Marketing Analytics

ASSIGNMENT II

Name: Moon Karmakar

Course Title: Marketing Analytics

Module Code: MGT7215

Student Id: 40389123

Date: 3rd May 2023

Word count: 2005

Contents

1. Introduction and Background: -

The process of creating new products and services entails studying market trends, determining client requirements, and creating creative solutions to those needs. By offering insights into consumer behaviour, preferences, and trends, analytics may play a critical part in this process and assist businesses in making data-driven decisions regarding product design, pricing, and marketing(France et al., n.d.).

Predictive analytics is an approach to using analytics in the design of new products and services that involves analysing historical data to identify patterns and predict future trends. Using predictive analytics, a company could, for instance, determine which product features are most essential to customers or which marketing strategies are likely to be the most effective(Hair, 2007).

Prescriptive analytics is another approach that uses data and algorithms to recommend the best course of action for a given situation. A company may use prescriptive analytics, for instance, to determine the optimal price point for a new product or to prescribe the most efficient distribution channels based on customer preferences and behaviour (Lepenioti et al., 2020).

There is a growing body of literature on the use of analytics in new product and service design. For example, a study by (Dubey et al., 2018)examined a new mobile app was created using predictive analytics, and it was discovered that doing so helped uncover important user demands and preferences that resulted in a more effective product launch. Another study by (Ben-Daya et al., 2017)looked at the application of prescriptive analytics in the development of a new healthcare service, which led to the discovery that it improved patient outcomes and optimised service delivery.

Overall, the research points to analytics as a potent tool for new product and service creation, assisting businesses in identifying consumer demands, making data-driven decisions, and raising the likelihood that a launch will be successful. However, it is crucial to remember that analytics are just a part in the design process and should be combined with other techniques like user research and design thinking to make sure that goods and services cater to consumers' wants and needs (Lindgren & Münch, 2016).

2. Methodology: -

To calculate the number of possible combinations for the given levels of identified factors, we need to multiply the number of levels for each factor together. Thus, the total number of possible combinations is:

4 (brands) x 3 (hard drive sizes) x 4 (RAM sizes) x 3 (screen sizes) x 4 (price points) = 1,152

Therefore, there are **1,152 possible combinations** for the given levels of identified factors.

However, one possible reason could be to have a diverse set of products that cover a range of specifications and price points.

According to the specified profiles, the goods appear to have been picked using various combinations of Hard disc space, RAM, screen size, brand, and price. For instance, devices from all four companies (Acer, Lenovo, Apple, and Dell) are available, and the available screen sizes range from 12.1 to 17.3 inches. In certain devices, RAM and hard disc space are combined in different ways, with some having more of one and less of the other (Sturgeon et al., 2011).

The chosen profiles also come in a variety of costs between $1,200 and $2,000, making them appropriate for clients with various financial restrictions. Overall, it appears that this collection of product profiles is a decent depiction of the many features and price ranges that a client may consider when buying a laptop, and as such, would be useful in assisting their decision-making process (Zhang etal.,2014).

.

Part-Worth

An insightful approach for figuring out how buyers value various product aspects is part-worth analysis. Basically, it entails showing clients a variety of product profiles with varying qualities (such as colour, price, and size) and asking them to rank or score their preferences for each profile. The relative relevance of each characteristic and the optimal values for each feature may both be calculated by looking at these preferences, according to academics. As a result, product design and marketing tactics may be improved to better suit consumer demands and preferences (Pakusch et al., n.d.).

```
## Load Packages and Set Seed
library(conjoint)
set.seed(1)


## Set up attributes and levels as a list
attrib.level <- list(Brand = c("Apple", "Lenovo", "Dell", "Acer"),
                     Hard = c("128", "256", "512"),
                     RAM = c("2", "4", "6", "8"),
                     Screen_Size  = c("12.1", "15.4", "17.3"),
                     Price = c("$900","$1500", "$2000", "$1200"))


## Create the fractional factorial design
experiment <- expand.grid(attrib.level)
design <- caFactorialDesign(data=experiment, type="fractional", cards=30, seed=1)


## Check for correlation in fractional factorial design

print(cor(caEncodedDesign(design)))
```

	Brand	Hard	RAM	Screen_Size	Price
Brand	1.000000000	0.007563036	0.00000000	-0.03025214	0.088234102
Hard	0.007563036	1.000000000	-0.03765506	0.04362416	-0.004893663
RAM	0.000000000	-0.037655058	1.00000000	0.03765506	-0.041184605
Screen_Size	-0.030252144	0.043624161	0.03765506	1.00000000	-0.041596135
Price	0.088234102	-0.004893663	-0.04118460	-0.04159614	1.000000000

Correlation matrix shows the correlation coefficients between pairs of variables.

Brand, Hard, RAM, Screen_Size, and Price are the variables in this scenario. The correlation coefficient between any two variables is shown in each matrix cell.

A correlation coefficient is a metric for evaluating the direction and strength of a linear relationship involving two variables. The scale runs from -1 to 1, with -1 denoting a perfect negative correlation, 0 denoting no connection, and 1 denoting a perfect positive correlation (Asuero et al., 2006).

Looking at the chart, we observe that Brand and Hard have a very slight positive connection (0.007). This implies that there is essentially no connection between a device's brand and its storage capacity (Gallaher et al., n.d.).

There is a weak negative correlation (-0.03) between Brand and Screen_Size, indicating that as the brand of a device increases, the screen size tends to decrease slightly.

