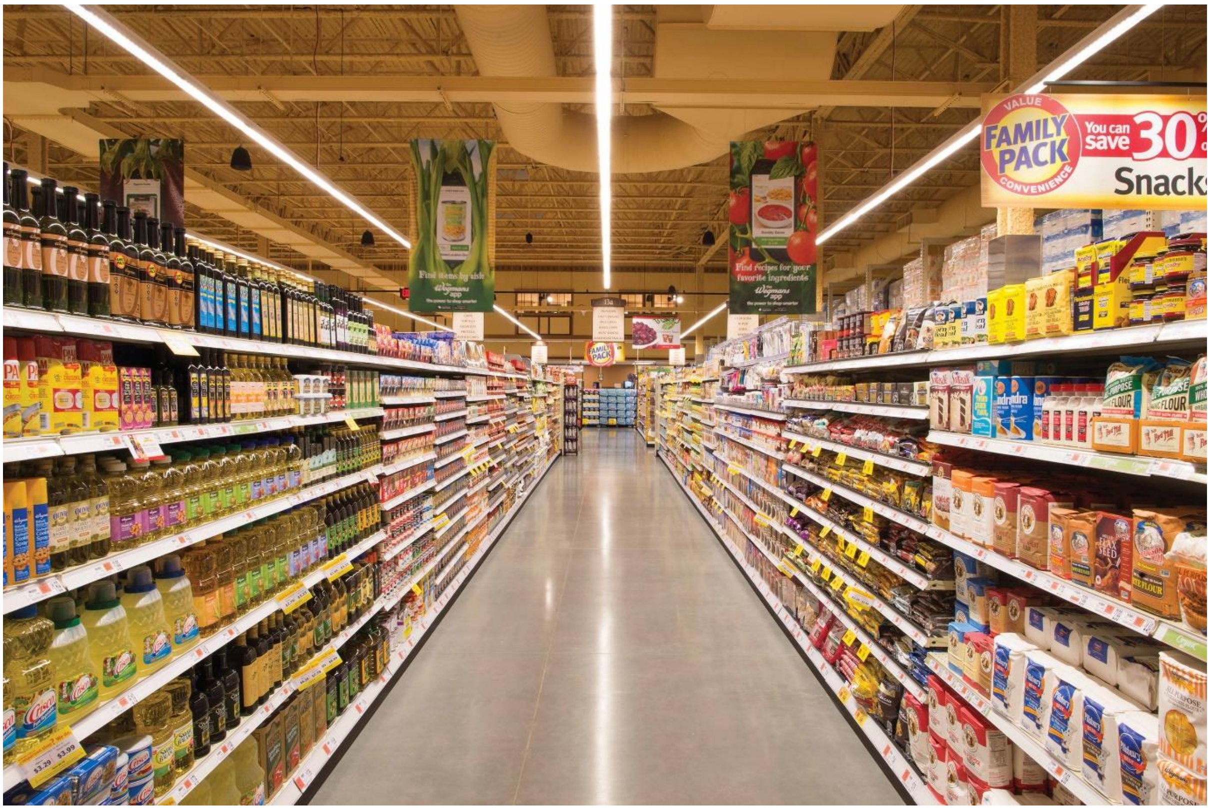
BUSINESS INTELLIGENCE

Instructed by: Farah Mehboob

SALES ANALYTICS REPORT



Prepared for **Supermarket Company**

REVISION HISTORY

Version #	Release Date	Comments	Author-ID
1.1	23rd October 2019	Initial draft	Mishaal Amin Hajiani-13050
1.2	20th November 2019	Reviewing phase	Mishaal Amin Hajiani-13050
1.3	5th December 2019	Final draft	Mishaal Amin Hajiani-13050

Last Updated: 5th December 2019

Contents

INTRODUCTION

The growth of supermarkets are increasing and market competitions are also very high. **Sales analytics** is the procedure used to identify, model, understand and predict **sales** trend. It helps to understand and improve these trends. It is used to govern the success of a previous **sales** drive and helps to forecast the future.

DATA SET INFORMATION

The dataset is one of the historical sales of supermarket company which was recorded in 3 different branches over 3 months of data.

OBJECTIVES

Predictive Data analysis can help in following areas to increase sales via:

1. Promotions
2. Shopper targeting
3. Marketing Campaign Management
4. Pricing
5. Inventory management

ABOUT THIS DATASET

This dataset has be picked from Kaggle. It contains 1000 rows and 17 columns.

DATA ATTRIBUTES

Invoice id: Computer generated sales slip invoice identification number
Branch: Branch of supercenter (3 branches are available identified by A, B and C).
City: Location of supercenters
Customer type: Type of customers, recorded by Members for customers using member card and Normal for without member card.
Gender: Gender type of customer
Product line: General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel
Unit price: Price of each product in $
Quantity: Number of products purchased by customer
Tax: 5% tax fee for customer buying
Total: Total price including tax
Date: Date of purchase (Record available from January 2019 to March 2019)
Time: Purchase time (10am to 9pm)
Payment: Payment used by customer for purchase (3 methods are available – Cash, Credit card and Ewallet)
COGS: Cost of goods sold
Gross margin percentage: Gross margin percentage
Gross income: Gross income
Rating: Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

DATA PROFILING REQUIREMENTS

This Dataset contains no missing values, outliers or any invalid values. Numeric columns are normalised. Invoice-ID column is deleted as it contains no useful information. Month from the date column is extracted into a new column. Dates columns was also standardised to one date format as dates were input with different data styles. All data cleaning, ETL is done in excel and PowerBI(using Query).

SOFTWARE TOOLS USED

1. Excel
2. PowerBI
3. Jupiter (python)

DATA USAGE POSSIBILITIES

1. **Challenge: To Increase Customer Loyalty**

Solution : Predictive Analytics can uses information gathered from these acts, matches them with the actual purchases, and aids retailers anticipate the customers' needs. We can gather information from Membership card. Thus, it helps to build well-targeted promotions and loyalty programs, and also helps build a marketing message that includes an array of products for those customers who have, in the past, demonstrated their inclination to spend.

Results: Promotions(Coupons) to customers helps to boost our sales. It gives us an idea of who are true customers, their spendings, their behaviour, all of which he can align with promotional campaigns.

