STOCKHOLM SCHOOL OF ECONOMICS

Decision support systems

Applying analytics to sales data

Oscar Andersin 50097

5/22/2013

Thesis supervisor: Mikael Hernant

Abstract

In this thesis I propose to test the applicability of certain statistical methods as decision support on a grocer's sales data. The sales data has been obtained from a grocer in Sweden and consists of a basket of 8 products and their sales data for 2012 on a daily basis. The method will be time series analysis using dependence, causality, trends and forecasting. As a result I obtained that there is a high dependence between products, the store general sales and number of customers affects products at a high rate, store past performance has an effect on products, products have differing trends and different characteristics from each other and that it is possible to forecast future sales. This means that analytics are well within the range of decision support and can be used by a sales agent to successfully monitor and analyse specific products. However, analytics have a limited use when analysing all the products in a retail store at once as the cost of gathering and using that information is too high. I therefore propose that further research could propose automated decision methods on how to simplify my proposed methods.

"Data is not information, Information is not knowledge, Knowledge is not understanding, Understanding is not wisdom."

- Cliff Stoll & Gary Schubert (Stoll, 1995)

"I never guess. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

- Sir Arthur Conan Doyle (1859 - 1930) (Doyle, 1890)

"We are awash in data and starved for information."

– A vice president of merchandise technology at a department store chain (Fisher, 2010)

Contents

Abstract 1
1.0 Introduction 5
1.1 Problem area 5
1.2 Purpose of thesis 6
1.3 Motivation of thesis 6
1.4 Demarcation of thesis 7
1.5 Thesis contribution 7
1.6 Composition of thesis 7
2. Theory 8
2.1 Decision support system theory 8
2.2 Statistical theory 12
2.3 Simple and Multiple Regression 17
2.4 Economic Forecasting 20
3. Earlier research 21
3.1 Decision support systems 21
3.2 Time series analysis on sales data 23
4. Data 25
4.1 The products 25
4.2 Descriptive statistics 25
4.3 Data source criticism 27
5. Method 29
5.1 Test of Dependence 29
5.3 Test of Data Stationarity 29
5.3. Test of Cointegration 29
5.2 Multiple Regressions on the Data 30
5.2 Test of Causality 30
5.5 Economic Forecasting 30
5.9 Hypothesis testing 31
6.0 Results 32
6.1 Test of Dependence 32
6.2 Test of Unit Root 33
6.3. Test of Cointegration 34

6.2 Multiple regression 34
6.2 Test of Causality 30
6.6 Economic Forecasting 31
6.7 Hypothesis results 35
7.0 Discussion 36
7.1 General discussion 36
7.2 Consequences for decision support 36
8.0 Conclusion 37
8.1 Further research 37
9.0 Attachments 38
10.0 Bibliography 41

1.0 Introduction

Product sales offer an interesting challenge for statistics and decision support systems. On one hand you have more data than you ever need about a product but on the other hand you need to weigh that information with what is practical to use by the sales agent within a limited time frame.

Until a few years ago all of this information was calculated on crude excel spread sheet from data recorded from the cash-registers. The people working with the data did not have the analytical tools available today or they did not understand it proficiently to use it (Gustke, 2013).

Enter the "Big Data" revolution. Business spectator defines it as "the process of examining large amounts (volume) of data of a variety of types (big data) to uncover hidden patterns, unknown correlations and other useful information". This data is used in real time and has led to increased competitive advantage and increased customer experience and higher revenue according to the newspaper (Wallace, 2013).

Fisher and Raman are on a similar path in their book "The New Science of Retailing". They propose that Analytics are transforming the retail atmosphere (Fisher, 2010). Paul Schottmiller, at Cisco Consulting Services estimates that Big Data could bring a 54 percent after-tax profit after Big Data is implemented (van Zanten, 2012).

In the article from Cisco the authors identify four characteristics of Big Data: volume, velocity, variety and value. Each characteristic offers some challenges but according to the authors the difference between praxis today and Big Data is:

- *"Today, retailers may know what customers want; with Big Data, they will predict what customers want.*
- *Today, retailers provide customers with the ability to research, decide, and buy; with Big Data, customers will be able to do this when and where they experience the need.*
- *Today, retailers provide products; with Big Data, retailers will offer an experience that is unique to the individual customer over time.*
- *Today, retailers strive to create a pre- and post- purchase relationship with the customer; Big Data will enhance and solidify this relationship immeasurably."* (van Zanten, 2012)

Failure to use this analytical information to review customer behavior and sales can be catastrophical as could be seen by J.C.Penney's failed store changes that created a massive downward crash in sales (Thau, 2013).

1.1 Problem area

All over the retail world data and data mining are revealing patterns that hitherto have remained elusive. By interpreting and analysing the data collected from an individual's interaction with the retail chain the analyst can make better estimates of the retail valuation, assortment planning, product life cycles, supply chains, aligning goals, product and associate availability and technological risk (Fisher, 2010).

The problem is that the store manager has limited tools and knowledge to make use of this data on an in-store level. Many of these methods remain complicated and time consuming to be made regularly and often this data is only analysed on an aggregated basis by the company's headquarters. Thus, the information to store managers is infrequent and lacks the speed and intuition that many store manager's use in their daily work in the store.

Aggregated information has many uses for a store chain. It helps the chain to determine the strategic decisions to be made for the whole company so as to be competitive in the market and it gives the store chain the means to calculate overall macroeconomic trends. It also gives the company managers decision support about the company's stores and their profitability. Decisions on aggregated information might be concerned with resource allocation and which stores to shut down and which stores to keep.

But a store manager is concerned with the sales and placement of certain products within his or her store. What his cash registers collect in data is sales of certain products, at certain dates over a time period for a certain price and margin. Often this information is just used on a limited basis by the store manager to check the level of sales for a few days. This is unfortunate as some statistical methods can show dependencies not yet known to the store manager as well as help in forecasting the sales better than the retail managers own experienced guesstimates. Statistical analysis can also help the store to reduce inventories, plan purchases and reduce waste. Even if the store uses purchase computers to estimate future inventory need the store manager can benefit from knowing how the computer calculates these estimates.

In this thesis I propose to use statistical methods on sales data to show what measures and methods can be made on this data and why it is of relevance to store managers as decision support. After I have proposed a method I will discuss its attributes, failures, speed, intuition and utility for a store manager within a limited time frame. The idea is to find measures and methods that can be used in praxis by the store manager. The data will be provided by a grocery store and contain a time series of sales volume data, margins and revenue of a basket of goods.

Generally I want to answer the question "What can time-series analysis tell us about products and the store?"

1.2 Purpose of thesis

The purpose of the thesis is to test the applicability of statistical measures and methods on in-store sales data and its use as decision support.

1.3 Motivation of thesis

My motivation to study time-series and decision support is for me to learn the characteristics and patterns that product time series might show.

1.4 Demarcation of thesis

The study will not focus on the products or the promotional effects of discounts, placements, or signage or shelf space height, width or number of facings. The thesis will also neglect other "Big Data" intermediaries such as Twitter, Facebook other social media sites and the internet. It is purely interested in the statistical sales time series and its applicability to provide knowledge.

The study is also not a comparison of different decision or managerial support systems. It is only interested in the information that can be given by statistical analysis to those systems and to what value it is to the systems.

1.5 Thesis contribution

The thesis will be useful to anyone designing decision support systems for retailing and trying to figure out what information is possible to gain from pure sales data. It can also be of relevance to producers trying to interpret the sales pattern for their products to gain additional information. Lastly the thesis summarizes many statistical methods that are useful in business academia.

1.6 Composition of thesis

The thesis first takes into account general theories in the subject of decision support systems as well as general theories about statistics regarding time series. After that I will present earlier research into the subject and highlight some studies that have prevalence for the subject of this thesis.

The data chapter will show descriptive statistics on the data received from a grocer. The descriptive statistics will focus on distributions, variance and trends in the data. A discussion on how reliable the data is and factors that could influence the basket and the individual products will be presented in the chapter.

The method chapter will show what statistical methods and measures are available to pursue on the data collected by the grocer. It will give a short introduction to each method and the general assumptions made for the applicability of the method.

Results will show what the methods revealed and the following discussion will present the pros and cons to each method for the store manager.

2. Theory

The theory presented in this thesis will be based on theories of decision support systems and statistical methods and theories. The idea is to use the theories about decision support systems to explain the need of statistics in decision making and discuss the demands that these systems place on statistics in a grocery store.

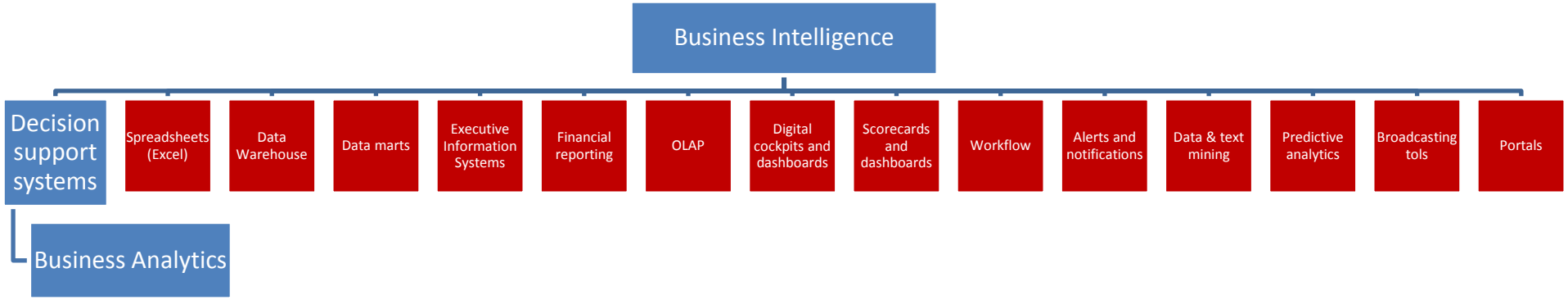
2.1 Decision support system theory

Decision support systems (DSS) can be defined as "interactive computer-based systems, which help decision makers utilize data and models to solve unstructured problems" (Gorry, 1971). Another definition is:

"Decision support systems couple the intellectual resources of individuals with the capabilities of the computer to improve the quality of decisions. It is a computer-based support system form management decision makers who deal with semi structured problems." (Keen, 1978)

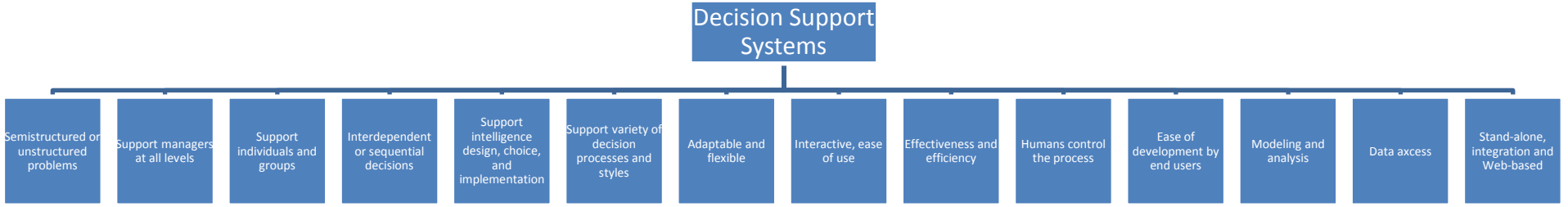
Keen et al (1978) defines the purpose of decision support systems as the process of building systems for key decisions and tasks to improve effectiveness of the manager's process of solving problems (Keen, 1978).

Decision support systems evolved in the 1990s to a part of a bigger picture of all the systems within a company called Business Intelligence. Business Intelligence is defined as "a umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies. The purpose of Business Intelligence is to give employees access to interactive data, enable changes in this data and to give employees data for analysis. In other words the purpose is to make employees better informed to make better decisions (Turban, 2011)



Figur 1:Evolution of Business Intelligence (Turban, 2011)

Today there is a lot more to Business Intelligence than just decision support systems. But the main focus of this thesis is to investigate how business analytics can provide support for decision making. We will therefore only be discussing the blue processes. Notice that Predictive Analytics have been given a own area of Business Intelligence. In my study it will be featured under Business Analytics.



Figur 2: Characteristics and Capabilities of Decision Support System (Turban, 2011)

2.1.1 Characteristics and Capabilities of Decision Support Systems

As seen in the in Figure 2 DSS have certain characteristics and capabilities according to Turban et al (2011). These are the following:

- DSS gives decision support for semi-structured and unstructured problems by bringing together human judgment and computerized information.
- They can give support to all managerial levels
- They can give support to individuals as well as teams
- They can give support for interdependent or sequential decisions once or on a repeated basis
- Support in all phases of the decision making process
- Support to various processes and styles
- The decision maker should be reactive and adapt the DSS to these changes
- The DSS should be user friendly with graphical capabilities
- It should improve the effectiveness of decision making
- The user should have complete control over all steps in the decision making process
- The end user should be able to modify and develop the system by themselves
- Models are used to analyse the decision making situation
- Access is provided to a wide variety of data sources, formats and types
- The DSS should be a stand-alone tool used by a single user at a single location (Turban, 2011)

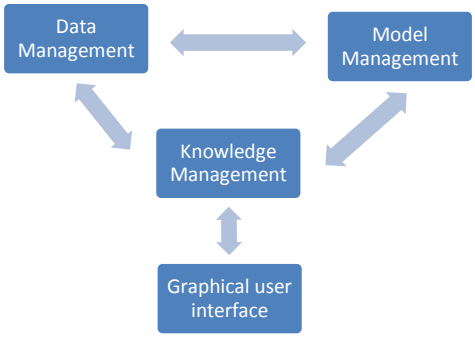
2.1.2 Criteria of successful decision support systems

Generally it can be said that in successful decision support systems the utility of information should exceed or equal the cost to gather said information.

$$U(I) \geq C(I)$$

This is very true in the retail industry where there is an onslaught of data pouring in from the cash registers almost at a minute-by-minute basis. To filter through all of this information a retailer needs to have an effective decision support system to analyse the data.

2.1.3 Frameworks for analysis



The framework above shows the interaction between the data, the models, the user knowledge and the user interface in a decision support system (Turban, 2011).

2.1.4 Business Analytics

According to Turban et al (2011) Business Analytics "implies the use of models and data to improve an organisation's performance or competitive posture". The author emphasises the use of models but recognizes that the use is limited since few users know how and when to apply models to interpret and describe a certain situation. That question will be answered by this thesis.

Predictive analytics is described by Turban et al (2011) as a "method of forecasting problems and opportunities rather than simply reporting them as they occur". The author then states that predictive analytics uses advanced forecasting and simulation models (Turban, 2011). Predictive analytics are used to give an estimate of what might happen with a product sales in the future so as to give a retail valuation, do assortment planning, estimate product life cycles, optimize supply chains as well as aligning goals, product and associate availability and technological risk.

Keen et al. (1978) argues that organisations can benefit from analytical thinking instead of just the models and applications. This analytical thinking constructs other models and uses logic to see the data in a different way. (Keen, 1978).

2.1.5 Effective decision-making

Effective decision-making can be defined as: "identifying what should be done and ensuring that the chosen criterion is the relevant one". This should be separated from efficient decision making that is "performing a given task as well as possible in relation to some predefined performance criterion". In other words you need to choose your performance metrics and then try to achieve them as well as possible (Keen, 1978).

It is not enough to provide a lot of decision support; you also need to know what is relevant. This can be described by the authors Fisher et al. (2010) when they encounter a vice-president of planning who produced a weekly eight-inch thick computer printout. No one can filter that much information in a week (Fisher, 2010).

Based on our first general conclusion we can extend it to an improved decision making utility formula:

$$\Delta U(I) = \Delta C(I)$$

In other words the decision making system should be improved until marginal utility equals marginal cost of information.

2.1.6 The management science approach and automated decision making

DSS should be differentiated from automated decision making which is done on routine problems facing the company. The management science (MS) approach has a ready four-step systematic process to solving these routine problems. This is a mathematical modelling

approach to a problem which can be used in programming. The steps are the following (Turban, 2011):

- Define the problem
- Classify the problem into a standard category
- Construct a model that describes the real-world problem
- Identify possible solutions to the modelled problem and evaluate the solutions
- Compare, chose, and recommend a potential solution to the problem

The automated decision systems (ADS) are a new way of dealing with repetitive managerial problem. It is a rule-based system that provides standardized solutions within a certain functional area. The final decision in a ADS is always left to a human. They are best suited for front-office employees who must make quick decisions from customer online information (Turban, 2011).

