

Contents Table

1	Introduction.....	3
2	The Sample.....	4
3	Description of Data.....	5
4	Bivariate analysis of data.....	12
5	Conclusion.....	19
	Bibliography & References.....	

1 Introduction

A random sample of 20 car users were consulted about four factors that determined their choice of car; miles per gallon, horsepower, service interval and price.

Utilising SPSS, the subsequent report will aid in describing, manipulating, analysing and interpreting this sample of collected data. It will see if there is a relationship between variables and if one variable can predict another.

The report will be split into the following sections:

- 2- Display of data that is to be used in SPSS.
- 3- Description of data using appropriate descriptive statistics
- 4- Using appropriate techniques to see if there is a relationship between variables and to see if one can predict another.
- 5- The report will conclude with a summary of all the findings.

2 The Sample

The following is the data drawn at random from 20 car users. The choice of car was determined by four factors- miles per gallon, horsepower, service interval and price. I will use SPSS to analyse the data for meaning.

Table 2: *Random Sample of 20 Car Users*

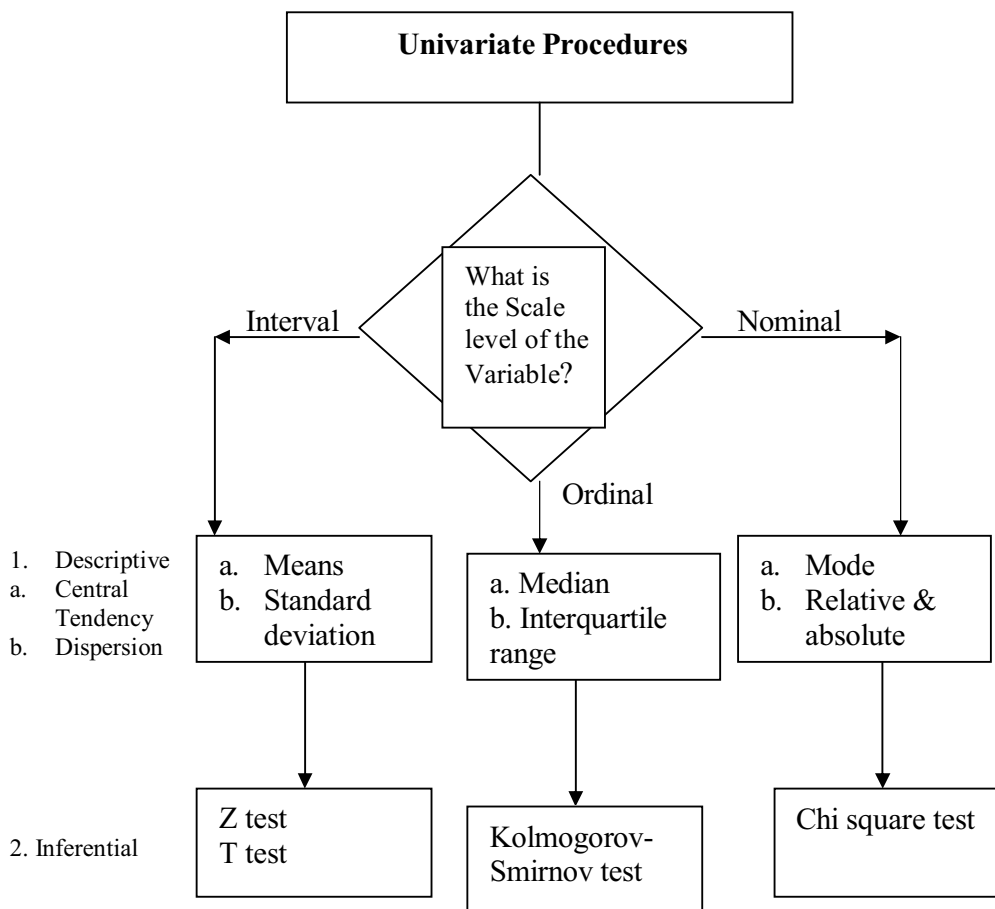
<i>Cases</i>	<i>MPG</i>	<i>HORSEPOWER</i>	<i>SERVICE INT.</i>	<i>PRICE</i>
<i>1</i>	24.00	118.00	8000.00	15000.00
<i>2</i>	26.00	120.00	10000.00	18000.00
<i>3</i>	22.00	135.00	8000.00	16000.00
<i>4</i>	28.00	120.00	10000.00	17000.00
<i>5</i>	24.00	125.00	12000.00	18000.00
<i>6</i>	25.00	130.00	8000.00	18000.00
<i>7</i>	29.00	136.00	10000.00	20000.00
<i>8</i>	30.00	135.00	8000.00	22000.00
<i>9</i>	25.00	140.00	12000.00	24000.00
<i>10</i>	24.00	135.00	10000.00	23000.00
<i>11</i>	26.00	120.00	10000.00	18000.00
<i>12</i>	24.00	118.00	8000.00	15000.00
<i>13</i>	24.00	125.00	12000.00	18000.00
<i>14</i>	25.00	130.00	8000.00	18000.00
<i>15</i>	22.00	135.00	8000.00	16000.00
<i>16</i>	30.00	135.00	8000.00	22000.00
<i>17</i>	29.00	136.00	10000.00	20000.00
<i>18</i>	25.00	140.00	12000.00	24000.00
<i>19</i>	24.00	135.00	10000.00	23000.00
<i>20</i>	28.00	120.00	10000.00	17000.00

