rule, KASPER calculates the four independent measures (TP, FP, FN, and TN), as described in Section 2.5, for a price change rule by straightforward counting.

Table 3.7 shows that if a positive instance is classified correctly then the true positive (TP) measure is incremented, if a positive instance is classified incorrectly then the false

Table 3.7: Independent measures for rule validation.

		Actual	
		Positive	**Negative**
Prediction	**Positive**	TP++	FP++
	Negative	FN++	TN++

positive (FP) measure is incremented, if a negative instance is classified incorrectly then the false negative (FN) measure is incremented, and if a negative instance is classified correctly then the true negative (TN) measure is incremented.

For each rule $r \in R$, the seventeen measures are calculated using Algorithm 6. This algorithm accepts as input all single and double-component rules. Steps 6 to 13 calculate the values of independent measures and steps 16 to 22 calculate the remaining thirteen measures.

3.6.4 Decision Rule Selection

This section describes the methodology for selecting the best single-component rule and best double-component rule for each price change category for the key component. For each price change category for the key component, the quality of the single-component and double-component rule is measured by seventeen quality measures shown in Table 3.3. The score for each rule is calculated by the utility function, as shown in Table 3.4.

Example:

There is an example to choose the best single and double-component rules among a group of single and double-component rules.

Single-component rule selection:

Consider three single-component rules that are used for making double-component rules:

Algorithm 6: Rule validation.

Input: set of rules R, set of days D (days in validating data set)
Output: For every rule $r \in R$ $\langle r, measures \rangle$
for each $r \in R$ **do**
    $TP = FP = FN = TN = 0$
    for each $d \in D$ **do**
        if positive instance classified correctly **then**
            $TP{+}{+}$
        else if positive instance classified incorrectly **then**
            $FP{+}{+}$
        else if negative instance classified incorrectly **then**
            $FN{+}{+}$
        else
            $TN{+}{+}$
        end if
    end for
    $TPR = \frac{TP}{TP+FN}$, $FPR = \frac{FP}{FP+TN}$
    $FNR = \frac{FN}{FN+TP}$, $TNR = \frac{TN}{TN+FP}$
    $P = \frac{TP}{TP+FP}$, $AC = \frac{TP+TN}{TP+TN+FP+FN}$
    $E = \frac{FP+FN}{TP+TN+FP+FN}$, $FDR = \frac{FP}{TP+FP}$
    $F = 2 \times \frac{P \times TPR}{P+TPR}$, $G = \sqrt{P \times TPR}$
    $LR{+} = \frac{TPR}{FPR}$, $LR{-} = \frac{FNR}{TNR}$
    $DOR = \frac{LR+}{LR-}$
    add $\langle r, measures \rangle$ to database
end for

s-rule1: "IF $EODPC_d$ of c_1 is in CAT1 THEN $EODPC_d$ of c_k will be in CAT1".

s-rule2: "IF $EODPC_d$ of c_2 is in CAT2 THEN $EODPC_d$ of c_k will be in CAT2".

s-rule3: "IF $EODPC_d$ of c_3 is in CAT4 THEN $EODPC_d$ of c_k will be in CAT4".

Table 3.8: Select one single-component rule using the utility function.

Measures	s-rule1	s-rule2	s-rule3	Parameters of utility function	s-rule1	s-rule2	s-rule3
TP	797	804	750	isMax(TP)	0	1	0
FP	10	11	11	isMax(-FP)	1	0	0
FN	98	91	95	isMax(-FN)	0	1	0
TN	186	185	235	isMax(TN)	0	0	1
P	98.76	98.65	98.55	isMax(P)	1	0	0
FDR	0.01	0.01	0.14	isMax(-FDR)	1	1	0
TPR	89.05	89.83	88.76	isMax(TPR)	0	1	0
FPR	5.1	5.61	4.47	isMax(-FPR)	0	0	1
F	93.65	94.04	93.73	isMax(F)	0	1	0
G	93.78	94.14	93.87	isMax(G)	0	1	0
AC	90.1	90.65	90.28	isMax(AC)	0	1	0
E	9.9	9.35	9.71	isMax(-E)	0	1	0
TNR	94.9	94.39	95.52	isMax(TNR)	0	0	1
FNR	10.95	10.17	11.24	isMax(-FNR)	0	1	0
LR+	17.45	16.01	22.25	isMax(LR+)	0	0	1
LR-	0.12	0.11	0.11	isMax(-(LR-))	0	1	1
DOR	151.27	148.59	202.27	isMax(DOR)	0	0	1
				Score	3	10	6

The score for each of these single-component rules is calculated using the utility function, as shown in Table 3.4. Here a rule is given a '1' value for a particular measure if it has the maximum value for that measure among the three competitive rules. From the first row of Table 3.8, we can see that s-rule2 has a '1' value, which means s-rule2 has the maximum value for TP among the three rules. A value of '1' or '0' is also determined for the other sixteen parameters in a similar way. The score for a rule is calculated by counting the number of '1' values for the seventeen measures. Thus, the calculated score for s-rule2 is 10, which is the highest score among three rules.

Double-component rule selection:

Double-component rules are selected similarly to single-component rules. For a specific PC category for key component c_k consider the following three double-component rules:

d-rule1: "IF $EODPC_d$ of c_1 is in CAT1 and $EODPC_d$ of c_2 is in CAT1 THEN $EODPC_d$ of c_k will be in CAT1."

d-rule2: "IF $EODPC_d$ of c_3 is in CAT2 and $EODPC_d$ of c_4 is in CAT2 THEN $EODPC_d$ of c_k will be in CAT2."

d-rule3: "IF $EODPC_d$ of c_5 is in CAT4 and $EODPC_d$ of c_6 is in CAT4 THEN $EODPC_d$ of c_k will be in CAT4."

Table 3.9: Selectin of one double-component rule using the utility function.

Measures	d-rule1	d-rule2	d-rule3	Parameters of customized utility function	d-rule1	d-rule2	d-rule3
TP	763	705	703	isMax(TP)	1	0	0
FP	8	6	6	isMax(-FP)	0	1	1
FN	132	190	192	isMax(-FN)	1	0	0
TN	188	190	190	isMax(TN)	0	1	1
P	98.96	99.16	99.15	isMax(P)	0	1	0
FDR	1.04	0.84	0.85	isMax(-FDR)	0	1	0
TPR	85.25	78.77	78.55	isMax(TPR)	1	0	0
FPR	4.08	3.06	3.06	isMax(-FPR)	0	1	1
F	91.6	87.8	87.66	isMax(F)	1	0	0
G	91.85	88.38	88.25	isMax(G)	1	0	0
AC	87.17	82.03	81.85	isMax(AC)	1	0	0
E	12.83	17.97	18.15	isMax(-E)	1	0	0
TNR	95.92	96.94	96.94	isMax(TNR)	0	1	1
FNR	14.75	21.23	21.45	isMax(-FNR)	1	0	0
LR+	20.89	25.73	25.66	isMax(LR+)	0	1	0
LR-	0.15	0.22	0.22	isMax(-(LR-))	1	0	0
DOR	135.84	117.5	115.95	isMax(DOR)	1	0	0
				Score	10	7	4

From Table 3.9, we can see that d-rule1 has the highest score among the three d-rules and thus it is selected. Either the single-component or the double-component rule is selected as the potential decision rule for each PC category for the key component. If one of them has a higher score, it is selected. If both of them have the same score, then the single-component rule is chosen. From Table 3.8, we can see that s-rule2 has the highest score (10) among the single-component rules. Similarly, from Table 3.9, d-rule1 has the highest score (10) among the double-component rules. Since there is a tie, s-rule2 is selected as the potential decision rule.

If the accuracy of a potential decision rule is greater than or equal to 80%, then the winning rule is considered to be a decision rule. As the accuracy of s-rule2 is 90.65%, this rule is a decision rule.

3.7 The Rank-Based Method

The rank-based method uses a predefined number of selected single-component rules for building double-component rules, whereas the previous complete-component method uses all of the single-component rules. It was hypothesized that the rank-based method might give similar results but run substantially faster. From profile construction to decision rule generation, every step of the rank-based method is the same as the previous complete-component method except for the training phase.

The training phase is shown in Figure 3.8. Recall that for every PC category for the key component, a group of unconflicting SC rules is used as the basis for building DC rules. The overall difference is that with the rank-based method, only a subset of the unconflicting SC rules are used to build DC rules instead of all unconflicting rules. In more detail, the unconflicting SC rules are first sorted by their FPR values to give the sorted SC rules. A number of rules with the highest FPR values are selected. Suppose n is the total number of unconflicting SC rules and a is the threshold. If n is higher than a, then only a rules with the highest FPR values are selected; otherwise, all SC rules are selected. The DC rules are then constructed from the selected SC rules. With the rank-based method,

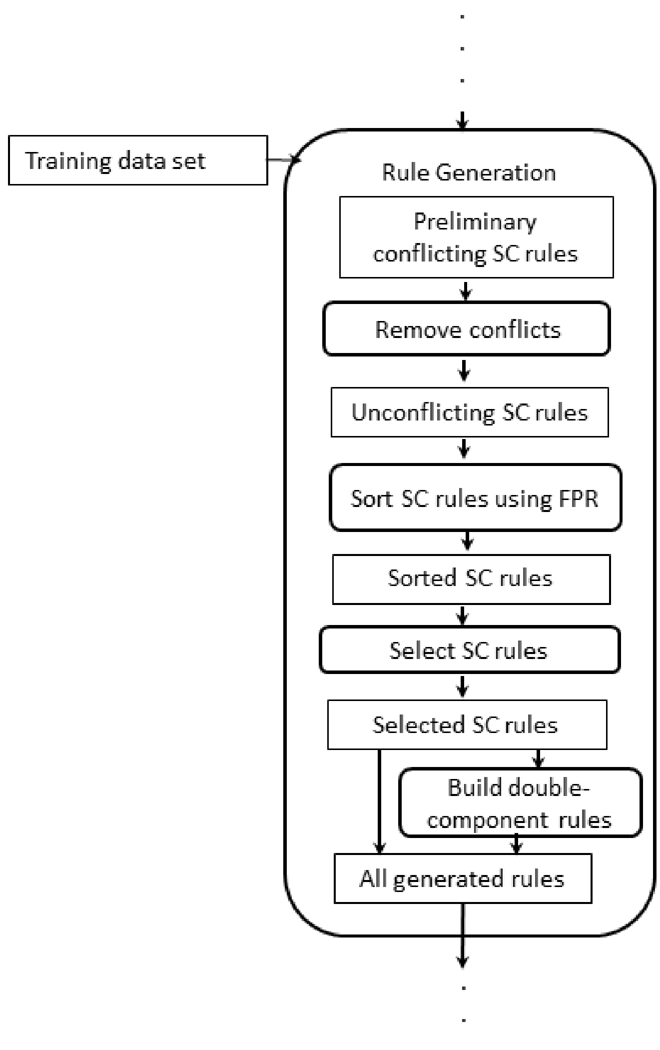


Figure 3.8: Rank-based method for generating and testing price change rules.

the group consists of a selected rules, whereas with the complete-component method, the group consists of n unconflicting rules.

Chapter 4

EXPERIMENTAL RESULTS AND COMPARISON

This chapter presents experimental results concerning KASPER, using data from several cities, brands, and stores. Section 4.1 describes the experimental setup, including the data set and the environment used for the experiments. Section 4.2 shows sample generated rules and quality measures for each rule. Section 4.3 discusses the behaviour of the average end of day price changes (EODPC) for four cities. Section 4.4 covers the overall quality of store-to-store and brand-to-brand directional and categorical rules for four cities. Section 4.5 gives detailed explanation of frequency, accuracy, precision, true positive rate, and F-measure for brand-based store-to-store rules. Section 4.6 analyzes the mean and meadian number of potential competitors for several distances. Section 4.7 evaluates the heuristic rank-based variation of KASPER on stores for two brands of two cities. Finally, Section 4.8 compares our research to other previous research.

4.1 Experimental Setup

Experiments were conducted on a historical data set of motor fuel prices for four cities. The data covered period of five years and four months for all cities. Besides price reports, information is provided about the stores in each city and the brand of each store. The *profile* of a price report consists of eight attributes: city id, brand id, store id, latitude, longitude, fuel type, dateTime, and price.

An evaluation data set was created by selecting data only for stores with at least 100 reports for each training, validation, and testing phase, and brands with at least 5 such stores. This data set consists of 9 brands and 189 stores for City1, 10 brands and 177 stores for City2, 9 brands and 317 stores for City3, and 16 brands and 538 stores for City4, as shown in Table 4.1. The 100 series of brands is used for brands in City1, the 200 series for City2, and so forth. The number of tuples in the evaluation data set is 1,091,270 for City1, 1,523,353 for City2, 1,435,548 for City3, and 1,802,016 for City4. For this research, we use the regular ('A') type of motor fuel; we do not consider high-octane or diesel fuel.

When rules are being made with respect to potential competitors, any store with at least one competitor is eligible to act as a key store.

The experiments were conducted on a personal computer with an Intel(R) Core(TM) i7 and 8.00 GB RAM, using software implemented in the Java programming language on the Netbeans IDE 7.4 platform with databases stored in Microsoft SQL Server Management Studio 2012.

Table 4.1: List of brands and number of stores for City1, City2, City3, and City4.