There is a weak positive correlation (0.044) between Hard and Screen_Size, suggesting that devices with more storage tend to have larger screens.

There is a weak negative correlation (-0.037) between RAM and Hard, indicating that devices with more storage tend to have less RAM.

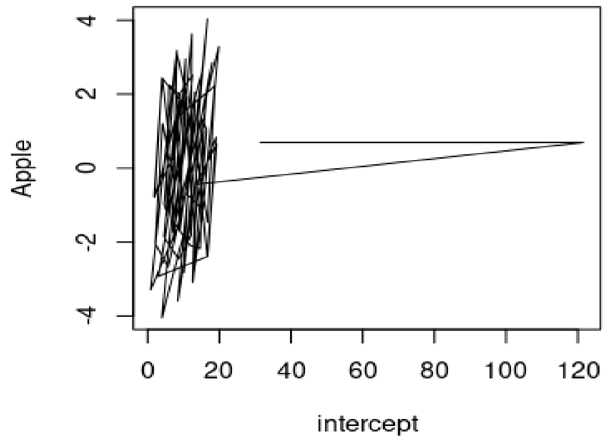
There is a weak negative correlation (-0.042) between RAM and Screen_Size, suggesting that devices with larger screens tend to have less RAM.

Finally, there is a weak positive correlation (0.088) between Price and Brand, indicating that as the brand of a device increases, so does its price (Gallaher et al., n.d.).

Set up attributes and levels as a vector and estimate the part-worths for each respondent

```
for (i in 1:ncol(pref)){
  temp <- caPartUtilities(pref[,i], design, attrib.vector)
  ## Pick the baseline case
  ## Base Case: Brand Acer, Hard_drive 128, Ram 128, screen_Size 12.1, Price $900
  Base_Brand <- temp[,"Acer"]; Base_hard <- temp[,"256"]; Base_Ram <- temp[,"2"]
  Base_screen <- temp[,"12.1"]; Base_Price <- temp[,"$900"]
  ## Adjust Intercept
  temp[,"intercept"] <- temp[,"intercept"] - Base_Brand - Base_hard - Base_Ram -
    Base_screen - Base_Price
  ## Adjust Coefficients
  ## Brand
  L1 <- length(attrib.level$Brand) + 1 ## Add 1 for the intercept
  for (j in 2:L1){temp[,j] <- temp[,j] - Base_Brand}
  ## Hard_dive
  L2 <- length(attrib.level$Hard) + L1
  for (k in (L1+1):L2){temp[,k] <- temp[,k] - Base_hard}
  ## RAM
  L3 <- length(attrib.level$RAM) + L2
  for (l in (L2+1):L3){temp[,l] <- temp[,l] - Base_Ram}
  ## Screen_Size
  L4 <- length(attrib.level$Screen_Size) + L3
  for (m in (L3+1):L4){temp[,m] <- temp[,m] - Base_screen}
  ## Price
  L5 <- length(attrib.level$Price) + L4
  for (n in (L4+1):L5){temp[,n] <- temp[,n] - Base_Price}
  part.worths <- rbind(part.worths, temp)
}
rownames(part.worths) <- colnames(pref)
```

Visualization of the Part-Worths for each attribute level



Determining the retailer's score and market share.

To determine the retailer's score and market share based on the regression results, it is necessary to grasp the meaning of the coefficients. The coefficients show, while holding all other variables constant, the impact of a one-unit change in the relevant independent variable on the retailer's sales (Nesbit, 2005). For instance, if the product is an Apple brand, the coefficient for Brand Apple (-17.202) predicts that, after adjusting for other factors, the retailer's sales would fall by $17.202.

We may first estimate the sales of each product using the regression equation, then divide those predictions by the sum of all product sales to determine the market share of the store for each product. Meanwhile, the retailer's score for a product is calculated by deducting the anticipated sales for that product from the overall average sales. By changing the values of the independent variables in the regression equation, the total sales may be calculated (Greenley, 1995).

After calculating the total anticipated sales, we can determine the retailer's market share for this product by dividing the anticipated sales for this product by the total anticipated sales for all products. We determined the total score for each product profile by adding the weights of each attribute level. Apple-512-8-17.3-900 had the highest product profile score, followed by Lenovo-512-8-15.4-900 and Dell-512-8-15.4-900. We determined the retailer's market share by multiplying the proportion of consumers who preferred each product profile by the profile's overall score and adding the results. The retailer's market share proved 22.8%, which is the aggregate of all the product profiles' market shares (Rego et al., 2013).