2.1.7 Distinction to management accounting systems

Management accounting can be defined as "*Accounting concerned with the provision of information to people within the organization to aid decision-making and improve the efficiency and effectiveness of existing operations"* (Drury, 2012). As you can see it has almost the same goal as decision support systems.

The difference is that decision support systems are more operational than management accounting systems. In a way you can say that management accounting systems deals with more aggregated data than decision support systems.

2.1.8 Technological risk

As with any technology there is some technological risk. This risk comes mainly from the fact that technology develops so fast that a system today might be obsolete tomorrow.

Another source of risk is that the interactivity with other systems is poor. A system should be developed so as to make it adaptable to technological development and easily transformed into the other format. It should also be constructed is such a way that it has a high ability to interact with other systems.

2.2 Statistical theory

2.2.1 Statistical Dependence

Two variables are dependent when the values of one variable influence the values of another variable. Dependent variables can be either the same variable tested twice or matched pairs. Matched pairs should be as alike as possible so that a comparison can be made (Newbold, 2010).

According to Newbold et al (2010) covariance is a measure of the linear relationship or the direction of a linear relationship between two variables. If it is positive it indicates a direct or increasing relationship and if it is negative it indicates a decreasing linear relationship. Covariance does not give us a measure of the strength of the relationship between two variables.

Newbold et al (2010) suggest that the correlation coefficient gives us a standardized measure of the relationship between two variables and overcomes the shortcomings of covariance. Correlation thus gives us both the strength and the direction of a relationship.

2.2.2 Causality

But the existence of a dependency between variables does not imply causation. According to the Oxford Dictionary causality is 1) the relationship between cause and effect and 2) the principle that everything has a cause (Oxford Dictionaries, 2013). This means that one variable X might be the cause of variable Y and thus the source of the dependence. According to Gujarati et al (2009) causality is a deeply philosophical debate and should be founded in theories and common sense outside the realm of statistics.

But because time does not run backwards we can assume that if event X happened before event Y there is a chance that X influenced Y. The same does not hold the other way around since X happened before Y. A more precise statement might be that X might contain useful information for predicting Y (Koop, 2008).

2.2.3 Stochastic processes

According to Gujarati & Porter a stochastic process is a collection of random variables ordered in time. If Y is the random variable then Y_1 is that random variable at point 1 in time. This can be extended to include $Y_2 Y_3 \ldots . Y_{52}$ and so on. A particular variable Y_{182} is a realisation of the stochastic process or a sample of that process. By using these realizations we can draw inferences about the underlying stochastic process.

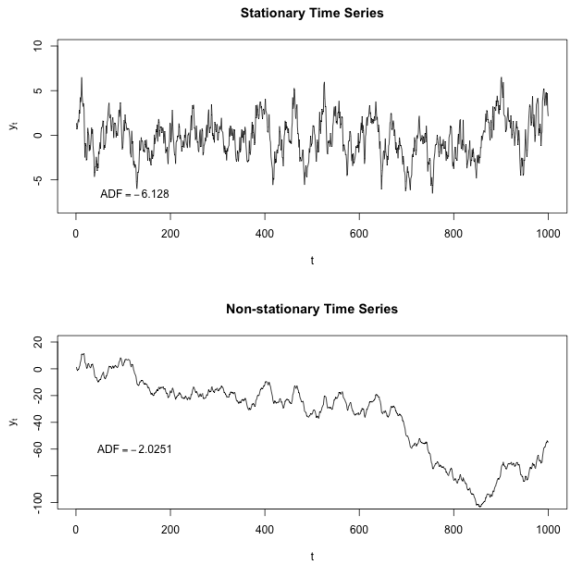


Figure 1: Examples of Stationary and Non-stationary time series. Picture from Wikipedia.

A stochastic process can be either stationary or non-stationary. A stochastic process can be defined as stationary *"if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed"* (Damodar Gujarati, 2009).

What this means is that the mean and variance should be constant while the covariance measure should not depend at what time T the covariance is measured. The measure k measures the time dependence, the distance between two values of Y in time. For the stochastic process Y_t above (Damodar Gujarati, 2009):

$$Mean: E(Y_t) = \mu$$

$$Variance: var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$

$$Covariance: \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$$

A stationary time series tends to fluctuate around its mean and mean revert, which means that as a function of time the stochastic process returns from outliers to its mean. The reason is mainly because the variance is finite. But the speed of the mean reversion is a function of the autocovariance (covariance at various lags). If the autocovariance is small the process is quick and if they are large the process is slow.

A non-stationary time series is defined as having a mean that varies with time or a variance that varies with time or both. That means that width and the direction of the time-series might vary with time.

The reason we are interested in the stationarity of a stochastic process is because of its usefulness in forecasting. If a process is non-stationary we cannot make assumptions of the data but for the particular episode. With a stationary process we can use the data to generalize past and future performances.

2.2.4 Random Walk

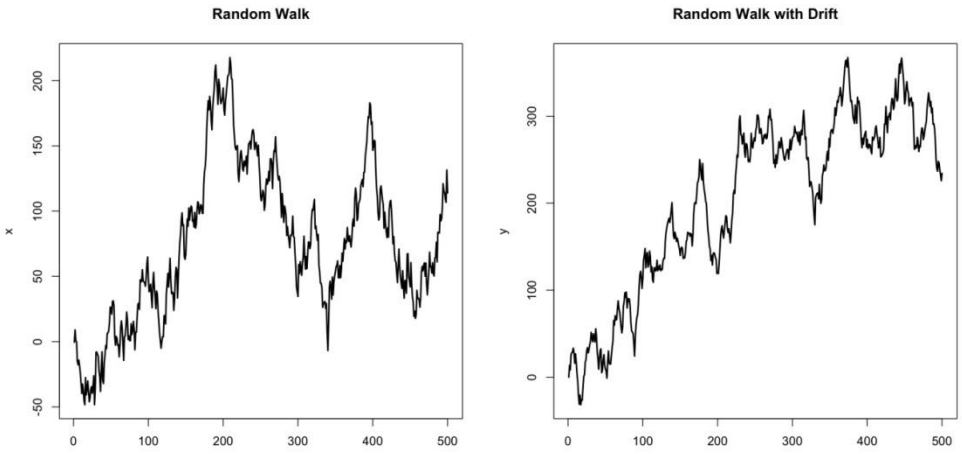


Figure 2: Examples of Random Walk without drift and Random Walk with drift. Picture from http://independentobservations.wordpress.com

There are two types of non-stationary stochastic processes. One is a random walk without drift and the other is a random walk with drift. The drift is defined as a rate of change in the time series.

2.2.5 Random walk without drift

A stochastic process is a random walk without drift if: (Damodar Gujarati, 2009)

$$Y_t = Y_{t-1} + u_t$$

Where the value of Y at time t is equal to its value at time (t – 1) plus a random shock u_t. From this logic it follows that:

$$Y_1 = Y_0 + u_1$$

$$Y_2 = Y_1 + u_2 = Y_0 + u_1 + u_2$$

$$Y_3 = Y_2 + u_3 = Y_0 + u_1 + u_2 + u_3$$

And from this we get:

$$Y_t = Y_0 + \sum u_t$$

$$E(Y_t) = E\left(Y_0 + \sum u_t\right) = Y_0$$

This means that the mean of Y is equal to its initial value at time t=0. The width of the time series thus varies around its starting value in a random walk (Damodar Gujarati, 2009). But the width or variance is an increasing function of time as can be shown by the expected variance:

$$var(Y_t) = \mathrm{t} \cdot \sigma^2$$

In a random walk there is an endurance of random chocks defined by u_t which are added to the last value of Y_{t-1}. This means that the chock lives on and the random walk never forgets a chock. A random walk will thus be forever upset by such a chock.

2.2.6 Random walk with drift

A random walk with drift is a random walk process but with a rate of change. Let δ be the drift parameter. Random walk with drift can then be written as (Damodar Gujarati, 2009):

$$Y_t = \delta + Y_{t-1} + u_t$$

And for clarity:

$$Y_1 = \delta + Y_0 + u_t$$

$$Y_2 = \delta + Y_1 + u_2 = \delta + \delta + Y_0 + u_1 + u_2$$

$$Y_3 = \delta + Y_2 + u_3 = \delta + \delta + \delta + Y_0 + u_1 + u_2 + u_3$$

And as such Y_t will drift upward or downward according to the parameter being positive or negative. And we can expect that:

$$E(Y_t) = Y_0 + t \cdot \delta$$

$$var(Y_t) = \mathrm{t} \cdot \sigma^2$$

In other words, the mean and the variance of Y_t increases over time violating the conditions of stationarity (Damodar Gujarati, 2009).

2.2.7 Unit Root Stochastic Process

A unit root is a process that evolves trough time and can cause problems in making inferences about the data. If we write a random walk as (Damodar Gujarati, 2009):

$$Y_t = \rho Y_{t-1} + u_t \qquad -1 \leq \rho \geq 1$$

And ρ has a value of 1 we have a random walk without drift and we have a non-stationary time series because Y_t is non-stationary (see above for explanation). If in fact $|\rho|$ is <1 then the process is stationary (Damodar Gujarati, 2009).

2.2.8 Deterministic or stochastic trend

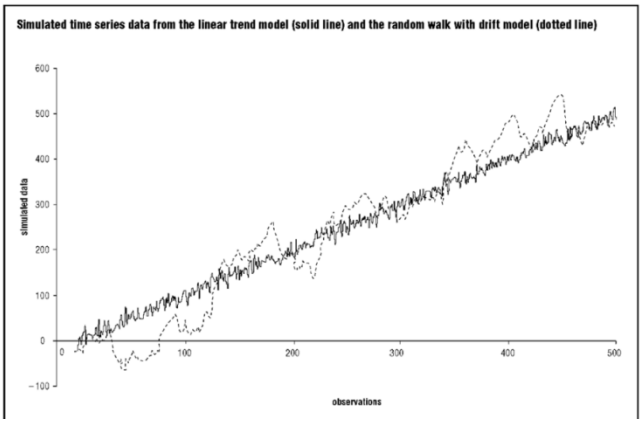


Figure 3: Dotted line is a stochastic trend; filled line is a deterministic trend. Picture from whatwhenhow.com

If a trend in the time series is a function of time it is called a deterministic trend. But if it is not a function of time and thus not predictable we call it a deterministic trend. In this chapter I will present some these trends for a model as suggested by Gujarati et al. (2009). For the model:

$$Y_t = \beta_1 + \beta_2 + \beta_3 Y_{t-1} + u_t$$

For a pure random walk $\beta_1 = 0, \beta_2 = 0, \beta_3 = 1$ so that:

$$Y_t = Y_{t-1} + u_t$$

Random walk with drift $\beta_1 \neq 0, \beta_2 = 0, \beta_3 = 1$ so that

$$Y_t = \beta_1 + Y_{t-1} + u_t$$

Deterministic trend $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 0$ so that

$$Y_t = \beta_1 + \beta_2 t + u_t$$

Random walk with deterministic trend $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 1$ so that

$$Y_t = \beta_1 + \beta_2 t + Y_{t-1} + u_t$$

Deterministic trend with a stationary component $\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 > 1$

$$Y_t = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + u_t$$

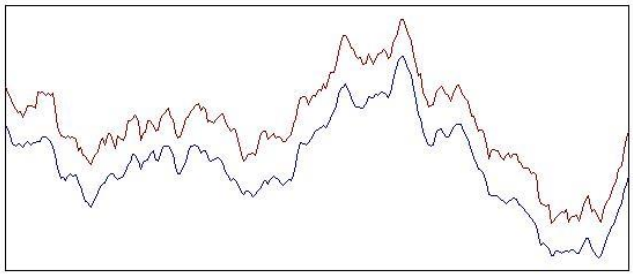
(Damodar Gujarati, 2009)

2.2.8.1 Spurious regression

A spurious or nonsense regression arises when regressions for two variables X and Y, that have no logical relationship, still produces results. A spurious regression can arise in nonstationary time series even when the sample is very large. It is also indicated by a high autocorrelation and thus by a low Durbin-Watson d value. A good rule of thumb is that the Durbin-Watson d value should be bigger than R^2 to not have a spurious regression (Damodar Gujarati, 2009).

2.2.9 Cointegration

Cointegration is when two variables have a long-term, or equilibrium, relationship between them. This relationship often takes the form of a similar stochastic trend (Damodar Gujarati, 2009).



Figur 3: Two cointegrated series

This is interesting since a regression of a nonstationary variable against another nonstationary variable might cancel out the stochastic trend in the two series. As a result a regression is possible and so the cointegration test might make us avoid spurious regressions (Damodar Gujarati, 2009).

2.3 Simple and Multiple Regression

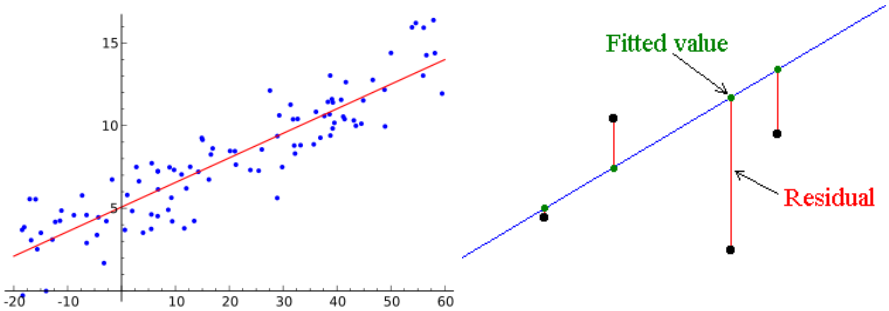


Figure 4: The best fitting line in a simple regression. Picture from http://www.shaunagm.net and http://www.le.ac.uk

A linear regression is a model that computes coefficients for the best fitting linear relationship given a set of data points in a coordinate system. Regressions can be used to compute how two or more variables interact and the strength and direction of that relationship (Koop, 2008) (Newbold, 2010).

2.3.1 Simple regression

The regression line is the best possible indicator of a linear relationship between two variables but it is not perfect. It actually fits a variable to adjust to the rest of the data. Because of that we need to add an error variable or residual at the end of the function for all the data that the regression line does not touch and the distance of the line to the error variable. We write the model as (Koop, 2008):

$$Y = \alpha + \beta X + \varepsilon$$

Where Y is the dependent or the output variable and where X is the independent or the cause variable. The intercept α is the starting value on the Y axis (5 in the above example) and β is the slope of the curve. β is also a measure of the sensitivity and influence of X on Y. The error variable is ε. Since we can never find the true line and only an estimate for them we need models to estimate the fit of the regression line. For a simple regression this method is called ordinary least squares. The assumptions for a simple regression are: (Koop, 2008).

1. $E(Y_i) = \beta X_i$ We expect Y_i to lie on the regression line on average.
2. $var(Y_i) = \sigma^2$ We expect all observations to have the same variance or homoscedasticity.
3. $cov(Y_i, Y_j) = 0 \; for \; i \neq j$ The observations are uncorrelated
4. $(Y_i) \; is \; normally \; distributed$
5. $X_i \; is \; fixed. It \; is \; nota \; random \; variable$

2.3.2 Multiple regression

A multiple regression is a simple regression with two or more independent variables. Basically it is similar but with important differences in interpretation and in choosing the variables.

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + \varepsilon_i$$

Where Y is the dependent or the output variable and where X_j are the independent or the cause variable. The intercept is still α and β_j is still the slope of the curve. Contrary to intuition β_j is calculated as the marginal effect of X_j on Y holding all other explanatory variables constant or equal. For a multiple regression R^2 is calculated differently by Adjusted R^2.

2.3.3 Normality

A normal distribution is a probability distribution of a continuous random variable. In other words it is the frequencies of values around a mean (σ) that arises with a certain probability when the sample size gets larger. Thus forming a "bell" formation.

The highest amount of frequencies usually arises around the mean and less frequencies the further you go from the mean. One standard deviation can be measured as the mean of variation from the sample mean. If you sum two means of the sample you get the variance (σ^2) of the sample.

