Leeds University Business School

UNIVERSITY OF LEEDS

Assessed Coursework Coversheet

For use with *individual* assessed work

Please Note:

Your declared word count must be accurate, and should not mislead. Making a fraudulent statement concerning the work submitted for assessment could be considered academic malpractice and investigated as such. If the amount of work submitted is higher than that specified by the word limit or that declared on your word count, this may be reflected in the mark awarded and noted through individual feedback given to you.

It is not acceptable to present matters of substance, which should be included in the main body of the text, in the appendices ("appendix abuse"). It is not acceptable to attempt to hide words in graphs and diagrams; only text which is strictly necessary should be included in graphs and diagrams.

By submitting an assignment you confirm you have read and understood the University of Leeds **Declaration of Academic Integrity** (http://www.leeds.ac.uk/secretariat/documents/academic_integrity.pdf).

Part 1
1st Task

AutonomousShipment has decided to run a test trial for a month to test one of the four prototype robots that they developed to have a successful delivery of products to maximum number of customers. The use of autonomous drones will ensure faster delivery of the products meanwhile this will also help the company to benefit from cost reduction. This report involves the crucial decisions that are to be made to conduct a successful trial in the city of Leeds. To accomplish the goal of this trial it is important to identify which robot is most suitable and how many robots need to be allocated to different stores like grocery, clothing and sports equipment stores.

The first task is to pick the most suitable robot for the trial depending on the five criteria identified by the management team to which the importance was assigned during the team meeting in July 2023.
The data set shown below provides information about the robots and the corresponding characteristics which are:

Robot_Prototype	Archer	Bowler	Corner	Deviant
Carrying Capacity	45	50	60	40
Battery Size	18	18	12	24
Average Speed	6	4	4	10
Cost Per Unit	5210	6250	4500	7100
Reliability	22	24	24	32

Figure 1

- Robot Prototype that contains the four prototypes of robots namely Archer, Bowler, Corner, Deviant.
- Carrying Capacity provides information about the maximum load each robot can carry in litres.
- Battery Size indicates the battery capacity of the robots in hours.
- Average Speed contains information pertaining to the average speed of the robot in km/h.
- Cost Per Unit gives the details about the cost per unit of each robot in GBP.
- Reliability provides information of average time in hours between each breakdown of the robot.

Appendix1.1a provides information about the importance given to each criterion. Reliability appears to be the most important criterion. The second most important criterion is cost per unit with at least 25% importance. Battery size emerges as the third most important with 3 out of 5 stars followed by average speed and carrying capacity. Since the focus is on making an optimal choice of robot for the trial by maximizing the utility in line with preferences, Weighted Sum Method (WSM) seems to be one of the most suitable approaches. WSM is a flexible method that can be used with both quantitative and qualitative data. It also ensures transparency by allowing the decision maker to assign weights to each criterion depending on the importance.

In accordance with the order of importance, weights are assigned to each criterion as shown below.

Robot Prototype	Weight	Archer	Bowler	Corner	Deviant
Carrying Capacity(L) - Max	14%	45	50	60	40
Battery SIze(hours) - Max	16%	18	18	12	24
Average Speed(km/h) - Max	15%	6	4	4	10
Cost Per unit (GBP) - Min	25%	5210	6250	4500	7100
Reliability (h) - Max	30%	22	24	24	32

Figure 2

To proceed with WSM, identification of the minimizing and maximizing criteria are important. Reliability, Average Speed, Battery Size and Carrying Capacity are the criteria that have to be maximized. Cost per unit is a minimizing criterion since this trial needs to be cost effective. Hence, we take an inverse of the values for cost per unit.

After normalizing the weights and the criteria values to use same measurements (calculations are shown in Appendix1.1b , Appendix1.1c), the final WSM scores for each prototype robots are obtained as shown below:

