Data Handling Course work - Used Cars Sales.

Aim

I am aiming to find out whether used cars have a correlation between there engine size and there insurance group. I will also aim to find out the comparisons between the most popular colours of the cars shown on my data sheet provided.

Introduction

During this investigation, I am investigating correlations and comparisons between used cars sales.

For each I will chose 18 cars at random, this will therefore calculate to 36 cars. The reason for choosing 36 cars is because there are 360 on a pie chart and consequently it will be easier to portray on the pie chart. I will use the random button on my calculator; the sequence I used to do this was;

$$66 \longrightarrow \text{SHIFT} \longrightarrow \text{RAM}$$

This sequence will make my investigation unbiased and a fair test and additionally more quick and efficient. Once my numbers are highlighted on the data sheet, I will go on to write a hypothesis for both correlation and comparisons.

Furthermore, all of the results and data I collect during the investigation, will be put into clear and well laid out tables, which can be used to produce other graphs such as scatter diagrams and bar charts. If there is any data missing from my results I will have to show this clearly on my graphs and tables, scatter diagrams will be very helpful if this occurs for the anomalies will clearly state an incorrect or missing piece of data. Moreover if I were to add a line of best fit this, would help me to predict and calculate the differences between the low data and the high data.

I can then analyze all of the data carefully and accurately and then I will be able to write a constructive summary of the investigation in general.

Rebecca Stewart. Mrs Stretch, Maths (Handling Data.) 07/05/2007

Hypothesis

I predict that if as I investigate the comparisons of cars I would state that I believe the most popular coloured car are red the reason for this is for as I glance at the first page of data this to me is the modal number. I am also going to go ahead and collect a numerical piece of data for another comparison. For this, I am going to use the car ages, and therefore I predict that the modal amount of cars will be one year old. I choose one year, because many people frequently change their cars to keep up with the fashion, and for that reason in my opinion there would be a vast amount of new cars.
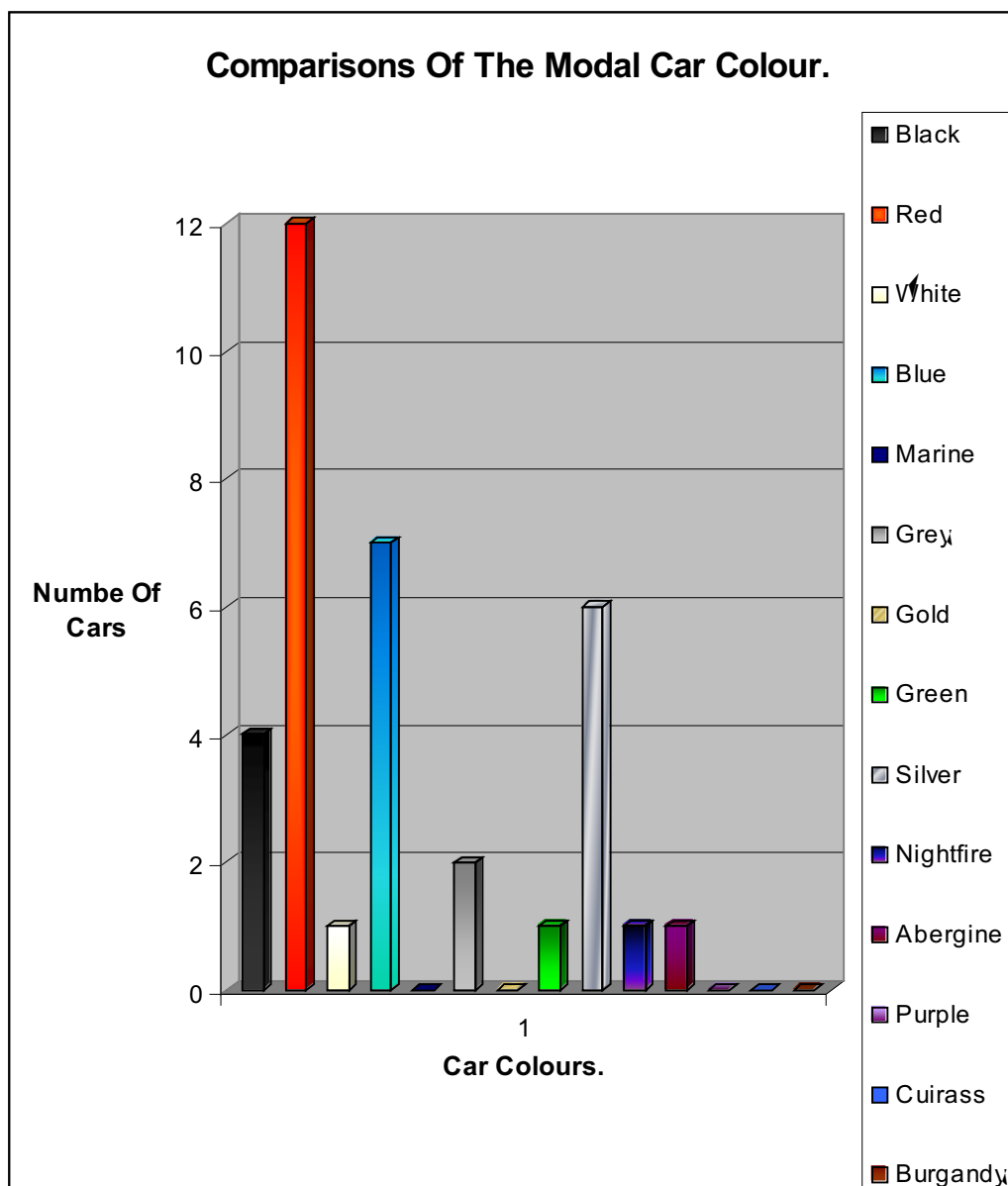
If I were to predict the correlation, I would state that the greater the size of the cars engines the increase in the insurance group. My reasons for this is that the larger and more powerful the engine, the more dangerous and fast the car must therefore be.

