Spring 2020 **Stat 305 (Section 4)** **Quiz 1**

Name: Solution

Total points for the exam is 50. Points for individual questions are given at the beginning of each problem. Show all your calculations clearly. Put final answers in the box at the right (except for the diagrams!).

1. [8+6+6+6=26 points]

The following are the closing prices of Facebook from Jan 21, 2020 to Jan 31, 2020:

221.44, 221.32, 218.76, 217.94, 214.87, 217.79, 223.23, 209.53, 201.91.

Sorted data : 201.91, 209.53, 214.87, 217.79, 217.94, 218.76, 221.32, 221.44, 223.23

(a) Find the 0.82 quantile, the median, first and third quartile for the above data.

$Q(0.82) =$ 221.4256

$Med. =$ 217.94

1st Quart.= 213.535

3rd Quart.= 221.35

$\frac{i'-0.5}{9} = 0.82 \Rightarrow i' = 7.88$

$\Rightarrow Q(0.82) = 0.88 \times 221.44 + (1-0.88) \times 221.32$

$= 221.4256$

$Q(0.5) = 217.94$

$\frac{i'-0.5}{9} = 0.25 \Rightarrow i' = 2.75.$

$\Rightarrow Q(0.25) = 0.75 \times 214.87 + 0.25 \times 209.53 = 213.535$

$\frac{i'-0.5}{9} = 0.75 \Rightarrow i' = 7.25$

$\Rightarrow Q(0.75) = 0.25 \times 221.44 + 0.75 \times 221.32 = 221.35$

(b) Give the coordinates (on a regular graph paper) of the upper right and lower left point that would appear on a normal plot of the data.

upper right point = (201.91 , −1.59)

lower left point = (223.23 , 1.59)

lower left: $\frac{1-0.5}{9} \approx 0.056$

$Q_{SN}(0.056) \approx 4.9\,(0.056^{0.14} - (1-0.056)^{0.14}) = -1.59$

$\Rightarrow$ Lower left point : $(201.91, -1.59)$

upper right : $\frac{9-0.5}{9} \approx 0.944$

$Q_{SN}(0.94) \approx 4.9\,(0.944^{0.14} - (1-0.944)^{0.14}) = 1.59$

$\Rightarrow$ upper right point : $(223.23, 1.59)$

(c) Draw a boxplot for this data. Carefully label numbers on the plot

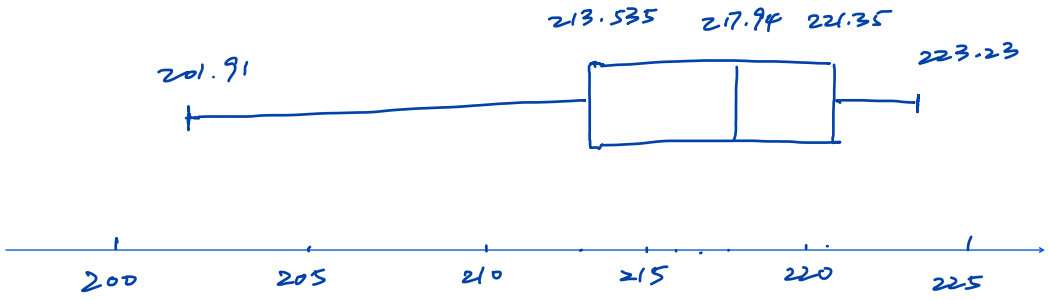
$IQR = Q(0.75) - Q(0.25) = 221.35 - 213.535 = 7.815$

$Q(0.25) - 1.5\,IQR = 213.535 - 1.5 \times 7.815 = 201.8125$

$Q(0.75) + 1.5\,IQR = 221.35 + 1.5 \times 7.815 = 233.0725$

$\Rightarrow$ No outlier.

Therefore the Boxplot:



(d) Find the sample mean and standard deviation for this data. Show calculations.

$\bar{x} =$	216.31
$s =$	6.76

$$\bar{x} = \frac{1}{9}\sum_{i=1}^{9} x_i = \frac{1}{9} \times (201.91 + 209.53 + \cdots + 223.23) = 216.31$$

$$s = \sqrt{\frac{1}{9-1}\sum_{i=1}^{9}(x_i - \bar{x})^2}$$

$$= \sqrt{\frac{1}{8}\left[(201.91 - 216.31)^2 + \cdots + (223.23 - 216.31)^2\right]}$$

$$= 6.76$$

2. [8×3=24 points]

For each of the following questions, choose only one (the best) answer. No credit will be given if more than one is chosen.

(a) Which of the following is the best numerical summary that is insensitive to outliers if we want to assess the precision of a measurement system?

precision is associated with variation in repeated measurements. so a numerical summary for variation that is insensitive to outliers is IQR.

(A) sample variance (B) IQR (C) sample mean (D) median **B**

(b) Which of the following best describes the methods for handling extraneous variables:

(A) blocking and replication (B) randomization and replication
(C) randomization and blocking (D) randomization, blocking, and replication **C**

(c) For a complete factorial study with 8 factors, each with 3 levels, the number of observations is at least

(A) 1024 (B) 6561 (C) 512 (D) none of these **B**

(d) For a 3 × 3 full factorial study with two factors A and B, where A has three levels (low, medium and high) and B has three levels (low, medium, and high), the nine experimental runs are labeled as:

No. 1: A low B low, No. 2: A low B medium, No. 3: A low B high,
No. 4: A medium B low, No. 5: A medium B medium, No. 6: A medium B high,
No. 7: A high B low, No. 8: A high B medium, and No. 9: A high B high.

Based on the following random digits

97437 52922 80739 59178 50628

Which experiment should be done last?

(A) No. 6 (B) No. 7 (C) No. 8 (D) No. 9 **A**

(e) Based on the following random digits

61017 51652 40915 94696 67843 58009

the second widget selected from 99 widgets labeled 1,2,. . . ,99 is

(A) 51 (B) 10 (C) 1 (D) 75 **C**

(f) In a series of experiments to study the purity of a chemical product, the effect of reactant A on the purity was of primary interest and three levels of A were used in the experiments. We also know these experiments are done in two different labs. The variable lab is a

lab is an extraneous variable that could affect the response in the experiments. so it should be a blocking variable.

(A) concomitant variable (B) controlled variable
(C) blocking variable (D) experimental variable **C**

(g) What is the relationship between the median and mean of the data set that is exponentially distributed and skewed to right?

(A) median < mean (B) median > mean
(C) median = mean (D) all above are possible **A**

(h) If one wished to assess if a data set is normally distributed, one should use

(A) a dot diagram (B) a histogram (C) a boxplot (D a normal Q-Q plot **D**