Archer	Bowler	Corner	Deviant
0.241624	0.22742877	0.24303077	0.287914393

Figure 3

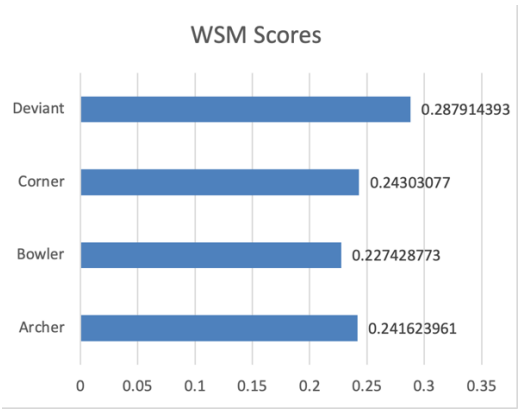


Figure 4

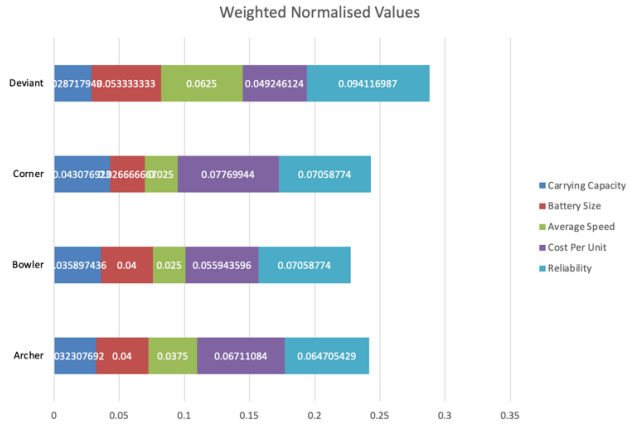


Figure 5

From the result in figures above it can be inferred that Deviant with a score of 0.28 is the highest scoring robot. From Figure 1 and Figure 5, it can be observed that Deviant has a high value of Reliability as compared to all other alternatives. Since Reliability is also the criteria with highest weight assigned as per the priority, Deviant would emerge as a distinguishable winner. Corner appears to be the suboptimal choice with a minute difference from Archer. But if maximum priority was given to cost per unit then Corner would be the optimal choice and Deviant would be a bad option. Corner would also be the best choice if carrying capacity was of greater importance.

WSM is a multi-attribute method that is sensitive to changing weights. Therefore, it is beneficial to perform sensitivity analysis to understand how changing the importance of each criteria can affect the results. Sensitivity analysis assists in identifying the critical criteria and helps in making informed decisions.

If sensitivity analysis is conducted by assigning equal weights to each criterion then it can be inferred that Deviant still remains the best choice and Archer is the suboptimal choice.

Robot Prototype	Weight
Carrying Capacity(L) - Max	20%
Battery SIze(hours) - Max	20%
Average Speed(km/h) - Max	20%
Cost Per unit (GBP) - Min	20%
Reliability (h) - Max	20%

Figure 6

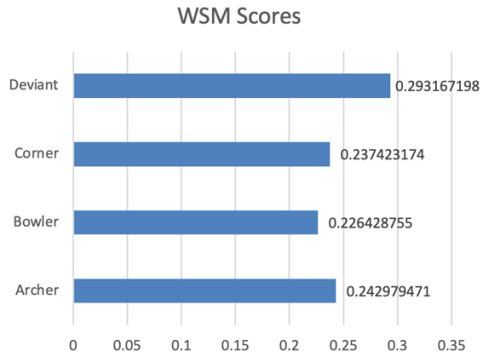


Figure 7

Therefore, even if all the criteria were to be equally important, Deviant continues to have the highest score and proves to be an optimal choice by maximizing utility.

2nd Task

The next task is to allocate the number of robots required to execute this one month trial across each of grocery, clothing and sports equipment stores. The goal is to keep the trial within a budget and complete maximum number of orders in a day.

The available data for the given task are as follows:

- Each robot has to execute 9,4,6 orders per day for grocery, clothing and sports equipment stores respectively.
- The cost of operation of the robot is 1600 GBP, 1000 GBP, 600 GBP per month grocery, clothing and sports equipment stores respectively.
- The technician labour hours required per week per robot are 10,7,5 hours for grocery, clothing and sports equipment stores respectively.
- Total technician work hours available to assist each store is 250 hours per week.
- The total budget for the trial is 250000 GBP.

Considering x1, x2, x3 correspond to the number of robots assigned to grocery, clothing and sports equipment stores respectively, Y is the maximum number of orders that can be delivered, d1+ and d1- are the positive and the negative deviations that total operating cost can take, d2+ and d2- are the positive and the negative deviations that technician man hours can take. Taking into account that Deviant is the most optimal choice of robot, Appendix 1.2a shows the equation for maximizing the number of orders, Appendix 1.2b is the equation for the operating cost for the whole trial which is the sum of the operating cost and acquisition cost of 7100 GBP for

Deviant, Appendix 1.2c shows the technician labour hours for a month's trial. All the equations are framed assuming 1 month is equal to 4 weeks.

It is known that minimum number of robots per store has to be greater than 5.
The appropriate allocation of robots to each store will help in maximizing the orders while adhere to the budget and technician work hour constraints. Since this task focuses on multiple objectives the best optimization technique would be goal programming. Goal programming deals with multiple contradictory objectives while minimizing the deviations. This technique realizes and balances all the objectives and tries to obtain a compromised solution that satisfies all the constraints.

