INTRODUCTION

The purpose of the of this report is elementing statistics in the company "Amazing Sports Australia Ltd (ASAL)". An online e-store. ASAL specializes in selling a wide range of branded and non-branded sports products, which are broadly categorized as (i) Equipment, (ii) Apparel, and (iii) Footwear. The company has recently launched a shopping mobile app and is concerned about its effectiveness in increasing sales and promoting its products. The management team aims to understand customer spending patterns and behavior with the ultimate goal of optimizing the app usage and enhancing sales by means of **supported by regression and visual representation**, with recommendations and solutions to be offered to the company.

SUMMARY STATISTIC

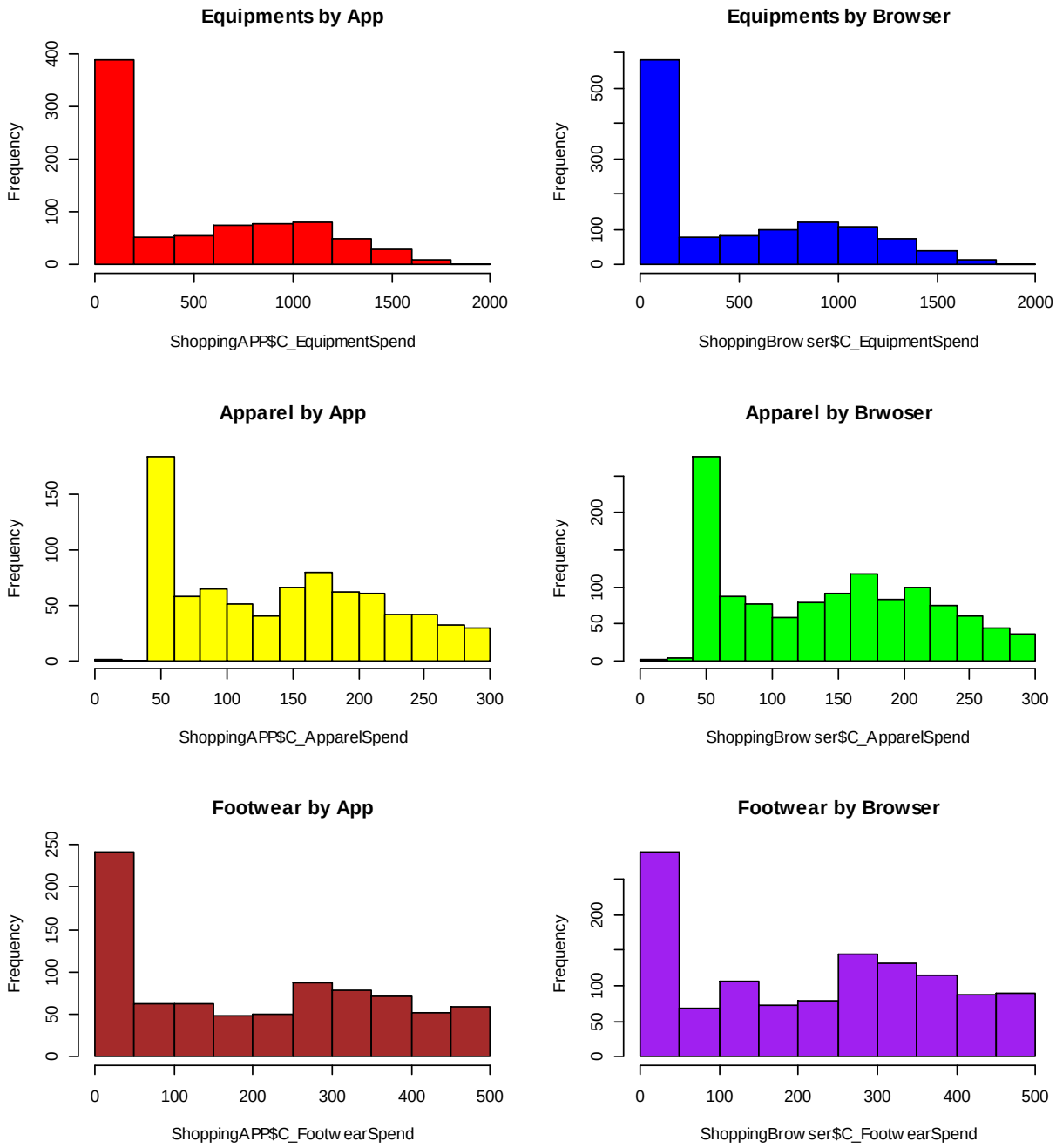
In summary statistics can be seen using descriptive statistics that function to see patterns and behavior from the data used, in this case data obtained from the company "Amazing Sports Australia Ltd (ASAL)" for analysis used 4 selected variables, namely, Shooping Cart, Equipment Spend, Apparel Spend, and Footwear Spend. The first thing to do is look at the following descriptive statistics:

