EGH404 Portfolio (Part 1)

The EGH404 Portfolio is based on an individual dataset that can be downloaded at:

www.egh404.com

Please note that the dataset is too big to be loaded into Excel, and while it opens, it will strip about 1/3 of the data.

The structure of the dataset is as follows:

Column 1: *Point Location in WGS84*
Column 2: *Temperature in degree Celsius*
Column 3: *Rainfall in mm*
Column 4: *Number of people in the location*
Column 5: *Sensor value A*
Column 6: *Sensor value B*

Tasks:

1. Import the dataset into Matlab
2. Remove any outliers using three standard deviations from the mean as cut-off
3. Provide a table with the minimum, maximum, and average for columns 2, 3, 4, 5 and 6
4. Provide a box and whiskers plot for column 3 and 4
5. Create a scatter plot to look for a correlation between column 3 and 6
6. Create a histogram for sensor value A using 10 bins of equal size
7. Determine the average number of people for all locations and plot the value for the top 100 locations with the most people as a descending line graph
8. Determine the minimum and maximum temperature and plot both values for the locations from 6 as line graphs
9. Is there a correlation between temperature and rainfall? Provide a plot and written answer (3 sentences max, let the plot do the talking)
10. Is there a correlation between Sensor value B and the rainfall? Provide a plot and written answer (3 sentences max, let the plot do the talking)
11. What is the expected temperature when rainfall is between 15 and 25mm?

Submission:

For each tasks provide the following:

- Verbal explanation of what you are doing
- Matlab code
- Result as requested in task (i.e. plot, table…)