AS EXTENDED ASSIGNMENT

BY

SUDEEP KUMAR DAS

PGPDSBA.O. SEP22.B

**TABLE OF CONTENTS:**

List of Figures\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 04

Problem 1\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 05

**1.1. For this data, construct the following contingency tables (Keep Gender as row variable)**

**1.1.1. Gender and Major\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 06**

**1.1.2. Gender and Grad Intention\_\_\_\_\_\_\_\_\_\_\_\_ 07**

**1.1.3. Gender and Employment\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 07**

**1.1.4. Gender and Computer\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 07**

**1.2. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question: \_\_\_\_\_\_\_\_ 07**

**1.2.1. What is the probability that a randomly selected CMSU student will be male? \_\_\_07**

**1.2.2. What is the probability that a randomly selected CMSU student will be female?\_\_ 08**

**1.3. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08**

**1.3.1. Find the conditional probability of different majors among the male students in CMSU.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08**

**1.3.2 Find the conditional probability of different majors among the female students of CMSU.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08**

**1.4. Assume that the sample is a representative of the population of CMSU. Based on the data, answer the following question:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08**

**1.4.1. Find the probability That a randomly chosen student is a male and intends to graduate. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08**

**1.4.2 Find the probability that a randomly selected student is a female and does NOT have a laptop. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09**

**1.5. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09**

**1.5.1. Find the probability that a randomly chosen student is a male or has full-time employment?\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09**

**1.5.2. Find the conditional probability that given a female student is randomly chosen, she is majoring in international business or management. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 10**

**1.6.  Construct a contingency table of Gender and Intent to Graduate at 2 levels (Yes/No). The Undecided students are not considered now and the table is a 2x2 table. Do you think the graduate intention and being female are independent events? \_\_\_\_\_ 10**

**1.7. Note that there are four numerical (continuous) variables in the data set, GPA, Salary, Spending, and Text Messages.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 11**

**1.7.1. If a student is chosen randomly, what is the probability that his/her GPA is less than 3? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 11**

**1.7.2. Find the conditional probability that a randomly selected male earns 50 or more. Find the conditional probability that a randomly selected female earns 50 or more. \_\_\_ 11**

**1.8. Note that there are four numerical (continuous) variables in the data set, GPA, Salary, Spending, and Text Messages. For each of them, comment whether they follow a normal distribution. Write a note summarizing your conclusions. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 11**

Problem 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 13

**2.1 Do you think there is evidence that means moisture contents in both types of shingles are within the permissible limits? State your conclusions clearly showing all steps. \_\_\_ 14**

**2.2 Do you think that the population mean for shingles A and B are equal? Form the hypothesis and conduct the test of the hypothesis. What assumption do you need to check before the test for equality of means is performed? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 15**

Problem 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 16

* 1. **State the null and the alternate hypothesis for conducting one-way ANOVA for both Education and Occupation individually. \_\_\_\_\_\_\_ 17**
  2. **Perform a one-way ANOVA on Salary with respect to Education. State whether the null hypothesis is accepted or rejected based on the ANOVA results. \_\_\_\_\_\_\_\_ 17**

**3.3 Perform a one-way ANOVA on Salary with respect to Occupation. State whether the null hypothesis is accepted or rejected based on the ANOVA results. \_\_\_\_\_\_\_\_\_\_\_\_\_ 17**

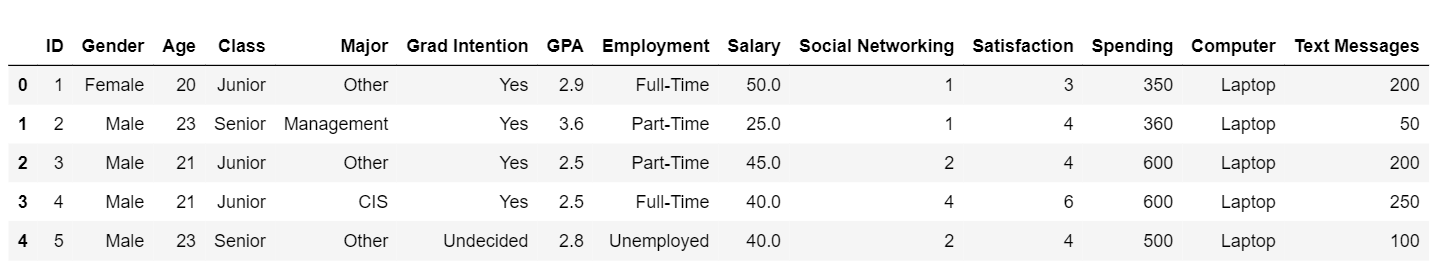
* 1. **If the null hypothesis is rejected in either (2) or in (3), find out which class means are significantly different. Interpret the result. (Non-Graded)**
  2. **Perform a two-way ANOVA based on Salary with respect to both Education and Occupation (along with their interaction Education\*Occupation). State the null and alternative hypotheses and state your results. How will you interpret this result? \_\_\_\_ 18**
  3. **Explain the business implications of performing ANOVA for this particular case study**. \_18