Table 1. Deskriptive Statistic

```
summary(ShoppingAPP$C_EquipmentSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0    40.0   300.0   494.6   950.0  2000.0
summary(ShoppingBrowser$C_EquipmentSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0    40.0   250.0   481.9   900.0  2000.0
summary(ShoppingAPP$C_ApparelSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.00   67.25  143.00  142.31  201.75  299.00
summary(ShoppingBrowser$C_ApparelSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0    63.0   143.0   142.3   204.0   300.0
summary(ShoppingAPP$C_FootwearSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.0    40.0   192.0   200.3   331.0   500.0
summary(ShoppingBrowser$C_FootwearSpend)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.00   57.25  238.50  220.24  349.00  500.00
```

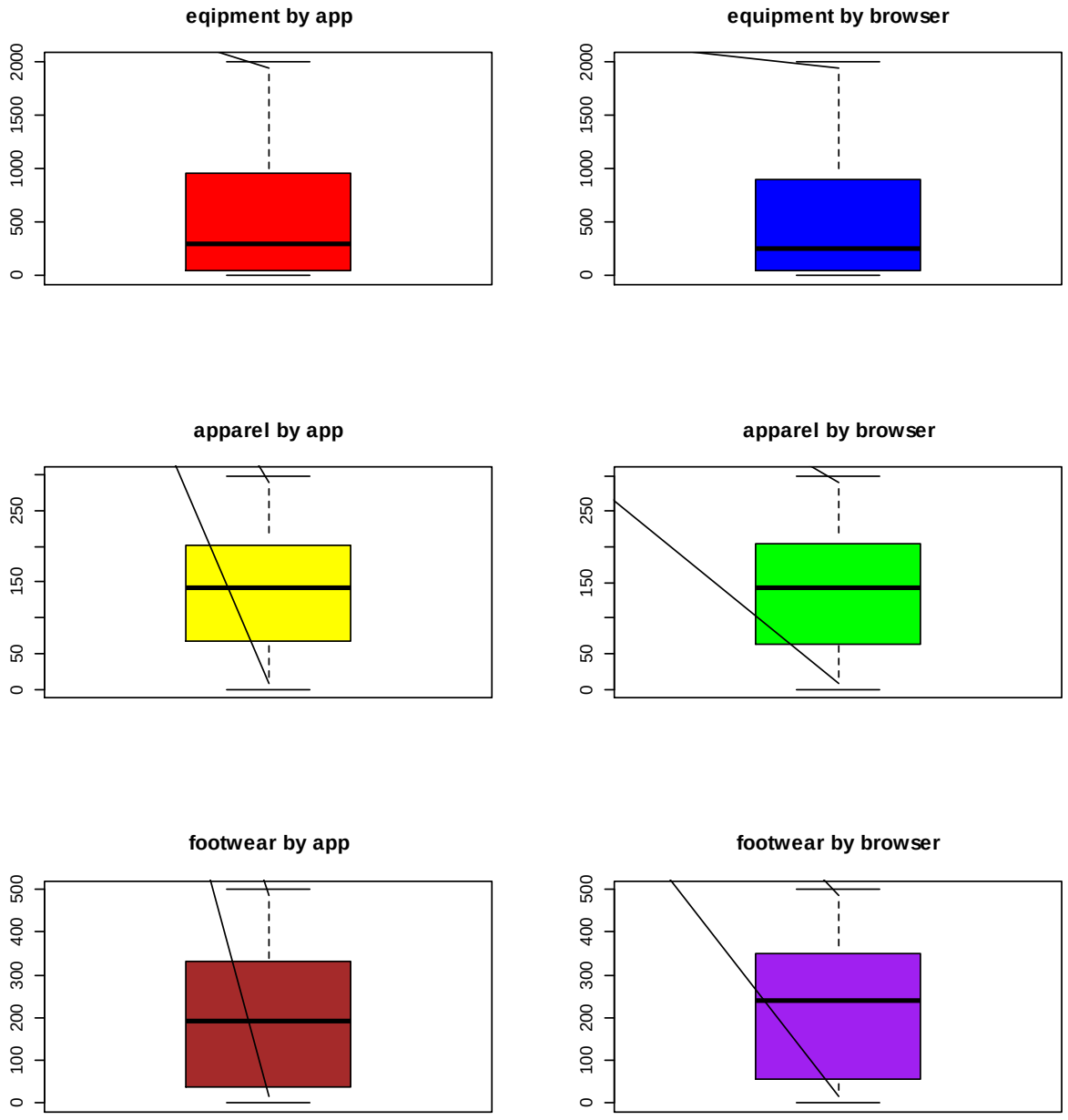
From Table 1, it can be seen how the distribution of data from the four variables used, such as median, Mean, Maximum, minimum, Quartile 1, and Quartil 3 from the selected data. More details can be seen in Figure 1 below.