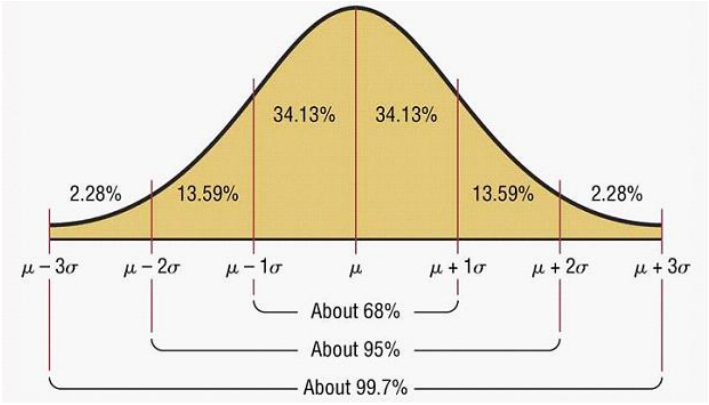


Figure 5: Properties of Normal-distribution. Picture from rchsbowman.wordpress.com

The normal distribution is used in economics and business situations because it approximates the probability distribution of many random variables. The sample also approaches a normal distribution if it is large enough and its probabilities are easy to calculate (Newbold, 2010).

2.3.4 Multicollinearity

Multicollinearity arises when two or more independent variables in a regression are highly correlated. This creates a problem because the two variables move at the same pace and captures the same effects. The probability of multicollinearity can be examined in a correlation matrix but in a regression it can be indicated by regression coefficient differing considerably from expected values or signs, a variable has a small student t statistics but strong influence, the t statistics indicate low effect but the F statistic is high (Newbold, 2010).

2.3.5 Heteroskedasticity and Homoskedasticity

A linear regression models where the error terms do not have the same variance and exhibit time varying trends are said to be heteroskedastic. Homoskedastic is when the error terms do have the same variance (Newbold, 2010).

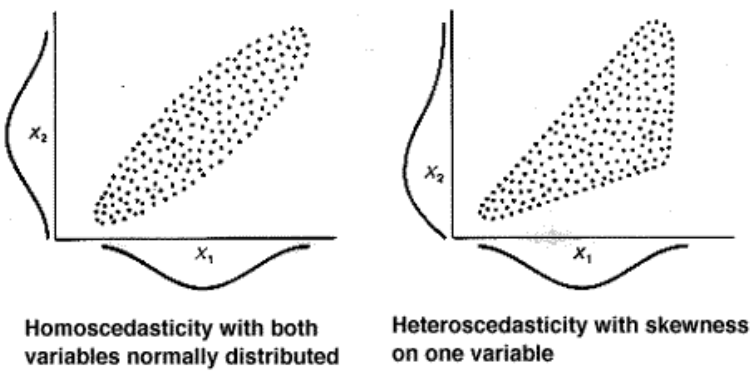


Figure 6: Heteroskedsticity and Homoskedasticity of data. Picture from http://bobhall.tamu.edu/Correlation/

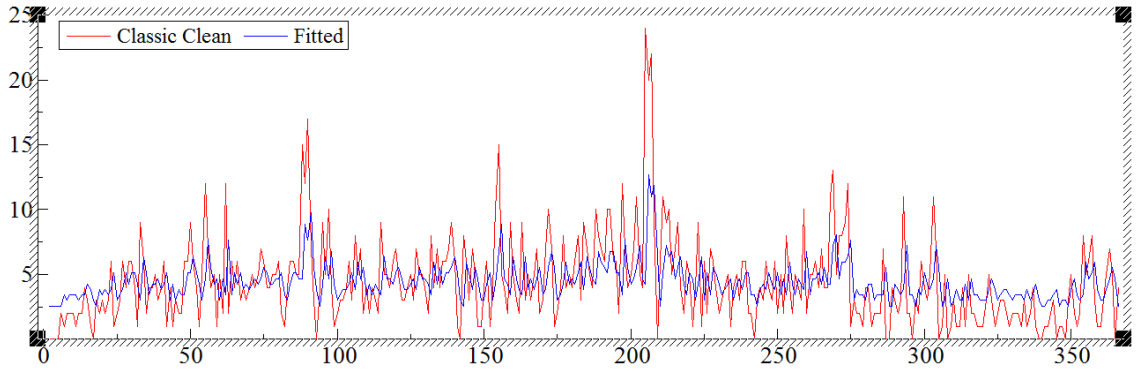
While the presence of heteroskedasticity changes the linear model it is only an indication that the ordinary least squares might not give the unbiased and consistent estimation (Damodar Gujarati, 2009).

2.3.6 Autocorrelation

Autocorrelation is correlation patterns between adjacent periods of the same data. If we assume that all variables in a time series are a random function of time the series has no autocorrelation. But if the time series starts to exhibit dependence on its past levels the series has autocorrelation. If a random variable does not exhibit autocorrelation it is not predictable. Extreme autocorrelation would mean that the past values are a straight line. Some models assume that there is no autocorrelation (Newbold, 2010).

2.4 Economic Forecasting

The goal of economic forecasting is to create a model that fits the actual data as well as possible and then rolling that model forward from the last available entry to predict the future. The problem with forecasting is that a forecast becomes worse the further it gets from the last available entry.



There are many forecasting models and every model takes into account special circumstances in the data. Examples of such models are ARCH, GARCH, ARMA, ARFIMA and so on. But they all have one thing in common, that is that they all start with a simple autoregression. To keep my thesis simple I will only attempt to forecast by rolling forward a simple AR(1) model as explained in the next chapter.

2.4.1 Autoregressive models

An autoregressive model is regression model where the explanatory variables are lags of the dependent variable. In other words the autoregression is a regression of a variable of itself with the historical values at certain intervals. A simple AR (1) model can be written as (Koop, 2008):

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

As you can see there is no independent variable X in this regression. Only Y lagged one time. The variable ρ represents the autocorrelation function between the variable Y_t and the variable lagged one time Y_{t-1} (Newbold, 2010). Autoregressive models are used to find repeating patterns in data or hidden patterns that are buried in the data.

3. Earlier research

Here I will present earlier research into decision support systems and time series analysis on sales data. What is interesting to note is that articles on Decision Support Systems are not published often and that the research into time series analysis has progressed so far that it is well beyond what we have learned in school. Many articles focusing on time series today incorporate computers which makes the academic discussion very mathematic and algorithm based.

3.1 Decision support systems

3.1.1 Gorry and Scott-Morton (1971)

Gorry and Scott-Morton first started the discussion by establishing a framework for management information systems in 1971. It discussed the categories of decisions are made within an organisation and what system can best support each category.

The decision categories according to Gorry and Scott-Morton are:

- Strategic planning
- Management control
- Operational control

They define strategic planning as "the process of deciding on objectives of the organization, on changes in these objectives, on resources used to attain these objectives, and on the policies that are to govern the acquisition, use, and disposition of these resources". Management control is defined by the authors as "the process by which managers assure that resources are obtained and used effectively and efficiently in the accomplishment of the organisation's objectives". Lastly they define operational control as "the process of assuring that specific tasks are carried out effectively and efficiently.

To this research they developed two models:

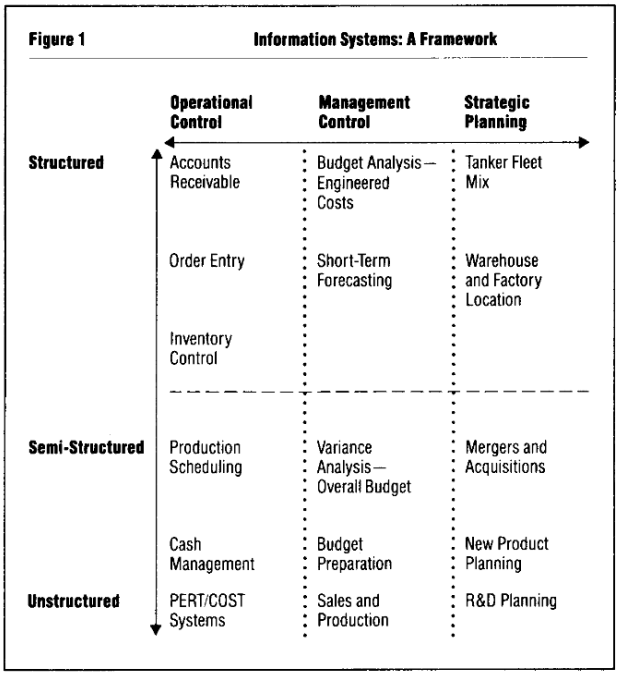
Table 1 Information Requirements by Decision Category

Characteristics of Information	Operational Control	Management Control	Strategic Planning
Source	Largely internal		External
Scope	Well defined, narrow		Very wide
Level of Aggregation	Detailed		Aggregate
Time Horizon	Historical		Future
Currency	Highly current		Quite old
Required Accuracy	High		Low
Frequency of Use	Very frequent		Infrequent

Figur 4: Gorry and Scott-Morton's Information Requirements by Decision Category - model

The first model discusses the where the different information and its scope falls within the presented three categories. Operational control needs faster, more frequent and specific

information while strategic planning needs aggregated information and forecasts more seldom.



Figur 5: Gorry and Scott-Morton's Information Systems Framework

The second model links the decision making categories with the type of data. By doing this they can place each organisational function within this model. They do stress that the borders between each data and decision category is not strict.

3.1.2 Steven Alter (2004)

Steven Alter claims that decision support systems are ancient history. He proposes to use the term "decision support" within a "work system" is more applicable in the real world as decision support is more of a help tool than a system within itself. He claims that this "reduces the chance being misled by techno-hype, vendor sales pitches, and incomplete understanding of determinants of success in organizations." (Alter, 2004) This confirms that Decision Support Systems should be a part of a larger whole.

3.1.3 Watson (2005)

Watson (2005) attempts to summarize in his article what has happened in decision support systems for the last five years. According to Watson new systems and praxis emerge constantly and they are always marketed as a new and improved approach when they just might be an evolution of an already established approach.

The author mentions Business Performance Management (BPM) as one good example of this. Digital dashboards and balanced scorecards are the two of the most used BPM methods to date but balanced scorecards go further by incorporating other qualitative metrics. The problem is, according to the author, that these approaches are not new.

In the 1970s they were called critical success factors and in the 1980s they were incorporated into Executive Information Systems. The author still feels like BPM should be considered new because of the new methodology, the focus on business strategy and the new use of technologies.

Other developments he considers new are:

- Real-Time Data Warehousing makes it possible to access company data anywhere at any time
- Business Activity Monitoring (BAM) is the ability of a company to monitor progress (sales) in real time as a purchase happens
- Customer Relationship Management (CRM) links business and IT strategy with customer-centric data, data cleansing, online analytical processing mining
- Guided business intelligence is a program that suggest to the analyst the best possible analyse method
- Text mining is when computers analyses texts and documents to find the appropriate patterns (Watson, 2005)

3.1.4 Summary of Decision Support System research

Researcher	Title	Data	Method	Result
Rania Shibl, Meredith Lawley, Justin Debuse	Factors influencing decision support system acceptance	Interviews with 37 general practitioners from a mix of education backgrounds, experience and gender.	Depth interviews	The developed model indicated that four main factors influence DSS acceptance and use including usefulness, facilitating conditions, ease of use and trust in the knowledge base.
Soussan Djamasbia, Eleanor T. Loiaconob,	Do men and women use feedback provided by their Decision Support Systems (DSS) differently?	Laboratory experiment	A laboratory experiment using a previously validated Decision Support System (DSS) and a commonly used and negatively framed feedback.	Outcome feedback improved the decision accuracy of the female users compared to their male counterparts. Outcome feedback affects the overall mood of men and women differently as well.

3.2 Time series analysis on sales data

Many of the modern approaches to forecasting are very complicated and consist of computer programmed neural networks that function like a simplified brain reacting and adapting to outside stimuli and changing the model accordingly (Philip Doganis, 2006) (R.J. Kuoa, 1999) (Xueb, 1998).

These methods are sadly too complicated to discuss in this article. To find a discussion of the simple models I will use in this thesis I have to go back around 50 years before computer programming. A short summary of the findings are given in the table below. For more specific details please visit the theory or method chapter.

Researcher	Title	Data	Method	Result
Charles C. Holt (1957)	Forecasting seasonals and trends by exponentially weighted moving averages	Method discussion	Forecasting expressions for exponential weighted moving averages. Methods for series with no trend, or additive or multiplicative trend are examined. Similarly, the methods cover non-seasonal, and seasonal series with additive or multiplicative error structures.	This exploratory analysis indicates the great flexibility of exponentially weighted moving averages in dealing with forecasts of seasonals and trends.
Peter R. Winters (1960)	Forecasting Sales by Exponentially Weighted Moving Averages	Monthly sales of a cooking utensil, Bi-monthly sales of paint, prefabricated houses manufactured and sold by Admiral Homes,	Exponentially Weighted Moving Averages	They conclude, then, that the exponential forecasting model has several advantages over more conventional forecasting models: (1) it gives better forecasts, (2) it requires less information and storage space, (3) it responds more rapidly to sudden shifts in the time series so that is routinely protects the forecaster
Kapoor, Madhok & Wu (1981)	Modeling and Forecasting Sales Data by Time Series Analysis	Two series of 71 observations (periods) taken at intervals of every four weeks covering a period from February 1971 to August 1976.	ARMA-model	Actual values are well within the 95% probability limits of the forecasts.
Fuller & Dickey (1979)	Distribution of the Estimators for Autoregressive Time Series with a Unit Root	Method Discussion	AR model with a unit root	A method for testing if there is a unit root in time series
Granger C.W.J (1969)	Investigating Causal Relations by Econometric Models and Cross-spectral Methods	Method discussion on causal relations	Granger causality	Granger causality method
Engle-Granger (1987)	Co-Integration and Error Correction: Representation, Estimation, and Testing	Method discussion	Estimation and testing of co-integration	Engle-Granger two step method

4. Data

Data has been provided by the grocer ICA Flygfyren in Norrtälje, Sweden. The grocer is an ICA Kvantum that has about 59 000 stock keeping units on a space of 5 500 square meters. It is close to the motorway leading further North to Gävle, West to Uppsala and South to Stockholm.

4.1 The products

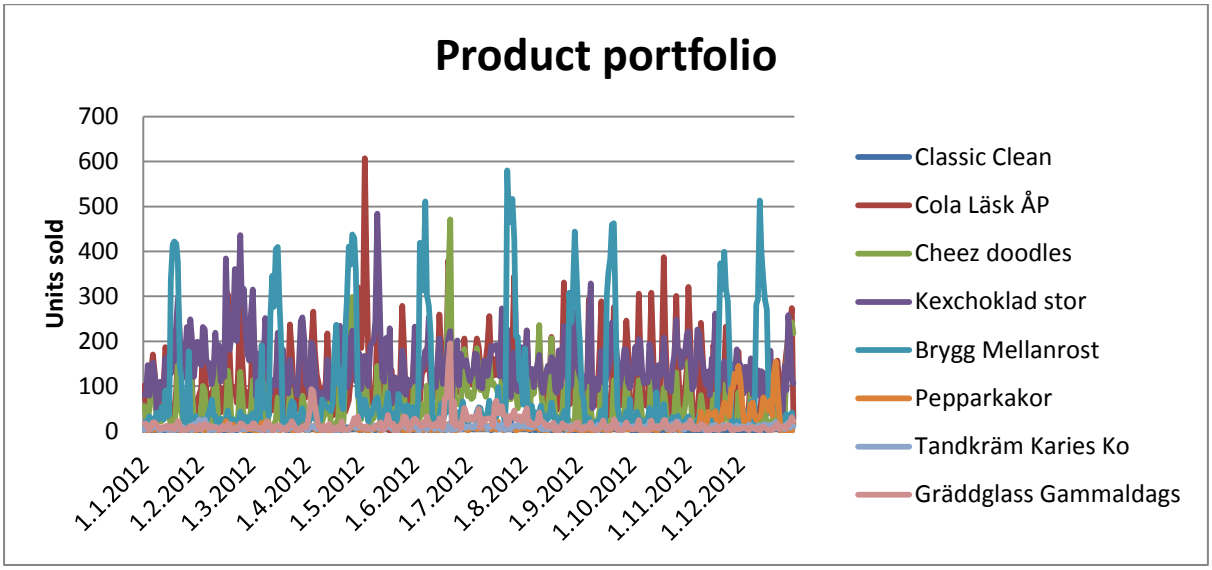
The products have been chosen to represent a wide variety of products with different sales patterns through time. Five products have been chosen in hope that they exhibit a stable sales pattern trough time, one product pair has been chosen for their high expected correlation and two products have been chosen because of high seasonal sales.