LIST OF FIGURES

1. Figure 1: Survey Dataset \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 05
2. Figure 2: Info of survey dataset \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 05
3. Figure 3: Description of Survey Dataset \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 06
4. Figure 4: Contingency Table of Gender & Major \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 06
5. Figure 5: Contingency Table of Gender & Grad Intention\_\_\_\_\_\_\_\_\_ 07
6. Figure 6: Contingency Table of Gender & Employment\_\_\_\_\_\_\_\_\_\_\_ 07
7. Figure 7: Contingency table of Gender & Computer\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 07
8. Figure 8: Table Gender and Major \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08
9. Figure 9: Table Gender and Major \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 08
10. Figure 10: Gender & Grad Intention\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09
11. Figure 11: Gender & Grad Intention\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09
12. Figure 12: Gender & Computer\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09
13. Figure 13: Gender & Computer\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 09
14. Figure 14: Gender & Major\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 10
15. Figure 15: Gender & Grad Intention\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 10
16. Figure 16: Gender & Salary\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 11
17. Figure 17: Gender & Salary\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 11
18. Figure 18: Histogram Plot of the Variables\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 12
19. Figure 19: Probability Plot of the Variables\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 12
20. Figure 20: First few rows of A & B Shingles Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 13
21. Figure 21: Info of A & B Shingles Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 13
22. Figure 22: Description of A & B Shingles Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 14
23. Figure 23: First few rows of Salary Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 16
24. Figure 24: Info of Salary Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 16
25. Figure 25: Description of Salary Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 17
26. Figure 26: ANOVA Table of Salary Dataset\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 18

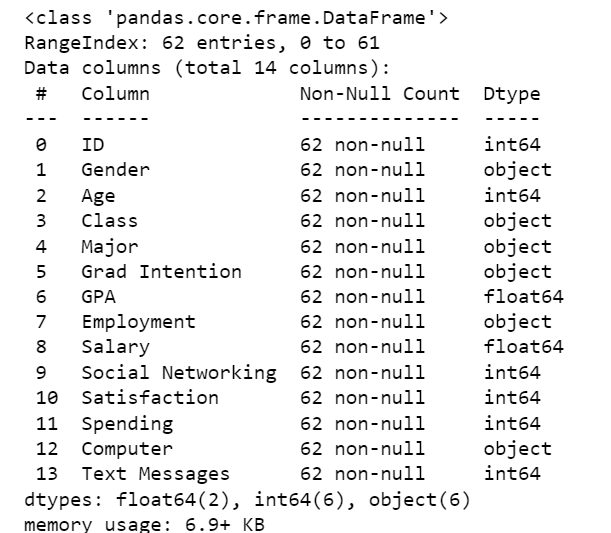
**Problem 1:  
The Student News Service at Clear Mountain State University (CMSU) has decided to gather data about the undergraduate students that attend CMSU. CMSU creates and distributes a survey of 14 questions and receives responses from 62 undergraduates (stored in the Survey data set).**

Answer:

After importing the necessary libraries, we will import the dataset and perform the basic EDA. The first few rows of the dataset are as follows: 

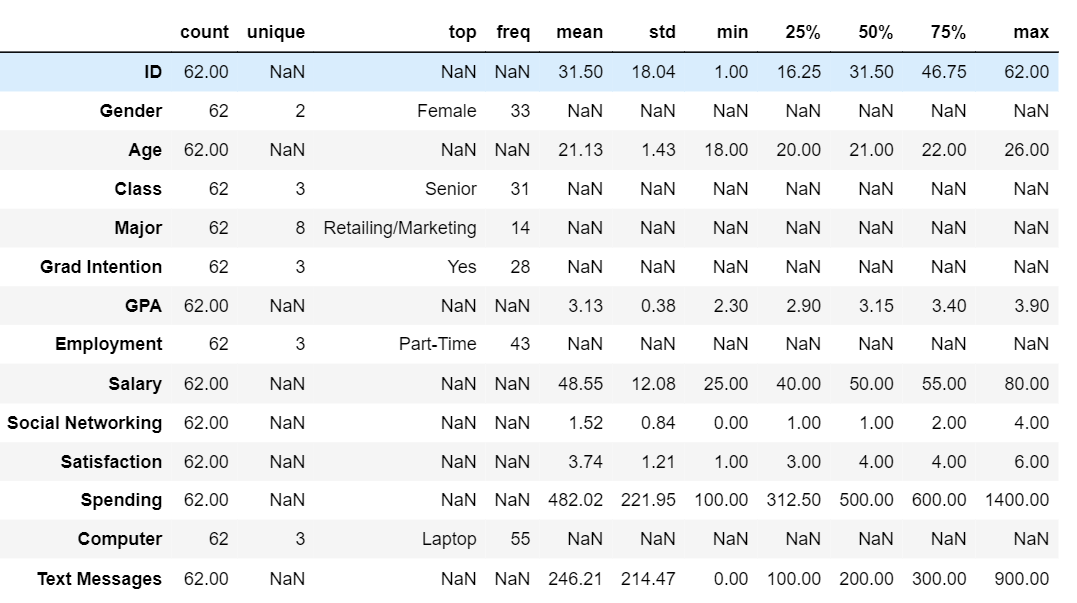
*Figure 1: Survey Dataset*

The dataset contains 62 rows and 14 columns. The info of the data is as follows :