Ficture 1. Histogram



From figure 1, a distribution pattern is obtained from the number of shopping using applications and browsers to buy Equipment, Apparel, to Footweaar. Look at the red and blue images, a comparison of Equipment purchases using the App and Browser concludes that Equipment purchases using the two methods are not too significantly different, it can be seen that the average purchase of Equipment using the App is 494.6 more than the Browser is 481.9 with both medians of 300 and 250 respectively, this indicates that the distribution of equipment purchases tends to be centered on the left Because the mean value is more than the median value. In addition, to prove our assumptions we can also use boxplot as a medium, where by using box plot displays the five-number summary of a data set (minimum = Q1-1.5 *IQR, first quartile Q1, median, third quartile Q3 and maximum = Q3 + 1.5 *IQR). In a box plot, we draw a box from the first quartile Q1 to the third quartile Q3. A yellow line goes through the box at the median. Green circles represent outliers. The shopping by App and Browser box plot does not indicate that equipment purchases with both are not normally distributed and tend to center to the left or bottom of the median, this is evidenced by the whisker towards 0 tends to be short and the line of the whisker is in the form of a dashed abosulut line and both have no outliers. Furthermore, for Apparel purchases using App and Browser tend to be normally distributed with the median located in the middle of the box in Apparel purchases using both methods also have no outliers due to the tendency of centralized data in the middle and the distribution of data is located around the median and mean of Apparel data. Finally, the comparison of Footwear purchase data with App and Browser, where the data tends to follow the pattern of Equipment purchases but the median is located in the middle of the box and the distribution of data tends to follow the mean and median so that Footwear purchase data has no outliers. So it can be concluded that the average purchase of Apparel and Footwear uses a browser more than purchases using the App, this can happen maybe due to the lack of optimization, discounts, tenure, and referrals using the application so that buyers think using the application and browser there is no significant difference so that purchases using a browser are more widely used. More details can be seen in Figure 2 Box Plot below.

Picture 2. Box Plot



Furthermore, by modeling Shopping Cart as a response variable and predictor variable (Equipment spend, apparel spend, and Footwear spend) to see the relationship of these variables so that the model is obtained as follows:

Figure 3. Output R (Multiple Regression)

```
Call:
lm(formula = model)

Residuals:
    Min      1Q  Median      3Q     Max
-19.456  -7.656  -0.573   5.606  33.768

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 32.9022921  0.8406047  39.141   <2e-16 ***
ShoppingAPP$C_EquipmentSpend  0.0156326  0.0006449  24.241   <2e-16 ***
ShoppingAPP$C_ApparelSpend   -0.0049699  0.0044094  -1.127    0.260
ShoppingAPP$C_FootwearSpend  -0.0007915  0.0020271  -0.390    0.696
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.166 on 810 degrees of freedom
Multiple R-squared:  0.4301,    Adjusted R-squared:  0.4279
F-statistic: 203.7 on 3 and 810 DF,  p-value: < 2.2e-16
```

Based on Figure 3 above, the relationship between purchases is obtained Equipment (x_1) Apparel (x_2), and Footwear (x_3) by using or against App (y) The following model is obtained:

$$y = 32.90 + 0.0156x_1 - 0.0049x_2 - 0.0008x_3$$

The model above indicates that the relationship of the Equipment variable is positive, which means that if the purchase of equipment is purchased using the App, then the apparel and footwear variables are both negative, it indicates that the purchase of both using the application is not much, most use the browser, it is in accordance with our statement above that purchases using applications are widely used when buying equipment, while the purchase of Apparel and Footwear uses a lot of browsers.

Furthermore, to see the effect of predictor variables on response variables can be seen using the value of multiple-r squared, multiple r sqare on multiple linear-regression values of 0.4301 which means that the equipment, apparel, and footwear variables affect the response variable simultaneously by 43.01%. The remaining 56,99% influenced by variables from outside the study.

SOLUTIONS

This time the task uses the R studio application to analyze the sales results of a sports company "ASAL (Amazing Sports Australia Ltd)," ASAL is a company that sells various sports equipment. ASAL wants to see the performance of purchases using the App and Browser and how the results of the increase in sales and promotions provided. The results of the analysis are different, namely the purchase of equipment is more widely used using the App, while for the purchase of Apparel and footwear using more browsers, this can be proven by the findings using histograms and boxplots reinforced again findings using statistical models, namely multiple linear regression with the following models:

$$y = 32.90 + 0.0156x_1 - 0.0049x_2 - 0.0008x_3$$

With Equipment (x_1) Apparel (x_2), and Footwear (x_3) by using App (y). Furthermore, to see the effect of predictor variables on response variables can be seen using the value of multiple-r squared, multiple r-square on multiple linear-regression values of 0.4301 yang berarti variable equipment, apparel, and footwear affect response variables simultaneously by 43.01%. The remaining 56.99% was influenced by variables from outside the study.

