# BUSINESS UNDERSTANDING

Vaya Salon based in downtown Chandler,AZ.The business provides various salon services and professional products as demanded.Provides contractual services to models.Mostly Aveda products are used and are also available on retail.Salon has a team of professionally trained artists specializing in different services.

This company acquires contracts from different apparel manufacturing/designing companies or designers. The main origin of the company is to serve the stakeholders or customers/models with different products and services according to the company standards. The company invests its capital in purchasing products, maintaining staff for designing the models and also for the individual customers for all types of special occasions. Company gets the returns in terms of profit/gain after the complete of the period of contract.

# DATA UNDERSTANDING

8 data sets with information about customers,sales,products,date and time,retail price,type of product and manufacturer.

Removed various insignificant variables like salon name,customers contact information in order to get less skewed data for modelling.

Imported data in R Studio to further explore and train AI model.

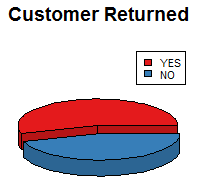
# DATA EXPLORATION

NEW CUSTOMER LIST

8 variables with 1045 rows.Removed rows with missing data.

CUSTOMERS RETURNED

|  |  |
| --- | --- |
| YES | NO |
| 473 | 572 |



PRODUCT DETAIL LIST

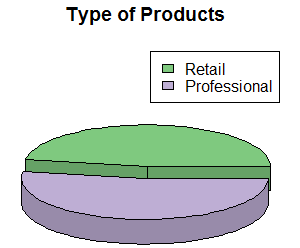
18 variables with 1208 rows.Removed missing values.

VARIOUS MANUFACTURERS

|  |  |  |  |
| --- | --- | --- | --- |
| AVEDA | NOT LISTED | T3 | WET BRUSH PRO |
| 1189 | 12 | 2 | 5 |

TYPE OF PRODUCT SOLD

|  |  |
| --- | --- |
| RETAIL | PROFESSIONAL |
| 572 | 636 |



COST OF PRODUCTS W.R.T. MANUFACTURERS

|  |  |  |  |
| --- | --- | --- | --- |
| NOT LISTED | AVEDA | W3 | WET BRUSH PRO |
| $142.60 | $1572.19 | $120 | $32.97 |

PRODUCT PRICE LIST

7 variables with 1214 rows.No missing values encountered.

“SHAMPOO” is the product sold with highest price at $1424.

76 products are not sold while 59 are sold when grouped with similar categories.

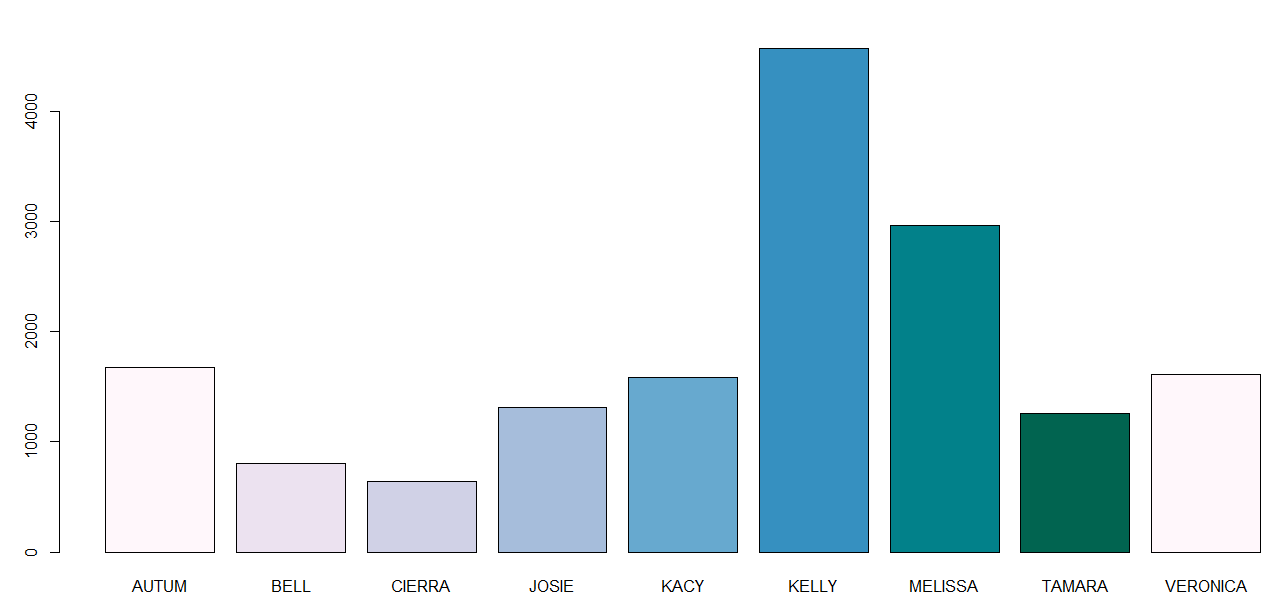
SALES DETAIL SUMMARY

13 variables with 309 rows.