*Figure 2: Info of survey dataset*

There are no null values in the dataset. The description of the dataset is as follows:



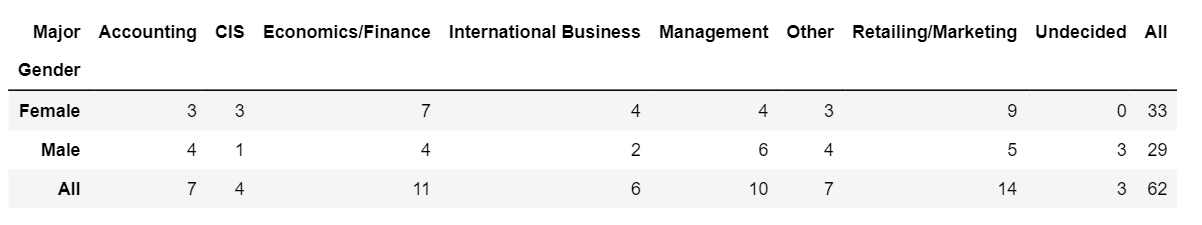
*Figure 3: Description of Survey Dataset*

Basic Inferences:

1. There are 62 rows and 14 columns in the dataset.
2. The dataset consists of 33 female students and 29 male students.
3. The age ranges from 18 years to 26 years with an mean age of 21.13 years.
4. There are 8 major subjects with Retailing/Marketing being the most popular subject.
5. More than 50% of students have Part-Time jobs.
6. Almost 89% of students have Laptop.

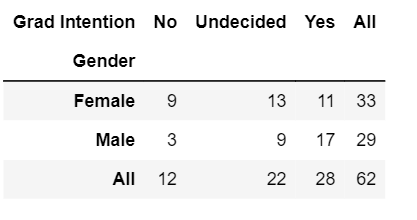
**1.1. For this data, construct the following contingency tables (Keep Gender as row variable)**

**1.1.1. Gender and Major**



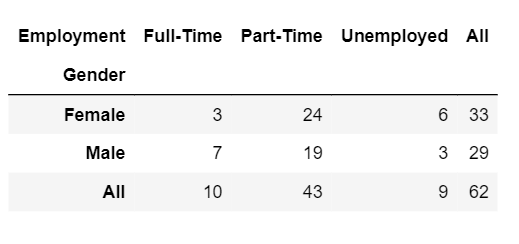
*Figure 4: Contingency Table of Gender & Major*

**1.1.2. Gender and Grad Intention**



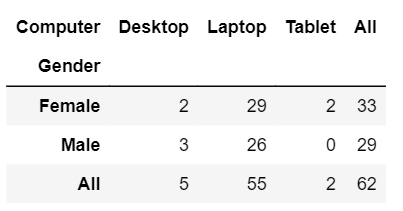
*Figure 5: Contingency Table of Gender & Grad Intention*

**1.1.3. Gender and Employment**



*Figure 6: Contingency Table of Gender & Employement*

**1.1.4. Gender and Computer**



*Figure 7: Contingency table of Gender & Computer*

**1.2. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question:**

**1.2.1. What is the probability that a randomly selected CMSU student will be male?**

Answer:

Number of male (A) = 29

Total Number of students (B) = 62

P (A/B) =29/62

The probability that a randomly selected CMSU student will be male is 46.774%

**1.2.2. What is the probability that a randomly selected CMSU student will be female?**

Answer:

Number of female (A) = 33

Total Number of students (B) = 62

P (A/B) =33/62

The probability that a randomly selected CMSU student will be female is 53.226%

**1.3. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question:**

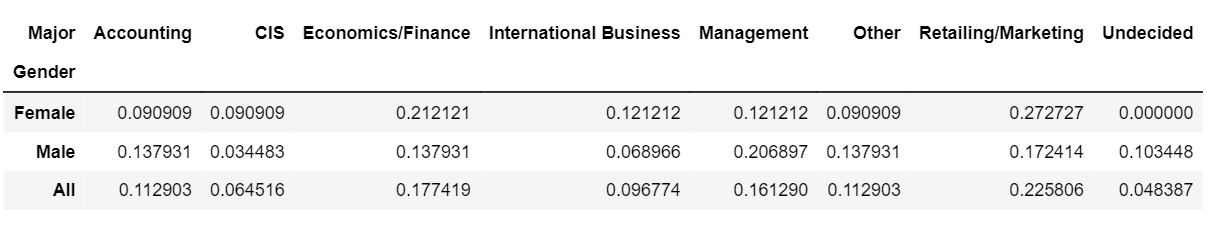
**1.3.1. Find the conditional probability of different majors among the male students in CMSU.**

Answer:

Conditional probability of different Majors

P (Different Majors/ Male).