STAKEHOLDERS

One of the factors contributing to the success of delivering these reports/ dashboards is the overall stakeholders of the project. Stakeholder list for our BI process includes:

1. Sales Department
2. Marketing Department
3. Business Analyst
4. Database Administrator
5. CEO
6. Chief technology officer (CTO)
7. Buyer

WORKFLOW

Data Mining Process: CRISP-DM

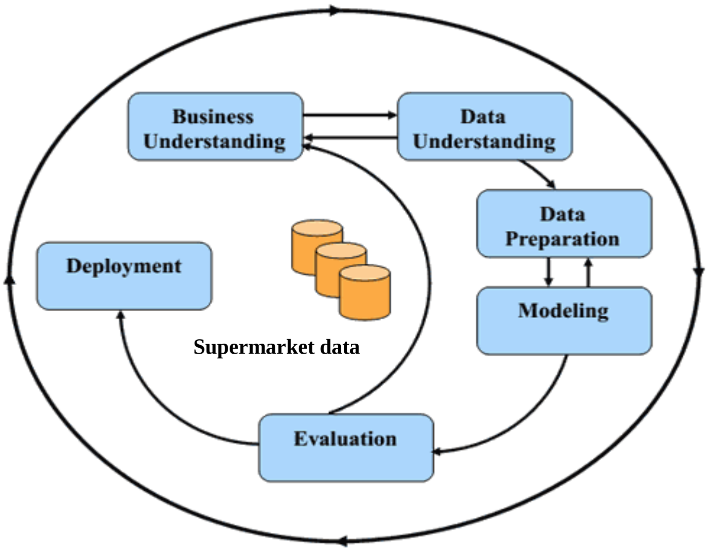


Figure 1

DATA PRE_PROCESSING/ MODELING:

The model run on this dataset is **Linear regression** to predict customer ratings. Linear regression is a basic and commonly used type of predictive analysis.Regression estimates are used to explain the relationship between one dependent variable(Ratings) and one or more independent variables. Data modeling is done in Jupiter Software.

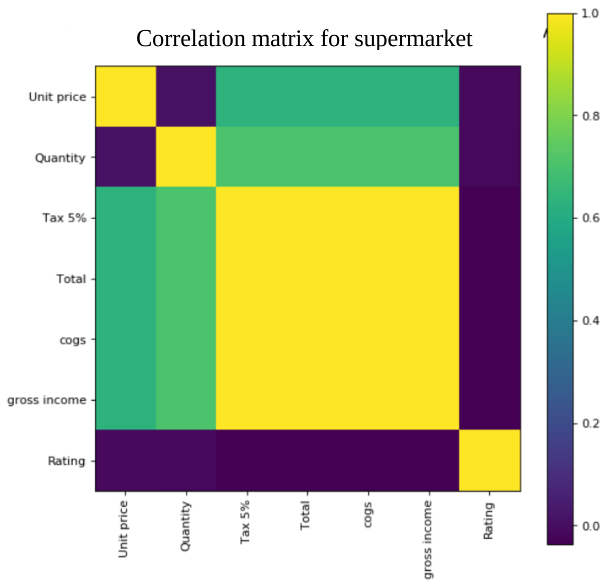


Figure 2

Above Figure 2 shows correlation between each variable. This matrix shows that all positive correlation with each other.

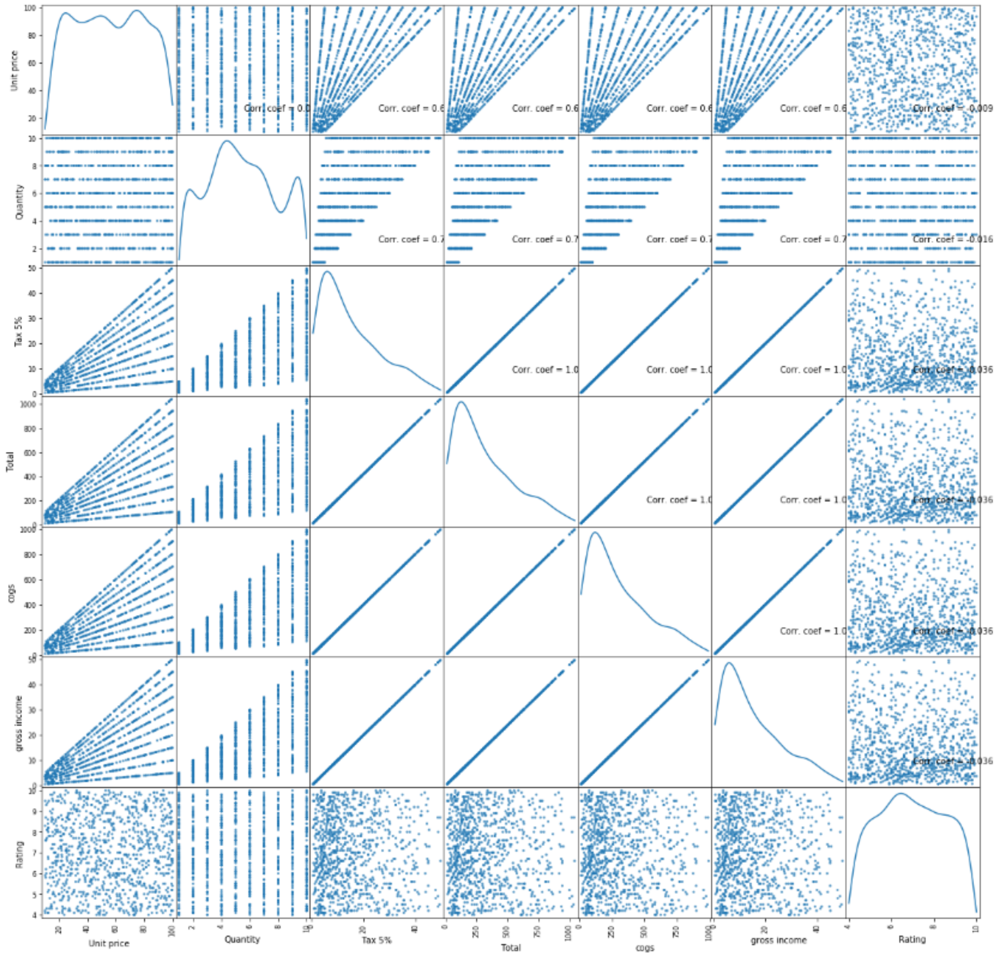


Figure 3

Figure 3 shows scatter plot describing the correlation between each variables.

DATA AUGMENTATION:

Another dataset with columns like discounted price, shipping date and shipping cost can create an impact on all over customer shopping experience; customer ratings. Thus, we can augment these columns in our dataset.

DEVELOPING KPIs

Indicator 1.1.1 : Measuring the conversion rate.The conversion rate is the proportion of store visits to the number of shoppers who made a purchase
Target 1.1.1: Turn the visitors into buyers
Goal 1.1.1: Improve sales performance
Rationale and interpretation: Conversion rate tells you how good you are at turning visitors into buyers. Driving store visits is great, but traffic alone won't add much to your bottom line if your visitors don't convert. If our conversion rate is below 5% we need to take actions to increase conversion rate. We can train and empower our salesperson to:

- Build rapport with customers
- Provide product information and insights
- Be convincing without being pushy

Computation Method: To calculate it, use the formula:

$$\textit{Conversion rate} = \textit{number of sales}(N) \,/\, \textit{total number of visitors}(N_p)$$

Data Disaggregation: Disaggregation should be done by branches, gender, type of customers to produce better results.