3 Description of Data

In order to identify which descriptive analytical tools to use I must identify the type of data being analysed. The data is Univariate as is a single set of data. This then gives way to the question what is the scale level of the variable. There are 3 possible routes here :

- Nominal – Categorical data
- Ordinal – Ranked data
- Interval – collected with random answer

The data in this case is collected in interval mode as replies not categorised or given in rank but actual answers given randomly dependent on preference. Therefore using the flow chart below it can be seen that in order to do a descriptive analysis of data using interval method must first calculate mean and standard deviation then carry out either a z or t test.



The table below is the descriptive analysis of the data.

The range shows the value between the minimum and maximum data number in that field this will help to show if preference is more spread i.e. wide range or concentrated i.e. small range.

The minimum shows the least value in that specific area of preference e.g. lowest price for car.

The maximum shows the highest value for that area of preference e.g. highest car price.

The above are all measures of dispersion i.e. the spread of the data.

Saunders et al 1997 believes “The central tendency usually provides some general impression of values that could be seen as common, middling or average.”

The mean is the most frequently use measure of central tendency and is described as the average that includes all data values in its calculation. As in the case it be done by adding all values together and dividing the by the number of samples. This will provide a figure that will allow for the average amount that of all replies to be seen e.g. the average car price from all responses.

Saunders et al 1997 “As well as describing the central tendency for a variable, it is important to describe how the data values dispersed around the central tendency.” The method of calculating dispersion in this case will be by standard deviation. Standard deviation is the extent to which value differ from the mean. This will show how far the data is spread. This will tell us whether preferences vary from small or large range on overall responses.

Table 1 Descriptive Statistics

	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
SERVICE	20	4000.00	8000.00	12000.00	9600.0000	343.3580	1535.5438	2357894.737
PRICE	20	9000.00	15000.00	24000.00	19100.0000	660.5420	2954.0338	8726315.789
MPG	20	8.00	22.00	30.00	25.7000	.5530	2.4730	6.116
H.POWER	20	22.00	118.00	140.00	129.4000	1.7417	7.7893	60.674
Valid N	20							

Table 1 above shows:

Service – It shows that the range is large in this section at 4000.00 so opinion on this option is Spread on what consumers prefer. Also the average service required by a consumer is 9600.00 so this is the figure that will satisfy most consumers. The standard deviation is 1535.5 which tell us the variation of preference is large. On a whole what can be seen is that there is a spread opinion amongst consumers on preference in this area.

Price – The range is £9000 this means the value between highest and lowest preference value is large. The mean of the values is £19100 this is the average value that a consumer would prefer to pay for the car. The variance from the mean of the data at 95% confidence is £2954. This shows that 95 % of replies are within that figure of the mean. This figure is quite large so it can be seen that consumer preference in this area also is varied.

MPG – It can be seen that range of the replies was 8. Although smaller figure than other areas is still significant range if size of replies taken into account. The mean of the data is 25.7 so this is the average consumer preference in this area. The standard deviation of this mean is 2.473 which is quite a small deviation from the mean. Thus showing that data is clumped close together and preference of most consumers is in quite a small range.

H.Power – the range of the replies was 22. This is quite a small range. The mean of the data is 129.4 for consumer preference. The standard deviation of the data is 7.7893 which shows most data is in a very small range and consumer preference is mostly based close to mean.

The next test will be a t test instead of a z test. The first reason being there is only a small sample. Secondly a t test is doing parametric testing which this data is.

Although the author is believes that he will not find any significance by doing this test he is doing it to test this hypothesis that there is no significance on a one sample test t-test. But will be interested to look at 95% Confidence Interval.

