Department of Computer Technology and Information Systems

CTIS 365: Applied Data Analysis

Semester: Fall 2019-2020

Lab Guide #3 – Week 04

OBJECTIVES: Frequency Distribution Tables & Graphs, Histogram, Stem and Leaf Plot and Bar Plot
Instructor: Dr. Seyid Amjad ALI **Assistant:** Burcu ALPER

Q1. In **R** there is a dataset named **faithful** that includes two columns consists of Waiting time between eruptions and the duration of the eruption for the Old Faithful geyser in Yellowstone National Park, Wyoming, USA.

You have been asked to construct a grouped frequency distribution table using an interval width of 5 points on eruptions data and find the max. value for the intervals' frequency.

Example run:

```
duration.cut
[1.5,2) [2,2.5) [2.5,3) [3,3.5) [3.5,4) [4,4.5) [4.5,5)
     51      41       5       7      30      73      61
[5,5.5)
      4

[1] "Maximum of the intervals' frequency is  73"
```

Q2. A dataset named **"ctiesTemperature.txt"** is given to you in the form of a text file. First read the data from the file and then construct a frequency distribution table for the temperature values you have read from the file. Include columns for cumulative frequency, proportion, proportion percentage, cumulative proportion and cumulative proportion percentage to the end of this frequency distribution table (see Figure 1).

x	freq	cum.freq	prop	prop.per	cum.prop	cum.prop.perc
37.9	1	1	0.02	2	0.02	2
38.4	1	2	0.02	2	0.04	4
38.8	1	3	0.02	2	0.06	6
39.3	2	5	0.03	3	0.09	9
39.4	2	7	0.03	3	0.12	12
39.6	1	8	0.02	2	0.14	14

Figure 1

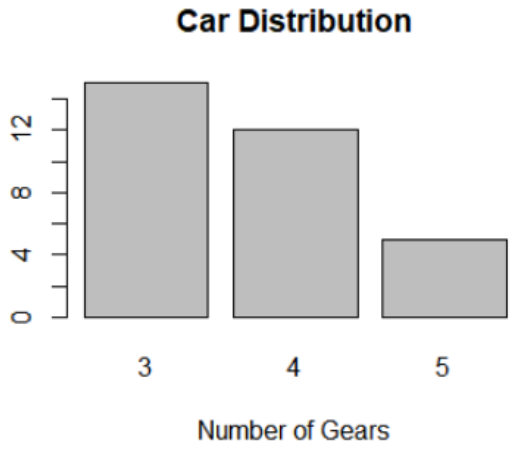
Q3. By using the same data from Q1.

- What is the percentile rank for X = 39.4?
- What is the percentile rank for X = 44.6?
- What is the temperature in the 20th percentile?
- What is the temperature in the 38th percentile?

Q4. Draw a stem and leaf graph for the temperatures in the previous question.

```
37 | 9
38 | 48
39 | 33446889
40 | 224566678899999
41 | 00113445566777788899 99
42 | 00136678
43 | 01
44 | 06
```

Q5. Use the built-in data set **mtcars** and draw a bar plot like the below graph.



Q6. Use the Chick Weights from the **ChickWeight** built-in dataset to create a histogram.