City1		City2		City3		City4	
Brand	# of stores	Brand	# of stores	Brand	# of stores	Brand	# of stores
B101	40	B201	11	B301	67	B401	71
B102	9	B202	36	B302	30	B402	8
B103	11	B203	8	B303	67	B403	166
B104	26	B204	5	B304	5	B404	9
B105	5	B205	8	B305	5	B405	44
B106	15	B206	28	B306	19	B406	6
B107	26	B207	15	B307	64	B407	25
B108	47	B208	21	B308	54	B408	11
B109	10	B209	16	B309	6	B409	7
		B210	29			B410	23
						B411	5
						B412	88
						B413	23
						B414	7
						B415	36
						B416	9
	Total=189		Total=177		Total=317		Total=538

Table 4.2: Duration of each phase.

Training data set	**Validating data set**	**Testing data set**
01 January 2010 to 31 December 2011	01 January 2012 to 31 December 2013	01 January 2014 to 30 April 2015

The data set for each city is divided into three sections for training, validation, and testing, as shown in Table 4.2. The three phases are named in this manner to be consistent with previous research [8].

KASPER generates rules for a key store or a key brand and it is capable of generating directional or categorical rules. Directional rules use two categories of price changes and categorical rules use six categories of price changes. As well, in both cases, there is one category representing "no change". This "no change" category means that the price difference is exactly zero. For convenience of reference, Table 3.2, which shows the possible directions of price change and the possible categories of price change, is reproduced here as Table 4.3.

Table 4.3: Price changes for directional and categorical rules.

(a) Directions of price change ($z = 2$).

Direction	**Meaning**	**Price Change (PC)**
		City1, City2, City3, and City4
NC	No Change	$= 0$
INC	Increase	> 0
DEC	Decrease	< 0

(b) Categories of price change ($z = 6$).

Category	**Price Change (PC)**	
	City1, City2, and City4	City3
CAT0	$= 0$	$= 0$
CAT1	> 0 and ≤ 0.05	> 0 and ≤ 0.02
CAT2	> 0.05 and ≤ 0.10	> 0.02 and ≤ 0.04
CAT3	> 0.10	> 0.04
CAT4	< 0 and ≥ -0.05	< 0 and ≥ -0.02
CAT5	< -0.05 and ≥ -0.10	< -0.02 and ≥ -0.04
CAT6	< -0.10	< -0.04

We used the testing data set mentioned in Table 4.2 to evaluate the rules generated by the KASPER system. Among other results, we report in the following sections on the applications of all measures listed in Table 2.6 to evaluate the quality of store-to-store and brand-to-brand directional and categorical rules for City1, City2, City3, and City4.

4.2 Sample Generated Rules

Before presenting an evaluation of the KASPER system, we first show some examples of the kinds of rules generated by the system. Twelve sample rules are shown in Table 4.4. Rules 1 and 2 are brand-to-brand directional rules and rules 3 and 4 are brand-to-brand categorical rules. Rules 5 to 8 are store-to-store directional rules and rules 9 to 12 are store-to-store categorical rules. The table includes single and double-component brand-to-brand rules and store-to-store rules. Rules 1, 3, 5, 6, 9, and 10 are single component rules and rules 2, 4, 7, 8, 11, and 12 are double-component rules. To illustrate further, the detailed meanings of example single and double-component rules from Table 4.4 are given below:

Meaning of an example single-component rule:

Rule 1. "IF $EODPC_d$ of B103 is in DEC THEN $EODPC_d$ of B107 (key brand id) will be in DEC"

Meaning of an example double-component rule:

Rule 7."IF $EODPC_d$ of S1137 is in INC and S1165 is in INC THEN $EODPC_d$ of S1167 (key store id) will be in INC"

Table 4.4: Twelve sample generated rules.

		Rule No.	City	IF $EODPC_d$ is	THEN $EODPC_d$ is
Brand	Directional	1	City1	B103 DEC	B107 DEC
		2	City2	B208 DEC B203 DEC	B205 DEC
	Categorical	3	City1	B103 CAT4	B107 CAT4
		4	City2	B206 CAT4 B208 CAT4	B209 CAT4
Store	Directional	5	City3	S3122 DEC	S3125 DEC
		6	City4	S4491 DEC	S4493 DEC
		7	City1	S1137 INC S1165 INC	S1167 INC
		8	City2	S2147 INC S2140 INC	S2142 INC
	Categorical	9	City1	S1116 CAT4	S1113 CAT4
		10	City2	S2157 CAT4	S2156 CAT4
		11	City1	S1029 CAT3 S1006 CAT3	S1014 CAT3
		12	City2	S2147 CAT3 S2140 CAT3	S2142 CAT3

Recall from Section 2.5 that we use seventeen measures to evaluate the quality of rules. We now provide a detailed explanation of each measure with respect to the second rule from Table 4.4. First consider the 2×2 contingency table shown in Table 4.5, which records the results of applying rule 2 to City2 data during the testing phase. Table 4.5 shows that positive instances are classified correctly by this rule 378 times, i.e. the true positive (TP) value is 378. A positive instance is incorrectly classified once, and thus, the false positive (FP) value is 2. Similarly, negative instances are classified incorrectly

Table 4.5: 2×2 contingency table for rule 1.

		Actual	
		Positive	**Negative**
Prediction	**Positive**	TP = 378	FP = 2
	Negative	FN = 16	TN = 76

16 times (FN is 16) and negative instances are classified correctly 76 times (TN is 76). The results for the other 13 measures, with respect to applying rule 2 to City2 data, are calculated from these four measures. Table 4.6 shows the values for all measures for all 12 example rules with respect to the data relevant to the rules.

4.3 Average EODPC for Four Cities

This section characterizes the manner in which prices changed during five years and four months for City1, City2, City3, and City4. In particular, we study the average EODPC. This average is computed as the mean of the EODPC at all stores that have non-null EODPC values for that day. For all cities, we describe the price changes according to chronological order of days, and sorted order.

Table 4.7 gives information about the percentage of average EODPC values in each of three categories (no price change, increased price, and decreased price) for four cities and the greatest magnitudes of price changes for two directions (increase and decrease).

From Table 4.7, we can see that the average EODPC values for City1 and City2 act similarly. The percentages are similar for each of the three categories. The largest price increases, which are 0.33 and 0.32 for City1 and City2, respectively, as well as similar largest price decreases, which are 0.09 and 0.07, respectively, are both similar. The average EODPC values for City3 and City4 act similar, but differently from those for City1 and City2. The percentages of no price change and price decreases for City3 and City4 are approximately 35-40% different from those for City1 and City2.

Table 4.6: Values of quality measures for sample generated rules.

(a) Values of quality measures for sample generated rules.

Rule No.	TP	FP	FN	TN	FDR (%)	P (%)	AC (%)	E (%)	TPR (%)
1	396	12	12	52	2.94	97.06	94.92	5.08	97.06
2	378	2	16	76	0.53	99.47	96.19	3.81	95.94
3	317	33	58	64	9.43	90.57	80.72	19.28	84.53
4	340	15	29	88	4.23	95.77	90.68	9.32	92.14
5	59	2	3	383	3.28	96.72	98.88	1.12	95.16
6	24	2	3	171	7.69	92.31	97.50	2.50	88.89
7	39	0	7	390	0.00	100.00	98.39	1.61	84.78
8	46	0	4	414	0.00	100.00	99.14	0.86	92.00
9	208	18	23	223	7.96	92.04	91.31	8.69	90.04
10	272	25	30	145	8.42	91.58	88.35	11.65	90.07
11	38	0	4	415	0.00	100.00	99.12	0.88	90.48
12	42	0	2	420	0.00	100.00	99.57	0.43	95.45

(b) Values of quality measures for sample generated rules (continued).

Rule No.	TNR (%)	FPR (%)	FNR (%)	F	G	LR+	LR-	DOR
1	81.25	18.75	2.94	97.06	97.06	5.18	0.04	143.00
2	97.44	2.56	4.06	97.67	97.69	37.42	0.04	897.75
3	65.98	34.02	15.47	87.45	87.50	2.48	0.23	10.60
4	85.44	14.56	7.86	93.92	93.94	6.33	0.09	68.78
5	99.48	0.52	4.84	95.93	95.94	183.19	0.05	3766.17
6	98.84	1.16	11.11	90.57	90.58	76.89	0.11	684.00
7	100.00	0.00	15.22	91.76	92.08	84.78	0.15	557.14
8	100.00	0.00	8.00	95.83	95.92	92.00	0.08	1150.00
9	92.53	7.47	9.96	91.03	91.03	12.06	0.11	112.04
10	85.29	14.71	9.93	90.82	90.82	6.12	0.12	52.59
11	100.00	0.00	9.52	95.00	95.12	90.48	0.10	950.00
12	100.00	0.00	4.55	97.67	97.70	95.45	0.05	2100.00

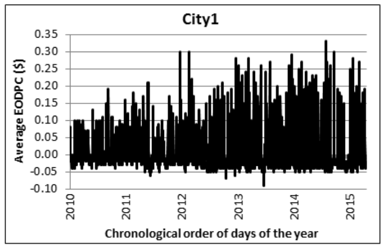
Table 4.7: Distribution of price changes and extrema of price changes for four cities.

City	No Change (%)	Increase (%)	Decrease (%)	Largest increase ($)	Largest decrease ($)
City1	5.43	20.37	74.20	0.33	0.09
City2	5.48	18.05	76.47	0.32	0.07
City3	41.21	25.49	33.30	0.18	0.06
City4	46.43	23.01	30.56	0.06	0.04

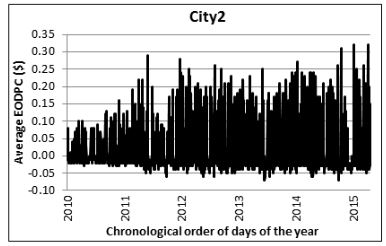
Figure 4.1 shows the average EODPC for each day in chronological order for the four cities as well as the average EODPC for all days sorted in ascending order for the four cities. The graphs in Figure 4.1a to Figure 4.1d show the behaviour of the average EODPC for City1, City2, City3, and City4. Days in chronological order are plotted on the X-axis and the average EODPC for each day is plotted on the Y-axis. From Figure 4.1a and Figure 4.1b, we can see that many price increases are more than 10 cents, with some higher than 30 cents, for both City1 and City2. In contrast, Figure 4.1c and Figure 4.1d show that few price increases are more than 10 cents for City3 and City4. Clearly, the size of price increases is larger in City1 and City2 than in City3 and City4.

The graphs in Figure 4.1e to Figure 4.1h represent the same data that was shown in Figure 4.1a to Figure 4.1d, respectively, but with the average EODPC values sorted in ascending order. The magnitudes of the price changes for City1 and City2 vary more than those of City3 and City4. From Table 4.7, we can also see that the number of days with no price changes is larger for City3 and City4 than for City1 and City2.

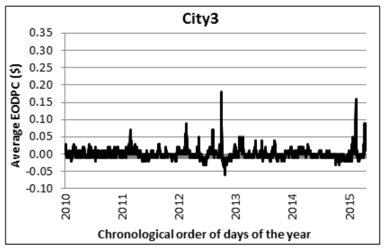
Overall, from the table and figures, we conclude that City1 and City2 show more variations in price changes than City3 and City4. Thus, for our thesis, we consider City1 and city2 as the high variability cities and City3 and City4 as the low variability cities.



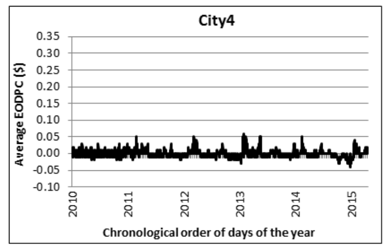
(a) Average EODPC during five years and four months for City1.



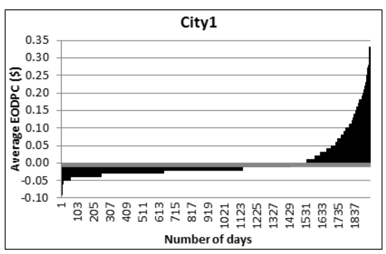
(b) Average EODPC during five years and four months for City2.



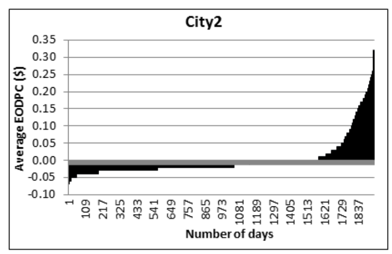
(c) Average EODPC during five years and four months for City3.



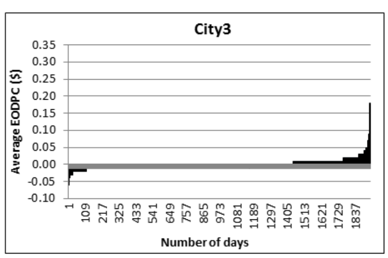
(d) Average EODPC during five years and four months for City4.



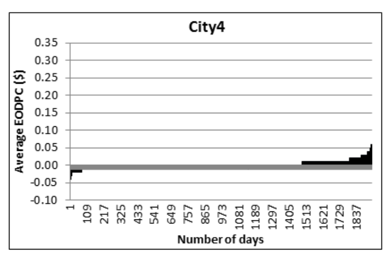
(e) Average EODPC during five years and four months for City1 (sorted).



(f) Average EODPC during five years and four months for City2 (sorted).



(g) Average EODPC during five years and four months for City3 (sorted).



(h) Average EODPC during five years and four months for City4 (sorted).

Figure 4.1: Average EODPC for City1, City2, City3, and City4.

4.4 Overview of Results

This section gives a detailed explanation of the coverage for brand-based store-to-store rules for four cities, and analyzes the coverage and the percentage of good rules for sets of rules generated by KASPER.