Table 2 T-test One-Sample Test

Test Value = 0						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
SERVICE	27.959	19	.000	9600.0000	8881.3434	10318.6566
PRICE	28.916	19	.000	19100.0000	17717.4696	20482.5304
MPG	46.475	19	.000	25.7000	24.5426	26.8574
H.POWER	74.293	19	.000	129.4000	125.7545	133.0455

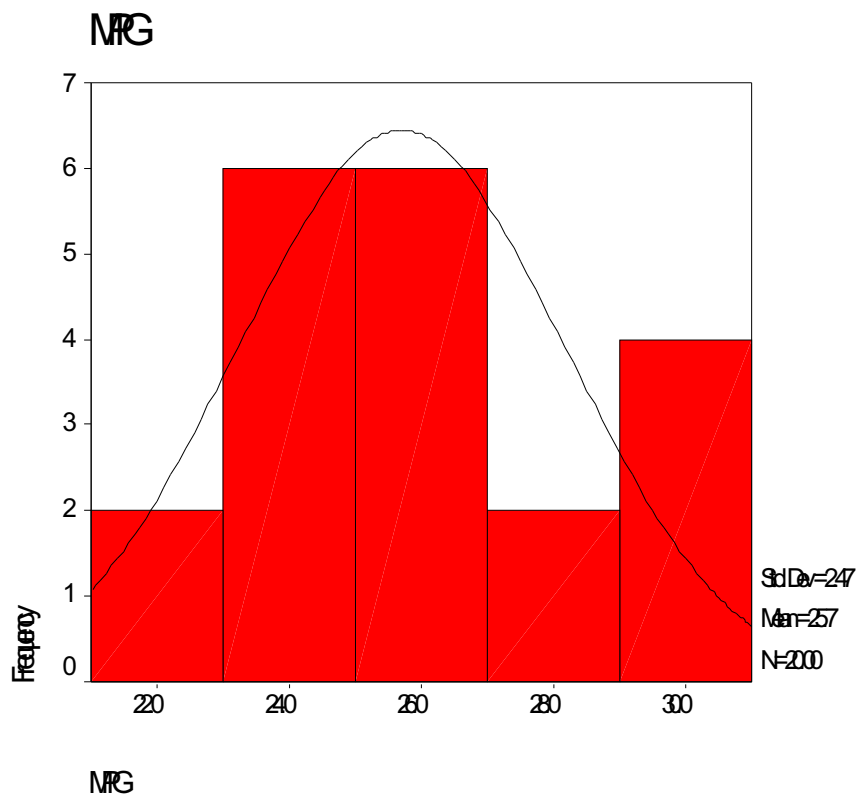
The table confirms the belief that there is no significance in any of the four variables. But what can be seen is that the 95% Confidence intervals for Price and Service have a large spread. Where as the 95% Confidence Interval for MPG and H.Power is very small. This shows that data for Price and Service is spread over a wider range thus preference in these areas is more varied than the latter two. The data for MPG and H.Power is more concentrated thus consumer preference in these areas is more agreed upon. More consumers want roughly the same level for these two things. The author has spent very little time defining this table as shows up very little in results.

Graph 1: Miles-Per-Gallon as a Determining Purchasing Factor**MPG**

N	Valid	20
	Missing	0

MPG

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	22.00	2	10.0	10.0	10.0
	24.00	6	30.0	30.0	40.0
	25.00	4	20.0	20.0	60.0
	26.00	2	10.0	10.0	70.0
	28.00	2	10.0	10.0	80.0
	29.00	2	10.0	10.0	90.0
	30.00	2	10.0	10.0	100.0
	Total	20	100.0	100.0	



The table shows that two most popular groups are 24 (30%) and 25 (20%) accounting for 50% of the data between them. This shows that 2 groups that are very close to each other can satisfy 50% of consumers. The graph shows that data can be mapped in bell shape curve.

Graph 2: Horse-Power as a Determining Purchasing Factor**H.POWER**

N	Valid	20
	Missing	0

H.POWER

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	118.00	2	10.0	10.0	10.0
	120.00	4	20.0	20.0	30.0
	125.00	2	10.0	10.0	40.0
	130.00	2	10.0	10.0	50.0
	135.00	6	30.0	30.0	80.0
	136.00	2	10.0	10.0	90.0
	140.00	2	10.0	10.0	100.0
	Total	20	100.0	100.0	