Goals	x1	x2	x3	d1+	d1-	d2+	d2-	LHS	RHS	
OP cost	8700	8100	7700	-1	1	0	0	244300	250000	
Tech man hours	40	28	20	0	0	-1	1	1000	1000	
Lower Limit	5	5	5	0	0	0	0			
									Objective Function	0
Decision Variable	19	5	5	0	0	0	0			
Orders	252	168	112	0	0	0	0	6188		

Figure 8

Solving the equations using solver, the recommended allocation of robots is 19, 5, 5 robots to grocery, clothing and sports equipment stores respectively. It is observed that a maximum of 6188 orders can be delivered in a month while keeping the trial under a budget of 244300 GBP and technician work hour of 1000 hours. The allocation of robots in this format creates a balance and helps in achieving the goal of maximizing the number of orders, minimizing the budget. It also complies to the constraints of minimum number of robots that each store can be allocated with and manages the trial within the available technician work hours.

The conclusion from the analysis is that the optimal choice of robot that maximizes utility is Deviant. It offers the highest reliability, average speed, battery size and a competitive carrying capacity and cost per unit as per the set requirement. Allocating 19 robots to grocery store, 5 robots to clothing store and 5 robots to sports equipment store helps in minimizing the budget and technician man hours for trial to 244300 GBP and 1000 hours respectively thereby assisting the company to execute the trial smoothly and achieve the desired goal.

Part 2

1st Task

Drinks@home.uk wants to understand the spending of each customer on the website. The data includes information about 400 customers related to revenue, advertisement channel, estimated income, estimated age and time spent on the website. The brief of each of the parameters present in the data is as follows:

- Revenue – It signifies the revenue generated from the latest order made by a customer in GBP.

- Advertisement Channel – This corresponds to the advertising medium that brought the customer to the website. There are four channels (1. Leaflet, 2. Social media, 3. Search engine, 4. Influencer).
- Estimated Age – It determines the approximate age of the customer.
- Estimated income – This is the approximate income of the customer as per the website tracking software.
- Time on website – The average time spent (in seconds) on the website by the customer in a week.
- Seen Voucher – This signifies if the customer has a seen a discount voucher popup.

The goal is to identify the factors that positively or negatively impact the revenue based on the data and suggest the best method to increase profit. To establish business understanding it is important to focus on the metric revenue in order to understand the dependency of revenue on other metrics included in the data and identify their impacts. To learn more about the significance it is optimal to pursue regression to analyse the data set.

It is important to perform a detailed inspection of the relevant data for the analysis and identify if it is categorical or numerical. Various plots can be used to examine the distribution of the variables and to learn more about the pattern of the data.
To begin with, a test for any missing data is conducted in order to ensure an effective analysis. It is observed that there is no missing data in the given data set. From the data it can be noted that the variables Seen_Voucher and Advertisement_Channel are categorical variables. It is important to convert these categorical variables into factors in order to analyse.

Figure 9

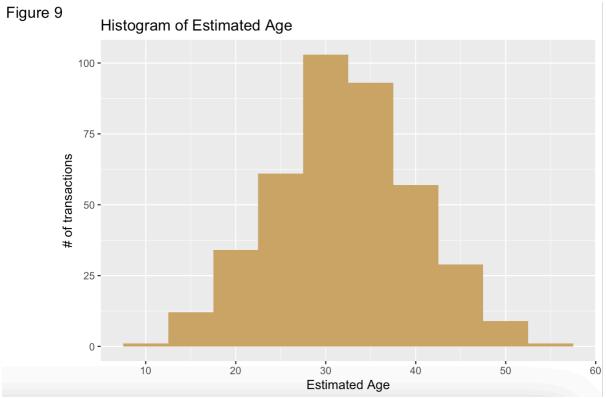


Figure 10

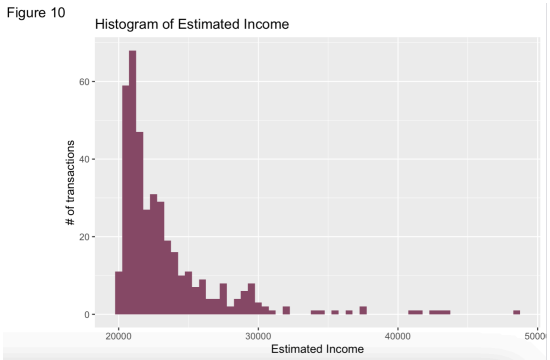


Figure 9 shows that maximum number of transactions are made by the customers in the age group of 27-37 years. Figure 10 gives an estimate that the estimated income of customers who made maximum transactions are in between 20000-23000 GBP.

