**BAN 5733**

**Individual Exercise 1**

**10 Points**

**Data Set Information:**

You will work with Forest Fires data set that contains 13 variables and over 500 observations. The variables in the data set are shown below with their appropriate description.

**Attribute Information:**

|  |  |
| --- | --- |
| **Name** | **Description** |
| **X** | x-axis spatial coordinate within the Montesinho park map: 1 to 9 |
| **Y** | y-axis spatial coordinate within the Montesinho park map: 2 to 9 |
| **month** | Month of the year: “jan” to “dec” |
| **day** | Day of the week: “mon” to “sun” |
| **ffmc** | FFMC index from the FWI system: 18.7 to 96.20 |
| **dmc** | DMC index from the FWI system: 1.1 to 291.3 |
| **dc** | DC index from the FWI system: 7.9 to 860.6 |
| **isi** | ISI index from the FWI system: 0.0 to 56.10 |
| **temp** | Temperature in Celsius degrees: 2.2 to 33.30 |
| **rh** | relative humidity in %: 15.0 to 100 |
| **wind** | wind speed in km/h: 0.40 to 9.40 |
| **rain** | outside rain in mm/m2 : 0.0 to 6.4 |
| **area** | the burned area of the forest (in ha): 0.00 to 1090.84 |

1. Explore the distribution of the **RH** via histogram and moments.
   1. Overlay a Normal curve on the histogram.
   2. Report the following information for the **RH** variable: mean, median, interquartile range, kurtosis and skewness values (**2 points**)
   3. Include a screenshot or copy of the histogram with the superimposed Normal curve and provide comments about the distribution of the variable. (**2 points**)
   4. Create a Normal Probability plot for the variable **RH** and then interpret this plot. (**2 points**)
2. Relationship between **2 variables**.
   1. Explore the relationship between **AREA** (Y) and **TEMP** (X) with a Fit Y by X graph or graph builder. Does there appear to be a relationship between these two variables? Write a few lines about what you observe in this data. **(2 points)**
   2. Explore the relationship between **AREA** and **DAY** with a Fit Y by X graph and the grouped means. Does there appear to be a relationship between these two variables write a few lines about what you observe in this data. **(2 points)**

Deliverables (please follow these instructions):

* As you complete the exercise, create a report in Microsoft Word. In this report, answer the questions in the exercise description.
* Make sure you comment or explain and not just provide snapshots of data.
* Limit your report to no more than 5 pages including tables and diagrams.
* Copy and paste or screen shot supporting tables/diagrams as needed to justify any of your answer. You may need to shrink your table/ diagrams but please ensure they are readable.
* Make sure you print your name, student ID#, student email on the cover page of the report and turn-in the report as communicated by your instructor.
* Please also put a running header/footer with your name, on each page of your exercise solution report.

Failure to follow these instructions will result in deduction of points