NO missing values.

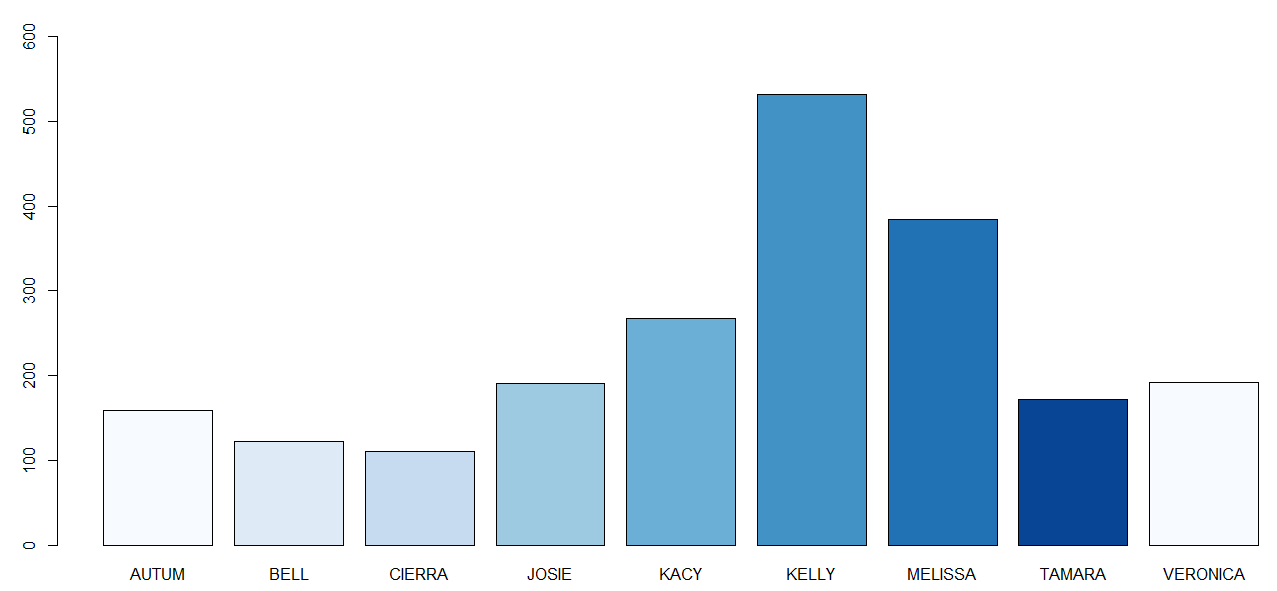
Net amount generated by each staff member

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AUTUM | BELL | CIERRA | JOSIE | KASY | KELLY | MELISSA | TAMARA | VERONCA |
| $1676.1 | $802.50 | $639 | $1315 | $1587 | $4568 | $2962 | $1261 | $1615.5 |



NUMBER OF TICKETS ATTENDED BY EACH STAFF MEMBER

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AUTUM | BELL | CIERRA | JOSIE | KASY | KELLY | MELISSA | TAMARA | VERONICA |
| 159 | 122 | 111 | 191 | 268 | 538 | 384 | 172 | 192 |



SERVICE LIST

14 variables with 118 rows.No missing values.

CATEGORIES OF SERVICES

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| COLOR | COLOR WITH CUT | | DONATION | EXTENSION | | FACIAL | | HAIR CUT | HAIR THERAPY | | MAKEUP |
| 18 | 13 | | 19 | 6 | | 7 | | 5 | 2 | | 8 |
| MASSAGE | | PERMS | | | STYLING | | TEXTURE | | | WAXING | |
| 15 | | 1 | | | 10 | | 1 | | | 13 | |

COMMISONABLE

|  |  |
| --- | --- |
| YES | NO |
| 116 | 2 |

TICKET DETAIL SUMMARY

15 variables with 309 rows.No missing values.

