San Francisco Bay University

MATH208 - Probability and Statistics
2023 Fall Homework #4

Due day:
12/7/2023

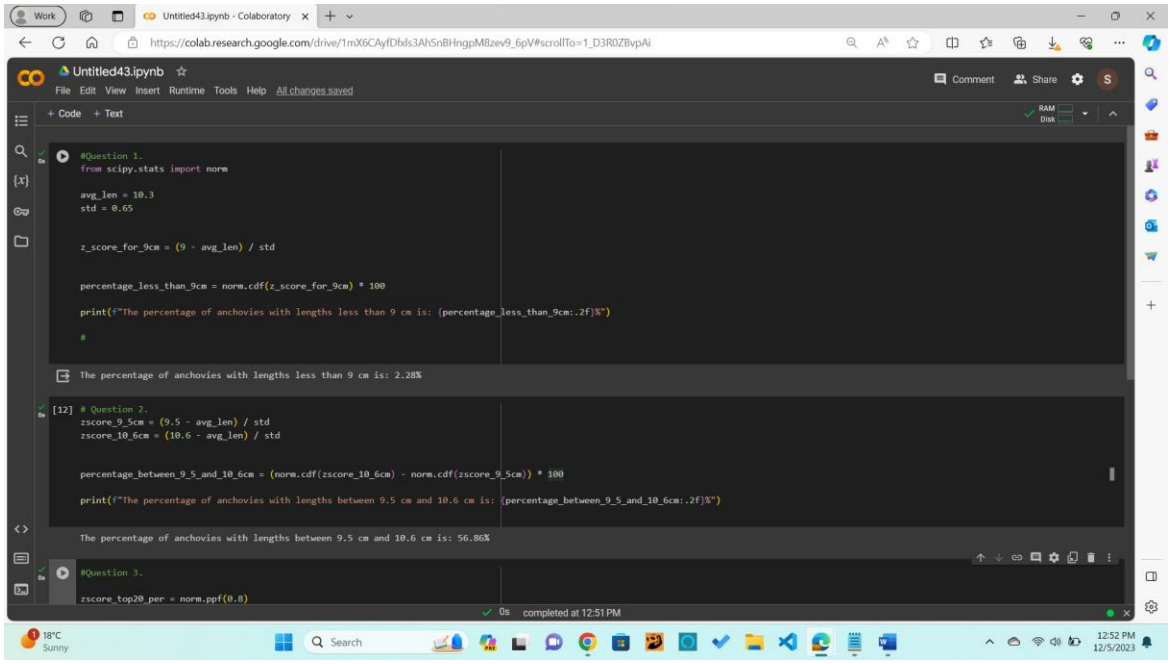
Instruction:

1. **Homework answer sheet should contain the original questions and corresponding answers.**
2. **Answer sheet must be in PDF file format with Github links for the programming questions, but MS Word file can't be accepted. As follows is the answer sheet name format.** ***<course_id>_week<week_number>_StudentID_FirstName_LastName.pdf***
3. **The program name in Github must follow the format like** ***<course_id>_week<week_number>_q<question_number>_StudentID_FirstName_LastName***
4. **If the calculation in Excel is needed, the original file must be provided.**
5. **Show screenshot of all running results, including the system date/time.**
6. **The calculation process must be printed if needed, handwriting can't be accepted.**
7. **Only accept homework submission uploaded via Canvas.**
8. **Overdue homework submission can't be accepted.**
3. **Takes academic honesty and integrity seriously (Zero Tolerance of Cheating & Plagiarism)**