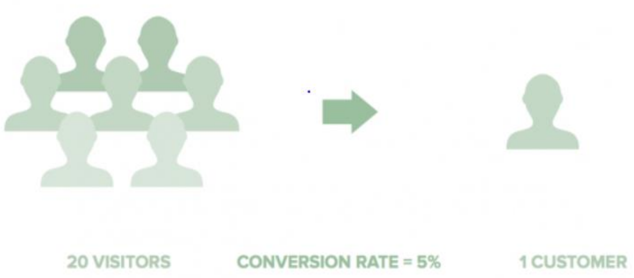


Figure 4

Indicator 1.1.2 : Calculating Gross Income
Target 1.1.2: To make enough profit each year. There should be a threshold matched each year
Goal 1.1.2: Improve sales performance
Rationale and interpretation: Gross profit tells you how much you have made after deducting the costs of creating and selling the product. Your gross and net profit will indicate whether or not you're making any money or not. Generating sales and revenue is good, but at the end of the day, you need to make money out of those sales.
Tracking these KPIs will help you make smarter decisions in various aspects of your business. For instance, if your gross profit is on the low side, then you may want to look into product sourcing and determine if there's a way to lower your cost of goods. Perhaps you can find ways to lower your operating expenses.
Few ways we can improve gross and net profit :

- Raise your prices
- Increase your average order value
- Optimize your vendor relationships

Computation Method:

***Gross income= sales revenues(Total) – cost of goods sold* (COG)**

Data Disaggregation: Disaggregation should be done by branches, gender and product line to produce better results.

Indicator 1.1.3 : Calculate monthly sell-through
Target 1.1.3: To order products or keep product line in stock by reading the sell-through pattern in order to save money
Goal 1.1.3: Improve sales performance
Rationale and interpretation: Sell through is the percentage of units sold versus the number of units that were available to be sold. Sell through is a great way to evaluate merchandise performance. It also helps you figure out the speed at which a product is selling so you can make the right purchasing decisions.

Computation Method:

Sell-through rate= [number of units sold(Quantity) / inventory(beginning)] x 100

Data Disaggregation: Disaggregation should be done by branches and product line to produce better results.

Figure 5

Indicator 1.1.4 : Customer Retention
Target 1.1.4: To increase and retain the number of membership.
Goal 1.1.4: Improve sales performance
Rationale and interpretation: customer retention rate tells the amount of customers that return to the store. This metric can gauge for customer service, product performance, and loyalty.Getting people to come back boils down to how well you manage your customer relationships. Tracking customer purchases and offering personalized recommendations. For this purpose we need to have a count of customer consisting membership card.

Computation Method:

$$\text{Customer Retention}=((CT\text{-}CN)/CM))\ x\ 100$$

CT = number of customers at the end of month (total transactions)
CN = number of new customers (type: normal)
CM = number of customers(type: membership)registered.

Data Disaggregation: Disaggregation should be done by branches to produce better results.

Indicator 1.1.5 : Month over Month (MOM)Growth
Target 1.1.5: Continuous Monthly growth
Goal 1.1.5: Improve sales performance
Rationale and interpretation: How better are you performing compared to your previous months in business? Helps in tracking progress to measure current results against the previous period. To improve supermarket's MOM check, If the growth has been stalled or not. It helps to drill down the reason behind it.

Computation Method:

$$\text{Month over Month}= [(\text{current month total} - \text{prior month total}) / \text{prior month total}]\ x\ 100$$

Data Disaggregation: Disaggregation should be done by branches, product line to produce better results.

COMPUTING AVERAGES

a) **Arithmetic Mean:** The average of a set of numerical values, as calculated by adding them together and dividing by the number of terms in the set.

$$A = \frac{1}{n} * \sum_{i=1}^{n} x_i$$

Formula 1

Where:

- A = average (or arithmetic mean)
- n = the number of terms (e.g., the number of items or numbers being averaged)
- x_1 = the value of each individual item in the list of numbers being averaged

Using formula 1 arithmetic for the following column is:

Column name: Ratings

Overall Rating average= $\frac{SUM(Ratings)}{Number\ of\ rows} = \frac{6972.7}{1000} = \mathbf{6.972}$

b) Geometric Mean: It is defined as the nth root of the product of n numbers

$$\bar{x}_{geom} = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

Formula 2

Where:

- n is the total number observations
- $\sqrt[n]{\prod_{i=1}^{n} x_i}$ is the n^{th} square of the product of the given numbers

Using Formula 2 geometric mean for the following column is:

Column name: Total (Revenue), Ratings

Overall Total(Revenue) average= $\sqrt[no.\ of\ rows]{product\ of\ Revenue}$ = **226.9243**

Overall Ratings average= $\sqrt[no.\ of\ rows]{product\ of\ Ratings}$ =**6.75**

c) Harmonic mean: It is used to calculate the average of the ratios or rates. It is calculated by dividing the number of values in the data series by the sum of reciprocals ($1/x_i$) of each value in the data series.

$$H = \frac{n}{\sum \left[\frac{1}{x_i} \right]}$$

Formula 3

Where:

- n is the number of the values in a dataset
- x_i is the point in a dataset

Note: There is no such rates or ratio columns in the dataset where harmonic mean could be useful.

d) Rolling Averages: A simple rolling average (also called a moving average) is the unweighted mean of the last n values. In our case n=10

Branch A	Ratings	#N/A
A	6.9	#N/A
A	7.7	#N/A
A	6.9	#N/A
A	5.7	#N/A
A	4.3	#N/A
A	8.4	#N/A
A	4.4	#N/A
A	8.5	#N/A
A	4.8	6.73
A	9.7	6.48
A	4.4	6.18
A	4.7	6.35
A	8.6	6.35
A	5.7	6.48
A	5.6	6.07
A	4.3	6.23
A	6	6.34
A	9.6	6.27
A	4.1	5.85
A	5.5	5.91

Table 1

Table 1 shows rolling out averages for Branch A at the interval of 10. We can find out rolling out averages for different Branches.

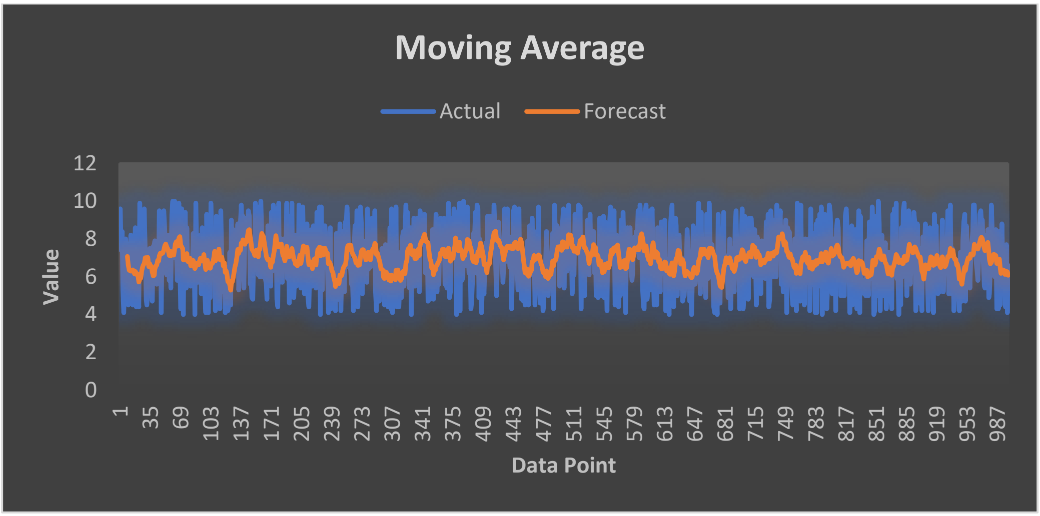


Figure 6

Figure 6 shows the graphical depiction of rolling out averages for branch A. It shows how the actual rating deviates from the forecasted rating value.