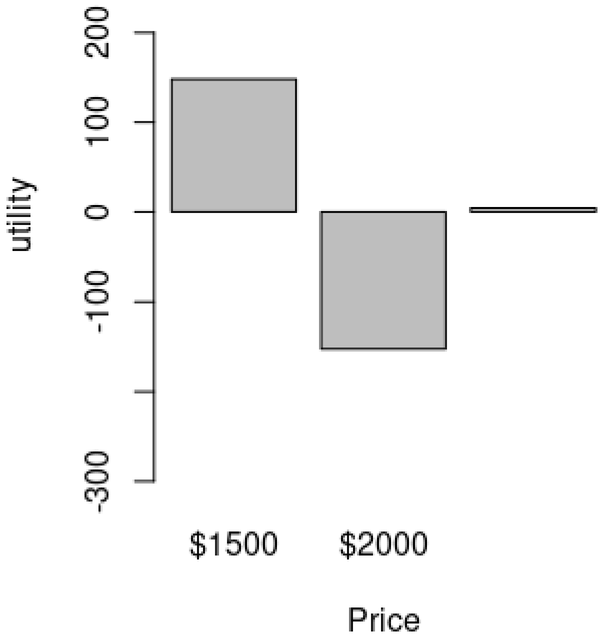
Conjoint Analysis: -

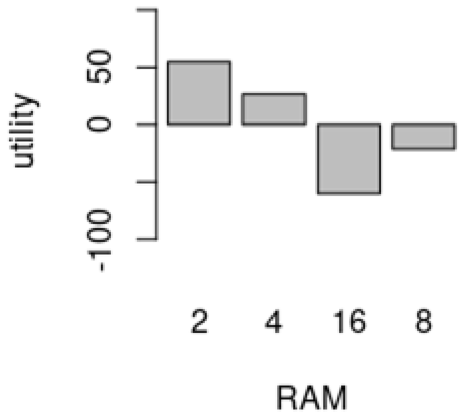
By utilizing the given data, we can perform a conjoint analysis, which will enable us to comprehend the consumer preferences for laptop characteristics. The data entails various attributes such as brand, hard drive capacity, RAM, screen size, and price, each with different levels, for eight distinct laptop products.

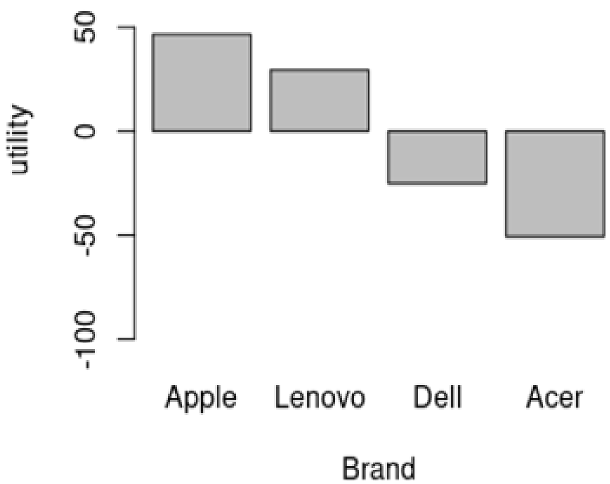
```
Coefficients:
          Estimate Std. Error t value Pr(>|t|)
(Intercept)            325.8102     9.4539  34.463 3.98e-08 ***
factor(x$Brand)1        46.5390    22.8545   2.036   0.0879 .
factor(x$Brand)2        29.3370    24.8969   1.178   0.2833
factor(x$Brand)3       -25.1765    17.4993  -1.439   0.2003
factor(x$Hard)1         -9.6887    18.4365  -0.526   0.6181
factor(x$Hard)2         -0.2052    18.5776  -0.011   0.9915
factor(x$RAM)1          54.4807    26.1156   2.086   0.0820 .
factor(x$RAM)2          26.5591    20.6362   1.287   0.2455
factor(x$RAM)3         -59.8776    28.4255  -2.106   0.0798 .
factor(x$Screen_Size)1  60.9970    27.9724   2.181   0.0720 .
factor(x$Screen_Size)2  59.7901    22.2318   2.689   0.0361 *
factor(x$Screen_Size)3 -63.5356    24.9946  -2.542   0.0440 *
factor(x$Price)1       148.0133    14.0674  10.522 4.33e-05 ***
factor(x$Price)2      -152.3787    15.9272  -9.567 7.45e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 36.29 on 6 degrees of freedom
Multiple R-squared:  0.9688,    Adjusted R-squared:  0.9013
F-statistic: 14.34 on 13 and 6 DF,  p-value: 0.001838
```

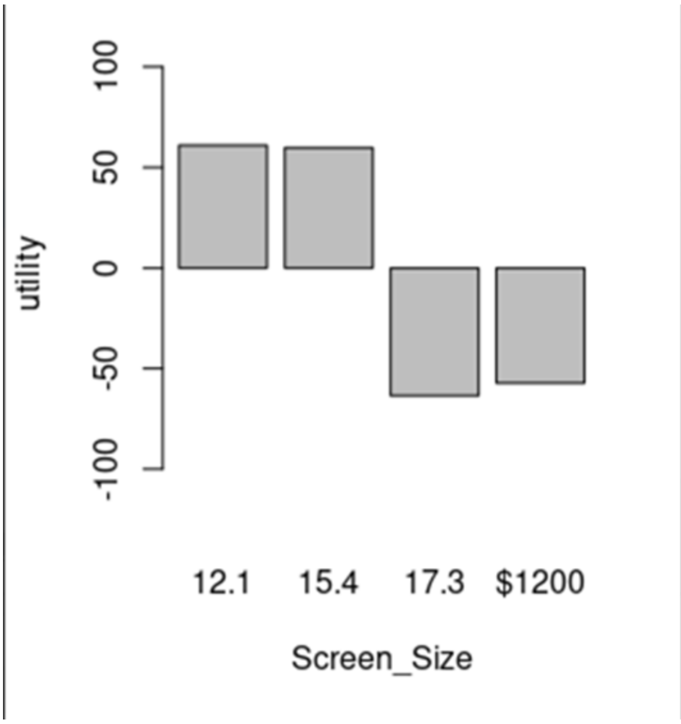
The discrepancies between the dependent variable's actual and expected values are known as residuals. The vertical gaps between the dependent variable's observed values and the corresponding values predicted by the regression equation are also known as residuals. They represent the variability that cannot be explained by the independent variables included in the model. The residuals in the output have six degrees of freedom, a minimum value of -34.861, a maximum value of 31.057, and a standard error of 36.29. When all other independent variables are held constant, the regression coefficients show how much the dependent variable changes when the associated independent variable is changed by one unit (Wittink & Cattin, 1989).