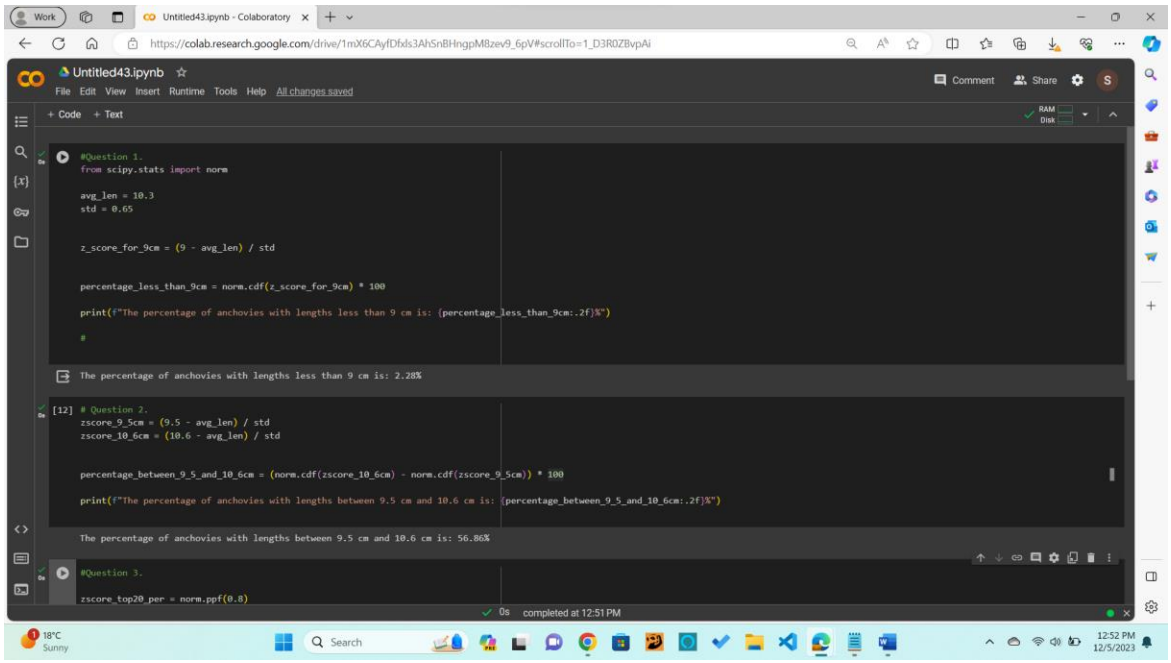
1. Assuming that the lengths of American anchovies appease the normal distribution with the mean $\mu = 10.3cm$ and standard deviation $\sigma = 0.65cm$, please find the percentages of the lengths in the population of American anchovies.

a. Less than *9cm*.