SIMPSON'S PARADOX CASE:

Simpson's paradox is a phenomenon in statistics, where the trend appears in several different groups of data but changes when these groups are combined. We found unreliable average as shown below.

Branch	Average of Rating
A	7.03
B	6.82
C	7.07
Overall	**6.97**

Table 2

Table 2 statistics shows that overall rating of this supermarket company is 6.97 but individual branches shows different statistics indicating that branch B consist lowest rating thus, it deviates the overall statistics.

Average of Rating	
A	
Electronic accessories	
Member	7.11
Normal	6.70
Fashion accessories	
Member	6.65
Normal	7.05
Food and beverages	
Member	7.03
Normal	7.48
Health and beauty	
Member	7
Normal	6.81
Home and lifestyle	
Member	6.80
Normal	7.07
Sports and travel	
Member	7.34
Normal	7.18

B	
Electronic accessories	
Member	7.17
Normal	7.06
Fashion accessories	
Member	6.46
Normal	7.00
Food and beverages	
Member	7.02
Normal	6.95
Health and beauty	
Member	7.18
Normal	7.02
Home and lifestyle	
Member	6.75
Normal	6.35
Sports and travel	
Member	6.16
Normal	6.84
C	
Electronic accessories	
Member	6.4
Normal	6.96
Fashion accessories	
Member	7.63
Normal	7.25
Food and beverages	
Member	6.95
Normal	7.24
Health and beauty	
Member	7.04
Normal	6.96
Home and lifestyle	
Member	7.03
Normal	7.1
Sports and travel	
Member	6.98
Normal	7.11

Table 3

Table 3 also show different ratings depending upon the type of customer and product line producing different Customer stratification rating on their overall shopping experience compared to the combined rating average.

DASHBOARDING

OBJECTIVE OF DASHBOARDING

The objective of dashboarding is to visualize the Key Performance Indicators(KPIs) and other strategic data for organizations. Dashboards provide an unbiased view not only of the company's performance overall, but each department. For example, in this case sales and marketing department can use this dashboard and past experiences to increase customer acquisitions and improve demand generation. Business dashboards provide a good starting point for these decisions.

TYPE OF DASHBOARDING : Analytical Dashboarding

For Sales analytics to gain insight on our past, present, and predictive data. we can use this information to understand the current strategy, and determine what adjustments need to be made in the future. To Identify patterns and opportunities in our data for better predictions.

DATA VISUALISATION ROLE WISE

a) Business Analyst and CEO
These two stakeholders have a full view of the dashboard.

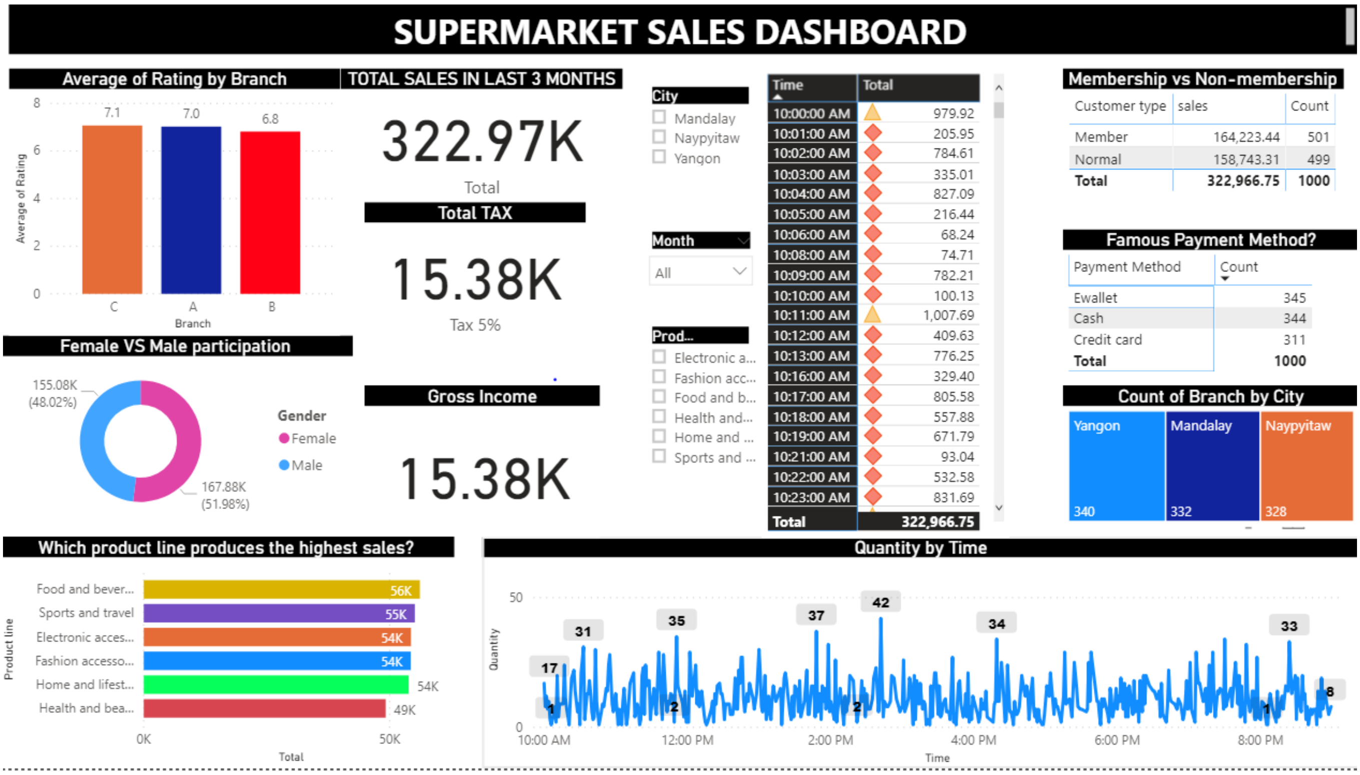


Figure 7: Dashboard preview 1 using powerBI

b) Sales Department
Sales Department will view gross income, branch wise total revenues and Monthly sales.

c) Marketing Department
Marketing Department will view at what time highest sales are done, which type of customer contributes more to the revenue and which products are sold most and which are sold least so that marketing strategy can be developed accordingly(promotions.)