Coverage for brand-based store-to-store rules

Recall from Section 3.1 coverage is the fraction of possible combinations of stores and price change categories for which we have decision rules. Table 4.8 presents statistics concerning brand-based store-to-store directional rules generated by KASPER for the four cities. The table shows the number of rules at each stage of the process shown in Figure 3.4. We will describe the results for City1 in detail; the results for the other cities are organized in the same way. The number of SC rules generated for City1 is 9,241. Recall from Chapter 3 that the number of DC rules is $\frac{n(n-1)}{2}$, where n is the number of SC rules. We can calculate $\frac{9{,}241(9{,}241-1)}{2} = 42{,}693{,}420$. However, the number of DC rules is 139,974, which is not the same because the number of DC rules here is the summation of all generated DC rules for all brand-based stores for two categories of price change for a city. The statement "the number of DC rules will be $\frac{n(n-1)}{2}$, where n is the number of SC rules" is true for every PC category for a key store.

The total number of stores for City1 is 189 and thus, the maximum possible number of directional rules is $189 \times 2 = 378$, if every store makes rules for two categories (increase and decrease) of price changes. From Table 4.8, we can see that the total number of potential decision rules for City1 is 372. Of these 372 rules, 201 are chosen to be decision rules after applying an accuracy filter. The coverage is therefore $201/378 = 53.17\%$. The coverage for City2, City3, and City4 is 53.11%, 72.08%, and 46.00% respectively. The coverage for brand-based store-to-store directional rules for City3 is 72.08% because of a higher value of TN, which records how many negative instances have been correctly classified. Accuracy is the proportion to the sum of TP and TN, where TP tells how many positive instances have been correctly classified. So, accuracy can be higher because of a higher value of TP or TN. We use an accuracy filter for selecting actual decision rules from potential decision rules.

Table 4.8: # of brand-based store-to-store directional rules for four cities.

	City1			**City2**		
	SC Rules	DC rules	Total	SC Rules	DC rules	Total
Preliminary conflicting rules	13,768	-	-	13,508	-	-
All generated rules	9,241	139,974	149,215	7,836	98,768	106,604
Rules with measure values	9,241	139,974	149,215	7,836	98,768	106,604
Potential decision rules	320	52	372	304	50	354
Decision rules	179	22	201	167	21	188
Maximum possible rules			378			354
Coverage (%)			53.17			53.11
	City3			**City4**		
	SC Rules	DC rules	Total	SC Rules	DC rules	Total
Preliminary conflicting rules	22,996	-	-	13,296	-	-
All generated rules	19,858	394,230	414,088	11,835	158,729	170,564
Rules with measure values	19,858	394,230	414,088	11,835	158,729	170,564
Potential decision rules	443	153	596	483	236	719
Decision rules	352	105	457	331	164	495
Maximum possible rules			634			1076
Coverage (%)			72.08			46.00

City1 has 189 stores and for six PC categories, the maximum number of possible rules is $189 \times 6 = 1134$. The total number of categorical decision rules for City1 is 364, as shown in Table 4.9. Apparently, there was insufficient data to make rules for the other

Table 4.9: # of brand-based store-to-store categorical rules for four cities.

	City1			**City2**		
	SC Rules	DC rules	Total	SC Rules	DC rules	Total
Preliminary conflicting rules	15,516			21,986		
All generated rules	10,117	141,380	151,497	12,728	184,383	197,111
Rules with measure values	10,117	141,380	151,497	12,728	184,383	197,111
Potential decision rules	400	184	584	448	172	620
Decision rules	271	93	364	306	89	395
Maximum possible rules			1134			1062
Coverage (%)			32.10			37.19

	City3			**City4**		
	SC Rules	DC rules	Total	SC Rules	DC rules	Total
Preliminary conflicting rules	11,740			5,638		
All generated rules	9,848	10,2008	111,856	4,893	35,760	40,653
Rules with measure values	9,848	10,2008	111,856	4,893	35,760	40,653
Potential decision rules	444	211	655	256	161	417
Decision rules	360	103	463	170	77	247
Maximum possible rules			1902			3228
Coverage (%)			24.34			7.65

six cases and sometimes a store never changed its price in a PC category. So, coverage is lower categorical rules for the four cities than directional rules.

The coverage and the percentage of good rules

KASPER generates directional and categorical store-to-store and brand-to-brand rules for four cities. A rule is considered *good* if the precision is greater than or equal to 60% and the accuracy is greater than or equal to 80%. Table 4.10 shows the number of possible rules, number of decision rules, number of good rules, coverage, and percentage of good rules for brand-based and distance-based store-to-store rules, and brand-to-brand rules for four cities. Coverage is the ratio of the number of decision rules and the number of possible rules. So, Coverage (%) $= \frac{\#\ of\ decision\ rules}{\#\ of\ possible\ rules}$. Percentage of good rules is the ratio of the number of good rules and the number of decision rules. So, the percentage of good rules $= \frac{\#\ of\ good\ rules}{\#\ of\ decision\ rules}$. From Table 4.10a, we can see that the percentage of coverage for directional rules for City1 is $\frac{201}{378} = 53.17\%$ and the percentage of good rules is $\frac{145}{201} = 72.14\%$. All values are calculated in similar way.

Table 4.10a and Table 4.10b show that the number of possible rules for distance-based is lower than brand-based because not all stores have competitors within a 2 km distance. KASPER generates distance-based store-to-store rules if every store has at least one potential competitor.

Table 4.10 shows that the percentage of good rules for City1 and City2 is more than 70% for all store-to-store and brand-to-brand directional rules. The brand-to-brand directional rules for both cities are of high quality (approximately 95% of them are good) and coverage is 100%. As well, coverage is more than 50% in all cases for directional rules except the distance-based store-to-store rules for City2. From the percentage of good rules, we can conclude that stores and brands for City1 and City2 are highly responsive to each other for the two categories of price changes.

Table 4.10: Directional and categorical rules for four cities.

(a) Brand-based store-to-store rules.

	Directional				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	378	201	145	53.17	**72.14**
City2	354	188	160	53.11	**85.11**
City3	634	457	90	72.08	19.69
City4	1076	495	70	46.00	14.14
	Categorical				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	1134	364	158	32.10	43.41
City2	1062	395	181	37.19	45.82
City3	1902	463	91	24.34	19.65
City4	3228	247	40	7.65	16.19

(b) Distance-based store-to-store rules ($d = 2$ km).

	Directional				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	342	171	137	50.00	**80.11**
City2	320	137	116	42.81	**84.67**
City3	626	346	75	55.27	21.68
City4	1050	250	56	23.81	22.40
	Categorical				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	1026	292	153	28.46	52.40
City2	960	277	175	28.85	63.18
City3	1878	247	59	13.15	23.89
City4	3150	126	26	4.00	20.63

(c) Brand-to-brand rules.

	Directional				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	18	18	17	100.00	**94.44**
City2	20	20	19	100.00	**95.00**
	Categorical				
City	Possible rules	Decision rules	Good rules	Coverage (%)	Good rules (%)
City1	54	34	15	62.96	44.12
City2	60	42	24	70.00	57.14

The coverage is lower for six categories price changes because of a smaller fraction of price changes. Percentage of coverage and good rules are lower for City3 and City4 and KASPER does not generate any brand-to-brand rules for City3 or City4 because of the low variability of price changes. From Table 4.7, we can also see that more than 40% of the time EODPC is zero for City3 and City4. KASPER does not generate rules with zero values of EODPC.

Overall, KASPER generates good rules for stores and brands where prices fluctuate more, and stores and brands which are highly responsive to each other for changing their prices.

4.5 Detailed Explanation of Measures

This section gives a detailed explanation of the results of applying the frequency, accuracy, precision, TPR, and F-measure measures to brand-based store-to-store rules generated by KASPER for four cities. The results of applying these measures to the distance-based store-to-store rules and brand-to-brand rules give similar patterns, are shown in Appendix B. We show results for 1 km, 1.6 km, 2 km, 3 km, and 4 km radiuses for the distance-based store-to-store rules.

Frequency

Table 4.11 shows the values for the four independent measures (TP, FP, FN, and TN) for brand-based store-to-store rules for the four cities. All of these measures represent simple frequencies of predictions made.

The number of directional rules for City1 is 201, as shown in Table 4.8, and the total number of days in the testing data set is 486. So, if a prediction were made for every day

Table 4.11: Values of independent measures for brand-based store-to-store rules.

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	9,229	9,536	6,702	6,149	8,457	8,920	4,011	2,437
FP	2,442	1,801	9,142	8,591	6,171	6,793	6,323	4,687
FN	3,349	2,167	1,1661	1,1770	8,479	8,714	7,999	6,761
TN	62,368	67,679	128,209	117,948	119,093	137,993	141,748	71,520
Total	77,388	81,183	155,714	144,458	142,200	162,420	160,081	85,405

by every rule, then the maximum possible number of predictions generated by the rules would be $201 \times 486 = 97{,}686$. However, the total frequency is 77,388, as shown in Table 4.11. The observed frequency is lower than the maximum possible number because of the two reasons described below.

The first reason is because of null EODPC values. In our method, if there is no EODP value on a day for a store then the EODPC value is considered a null value on that day. If the key store has an EODPC value on a specific day and the other store does not or the other store has one on a specific day but the key store does not, then that day is not considered for generating rules. The second reason is because of zero EODPC values. If EODPC is zero for a store that means two consecutive days have the same EODP for that store. Because KASPER does not make any rules for price change category zero, any day with a zero EODPC value is not considered for generating rules.

Accuracy

Table 4.14 shows the first quartile and median values for the 13 dependent measures for brand-based store-to-store rules for four cities. Table 4.15 shows the corresponding third quartile and maximum values for these measures. From the first quartile value in Table 4.14a, we see that the accuracy is more than 80% for at least 75% of the directional and categorical rules for all cities. Table 4.14b shows that the median accuracy for half of the directional and categorical rules for City1 and City2 is above 90%. The median accuracy for distance-based store-to-store and brand-to-brand rules are also more than 90%, as shown in Table 4.16 and Table 4.17. Recall from Chapter 2 that the accuracy is the proportion of the total number of true values (correct classifications) to the total number of all values (correct and incorrect classifications). Table 4.12 also shows the lower bound, upper bound, and percentage of outliers for accuracy for brand-based store-to-store rules. From Table 4.12 we can see that the categorical rules for City2 have no outliers. In most cases, the number of outliers is less than 1% of the total number of decision rules.

Table 4.12: LB, UB, and outlier values for accuracy for brand-based store-to-store rules.

	Directional			**Categorical**		
	LB	**UB**	**Outlier (%)**	**LB**	**UB**	**Outlier (%)**
City1	70.50	100	1.99	65.28	100	0.84
City2	82.59	100	2.65	65.74	100	none
City3	73.36	100	none	79.92	100	0.93
City4	69.97	100	0.82	73.47	98.57	4.89

Precision, TPR, and F-measure

Recall from Chapter 1 that the goal of our research is to generate rules with high accuracy and precision. Recall that Precision is a measure that provides information about the number of correct classifications of positive instances in comparison to the number of correct and incorrect classifications of positive instances. From Table 4.14b, we can see that the median precision for directional rules for City1 is 79.17%, which tells us that at least 50 percent of the generated brand-based store-to-store directional rules are acceptably precise. The median precision for rules for City2 is 84.44%. The median precision for distance-based directional rules for City1 and City2 is more than 80%, as shown in Table 4.16. This value is more than 90% for brand-to-brand rules, as shown in Table 4.17.

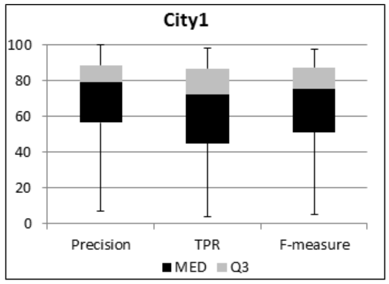
Table 4.13: Number of brand-based store-to-store directional rules in specific precision ranges.

Precision (%)	**City1**	**Precision (%)**	**City2**
[79.17 – 100]	103	[84.44 – 100]	96
[56.36 – 79.17)	48	[72.55 – 84.44)	46
[40 – 56.36)	25	[50 – 72.55)	38
[6.90 – 40)	25	[14.9 – 50)	8
Total rules	200	Total rules	188

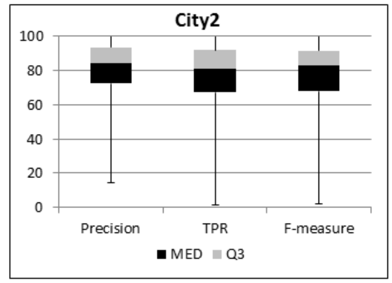
Sometimes the first quartile value is high but the minimum value is low because of a small number of low precision rules. Table 4.13 shows the number of rules in various precision ranges. The minimum value for directional rules for City2 is 14.9, but only 4.25% of the rules are within the 14.9 to 50 range and 20.21% are in the 50 to 72.55 range among 25% of the rules.

Highly precise rules generated by KASPER have also high TPR values. We also mention TPR and F-measure values with precision because sometimes the precision of a rule shows a high value but the TPR is too low. In this case, that rule will not be considered a good rule, as shown in an example in Section 2.5. From Table 4.14b, we can see that the median TPR for directional rules for City1 is 72.22% and that for City2 is 81.25%. The F-measure is another measure that represents a combined value of precision and TPR. The F-measure is calculated as the harmonic mean of the precision and the TPR. The median F-measure for directional rules for City1 and City2 also gives a higher value which is 75% and 82.98% for City1 and City2 respectively.

