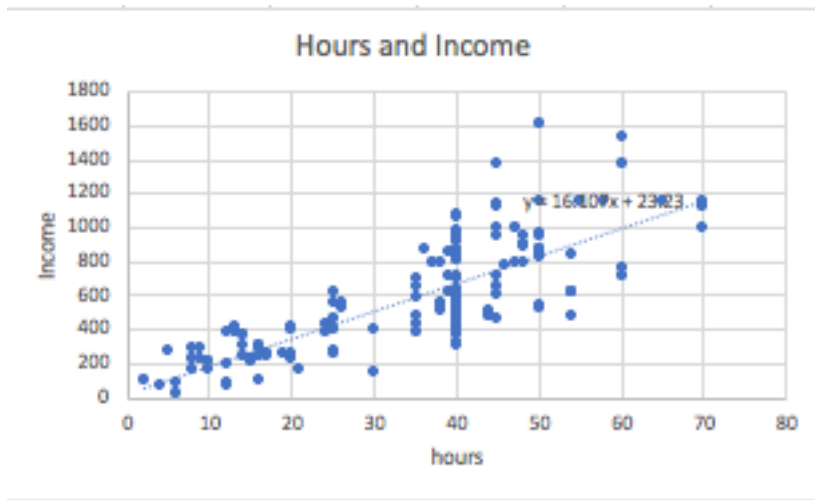


1. Draw a scatterplot of the Hours and Income variables. Explain why you should put Hours on the x-axis.

Hour should be on the x axis as it is an independent variable - a variable which we change or control to observe the effect on y (hours)



2. For the Hours and Income variables:

- Calculate Pearson's correlation and interpret this figure: 0.789045237
Because r is $>$ than 0 we can interpret a positive relationship between the two variables, additionally because r is $>$ than 0.7 this suggest a strong linear relationship between the two variables, hour and income
- Calculate Spearman's correlation, and compare it to Pearson's. Which do you think is most appropriate for your data, and why? 0.82436661
Spearman's correlation is more appropriate as the variables, hour and income, have a stronger non linear relationship (0.82436661) than a linear relationship (0.789045237)
- Calculate the regression line, **write down its equation**, and interpret the slope coefficient.
 $y = 16.107x + 23.23$

For each hour worked income increases by \$16.107

- What proportion of the variation in Income is explained by Hours?
 $R^2 = 0.6226$

This suggest a moderate linear relationship between the change in hours worked and income as R is greater than 0.3 but less than 0.7

- Predict the Income of someone who works 32.5 hours per week, and of someone who works 90 hours per week. Comment on the quality of these predictions.

32.5 hours

$$\text{income} = 16.107(32.5) + 23.23$$

$$\text{income} = \$546.7075$$

$$\text{Income per hour} = \$546.7075 / 32.5$$

$$\text{Income per hour} = \$16.82$$

This seems like a good prediction as this is a logical amount of income to earn per hour

90 hours

$$\text{Income} = 16.107(90) + 23.23$$

income=\$1,472.86

income per hour= \$1,472.86/90

Income per hour=\$16.365111111111111

This seems like a good prediction as this is a logical amount of income to earn per hour

3. Download the spreadsheet containing the Visitor Arrivals data. Identify your allocated country using the last digit of your student ID number. (0 = Australia; 1 = UK, 2 = USA; 3 = Japan; 4 = Korea; 5 = China; 6 = Germany; 7 = Other, 8 = Australia, 9 = UK).

- Draw a time series graph of your data. Comment on any obvious features
 - Trend is generally upwards, became slightly downwards between July 94- Apr98 but returned to upwards trend after this period
 - Peaks suggest seasonality



Calculate the proportion your country represents of the Total arrivals. [Hint: use a formula such as =C2/J2, where your data would be in column C, and the Total in column J.] Draw a time series graph of this proportion, and comment on any obvious features.

- Trend is initially constant between Apr 78 - Jan 87, became upwards after this period until July 94 where the trend became downwards
- Peaks suggest seasonality