Product type	Actual product	Expected pattern	Actual visual pattern
Coffee	Brygg Mellanrost	Stable pattern	Almost monthly very large chocks
Chocolate	Kexchoklad stor	Stable pattern	Stable pattern with two chocks and one trend
Toothpaste	Tandkräm Karies KO	Stable pattern	Stable pattern with one chock
Coca-Cola	Coca-Cola 0,5 l	Stable pattern, highly correlated with Chips	Stable pattern with one chock, highly correlated with Chips
Chips	Cheese doodles	Stable pattern, highly correlated with Coca-Cola	Stable pattern with three chocks, highly correlated with Coca-Cola
Shampoo	Head & Schoulders Classic Clean	Stable pattern	High volatility, low volume, some seasonality
Ice-cream	Gammaldags Gräddglass	High seasonality	Stable pattern with two chocks and a trend
Ginger biscuit	Pepparkakor	High seasonality	Stable pattern with one trend

4.2 Descriptive statistics

Descriptive statistics is given I graphical and table form for the reader to review. To save space in the thesis and to make it more readable individual graphs have been but in as an attachment.

4.2.1 Descriptive statistics by charts



As is visible from the graph some products have visible trends and others have visible chocks or both. The trends are because of seasonality. Ice-cream tends to sell better in the summer and ginger biscuits around Christmas and Lucia. Chocks are probably due to better product placement in the stores or discounts to the products. Especially coffee seems to have regular, almost monthly promotions in the store which will be quite interesting to analyse.

4.2.2 Descriptive statistics by tables

Variable	N	Min	Max	Mean	Standard deviation	Variance
Classic Clean	366	0	24	4,43	3,35	11,25
Cola	366	0	607	122,40	77,50	6006,11
Cheese doodles	366	5	471	61,21	57,17	3268,11
Kexchoklad stor	366	12	484	148,77	60,56	3667,84
Brygg Mellanrost	366	9	580	90,11	123,12	15157,90
Pepparkakor	366	0	155	11,81	21,04	442,84
Tandkräm Karies KO	366	0	26	9,12	4,33	18,78
Gammaldags Gräddglass	366	0	196	15,06	17,85	318,63
Total product basket	366	69	1336	462,90	216,85	47024,37
Total store sales	366	563011	5075556	1583602	593447	3,522E11
Total customers	366	2498	10583	5590,46	1113,98	1240950,22
Total units sold	366	18978	220219	70731,46	24685,52	6,094E8

The number of days in the sample is 366 since 2012 was a leap year. Of the sample presented the coffee showed the highest standard deviation and shampoo the smallest. This can be attributed to monthly promotions to coffee and small volume of daily sales for the shampoo.

What is interesting to note is that the chosen basket represents on average 463 units of the total 70731 units sold per day or roughly 0,7 %.

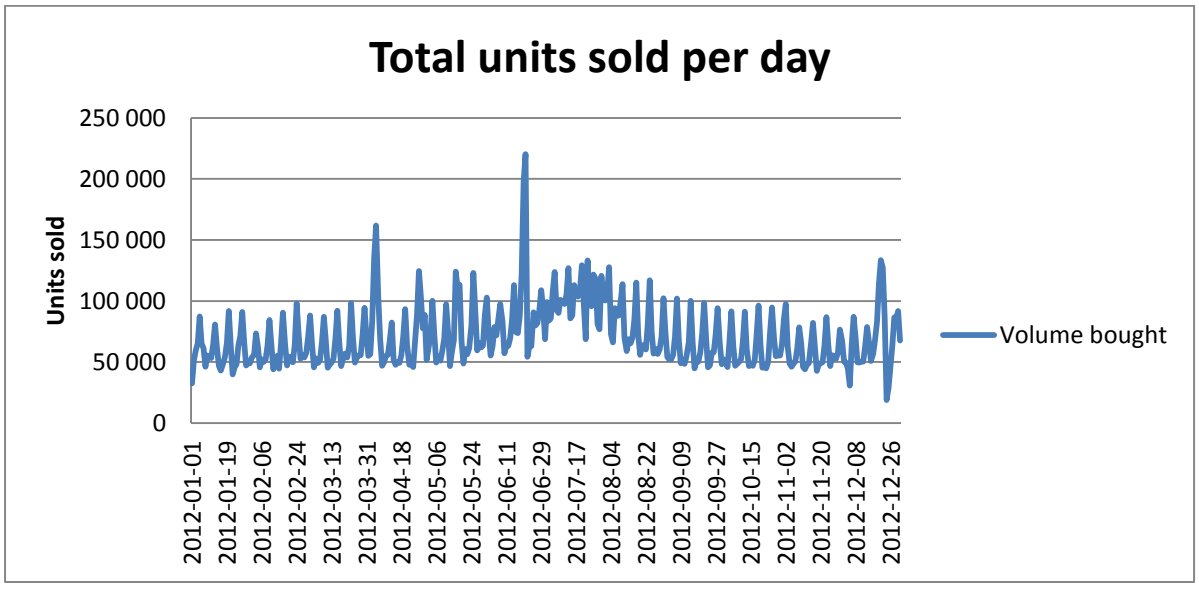
4.2.3 Data distribution

Variable	Skewness	Kurtosis	Kolmogorov-Smirnova	Shapiro-Wilk	Student-t
Classic Clean	1,796	6,082	0,137**	0,870**	25,242**
Cola	1,548	4,399	0,117**	0,889**	30,214**
Cheese doodles	2,301	9,262	0,166**	0,793**	20,484**
Kexchoklad stor	1,483	4,251	0,086**	0,909**	46,996**
Brygg Mellanrost	2,016	2,922	0,320**	0,629**	14,002**
Pepparkakor	4,119	19,013	0,341**	0,471**	10,733**
Tandkräm Karies KO	0,762	1,101	0,092**	0,963**	40,262**
Gammaldags Gräddglass	5,016	38,580	0,205**	0,589**	16,138**
Total product basket	1,1	1,181	0,094**	0,918**	40,838**
Total store sales	1,58	4,316	0,123**	0,881**	51,051**
Total customers	0,676	0,711	0,087**	0,964**	96,009**
Total units sold	1,644	5,067	0,129**	0,876**	54,816**

A normal distributed sample has a skewness of 0 and a kurtosis of 3. As can be seen by the table the ginger biscuits and the ice-cream exhibit very large diversions from this. But according to the normality tests the sample is large enough to assume normality.

4.3 Data source criticism

The grocer has been chosen because of its willingness to cooperate with the study. This leads the data to be highly specified with one store at one location in one city. I want to remind the reader that the purpose of the thesis is to test the applicability of statistical methods, not a study of the products sales patterns or sales methods. For that purpose the data is sufficient. Even so, there are some problems with the data source.



The grocer is in an area exhibiting seasonal changes as the population of Stockholm move out to their summer cottages during the summer holidays. This is expected to generate an upward fluctuation in the sales data as the population of the surrounding area increases. This is however only for a short time and should generate a predictable pattern in the long run.

The grocer has been undergoing renovation for the last six months which might influence the likability of the store. According to Mikael Jarleskog, who is working at the financial and statistical group of the grocer, the sales have gone down 5 % from September 2012 onwards as a reaction (Jarlaskog, 2013). The customers might be negatively influenced by the constant work going on inside and outside the store and frequent other stores instead. The store is however one of the only two stores in the city and the other store has a far lesser supply of goods for sale. This would favour the study because of a smaller impact on sales.

5. Method

By pursuing to test the applicability of the following statistical measures and methods on the sales data of certain products for a grocery store I hope to gain insights into the applicability of the measures and methods as decision support systems for the grocer. This chapter is dedicated to explaining how each measure and method is computed in the statistical programs and which assumptions, data and results can be expected from each measure and method.

5.1 Test of Dependence

Correlations are calculated using a two-sided Pearson product-moment correlation coefficient in SPSS. The method was developed in the 1880s by Karl Pearson. To do that we first need to calculate the covariance for the sample period (Koop, 2008):

$$Cov(x,y) = s_{xy} = \frac{\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)}{n-1}$$

A sample correlation coefficient is then calculated by:

$$r = \frac{Cov\,(x,y)}{s_x s_y}$$

5.3 Test of Data Stationarity

The Augumented Dickey-Fueller (ADF) Method test of unit root gives some indications if the data is stationary. Other viable tests are just a visual scanning of the data graph to see if it shows signs of stationarity or not.

The ADF method studies if the AR process follows a t-distribution or a Dickey-Fueller distribution. The augumented part simply means more lags on the data than one. The test also varies if the data shows signs for a deterministic trend. To estimate the ADF model you first need to estimate an AR (p) model with a deterministic trend:

$$\Delta Y_t = \alpha + \emptyset Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_{p-1}\Delta Y_{t-p+1} + \delta t + \varepsilon_t$$

Afterwards you can use hypothesis testing procedures or information criteria to select a lag length or decide if a deterministic trend should be included. The model must then exceed the values of -2,80 for a 5 % critical value if it has a deterministic trend and -3,41if it does not in a sample over 100 entries. A unit root exists if $\emptyset$ is in between -2 and 0. (Damodar Gujarati, 2009)

The lag length is chosen by the best information criterion. There are four information criterion test that test how much of the data is lost by the method. These are AIC, Schwartz, Shibata and Hannan-Quinn. I will chose the lags with the lowest information criterion.

5.3. Test of Cointegration

Testing for cointegration is done by the Engle-Granger two step method. You start by regressing variable X against variable Y and then testing the residuals ε for a unit root.

$$Y = \alpha + \beta X + \varepsilon$$

For simplicity and to save time I will test for the unit root with the Dickey-Fueller (DF) method with only one lag.

5.2 Multiple Regressions on the Data

The multiple regressions are made using total volume of products sold in the store per day as the dependent variable and the individual products as the independent variables.

This regression will give us insight into how each product in the portfolio affects the store sales. The multiple regression is in the form of:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \varepsilon_i$$

Where Y is the total units sold in the store per day and X_{1-8} is one of the eight products.

5.2 Test of Causality

The Granger test of causality gives us a suggestion of causality. It is done by lagging the variable X against past values of itself X_{t-1} and another variable Y and its past values Y_{t-1}. If the independent variable Y_{t-1} shows signs of significantly affecting X then there might be some causality between the variables. There might also be contemporaneous causality where X_{t-1} influences Y. For simplicity I will assume that all the series are stationary and the lag length is chosen by the best information criterion. The information criteria are Akaike, Hannan–Quinn, Bayesian- Schwarz and Akaike's Final Prediction Error. Allowing for multiple lags (Damodar Gujarati, 2009):

$$X_t = \alpha + \delta t + \rho_1 X_{t-1} + \cdots + \rho_p X_{t-p} + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + \varepsilon_t$$

$$Y_t = \alpha + \delta t + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + \rho_1 X_{t-1} + \cdots + \rho_p X_{t-p} + \varepsilon_t$$

5.5 Economic Forecasting

An autoregression is a regression of a variable of itself with the historical values at certain intervals. A simple AR (1) model with one lag can be written as (Koop, 2008):

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$$

By rolling the model forward from the last possible data point it is possible to get some suggestions about where the data is heading. I will compare the forecast to the actual outcome in 2013 to see if the mean differs significantly. For simplicity the process is as follows.



This leads to a forecast of 30 days which will be compared to the actual units' sales the next 30 days. The mean of the two series will be compared using a paired samples t-test to test if they differ significantly. The null hypothesis is that the difference in the mean values is zero. If the null hypothesis cannot be rejected then the forecast is successful i.e. the mean does not differ significantly.

5.9 Hypothesis testing

	H0	H1
Hypothesis 1	There is no correlation between products	There is correlation between products
Hypothesis 2	There is no unit root for a product	There is a unit root for a product
Hypothesis 3	Products do not share a similar drift	Products might share a similar drift
Hypothesis 4	Product sales do not influence store sales	Product sales do influence store sales
Hypothesis 5	There is no Granger causality for the products	There is Granger causality for the products
Hypothesis 6	It is not possible to forecast a products volume sale	It is possible to forecast a products volume sale

6.0 Results

The following chapter shows the results for the given tests. The chapter will present the statistical results and briefly discuss the model diagnostics. For a more general discussion on the implications for decision support the reader is referred to the next chapter discussion.

6.1 Test of Dependence

	Classic	Cola	Cheese	Kexchoklad	Bryg g	Pepparkakor	Tandkräm	Gräddglass	Basket	Sales (SEK)	Customers	Volume Sales
Classic Clean	1											
Cola	0,235**	1										
Cheese doodles	0.240**	0,678**	1									
Kexchoklad	0.166**	0,412**	0,337**	1								
Brygg Mellanrost	0,124*	0,036	0,092	0,036	1							
Pepparkakor	-0,116*	-0,044	-0,063	-0,086	0,047	1						
Tandkräm	0.136**	0,277**	0,278**	0,185**	0,044	0,166**	1					
Gräddglass	0,228**	0,440**	0,622**	0,141**	0,004	0,094	0,208**	1				
Basket	0,290**	0,713**	0,707**	0,545**	0,623**	0,061	0,304**	0,444**	1			
Sales (SEK)	0,382**	0,673**	0,768**	0,318**	0,123*	0,142*	0,387**	0,694**	0,686**	1		
Customers	0,364**	0,709**	0,738**	0,342**	0,116*	-0,02	0,342**	0,627**	0,673**	0,930**	1	
Store unit sales	0,389**	0,675**	0,773**	0,341**	0,123*	0,098	0,371**	0,711**	0,692**	0,992**	0,931**	1

Table 1: ** Correlation is significant at 0,01 level, * Correlation is significant at the 0,05 level,

As the reader can see from the table many of the series are highly correlated. All of them except pepparkakor seem to be dependent on the general level of stores sales, customers and unit sales. This means that the more there are people in the store, and the more kronor and units sold in a day the higher the unit sales of a product.

Another interesting aspect is that self-indulgent products such as kexchoklad, cola and ice-cream seem to be highly correlated. This would indicate that there are special days when these products rise more than others and rise together such as weekends. Coffee seems also to be more independent from other products though highly correlated with general store variables.

Store sales, total customers and store unit sales are also highly correlated as can be expected. For this test we can conclude that for hypothesis 1: H1 There is correlation between products is accepted.

6.2 Test of Unit Root

Variable	Lags	ADF-statistic	Critical value 1 %	Critical value 5%	AIC	Schwartz	Shibata	Hannan-Quinn
Classic Clean	5	-1.96487	-2.56572	-1.94093	5.099491	5.164259	5.098947	5.125244
Cola	7	-1.00324	-2.56572	-1.94093	10.790645	10.877361	10.789675	10.825132
Cheese doodles	7	-1.3871	-2.56572	-1.94093	10.243609	10.330325	10.242639	10.278096
Kexchoklad stor	6	-0.838984	-2.56572	-1.94093	10.848519	10.924238	10.847778	10.878630
Brygg Mellanrost	3	-5.98912	-2.56572	-1.94093	11.373253	11.416255	11.373013	11.390348
Pepparkakor	3	-3.51562	-2.56572	-1.94093	7.560812	7.603813	7.560571	7.577906
Tandkräm Karies KO	6	-1.07477	-2.56572	-1.94093	5.732127	5.807846	5.731385	5.762237
Gammaldags Gräddglass	6	-2.811	-2.56572	-1.94093	8.080014	8.155734	8.079273	8.110125
Total product basket	7	-1.26799	-2.56572	-1.94093	12.974478	13.061194	12.973508	13.008965
Total store sales	8	-0.508698	-2.56572	-1.94093	28.570618	28.668376	28.569388	28.609501
Total customers	7	-0.172719	-2.56572	-1.94093	15.885674	15.972390	15.884704	15.920161
Total units sold	8	-0.540386	-2.56572	-1.94093	22.242246	22.340004	22.241016	22.281128

In the unit root test we can see some differences between the series. The optimal lag length was chosen according to information criterions. I will use the 5 % critical value as the cut-off rate. To explain the results further I have constructed the following table.

