STAT 8010-006 Exam I, Summer 2023

May 31, 2023

Name:_____

Directions

1. Show your work on ALL questions. Unsupported work will NOT receive full credit.
2. Decimal answers should be exact, or to exactly 4 significant digits.
3. Please write legibly. If I cannot read your writing, NO credit will be given.
4. Put your work into a **single file** and upload it to Canvas before 11:59pm ET.

Use your time wisely. Good Luck!!!

Problem	Points Possible	Points Earned
1	30	
2	20	
3	15	
4	10	
5	15	
6	10	
Total	100	

Problem 1

(30 points: 6 points each)

The data file, `NY_AirQuality.csv`, contains the daily readings of the following variables for May 1, 1973 to September 30, 1973:

- `Ozone`: Mean ozone in parts per billion (ppm) from 1300 to 1500 hours at Roosevelt Island
- `Temp`: Maximum daily temperature in degrees Fahrenheit at La Guardia Airport.
- `Month`
- `Day`

Source: The data were obtained from the New York State Department of Conservation (ozone data) and the National Weather Service (meteorological data).

(a) Is this study *experimental* or *observational*?

(b) Summarize `Ozone` and `Temp` graphically (separately) and describe their shapes.

(c) Compute the mean, median, range, standard deviation and IQR for `Ozone` and `Temp`.

(d) Make a scatterplot of `Ozone` and `Temp` and comment on their potential relationship.

(e) Create a time series plot of `Temp` and provide a concise summary of your observations and conclusions.

Problem 2

(20 points)

This data set, `Heart Rate.csv`, provides heart rates of male and female runners and generally sedentary participants following 6 minutes exercise. (**Data source**: JASP built-in data.)

(a) How many units and variables in the dataset? **(5 points)**

(b) Create a 2×2 frequency table of `Gender` and `Group` and infer whether this is an *experimental* or *observational* study? **(5 points)**

(c) Make side-by-side boxplot of `Heart Rate` by `Gender` and by `Group`, respectively. Summarize your findings based on these two side-by-side boxplots **(10 points)**

Problem 3

(15 points)

Use the data set `MaritalGender.csv` that records the marital status and gender to answer the following questions.

(a) What percent of the men were single? **(5 points)**

(b) Plot `Gender` and `MaritalStatus` together and describe your findings. **(10 points)**

Problem 4

(10 points)

A recent survey indicated that 60% of mobile devices are iPhones. Suppose you randomly ask 100 people if they use an iPhone. Assume all individuals are independent, let X be the number of people asked that use an iPhone.

(a) What are the distribution and parameters of X? **(3 points)**

(b) What is the standard deviation of X? **(3 points)**

(c) Given no more than 80 people surveyed use an iPhone, what is the probability exactly 60 people asked use an iPhone? **(4 points)**

Problem 5

(15 points, 5 points each)

Denver Downs has a large pumpkin patch, where the weight of the pumpkins follows a normal distribution with an average of 13.5 pounds, and a variance of 9. Each pumpkin's weight is independent of all other pumpkins.

(a) What is the probability that a randomly selected pumpkin weighs over 14 pounds?

(b) What is the probability that a randomly selected pumpkin weighs over 15 pounds given that the selected pumpkin weighs over 12 pounds?

(c) Find the cutoff for the top 2.5% of pumpkin weights at Denver Downs.

Problem 6

(10 points)

Explain clearly why switching would increase the chance of winning a car in the Monthy Hall problem.