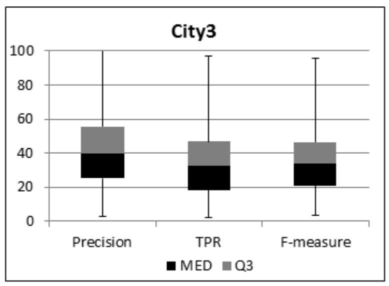
Figure 4.2 shows a graphical representation of the minimum value, first quartile, median, third quartile, and maximum value for precision, TPR, and F-measure for brand-based store-to-store directional rules for four cities by boxplot and whiskers. The lower horizontal strike gives information about the minimum value and the upper horizontal strike shows the maximum value for all rules for a city. The black part of the boxplot



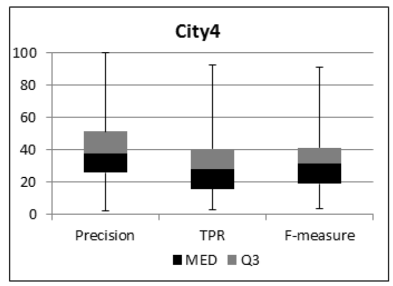
(a) Directional (City1)



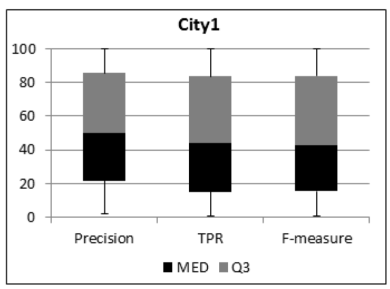
(b) Directional (City2)



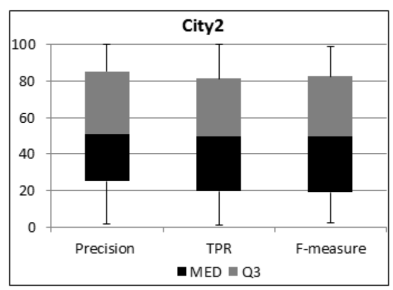
(c) Directional (City3)



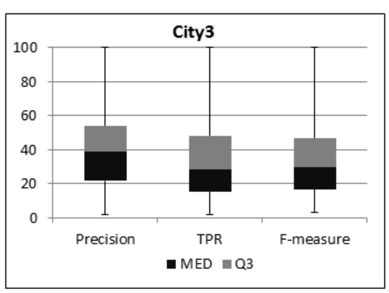
(d) Directional (City4)



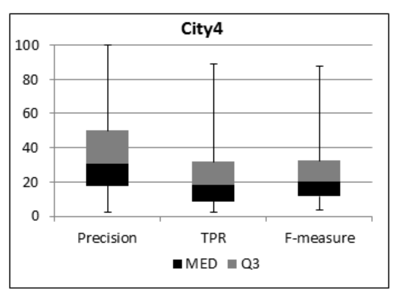
(e) Categorical (City1)



(f) Categorical (City2)



(g) Categorical (City3)



(h) Categorical (City4)

Figure 4.2: Analysis of precision, TPR, and F-measure values for brand-based store-to-store rules.

covers the area from the first quartile to the median value of the precision, TPR, and F-measure for all rules for the city. Similarly, the grey area covers the part from the median to the third quartile. All values here are percentages. From the distribution shown in Figure 4.2, we can see that the median precision, TPR, and F-measure values for directional rules are more than 70% for City1 and 80% for City2. From Table 4.16 and Table 4.17, we can see similar values for the median precision, TPR, and F-measure for distance-based store-to-store and brand-to-brand rules. From the analysis of these values, we can conclude that KASPER generates highly precise directional rules for all cases for City1 and City2.

Table 4.14: First quartile and median values for measures for four cities.

(a) First quartile for 13 dependent measures for four cities.

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	56.36	72.55	25.00	25.93	21.74	25.00	21.43	17.86
FDR	11.63	6.25	44.44	48.65	14.63	14.70	45.83	50.00
TPR	44.83	67.31	18.18	15.79	15.00	19.58	15.38	8.61
FPR	1.41	0.75	3.44	3.40	1.45	1.23	1.94	2.09
F	51.06	67.86	20.79	18.90	15.38	19.08	16.67	11.49
G	51.80	68.69	22.13	20.44	16.17	20.29	18.57	14.36
AC	86.52	91.89	83.38	81.54	83.98	84.26	88.14	82.88
E	2.80	1.91	9.94	10.76	3.55	3.39	6.38	10.85
TNR	93.68	96.14	90.98	90.54	92.61	92.31	93.98	90.94
FNR	13.33	8.00	53.20	59.54	16.48	18.71	51.85	68.30
LR+	6.97	13.75	2.74	2.60	2.17	2.62	3.78	2.17
LR-	0.14	0.08	0.58	0.66	0.17	0.19	0.55	0.76
DOR	13.90	47.42	3.32	3.09	2.38	2.90	4.36	2.39

(b) Median for 13 dependent measures for four cities.

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	79.17	84.44	39.53	37.70	50.00	51.19	38.46	30.91
FDR	20.83	15.56	60.48	62.30	50.00	48.82	61.54	69.09
TPR	72.22	81.25	32.44	28.00	44.19	50.00	28.57	18.52
FPR	2.94	1.91	6.07	6.18	3.61	2.90	3.88	5.67
F	75.54	82.81	35.63	32.13	46.92	50.59	32.79	23.16
G	75.02	83.00	34.83	32.41	42.93	50.51	31.62	22.65
AC	92.58	95.94	87.26	85.59	90.21	92.33	91.30	86.53
E	7.42	4.06	12.74	14.41	9.79	7.67	8.70	13.47
TNR	97.06	98.09	93.94	93.82	96.39	97.11	96.12	94.33
FNR	27.78	18.75	67.57	72.00	55.81	50.00	71.43	81.48
LR+	20.33	39.55	5.21	4.22	7.80	12.62	7.14	3.40
LR-	0.29	0.19	0.73	0.78	0.61	0.53	0.75	0.87
DOR	59.00	175.22	7.25	5.56	12.53	26.47	9.90	4.19

Table 4.15: Third quartile and maximum values for measures for four cities.

(a) Third quartile for 13 dependent measures for four cities.

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	88.37	93.75	55.56	51.35	85.37	85.31	54.17	50.00
FDR	43.64	27.45	75.00	74.07	78.26	75.00	78.57	82.14
TPR	86.67	92.00	46.80	40.46	83.52	81.29	48.15	31.70
FPR	6.32	3.86	9.02	9.46	7.39	7.70	6.02	9.06
F	87.23	91.84	46.26	41.11	83.72	82.70	46.88	32.56
G	87.23	91.86	46.86	41.96	83.93	82.89	48.28	33.60
AC	97.20	98.09	90.07	89.25	96.45	96.61	93.62	89.16
E	13.48	8.11	16.62	18.47	16.02	15.74	11.86	17.12
TNR	98.59	99.25	96.57	96.60	98.55	98.77	98.06	97.91
FNR	55.17	32.69	81.82	84.21	85.00	80.42	84.62	91.40
LR+	59.90	92.00	9.56	7.77	43.07	48.78	14.52	6.81
LR-	0.58	0.34	0.87	0.89	0.93	0.89	0.88	0.95
DOR	408.52	993.78	16.71	11.92	227.75	217.87	27.71	8.32

(b) Maximum value for 13 dependent measures for four cities.

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
FDR	93.10	97.14	97.50	97.83	98.11	98.59	98.15	97.78
TPR	97.62	100.00	96.72	92.31	100.00	100.00	100.00	89.06
FPR	48.30	57.14	18.77	24.44	24.20	22.32	15.38	23.60
F	97.10	100.00	95.93	90.57	100.00	98.88	100.00	87.69
G	97.10	100.00	95.94	90.58	100.00	98.88	100.00	87.70
AC	99.12	100.00	98.88	97.50	100.00	99.79	100.00	98.87
E	45.16	54.49	25.12	32.83	49.06	32.69	21.28	29.74
TNR	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
FNR	96.30	98.82	98.25	97.22	99.51	98.68	97.96	97.92
LR+	383.71	426.00	183.19	76.89	429.00	424.00	393.30	180.00
LR-	1.08	1.11	1.08	1.06	1.07	1.10	1.09	1.23
DOR	5231.33	19504	3766.17	684.00	17589	18017	3924	574.67

Table 4.16: Median values for measures of distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	81.48	86.96	41.35	41.32	65.77	70.94	40.00	35.26
FDR	18.52	13.04	58.65	58.68	34.24	29.06	60.00	64.74
TPR	78.29	84.00	29.34	25.32	58.33	65.00	24.32	19.38
FPR	2.93	2.02	5.30	4.93	3.32	2.59	3.21	4.49
F	80.88	84.48	33.63	29.42	58.04	66.67	27.03	23.94
G	80.89	84.54	34.61	31.99	58.56	66.69	29.62	27.25
AC	93.57	95.09	87.41	87.26	91.43	94.12	90.89	86.89
E	6.43	4.91	12.59	12.74	8.58	5.88	9.11	13.11
TNR	97.07	97.98	94.71	95.07	96.69	97.41	96.79	95.52
FNR	21.71	16.00	70.66	74.68	41.67	35.00	75.68	80.63
LR+	20.31	33.94	5.67	5.11	12.44	21.49	7.48	4.50
LR-	0.24	0.19	0.76	0.80	0.45	0.38	0.80	0.85
DOR	80.35	141.30	7.67	6.61	28.72	50.19	9.95	5.78

Table 4.17: Median for measures of brand-to-brand rules for City1 and City2.

	Directional		Categorical	
	City1	**City2**	**City1**	**City2**
P	91.81	96.63	63.34	71.97
FDR	8.19	3.38	36.67	28.03
TPR	81.90	94.96	46.70	60.48
FPR	4.62	3.55	2.14	2.16
F	82.16	91.24	49.32	54.70
G	83.12	91.36	50.62	56.01
AC	93.64	95.87	91.00	95.35
E	6.36	4.13	9.01	4.655
TNR	95.38	96.46	97.86	97.84
FNR	18.11	5.04	53.3	39.52
LR+	16.56	27.54	6.11	16.04
LR-	0.19	0.06	0.55	0.41
DOR	121.85	411.01	13.82	51.56

4.6 Potential Competitors

This section gives a brief discussion about the mean and median value for the number of potential competitors for several distances. For generating distance-based store-to-store rules, our approach determines a set of potential competitors according to a given distance.

Table 4.18: Mean and median for number of potential competitors for four cities.

City	Distribution	1 km	1.6 km	2 km	3 km	4 km
City1	Mean	2.33	4.23	5.80	10.78	17.31
	Median	2.00	4.00	6.00	11.00	16.00
City2	Mean	1.80	2.80	3.93	6.95	11.66
	Median	2.00	3.00	3.00	7.00	11.00
City3	Mean	2.57	4.88	7.44	16.04	27.21
	Median	2.00	5.00	8.00	16.00	29.00
City4	Mean	2.87	4.99	7.68	16.31	27.83
	Median	3.00	5.00	7.00	16.00	28.00

Table 4.18 shows the mean and median value for the number of competitors for several distances for stores for City1, City2, City3, and City4. It is clear that when the distance increases the number of competitors also increases. The number of competitors for the 1.6 km and 2 km radiuses is similar, as shown in Table 4.18. Hasan et al. found stores have a strong price change relation within 1.6 km radius [2]. We found a similar number of competitors for the 2 km radius. We consider the 2 km radius for analyzing distance-based store-to-store rule quality as a sample radius. We also run our system for 1 km, 1.6 km, 3 km, and 4 km and analyze the rule quality for the four cities. The median, mean, and STDEV value for the seventeen measures are presented in Appendix B. From the analysis of the median, mean, and STDEV value of the accuracy, precision, TPR, and F-measures are similar for all radiuses. We consider spatial distance for calculating a set of potential competitors. There are several factors related to competitors. Just because a store has a small distance, does not necessarily mean it is a strong competitor. Sometimes spatial distance is closer but driving distance is not. From Table 4.18, we

can see that the median number of competitors for 4 km distance is 16, 11, 29, and 28 for City1, City2, City3 ,and City4 respectively. The greater the distance, the higher the number of potential competitors. So, we do not consider a distance of more than 4 km for calculating potential competitors.

4.7 Heuristic Rank-Based Variation

KASPER is able to generate brand-based store-to-store price change rules for stores for a brand for a city. KASPER generates rules for a key store from the relation of all other stores in the same brand. A key store generates SC rules if it has relation with other stores in a PC category. For every PC category for key store, a number of selected SC rules build DC rules. Every SC rule has a key store and other store. Every DC rule has a key store and two other stores. If there are h other stores, the number of possible DC rules will be $\frac{h(h-1)}{h}$. Table 4.19 shows the median, mean, and STDEV of number other stores for directional rules for all key stores for four cities. From Table 4.19, we can see that the median number of other stores is 25 for City1, which means approximately half of the key stores have a PC relation with at least 25 other stores that are used to build DC rules for a PC category. Similarly, the median number of other stores for City2, City3, and City4 are 27, 33.50, and 10 respectively.

Table 4.19: Distribution of number of other stores for directional rules for all key stores for a City.

Distribution	**City1**	**City2**	**City3**	**City4**
Median	25.00	27.00	33.50	10.00
Mean	24.78	22.14	32.51	14.97
STDEV	12.89	9.55	17.27	15.44

For this experiment, we run the system three times and consider average execution time. From Table 4.20, we can see that the average execution time is calculated for generating directional rules for 47 key stores for brand B108 for City1 from the price change

Table 4.20: Average execution time for directional rules for 47 key stores with variable number of other stores.

# of other stores	Combination	Average execution time (minute)	STDEV
5	47 × 5	22.27	4.33
10	47 × 10	24.77	4.45
15	47 × 15	27.08	4.14
20	47 × 20	33.57	5.68
25	47 × 25	37.11	4.92
30	47 × 30	39.41	0.19
35	47 × 35	41.91	4.22
40	47 × 40	41.82	2.26
45	47 × 45	43.06	1.70
47	47 × 47	43.58	0.21

Table 4.21: Distribution of number of other stores for directional rules for all key stores for brand B108 for City1.