ANOTHER DISPLAY TO UNDERSTAND SUPERMARKET SALES

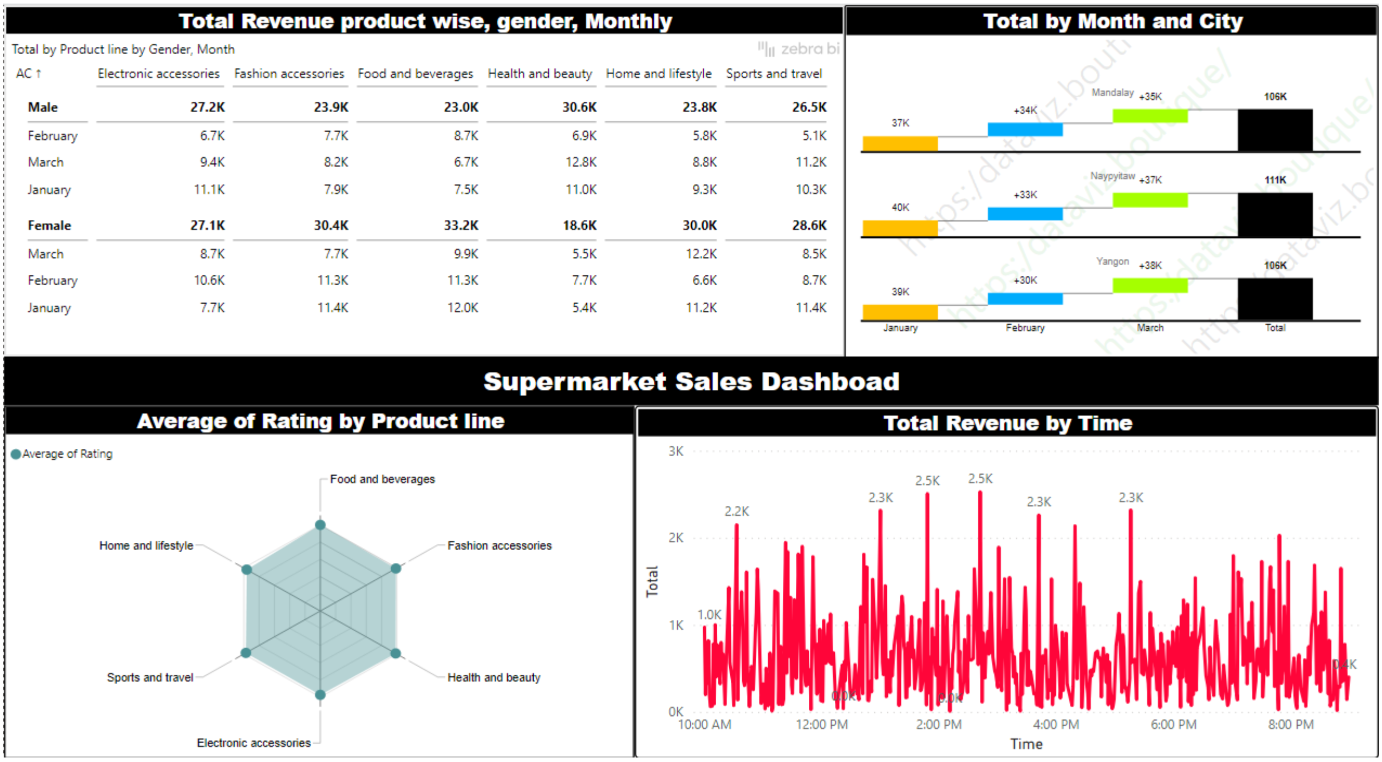


Figure 8: Dashboarding preview 2 using powerBI

Figure 8 shows in-depth analysis of transactions done by customers. Total Revenue is break down according to product, gender and monthly basis using Zebra chart visual. This shows that male spends more money on electronic accessories where as, female highest spending was observed for fashion products. The Radar chart (spider web) shows the average rating for each product line, it was noticed that Food and beverages produces the highest average rating that is 7.1 and the lowest that was 6.8 observed for Home and lifestyle products. Waterfall representation was used to analyze total revenue generated in each city monthly. We observed that highest number of sales were done in the month of march. So sales slowly and gradually increases each month. We can expect increase of sales in February and so on. Lastly, Line graph shows that at what time highest revenue was generated. Over all Highest sales were recorded was 2.5 K in a day taken place between 2.00 PM to 4.00 PM.

BAYESIAN NETWORK

Software used: **GeNIe modeler**

Algorithm used: **Bayesian Network**

Most algorithms for **Bayesian Network** are designed for discrete variables. Before loading our data, we have discretized few predictors.

Preparing the Data:

Discretization of variable 'Total' spending of customers into 3 labels as following:

Descriptive statistics for the intervals					
Lower bound	Upper bound	Frequency	Relative frequency	Density	Label
[10.6785,	354.669[	636	0.636	0.002	State 1
[354.669,	698.66[	255	0.255	0.001	State 2
[698.66,	1042.65]	109	0.109	0.000	State 3

Discretization of variable 'Quantity' spending of customers into 2 labels as following:

Descriptive statistics for the intervals:					
Lower bound	Upper bound	Frequency	Relative frequency	Density	Label
[1,	5.5[	504	0.504	0.112	State 1
[5.5,	10]	496	0.496	0.110	State 2

Discretization of variable 'Ratings' spending of customers into 5 labels as following:

Descriptive statistics for the intervals:					
Lower bound	Upper bound	Frequency	Relative frequency	Density	Labels
[4,	5.2[	195	0.195	0.163	State 1
[5.2,	6.4[	196	0.196	0.163	State 2
[6.4,	7.6[	209	0.209	0.174	State 3
[7.6,	8.8[	204	0.204	0.170	State 4
[8.8,	10]	196	0.196	0.163	State 5

BUILDING A BAYESIAN NETWORK

Bayesian networks are a type of probabilistic graphical model that uses inference for probability computations. **Bayesian networks** aim to model joint probability distribution