Figure 11

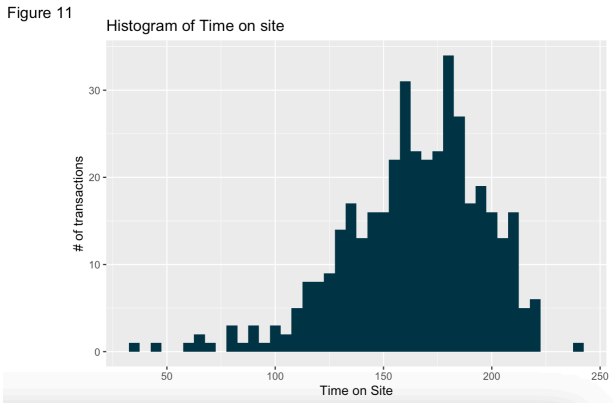


Figure 12

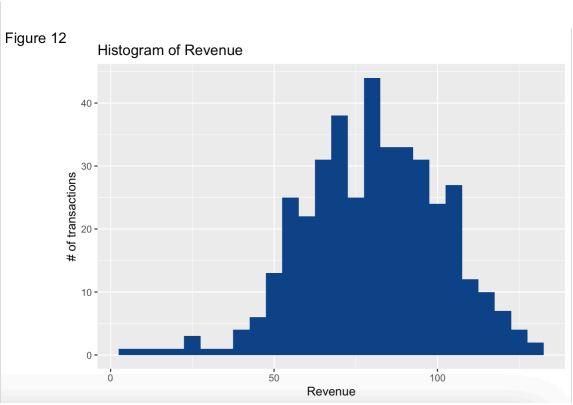


Figure 11 shows that highest number of transactions are made by customers who spent more time on the website. Figure 12 shows that higher the number of transactions higher is the revenue.

Further scatter plot is used to check the relationship of the dependent variable revenue with each independent with the impact of categorical variables.