Min	Q1	Median	Q3	Max
1	33	39	42	47

Figure 4.3: Number of other stores vs. average execution time for directional rules for all key stores.

relation with 5, 10, 15, 20, 25, 30, 35, 40, 45, and 47 other stores in the same brand.

In Figure 4.3, the number of other stores are plotted on the X-axis and the average execution time in minutes is plotted on the Y-axis. From Figure 4.3, we can see that the average execution time for 5 other stores is 22.27 minutes, for 10 other stores is 24.77 and so on. The average execution time is approximately linear for directional rules for all key stores for brand B108. From Table 4.21, the median number of other stores for half of the key stores for brand B108 is 39. When we use 5 stores that means the top 5 stores will be chosen from 39 stores for all key stores to build DC rules. Similarly for 10, 15, 20 and so on.

The average execution time is not always linear. It depends on the PC relation of a key store to number of other stores. If the maximum number of other stores is less than a threshold then execution time will not be affected. Table 4.22 shows the execution time for directional rules for 166 key stores for brand B403 for City4 with variable number of other stores that are used to build DC rules.

Table 4.22: Execution time for directional rules for 166 key stores with variable number of other stores.

# of other stores	Combination	Execution time (hour)
10	166×10	3.43
20	166×20	3.58
30	166×30	3.71
40	166×40	3.76
50	166×50	3.8
60	166×60	3.83
70	166×70	3.97
80	166×80	4.01
100	166×100	3.98
120	166×120	4.00
140	166×140	3.99
166	166×166	4.13

From Figure 4.4, we can see that the execution time is grows until 75 other stores and

Table 4.23: Distribution of number of other stores for directional rules for all key stores for brand B403 for City4.

Min	Q1	Median	Q3	Max
1.00	4.75	15.00	34.00	75.00

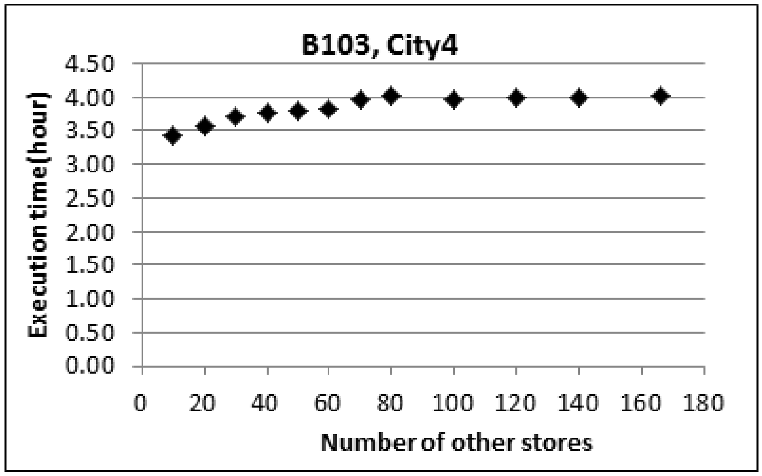


Figure 4.4: Number of other stores vs. execution time for brand B403 for City4.

then plateaus. From Table 4.23, we can see that the maximum number of other stores is 75. Execution time will be almost the same for more than 75 stores. If we increase the threshold value and there is less possible other stores for making PC for the key store, then the exeution time will be the same. Though there are 166 stores and the maximum number of possible other stores is 75, using 80, 100 through 166 we found that the execution times were similar.

The growth rate of the execution time is not linear with an increasing number of other stores because the number of other stores that have a PC relation with a key store varies. From Table 4.23, we can see that half of the key stores have a PC relation with less than or equal to 15 other stores, and half with more than 15 stores. The median number of other stores is 15 and the third quartile is 34, which means that 25% of the key stores have a PC relation with 15 to 34 other stores.

Accuracy and variable number of other stores

In the complete-component method, we use all other stores that are related to the key store when generating rules. This method ensures that the result is correct, but it needs more execution time than the ranking method. For the ranking method, we analyze some measures to compare results for a variable number of stores. Table 4.24 shows the median, mean, accuracy, precision, TPR, F-measure, and STDEV values for directional rules for 47 key stores with brand B108 in City1 for a variable number of other stores. We can see that STDEV values are similar for all cases and the median and mean values for the accuracy, precision, TPR, and F-measure are also similar. So, the ranking method provides similar results with lower execution time.

Table 4.24: Median, mean, and STDEV of some measures for directional rules for variable number of other stores for brand B108 for City1.

Number of stores: 47				
Distribution	**Accuracy**	**Precision**	**TPR**	**F-measure**
Median	89.08	66.03	52.48	55.25
Mean	89.26	64.60	53.95	57.12
STDEV	6.02	22.60	25.63	24.11
No. of stores: 40				
Median	89.08	66.03	52.48	55.25
Mean	89.26	64.60	53.95	57.12
STDEV	6.02	22.60	25.63	24.11
No. of stores: 30				
Median	89.08	63.01	53.18	56.33
Mean	88.91	63.86	53.80	56.63
STDEV	6.40	21.92	25.12	24.03
No. of stores: 20				
Median	89.56	63.47	52.48	55.62
Mean	89.51	64.35	52.66	56.27
STDEV	6.37	21.64	24.91	23.55
No. of stores: 10				
Median	89.19	66.67	48.28	53.16
Mean	88.65	64.75	50.34	54.13
STDEV	6.90	23.10	24.08	22.52

The performance of the ranking method is affected by the number of other stores that are required to give correct results for a key store of a brand in a city. We recommend that this method be used when low execution time is important and the complete-component method be used when accuracy is important.

4.8 Comparison With Other Research

This section compares our research to existing research. The comparison is summarized with respect to data collecion, duration, type of data, data transformation, analysis problem, method, and measures, as shown in Table 4.25 and Table 4.26.

The general similarities of our research with four previous research efforts involving data sets of motor fuel prices, are shown in Table 4.25 and Table 4.26. Hosken et al. showed three findings from an experiment involving weekly analysis of retail margin and retail prices [19], whereas our approach generates hundreds of price change rules from daily retail prices. Hosken et al. claim that the primary sources of retail price variation results from (a) a store changing its price in response to a change in the wholesale price and (b) a store changing its price relative to other stores. In our research, we also use a store changing its price relative to other stores.

Finding 2 of Hosken et al. stated that stores do not follow simple pricing rules. For finding 2, Hosken et al. observed the difference between weekly prices for stores and the region's mean. They found 56% of prices are within 2.5 cents per gallon of the region's mean and 71% of prices are within 3.5 cents per gallon. Hosken et al. also found that 3.5% of prices are more than 10 cents per gallon from the mean. They also found that the distribution of the residuals is not normal. If the residuals were normal, they would have expected that between 47% and 62% of prices would be within 2.5 and 3.5 cents per gallon of the mean, and 1.2% of prices would be more than 10 cents per gallon from the mean. Here, we observe the percentage of daily prices above, below or equal to the

Table 4.25: Comparison of our research with existing research.

research	Hosken et al. [19]	AlGudhea et al. [3]	our research
data collection	(a) prices from fleet card transaction from 272 motor fuel stores (OPIS) (b) survey of 600 store's address, attributes, etc. (c) neighborhood characteristics	spot, crude oil price (U. S. dept. of energy)and wholesale, retail price (OPIS)	motor fuel prices at 1221 motor fuel stores
collection period	1997 to 1999	December 1998 to January 2004	1 January 2010 to 30 April 2015
type of data	weekly	daily	daily price report
data transformation	branded rack price (wholesale price), other store prices	spot, crude oil, wholesale, and retail price changes	brand-to-brand relation; store-to-store relation
analysis problem	three findings	pairwise relations among several levels of distribution chain	motor fuel price change rules
method	regression	threshold and momentum model [12, 13]	KASPER
measures		F-statistics	17 rule quality measures and statistical analysis

Table 4.26: Comparison of our research with existing research.

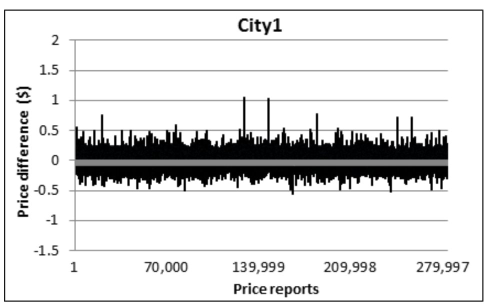
research	Eckert and West [11]	Lewis et al. [27]	our research
data collection	426 stores prices were reported by consumers to the website www.gastips.com	(a) daily average price of 280 cities (AAA, OPIS), (b) store-specific data for 165 cities (OPIS) (c) prices of speedway stores every 3 hours.	motor fuel prices at 1221 motor fuel stores
collection period	March 1 to August 31, 2000	(a) October 2004 to July 2010, (b) July 2008 to July 2010, (c) August 2008 to July 2010	1 January 2010 to 30 April 2015
type of data	real time	daily and real time	daily price report
data transformation	brand effects, spatial and product characteristics, market structure, etc.	dependent, independent brands, market share	brand-to-brand relation; store-to-store relation
analysis problem		price cycles, leader of price restoration	motor fuel price change rules
method	econometric model	median, mode price [26], regression	KASPER
measures	coefficient (S.E.), probability	coefficient (S.E.)	17 rule quality measures and statistical analysis

city mean. The main difference of these two is that Hosken et. al. considered exact price prediction, whereas our approach consider price change category for generating rules.

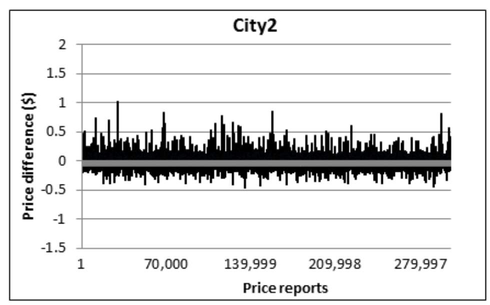
Table 4.27 shows the distribution of prices relative to the city mean price and Figure 4.5 shows the daily price difference from the city mean price for every store for four cities. Every day price report is plotted on the X-axis and the daily price difference from the city mean is plotted on the Y-axis.

Table 4.27: Distribution of prices relative to the city mean price.

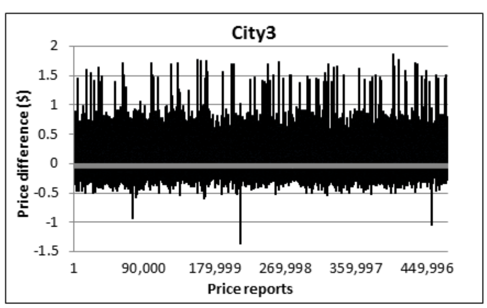
City	**Above Mean (%)**	**Below Mean (%)**	**Equal Mean (%)**
City1	55.50	44.49	0.01
City2	56.77	43.23	0.00
City3	51.31	48.69	0.00
City4	46.84	53.16	0.00



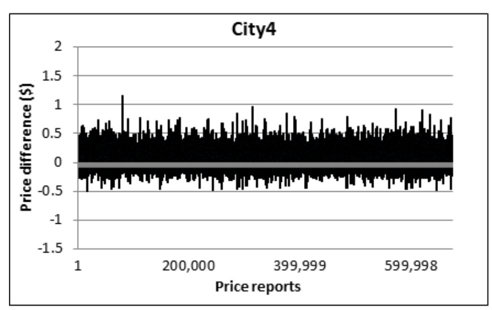
(a) City1.



(b) City2.



(c) City3.



(d) City4.

Figure 4.5: Daily price difference for every store from the city mean price for four cities.

From Table 4.27, we can see that almost every price report for every store is above or below the city mean price. Figure 4.5a and Figure 4.5b show that most prices are between 30 to 50 cents per gallon of the city mean for the high variability cities. Figure 4.5c and Figure 4.5d show that most prices are more than 50 cents per gallon above the city mean for low variability cities. In our data set, we find more fluctuation in daily prices from the city mean than Hoskel et al. found. Hosken et al. concluded that there is no simple rule that can define the pricing behaviour of all stores [19]. Our experiment results support this statement in the form it is given. Although there is no simple rule that defines the pricing behaviour for all stores for a specific price prediction, our research gives an idea to generate a simple price change rule for a PC category for every individual store for a city.

Al-Gudhea et al. analyzed pairwise relations among crude oil and retail motor fuel price, crude oil and spot motor fuel price, spot and wholesale motor fuel price, and wholesale and retail motor fuel price to determine the response time of price hikes [3]. Our research generates rules from the pairwise price change relation from one store to another store and from one to another. Both approaches use retail motor fuel prices. From their experiments, Al-Gudhea et al. showed the dissimilar behaviour of the response time of upstream and downstream prices at different stages in the distribution chain [23]. Our research generates price change rules from the pairwise relation between a key store or brand to another store or brand.

Eckert and West, and our research use same belief of assumptions about competitors in the retail motor fuel market. Our approach determines a set of potential competitors for generating distance-based store-to-store rules within a certain radius. Both of Eckert and West, and our research use store-specific daily retail motor fuel prices. Eckert and West used six months of daily prices for 426 stores, whereas our research uses five years and four months of daily prices for 1221 stores.

Lewis and our research analyze retail motor fuel prices for different regions. Lewis analyzed data to search for pricing cycles in a particular region and if it existed then deduced the pricing behaviour [27]. From the analysis of data he claimed that *"Midwestern U.S. retail prices often fall at an average of a cent per day or more a week or two and jump 10 to 20 cents in one day before starting fall again"*. Instead of demonstrating general pricing behaviour of a city, our research provides store-specific price change rules for a city.