CUSTOMERS WITH PRIVLEDGES

|  |  |
| --- | --- |
| NONE | PRIVLEDGED |
| 299 | 10 |

TOTAL DISCOUNT GIVEN

|  |  |
| --- | --- |
| NONE | PRIVLEDGED |
| $444.65 | $149.40 |

TICKET LIST

11 variables with 369 rows.17 missing in Customer,30 in Ticket Staff.

Omitted rows with missing values.

NUMBER OF TICKETS ATTENDED BY EACH STAFF MEMBER

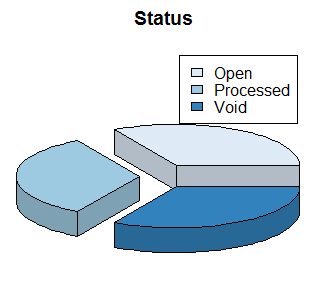
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AUTUM | BELL | CIERRA | JOSIE | KACY | KELLY | MELISSA | TAMARA | VERONICA |
| 173 | 141 | 113 | 186 | 318 | 582 | 352 | 260 | 204 |

TOTAL AMOUNT GENERATED ON ALL TICKETS BY EACH STAFF MEMBER

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AUTUM | BELL | CIERRA | JOSIE | KACY | KELLY | MELISSA | TAMARA | VERONICA |
| $1713.1 | $863.50 | $539.00 | $1340.0 | $2286.5 | $4484 | $2877.5 | $1620.0 | $1314.5 |

AMOUNT OF TICKET WRT EACH STATUS

|  |  |  |
| --- | --- | --- |
| OPEN | PROCESSED | VOID |
| $150.0 | $16452.45 | $454.00 |



# DATA MODELLING

1. means Clustering

Numerical data is needed for K-means model.

“Ticket Detail Summary” dataset with 15 variables.

Using dimension reduction to eliminate insignificant variables having categorical and character values.

Determined suitable number of clusters using sum of WITHIN SUM OF SQAURES over random 15 clusters.

5 is the suitable number of clusters.

Between sum of squares is 955.8086

Size of every cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1st Cluster | 2nd Cluster | 3rd Cluster | 4th Cluster | 5th Cluster |
| 40 | 124 | 65 | 64 | 16 |

WITHIN SUM OF SQUARES

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1st Cluster | 2nd Cluster | 3rd Cluster | 4th Cluster | 5th Cluster |
| 35.651 | 60.631 | 85.367 | 324.462 | 78.08 |

1st  Cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Discount | Average Product Sales | Average Service Sales | Average Other Sales | Average Total Sales |
| $1.011 | $0.0 | $64.164 | $0.0 | $64.164 |

2nd  Cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Discount | Average Product Sales | Average Service Sales | Average Other Sales | Average Total Sales |
| $0.51 | $2.931 | $13.411 | $0.0 | $16.343 |

3rd Cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Discount | Average Product Sales | Average Service Sales | Average Other Sales | Average Total Sales |
| $0.0 | $29.421 | $0.0 | $0.0 | $29.421 |

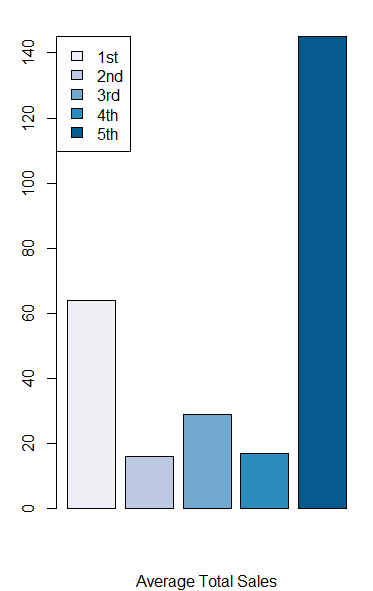
4TH Cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Discount | Average Product Sales | Average Service Sales | Average Other Sales | Average Total Sales |
| $30.0 | $0.0 | $16.43 | $0.0 | $17.142 |

5th Cluster

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Average Discount | Average Product Sales | Average Service Sales | Average Other Sales | Average Total Sales |
| $0.91 | $0.0 | $145.06 | $0.0 | $145.06 |

1. AVERAGE TOTAL SALES in 5th Cluster is the highest followed by 1st Cluster.
2. Highest AVERAGE SERVICE SALES is for the 5th Cluster at $145.
3. AVERAGE DISCOUNT highest for 4th Cluster I.e. $30.
4. AVERAGE PRODUCT SALES maximum in 3rd Cluster.
5. AVERAGE OTHER SALES for all Clusters is $0.0.