Figure 13

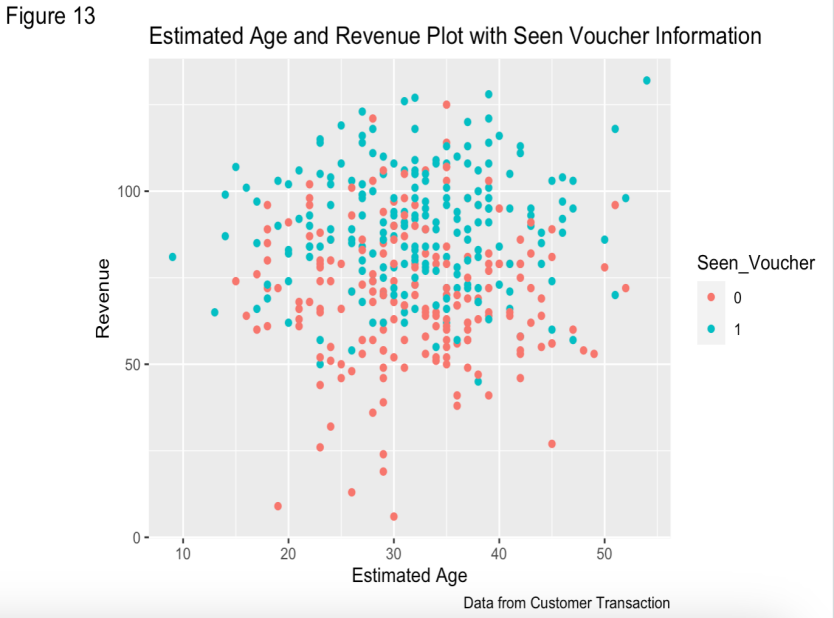
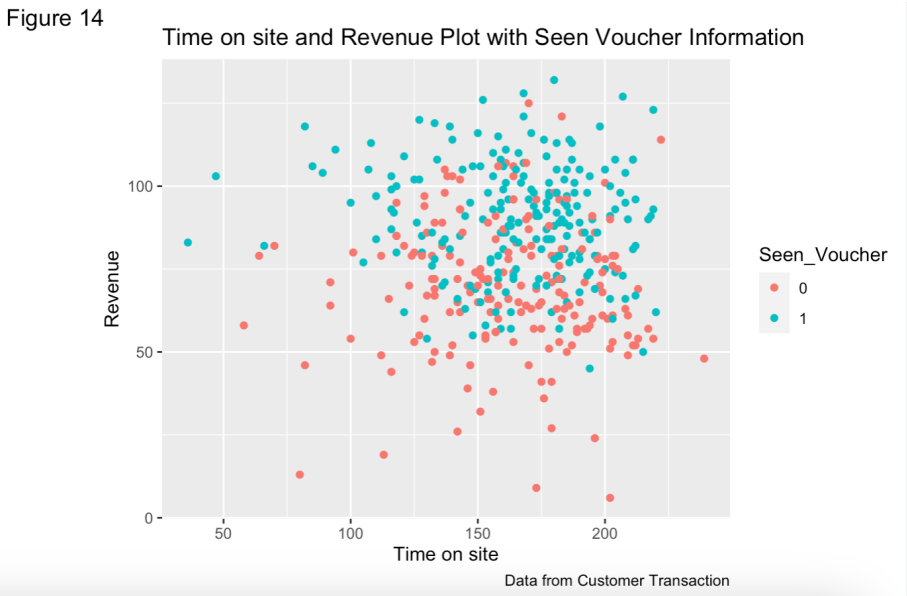
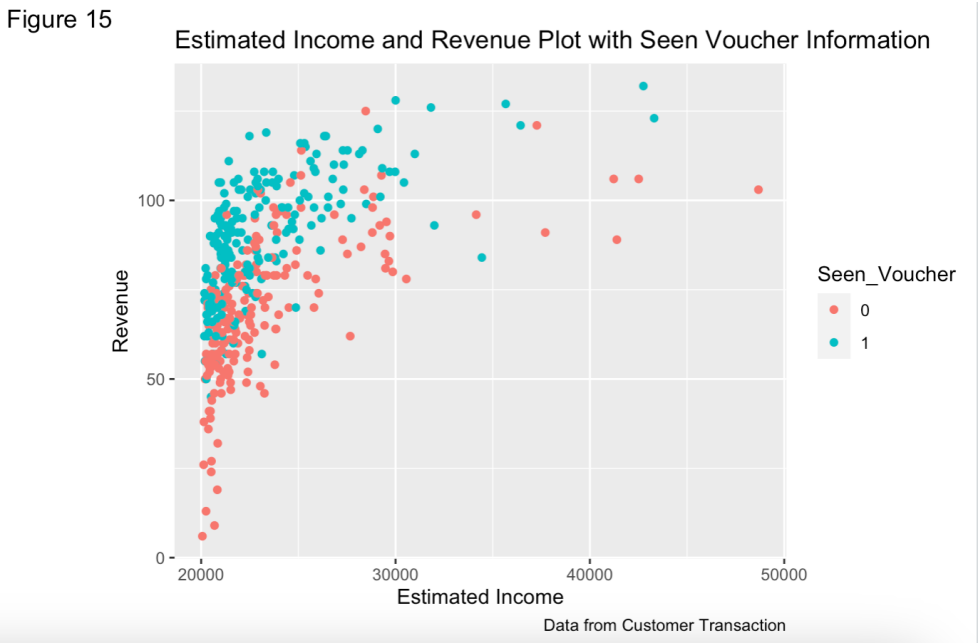


Figure 14



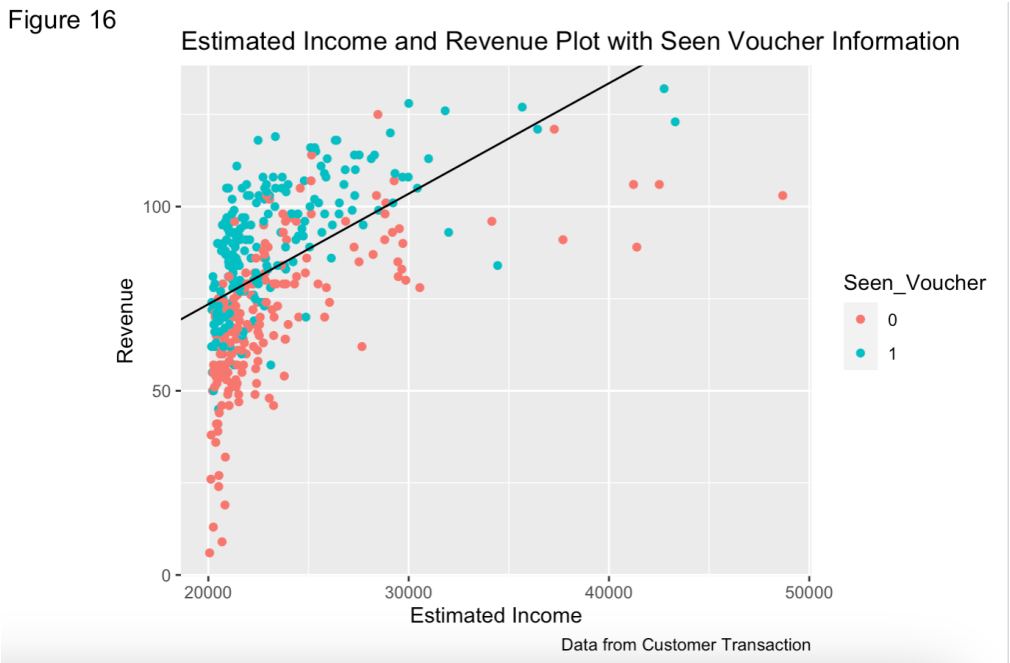
From Figure 13 and Figure 14, it can be concluded that there is no linear relationship between revenue and estimated age, time on site since scatter plot is a blob.