Our research generates rules at the store level and the brand level, while other systems mostly summarized pricing behaviour for a city. This research provides a novel approach of using seventeen existing measures to evaluate the rule quality. We did not find any other research that uses all of these measures together for evaluating the quality of a rule. Although previous research suggested ideas for developing our method, our method is a new and unique approach to the problem of generating price change rules.

Chapter 5

CONCLUSION

Section 5.1 of this chapter summarizes our research work, including our experimental results. Section 5.2 describes possible future directions for this research.

5.1 Summary

The general goal of our research is to generate and test price change rules for setting retail prices for commodities that change in price frequently. The main target of this research is to provide a specific decision rule that explains, with high predictive accuracy, why a specific store or brand made a price change in a specific category in terms of price changes at other stores or brands in a city.

Our research developed the KASPER software system for generating meaningful and high quality decision rules that predict price changes for motor fuel for a specific store or brand from the relation of price changes of other stores or brands in a city. The generated rules provide information about the direction and amount of the predicted price change.

Our approach uses pairwise relations at the store (retail) level and generates price change rules at the store level, while other systems use pairwise relations at other stages in the distribution chain [19] and summarize retail pricing behaviour at the city level [11,27]. Our system is capable of generating thousands of rules while Olvarrieta et al. manually devise eleven hypotheses [34]. The overall process from profile construction to decision rule generation of KASPER is different from that of any other rule generation model. Our research implements a new approach for selecting decision rules using a combination of seventeen rule quality measures; we did not find any other system that uses all these measures together.

KASPER was tested on data from motor fuel prices of stores and brands from four cities. KASPER generates rules from three relations (brand-based store-to-store, distance-based store-to-store, brand-to-brand) for a specific set of PC categories.

According to the evaluation of the directional decision rules on unseen data, they have high predictive accuracy and precision for most of the store-to-store, and brand-to-brand rules for high-variability cities (City1 and City2). This system is more effective with directional PC categories ($z = 2$) than categorical PC categories ($z = 6$) for high-variability cities than for low-variability cities. This system did not generate any brand-to-brand rules for low-variability cities because of the consistency of price changes for these two cities.

Overall, KASPER generates good rules for stores and brands where prices often fluctuate, and stores and brands which are highly responsive to each other for changing their prices.

5.2 Future Work

There is a good potential for performing other research related to the topic of this thesis. In this section, we discuss possible research topics related to combining multiple rule sources, improving the rank-based method, duration of training, validation and testing phases, the tacit collusion model, other features, and other distance measures.

Our approach could be extended by combining multiple sources of rules. We obtained results for a city in the form of a two-dimensional table with rows for stores and columns for price categories and at most one decision rule for each city-category combination. We collected several such tables for directional price changes, i.e. one for brand-based store–to-store rules, several for distance-based store-to-store rules (with different distances), and one for brand-to-brand rules. We also collected a similar set for categorical price changes. Future research could devise a strategy for combining information from several such tables into a single table. For example, S1001 is included in brand B101 and suppose store S1001 has a store-to-store rule for PC category 1 and brand B101 also has a brand-to-brand rule for PC category 1. From these two rules, the system could select one. One idea is that the store manager could choose whichever rule has higher values for precision and accuracy.

Information about the distribution of the number of other stores could be used to improve the rank-based method of generating DC rules described in Section 4.7. Recall that KASPER creates DC rules using either the complete-component or the rank-based method. With the complete-component method, DC rules are built for all possible stores related to the key store for a specific PC category, whereas with the rank-based method DC rules are built for only a selected stores.

As given, the rank-based method lacks an effective means of choosing a, the number of other stores to select. The value of number of other stores effects the execution time of a method. There is no option for the complete-component method to tune the number of other stores but the rank-based method has this option. Section 4.7 shows that the rank-based method is faster than the complete-component method for generating brand-based store-to-store price change directional rules. The lower number of other stores makes the brand-based method faster and the higher number of other stores makes the brand-based method slower. On the other hand, the complete-component method uses the maximum number of other stores which is slower but makes sure to provide the correct result in all cases. If we use few number of stores then sometimes good rules can be deducted, so the goal is to select an efficient number of other stores that will provide correct results with lower execution time. So, selecting the number of other stores will be a challenge for future reseach.

One simple idea could be extracting the minimum, first quartile, median, third quartile, and maximum value of number of other stores from the distribution of number of other stores for a brand and make a choice to select the number of other stores. From Table 4.23, we can see the distribution of the number of other stores for brand B403. The total number of stores for brand B403 is 166. In our research, for brand-based store-to-store rules for a key store, the number of other stores for that key store is equal to the total number of stores of that brand less one. So, the possible number of other stores is 165. The maximum number of other stores is 75 which indicates that no key store for brand B403 have price change relations more than 75 other stores . Third quartile indicates that 75 percent of key stores have price change relations at most 34 other stores. Here 34 other stores could be chosen for generating brand-based store-to-store DC rules instead of taking all 165 stores. So, future research could examine the median or the third quartile value of the number of other stores for choosing the value of a.

For low variability cities (City3 and City4), restricting the duration of the training, validation, and testing phases could lead to better rules using the KASPER system. In our research, we used two years of data for training, two years for validation, and one and a half years for testing for all cities. We found good rules for high variability cities (City1 and City2). Low variability cities may follow a rule for a shorter period of time and then change to another rule. For example, perhaps one key store maintains a PC relation with store A for two months and then with store B for two months. Instead of using a five and a half year period, future research could investigate whether using KASPER with (say) two months as the duration for the training, validation, and testing phases would lead to better results for low variability cities.

Future work could apply the concept of a tacit collusion model to stores in low variability cities (City3 and City4) to evaluate how effective the concept is for these cities. In the tacit collusion model, if one store sets a price for a product, then a group of stores will also set the same price for the same product. This price is called the *tacitly collusive price*. “Regions that have [a] higher station concentration or a smaller number of stations are expected to be more likely to sustain tacit collusion than regions with many firms and stations” [11]. From Section 4.3, we see that the EODPC for stores for low variability cities (City3 and City4) vary less than for the EODPC for stores in high variability cities (City1 and City2). Also from the graph with the EODPC values in sorted order, shown in Figure 4.1g and Figure 4.1h, we see that the PC category does not change for large parts of a year. From Table 4.7, we see that the stores in low variability cities (City3 and City4) do not change prices on more than 40% of the days. These characteristics of the observed price changes for low variability cities (City3 and City4) suggest that the tacit cullision model could be applied effectively to predict price changes at stores within a certain distance of a key store. The prices could be predicted assuming that these stores will make a certain price change when the key store makes a change.

Some other features may play a role in a pricing strategy and thus could be added to our approach. One such feature that might be added is whether or not a store is branded. Future research could then investigate whether branded stores behave differently from unbranded ones. The percentage of market share might also be a useful feature for a brand. Future research could also investigate whether the behaviour of a brand for setting prices depends on the size of its market share. As well, other factors, such as wholesale price, rack price, spot price, and margin could be added to make our approach more comprehensive.

KASPER generates distance-based store-to-store rules where the competitors are selected according to spatial distances. However, spatial distances may be considerably different from driving distances because of traffic rules and obstacles, such as boulevards and overpasses. In the future, potential competitors could be selected using driving distances or driving times instead of spatial distances.

REFERENCES

[1] AÅKERMAN, N. Knowledge-Acquisition Strategies and the Effects on Market Knowledge Profiling the Internationalizing Firm. *European Management Journal 33*, 2 (2015), 79–88.

[2] AHMED, H. W. Prediction of Retail Prices Using Local Competitors. Master's thesis, University of Regina, December 2013.

[3] AL-GUDHEA, S., KENC, T., AND DIBOOGLU, S. Do Retail Gasoline Prices Rise More Readily Than They Fall?: A Threshold Cointegration Approach. *Journal of Economics and Business 59*, 6 (2007), 560 – 574.

[4] ANDERSON, D. R., DENNIS J. SWEENEY, AND THOMAS A. WILLIAMS. *Statistics for Business and Economics*, 10^{th} ed. Kindle, 2009.

[5] BULLINARIA, J. A. IAI : Expert Systems. `http://www.cs.bham.ac.uk/~jxb/IAI/w9.pdf`, Accessed February, 2015.

[6] CAIRÓ, O., AND GUARDATI, S. The KAMET II Methodology: Knowledge Acquisition, Knowledge Modeling and Knowledge Generation. *Expert Systems with Applications 39*, 9 (2012), 8108 – 8114.

[7] CANADIAN FUELS ASSOCIATION. Fuel Markets and Pricing. `http://canadianfuels.ca/en/faq`, Accessed February, 2015.

[8] Crocker, M. W. Connectionist and Statistical Language Processing. http://www.coli.uni-saarland.de/~crocker/Teaching/Connectionist/lecture11_4up.pdf, Accessed April, 2015.

[9] Doyle, J., Muehlegger, E., and Samphantharak, K. Edgeworth Cycles Revisited. *Energy Economics 32*, 3 (2010), 651 – 660.

[10] Ebell, M. H. The University of Georgia, Chapter 4: Diagnosis I - Likelihood Ratios. http://ebp.uga.edu/courses/, Accessed March, 2015.

[11] Eckert, A., and West, D. S. Price Uniformity and Competition in a Retail Gasoline Market. *Journal of Economic Behavior & Organization 56*, 2 (2005), 219 – 237.

[12] Enders, W., and Granger, C. W. J. Unit-Root Tests and Asymmetric Adjustment With an Example Using the Term Structure of Interest Rates. *Journal of Business & Economic Statistics 16*, 3 (1998), 304–311.

[13] Enders, W., and Siklos, P. L. Cointegration and Threshold Adjustment. *Journal of Business & Economic Statistics 19* (2001), 166–176.

[14] Engel, B. Knowledge Acquisition and Validation. www.prenhall.com/divisions/bp/app/turban/dss/6e/ppt/short-ch11.ppt, Accessed February, 2015.

[15] He, Q., Ghobadian, A., and Gallear, D. Knowledge Acquisition in Supply Chain Partnerships: The Role of Power. *International Journal of Production Economics 141*, 2 (2013), 605 – 618.

[16] Hibbard, B. Model-based Utility Functions. *Journal of Artificial General Intelligence 3*, 1 (2012), 1–24.

[17] Hilderman, R. J. CS 831: Knowledge Discovery in Databases. http://www2.cs.uregina.ca/~hilder/previous_courses.html, Accessed February, 2015.

[18] Hopkins, E. Price Dispersion. In *The New Palgrave Dictionary of Economics*, S. N. Durlauf and L. E. Blume, Eds. Palgrave Macmillan, Basingstoke, 2008.

[19] Hosken, D. S., McMillan, R. S., and Taylor, C. T. Retail Gasoline Pricing: What Do We Know? *International Journal of Industrial Organization 26*, 6 (2008), 1425 – 1436.

[20] Investopedia. What is the difference between upstream and downstream oil and gas operations? `http://www.investopedia.com/ask/answers/060215/what-difference-between-upstream-and-downstream-oil-and-gas-operations.asp`, Accessed August, 2015.

[21] Jahan, M. Application of a Genetic Algorithm Decision Tree to Classifying Motor Fuel Prices and Finding the Best Classification Algorithm, August 2012. CS 831 Course Project, Department of Computer Science, University of Regina.

[22] Jensen, B. B., and Grunert, K. G. Price Knowledge During Grocery Shopping: What We Learn and What We Forget. *Journal of Retailing 90*, 3 (2014), 332 – 346.

[23] Karrenbrock, J. D. The Behavior of Retail Gasoline Prices: Symmetric or Not? *Federal Reserve Bank of St. Louis 73* (1991), 19–29.

[24] Kohavi, R., and Provost, F. Glossary of Terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. *Machine Learning 30*, 2-3 (1998), 271 – 274.

[25] Kubat, M., Holte, R. C., and Matwin, S. Machine Learning for the Detection of Oil Spills in Satellite Radar Images. *Machine Learning 30*, 2-3 (1998), 195 – 215.

[26] Lewis, M. S. Temporary Wholesale Gasoline Price Spikes Have Long Lasting Retail Effects: The Aftermath of Hurricane Rita. *Journal of Law and Economics 52*, 3 (2009), 581–605.

[27] Lewis, M. S. Price Leadership and Coordination in Retail Gasoline Markets with Price Cycles. *International Journal of Industrial Organization 30*, 4 (2012), 342 – 351.

[28] Lucas, P. J., and van der Gaag, L. C. *Principles of Expert Systems.* Addison-Wesley, Amsterdam, 1991.

[29] Macaskill, P., Gatsonis, C., Deeks, J., Harbord, R., and Takwoingi, Y. Chapter 10: Analysing and Presenting Results. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*, J. Deeks, B. PM, and C. Gatsonis, Eds. The Cochrane Collaboration, 2010.

[30] Martínez-Ballesteros, M., Martínez-Álvarez, F., Troncoso, A., and Riquelme, J. Selecting the Best Measures to Discover Quantitative Association Rules. *Neurocomputing 126* (2014), 3 – 14.

[31] Matuszak, A. The Economist at Large: Economic and Financial Analysis for Real World. `http://economistatlarge.com/finance/applied-finance/differences-arithmetic-geometric-harmonic-means`, Accessed February, 2015.

[32] Moore, D. S., McCabe, G. P., and Craig, B. A. *Introduction to the Practice of Statistics*, 3^{rd} ed. New York: W. H. Freeman, 1999.