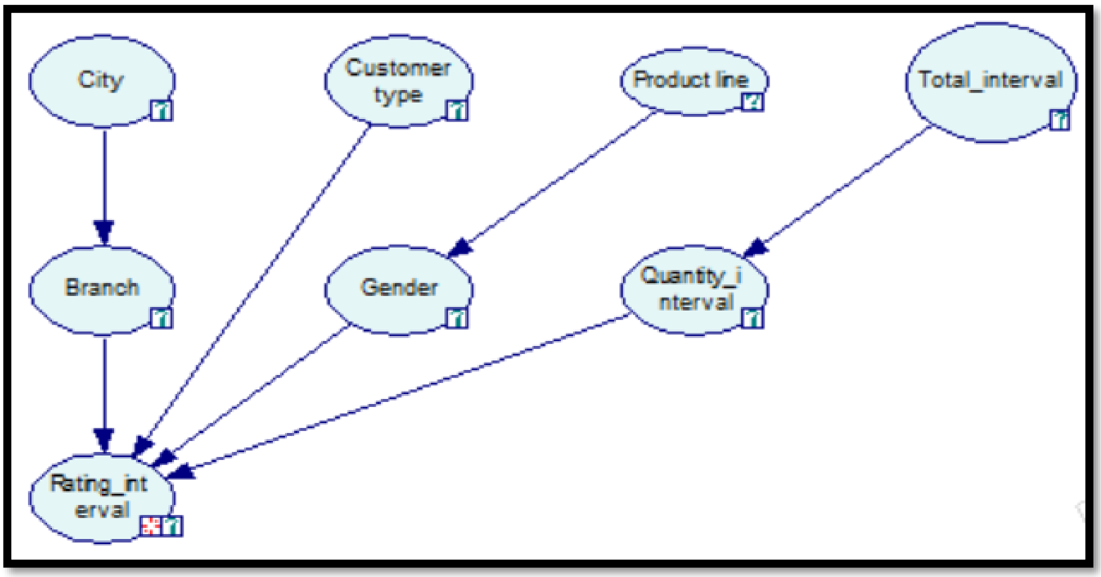


Figure 9: A directed acyclic graph(DAG)

Parameters in Figure 1 are then learned before yielding a fully parametrized Bayesian network.

Now the data is selected to **learn the network**. To invoke **structural learning**, algorithm used is 'Bayesian Search'. It follows essentially a hill climbing procedure (guided by a scoring heuristic) with random restarts. The algorithm produces an acyclic directed graph that gives the maximum score.

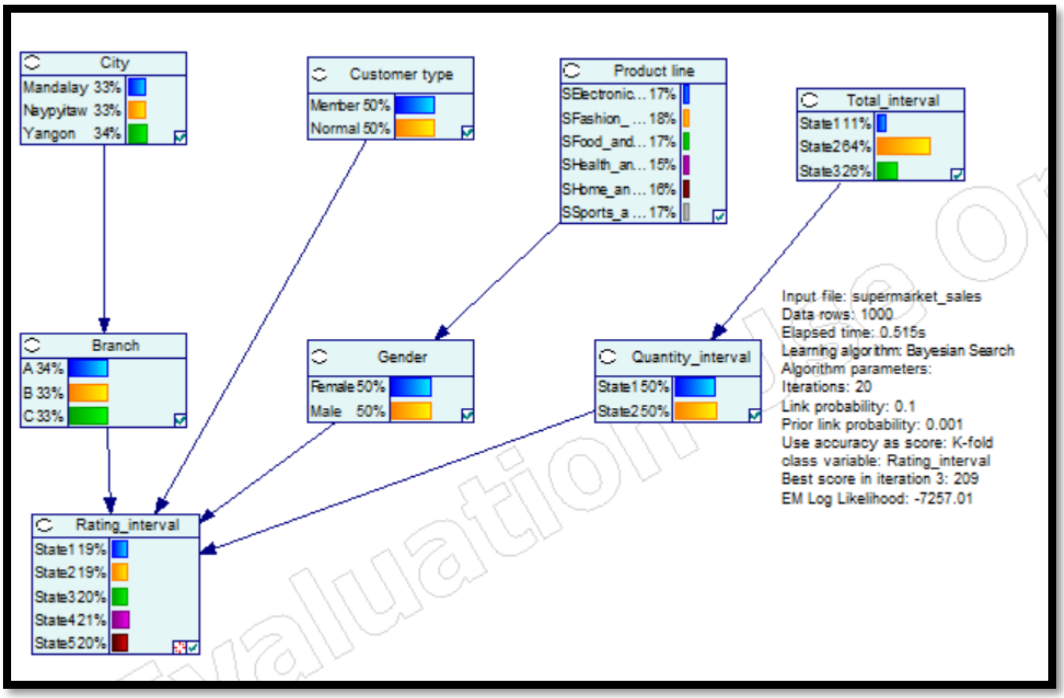


Figure 10: Structure Learning Output

Then **learn the parameters** of an existing network (i.e., one for which the structure is already defined). *Log(p)*, ranging from minus infinity to zero, is a measure of fit of the model to the data. Learning parameters functionality focuses on learning parameters, not the structure. EM algorithm updates the network parameters following the options chosen and comes back with the following dialog:

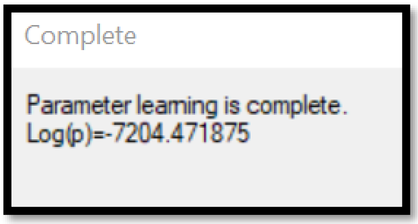


Figure 11: Parameter Learning dialog

Now we'll run the **validation process.** The Evaluation method used here is 'K-fold cross validation'. *K-fold cross validation*, which divides the data set into *K* parts of equal size, trains the network on *K*-1 parts, and tests it on the last, *K*th part. The process is repeated *K* times, with a different part of the data being selected for testing.

Over all Accuracy achieved for the model:

Rating_interval = 0.684 (684/1000) accuracy calculated. In this case, the mode achieved 68.4% accuracy in predicting the correct Rating Intervals i.e. State 1,2,3,4,5.

Correlation behaviour on evidence on a particular node

Target Node: Rating_interval

Sensitivity analysis is technique that can help validate the probability parameters of a Bayesian network. Highly sensitivity affects the reasoning results more significantly. This leads to changing the colouring of the network to indicate where the sensitive parameters are located.

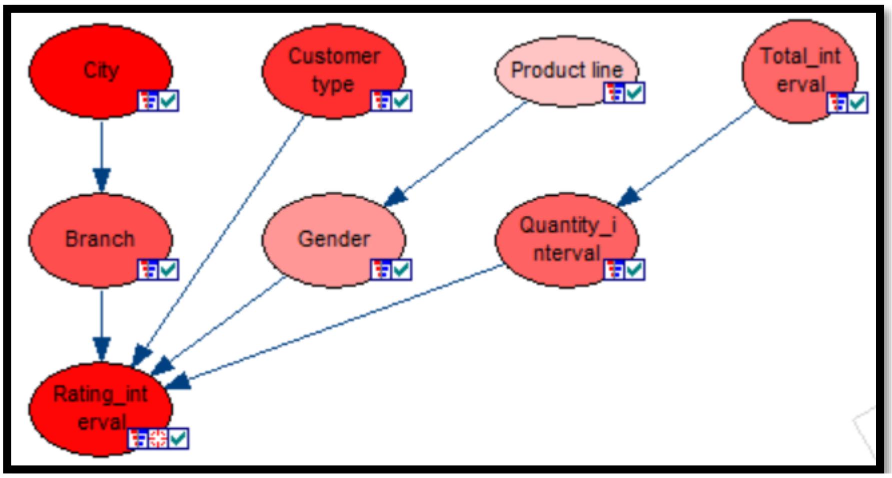


Figure 12:Sensitivity analysis Output

Dark red shows they are highly correlated where as, light pinkish tone shows that they are not strongly correlated. The colouring of the individual elements of the definition shows those individual parameters that are important. In this case city affects branch choices that automatically affects the Rating_interval strongly. Customer type is also strongly correlated to Ratings and Total spending and quantity bought also affects the all over customer experience.