The following figure shows the probability of male choosing different majors



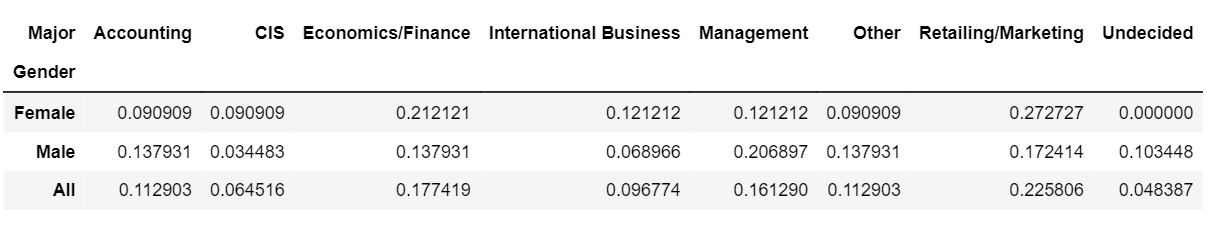
*Figure 8: Table Gender and Major*

**1.3.2 Find the conditional probability of different majors among the female students of CMSU.**

Answer:

P (Conditional Majors/ Female).

The figure shows the probability of female choosing different majors



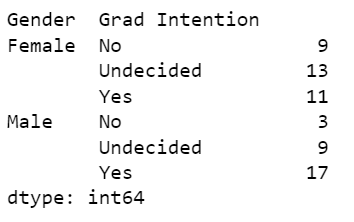
*Figure 9: Table Gender and Major*

**1.4. Assume that the sample is a representative of the population of CMSU. Based on the data, answer the following question:**

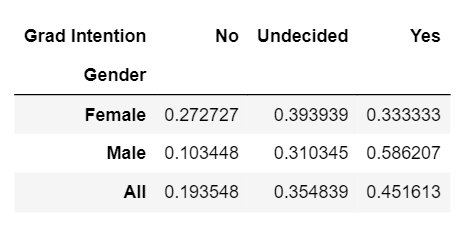
**1.4.1. Find the probability That a randomly chosen student is a male and intends to graduate.**

Answer:

P (Graduation Intention Yes/ Male) = 17/29



*Figure 10: Gender & Grad Intention*



*Figure 11: Gender & Grad Intention*

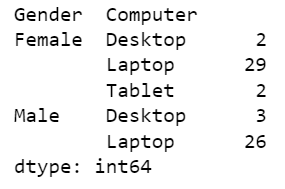
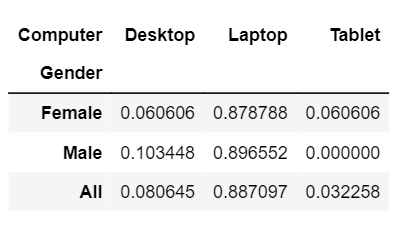
The probability that a randomly chosen student is a male and intends to graduate is 58.62%

**1.4.2 Find the probability that a randomly selected student is a female and does NOT have a laptop.**

**Answer:**

P (Have a laptop/ female) = 29/33

P (does not have a laptop/ female) = 1- P (Have a laptop/ female) = 1-0.88=12%

*Figure 12: Gender & Computer Figure 13: Gender & Computer*

The probability that a randomly selected student is a female and does NOT have a laptop is 12%

**1.5. Assume that the sample is representative of the population of CMSU. Based on the data, answer the following question:**

**1.5.1. Find the probability that a randomly chosen student is a male or has full-time employment?**

Answer:

Probability of randomly selected student is male P (A) = 46.77%

Probability of randomly selected student has a fulltime job P (B) = 16.13%

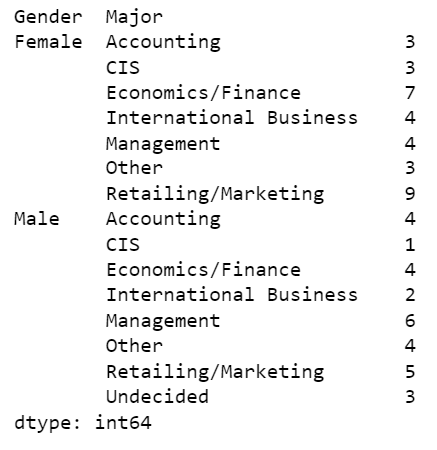
Probability of male having a fulltime job P (A and B) = 11.29%

p\_male\_or\_fulltime= p\_male\_student+p\_fulltime-p\_male\_fulltime= 51.61%

The probability that a randomly chosen student is a male or has full-time employment is 51.61 %

**1.5.2. Find the conditional probability that given a female student is randomly chosen, she is majoring in international business or management.**

**Answer:**

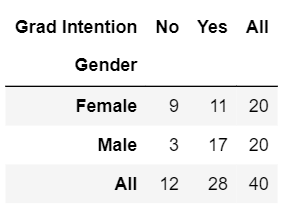


*Figure 14: Gender & Major*

Probability that given a female student is randomly chosen, she is majoring in international business or management is 24.24 %

**1.6.  Construct a contingency table of Gender and Intent to Graduate at 2 levels (Yes/No). The Undecided students are not considered now and the table is a 2x2 table. Do you think the graduate intention and being female are independent events?**

**Answer:**



*Figure 15: Gender & Grad Intention*

The probability that a randomly selected student is female and intends to graduate 55.0 %

They are not independent events

**1.7. Note that there are four numerical (continuous) variables in the data set, GPA, Salary, Spending, and Text Messages.**