Variable	Unit root	Stationary	Chocks affect process	Mean might change	Volatility might change
Classic Clean	No	Yes	No	No	No
Cola	Yes	No	Yes	Yes	Yes
Cheese doodles	Yes	No	Yes	Yes	Yes
Kexchoklad stor	Yes	No	Yes	Yes	Yes
Brygg Mellanrost	No	Yes	No	No	No
Pepparkakor	No	Yes	No	No	No
Tandkräm Karies KO	Yes	No	Yes	Yes	Yes
Gammaldags Gräddglass	No	Yes	No	No	No
Total product basket	Yes	No	Yes	Yes	Yes
Total store sales	Yes	No	Yes	Yes	Yes
Total customers	Yes	No	Yes	Yes	Yes
Total units sold	Yes	No	Yes	Yes	Yes

As you can see some of the series have a unit root and some do not. This gives an indication if the series are stationary. A series with a unit root have a changing mean and volatility when the time t changes. Also such a change, in the form of a chock, affects the series permanently. Products with a unit root would respond well to promotions, optimal product placement, discounts and have a positive or negative effect on the series from then onwards. Of course this effect might be offset by other chocks. Cola, Cheese doodles, Kexchoklad, Tankräm Karies are such products. Another interesting conclusion we can make is that the general store variables also exhibit a unit root. That means chocks and changes in the store give rise to a “new normal” in the trends.

For the unit root test we can conclude that for hypothesis 2: H1There is a unit root for a product is accepted.

6.3. Test of Cointegration

	Classic	Cola	Cheese	Kexchokla d	Brygg	Pepparkakor	Tandkräm	Gräddglass	Basket	Sales (SEK)	Customers	Volume Sales
Classic Clean	X											
Cola	-10,06	X										
Cheese doodles	-9,81	-9,98	X									
Kexchoklad	-9,80	-11,63	-7,91	X								
Brygg Mellanrost	-9,78	-14,40	-9,65	-10,20	X							
Pepparkakor	-10,04	-14,63	-9,93	-10,48	-6,24	X						
Tandkräm	-9,68	-13,57	-9,10	-9,65	-6,23	-5,62	X					
Gräddglass	-10,05	-13,91	-9,41	-9,93	-6,25	-5,92	-9,99	X				
Basket	-10,08	-9,21	-5,56	-7,63	-8,34	-5,67	-9,46	-6,68	X			
Sales (SEK)	-10,63	-10,80	-6,11	-8,40	-6,11	-5,35	-9,92	-8,31	-7,29	X		
Customers	-10,67	-10,32	-5,87	-8,25	-6,14	-5,81	-9,70	-8,16	-7,14	-7,74	X	
Volume sales	-10,75	-11,16	-6,48	-8,37	-6,09	-5,44	-9,78	-8,88	-7,40	-6,19	-8,44	X

As is visible from the table no ADF statistic is over the 5 % critical value of -1,94093. This means that no cointegration was found between the series which in turn means that they do not share a common stochastic drift. Instead the differences between the series seem to revert back to its mean. This is what we would expect from a mature company that does not see a long term trending decline or growth in the general level of sales. On the other hand some of the series where not stationary so you would expect the residuals to have a unit root. The results where the same for lags of 5 and 10.

For the cointegration test we can conclude that for hypothesis 3: H0 Products do not share a similar drift is accepted.

6.2 Multiple regression

Variables	(Total store sales in krona)	(Customers)	(Total store unit sales)
Products			
Classic Clean	0,171(6,684)**	0,143(4,668)**	0,170(6,669)**
Cola	0,221(6,446)**	0,334(8,130)**	0,212(6,202)**
Cheese doodles	0,332(8,667)**	0,282(6,159)**	0,321(8,415)**
Kexchoklad stor	0,037(1,380)	0,038(1,167)	0,063(2,354)*
Brygg Mellanrost	0,046(1,883)	0,052(1,759)	0,050(2,030)*
Pepparkakor	0,21(8,424)**	0,058(1,953)	0,172(6,924)**
Tandkräm Karies KO	0,095(3,647)**	0,080(2,571)**	0,080(3,104)**
Gammaldags Gräddglass	0,346(11,040)**	0,255(6,807)**	0,370(11,828)**
F value	170,811**	105,719**	172,27**
Adjusted R^2	0,793	0,703	0,79
N	365	365	365
Durbin Watson	0,958	0,89	1,038

As is visible from the table the product basket seems to have an effect on the store sales, customers and store unit sales. Again the indulgent products seem to have a higher effect on store variables than others as they are bought in higher volumes.

Because the product basket constitutes on average 463 of the total average 70 731 products sold per day or 0,7 % it is more likely that the products share the same trend as the total store.

A word of warning about the multiple regressions. The product series are highly correlated with other products which might result in multicollinearity. It is also impossible to know if the product unit sales caused the high store unit sales or vice versa. To test this if the general store trend affects the product we have to test for granger causality.

For the multiple regressions we can conclude that for hypothesis 4: H1 Product sales do influence store sales is accepted.

6.2 Test of Causality

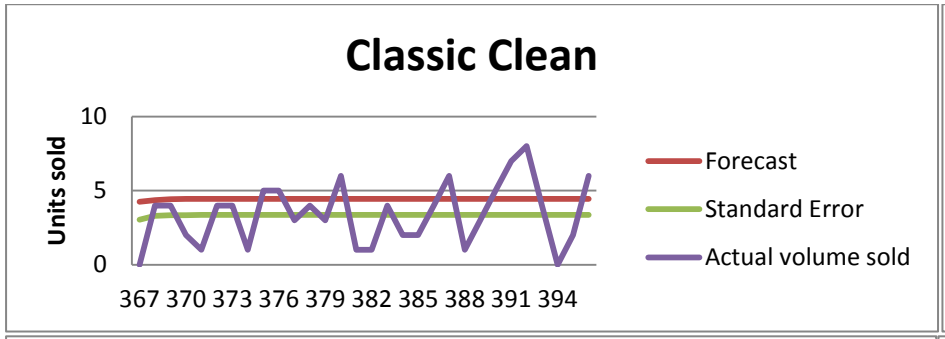
Variables	Classic Clean	Cola	Cheese doodles	Kexchoklad stor	Brygg Mellanrost	Pepparkakor	Tandkräm Karies KO	Gräddglass	Total product basket	Sales per day	Customers per day
Constant	-0,429942(-0,639)	-6.31629(-0,670)	-15.2063(-2,23)*	62.0888(4,87)**	26.4129(1,70)	4.99140 (2,05)	2.53776(2,76)**	-20.0747(-6,27)**	54.0639(2,16)*	-14999.9(-1,37)	493.287(3,17)**
Lags											
X_1	0.335214(6,41)**	0.645312(12,4)**	0.634093 (11,7)**	0.386583(7,25)**	0.830914(16,1)**	0.937527(17,8)**	0.220022(4,18)**	0.338119(6,54)**	0.736681(13,9)**	0.583142(11,1)**	0.453139(7,84)**
X_2	0.000689572(0,0124)	-0.150132(-2,78)**	0.167501(2,65)**	0.102735(1,82)	0.129982(1,92)	-0.0454386(-0,635)	0.102037(1,90)	0.225029(4,10)**	0.0368785(0,559)	0.349500(5,73)**	-0.0246279(-0,384)
X_3	-0.0569199(-1,03)		0.0294822(0,457)	0.109893(2,11)*	-0.00202234(-0,0299)	-0.240508(-3,35)**	0.174765(3,32)**	0.0637568(1,17)	0.0596394(0,915)	-0.208490(-3,37)**	0.202884(3,23)**
X_4	0.159108(3,04)**		-0.180730(-2,87)**		-0.233851(-4,55)**	0.227969(4,31)**		-0.239948(-4,83)**	-0.169123(-3,24)**	-0.0881566(-1,46)	0.165809(2,62)**
X_5			0.0880397(1,60)							0.253626(5,18)**	-0.0309114(-0,540)
Total store volume sales	4.81608e-005(5,91)**	0.00220178(16,2)**	0.00186653(25,0)**	0.00121126(9.,02)**	0.000650325(3,50)**	0.000123433(4,22)**	7.93580e-005(7,32)**	0.000415712(14,2)**	0.00693801(23,8)**	23.3451(194)**	0.0406855(41,8)**
Total store volume sales_1	-2.67102e-005(-2,51)*	-0.000943650(-4,40)**	-0.000860657(-5,87)**	-0.000467562(-2,49)*	-0.000842528(-3,49)**	-0.000137061(-3,67)**	-3.00033e-005(-2,03)*	-4.11011e-005(-0,958)	-0.00447348(-8,23)**	-12.2811(-9,75)**	-0.0143926(-4,92)**
Total store volume sales_2	2.72954e-005(2,45)*	-0.000297310(-1,63)	-0.000311849(-1,99)*	-0.000429277(-2,24)*	-0.000342153 (-1,34)	-5.50631e-006(-0,140)	7.83816e-006(0,527)	-1.35067e-00(-0,301)	-0.00145907(-2,42)**	-8.99065(-6,30)**	-0.00179584(-0,569)
Total store volume sales_3	-1.68776e-005(-1,58)		-0.000425507(-2,75)**	-0.000345819(-2,24)*	0.000577331(2,34)*	1.58455e-005(0,417)	-2.78978e-005(-2,40)*	-5.65187e-005(-1,31)	-0.000621280(-1,02)	4.92323(3,37)**	-0.00539727(-1,74)
Total store volume sales_4	9.93122e-006(1,17)		0.000521469(3,38)**		-6.15735e-005(-0,319)	-4.64871e-005(-1,55)		0.000109583(2,93)**	0.00105570(2,22)*	2.01958(1,40)	-0.00701946(-2,26)*
Total store volume sales_5			-0.000346188(-2,75)**							-6.31740(-5,43)**	-0.000576544(-0,216)
Model diagnostics											
F value	15,79**	132,8**	140,1**	36,44**	103,6**	135,5**	17,07**	83,94**	144,6**	7103**	393,2**
Adjusted R^2	0,26992	0,646794	0,80948	0,408	0,719438	0,770743	0,238081	0,674643	0,782142	0,995413	0,922989
N	361	361	361	361	361	361	361	361	361	361	361
AR 1-2 test:	0,2225	0,92723	3,6325*	0,68153	0,77014	3,198*	0,8671	0,21075	0,55678	14,355**	3,2292*
ARCH 1-1 test:	10,129**	2,1997	62,714**	2,8283	0,035215	26,764**	34,539**	191,56**	0,55953	0,25929	33,636**
Normality test:	60,719**	125,67**	59,813**	60,000**	142,69**	175,89**	11,275**	341**	15,421**	19,343**	62,714**
Hetero test:	2,8839**	4,7114**	4,7545**	2,3910**	2,5835**	8,9962**	5,7067**	47,572**	4,9725**	4,5212**	18,097**
Hetero-X test:	3,7679**	8,2859**	5,5397**	1,8592**	3,3678**	23,499**	3,7218**	21,751**	4,9251**	4,1984**	9,0128**
RESET23 test:	3,6543*	22,040**	9,0066**	5,4455**	10,261**	18,110**	0,26563	54,347**	9,5628**	3,1926*	36,233**
Information criteria											
AIC	2,1227	7,68673	6,47651	7,70794	8,3922	4,65851	2,68768	4,68054	9,26627	21,2388	11,4969
HQ	2,16553	7,71243	6,5279	7,7422	8,43503	4,70134	2,72194	4,72337	9,30909	21,2902	11,5483
SC	2,23043	7,75137	6,60578	7,79412	8,49992	4,76624	2,77386	4,78826	9,37399	21,3681	11,6261
FPE	8,35379	2179,25	649,715	2225,97	4412,57	105,48	14,6976	107,83	10575,3	1.67464e+009	98410,3
									Note high correlation!	Note high correlation!	

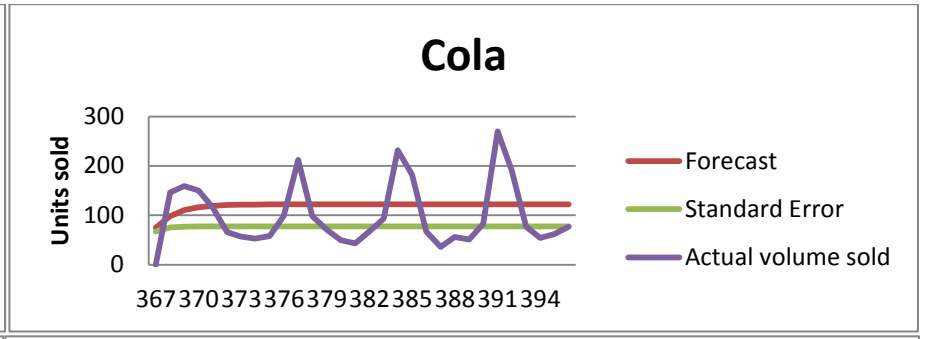
In the causality test we can witness that the product unit sales, except ice-cream, are all dependent on current and past levels of store total unit sales. That would give some indication that a general upswing in the store sales has a positive effect on product sales. For some products like ice-cream and kexchoklad this effect lasts longer than others. All products are dependent on the current store sales. The effect is, however, very small but that can be expected since the product is only one of 59 000 products.

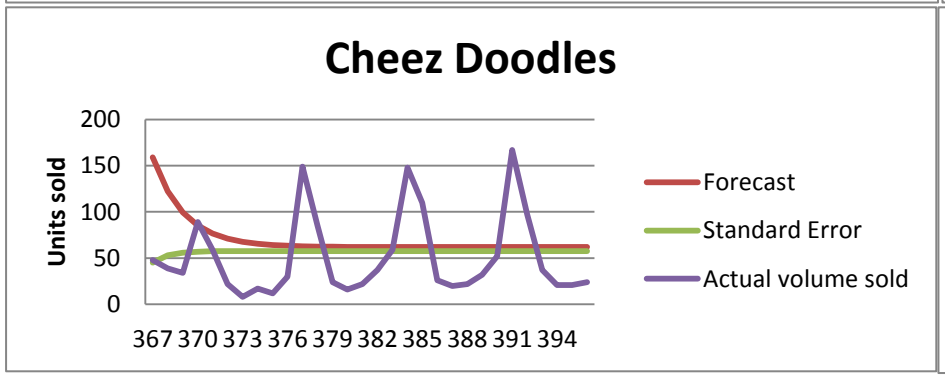
A word of warning for the reader. While a test of this kind might give an indication of causality it is not a definite proof of causality. This problem exists for all causality tests. Also note that highly correlated variables produce a lower information criterion for every lag added. They are therefore somewhat nonsensical since adding more of a highly correlated variable would produce a better estimate.

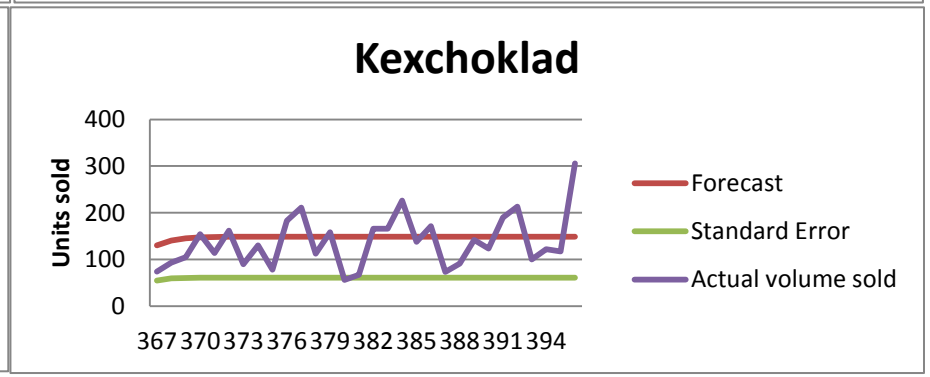
For the Granger causality test we can conclude that for hypothesis 5: H1 There is Granger causality for the products is accepted.

