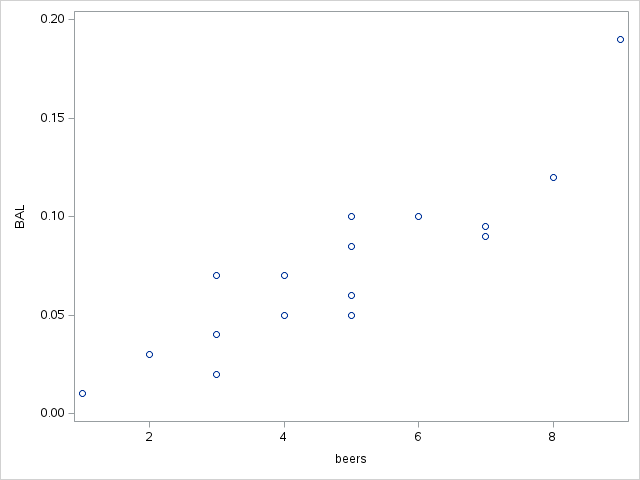
Elijah McLaughlin

1(a)



There appears to be a linear relationship between beers consumed and blood alcohol level (BAL)

1(b)

linear regression equation: BAL = -0.01270 + 0.01796\*beers

1(c)

A graph of a line with blue dots

Description automatically generated with medium confidence

The regression line appears to fit the data on the scatterplot decently well.

1(d)

A diagram of a function

Description automatically generated with medium confidence

Analyzing the residual plots gives better insight into the data. The conditions for the regression model are met, and there is a strong correlation of beers consumed to BAL according to these plots.

2(a)

A diagram of a number of dots

Description automatically generated with medium confidence

The scatterplot reveals a slight negative correlation, with a few candidates for outliers. However, it does appear that the higher the family incomes means the less the financial aid, and the higher the financial aid means the lower the family income.

2(b)

linear regression equation: Aid = 24.31933 + (-0.04307)\*FamInc

2(c)

A diagram of a function

Description automatically generated with medium confidence

The fit diagnostics and residual plots reveal that the conditions are met for the regression model inference as most of the data fits the graphs. However, there are still some outliers that can happen.

3(a)

A graph with numbers and lines

Description automatically generatedA diagram of a graph

Description automatically generated

3(b)

Diameter have a stronger positive linear correlation to volume than height. In the hieght plot, the data is more loosely correlated and appears more random than the data in the diameter plot.

3(c)

A graph of a graph with blue dots

Description automatically generated with medium confidence

3(d)

A diagram of a function

Description automatically generated with medium confidence

Using these graphs, we can visualize how the data fits for diameter being able to predict volume. Because the data fits, we can assume that the adequacy of the model is fairly accurate.

4(a)

Linear regression equation: SRA = -1732400 + 868.00256\*Year

4(b)

A graph showing the growth of a number of years

Description automatically generated

The scatterplot with the linear regression line visualizes the data. In it, we can see that it mostly fits. However, there appears to be two outliers around the years 2003 and 2011, respectively.

4(c)

A diagram of a function

Description automatically generated with medium confidence

Analyzing these two graphs we can see that there is consistently two outliers in each of the graphs.

4(d)

Using the graphs above as evidence, I double checked the data itself to confirm my suspicions.

1997 787.08

1998 968.16

1999 1975.08

2000 3990.00

2001 5455.80

2002 6338.60

**2003 566.25**

2004 7014.90

2005 10500.00

2006 10945.06

2007 12250.80

2008 13035.45

2009 13053.15

2010 14993.60

**2011 952.04**

2012 17349.69

In this, we see that the bolded years (2003, 2011) are much lower than the surrounding years. Because of this and the graphical evidence from 4(b) and 4(c), I can confirm that these two years are sabbatical years and can exclude them and refit the data.

4(e)

A diagram of a function

Description automatically generated with medium confidenceA graph showing the growth of a number of years

Description automatically generated

The data fits much more accurately now. Since it fits with better accuracy, this means that the prediction plots will have higher accuracy as well.