**Answer the following questions based on the data**

**1.7.1. If a student is chosen randomly, what is the probability that his/her GPA is less than 3?**

**Answer:**

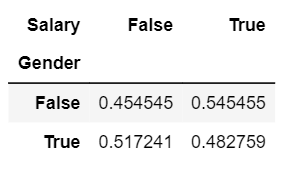
No of students less than 3GPA is 17

Total number of students is 62

The probability that a student's GPA is less than 3 is 27.419 %

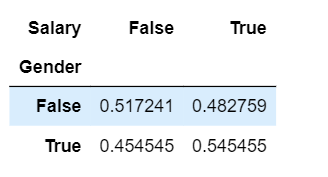
**1.7.2. Find the conditional probability that a randomly selected male earns 50 or more. Find the conditional probability that a randomly selected female earns 50 or more.**

**Answer:**



*Figure 16: Gender & Salary*

Probability that a randomly selected male earns 50 or more is 48%



*Figure 17: Gender & Salary*

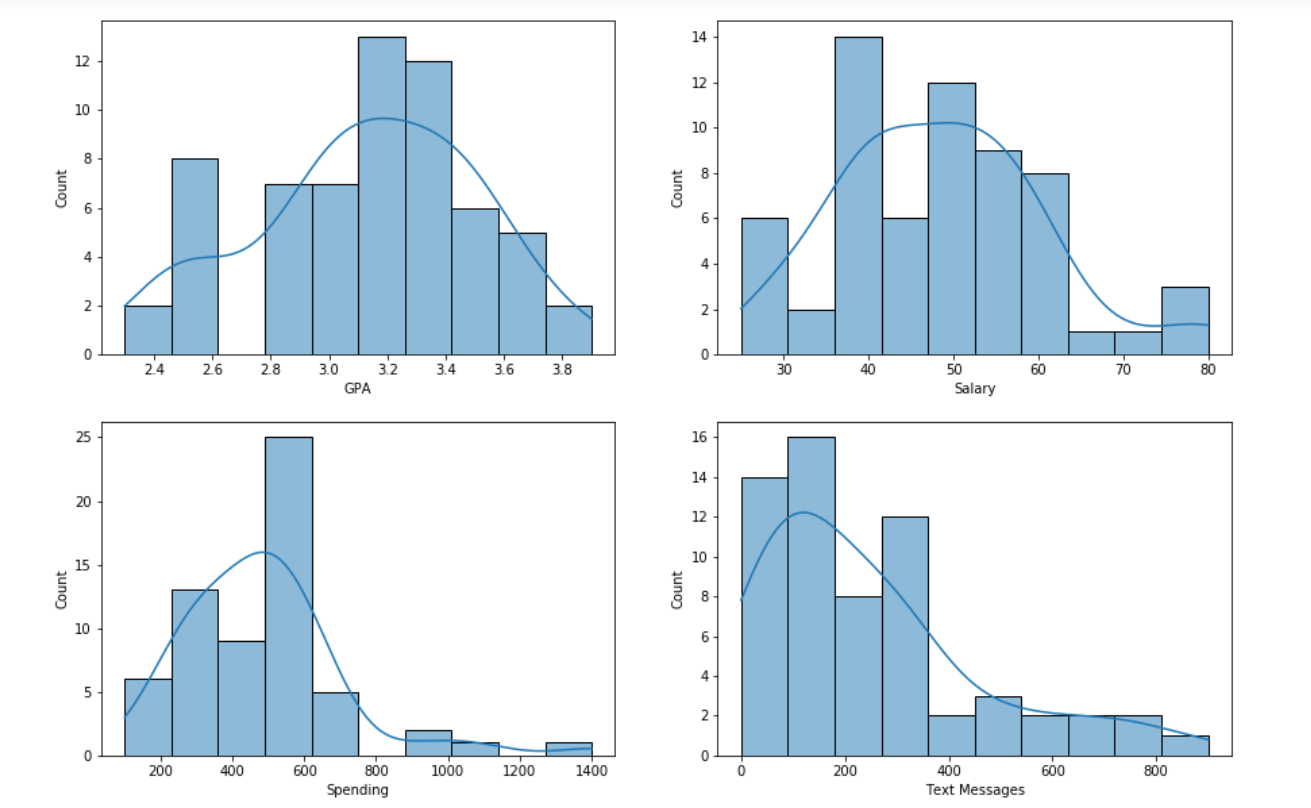
Probability that a randomly selected female earns 50 or more is 54%.