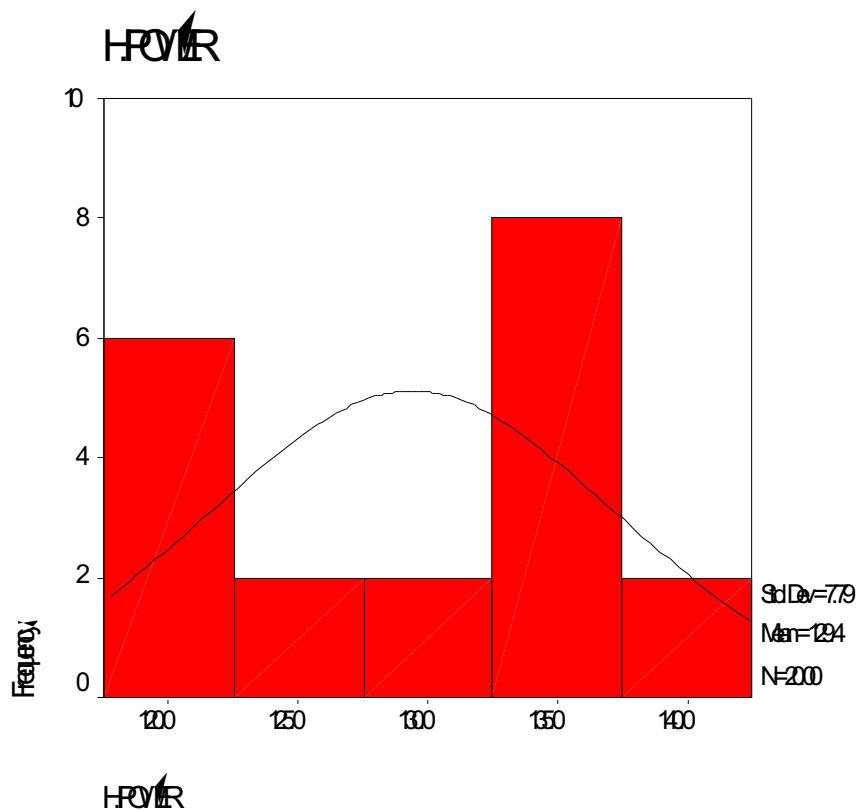


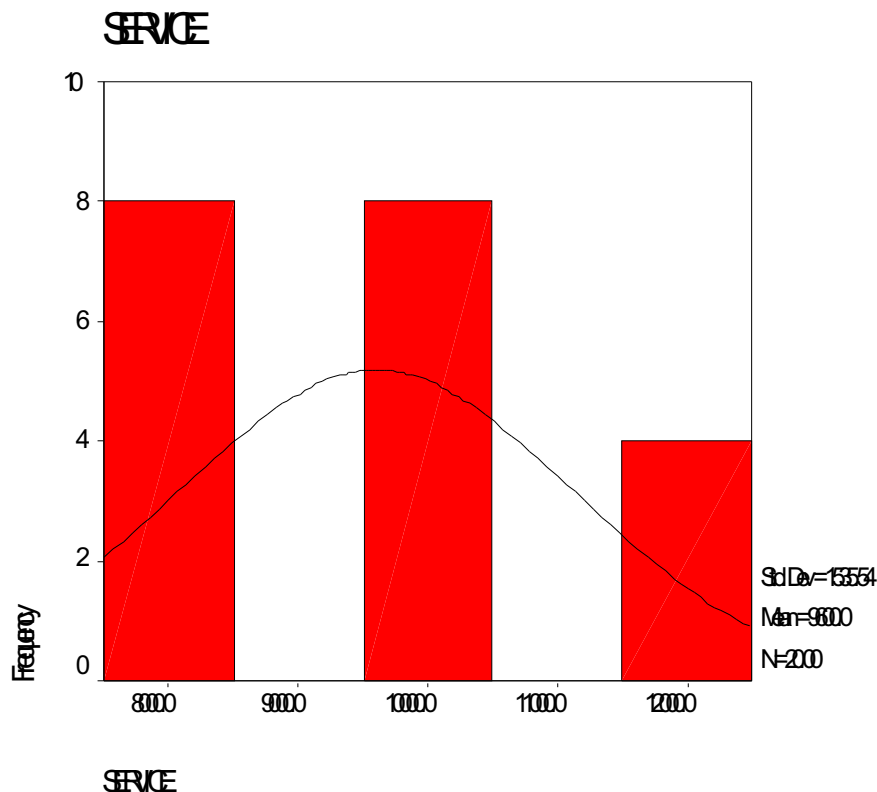
Table shows data to be spread. But two groups 120 (20%) and 135 (30%) dominate 50% of the data. But there is a gap between these two groups thus spreading the data and making it harder to predict consumer preference. The graph represents this with the curve being flatter and wider than for MPG graph.