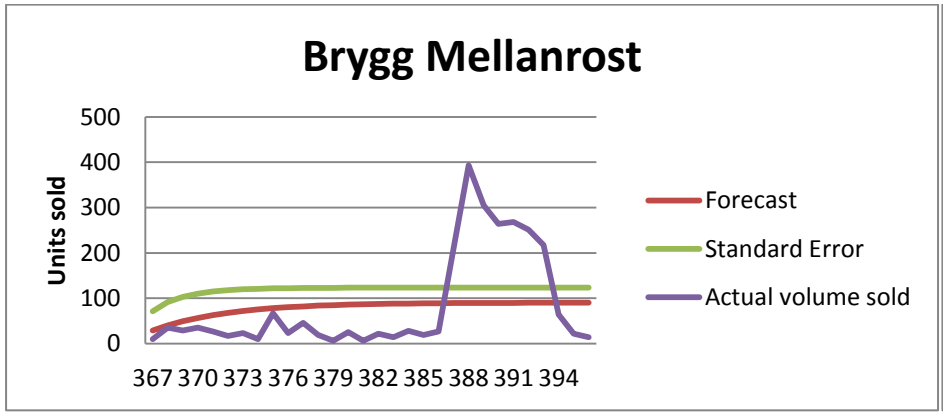
6.6 Economic Forecasting

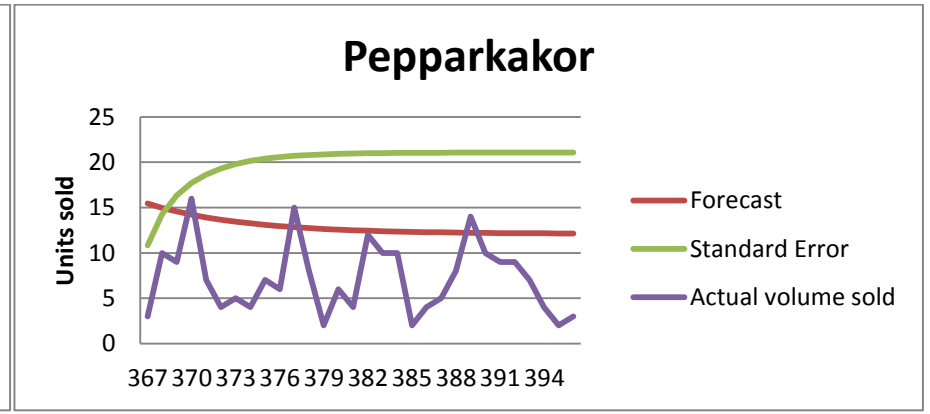


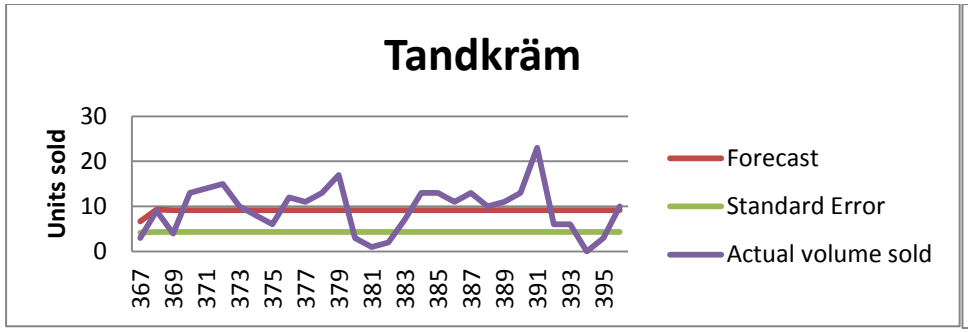


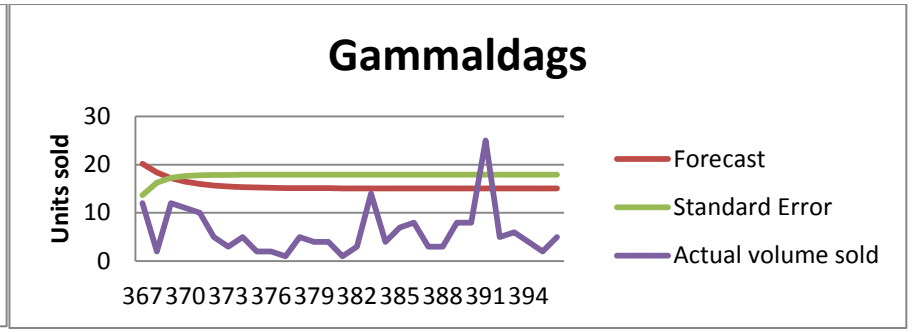


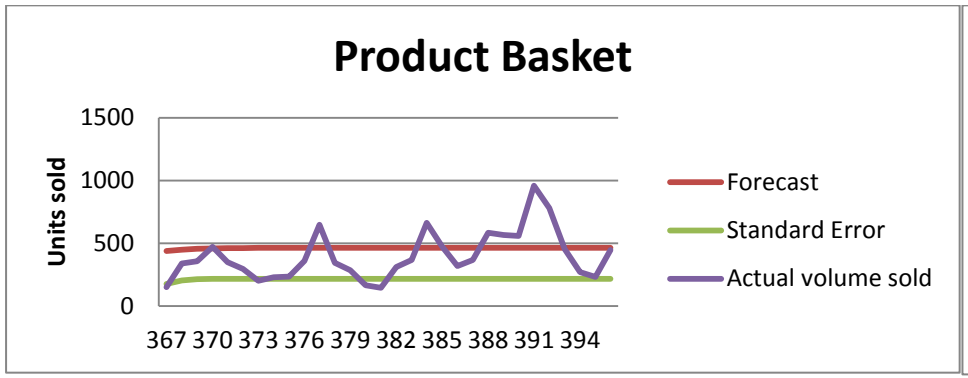


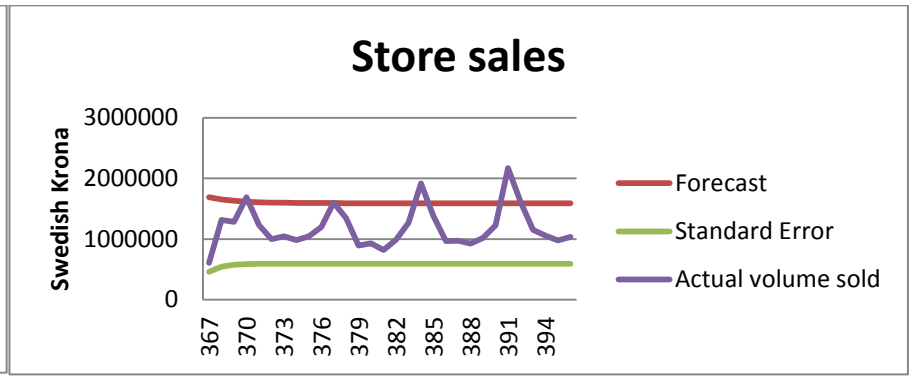




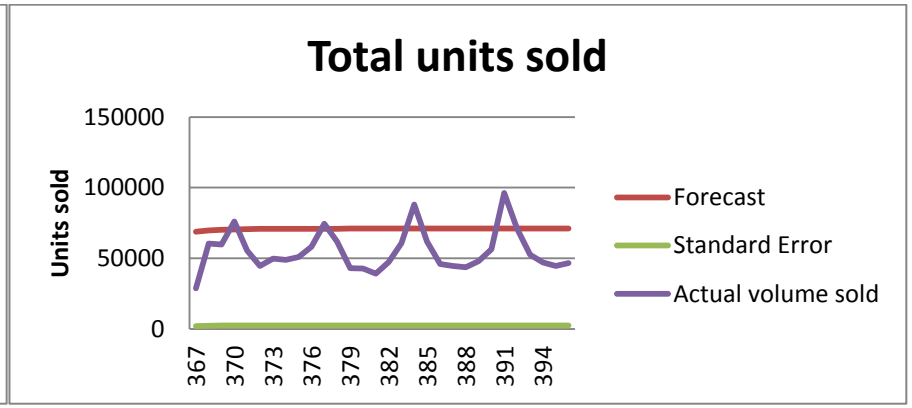

Tandkräm
Units sold
30
20
10
0
367 369 371 373 375 377 379 381 383 385 387 389 391 393 395
Forecast
Standard Error
Actual volume sold


Gammaldags
Units sold
30
20
10
0
367 370 373 376 379 382 385 388 391 394
Forecast
Standard Error
Actual volume sold


Product Basket
Units sold
1500
1000
500
0
367 370 373 376 379 382 385 388 391 394
Forecast
Standard Error
Actual volume sold


Store sales
Swedish Krona
3000000
2000000
1000000
0
367 370 373 376 379 382 385 388 391 394
Forecast
Standard Error
Actual volume sold


Customers
Number of customers
8000
6000
4000
2000
0
367 370 373 376 379 382 385 388 391 394
Forecast
Standard Error
Actual volume sold


Total units sold
Units sold
150000
100000
50000
0
367 370 373 376 379 382 385 388 391 394
Forecast
Standard Error
Actual volume sold

The graphs show the forecast, the standard error and the actual unit sold for the period 1.1.2013 – 30.1.2013. The x- axis is the 30 days beginning from the last entry in 31.12.2012 (day 366) to the last forecast day 30.1.2013 (day 396). The graphs show how much the actual results perceive intra week fluctuations with peaks during the weekends. The simple AR (1) forecast cannot replicate this pattern but tries to cut through the peaks on average. Observe how the model fails to predict the promotional effects on coffee. Model diagnostics and actual figures are given in the tables below.

Variable	Classic Clean			Cola			Cheez Doodles			Kexchoclad			Brygg			Pepparkakor		
Day	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold
367	4,25817	3,041	0	74,7581	67,4	0	159,083	45,16	48	130,131	54,45	74	28,6445	70,81	10	15,4471	10,82	3
368	4,36701	3,301	4	98,5698	75,32	146	122,412	53,19	39	140,726	59,47	93	39,8181	91,53	35	14,9726	14,26	10
369	4,41290	3,345	4	110,442	77,16	159	99,5858	55,99	34	145,377	60,39	105	48,9684	103,1	29	14,5655	16,33	9
370	4,43224	3,352	2	116,362	77,61	150	85,3783	57,04	89	147,419	60,56	154	56,4616	110,2	35	14,2161	17,7	16
371	4,44040	3,354	1	119,313	77,72	116	76,535	57,44	59	148,316	60,6	114	62,598	114,7	27	13,9163	18,65	7
372	4,44384	3,354	4	120,784	77,75	66	71,0306	57,6	22	148,709	60,6	162	67,6231	117,6	17	13,6591	19,32	4
373	4,44529	3,354	4	121,518	77,75	57	67,6045	57,66	8	148,882	60,61	90	71,7383	119,6	23	13,4383	19,8	5
374	4,44590	3,354	1	121,884	77,76	53	65,4719	57,68	17	148,958	60,61	130	75,1083	120,8	10	13,2489	20,14	4
375	4,44616	3,354	5	122,066	77,76	58	64,1446	57,69	12	148,991	60,61	78	77,868	121,7	67	13,0864	20,39	7
376	4,44626	3,354	5	122,157	77,76	99	63,3184	57,69	30	149,012	60,61	183	80,1279	122,2	23	12,9469	20,57	6
377	4,44631	3,354	3	122,203	77,76	212	62,8041	57,7	149	149,015	60,61	211	81,9787	122,6	46	12,8272	20,7	15
378	4,44633	3,354	4	122,225	77,76	98	62,484	57,7	85	149,016	60,61	112	83,4942	122,9	19	12,7245	20,8	8
379	4,44634	3,354	3	122,237	77,76	71	62,2848	57,7	24	149,017	60,61	158	84,7354	123	6	12,6364	20,87	2
380	4,44634	3,354	6	122,242	77,76	50	62,1608	57,7	16	149,017	60,61	56	85,7518	123,2	25	12,5607	20,92	6
381	4,44634	3,354	1	122,245	77,76	43	62,0836	57,7	22	149,017	60,61	67	86,35841	123,2	6	12,4958	20,96	4
382	4,44634	3,354	1	122,246	77,76	67	62,0355	57,7	37	149,017	60,61	166	87,2657	123,3	22	12,4402	20,99	12
383	4,44634	3,354	4	122,247	77,76	93	62,0056	57,7	59	149,017	60,61	166	87,8239	123,3	14	12,3924	21,01	10
384	4,44634	3,354	2	122,247	77,76	232	61,987	57,7	148	149,017	60,61	226	88,281	123,3	28	12,3514	21,03	10
385	4,44634	3,354	2	122,248	77,76	182	61,9754	57,7	110	149,017	60,61	138	88,6553	123,4	19	12,3162	21,04	2
386	4,44634	3,354	4	122,248	77,76	67	61,9682	57,7	26	149,017	60,61	171	88,9618	123,4	27	12,286	21,05	4
387	4,44634	3,354	6	122,248	77,76	36	61,9637	57,7	20	149,017	60,61	73	89,2192	123,4	210	12,2601	21,05	5
388	4,44634	3,354	1	122,248	77,76	56	61,9609	57,7	22	149,017	60,61	91	89,4184	123,4	393	12,2379	21,06	8
389	4,44634	3,354	3	122,248	77,76	51	61,9592	57,7	32	149,017	60,61	142	89,5868	123,4	305	12,2188	21,06	14
390	4,44634	3,354	5	122,248	77,76	83	61,9581	57,7	52	149,017	60,61	123	89,7246	123,4	264	12,2024	21,06	10
391	4,44634	3,354	7	122,248	77,76	270	61,9575	57,7	167	149,017	60,61	190	89,8375	123,4	268	12,1884	21,07	9
392	4,44634	3,354	8	122,248	77,76	191	61,957	57,7	97	149,017	60,61	213	89,93	123,4	251	12,1763	21,07	9
393	4,44634	3,354	4	122,248	77,76	77	61,9568	57,7	37	149,017	60,61	100	90,0057	123,4	217	12,166	21,07	7
394	4,44634	3,354	0	122,248	77,76	54	61,9566	57,7	21	149,017	60,61	122	90,0677	123,4	64	12,1571	21,07	4
395	4,44634	3,354	2	122,248	77,76	62	61,9565	57,7	21	149,017	60,61	117	90,1185	123,4	22	12,1495	21,07	2
396	4,44634	3,354	6	122,248	77,76	77	61,9564	57,7	24	149,017	60,61	306	90,16	123,4	14	12,143	21,07	3
Mean	4,44	3,34	3,40	119,09	77,31	99,20	70,53	57,04	50,90	147,90	60,36	137,70	79,01	118,59	83,20	12,88	19,93	7,17
Median	4,45	3,35	4,00	122,25	77,76	74,00	62,06	57,70	33,00	149,02	60,61	126,50	86,81	123,25	27,00	12,47	20,98	7,00
Model diagnostics																		
F value	78,92**			119,5**			221,9**			86,91**			739,2**			1017**		
R^2	0,178586			0,247605			0,379428			0,193168			0,670658			0,736891		
N	365			365			365			365			365			365		
AR 1-2 test:	1,3069			33,281**			5,9973**			0,65166			6,2645**			3,3437*		
ARCH 1-1 test:	4,3372*			0,48826			171,43**			0,87808			0,0040706			24,617**		
Normality test:	79,675**			118,39**			59,161**			57,664**			181,07**			176,55**		
Hetero test:	6,3846**			7,8798**			116,32**			6,6326**			3,9872*			38,635**		
Hetero-X test:	6,3846**			7,8798**			116,32**			6,6326**			3,9872*			38,635**		
RESET 23 test:	5,0540**			7,0832**			7,3918**			0,2207			7,1631			2,6832		