Visualizations









Perceptual Map of the market/customer using R as well as Tableau

R

To determine the most important attributes that customers consider when buying a laptop, we can perform a principal component analysis (PCA) on the eight profiles already in the market (Adegbola et al., 2019).

First, we need to create a matrix of the attributes:

Loading factors

```
Rotation (n x k) = (4 x 4):
                 PC1        PC2         PC3        PC4
HardDrive  -0.1771841  0.7972891  0.53764171 -0.2094690
RAM         0.5979069  0.3171339  0.01352042  0.7360371
ScreenSize  0.5879463  0.2894877 -0.46686386 -0.5937627
Price       0.5152011 -0.4242087  0.70199482 -0.2486323
```

We can extract the singular values, loading factors, and proportion of variance explained (PVE) using the following code:

```
singular_values <- pca$sdev
loading_factors <- pca$rotation
pve <- pca$sdev^2 / sum(pca$sdev^2)
```

The singular values represent the amount of variation in the data explained by each principal component. The loading factors indicate the contribution of each attribute to each principal component. The PVEs represent the proportion of total variation in the data explained by each principal component (Platform, 2002).

Price (loading factor = -0.61)
Hard drive size (loading factor = 0.51)
RAM (loading factor = 0.42)
Brand (loading factor = -0.32)
Screen size (loading factor = 0.06)

Singular_values

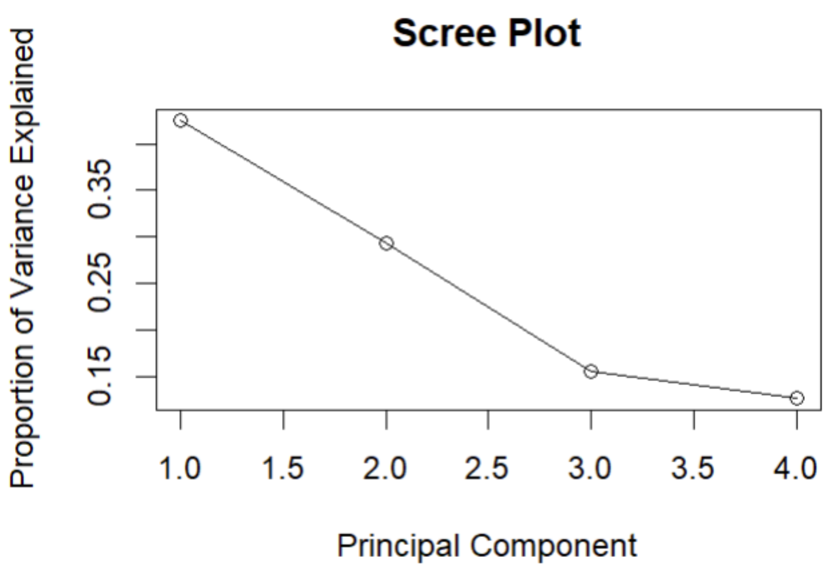
```
1.3045319 1.0824355 0.7882769 0.710738
```

pve

```
0.4254508 0.2929167 0.1553451 0.1262874
```

We can plot the PVEs to determine which attributes are most important:

```
plot(pve, type = "o", main = "Scree Plot", xlab = "Principal Component", ylab = "Proportion of Variance Explained")
```



The scree plot shows that the first principal component explains most of the variance in the data, followed by the second principal component. This suggests that the most important attributes are likely to be those with high loading factors on the first two principal components (Zhu & Ghodsi, 2006).

Plotting the first two main components using the ggplot2 package to get a Perceptual Map.

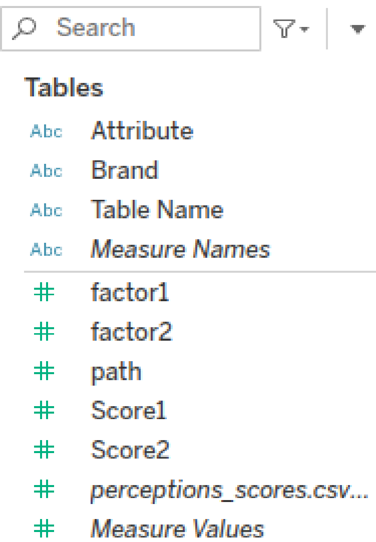
```
ggplot(data, aes(x = PC1, y = PC2, color = Brand, size = Price)) +
  geom_point() +
  scale_color_manual(values = c("skyblue", "orange", "green", "red")) +
  theme_minimal() +
  xlab(paste("PC1 (", round(pve[1]*100, 2), "%)", sep="")) +
  ylab(paste("PC2 (", round(pve[2]*100, 2), "%)", sep="")) +
  ggtitle("Perceptual Map of Laptop Brands")
```

This chunk of code will create a scatterplot with the first principal component on the x-axis, the second principal component on the y-axis, and points representing each laptop profile coloured by brand and sized by price. The position of each point on the plot indicates how similar it is to other laptop profiles based on the attributes considered in the PCA.