Comparisons

Comparisons of most popular car colours.

Black	Red	White	Blue	Marine	Grey	Gold
IIII	IIII IIII II	I	IIII I	0	II	0

Green	Silver	Night fire	Aubergine	Purple	Cuirass	Burgundy
I	IIII II	I	I	0	0	0

Comparisons Of The Modal Car Colour.

Numbe Of Cars

Legend:
- Black
- Red
- White
- Blue
- Marine
- Grey
- Gold
- Green
- Silver
- Nightfire
- Abergine
- Purple
- Cuirass
- Burgandy

1

Car Colours.

The simple bar chart above shows, which colour car, is the most popular and is clearly labelled neatly.

I have begun to realise that I could not detain enough information from collecting the most popular colours and so therefore I am now stopping after this bar chart to go on and investigate a numerical part of the data sheet such as the age of the cars, this will give me better results and information that is more precise.
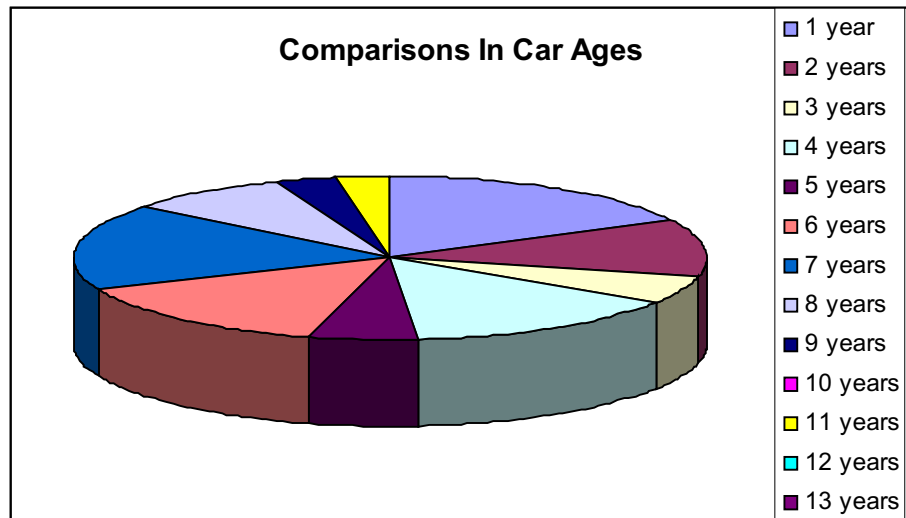
Rebecca Stewart. Mrs Stretch, Maths (Handling Data.) 07/05/2007