**1.8. Note that there are four numerical (continuous) variables in the data set, GPA, Salary, Spending, and Text Messages. For each of them, comment whether they follow a normal distribution. Write a note summarizing your conclusions.**

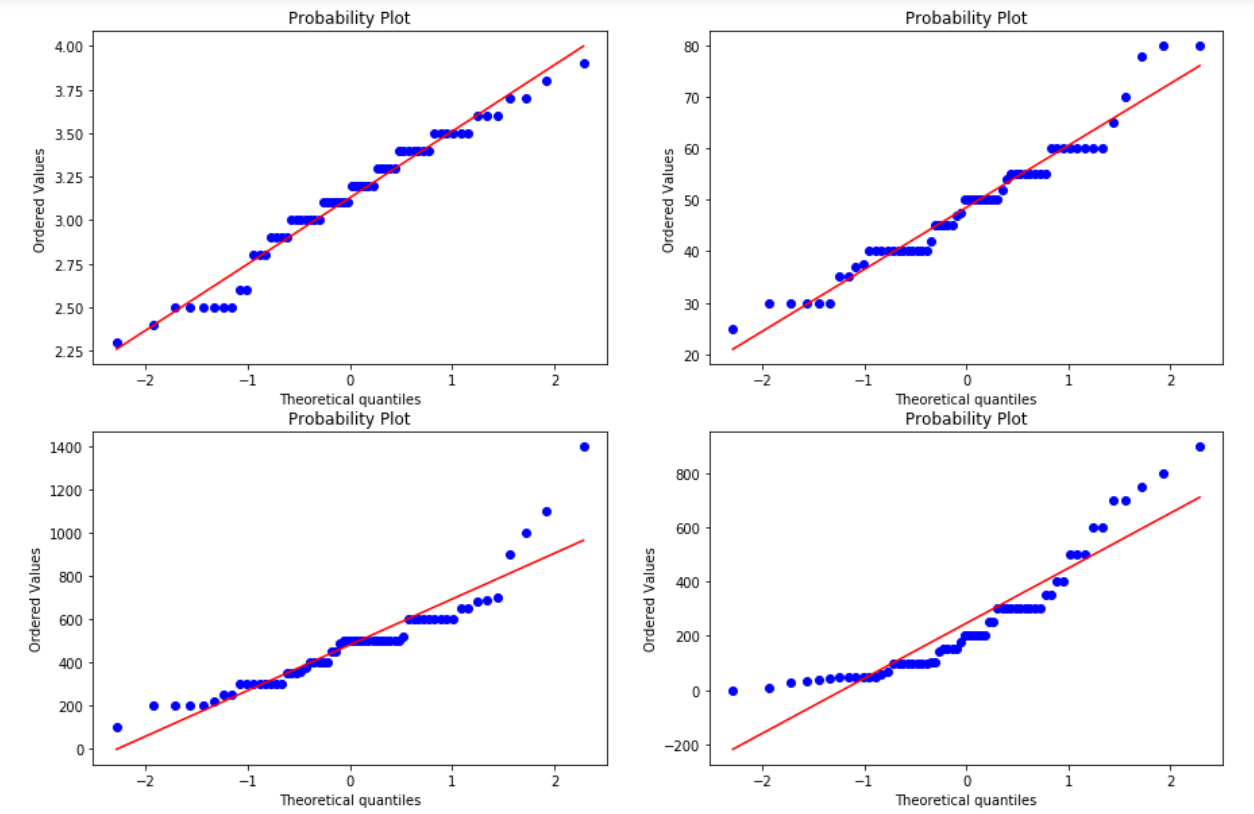
**Answer:**

The probability plot can be used to find the dataset follows a normal distribution or not, in our dataset we can find the points follow a straight line and we can say that all the GPA, salary, spending and text messages follow a normal distribution. Looking at the skew value if the value is zero it is symmetric data, if we have a negative value for the skew that indicates that the data are skewed left and positive value of skew indicates the data are skewed towards right.

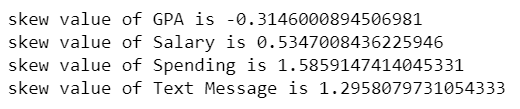
The Histplot, probability plot and skewness of the variables is as follows.



*Figure 18: Histogram Plot of the Variables*



*Figure 19: Probability Plot of the Variables*

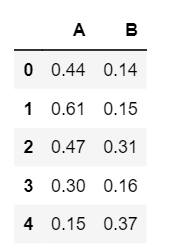


**Problem 2**

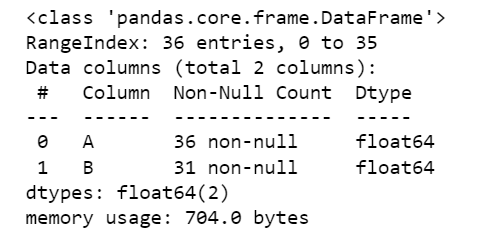
**An important quality characteristic used by the manufacturers of ABC asphalt shingles is the amount of moisture the shingles contain when they are packaged. Customers may feel that they have purchased a product lacking in quality if they find moisture and wet shingles inside the packaging.   In some cases, excessive moisture can cause the granules attached to the shingles for texture and coloring purposes to fall off the shingles resulting in appearance problems. To monitor the amount of moisture present, the company conducts moisture tests. A shingle is weighed and then dried. The shingle is then reweighed, and based on the amount of moisture taken out of the product, the pounds of moisture per 100 square feet is calculated. The company would like to show that the mean moisture content is less than 0.35 pounds per 100 square feet.**

**The file includes 36 measurements (in pounds per 100 square feet) for A shingles and 31 for B shingles.**

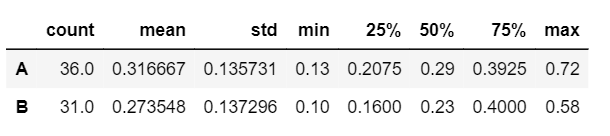
After Importing the necessary libraries and the dataset, the following are the first few rows of the dataset:



*Figure 20: First few rows of A & B Shingles Dataset*

The info of the data is as follows:  


*Figure 21: Info of A & B Shingles Dataset*

The description of the data is as follows:  


*Figure 22: Description of A & B Shingles Dataset*

**2.1 Do you think there is evidence that means moisture contents in both types of shingles are within the permissible limits? State your conclusions clearly showing all steps.**

Answer:

Here we are provided with two independent samples of shingles A and B population standard deviation is unknown and hence we can’t perform z test. So, we have to go with t-test. Since we have to find the mean moisture level is less than the permissible limit for the both samples, we have performed one sample t-test for sample A and sample B.