“PRODUCT DETAIL LIST” dataset with 18 variables.

Determined suitable number of clusters using sum of WITHIN SUM OF SQAURES over random 15 clusters.

4 is the suitable number of clusters.

Between sum of squares is 1823.465

Size of every cluster

|  |  |  |  |
| --- | --- | --- | --- |
| 1st Cluster | 2nd Cluster | 3rd Cluster | 4th Cluster |
| 59 | 605 | 458 | 92 |

WITHIN SUM OF SQUARES

|  |  |  |  |
| --- | --- | --- | --- |
| 1st Cluster | 2nd Cluster | 3rd Cluster | 4th Cluster |
| 477.932 | 279.2 | 911.30 | 147.11 |

1ST Cluster

|  |  |  |
| --- | --- | --- |
| AVERAGE ON HAND-QTY | AVERAGE COST | AVERAGE RETAIL PRICE |
| -2 | $4.073 | $8.40 |

2ND Cluster

|  |  |  |
| --- | --- | --- |
| AVERAGE ON HAND-QTY | AVERAGE COST | AVERAGE RETAIL PRICE |
| 0 | $7.862 | $0.5611 |

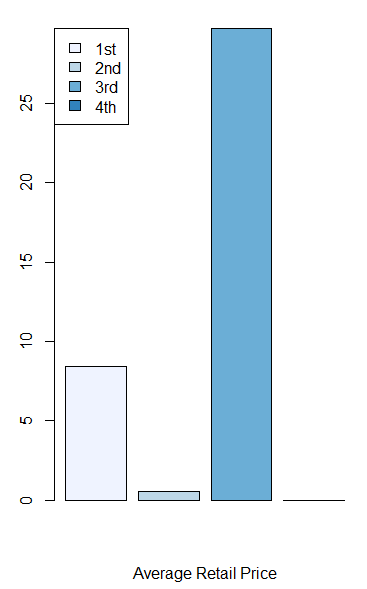
3RD Cluster

|  |  |  |
| --- | --- | --- |
| AVERAGE ON HAND-QTY | AVERAGE COST | AVERAGE RETAIL PRICE |
| 1 | $14.804 | $29.784 |

4th Cluster

|  |  |  |
| --- | --- | --- |
| AVERAGE ON HAND-QTY | AVERAGE COST | AVERAGE RETAIL PRICE |
| 0 | $47.300 | $0 |

1. Average ON HAND-QTY is highest for the 3rd Cluster whereas lowest for 1st Cluster,being -1.
2. Average COST is highest for 4th Cluster and lowest for the 1st Cluster.
3. Average RETAIL PRICE is highest for 3rd Cluster being $29.8.



# F-M Analysis

This analysis tells us about the visits by customers.

It answers various questions like how recently a customer visited,how frequently they visit and average money they spend over every visit.

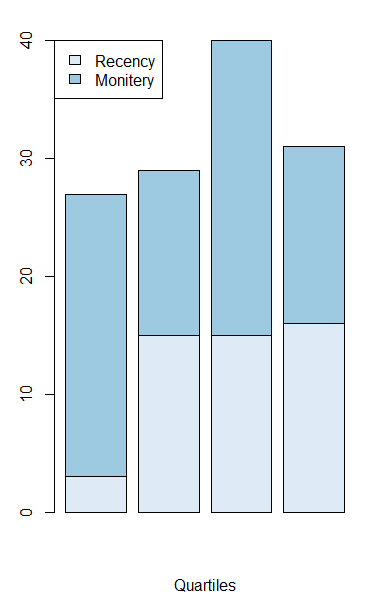
“TICKET LIST” dataset with 11 variables.

AVERAGE RECENCY(in days)

|  |  |  |  |
| --- | --- | --- | --- |
| 1ST QUARTILE | 2ND QUARTILE | 3RD QUARTILE | 4TH QUARTILE |
| 3 | 15 | 24 | 25 |

AVERAGE MONEY SPEND

|  |  |  |  |
| --- | --- | --- | --- |
| 1ST QUARTILE | 2ND QUARTILE | 3RD QUARTILE | 4TH QUARTILE |
| $15 | $16 | $14 | $15 |



“NEW CUSTOMER LIST” with 10 variables.

AVERAGE RECENCY(in days)

|  |  |  |  |
| --- | --- | --- | --- |
| 1st QUARTILE | 2ND QUARTILE | 3RD QUARTILE | 4TH QUARTILE |
| 3 | 11 | 19 | 26 |