Graph 3: Service Interval as a Determining Purchasing Factor**SERVICE**

N	Valid	20
	Missing	0

SERVICE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	8000.00	8	40.0	40.0	40.0
	10000.00	8	40.0	40.0	80.0
	12000.00	4	20.0	20.0	100.0
	Total	20	100.0	100.0	



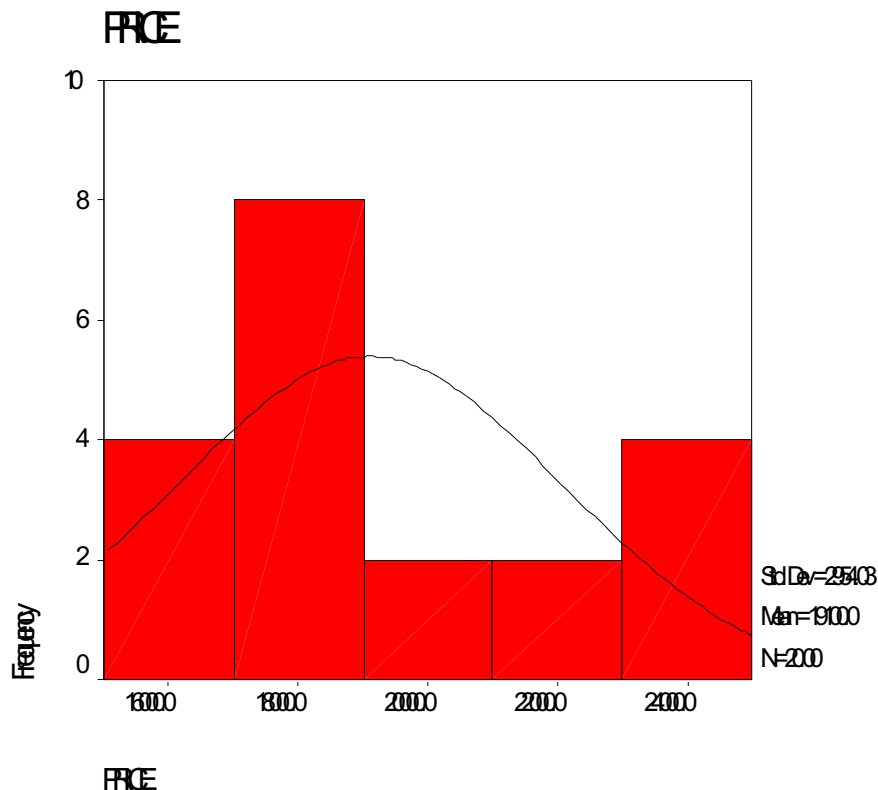
The table shows that there are only 3 groups of data. With 8000 (40%) and 10000 (40%) dominating the responses but the range between 3 responses is large. This would explain why graph curve is flatter. This demonstrates 80% of consumers have preference within 2 ranges.

Graph 4: Price as a Determining Purchasing Factor**PRICE**

N	Valid	20
	Missing	0

PRICE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	15000.00	2	10.0	10.0	10.0
	16000.00	2	10.0	10.0	20.0
	17000.00	2	10.0	10.0	30.0
	18000.00	6	30.0	30.0	60.0
	20000.00	2	10.0	10.0	70.0
	22000.00	2	10.0	10.0	80.0
	23000.00	2	10.0	10.0	90.0
	24000.00	2	10.0	10.0	100.0
	Total	20	100.0	100.0	