[33] Mwemezi, J. J., and Huang, Y. Optimal Facility Location on Spherical Surfaces: Algorithm and Application. *New York Science Journal 4*, 7 (2011), 21–28.

[34] Olavarrieta, S., Hidalgo, P., Manzur, E., and Farías, P. Determinants of In-Store Price Knowledge for Packaged Products: An Empirical Study in a Chilean Hypermarket. *Journal of Business Research 65*, 12 (2012), 1759 – 1766.

[35] Renze, J. Outlier: From Mathworld–A Wolfram Web Resource, created by Eric W. Weisstein. `http://mathworld.wolfram.com/Outlier.html`, Accessed February, 2015.

[36] SAMADDAR, D. S. MGS 9920: PROB & STAT THEORY I: Descriptive Statistics - Numerical Measures. `http://www2.gsu.edu/~dscsss/teaching/mgs9920/`, Accessed February, 2015.

[37] SANDERS, D. H., SMIDT, R. K., ADATIA, A., AND LARSON, G. A. *Statistics: A First Course*, 1^{st} ed. McGraw-Hill, 2001.

[38] SASAKI, Y. Properties of Rule Interestingness Measures and Alternative Approaches to Normalization of Measures. *Information Sciences 216* (2012), 1–16.

[39] ZARRI, G. P. Knowledge Representation and Inference Techniques to Improve the Management of Gas and Oil Facilities. *Knowledge-Based Systems 24*, 7 (2011), 989 – 1003.

[40] ZHOU, K. Z., ZHANG, Q., SHENG, S., XIE, E., AND BAOD, Y. Are Relational Ties Always Good for Knowledge Acquisition? Buyer-Supplier Exchanges in China. *Journal of Operations Management 32*, 3 (2014), 88 – 98.

Appendix A

PRODUCT PRICING KNOWLEDGE

Appendix A describes two research efforts related to product pricing and Table A.1 shows a comparison of our research with these two research efforts.

Olavarrieta et al. conducted a survey to determine shoppers' in-store price knowledge [34]. This study collected data from 585 shoppers' in-store price knowledge for packaged food products in a Chilean hypermarket and compared the estimated prices given by shoppers with the actual product prices to determine how many times prices were correct. Olavarrieta et al. used five product factors and three shopper factors to make eleven hypotheses, eight related to products and three to shoppers. The product factors were price signs, product bundling, store brands, product category, and item price. The shopper factors were in-store price comparison activity, shopping frequency, and brand loyalty. One hypothesis concerning price signs was as follows:

H9. The higher the in-store price-comparison activity by shoppers, the higher her (his) in-store price knowledge [34].

This hypothesis is particularly relevant to our research.

The eleven hypotheses were tested on the hypermarket data set. The accuracy of shoppers' in-store price knowledge was determined by the comparison of respondents' price knowledge to actual price in-store, the cumulative percentage of respondents within several accuracy levels, calculating the p-value of correct and incorrect respondents, calculating the odds-ratio, and calculating the p-value of Chi-square test for all hypotheses. The results showed that four of the eight hypotheses created from product factors and one of the three hypotheses created from shoppers factors were statistically significant.

Jensen and Grunert described a multi-point and multi-measure approach of measuring shoppers' price knowledge about grocery shopping [22]. Jensen and Grunert used the time and type of measurement of the price knowledge. The times of measurement were before, during, and after a store visit. The types of measurement were price recall, price recognition, and deal spotting. Three questions were investigated; the most relevant one was:

Question 3. How do brand and store loyalty, category purchase frequency, price range, and deal share add to the explanation of the variance in consumer price knowledge? [22]

This study collected data from interview of 1,204 shoppers (395 before the store visit, 420 at the shelves, and 389 after the store visit) during a two-week period in August in Denmark. The study used two stores (hypermarket and supermarket) and identified 29 potential product categories. The data set consisted of a combination of personal interviews and a questionnaire completed at home. Sample questions before, during, and after leaving store are given below:

Before Visit: What do you think is the price of this brand of [category] at [store name] *today?*

During Visit (after purchase): What was the price of this brand of [category] at [store name] today?

After Visit: What was the price of this brand of [category] you just chose?

The study correctly calculated price recall, price recognition, and deal spotting for respondents before, during, and after the store visit. The results showed that the price recall, accuracy, which is the deviation relative to the correct price was 7.9% during the store visit. that was the lowest deviation among three times. This paper also reported the significance (p-value) of these measures for matching features such as demographics (mean age, gender, etc.), customer characteristics (brand loyalty, store loyalty, etc.), and design variables (hypermarket proportion, margarine proportion, juice proportion etc.) before, during, and after measures of price knowledge.

Comparison:

The two research efforts just discussed are related to our work. Table A.1 shows the overview of the comparison of them with our research. Olvarrieta et al. created hypotheses about shoppers' in-store price knowledge for packaged products based on product and shopper factors [34]. Jensen and Grunert emphasized time and type of measurement of shoppers' price knowledge [22]. Although our approach has similarities with both Olvarrieta et al., and Jensen and Grunert in terms of pricing knowledge and hypotheses, they used grocery products prices and we use motor fuel prices. Our approach uses daily price report while Jensen and Grunert used customers' price knowledge before, during, and after a store visit. Our approach generates hundreds or thousands of rules from the relationship of stores or brands, whereas Olvarrieta et al. created eleven hypotheses from product and shoppers factors. Olvarrieta et al. evaluated the correctness of hypotheses on

Table A.1: Comparison of our research with existing research.

Research	Olavarrieta et al. [34]	Jensen and Grunert [22]	Our research
data collection	585 shoppers in-store price knowledge for packaged food products in a Chilean hypermarket.	interview from 1,204 shoppers (Two stores and 29 potential product categories)	motor fuel prices at 1221 motor fuel stores
collection period		two-week period	1 January 2010 to 30 April 2015
type of data	in-store price knowledge	consumers' price knowledge before, during, after visiting store	daily price report
data transformation	product factors (store brands, item's price etc.); shopper factors (price comparison activity, shopping frequency, etc.)	purchase frequency, price range	brand-to-brand relation; store-to-store relation
analysis problem	eleven hypotheses (made by authors)	three research questions	motor fuel price change rules
method	Chi-square test	matching of shoppers' in-store price knowledge to actual price	KASPER
measures	odds-ratio and p-value	price recall, recognition, judgement, and p-value (significance test)	17 rule quality measures and statistical analysis

the data and found that five were statistically significant, among the eleven they studied. Jensen and Grunert measured the percentage of price recall, price recognition, and deal spotting for different products from customers' price knowledge across product categories before, during, and after a store visit. Our research uses a new method of the combination of seventeen measures for generating rules which is different from that of Olvarrieta et al. and Jensen and Grunert.

Appendix B

DETAILED EXPERIMENTAL RESULTS

Table B.1: Values of 4 independent measures for distance-based store-to-store rules for four cities ($d = 1$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	3,438	2,247	1,516	693	2,899	2,067	512	228
FP	600	430	2,170	894	1,609	867	612	353
FN	1,084	599	3,902	2,190	2,218	1,222	1,738	928
TN	16,384	10,920	40,415	17,600	31,950	22,945	23,427	7,301
Total	21,506	14,196	48,003	21,377	38,676	27,101	26,289	8,810

Table B.2: Values of 4 independent measures for distance-based store-to-store rules for four cities ($d = 1.6$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	8,961	7,916	3,366	2,034	6,882	7,460	1,276	966
FP	1,940	1,279	4,887	2,481	3,657	2,819	1,899	1,316
FN	2,690	1,656	8,277	5,303	5,047	3,746	3,694	2,627
TN	42,246	33,505	85,187	47,474	76,273	70,226	53,156	26,218
Total	55,837	44,356	101,717	57,292	91,859	84,251	60,025	31,127

Table B.3: Values of 4 independent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	12,583	11,055	4,739	3,477	9,477	10,676	2,189	1,621
FP	2,603	1,757	6,398	3,981	4,716	4,215	3,184	2,216
FN	3,293	2,222	10,032	6,885	6,300	5,482	5,355	3,503
TN	51,066	45,593	105,705	69,083	98,425	102,948	82,810	38,250
Total	69,545	60,627	126,874	83,426	118,918	123,321	93,538	45,590

Table B.4: Values of 4 independent measures for distance-based store-to-store rules for four cities ($d = 3$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	13,816	15,162	6,258	5,287	11,146	14,073	4,047	2,475
FP	3,057	2,460	8,486	6,284	6,112	5,695	5,861	3,562
FN	3,682	2,835	11,886	9,188	7,710	6,473	8,122	5,144
TN	59,701	57,251	128,612	95,104	118,837	132,169	138,882	55,986
Total	80,256	77,708	155,242	115,863	143,805	158,410	156,912	67,167

Table B.5: Values of 4 independent measures for distance-based store-to-store rules for four cities ($d = 4$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
TP	13,828	18,750	6,841	6,426	11,446	17,469	3,847	2,940
FP	3,062	2,830	9,307	7,626	6,560	6,418	5,929	4,272
FN	3,882	3,293	12,039	10,246	8,147	7,373	7,784	6,196
TN	61,886	68,191	132,983	108,002	125,344	156,822	132,461	67,889
Total	82,658	93,064	161,170	132,300	151,497	188,082	150,021	81,297

Table B.6: First quartile for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	68.81	76.09	27.90	28.57	43.09	49.21	26.01	18.65
FDR	10.87	7.89	42.86	44.25	13.86	12.50	42.27	44.01
TPR	58.68	68.00	15.81	15.00	33.62	42.11	12.50	8.55
FPR	1.41	1.10	2.92	2.44	1.18	1.02	1.27	2.07
F	63.45	71.58	20.69	20.36	36.70	46.51	17.45	11.43
G	63.79	71.68	22.98	21.69	37.62	46.77	19.36	13.96
AC	86.66	90.11	84.84	83.30	86.02	87.23	88.80	83.72
E	2.99	2.33	10.01	9.63	2.88	2.86	6.58	9.19
TNR	92.87	94.42	92.18	92.10	93.17	94.19	94.43	90.96
FNR	12.64	8.36	55.09	53.17	14.76	15.56	60.00	58.49
LR+	7.41	9.77	3.06	3.07	5.02	7.01	4.30	2.50
LR-	0.14	0.09	0.58	0.59	0.15	0.17	0.63	0.63
DOR	19.58	37.59	3.70	3.74	7.45	15.70	5.04	2.78

Table B.7: Third quartile for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	89.13	92.11	57.14	55.75	86.14	87.50	57.74	55.99
FDR	31.19	23.91	72.10	71.43	56.92	50.79	73.99	81.35
TPR	87.37	91.64	44.91	46.83	85.24	84.44	40.00	41.51
FPR	7.14	5.58	7.82	7.90	6.83	5.81	5.58	9.04
F	86.96	91.67	45.24	45.60	86.07	83.53	41.33	41.03
G	87.00	91.67	45.71	46.52	86.07	83.83	42.16	41.09
AC	97.01	97.67	90.00	90.37	97.12	97.14	93.42	90.81
E	13.34	9.89	15.16	16.70	13.98	12.77	11.20	16.28
TNR	98.59	98.90	97.09	97.57	98.82	98.98	98.73	97.93
FNR	41.32	32.00	84.19	85.00	66.38	57.89	87.50	91.46
LR+	58.56	74.17	9.93	9.30	55.17	58.64	15.81	7.98
LR-	0.44	0.33	0.88	0.88	0.72	0.60	0.90	0.96
DOR	432.59	619.50	16.86	15.00	295.50	234.00	24.38	13.64

Table B.8: Maximum value for 13 dependent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
FDR	66.67	66.67	95.00	92.31	95.83	83.61	97.14	98.21
TPR	98.98	100.00	96.72	93.33	100.00	100.00	94.74	84.91
FPR	50.43	38.78	16.55	20.47	29.21	28.76	15.55	20.45
F	98.81	100.00	95.93	89.36	100.00	100.00	94.74	84.11
G	98.81	100.00	95.94	89.44	100.00	100.00	94.74	84.12
AC	99.15	100.00	98.88	97.88	100.00	100.00	99.11	97.76
E	34.65	25.07	27.66	27.36	32.42	26.39	23.95	23.98
TNR	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
FNR	94.59	80.00	98.25	97.44	97.73	95.16	98.59	98.33
LR+	294.96	414.00	124.45	120.86	420.49	428.00	193.74	77.69
LR-	0.96	0.85	1.04	1.02	1.03	0.98	1.03	1.04
DOR	3447.50	19412.00	3766.17	840.00	17200.00	18744.00	3663.00	249.37

Table B.9: Mean for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 1$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	84.70	79.82	43.46	42.88	66.38	66.41	48.35	42.40
FDR	15.30	20.18	56.54	57.12	33.62	33.59	51.65	57.60
TPR	72.99	71.56	28.71	23.28	58.93	57.22	24.47	19.77
FPR	5.22	5.58	4.98	4.89	4.75	4.01	2.58	4.66
F	76.75	74.74	32.08	27.30	61.16	60.42	29.05	23.96
G	77.75	75.20	33.84	29.79	61.87	61.09	32.12	26.93
AC	92.25	92.90	87.33	85.20	90.34	92.34	90.82	85.33
E	7.75	7.10	12.67	14.80	9.66	7.66	9.18	14.67
TNR	94.78	94.43	95.02	95.11	95.25	95.99	97.42	95.34
FNR	27.01	28.44	71.29	76.72	41.07	42.78	75.53	80.23
LR+	58.05	46.76	9.10	7.86	58.37	43.69	13.91	6.84
LR-	0.28	0.30	0.75	0.80	0.44	0.45	0.77	0.84
DOR	352.02	503.97	16.33	18.46	919.34	542.02	23.79	12.19