Figure 15



From Figure 15 it can be observed that there is a linear relationship between estimated income and revenue with the influence of the variable seen voucher. The variation in the dependent variable remains almost the same with the change in the independent variable.

Figure 16



Checking for the error of points, from the above graph it is observed that error of each data points are independent.

Figure 17

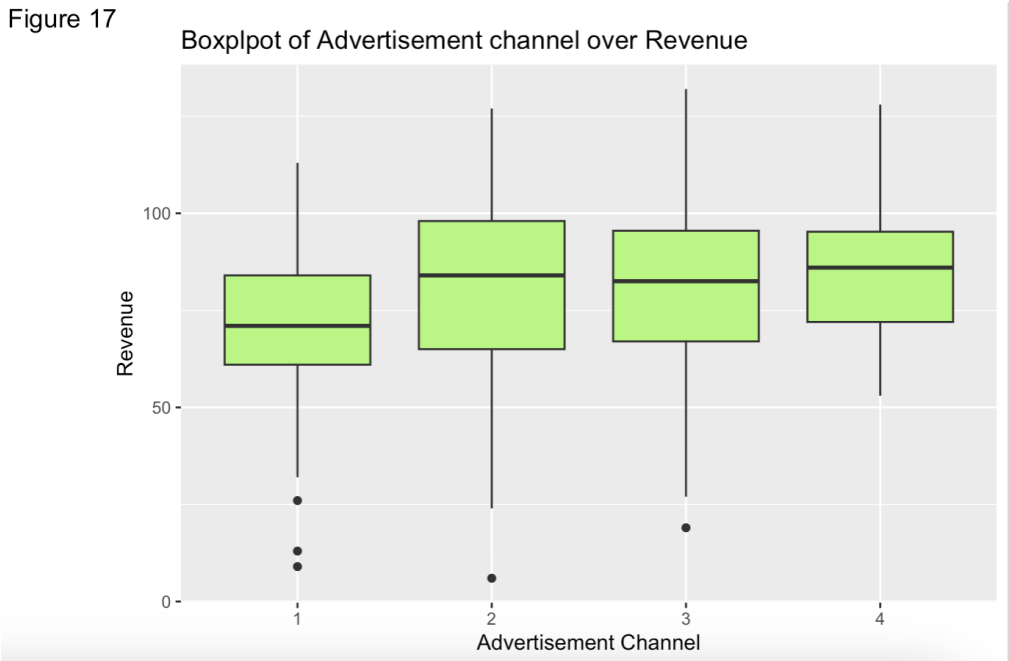
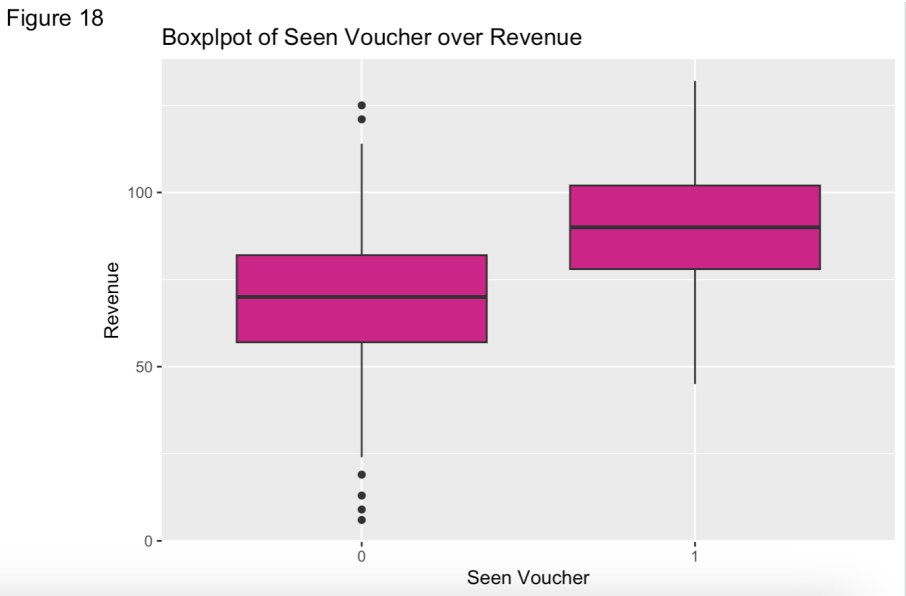