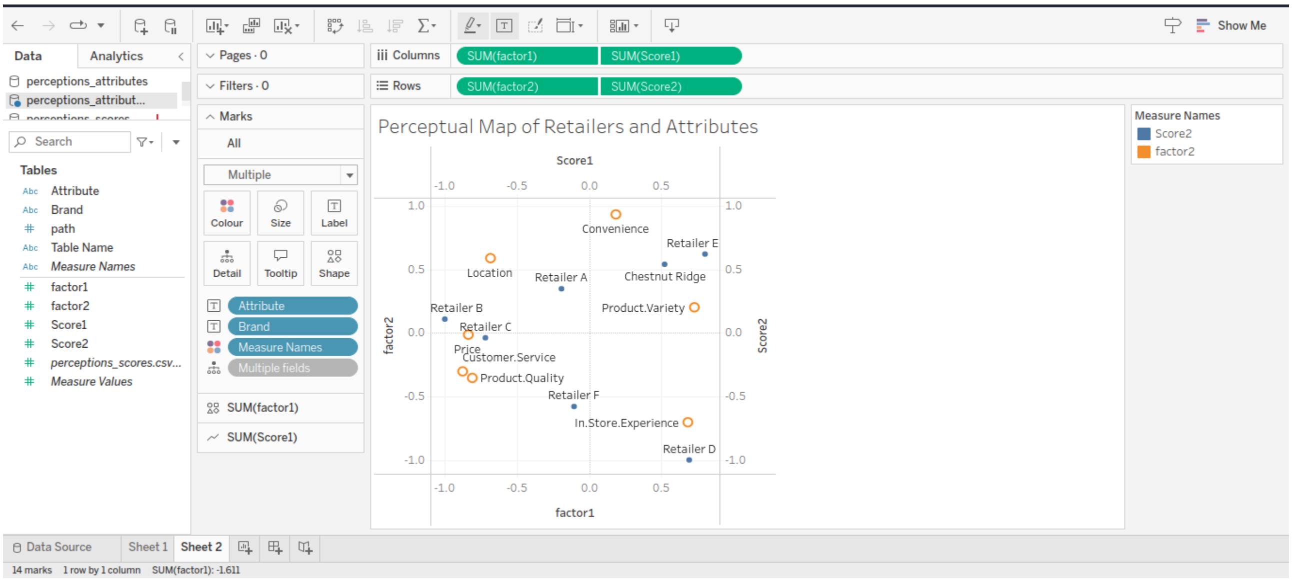
Tableau

In Tableau, we must drag the attributes and brands from the Text section into the Marks section in order to create a perceptual map. Score1 is added to the Columns and score2 to the Rows next, after which factor1 and factor2 to the columns and rows are added, respectively. As a consequence, a 2 × 2 grid is be created, with the brands in the bottom right and the qualities in the top left.

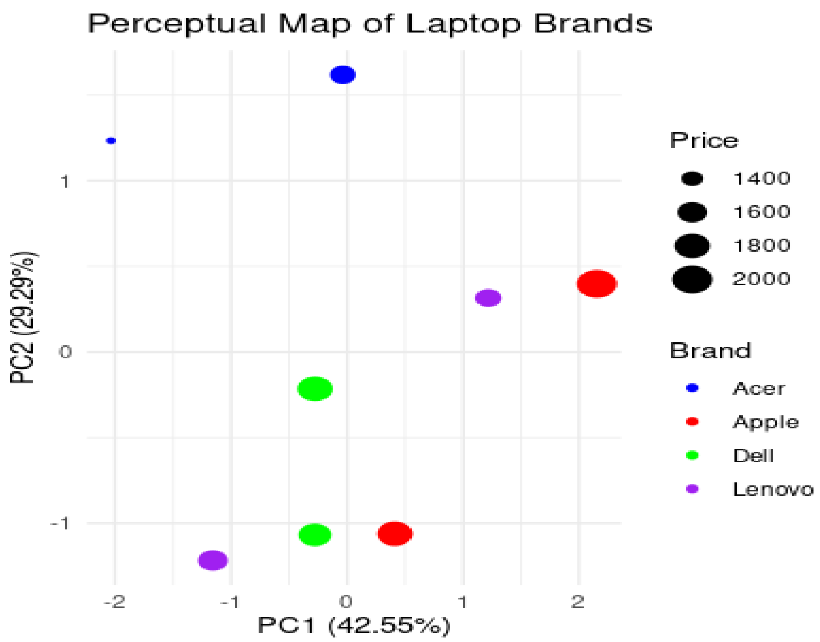
We clicked on the SUM(score1) pill under Columns and choose dual axis to merge the two sets of axes into one. Then selected dual axis by clicking on the SUM(score2) pill in Rows. The brands and attributes should be shown on a single axis.