Sample A:

Step 1: Defining Null and Alternate Hypothesis for Sample A and Sample B

The null hypothesis states that the moisture content of Sample A and Sample B is greater or than equal to the permissible limit, 𝜇 ≥ 0.35

The alternative hypothesis states that the moisture content of sample A and Aample B is less than permissible limit, 𝜇 < 0.35

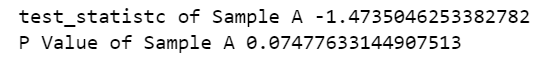
𝐻0 : 𝜇 ≥ 0.35

𝐻𝐴 : 𝜇 < 0.35

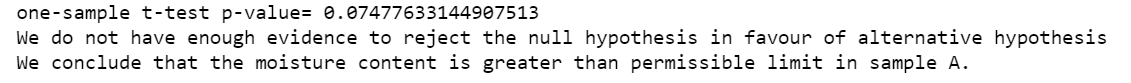
Step 2: Significance Level. We will take the significance level 𝛼 = 0.05 as the alpha value is not given in the question.

Step 3: Identify the test statistic We have two samples (A and B) and we do not know the population standard deviation. Sample sizes for both samples are not the same. The sample size is , n > 30. So we use the t distribution and the 𝑡𝑆𝑇𝐴𝑇 test statistic for one sample test for Sample A and Sample B. One tail test for Sample A and Sample B.

Step 4: Calculate the p - value and test statistic



Step 5: Decide to reject or accept null hypothesis



Sample B:

Step 1: Defining Null and Alternate Hypothesis for Sample A and Sample B

The null hypothesis states that the moisture content of Sample A and Sample B is greater or than equal to the permissible limit, 𝜇 ≥ 0.35

The alternative hypothesis states that the moisture content of sample A and Aample B is less than permissible limit, 𝜇 < 0.35

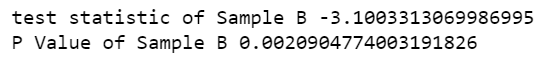
𝐻0 : 𝜇 ≥ 0.35

𝐻𝐴 : 𝜇 < 0.35

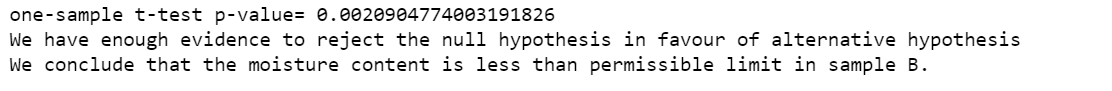
Step 2: Significance Level. We will take the significance level 𝛼 = 0.05 as the alpha value is not given in the question.

Step 3: Identify the test statistic We have two samples (A and B) and we do not know the population standard deviation. Sample sizes for both samples are not the same. The sample size is , n > 30. So we use the t distribution and the 𝑡𝑆𝑇𝐴𝑇 test statistic for one sample test for Sample A and Sample B. One tail test for Sample A and Sample B.

Step 4: Calculate the p - value and test statistic



Step 5: Decide to reject or accept null hypothesis



**2.2 Do you think that the population mean for shingles A and B are equal? Form the hypothesis and conduct the test of the hypothesis. What assumption do you need to check before the test for equality of means is performed?**

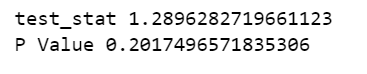
Answer: The following steps are followed:

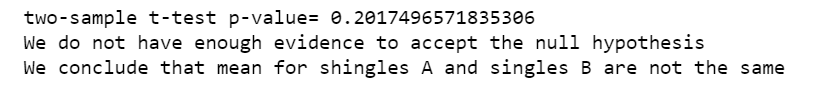
Step 1: Define Null and Alternate Hypothesis. The Null hypothesis states that the mean of shingle A to mean of shingle B are the same. The Alternate hypothesis states that the mean are different. Mean of Sample A in not equal to mean od sample B

Step 2:Decide the significance level. Here we select alpha= 0.05 as the population standard deviation is not known.

Step 3: Identify the test statistic We have two samples and we do not know the population standard deviation. Sample sizes for both samples are not the same. The sample size is , n > 30. So we use the t distribution and the 𝑡𝑆𝑇𝐴𝑇 test statistic for two sample test.