SUPERMARKET DATA MODEL VULNERABILITY AND ITS SOLUTION TO THE DATA SECURITY

At the time when customer makes monetary transaction at the store, their information is momentarily vulnerable, especially we don't have proper security measures in place. This applies to all superstores, but especially to those with multiple store locations. Security is a much larger issue that needs to be approached from multiple angles for success. Our dataset tells that a customer's personal information is collected if he/she is a member, also the payment method extracts information about any customer at the time of transaction even if the customer is not a loyal member.

IT team should **store the data off-site** so that in the event of a natural disaster, we don't lose customer data. Cloud security has come a long way, and is often better than a traditional data center. **Update your in-store software regularly.** Any software system used will release patches and updates that will correct security breaches and previous problems within the system. We should download updates as soon as possible for best results. We should **encrypt store's credit card swiper.** This will require a good POS software. It ensures that no credit card information is stored in your POS device, making it impossible for hackers to access credit card information. Product loss is a huge issue for stores which imbalances the in- stock inventory calculations, for this purpose stores should **Install visible security cameras,** which can be effective at stopping shoplifters. **Verify employees** before you hire them. Our employees can be our biggest advocate for store security or they can be your greatest downfall.Employees are considered the biggest digital and physical security threat to an organization. **Data poisoning** can be on of the issues where someone can change training data to manipulate model's predictions. To poison data, an attacker must have access to some or all of your training data. Some potential measures should be taken to encounter Data poisoning, on of the way is **disparate impact analysis, where it c**ould potentially discover intentional discrimination in model predictions. There are several great open source tools for detecting discrimination and disparate impact analysis, such as Aequitas, Themis, and AIF360. **Benchmark models**, an older but trusted interpretable modelling pipeline can be used to measure whether a prediction was manipulated by any number of means. If the difference between our complexed model and benchmark model is too large, we should refer to the analyst. Lastly, attackers can get the unauthorized data and build the **surrogate model** on it. An attacker could train a surrogate model between the inputs they used to generate the received predictions and the received predictions themselves. The surrogate model can become an accurate simulation of our model. The attacker can mimic examples which attacks against model's integrity, or the potential ability to start reconstructing aspects our sensitive training data. To protect the model against inversion by surrogate model we can have **authorized access,** which requires additional authentication (e.g., 2FA) to receive a prediction. We can use **throttle predictions** where we can restrict high numbers of rapid predictions from single users; consider artificially increasing prediction.

HOW CAN SUPERMARKETS' MEASURE UP ON GDPR?

The General Data Protection Regulation (GDPR) is a set of protection rules.When a customer buy goods and services, or sometimes even just visit a the store, organization collects and processes their information. GDPR allows individual to own their personal data. As a company that's involved in processing that personal data, must **disclose everything** that they do with it in our case informing the customers. The GDPR **sets the rules** about how personal data should be processed. This a way to show your customers that you can be trusted with their personal data increasing customer loyalty. Our dataset defines the column for customer type, this defines whether the customer is a member or not. The customer having membership must have provide their personal data like Email addresses, phone number, address and etc. The rules state that data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. The supermarket cannot pass this personal information to the retailers without their consent. If they wishes to use their personal data for direct marketing this must be explicitly brought to the attention of the customer and presented clearly and separately from any other information. Superstores can profile customers in a number of ways, whether through the use of membership card,Payment methods or using CCTV to record in-store images of known individuals – all of which should fall under GDPR regulations and then the special offers should be made available to the customers to falling under their criteria. The Company should not only take care of their customer data but also their employees' personal data. These superstores need to ensure that they have implemented appropriate procedures to ensure personal data breaches are detected, reported and investigated effectively.

They need to put mechanisms in place to assess and then report relevant breaches to the Information Commissioner's Office(ICO).ICO, publish the names and addresses of the data controllers. They also include a description of the type of processing each organisation performs. They also need to ensure that there are mechanisms in place to notify affected individuals where the breach is likely to result in a high risk to their rights and freedoms. The GDPR rules introduce mandatory data breach notifications to the ICO within 72 hours and in some cases to the data subjects too. The IT system needs to ensure data protection with great care.

COLLECTIVE INTELLIGENCE

Collective intelligence (CI) is shared or group **intelligence** that emerges from the collaboration, **collective** efforts, and competition of many individuals and appears in consensus decision making. It can enhance business outcomes by improving how organizations access the untapped knowledge and experience of their networks. Harnessing Collective Intelligence can play an important role in generating new ideas, solving age-old problems, disaggregating and distributing work in new and innovative ways, and making better, more informed decisions about the future.

HARNESSING COLLECTIVE INTELLIGENCE IN SALES ANALYTICS

Using collective intelligence requires *a lot* of data. To gain collective intelligence, you need to log all the relevant data that will be useful to users. This includes things like each call placed, each action taken on an email, at what timestamp highest transactions are done and what the title of the target person and their geography are. As marketing professionals communicate value and manage customer relationships, they must target changing markets, and personalize offers to individual customers. Traditional marketing methodologies struggled to produce actionable insights from such information quickly, emerging collective intelligence techniques enable marketing professionals to understand and act on the observed behaviours, preferences and ideas of groups of people. Marketing professionals apply collective intelligence technology to create behavioural models and apply them for targeting and personalization. As they analyse preferences, match products to customers, discover groups of similar consumers, and construct pricing models, they generate significant competitive advantage. By looking at sales history of products and services, accounting and production systems have related collective intelligence about the customer, so why not share this information with your sales team and allow them to come to their own insights? Is this customer buying different products, or more or less? Is the average price paid changing? We can also dig into membership data list and check their engagements on social platform and create user profiles for them and do churning of customers to target market. In short, Stay in touch and find out who's listening! Do not let customer feel left out or not communicated with on a timely basis. We can start a web site/blog where we can mention the importance of staying in touch and being **"Top of Mind"** with prospects and customers and to capture their feedback from surveys or forum discussions.