Figure 18



From Figure 17 and Figure 18 it can be noted that the categorical variable advertisement channel and seen voucher have an influence on the Revenue.

All the above plots help in deciding upon which variables have an impact on revenue. The independent variables Estimated income, Advertisement channel and Seen voucher seem to influence the dependent variable revenue.

It also important to perform a check for multicollinearity. From Appendix 2.1a it can be observed that the correlation between the independent variables is almost zero or negligible. Hence there is no issue of multicollinearity in the data.

As stated earlier it is known that this is a regression type of analysis. After analysing the data it is clear that the variables that create a dependency are estimated income, advertisement channel, seen voucher. But the information regarding the variables influencing revenue positively or negatively are still unknown.

To understand the influence of each estimated income, seen voucher and advertisement channel on revenue it is important to perform multiple linear regression. Time on site and estimated age do not show linearity with revenue and hence does not have any impact.

```
Call:
lm(formula = Revenue ~ Advertisement_Channel + Estimated_Income +
    Seen_Voucher, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-52.574  -7.599   1.230   8.118  40.637

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          -7.7260963  4.6491735  -1.662   0.0973 .
Advertisement_Channel  3.9978347  0.6300427   6.345 6.07e-10 ***
Estimated_Income       0.0029061  0.0001812  16.038  < 2e-16 ***
Seen_Voucher1         19.5288841  1.4094637  13.856  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.08 on 396 degrees of freedom
Multiple R-squared:  0.5504,    Adjusted R-squared:  0.547
F-statistic: 161.6 on 3 and 396 DF,  p-value: < 2.2e-16
```

Figure 19

From the summary of the regression model that considers revenue as the dependent variable, following inferences can be made:

- Advertisement channel has a positive influence on revenue. For every customer brought in by any of the 4 advertisement channels, revenue increases by 3.997 GBP.
- Estimated income has a positive influence on revenue which implies that for a single unit increase in the estimated income the revenue increases by 0.002GBP.
- The coefficient value of seen voucher is the highest which indicates that if a customer has seen the voucher there is a higher impact on revenue rather than those who have not. The revenue increases by 19.528 GBP if the customer has seen the voucher. This also indicates a positive influence.

From the respective p-values of the independent variables it can be observed that estimated income, advertisement channel and seen voucher are significant. Hence it can be concluded that time on site and estimated age have no impact on revenue and are also statistically insignificant. However, the independent variables seen voucher, advertisement channel and estimated income are statistically significant and also influence the revenue positively.

2nd Task

- Based on the regression analyses it is known that there is no proven linear relationship between estimated age and revenue and there is insufficient evidence to prove its significance. Hence running an advertisement to target the customer older that 45 years of age may not be a wise choice to make. However, it might be something that can be tested in the future with some additional data.

- Another observation from the analysis is that the variable seen voucher has significantly high positive influence on revenue. An increase of 19.62 GBP can be observed for every customer who made a transaction on the basis of seeing a discount voucher. This indicates that providing a voucher that gives the customer 20GBP off on their next buys can be a good way of enhancing the revenue. It can prove to be a profitable strategy since the customers who have seen a voucher happen to spend more.

- To know the influence of each channel individually it important to dummify the categorical variable.

```
Call:
lm(formula = Revenue ~ Advertisement_Channel_3 + Advertisement_Channel_2 +
    Advertisement_Channel_4 + Estimated_Income + Seen_Voucher,
    data = newdata)

Residuals:
    Min      1Q  Median      3Q     Max
-54.575  -7.738   1.029   8.861  38.841

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)            -3.8215992  4.4645012  -0.856 0.392520
Advertisement_Channel_3  8.1139387  1.9943405   4.068 5.72e-05 ***
Advertisement_Channel_2  6.9039364  2.0132224   3.429 0.000669 ***
Advertisement_Channel_4 12.9103883  1.9901035   6.487 2.63e-10 ***
Estimated_Income         0.0028656  0.0001834  15.622  < 2e-16 ***
Seen_Voucher1           19.6270625  1.4113549  13.907  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.07 on 394 degrees of freedom
Multiple R-squared:  0.5534,    Adjusted R-squared:  0.5478
F-statistic: 97.66 on 5 and 394 DF,  p-value: < 2.2e-16
```