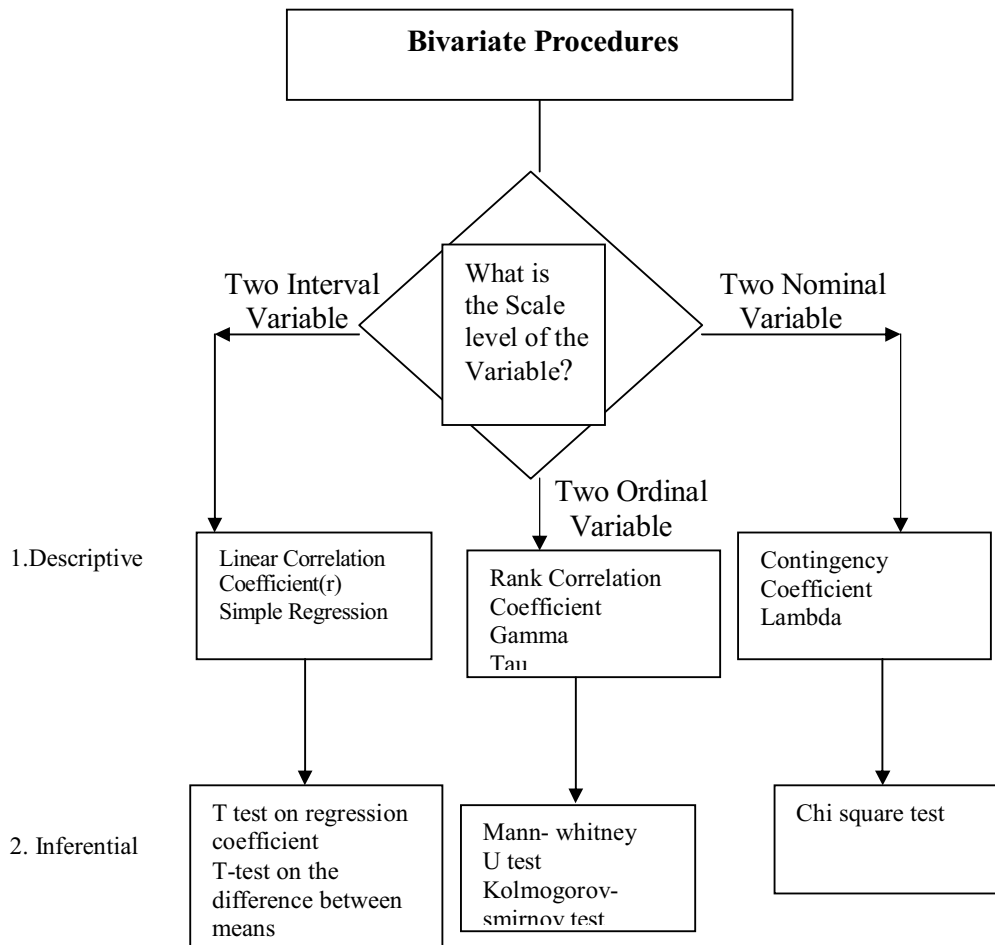


The table shows that consumer preference is spread over a large number of options. The data is spread evenly apart from at 18000 because 30% of consumer choose this as preference point. The data has the widest spread out of any option. Thus the graph curve is quite flat and long.

4 Bivariate Analysis of Data

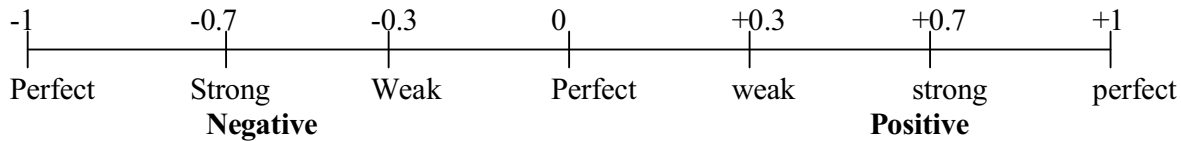
Using the appropriate techniques I will now find out if there is a relationship between the variables. Once that has been done I will attempt to find out if one variable can predict another.

The route that will be taken is the two interval variable on the table as the author has already identified as to why the interval route is to be taken. The reason for bivariate testing is that it explores and identifies relationship between variables.



Correlation's

A Correlation enables me to quantify the strength of the relationship between two variables. A perfect positive correlation is represented by the value +1. This means that the two variables are closely related and as value of one changes the other will change proportionately. Also able to get negative perfect correlation -1 so as one changes other will change negatively proportionately the correlation is best shown by diagram below out of Saunders 1997:



Descriptive Statistics

	Mean	Std. Deviation	N
SERVICE	9600.0000	1535.5438	20
H.POWER	129.4000	7.7893	20
MPG	25.7000	2.4730	20
PRICE	19100.0000	2954.0338	20

Correlations

Correlations					
		SERVICE	H.POWER	MPG	PRICE
SERVICE	Pearson Correlation	1.000	.102	-.033	.427
	Sig. (2-tailed)	.	.668	.889	.060
	N	20	20	20	20
H.POWER	Pearson Correlation	.102	1.000	.056	.730
	Sig. (2-tailed)	.668	.	.815	.000
	N	20	20	20	20
MPG	Pearson Correlation	-.033	.056	1.000	.307
	Sig. (2-tailed)	.889	.815	.	.188
	N	20	20	20	20
PRICE	Pearson Correlation	.427	.730	.307	1.000
	Sig. (2-tailed)	.060	.000	.188	.
	N	20	20	20	20

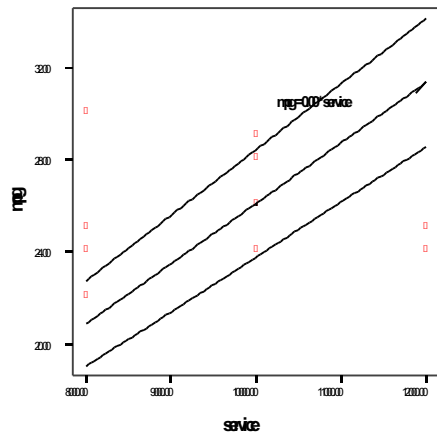
** Correlation is significant at the 0.01 level (2-tailed).