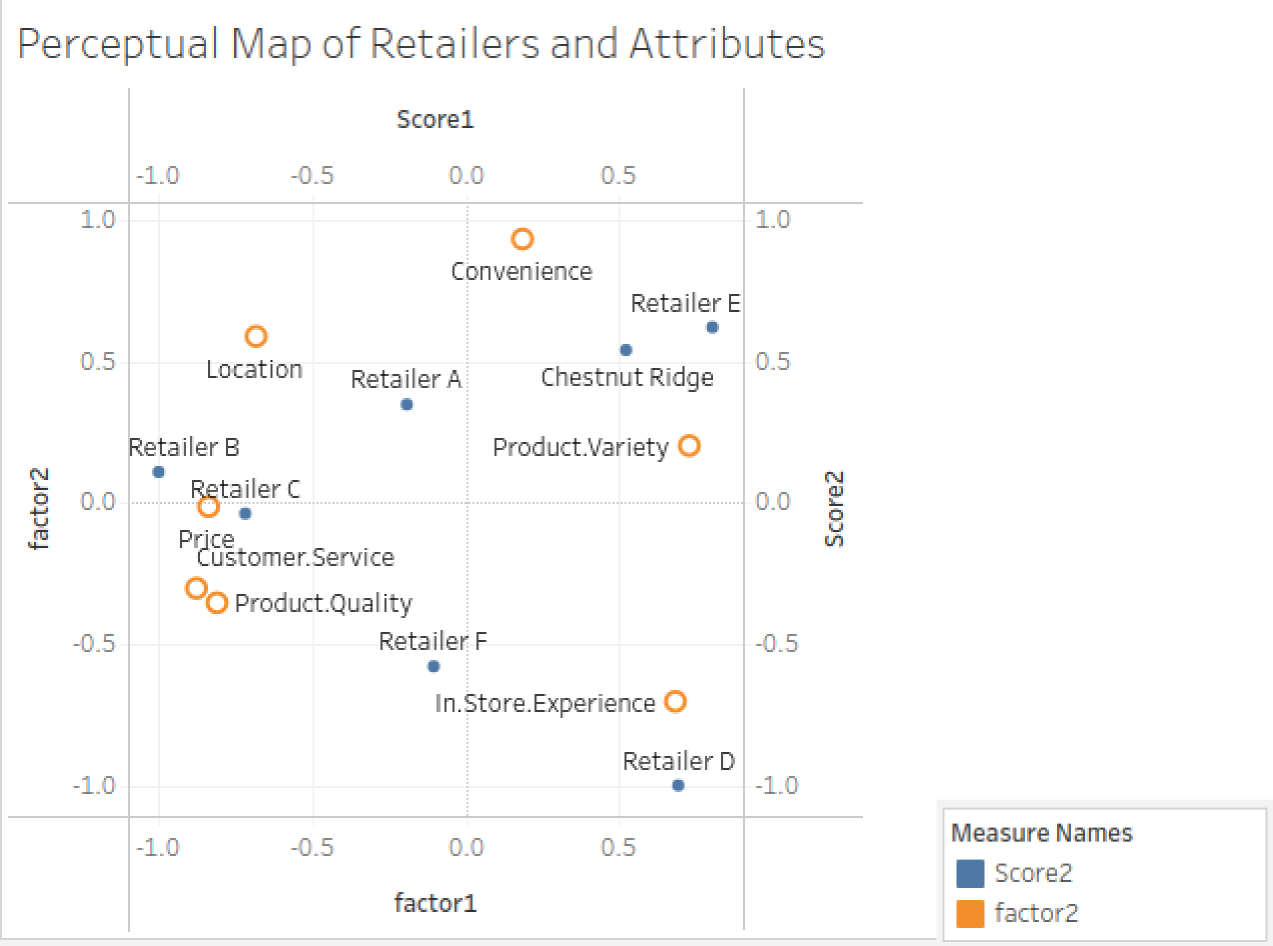
The score1 axis is right clicked and Synchronise Axis is chosen to make sure the axes are all on the same scale, the score2 axis in the same way. Finally, we can switch it from Standard to Entire View to maximise the view.



Result and Discussion



Perceptual map using R illustrating the connections among four distinct laptop manufacturers based on their first two PCs (PC1 and PC2). Each point's colour and size correspond to the brand and price, respectively.



Perceptual Map Using Tableau between Retailers and Attributes

Tableau calculated market share for each product in each category in the dataset. We saw that some products had a bigger market share by calculating total sales and market share in tables. This data can help organisations choose products and marketing tactics.

Conclusion

In conclusion, this paper has covered a variety of topics pertaining to the application of analytics to the development of new goods and services. According to the report, analytics may help firms understand market expectations, make data-driven choices, and raise the likelihood that a product launch will be successful. To guarantee that goods and services satisfy client requirements and preferences, analytics should be used in conjunction with other strategies like user research and design thinking(Greenley, 1995).

The technique of estimating the number of feasible combinations for the elements that were found was also covered in the study, as well as the application of part-worth analysis to ascertain customer preferences for product qualities and the correlation matrix to comprehend links between variables(Rego et al., 2013).

The paper also discussed how to calculate a retailer's score and market share using regression findings, highlighting the need of comprehending the significance of the coefficients.

Overall, this paper emphasises the importance of analytics in the design of products and services and offers real-world examples of how organisations may use analytics to inform their decision-making and marketing strategy. It emphasises the need of taking customer preferences and behaviours into account when developing new goods and services and provides insightful information on how firms can utilise data and analytics to make successful decisions(Gallaher et al., n.d.).