Age Comparisons

1	2	3	4	5
IIII I	IIII	II	IIII	II
6	**7**	**8**	**9**	**10**
IIII	IIII I	III	I	0
11	**12**	**13**	**14**	**15**
I	0	0	0	0

Comparisons of The Cars Ages.
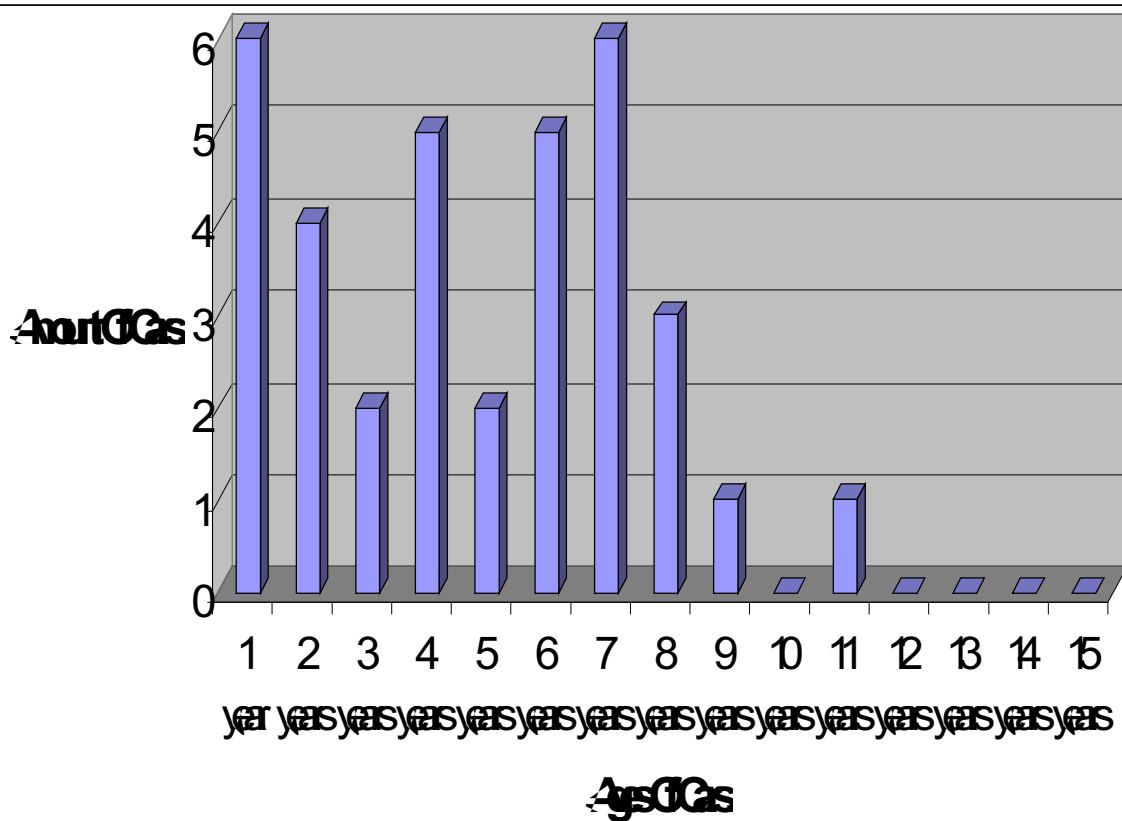
As well as the bar chart above, I have been able to produce a Pie chart which includes all the data I collected from the comparisons of car ages. Furthermore, have calculated the mean, mode, median, and range for this data.

Rebecca Stewart. Mrs Stretch, Maths (Handling Data.) 07/05/2007

Comparisons In Car Ages

- 1 year
- 2 years
- 3 years
- 4 years
- 5 years
- 6 years
- 7 years
- 8 years
- 9 years
- 10 years
- 11 years
- 12 years
- 13 years

This complex pie chart contains the many pieces of data that I have collected and neatly presents them accurately. I will use this as a reference later on in my investigation.

Mode = 1 year old and 7 years old.

Mean =
1 x 6 = 6
2 x 4 = 8
3 x 2 = 6
4 x 5 = 20
5 x 2 = 25
6 x 5 = 30
7 x 6 = 42
8 x 3 = 24
9 x 1 = 9
11 x 1 = 11

6 + 8 + 6 + 20 + 25 + 30 + 42 + 24 + 9 + 11 =

181 / 36 = 5.02

Rounded to the nearest whole number the above would be five. The reason for rounding this to its nearest whole number is that, the years of the cars are shown as whole numbers and it is therefore an easier way of seeing the exact year.

Median = 5 Years old.

Range = 10 (11 – 1)

Correlation

I am now going to go on to produce a far more sophisticated part of investigation, this particular piece will involve a various amount of information, together I have chosen insurance groups and the size of each engine. This will then be plotted onto three different scatter graphs, which will each contain a line of best fit. The reason for a line of best is to make it easy for me to see any anomalies that I may have plotted, and to make it straightforward for me to plot any information with a piece of missing data.

Some of the graphs contain less than 18-plotted marks (or 36 for the high and low insurance groups) the reason for this is that some of the data are the same and therefore land on the same point on my graph, this therefore makes my graphs neat and clear to read.

Correlation

All three of the graphs I have plotted are presented well and fairly accurate the only anomaly I have highlighted is one on the high insurance group. The reason for this anomaly is simply that the car was a classic Bentley; because the Bentley is a very valuable car, this adds to its insurance group, and makes it increase.

As this is my only anomaly, I feel my graphs are extremely accurate for this is not a mistake or a miscalculation but just an unusual piece of data.

Evaluation

If I were to go back and redo this investigation, I would not do the modal car colour for my comparison at all. This is because I feel this is not expressing my full ability and only showing that I can use simple and plain graphs. I would use a numerical piece of data, as this would therefore show my full ability, and I would be able to produce much better pieces of work that are more detailed and recognised in the time given.

During this investigation I have found that the used cars do have a correlation between there engine size and there insurance groups. Furthermore, I obtained enough data to see that the comparisons of the most popular coloured car proved that my hypothesis for compa risons was correct.

All three hypothesises that I predicted were found to be correct and I am obviously extremely satisfied with these findings. I predicted that as I investigate the comparisons of cars I believed the most popular coloured car was red the reason for this is for as I glanced at the first page of data this to me is the modal number. I also predicted for another comparison for a numerical piece of data, which was also correct. I predicted that the modal amount of cars would be one year old.

When I was predicting the correlation, I stated the greater the size of the cars engines the increase in the insurance group. My reasons for this is that the larger and more powerful the engine, the more dangerous and fast the car must therefore be.

I am extremely pleased with how I set out the entire project and I feel that I have worked hard to show as much of my potential ability as I possibly could have done.