Step 4: Calculate the p - value and test statistic



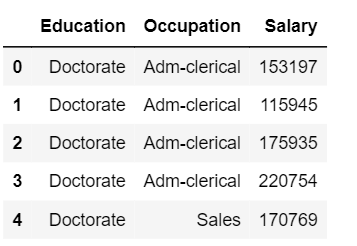
Step 5: Decide to reject or accept null hypothesis

**Problem 3:**

**Salary is hypothesized to depend on educational qualification and occupation. To understand the dependency, the salaries of 40 individuals are collected and each person’s educational qualification and occupation are noted. Educational qualification is at three levels, High school graduate, Bachelor's, and Doctorate. Occupation is at four levels, Administrative and clerical, Sales, Professional or specialty, and Executive or managerial. A different number of observations are in each level of education – occupation combination.**

Answer:

After importing the necessary libraries and the dataset, the first few rows of the dataset are as follows:

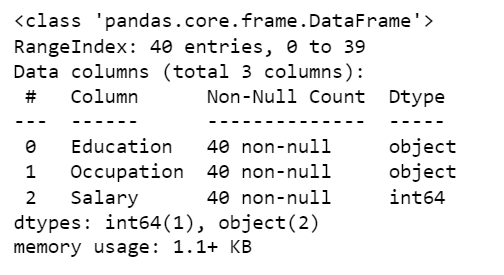


*Figure 23: First few rows of Salary Dataset*

The shape of the data is as follows:

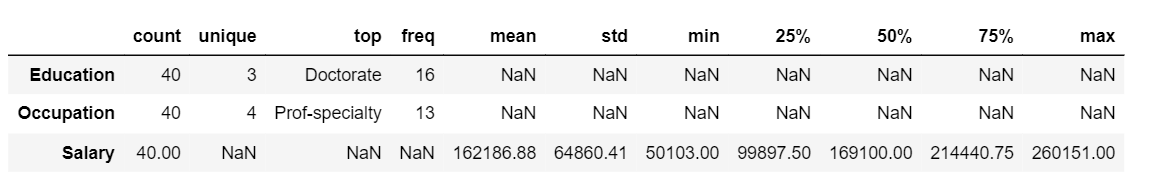


The info of the data is as follows:



*Figure 24: Info of Salary Dataset*

The description of the data is as follows:



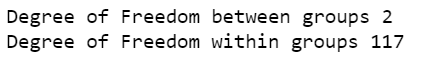
*Figure 25: Description of Salary Dataset*

**3.1 State the null and the alternate hypothesis for conducting one-way ANOVA for both Education and Occupation individually.**

Answer:

Null hypothesis: The salary is dependent on Education Qualification and Occupation.

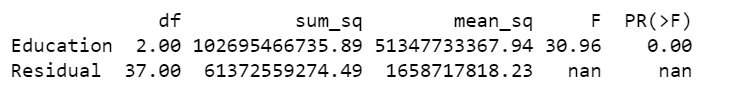
Alternate hypothesis: The salary is not dependent on Education Qualification and Occupation



**3.2 Perform a one-way ANOVA on Salary with respect to Education. State whether the null hypothesis is accepted or rejected based on the ANOVA results.**

Answer:

The output of the ANOVA test is as follows:

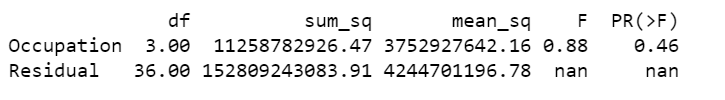


Inference: Since the value of p is less than the significance level, we can reject the null hypothesis and state thet salary does not depend on the educational qualification of the person.

**3.3 Perform a one-way ANOVA on Salary with respect to Occupation. State whether the null hypothesis is accepted or rejected based on the ANOVA results.**

Answer:

The output of the ANOVA test is as follows:



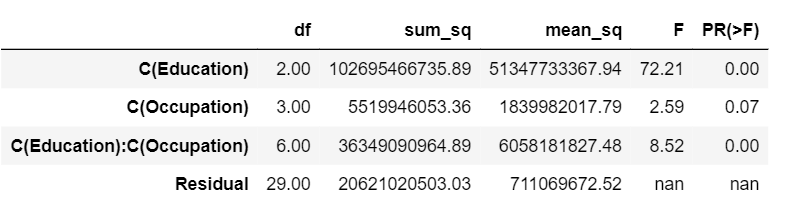
Inference: Since the value of p is greater than the significance level, therefore, we can accept the null hypothesis and state that salary is dependet on the occupation of the person.

**3.5 Perform a two-way ANOVA based on Salary with respect to both Education and Occupation (along with their interaction Education\*Occupation). State the null and alternative hypotheses and state your results. How will you interpret this result?**

Answer:

Null hypothesis: The salary is dependent on Education Qualification and Occupation.

Alternate hypothesis: The salary is not dependent on Education Qualification and Occupation



*Figure 26: ANOVA Table of Salary Dataset*

Inference: Since the value of p is less than the significance level, we can reject the null hypothesis and state that salary does not depend on both educational qualification and the occupation together.

**3.6 Explain the business implications of performing ANOVA for this particular case study.**

Answer:

By performing the ANOVA test for Salary case study we can say that Salary is not depended on educational qualification but in fact it is little dependent on Occupation.

But when considered both the class it says that Salary is moderately depended both with respect to Education and Occupation