Figure20

Advertisement channel is a categorical data with four categories which means that there are four different channels through which the customers were drawn to make transactions. To know if splurging on influencer marketing would be beneficial it is important to analyse the impact and the significance of each advertisement channel. From the analysis it is observed that channel 1 which is advertising through leaflets is considered as the base and from the summary of the model it can be seen that it is not significant and has a negative influence on revenue. Channel 2, advertising via social media proves to be statistically significant and also has a positive impact on revenue. Search engine being the third advertisement channel has a positive impact and is also significant. But, advertising through an influencer which is channel 4 stands out by being statistically significant and increasing the revenue by 12.91 GBP for every customer brought in by the influencer channel.

Considering the evidences from the analysis, advertising through an influencer has a pretty strong impact on revenue however providing the customer with a voucher of 20GBP off their next purchase would be a better choice as it helps in increasing the revenue by a higher amount. This is because per unit increase in revenue is observed to be higher for a customer who was brought in on viewing a voucher. Therefore, it proves to be a data driven choice.

Appendices

Appendix 1.1: Part 1 Task1

Appendix 1.1a: Management team priority order

Carrying Capacity	The carrying capacity is the least important criteria according to the majority of the management team.
Battery Size	After careful deliberation, the size of the battery is 3 out 5 stars important according the majority of the board. They believe that it is an important criteria but not as important as some others and that this would likely improved with better battery tech in the future anyway.
Average Speed	The speed is not as important as battery size but more important than carrying capacity.
Cost per Unit	The cost is more important that any other criteria except for the reliability. One of the management team considered it to be at least 25% of total consideration amongst all criteria.
Reliability	This is the most important consideration according to the management team and this is clearly favoured over all other criterias.

Appendix 1.1b: Converting the minimizing criteria to maximizing

Robot Prototype	Weight	Archer	Bowler	Corner	Deviant
Carrying Capacity(L) - Max	14%	45	50	60	40
Battery SIze(hours) - Max	16%	18	18	12	24
Average Speed(km/h) - Max	15%	6	4	4	10
Cost Per unit (GBP) - Min	25%	0.000192	0.00016	0.00022222	0.00014085
Reliability (h) - Max	30%	22	24	24	32

Appendix 1.1c: Solving using WSM

Robot Prototype	Weight	Archer	Bowler	Corner	Deviant
Carrying Capacity(L) - Max	0.14	0.2307692	0.25641026	0.307692308	0.205128205
Battery SIze(hours) - Max	0.16	0.25	0.25	0.166666667	0.333333333
Average Speed(km/h) - Max	0.15	0.25	0.16666667	0.166666667	0.416666667
Cost Per unit (GBP) - Min	0.25	0.2684434	0.22377439	0.310797758	0.196984495
Reliability (h) - Max	0.30	0.2156848	0.23529247	0.235292468	0.313723291

Appendix 1.2: Part 1 Task2

Appendix 1.2a: Calculating the maximum number of orders for a month's trial

$$252x1 + 168x2 + 112x3 \geq Y$$

Appendix 1.2b: Total operating cost

$$8700x1+8100x2+7700x3-d1+d1- \leq 250000$$

Appendix 1.2c: Technician labour hours

40x1+28x2+20x3-d2+d2- = 1000

Appendix 2.1: Part 2 Task1
Appendix 2.1a: Table for Correlation

```
> cor(data)
                      Estimated_Age Time_On_Site Seen_Voucher Estimated_Income
Estimated_Age           1.00000000   0.02774864  0.011600588      0.014782713
Time_On_Site            0.02774864   1.00000000  0.050760998     -0.038676812
Seen_Voucher            0.01160059   0.05076100  1.000000000     -0.009811076
Estimated_Income        0.01478271  -0.03867681 -0.009811076      1.000000000
Advertisement_Channel   0.09709081   0.02294742  0.020136949     -0.020151214
Revenue                 0.02628325  -0.02826230  0.465976667      0.531651842
                      Advertisement_Channel     Revenue
Estimated_Age                    0.09709081  0.02628325
Time_On_Site                     0.02294742 -0.02826230
Seen_Voucher                     0.02013695  0.46597667
Estimated_Income                -0.02015121  0.53165184
Advertisement_Channel            1.00000000  0.21240069
Revenue                          0.21240069  1.00000000
```