As can be seen from the table there are 3 strong correlation's that exist. So a strong relationship exists amongst three sets of variables. They are listed below in rank of highest correlation first:

- Service & MPG- .889
- H.Power & MPG- .815
- Service & H.Power- .668

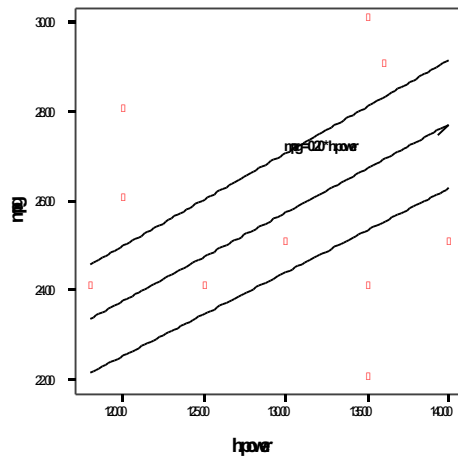
The list is topped with Service and MPG, which has a very strong correlation so are very closely related. The author would have expected MPG and H.Power to top the list but as can be seen it came second but has a strong correlation also. The third one is service and H.Power which are correlated by to a weaker extent. So it can be seen that consumer preference on Miles per Gallon. affects their preference on both Service and Horse Power of the car.

Regression



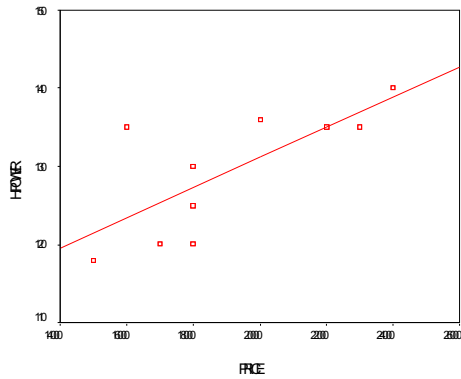
Linear Regression through the Origin with
95.0% Air Pollution Index

*



Linear Regression through the Origin with
95.0% Air Pollution Index

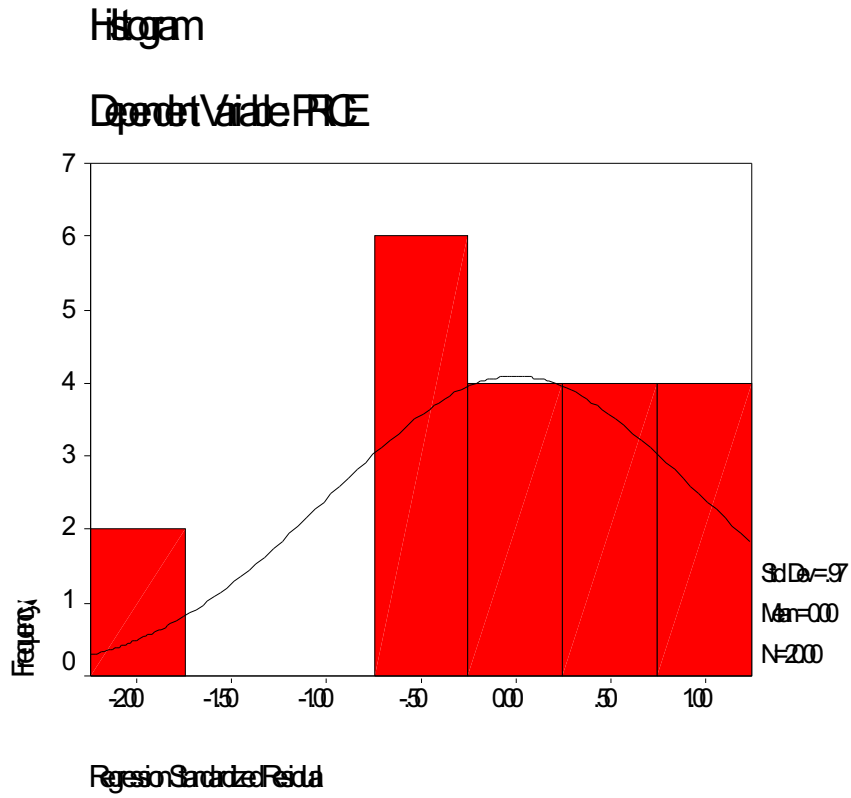
a Predictors: (Constant), H.POWER
b Dependent Variable: PRICE



Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	15943.4424	22035.0449	19100.0000	2156.7961	20
Std. Predicted Value	-1.464	1.361	.000	1.000	20
Standard Error of Predicted Value	465.1759	836.6033	645.3391	119.7738	20
Adjusted Predicted Value	16126.8135	21695.1563	19078.0141	2141.8498	20
Residual	-4650.5898	2349.4102	1.455E-12	2018.5505	20
Std. Residual	-2.242	1.133	.000	.973	20
Stud. Residual	-2.334	1.179	.005	1.020	20
Deleted Residual	-5039.6689	2545.9673	21.9859	2217.5194	20
Stud. Deleted Residual	-2.717	1.193	-.031	1.113	20
Mahal. Distance	.006	2.142	.950	.700	20
Cook's Distance	.000	.228	.049	.067	20
Centered Leverage Value	.000	.113	.050	.037	20

a. Dependent Variable: PRICE



Paired T-Test

The hypothesis for this test is that it will show Mpg and service to have a very strong relationship.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MPG	25.7000	20	2.4730	.5530
	H.POWER	129.4000	20	7.7893	1.7417
Pair 2	MPG	25.7000	20	2.4730	.5530
	SERVICE	9600.0000	20	1535.5438	343.3580
Pair 3	MPG	25.7000	20	2.4730	.5530
	PRICE	19100.0000	20	2954.0338	660.5420
Pair 4	H.POWER	129.4000	20	7.7893	1.7417
	SERVICE	9600.0000	20	1535.5438	343.3580
Pair 5	H.POWER	129.4000	20	7.7893	1.7417
	PRICE	19100.0000	20	2954.0338	660.5420
Pair 6	SERVICE	9600.0000	20	1535.5438	343.3580
	PRICE	19100.0000	20	2954.0338	660.5420

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	MPG & H.POWER	20	.056	.815
Pair 2	MPG & SERVICE	20	-.033	.889
Pair 3	MPG & PRICE	20	.307	.188
Pair 4	H.POWER & SERVICE	20	.102	.668
Pair 5	H.POWER & PRICE	20	.730	.000
Pair 6	SERVICE & PRICE	20	.427	.060

Paired Samples Test

		Paired Differences Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	MPG - H.POWER	-103.7000	8.0400	1.7978	-107.4629	-99.9371	-57.681	19	.000
Pair 2	MPG - SERVICE	-9574.3000	1535.6280	343.3769	-10292.9960	-8855.6040	-27.883	19	.000
Pair 3	MPG - PRICE	-19074.3000	2953.2758	660.3725	-20456.4756	-17692.1244	-28.884	19	.000
Pair 4	H.POWER - SERVICE	-9470.6000	1534.7682	343.1846	-10188.8936	-8752.3064	-27.596	19	.000
Pair 5	H.POWER - PRICE	-18970.6000	2948.3515	659.2714	-20350.4710	-17590.7290	-28.775	19	.000
Pair 6	SERVICE - PRICE	-9500.0000	2685.2423	600.4384	-10756.7321	-8243.2679	-15.822	19	.000

From the tables above it can be seen that none of the above pairs has a strong significance and correlation. There are strong significance held by MPG and H. Power & MPG and Service but both have a weak significance. The standard deviation variance range also shows very little that has not been discussed before. Therefore hypothesis for testing proven to be wrong as no links were found as although significance high the correlation was very low at .056.

5 Conclusion

To conclude the report it can be clearly identified that consumer preferences on Price and Service are very spread in range. Where as consumer preferences on MPG and Horse power are closer together enabling the company to focus roughly around the mean area as the standard deviation shows. This shows that consumers want the same thing in these areas.

The bivariate analysis showed that there is a close relationship between:

- Service & MPG- .889
- H.Power & MPG- .815
- Service & H.Power- .668

Therefore there is a close relationship between these. This correlation testing proves that they are very closely linked

But the secondary testing showed that one variable cannot directly predict the another. Even though have significant areas they are not correlated enough to sustain a link.

Bibliography & References

Bajpai, A. C. et al. (1974) *Engineering Mathematics*,

London: Wiley & Sons Ltd

Foster, J. J. (1993) *Starting SPSS/PC+ and SPSS for Windows – A beginners guide to data analysis*,
Wilmslow, United Kingdom: Wiley & Sons Ltd

Jankowicz, A.D, (1995) *Business Research Projects*. 2nd Edition,
Cornwall, United Kingdom: Thompson Business Press

Saunders, M. et al. (1997) *Research Methods for Business Students*,
Harlow, Essex: Financial Times - Prentice Hall