Name ___KEY___

Section _____

**Instructions:** This exam is in two parts: Part I is to be completed partly at home using the materials posted on Blackboard for Part I and you will answer questions about that work in class below; Part II is to be completed entirely in class. You may not use cell phones, and you may only access internet resources you are specifically directed to use. You may access your data file for Part I of the exam in Blackboard. You may access the data files posted to Blackboard for the Exam part II. Be sure you are using the data file that matches the exam version you are given.

Part I:

The following questions apply to problem #1 from Part I:

1. Report the mean and standard deviation of Experience. (6 points)

    Mean = 15.07
    St.dev = 8.69

2. Report the five-number summary of Experience. (8 points)

    Min = 1          Q3 = 21
    Q1 = 9
    Median = 14      Max = 43

3. Use this information to determine if there are any outliers in the data. If so, what are they? (5 points)

    There is one outlier. Upper fence is 39 but Max is 43.

4. Describe the boxplot of Experience. Does it appear to be skewed left, skewed right or approximately symmetric? Explain your reasoning. (4 points)

    Skewed right (larger values)
    median is smaller than mean

5.  What does the x in the center of the box signify?  (4 points)

*location of the mean*

6.  Consider your histograms of Age.  Describe the shape of the distribution.  Possible descriptions include one or more of the following: skewed left, skewed right, symmetric, normal, monomodal, bimodal, multimodal, uniform, or none of these.  Do the two graphs appear to tell the same story about the data or a different one?  Explain your reasoning.  (10 points)

*they both appear to be bimodal w/ no strong skew*

*They appear similar*

7.  Describe your scatterplot of Experience vs. Annual Salary.  Does the data appear to have a trend?  If so, is it positive or negative?  (5 points)

*there does appear to be a positive trend (as experience increases, so does salary) the relationship appears strong and linear*

The following question relates to problem #2 from Part I:

8.  Describe your time series graph.  Do you notice any trends in either Manufacturing or Retail over time?  Do the two variables appear to be related to each other?  Explain.  (8 points)

*they appear stable, dip or increase at the same time, but the gap between them is getting wider*

*(answers may vary)*

The following questions relate to problem #3 from Part I:

9. Report the mean and median percentage on-time arrive for BWI. Label each one clearly. (4 points)

Mean BWI = 79.74

Median BWI = 80.37

10. Based on your graphs, which airport has the better long-term on-time arrival percentage? Explain your reasoning based on the data. Can you think of a real-world reason as to why this might be so? (10 points)

mean and median are better for BWI
but box plot shows that variability is also
greater and BWI has more extreme low values
                                        CLE is better at bad weather
(answers will vary) but gets more of it

11. Which graphs did you choose for this data and why? Were there some graphs types you rejected? (7 points)

Cluster column for mean/median comparison
box plot for distribution comparisons

answers will vary

Calculations in Excel: (1) 48 points, (2) 15 points, (3) 30 points

Part 2:

12. Order the seven steps of the modeling process in the appropriate order.  List the sequence in the column to the right. (14 points)

| Step | Order |
|---|---|
| Present the results to the organization. | 6 |
| Verify the model. | 4 |
| Implement model and update it over time. | 7 |
| Define the problem. | 1 |
| Develop the model. | 3 |
| Collect and summarize data. | 2 |
| Select one or more suitable decisions. | 5 |

13. Classify the following variables: (33 points)

| Variable | Categorical | Quantitative | Discrete | Continuous | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|---|---|---|---|
| Test Grade (Percentage) | | X | | X | | | | X |
| City | X | | | | X | | | |
| Day of the Month | | X | X | | | | X | |
| Volume (3D) | | X | | X | | | | X |
| Cost ($) | | X | | X | | | | X |
| Fruit | X | | | | X | | | |
| Nationality | X | | | | X | | | |
| Depth | | X | | X | | | | X |
| Birth Rate | | X | | X | | | | X |
| Race | X | | | | X | | | |
| Age (Years) | | X | X | | | | | X |

14. What does it mean if we say a value represents the 94th percentile? (5 points)

94% of the data is below That value

15. With symmetric/bell-shaped distributions, approximately what percent of the observations are within one standard deviation of the mean? (5 points)

**68%**

16. Expressed in percentiles, what does the third quartile represent? (5 points)

**75th percentile**

17. A screen capture of an Excel spreadsheet is shown below. A company bases their current salary for a particular job on a simple linear formula based on a base (entry-level salary) and the number of years of experience: i.e. $y = mx + b$ where $y$ is salary, $m$ is the increase, and $b$ is the base salary. The annual increase for each year of experience and the base salary are provided on the screen. What would you need to type in Cell B3 to calculate the expected salary for the given number of years' experience, so that you can copy the formula into cells B4 and B5 without having to update any cell references manually? Write the formula below. (10 points)

B3

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Years Experience | Salary | | | Each Year of Experience | $4,321 |
| 2 | 7 | $66,996 | | | Base Salary | $36,749 |
| 3 | 14 | | | | | |
| 4 | 25 | | | | | |
| 5 | 32 | | | | | |
| 6 | | | | | | |

$$= F\$1 * A3 + F\$2$$

18. In the data file for the exam, use the data set on the sheet #18 to answer the following questions using the Exercise data (in minutes).
    a. Some of the data in the column is missing. How does Excel treat those blank cells? (5 points)

    *it ignores it*

    b. Find and report the first quartile. (5 points)

    *4.00*

    c. Find and report the maximum. (5 points)

    *94.00*

    d. Find and report the sample variance and the population variance. How do they differ? (7 points)

    *Sample: 620.52*
    *Population: 619.87 ← Smaller by a bit*

    e. Find the 69th percentile. (5 points)

    *38.00*

19. Create a blank sheet and label it #19. Enter the following values into Excel in a table and use it to calculate the formula. Round your answer to 4 decimal places. (12 points)

| $X_1$ | $X_2$ | $Z$ | $S$ | $N$ |
|---|---|---|---|---|
| 58 | 73 | 1.96 | 7.4 | 104 |

$$\frac{Z(X_1 - X_2)}{S/\sqrt{N-1}}$$

*— 40.3213*

Upload your completed Excel files (**both of them!**) to the Exam #1 **to be graded** submission box in Blackboard and submit your completed paper exam to your instructor. You may not modify anything once the exam is submitted.