Variable	Tandkräm			Gammaldags			Basket			Store sales			Customers			Total units sold		
Day	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold	Forecast	Standard Error	Actual volume sold
367	6,66997	4,156	3	20,188	13,68	12	437,987	175,1	150	1690360	461800	607 650,10	5534,17	891	3 036	68892,1	1947	28 656
368	9,29389	4,319	9	18,3746	16,28	2	448,624	203,4	338	1653180	544700	1 312 196,53	5563,33	1036	5 142	69705,5	2283	60 340
369	9,18755	4,332	4	17,2052	17,25	12	454,915	212,4	356	1629920	573900	1 282 411,93	5580,62	1082	4 838	70203,3	2397	59 713
370	9,15748	4,333	13	16,4512	17,63	11	458,635	215,5	470	1615360	584900	1 688 439,88	5590,87	1098	5 933	70508	2438	76 160
371	9,14898	4,333	14	15,9649	17,79	10	460,835	216,5	348	1606260	589200	1 229 584,09	5596,95	1104	4 872	70694,5	2453	55 475
372	9,14657	4,333	15	15,6513	17,86	5	462,136	216,9	295	1600560	590900	996 244,58	5600,56	1105	3 955	70808,6	2459	44 541
373	9,1459	4,333	10	15,4491	17,88	3	462,905	217	200	1597000	591500	1 044 735,04	5602,69	1106	4 386	70878,4	2461	49 763
374	9,1457	4,333	8	15,3187	17,89	5	463,36	217,1	228	1594770	591800	982 460,17	5603,96	1106	4 357	70921,1	2462	48 800
375	9,14563	4,333	6	15,2346	17,9	2	463,629	217,1	235	1593380	591900	1 045 731,22	5604,71	1106	4 321	70947,3	2462	50 999
376	9,14563	4,333	12	15,1804	17,9	2	463,788	217,1	360	1592500	591900	1 199 672,05	5605,16	1107	4 477	70963,3	2462	58 256
377	9,14563	4,333	11	15,1454	17,9	1	463,882	217,1	648	1591960	591900	1 598 741,71	5605,42	1107	5 608	70973,1	2462	74 510
378	9,14563	4,333	13	15,1228	17,9	5	463,938	217,1	344	1591620	591900	1 346 356,85	5605,58	1107	5 066	70979,1	2462	61 707
379	9,14563	4,333	17	15,1083	17,9	4	463,971	217,1	285	1591400	591900	891 410,65	5605,67	1107	3 775	70982,8	2462	42 924
380	9,14563	4,333	3	15,0989	17,9	4	463,99	217,1	166	1591270	591900	931 658,98	5605,73	1107	4 145	70985	2462	42 716
381	9,14563	4,333	1	15,0929	17,9	1	464,002	217,1	145	1591190	591900	819 198,33	5605,76	1107	3 642	70986,4	2462	39 206
382	9,14563	4,333	2	15,089	17,9	3	464,008	217,1	310	1591130	591900	987 434,47	5605,78	1107	4 104	70987,2	2462	47 342
383	9,14563	4,333	7	15,0865	17,9	14	464,012	217,1	367	1591100	591900	1 269 725,28	5605,79	1107	4 630	70987,7	2462	60 689
384	9,14563	4,333	13	15,0848	17,9	4	464,015	217,1	663	1591080	591900	1 918 014,68	5605,8	1107	5 972	70988,1	2462	88 013
385	9,14563	4,333	13	15,0838	17,9	7	464,016	217,1	473	1591070	591900	1 379 514,80	5605,8	1107	4 931	70988,2	2462	61 710
386	9,14563	4,333	11	15,0831	17,9	8	464,017	217,1	318	1591060	591900	963 968,48	5605,81	1107	3 939	70988,4	2462	46 039
387	9,14563	4,333	13	15,0827	17,9	3	464,017	217,1	366	1591050	591900	971 904,31	5605,81	1107	4 147	70988,4	2462	44 559
388	9,14563	4,333	10	15,0824	17,9	3	464,018	217,1	584	1591050	591900	923 866,36	5605,81	1107	4 035	70988,5	2462	43 735
389	9,14563	4,333	11	15,0822	17,9	8	464,018	217,1	566	1591050	591900	1 019 633,49	5605,81	1107	4 162	70988,5	2462	48 090
390	9,14563	4,333	13	15,0821	17,9	8	464,018	217,1	558	1591050	591900	1 225 548,65	5605,81	1107	4 347	70988,5	2462	56 320
391	9,14563	4,333	23	15,082	17,9	25	464,018	217,1	959	1591050	591900	2 168 639,23	5605,81	1107	6 155	70988,5	2462	96 344
392	9,14563	4,333	6	15,082	17,9	5	464,018	217,1	780	1591050	591900	1 616 266,68	5605,81	1107	5 352	70988,5	2462	70 993
393	9,14563	4,333	6	15,082	17,9	6	464,018	217,1	454	1591050	591900	1 150 369,29	5605,81	1107	4 184	70988,5	2462	52 522
394	9,14563	4,333	0	15,0819	17,9	4	464,018	217,1	269	1591050	591900	1 053 704,40	5605,81	1107	4 232	70988,5	2462	46 983
395	9,14563	4,333	3	15,0819	17,9	2	464,018	217,1	231	1591050	591900	977 418,33	5605,81	1107	4 196	70988,6	2462	44 699
396	9,14563	4,333	10	15,0819	17,9	5	464,018	217,1	445	1591050	591900	1 035 170,95	5605,81	1107	4 190	70988,6	2462	46 613
Mean	9,07	4,33	9,33	15,56	17,67	6,13	461,89	215,00	397,03	1599889,00	585016,67	1187922,38	5599,94	1096,03	4537,63	70808,44	2435,47	54947,23
Median	9,15	4,33	10,00	15,09	17,90	5,00	464,01	217,10	352,00	1591160,00	591900,00	1049717,81	5605,77	1107,00	4334,00	70986,80	2462,00	50381,00
Model diagnostics																		
F value	31,72**			258,2**			195,8**			235,9**			200,6**			219,7**		
R^2	0,080353			0,415632			0,350429			0,393866			0,35598			0,377025		
N	365			365			365			365			365			365		
AR 1-2 test:	1,8048			2,2859			15,357**			17,232**			10,023**			17,914**		
ARCH 1-1 test:	20,334**			273,23**			3,1429			52,519**			1,7822			54,479**		
Normality test:	23,871**			971,64**			31,414**			82,880**			17,403**			107,04**		
Hetero test:	6,2306**			419,95**			14,451**			116,46**			25,098**			130,37**		
Hetero-X test:	6,2306**			419,95**			14,451**			116,46**			25,098**			130,37**		
RESET23 test:	0,73253			419,95**			0,73743			6,4377**			5,0092**			5,9734**		

As is visible from the graphs and the tables the forecast tends to revert to the mean. The forecast also seems to be better for series that are more stable than series that are not.

The reader can also observe that almost all series exhibit autocorrelation, heteroskedasticity and problems with normality. This is because of the simplicity of the AR (1) model used. Other, more advanced models take these problems into account and adjust for them by conditioning heteroskedasticity and moving the model average. These models are above the purpose of this study. But to test if the forecast is similar to the actual result I will do a statistical mean comparison.

Variable	Paired samples t-test	Forecast Possible
Classic Clean	2,742**	No
Cola	1,703	Yes
Cheese doodles	2,168*	No
Kexchoklad stor	1,016	Yes
Brygg Mellanrost	-0,216	Yes
Pepparkakor	8,052**	No
Tandkräm Karies KO	-0,280	Yes
Gammaldags Gräddglass	10,597**	No
Total product basket	1,864	Yes
Total store sales	6,639**	No
Total customers	8,047**	No
Total units sold	5,989**	No

From the above table we can see that using a basic AR(1) model and rolling it forward 30 days we can significantly forecast the sale of Cola, Chocolate, Coffee and Toothpaste. For the other products more complicated might give better results but that is above the scope of this thesis. For hypothesis 6: H1 It is possible to forecast a products volume sale is accepted.

6.7 Hypothesis results

	H0	H1	Result
Hypothesis 1	There is no correlation between products	There is correlation between products	H0 is rejected
Hypothesis 2	There is no unit root for a product	There is a unit root for a product	H0 is rejected
Hypothesis 3	Products do not share a similar drift	Products might share a similar drift	H1 is rejected
Hypothesis 4	Product sales do not influence store sales	Product sales do influence store sales	H0 is rejected
Hypothesis 5	There is no Granger causality for the products	There is Granger causality for the products	H0 is rejected
Hypothesis 6	It is not possible to forecast a products volume sale	It is possible to forecast a products volume sale	H0 is rejected

7.0 Discussion

7.1 General discussion

As was mentioned in the introduction retail analytics are very complex. Each product within the store shows both dependent development on the store general sales but also independent development on its own. Some products can be influenced by chocks like promotions while others do not.

To manage 49 000 products, each with both its dependent and independent part, requires a lot of time and effort. Even if it would be profitable to promote a certain product it might be at the expense of shelf-space, visibility or attractiveness of another. But I have showed that it is possible to gain new information from analytics for a few products. Using this approach we now know that:

- Store and product variables are highly dependent on each other
- Some products and store variables respond to chocks like promotions
- They do not however share a similar trend
- Products do influence store sales but they share roughly the same trend
- Products sales are dependent on past sales of itself and the total store which would indicate short term trends
- It is possible to forecast future sales

Therefore the approach adds value to a sales agent. The sales agent can use this information to:

- Measure how product A reacts to the change in sales of product B
- Choose which products to promote
- Investigate trends or lack thereof between products
- Measure the level of influence a product has on the store and vice versa
- Measure how long promotional effects on the store or products stay
- Forecast future sales and replenishment needs

7.2 Consequences for decision support

Using Business Analytics such as those presented in this study fall well into the realm of Decision Support Systems and not into the category of Automated Decision Making Systems nor in the category of Management Accounting Systems (Turban, 2011). Using Gorry and Scott-Morton's (1971) framework Business Analytics would fall into the category of operational control (Gorry, 1971).

The data is too product specific to be used in an aggregated Management Accounting System and too complex to be analysed by Automated Decision Making Systems. It is the Sales Agent himself who has the most use of seeing the dependencies, causalities, trends and forecasts available through Business Analytics.

To further prove this statement I can cross reference the characteristics and capabilities of Decision Support Systems with the Business Analytics approach I used (Turban, 2011).

Characteristics and capabilities of Decision Support Systems	My Business Analytics approach
DSS gives decision support for semi-structured and unstructured problems by bringing together human judgment and computerized information.	Yes
They can give support to all managerial levels	Yes
They can give support to individuals as well as teams	Yes
They can give support for interdependent or sequential decisions once or on a repeated basis	Yes
Support in all phases of the decision making process	Yes
Support to various processes and styles	Yes
The decision maker should be reactive and adapt the DSS to these changes	Yes
The DSS should be user friendly with graphical capabilities	Yes
It should improve the effectiveness of decision making	Yes
The user should have complete control over all steps in the decision making process	Yes
The end user should be able to modify and develop the system by themselves	Yes
Models are used to analyse the decision making situation	Yes
Access is provided to a wide variety of data sources, formats and types	Yes
The DSS should be a stand-alone tool used by a single user at a single location	Yes

There is however some Technological Risk to this statement. As technology progresses computers will be able to take over more and more of the analytics. We also know that in successful decision support systems the utility of information should exceed or equal the cost to gather said information.

$$U(I) \geq C(I)$$

When the product number increases from the basket of eight goods presented in the study to the 59 000 available it becomes too costly for the company to analyse the sales series for all products. That is not effective decision making. There are two potential solutions to this problem. One is to categorise all products into categories showing similar patterns or having similar characteristics and the other is to try to develop a simplified business analytics approach to Automated Decision Making. Neglecting analytics would be fatal for the company as I have proven that it adds value.

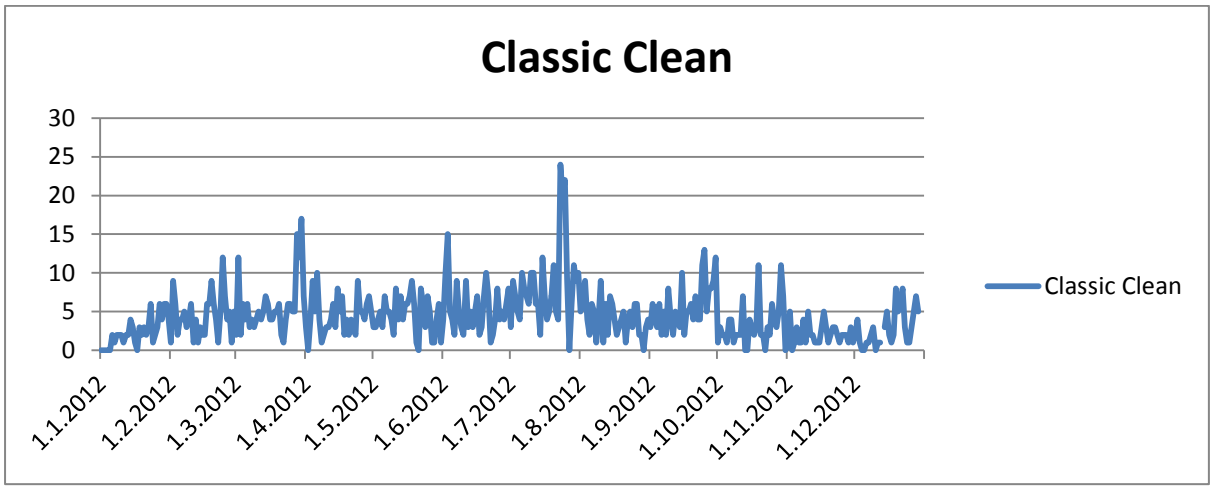
8.0 Conclusion

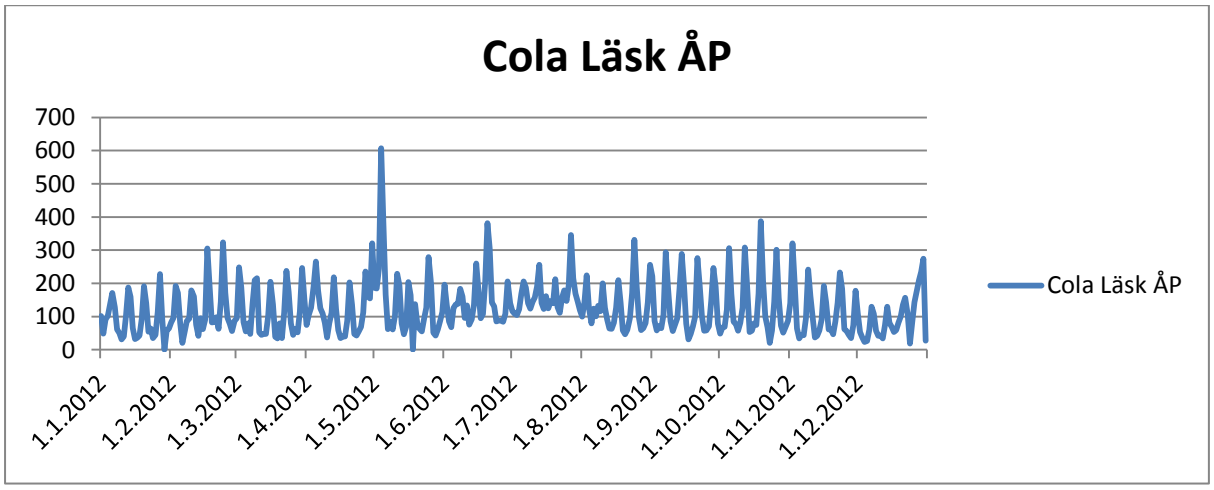
In this thesis I have analysed a basket of eight products and aggregated store variables using time-series analysis. I found that there is value as decision support in analysing the historical data that these products and store variables generate but the use is limited to the number of products. I therefore propose that these methods be used to analyse fewer products, aggregated product categories or the development of Automated Decision Systems that simplify the decision process.