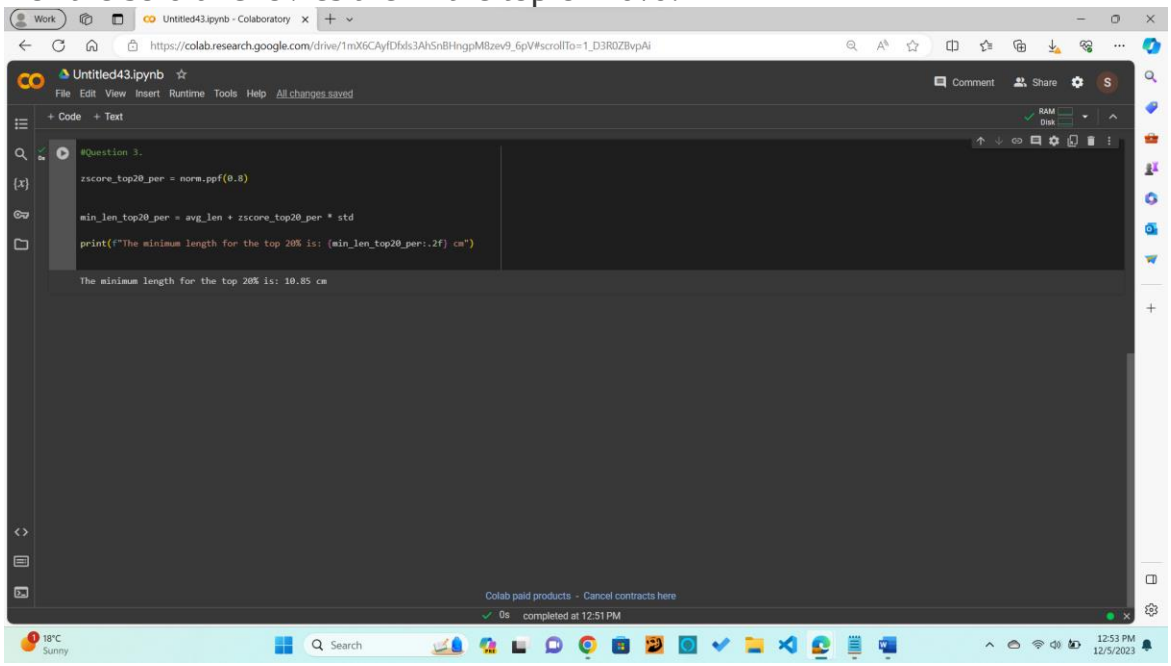
→



b. Between *9.5cm* and *10.6cm*.



c. What is the minimum length if a restaurant claimed that the lengths of the sold anchovies are in the top of 20%?



Hint: Python functions or functions in Excel for the calculation of each question is preferred.

2. If the random variables X and Y are normal distributions with $\mu = 10$ & $\sigma = 3$ and $\mu = 15$ & $\sigma = 8$, namely, $X \sim N(10, 3)$ and $Y \sim N(15, 8)$, and they are independent, what is the probability distribution and statistical parameters of :

→

(1) $X + Y$:

$\mu x + \mu y = 10 + 15$

Mean = 25

$\sigma^2_{x+y} = 3^2 + 8^2$

$= 9 + 64$

Variance = 73

The probability distribution is: $N(25, \sqrt{73})$

(2) $X - Y$:

$\mu x - \mu y = 10 - 15$

Mean = -5

$$\sigma^2_{x-y} = 3^2 + 8^2$$
$$= 9 + 64$$
Variance = 73

The probability distribution is : $\mathrm{N}\left(-5, \sqrt{73}\right)$

(3) $3X$:

$$\mu 3x = 3.\mu x$$
Mean = 3 . 10 = 30

$$\sigma^2_{3x} = (3.\sigma x)^2$$
$$= (3.3)^2$$
Variance = 81

The probability distribution is : $\mathrm{N}(30,9)$

(4) $4X + 5Y$:

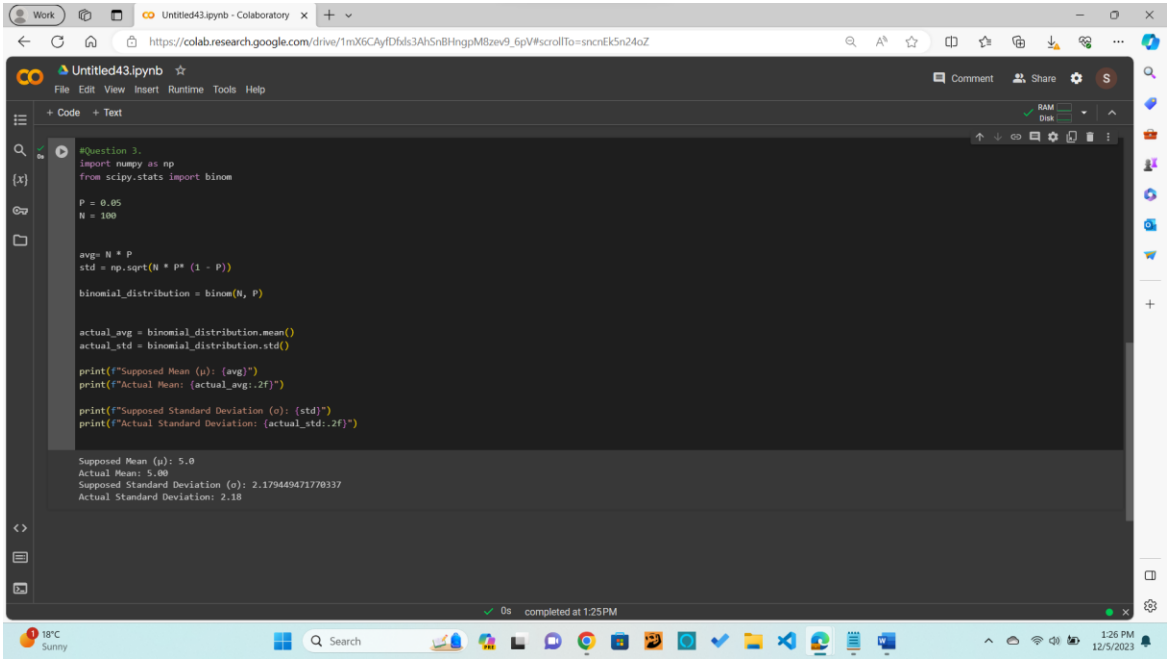
$$\mu 4x + 5y = 4.\mu x + 5.\mu y$$
$$= 4.10+5.15$$
$$= 40 + 75$$
Mean = 115