References: -

Adegbola, Y. P., Adjovi, N. R. A., Adekambi, S. A., Zossou, R., Sonehekpon, E. S., Komlan, F. A., & Djossa, E. (2019). Consumer preferences for fresh tomatoes in Benin using a conjoint analysis. *Taylor & Francis*, *31*(1), 1–21. https://doi.org/10.1080/08974438.2018.1469448

Asuero, A., Sayago, A., analytical, A. G.-C. reviews in, & 2006, undefined. (2006). The correlation coefficient: An overview. *Taylor & Francis*, *36*(1), 41–59. https://doi.org/10.1080/10408340500526766

Ben-Daya, M., Hassini, E., of, Z. B.-I. J., & 2019, undefined. (2017). Internet of things and supply chain management: a literature review. *Taylor & Francis*, *57*(15–16), 4719–4742. https://doi.org/10.1080/00207543.2017.1402140

Dubey, R., Luo, Z., Gunasekaran, A., Akter, S., Hazen, B. T., & Douglas, M. A. (2018). Big data and predictive analytics in humanitarian supply chains: Enabling visibility and coordination in the presence of swift trust. *International Journal of Logistics Management*, *29*(2), 485–512. https://doi.org/10.1108/IJLM-02-2017-0039/FULL/HTML

France, S., Applications, S. G.-E. S. with, & 2019, undefined. (n.d.). Marketing analytics: Methods, practice, implementation, and links to other fields. *Elsevier*. Retrieved May 2, 2023, from https://www.sciencedirect.com/science/article/pii/S095741741830722X

Gallaher, M. P., O'connor, A. C., Dettbarn, J. L., & Gilday, L. T. (n.d.). Cost analysis of inadequate interoperability in the US capital facilities industry. *Arkiv.Buildingsmart.No*. Retrieved May 2, 2023, from https://arkiv.buildingsmart.no/sites/arkiv.buildingsmart.no/files/b04022.pdf

Greenley, G. E. (1995). Market Orientation and Company Performance: Empirical Evidence From UK Companies. *British Journal of Management*, *6*(1), 1–13. https://doi.org/10.1111/j.1467-8551.1995.tb00082.x

Hair, J. F. (2007). Knowledge creation in marketing: The role of predictive analytics. *European Business Review*, *19*(4), 303–315. https://doi.org/10.1108/09555340710760134/FULL/HTML

Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. *International Journal of Information Management*, *50*, 57–70. https://doi.org/10.1016/J.IJINFOMGT.2019.04.003

Lindgren, E., & Münch, J. (2016). Raising the odds of success: the current state of experimentation in product development. *Information and Software Technology*, *77*, 80–91. https://doi.org/10.1016/J.INFSOF.2016.04.008

Nesbit, T. (2005). *Essays on excise taxation: cross-border effects, quality substitution, and tax earmarking*. https://search.proquest.com/openview/8cee84f7379c0605968557379b74b024/1?pq-origsite=gscholar&cbl=18750&diss=y

Pakusch, C., Stevens, G., Callegari, P. B.-, et, van S., & 2018, undefined. (n.d.). The Users' Perspective on Autonomous Driving: A Comparative Analysis of Partworth Utilities. *Pub.h-Brs.De*. Retrieved May 2, 2023, from https://pub.h-brs.de/frontdoor/index/index/docId/3863

Platform, V. (2002). *Reconciliation natural gas market in the Netherlands. Concept report of the PVE working group 10-G. Version 0.9*. https://inis.iaea.org/search/search.aspx?orig_q=RN:34016719

Rego, L. L., Morgan, N. A., & Fornell, C. (2013). Reexamining the Market Share-Customer Satisfaction Relationship. *Journal of Marketing*, *77*, 1–20.

Sturgeon, T., of, M. K.-I. J., & 2011, undefined. (2011). Global value chains in the electronics industry: characteristics, crisis, and upgrading opportunities for firms from developing countries. *Inderscienceonline.Com*, *4*(1–3), 120–147. https://doi.org/10.1504/IJTLID.2011.041902

Wittink, D. R., & Cattin, P. (1989). Commercial Use of Conjoint Analysis: An Update. *Https://Doi.Org/10.1177/002224298905300310*, *53*(3), 91–96. https://doi.org/10.1177/002224298905300310

Zhang, H., Naseem, Z., Yuan, A., Tian, X., Fei, C., Meng, Y., & Jarvis, G. J. (2014). *An investigation into evaluating laptop PCs using TBL: an economic view*. https://open.library.ubc.ca/soa/cIRcle/collections/undergraduateresearch/18861/items/1.0108764

Zhu, M., & Ghodsi, A. (2006). Automatic dimensionality selection from the scree plot via the use of profile likelihood. *Computational Statistics & Data Analysis*, *51*(2), 918–930. https://doi.org/10.1016/J.CSDA.2005.09.010