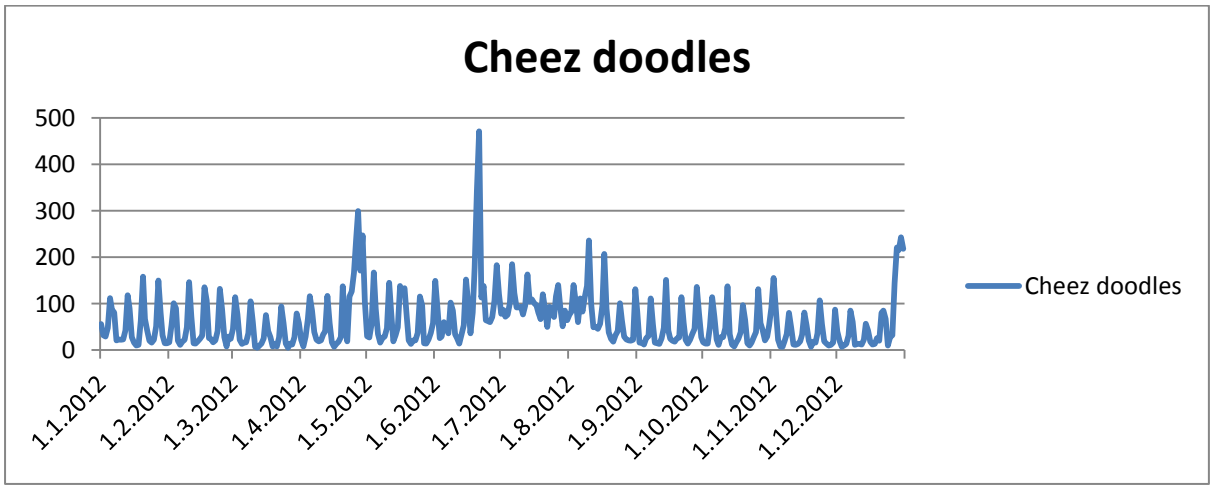
8.1 Further research

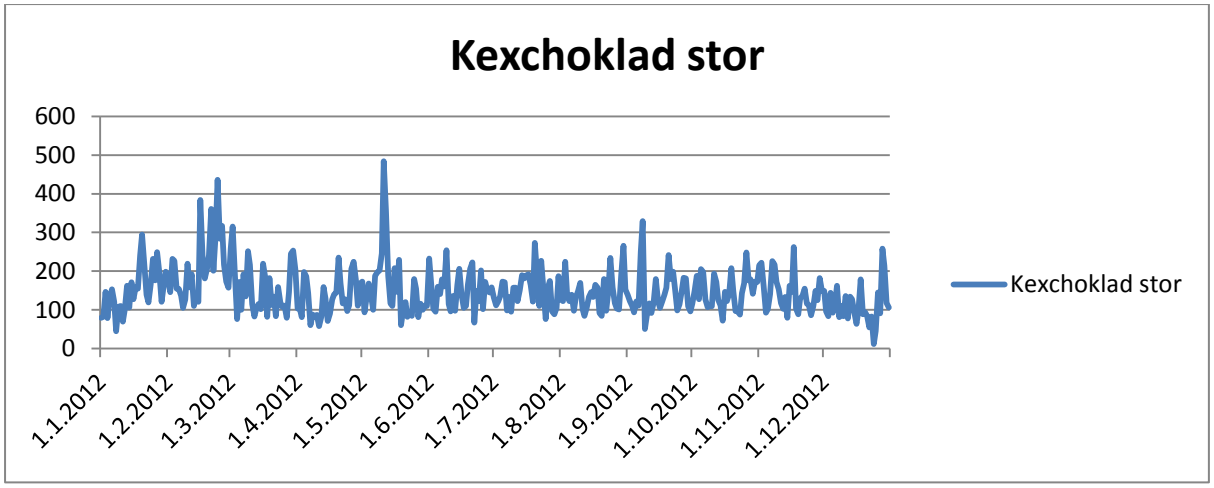
Further research should be focused on how to simplify the methods I proposed so as to be able to review them from an Automated Decision Making platform such as a handheld computer. While a computer can process data very fast through a model it is the algorithms concerning the model selection that is of uttermost importance in such a system. Another interesting research might be focused on how all the 59 000 products develop, change and evolve with each other.

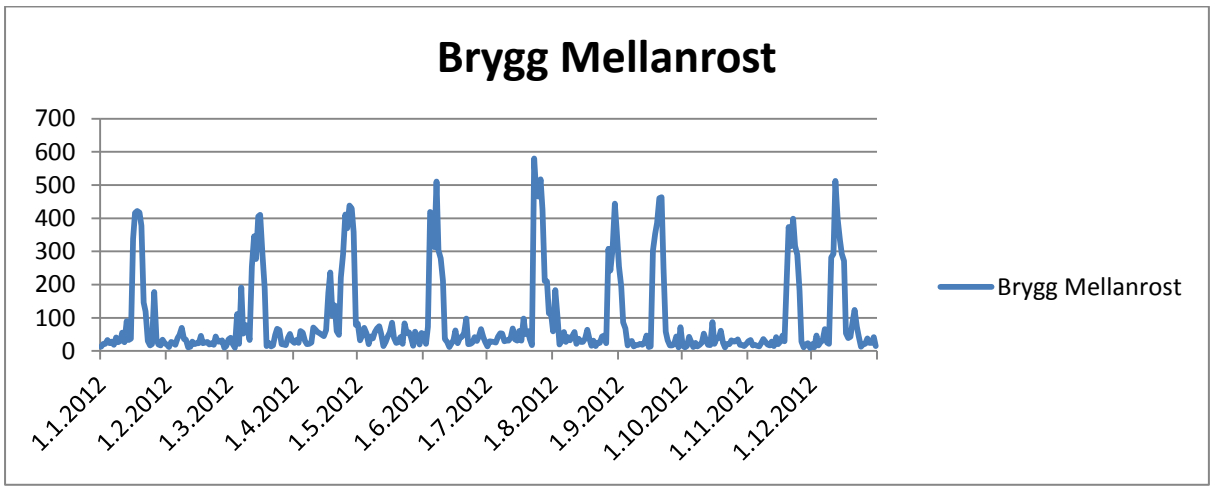
9.0 Attachments

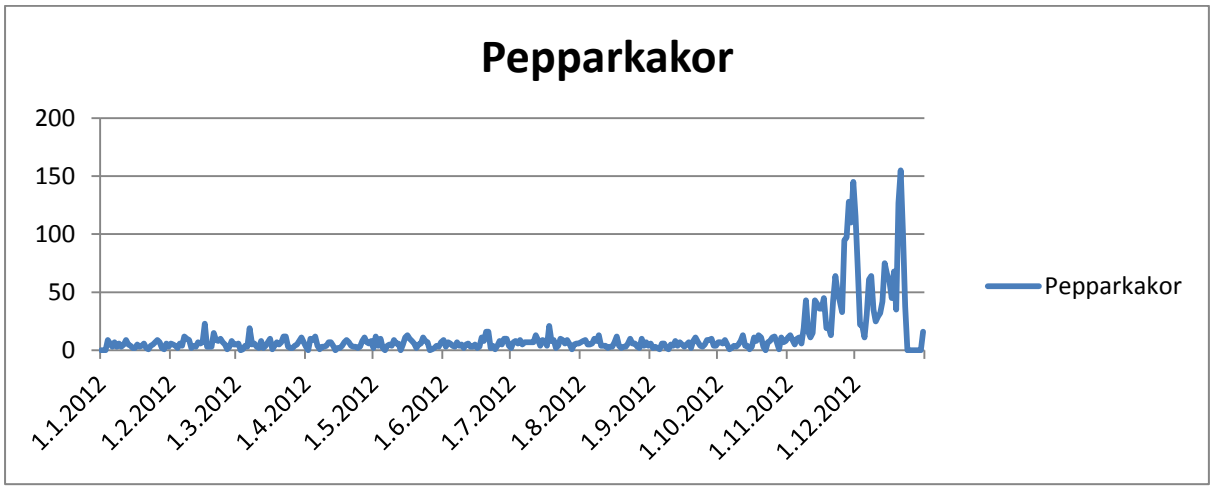


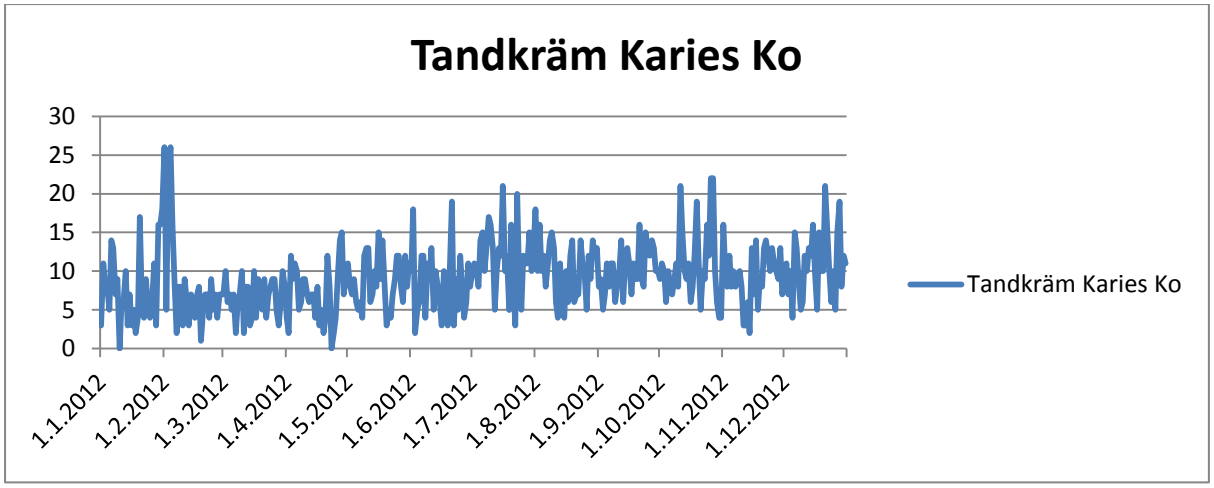


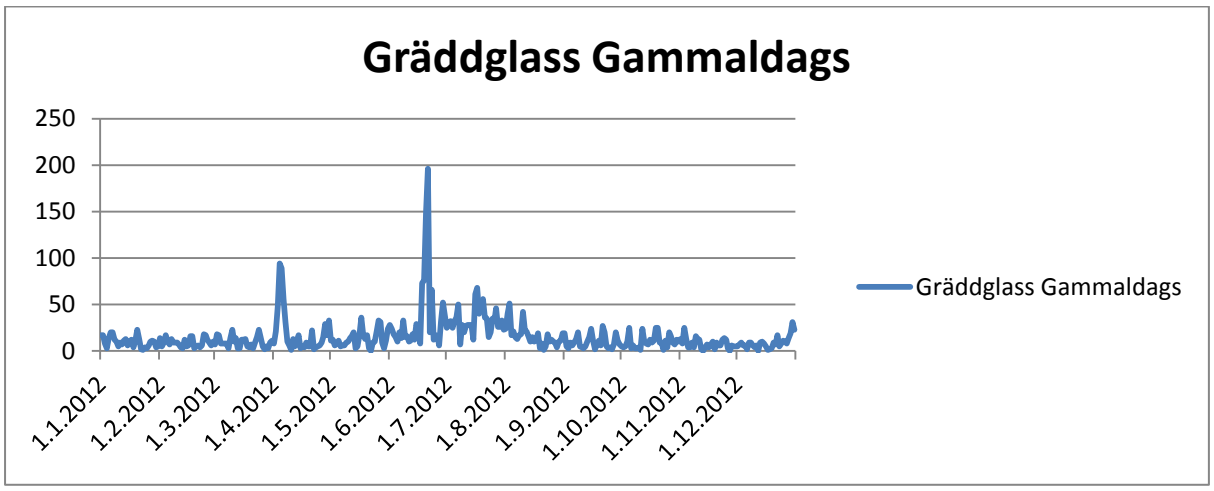



Kexchoklad stor
600
500
400
300
200
100
0
1.1.2012
1.2.2012
1.3.2012
1.4.2012
1.5.2012
1.6.2012
1.7.2012
1.8.2012
1.9.2012
1.10.2012
1.11.2012
1.12.2012
Kexchoklad stor


Brygg Mellanrost
700
600
500
400
300
200
100
0
1.1.2012
1.2.2012
1.3.2012
1.4.2012
1.5.2012
1.6.2012
1.7.2012
1.8.2012
1.9.2012
1.10.2012
1.11.2012
1.12.2012
Brygg Mellanrost


Pepparkakor
200
150
100
50
0
1.1.2012
1.2.2012
1.3.2012
1.4.2012
1.5.2012
1.6.2012
1.7.2012
1.8.2012
1.9.2012
1.10.2012
1.11.2012
1.12.2012
Pepparkakor


Tandkräm Karies Ko
30
25
20
15
10
5
0
1.1.2012
1.2.2012
1.3.2012
1.4.2012
1.5.2012
1.6.2012
1.7.2012
1.8.2012
1.9.2012
1.10.2012
1.11.2012
1.12.2012
Tandkräm Karies Ko


Gräddglass Gammaldags
250
200
150
100
50
0
1.1.2012
1.2.2012
1.3.2012
1.4.2012
1.5.2012
1.6.2012
1.7.2012
1.8.2012
1.9.2012
1.10.2012
1.11.2012
1.12.2012
Gräddglass Gammaldags

10.0 Bibliography

Alter, S. (2004). A work system view of DSS in its fourth decade. *Decision Support Systems*, 319-327.

Bhasin, K. (2012, July 16). *9 Trends That are Transforming The Retail World*. Retrieved 2 5, 2013, from Businessinsider.com: http://www.businessinsider.com/9-trends-that-are-transforming-the-retail-world-2012-7?op=1

Damodar Gujarati, D. P. (2009). *Basic Econometrics.* Singapore: McGraw-Hill.

Dimitrios Papakiriakopoulos, K. P. (2009). A decision support system for detecting products missing from the shelf based on heuristic rules. *Decision Support Systems*, 685-694.

Doyle, S. A. (1890). *A Scandal in Bohemia* .

Drury, C. (2012). *Management and Cost Accounting.* China: RR Donnelly.

Fisher, R. (2010). *The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance.* Harvard Business Press.

Fuller, D. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of American Statistical Organisation*.

Gorry, S.-M. (1971). A framework for Management Information Systems. *Sloan Management Review*.

Granger. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*.

Granger, R. F. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*.

Gustke, C. (2013, 4 15). *Retail Goes Shopping Through Big Data*. Retrieved 4 26, 2013, from CNBC: http://www.cnbc.com/id/100638141

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 5-10.

James J. Jiang, G. K. (1998). A marketing category management system: a decision support system using scanner data. *Decision support systems*, 259-271.

Jarlaskog, M. (2013, 3). Financial group at ICA Flygfyren. (O. Andersin, Interviewer)

Kapoor, M. a. (1981). Modeling and Forecasting Sales Data by Time Series Analysis. *Journal of Marketing Research*, 94-100.

Keen, S. M. (1978). *Decision Support Systems.* Addison-Wasley Publishing Company.

Koop. (2008). *Introduction to Econometrics.* Chichester: John Wiley & Sons.

Newbold, C. T. (2010). *Statistics for Business and Economics.* New Jersey: Pearson.

Oxford Dictionaries. (2013). *Causality*. Retrieved 2 19, 2013, from Oxford Dictionaries: http://oxforddictionaries.com/definition/english/causality?q=causality

Philip Doganis, A. A. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. *Journal of Food Engineering*.

R.J. Kuoa. (1999). Fuzzy neural networks with application to sales forecasting. *Fuzzy Sets and Systems*.

Rania Shibl, M. L. (2013). Factors influencing decision support system acceptance. *Decision Support Systems*, 953-961.

Se´bastien Thomassey, A. F. (2006). A hybrid sales forecasting system based on. *Decision Support Systems*, 408-421.

Soussan Djamasbi, E. T. (2008). Do men and women use feedback provided by their Decision Support Systems (DSS) differently? *Decision Support Systems*, 854-869.

Stoll, C. (1995). *Silicon Snake Oil: Second Thoughts on the Information Highway.* Doubleday.

Thau, B. (2013, 4 11). *J.C. Penney Shoppers React To CEO Ouster.* Retrieved 4 11, 2013, from Forbes: http://www.forbes.com/sites/barbarathau/2013/04/11/j-c-penney-shoppers-react-to-ceo-ouster/

Turban, S. D. (2011). *Decision Support and Business Intelligence Systems.* New Jersey: Pearson Education.

Wallace, C. (2013, 4 26). *Why retail should buy into 'Big Data'*. Retrieved 4 26, 2013, from Businesspectator: http://www.businessspectator.com.au/article/2013/4/26/technology/why-retail-should-buy-big-data

van Zanten, S. K. (2012, 8). *Surfing the Data Deluge - How Retailers Can Turn Big Data into Big Profits.* Retrieved 4 26, 2012, from Cisco: http://www.cisco.com/web/about/ac79/docs/retail/Big-Data-Retail.pdf

Watson, H. (2005). Sorting Out What's New in Decision Support. *Business Intelligence Journal*.

WINTERS, P. R. (1960). FORECASTING SALES BY EXPONENTIALLY WEIGHTED MOVING AVERAGES. *Management Science,*, 324-342.

Xueb, R. K. (1998). A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights. *Decision Support Systems*.

Z.X. Guo a, W. W. (2013). A multivariate intelligent decision-making model for retail sales forecasting. *Decision Support Systems*.

Zhan-Li Sun, T.-M. C.-F. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, 411-419.