

**STAT40720 Intro. to Data Analytics**

**Assignment 4**



**Name** Kevin O’Connell

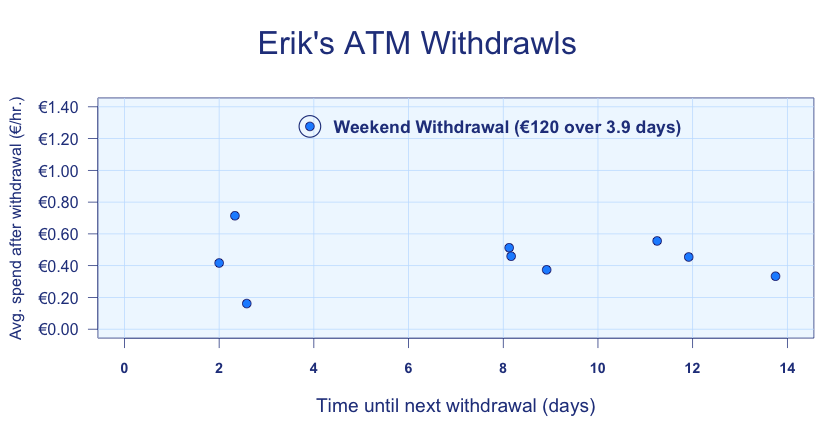
**Student ID** ########

**Submission Date** 16ndth November 2015

**Question 1**

To complete the analysis, a plot of the average spend per hour versus time to next withdrawal was computed.

**(a) Weekend Withdrawal**



Erik’s withdrawal prior to the weekend away is likely to have the highest average spend over a period of 3 to 4 days. The most likely candidate is observation 5, where €120 was withdrawn and spent over 94 hours.

**(b) Sick Withdrawal**



Erik’s withdrawal prior to getting sick and confined to bed is likely to have the lowest average spend over a relatively short period of 2 to 3 days. The most likely candidate is observation 2, where €10 was withdrawn and spent over 62 hours.

**Question 2**

1. **The equation of the regression line**

To find the value of *a*, the mean value of x and y must be found.

The above regression shows from the figure *b*, that for each minute after 8:08am the teacher leaves, the journey time increases by 1.84 minutes.

The value of *a* shows and extrapolation back to 8:00am from this data would yield a journey time of 2.6144 minutes. This is, of course, not practical, but it can be used to show the expected minimum journey time for an 8:08am departure is given by:

Therefore, the expected minimum travel time is 17.3 minutes.

1. **8:15am Departure Time**

For an 8:15am departure, the expected journey duration is 30.2 minutes.

1. **8:15am Departure Time 95% Confidence Interval**

1. **Interpret (c)**

The value from (c) states that we can be 95% confident the mean population value of journey time for an 8:15am departure lies within 2.82 minutes of 30.16 minutes.

1. **95% Prediction Interval**

1. **8:30am Prediction Feasibility**

Although a prediction of travel time for a departure time at 8:30am is physically possible, there is no evidence or observations to suggest this model holds for this departure. Therefore, this model should not be used to perform such a prediction.

**Question 3**

Case 1 – Removing case “+” should yield r = 0.285

Case 2 – Removing case “x” should yield r = 0.720

Case 3 – Removing case “+” and “x” should yield r = 0.925