Table B.10: STDEV for the 13 dependent measures for distance-based store-to-store rules for four cities (d = 1km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	12.53	14.02	19.16	22.07	24.60	22.01	21.36	20.93
FDR	12.53	14.02	19.16	22.07	24.60	22.01	21.36	20.93
TPR	20.12	18.09	16.20	20.85	26.71	24.29	17.00	18.28
FPR	8.80	8.33	3.25	3.65	4.63	4.66	2.31	4.38
F	16.25	15.57	15.29	19.86	25.61	22.89	16.73	17.05
G	14.87	15.12	14.83	19.18	25.06	22.46	15.93	16.46
AC	6.60	5.19	4.10	4.87	7.71	5.70	3.69	5.33
E	6.60	5.19	4.10	4.87	7.71	5.70	3.69	5.33
TNR	8.80	8.33	3.25	3.65	4.63	4.66	2.31	4.38
FNR	20.12	18.09	16.20	20.85	26.71	24.29	17.00	18.28
LR+	58.12	71.44	11.04	11.98	93.63	69.42	13.99	7.66
LR-	0.20	0.18	0.16	0.21	0.29	0.25	0.17	0.18
DOR	401.65	1235.00	38.04	44.70	2810.69	2452.31	48.60	24.26

Table B.11: Mean for the 13 dependent measures for distance-based store-to-store rules for four cities (d = 1.6 km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	77.94	81.59	42.10	43.16	64.56	68.94	42.42	41.86
FDR	22.06	18.41	57.90	56.84	35.44	31.06	57.58	58.14
TPR	69.23	75.08	29.01	26.57	58.02	61.23	26.49	25.70
FPR	6.22	5.53	5.30	5.12	4.78	4.34	3.38	4.91
F	71.85	77.44	31.63	30.02	59.63	63.26	29.12	27.28
G	72.65	77.87	33.34	32.18	60.37	64.10	31.36	29.67
AC	91.15	93.36	87.13	86.10	90.44	92.17	90.57	87.04
E	8.85	6.64	12.87	13.90	9.56	7.83	9.43	12.96
TNR	93.78	94.47	94.70	94.88	95.22	95.66	96.62	95.09
FNR	30.77	24.92	70.99	73.43	41.98	38.77	73.51	74.30
LR+	46.83	58.30	8.18	7.79	52.88	56.05	12.90	8.46
LR-	0.33	0.26	0.75	0.77	0.45	0.41	0.76	0.78
DOR	399.76	571.95	14.89	17.45	749.53	769.65	26.29	18.21

Table B.12: STDEV for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 1.6$ km).

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	16.92	14.39	19.23	19.75	24.18	22.97	20.21	23.91
FDR	16.92	14.39	19.23	19.75	24.18	22.97	20.21	23.91
TPR	23.08	19.82	18.02	20.29	27.47	25.91	19.35	22.52
FPR	9.03	8.14	3.46	3.85	4.69	5.02	2.61	3.97
F	20.54	17.22	16.66	18.34	26.23	24.28	17.42	21.01
G	19.45	16.63	16.07	17.72	25.51	23.55	16.66	19.85
AC	7.14	4.99	4.02	4.99	7.51	6.17	3.63	4.97
E	7.14	4.99	4.02	4.99	7.51	6.17	3.63	4.97
TNR	9.03	8.14	3.46	3.85	4.69	5.02	2.61	3.97
FNR	23.08	19.82	18.02	20.29	27.47	25.91	19.35	22.52
LR+	56.24	78.93	9.76	8.98	88.69	88.62	16.04	9.87
LR-	0.24	0.20	0.18	0.20	0.29	0.27	0.20	0.23
DOR	748.69	1200.13	31.28	41.50	2186.02	2588.82	64.64	36.39

Table B.13: Mean for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	78.03	82.12	43.10	44.34	63.57	67.61	42.75	39.29
FDR	21.97	17.88	56.90	55.66	36.43	32.39	57.25	60.71
TPR	71.32	77.69	32.02	31.65	57.71	61.39	29.22	28.15
FPR	6.87	5.76	5.63	5.64	4.79	4.44	3.66	5.60
F	73.41	79.46	34.13	34.18	59.06	63.02	31.51	29.33
G	74.00	79.68	35.60	35.87	59.77	63.70	33.41	31.06
AC	91.09	93.38	87.00	86.62	90.53	92.09	90.75	87.15
E	8.91	6.62	13.00	13.38	9.47	7.91	9.25	12.85
TNR	93.13	94.24	94.37	94.36	95.21	95.56	96.34	94.40
FNR	28.68	22.31	67.98	68.35	42.29	38.61	70.78	71.85
LR+	41.46	62.55	8.81	9.14	47.61	57.03	15.06	7.92
LR-	0.31	0.24	0.72	0.72	0.45	0.41	0.73	0.76
DOR	362.09	810.84	32.07	26.42	648.23	805.00	54.67	18.05

Table B.14: STDEV for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 2$ km).

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	15.14	15.52	20.06	20.21	25.13	23.72	21.85	23.12
FDR	15.14	15.52	20.06	20.21	25.13	23.72	21.85	23.12
TPR	21.57	17.37	19.68	21.98	28.52	26.43	21.21	23.12
FPR	10.04	8.37	3.45	3.98	4.60	5.10	2.77	4.28
F	19.00	16.09	18.27	20.14	27.33	25.05	19.99	22.08
G	18.12	15.89	17.64	19.38	26.62	24.46	19.28	21.22
AC	7.24	5.53	4.58	5.19	7.17	6.49	3.99	4.87
E	7.24	5.53	4.58	5.19	7.17	6.49	3.99	4.87
TNR	10.04	8.37	3.45	3.98	4.60	5.10	2.77	4.28
FNR	21.57	17.37	19.68	21.98	28.52	26.43	21.21	23.12
LR+	50.79	84.06	11.90	12.90	82.43	91.72	22.31	10.50
LR-	0.22	0.18	0.20	0.22	0.30	0.28	0.22	0.24
DOR	652.36	2087.94	219.49	84.41	2024.20	2465.46	282.12	36.59

Table B.15: Mean for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 3$ km).

	Directional				Categorical			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	77.57	83.04	41.64	43.66	62.29	67.35	42.92	38.79
FDR	22.43	16.96	58.36	56.34	37.71	32.65	57.08	61.21
TPR	72.10	79.93	34.26	33.53	57.37	63.75	34.25	29.74
FPR	6.83	6.72	6.21	6.24	5.09	4.75	4.03	5.92
F	73.97	81.19	35.42	35.27	58.27	64.29	35.31	30.76
G	74.39	81.34	36.53	36.70	58.92	64.83	36.67	32.19
AC	91.26	93.17	86.77	86.24	90.24	92.23	90.95	86.91
E	8.74	6.83	13.23	13.76	9.76	7.77	9.05	13.09
TNR	93.17	93.28	93.79	93.76	94.91	95.25	95.97	94.08
FNR	27.90	20.07	65.74	66.47	42.63	36.25	65.75	70.26
LR+	44.75	60.94	8.17	8.52	44.30	56.10	16.67	7.70
LR-	0.30	0.21	0.70	0.71	0.46	0.38	0.69	0.75
DOR	406.39	821.01	30.41	26.53	627.30	787.60	60.15	20.90

Table B.16: STDEV for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 3$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	16.67	12.80	19.67	20.46	25.14	23.71	23.14	21.24
FDR	16.67	12.80	19.67	20.46	25.14	23.71	23.14	21.24
TPR	20.14	15.12	20.52	22.03	28.38	25.34	22.91	22.31
FPR	9.85	10.27	3.68	4.03	4.92	5.62	2.85	3.91
F	18.11	13.65	19.20	20.46	27.43	24.49	21.71	21.07
G	17.69	13.49	18.62	19.64	26.70	23.93	21.03	20.28
AC	7.16	6.08	4.92	5.16	7.38	6.71	4.46	4.83
E	7.16	6.08	4.92	5.16	7.38	6.71	4.46	4.83
TNR	9.85	10.27	3.68	4.03	4.92	5.62	2.85	3.91
FNR	20.14	15.12	20.52	22.03	28.38	25.34	22.91	22.31
LR+	58.21	85.20	10.87	11.90	77.04	88.41	28.90	10.13
LR-	0.21	0.16	0.22	0.23	0.30	0.27	0.24	0.23
DOR	787.89	2011.50	197.61	87.79	2033.06	2397.80	294.33	54.57

Table B.17: Mean for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 4$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	76.57	83.54	41.46	43.62	60.32	68.09	41.40	37.59
FDR	23.43	16.46	58.54	56.38	39.68	31.91	58.60	62.41
TPR	70.07	79.82	35.54	35.60	56.42	64.25	33.61	28.90
FPR	6.50	6.38	6.61	6.67	5.20	4.47	4.27	5.89
F	72.20	81.30	36.41	36.84	57.07	65.18	34.27	29.78
G	72.73	81.49	37.35	38.06	57.64	65.65	35.60	31.18
AC	91.12	93.30	86.64	85.99	90.12	92.61	90.73	86.88
E	8.88	6.70	13.36	14.01	9.88	7.39	9.27	13.12
TNR	93.50	93.62	93.39	93.33	94.80	95.53	95.73	94.11
FNR	29.93	20.18	64.46	64.40	43.58	35.75	66.39	71.10
LR+	41.51	55.78	8.44	8.17	43.84	55.01	14.37	7.47
LR-	0.32	0.22	0.69	0.69	0.47	0.38	0.70	0.76
DOR	377.33	805.84	30.78	25.91	621.55	737.81	46.71	21.84

Table B.18: STDEV for the 13 dependent measures for distance-based store-to-store rules for four cities ($d = 4$ km).

	Directional				**Categorical**			
	City1	**City2**	**City3**	**City4**	**City1**	**City2**	**City3**	**City4**
P	17.59	13.80	20.05	19.95	26.08	23.57	22.47	21.67
FDR	17.59	13.80	20.05	19.95	26.08	23.57	22.47	21.67
TPR	22.75	16.63	20.27	21.59	28.95	25.71	22.05	23.02
FPR	9.30	9.73	4.00	4.05	4.96	5.31	2.97	4.05
F	20.37	15.01	19.34	20.02	28.24	24.55	20.90	21.76
G	19.74	14.77	18.89	19.32	27.62	24.13	20.25	20.98
AC	7.23	6.06	5.14	5.28	7.47	6.40	4.38	4.94
E	7.23	6.06	5.14	5.28	7.47	6.40	4.38	4.94
TNR	9.30	9.73	4.00	4.05	4.96	5.31	2.97	4.05
FNR	22.75	16.63	20.27	21.59	28.95	25.71	22.05	23.02
LR+	50.95	67.53	11.56	10.73	77.37	82.50	20.04	9.97
LR-	0.24	0.17	0.22	0.22	0.31	0.27	0.23	0.24
DOR	724.29	1918.41	193.31	83.56	2090.96	2148.37	226.74	58.19

Table B.19: Values of 4 independent measures for brand-to-brand rules for City1 and City2.

	Directional		**Categorical**	
	City1	**City2**	**City1**	**City2**
TP	3,739	4,186	2,506	18,968
FP	302	215	699	3,371
FN	495	241	912	385
TN	3,960	4,369	11,936	531
Total	8,496	9,011	16,053	14,681

Table B.20: First quartile for the 13 dependent measures for brand-to-brand rules for City1 and City2.

	Directional		**Categorical**	
	City1	**City2**	**City1**	**City2**
P	85.00	87.85	37.05	31.41
FDR	3.09	1.67	11.74	3.32
TPR	70.3	87.55	16.69	20.21
FPR	2.39	1.85	0.51	0.28
F	79.86	85.92	16.91	21.43
G	80.58	86.18	18.53	25.02
AC	85.86	95.14	86.10	92.68
E	5.67	3.81	4.19	2.06
TNR	77.733	92.56	92.62	96.27
FNR	2.97	3.22	12.60	7.59
LR+	4.38	12.52	2.44	7.64
LR-	0.05	0.03	0.19	0.09
DOR	49.20	248.82	6.10	13.92

Table B.21: Third quartile for the 13 dependent measures for brand-to-brand rules for City1 and City2.

	Directional		**Categorical**	
	City1	**City2**	**City1**	**City2**
P	96.91	98.33	88.26	96.68
FDR	15.00	12.15	62.95	68.59
TPR	97.03	96.79	87.40	92.41
FPR	22.27	7.44	7.39	3.74
F	94.76	97.55	86.28	94.07
G	94.85	97.57	86.29	94.07
AC	94.33	96.19	95.81	97.94
E	14.14	4.86	13.91	7.32
TNR	97.62	98.15	99.49	99.72
FNR	29.70	12.45	83.32	79.79
LR+	29.64	46.53	32.29	44.10
LR-	0.31	0.17	0.87	0.81
DOR	158.67	714.15	39.68	97.62

Table B.22: Maximum value for the 13 dependent measures for brand-to-brand rules for City1 and City2.

	Directional		**Categorical**	
	City1	**City2**	**City1**	**City2**
P	98.59	99.74	100.00	100.00
FDR	26.98	35.92	95.65	85.71
TPR	99.48	98.77	100.00	97.68
FPR	54.47	66.07	50.00	31.76
F	97.06	98.18	98.80	97.44
G	97.06	98.19	98.80	97.44
AC	95.13	97.25	99.79	99.58
E	25.21	33.07	23.09	13.39
TNR	99.73	99.75	100.00	100.00
FNR	53.72	30.14	98.31	96.67
LR+	259.51	279.45	431.00	421.90
LR-	0.55	0.30	0.99	0.99
DOR	845.47	2326.15	4280.00	16416.00