Assignment 1

STAT 230
UAEU

Part I: Short answer type questions.

Must Review: Review All the Quiz Problems.

1. Let us consider a simple linear regression betwen the response variable Y and the covariate X. The fitted regression line is $\hat{Y} = 3 + 2.5X$. Then idensity which if the following statement is correct?

   Ans:
   - The correlation between X and Y must be positive.
   - The correlation between X and Y must be negative.
   - The correlation between X and Y must be zero.
   - From the provided information, it is not possible to determine whether the correlation is positive, negative or zero.

2. **Statement:** If correlation between two continuous variables is zero then the variables must be unreleated.

   Ans: TURE FALSE

3. **Statement:** If two continuous variables are unrelated then the correlation between them must be zero.

   Ans: TURE FALSE

4. Consider the scatter plot and the corresponding regression line for the 'Adult Female Literacy' (x) and on 'Total Fertility Rate' (y) based on a sample of 35 countries worldwide.

   Based on the information provided on the plot, what can we tell about the correlation between the two two variables?

   Ans: 0.8952 -0.8952 0.8014 Can not be determined

5. If the Gini coefficient for Estonia income is 0.348 in 2015 and the Gini coefficient for Macedonia income is 0.434 in the same year, which country has the least unequal distribution of income among the two?

Ans: ☐ Macedonia ☐ Estonia ☐ Can be determined

6. Consider the Lorenz curves of wealth inequality for Japan and USA. Based on the plot below, which country has less wealth inequality among the two?

Ans: ☐ Japan ☐ USA ☐ Can be determined

7. Consider the Lorenz curves of wealth inequality for Brazil and Sweden. Based on the plot below, which country has less wealth inequality among the two?

Ans: ☐ Brazil ☐ Sweden ☐ Can be determined

8. Consider the histogram of 'Infant Mortality Rate'

Based on the provided histogram, what can we say about the type of the histogram of the variable? There are multiple corect answers, select all the correct options.

Ans:
☐ Skewed Right ☐ Skewed Left ☐ Skewed Positive
☐ Skewed Negative ☐ Mean is larger than Median
☐ Mean is smaller than Median

9. Consider the histogram of 'Life Expectancy at birth'

Based on the provided histogram, what can we say about the type of the histogram of the variable? There are multiple corect answers, select all the correct options.

Ans:
☐ Skewed Right ☐ Skewed Left ☐ Skewed Positive
☐ Skewed Negative ☐ Mean is larger than Median
☐ Mean is smaller than Median

10. According to the 2007 World Bank data, 5.21at $1.25 a day (PPP). When 5000 individuals were selected from Brazil, 6% were living below the $1.25 poverty line. What does the value 5.21% refer to?

Ans: ☐ Statistic ☐ Parameter ☐ Population ☐ Sample

11. According to the 2007 World Bank data, 5.21at $1.25 a day (PPP). When 5000 individuals were selected from Brazil, 6% were living below the $1.25 poverty line. What does the the value 6% refer to?

Ans: ☐ Statistic ☐ Parameter ☐ Population ☐ Sample

12. In the 2009, American Community Survey conducted to determine, among other things, the family size, 30366 American households were randomly sampled. The sample mean and the standard deviation of the number of persons in family were 3.11 and 1.723. Compute a 90% confidence the average family size in USA.

Ans:

13. According to the 2007 UNDP data, the mean and the standard deviation of the GDP index for 182 countries are 0.6925 and 0.2164 respectively. Compute a 99% confidence interval for the average GDP index over the last decade .

Ans:

14. Consider the following summary of two continuous variables X, Y of a dataset.

Sample Mean of X	Sample Mean of Y	Sample Standard Deviation of X	Sample Standard Deviation of Y	Correlation between X, and Y
45	30	5	7.5	.9

Finally we consider a simple linear regression model: $\hat{Y} = a + bX$ where a and b denotes the intercept and the slope corrrespondingly. Estimate a, b and writedown the equation of the fitted line.

Ans:

Part II: Descriptive Problems.

1. A car manufacturer provides cars with the following different variations:
   - Manual or automatic transmission
   - Three different stereo systems
   - Four possible exterior colors

   How many different types of car the manufacturer sells?

2. In a popular lottery game, a person may pick any 5 numbers from $\{1,2,\ldots,52\}$ for her ticket. How many different groups of 5 numbers can be chosen from the 52 **if the repeated selection of the numbers is NOT allowed, however the "order in which the numbers are selected are important" ?**

3. How many different 'words' (sequences of letters that may, or may not be a dictionary word) can be constructed by permuting the letters of the word "ALMADAM" ?

4. Consider 8 tosses of a coin. For example two typical sequences of outcomes that are considered different is 'TTHTTTTT', and 'HHTHTTTT'.
   (a) What is the total number of different outcomes from the experiment?
   (b) How many different ways there can be exactly 3 heads?

5. If a dice is thrown 8 times and all the the 8 numbers (8-tuple) that appear are recorded.
   (a) What is the total number of distinct possibilities of the 8 numbers (distinct 8-tuples) that can occur?
   (b) In how many such 8-tuples, there will be exactly three 6?
   (c) In how many such 8-tuples, there will be exactly three 6 and two 1 ?

6. One of the sections in the class STAT230 has 10 students. Assume the birthday of a student can be any one of days out of 365 days in a year. We say that the two students share a (same) birthday if they are born on the same day and the same month (for example two students are born on 12^{th} January). Answer the following questions.
   (a) How many different possibilities are there such that exactly five students have the same birthday?
   (b) How many different possibilities are there so that all the birthdays are on different days of the year.