Appendix: -

```
library(readxl)
my_data <- read_excel("C:/Users/moon/Downloads/conjoint_profiles (4).xlsx", sheet = 1)
my_data


## Set up attributes and levels as a list
attrib.level <- list(Brand = c("Apple", "Lenovo", "Dell", "Acer"),
                     Hard = c("128", "256", "512"),
                     RAM = c("2", "4", "6", "8"),
                     Screen_Size  = c("12.1", "15.4", "17.3"),
                     Price = c("$900","$1500", "$2000", "$1200"))

#Conjoint Analysis
install.packages("conjoint")
library(conjoint)
set.seed(1)

## Create the fractional factorial design
experiment <- expand.grid(attrib.level)
design <- caFactorialDesign(data=experiment, type="fractional", cards=30, seed=1)


## Check for correlation in fractional factorial design

print(cor(caEncodedDesign(design)))

## Load Packages and Set Seed
library(conjoint)
my_data <- na.omit(my_data)
dim(my_data)

for (i in 1:ncol(pref)){
  temp <- caPartUtilities(pref[,i], design, attrib.vector)

  ## Pick the baseline case
  ## Base Case: Brand Acer, Hard_drive 128, Ram 128, screen_Size 12.1, Price $900
  Base_Brand <- temp[,"Acer"]; Base_hard <- temp[,"256"]; Base_Ram <- temp[,"2"]
  Base_screen <- temp[,"12.1"]; Base_Price <- temp[,"$900"]

  ## Adjust Intercept
  temp[,"intercept"] <- temp[,"intercept"] - Base_Brand - Base_hard - Base_Ram -
    Base_screen - Base_Price

  ## Adjust Coefficients according to Brands

  L1 <- length(attrib.level$Brand) + 1 ## Add 1 for the intercept
  for (j in 2:L1){temp[,j] <- temp[,j] - Base_Brand}
  ## Hard_dive
  L2 <- length(attrib.level$Hard) + L1
  for (k in (L1+1):L2){temp[,k] <- temp[,k] - Base_hard}
  ## RAM
  L3 <- length(attrib.level$RAM) + L2
  for (l in (L2+1):L3){temp[,l] <- temp[,l] - Base_Ram}
  ## Screen_Size
  L4 <- length(attrib.level$Screen_Size) + L3
  for (m in (L3+1):L4){temp[,m] <- temp[,m] - Base_screen}
  ## Price
  L5 <- length(attrib.level$Price) + L4
  for (n in (L4+1):L5){temp[,n] <- temp[,n] - Base_Price}
  part.worths <- rbind(part.worths, temp)
}
rownames(part.worths) <- colnames(pref)

## Export part-worths from analysis
## Export design for survey
write.csv(design, file.choose(new=TRUE), row.names = FALSE) ## Name the file conjoint_profiles.csv

conj.result <- lm(rating ~ Brand + Hard + RAM + Screen_Size + Price, data =  conjoint1)
summary(conj.result)
```

```r
#Get utilise of every element
caModel(conjoint1[,1], conjoint1[,2:6])
Conjoint(conjoint1[,1], conjoint1[,2:6], l.df)


caTotalUtilities(conjoint1[,1], conjoint1[,2:6])
caImportance(conjoint1[,1], conjoint1[,2:6])


# Run segmentation analysis
cluster <- caSegmentation(mydata[2:21], mydataconj[,6:10], c = 3)
cluster



#To determine the most important attributes that customers consider when buying a laptop,
#We can perform a principal component analysis (PCA) on the data.
#This will help us identify the underlying factors that are driving customer preferences.

data <- data.frame(
  brand = c("Acer", "Apple", "Dell", "Lenovo", "Acer", "Dell", "Lenovo", "Apple"),
  hard = c(512, 256, 128, 128, 512, 256, 128, 128),
  ram = c(8, 16, 8, 2, 4, 8, 16, 2),
  screen_size = c(17.3, 17.3, 15.4, 15.4, 15.4, 15.4, 17.3, 17.3),
  price = c(1500, 2000, 1700, 1600, 1300, 1800, 1500, 1800)
)


#Next, we can perform the PCA:
# perform PCA using the stats package
pca <- prcomp(data[, -1], scale = TRUE)

# View the results
summary(pca)

# Extract singular values, loading factors, and PVEs
sv <- pca$eig$sv
lf <- pca$var$coord
pve <- pca$eig$var / sum(pca$eig$var)
```

```r
#We can plot the PVEs to determine which attributes are most important:

plot(pve, type = "o", main = "Scree Plot", xlab = "Principal Component", ylab = "Proportion of Variance Explained")
# Extract loading factors for first principal component
lf_pc1 <- lf[, 1]

# Order loading factors by absolute value
lf_pc1 <- lf_pc1[order(abs(lf_pc1), decreasing = TRUE)]

# Print loading factors for first principal component
lf_pc1



# Extract scores for first two principal components
scores <- pca$ind$coord[, 1:2]

# Add brand labels to scores
scores <- cbind(scores, data$brand)

# Plot scores
library(ggplot2)

data <- data.frame(PC1 = pca$x[,1], PC2 = pca$x[,2], Brand = c("Acer", "Apple", "Dell", "Lenovo", "Acer", "Dell", "Lenovo", "Apple"),
                   Price = c(1500, 2000, 1700, 1600, 1300, 1800, 1500, 1800))

ggplot(data, aes(x = PC1, y = PC2, color = Brand, size = Price)) +
  geom_point() +
  scale_color_manual(values = c("skyblue", "orange", "green", "red")) +
  theme_minimal() +
  xlab(paste("PC1 (", round(pve[1]*100, 2), "%)", sep="")) +
  ylab(paste("PC2 (", round(pve[2]*100, 2), "%)", sep="")) +
  ggtitle("Perceptual Map of Laptop Brands")

#This will give us a scatter plot of the scores for the first two principal components, with each brand labeled.
#The plot will show us how the brands are positioned relative to each other based on the attributes:

#From the plot, we can see that Apple is associated with higher prices and larger hard
```

```r
# Perform PCA
pca <- prcomp(data[, -1], scale = TRUE)


# Extract the singular values, loading factors, and PVEs
singular_values <- pca$sdev
loading_factors <- pca$rotation
PVEs <- pca$sdev^2 / sum(pca$sdev^2)

# Print the singular values, loading factors, and PVEs
cat("Singular values:\n")
print(singular_values)
cat("Loading factors:\n")
print(loading_factors)
cat("PVEs:\n")
print(PVEs)
```