Data Science – Assig #3: Data Exploration Using Multiple Data Sets **(25 Marks)**

**Problem:** The city of Cambridge is trying to promote tourism in their busy month of August by renting bicycles to tourist to cycle to the various attractions. Some of these tourist attractions include: museums, art galleries, parks, hiking and cycling trails, that are located along very scenic routes. For tourist that are unable to cycle, such as senior citizens, the city plans to utilize hop-on/hop-off buses to traverse the various attractions. These buses will also be utilized by cyclist when the routes become too difficult for cycling. To schedule this bus service adequately for the cyclist and the non-cycling tourist, the city requires you examine the weather (weatherData) and cyclist (Eco-Totem\_Broadway\_Bicycle\_Count) datasets to determine if temperature affects cycling in the city. To accomplish these tasks, the city has required you to do the following:

1. Plot and analyze per day for the month of August 2017 the: i) total cyclist, ii) iii) Average cyclist for the month, and iii) days when there were high temperatures i.e. above 60 (F). Superimpose these plots (i.e. all plots on the same graph, using different colours, so that trends can be easily spotted). From the plot discuss if high temperatures affected cycling in August 2017

Submit this assignment as a pdf document with the following sections:

1. **Problem:** The problem as stated above.
2. **Methodology:** which outlines how you are going to solve each task, showing: a) the major steps that has to be done to get the data, and b) flow chart(s) indicating how these major steps would be performed. **(12 marks)**
3. **Results:** Plots of the data obtained with a discussion of any trends observed. **(4 marks)**
4. **Conclusion:** answering the question above if temperature affected cycling in the city of Cambridge in August 2017. **(4 marks)**
5. **Appendix:** containing a copy of your code with comments corresponding to the major sections outlined in your methodology **(5 marks)**