TIME ANALYSIS

An expressive visual **storytelling** environment for presenting **timelines.**

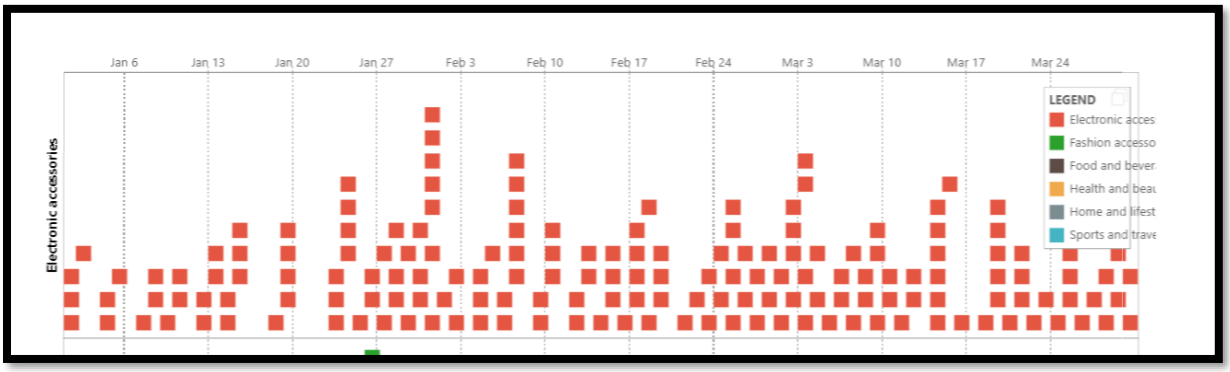


Figure 13

Figure 11 shows a snap of PowerBI representing a linear chronological representation of product 'Electronic accessories' from Jan to Feb 2019

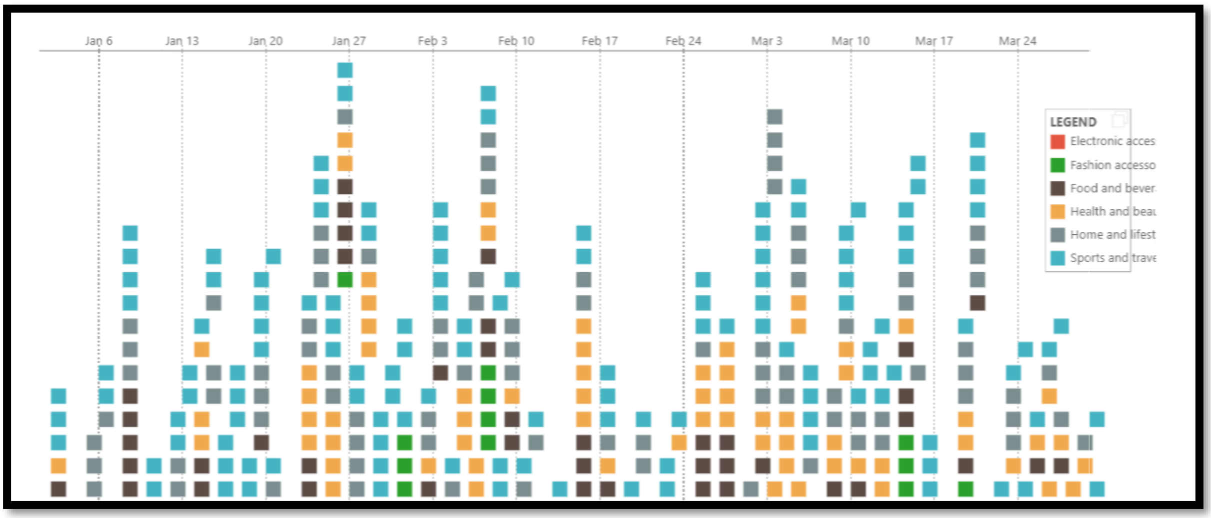


Figure 14

Figure 12 shows a snap of PowerBI representing a linear chronological unified layout of product-lines.

Figure 15

Figure 13 shows a snap of PowerBI representing a radical faceted layout of Electronic accessories from Jan to Feb 2019 with gross income.

INFLUENCE NET ANALYSIS

Tool used: PrecisonTree Palisade Add on in Excel

Influence analysis is an important component of data analysis. This approach is based on the observed data likelihood, which involves multidimensional integrals, directly applying it to develop influence analysis procedures for the factor analysis models with ranking data is difficult. Influence diagrams are a graphical tool for mapping the interaction of the various elements of a decision setting. They usually represent decisions with rectangles; chance events or uncertainties with ovals or circles; calculated or fixed inputs and outputs with rounded rectangles, and outcomes or values with triangles.

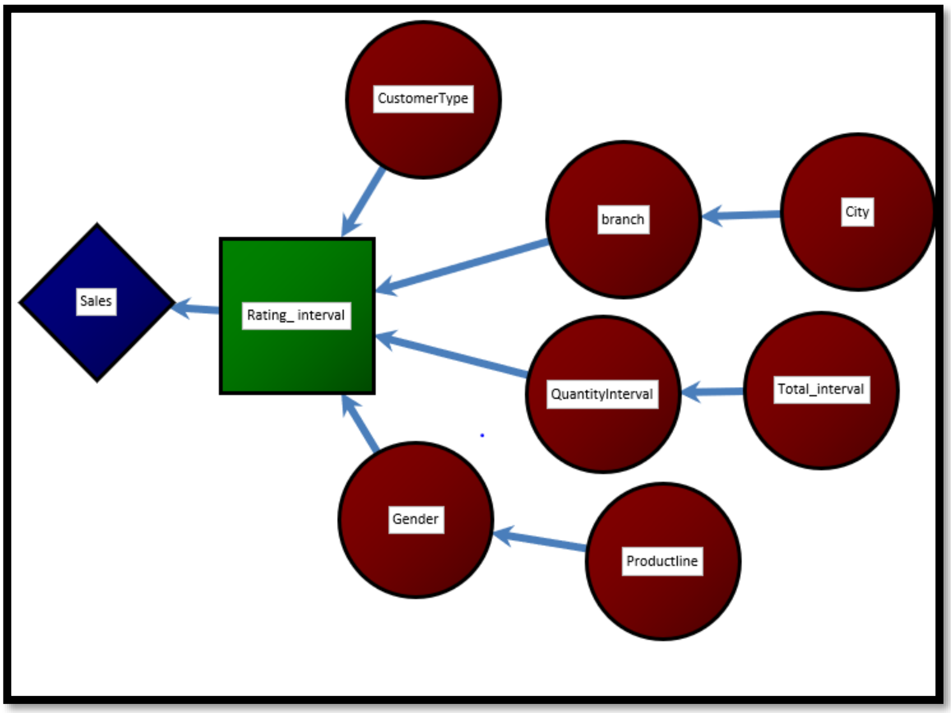


Figure 16: Influence tree

REFERENCES

https://support.bayesfusion.com/docs/GeNIe/bn_sensitivitybn.html

https://digital.hbs.edu/platform-rctom/submission/smart-sales-how-collective-intelligence-is-improving-enterprise-sales-at-pluralsight/

https://www.kaggle.com/aungpyaeap/supermarket-sales

https://progressivegrocer.com/growing-threat-data-breaches-grocery

https://lumina.com/technology/influence-diagrams/

https://unstats.un.org/wiki/display/SDGeHandbook/Indicator+1.1.1

Supermarket image: https://bit.ly/2qrwTgs