HW 24 Solutions

Problem 1

Below are data on speed and fuel consumption recorded for the British Ford Escort fit in a SLR model to predict fuel consumption (in liters/100 km) as a function of speed (in km/hr).

```
ford_data = data.frame( speed = c(10,20,30,40,50,60,70,80,90,100,110,120,130,140,150),
                        fuel = c(21.0,13.0,10.0,8.0,7.0,5.9,6.3,7.0,7.6,8.3,9.0,9.9,10.8,11.8,12.8))

reg = lm(fuel~speed, data = ford_data)
summary(reg)
```

(a) State the equation for predicting fuel consumption from speed. Is there evidence from the model output to imply that the mean fuel consumption changes with speed? Report the results of the relevant test.

$$\hat{fuel} = 11.062 - 0.016 * speed$$

No, the t-test for the coefficient of speed has a p-value of 0.541 which is too large to indicate that the mean fuel consumption changes with the speed. this seems contradictory to (somewhat) common knowledge. see prob (b).

(b) Use the code below to make a scatter plot of fuel vs speed and overlay the least squares line. Explain why the linear regression fit is not a good summary of this association.

```
library(ggplot2)
ggplot(ford_data, aes(x=speed, y=fuel)) +
  geom_point() +
  geom_smooth(method = "lm", se=FALSE)
```

The relationship between speed and fuel is non-linear, therefore we should consider a transformation of either fuel or speed before fitting a linear model.

Problem 2

Create a flow chart (or outline) that helps someone determine which methods to use (of those that we've covered) to answer a research question about some arbitrary data set.

Graded for completion.