Refferences

Alpar. R. (2003) Uygulamalı çok değişkenli istatistiksel yöntemlere giriş 1 (ikinci baskı). Ankara: Nobel Yayınları.

Purnamasari, I., & Hayati, M.N., 2018. Analisis Deskriptif Pada Faktor-Faktor Yang Mempengaruhi Minat Siswa Untuk Melanjutkan Pendidikan Ke Tingkat Perguruan Tinggi. Jurnal.unimus.ac.id. [6 :2]. Doi: https://doi.org/10.26714/jsunimus.v6i2.4315.g3994

Wilson, C., 2016. Mean vs median: What do they mean and when do you use them?, [online] Michigan State University. Available at: <https://www.canr.msu.edu/news/mean_vs.median what do they mean and when do you use them> [Accessed 10 March 20221]

Appendix

```
library(moments)

library(dplyr)

library(psych)

#read data

data<-read.csv("https://raw.githubusercontent.com/dat-analytics/data_assess_1_t2_2023/main/z5397591_z5397591-Assessment1Data.csv", encoding="UTF-8")

data

#-------------histogram--------------------#

par(mfrow=c(3,2))

#------ShoppingCart-------------------#

ShoppingAPP = subset(data, C_ShoppingCart =="App")

ShoppingBrowser = subset(data, C_ShoppingCart =="Browser")

hist(ShoppingAPP$C_EquipmentSpend,col = "red",main = "Equipments by App")

hist(ShoppingBrowser$C_EquipmentSpend,col = "blue",main = "Equipments by Browser")

hist(ShoppingAPP$C_ApparelSpend,col = "yellow",main = "Apparel by App")

hist(ShoppingBrowser$C_ApparelSpend,col = "green",main = "Apparel by Brwoser")

hist(ShoppingAPP$C_FootwearSpend,col = "brown",main = "Footwear by App")

hist(ShoppingBrowser$C_FootwearSpend,col = "purple",main = "Footwear by Browser")


#deskriptif statistik#

summary(ShoppingAPP$C_EquipmentSpend)

summary(ShoppingBrowser$C_EquipmentSpend)

summary(ShoppingAPP$C_ApparelSpend)

summary(ShoppingBrowser$C_ApparelSpend)

summary(ShoppingAPP$C_FootwearSpend)

summary(ShoppingBrowser$C_FootwearSpend)
```

```
#boxplot
boxplot(ShoppingAPP$C_EquipmentSpend, col = "red",main="eqipment by app")
boxplot.stats(ShoppingAPP$C_EquipmentSpend)$out
boxplot(ShoppingBrowser$C_EquipmentSpend,col="blue",main="equipment by browser")
boxplot.stats(ShoppingBrowser$C_EquipmentSpend)$out
boxplot(ShoppingAPP$C_ApparelSpend,col = "yellow", main="apparel by app")
boxplot.stats(ShoppingAPP$C_ApparelSpend)$out
boxplot(ShoppingBrowser$C_ApparelSpend,col = "green", main="apparel by browser")
boxplot.stats(ShoppingBrowser$C_ApparelSpend)$out
boxplot(ShoppingAPP$C_FootwearSpend,col = "brown", main="footwear by app")
boxplot.stats(ShoppingAPP$C_FootwearSpend)$out
boxplot(ShoppingBrowser$C_FootwearSpend,col = "purple", main="footwear by browser")
boxplot.stats(ShoppingBrowser$C_FootwearSpend)$out
lot(perthhouseprice,las=1,col="yellow",ylim=c(0,2000),names.arg=colnames,xlab = "Number of Bed Rooms", ylab = "Mean Price thousands $",        main="Perth's average housing price by Bed Rooms")


#regression
model<-
lm(ShoppingAPP$C_Age~ShoppingAPP$C_EquipmentSpend+ShoppingAPP$C_ApparelSpend
+ShoppingAPP$C_FootwearSpend, data = data)
fit<-lm(model)
summary(fit)#summary result
scoef(fit) #vie coefficient estimate
y=fitted(fit) #view fitted values
y
x=residuals(fit) #view ressiduals
x
plot(x, main = "qq residual",type = c("qq"))
#plot
```

```
plot(x,y)
plot(fitted(fit))
plot.residuals(residuals(fit))
```