$$\sigma^2_{4x+5y} = (4.\sigma x)^2 + (5.\sigma y)^2$$
$$= (4.3)^2 + (5.8)^2$$
$$= 144 + 200$$
$$= 344$$

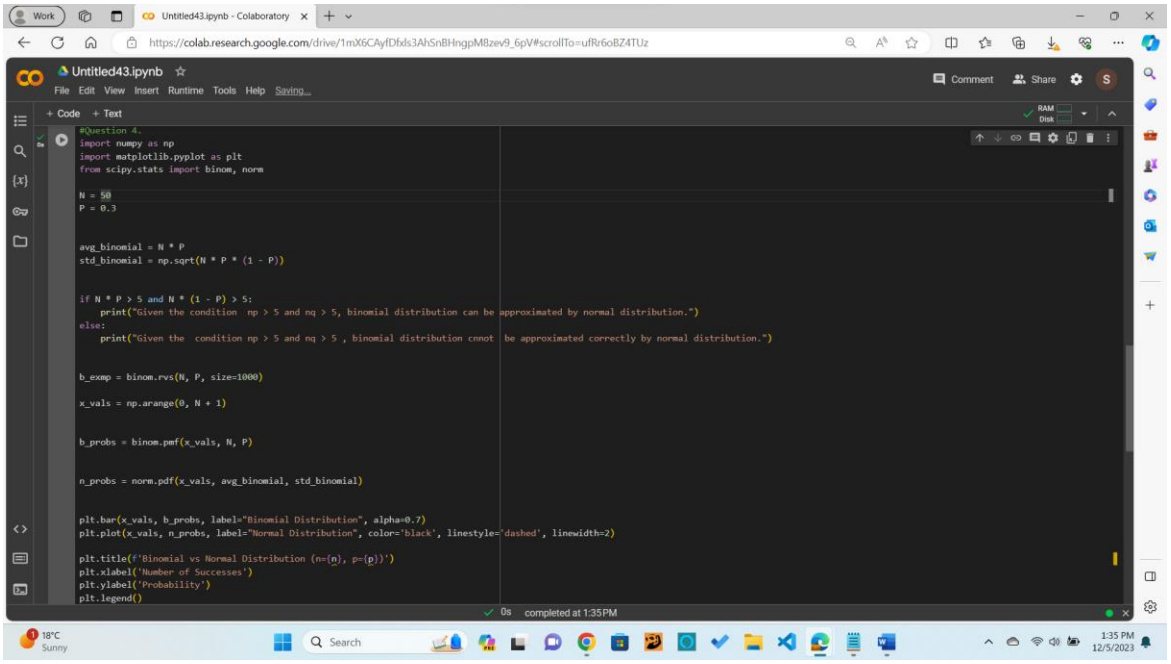
Variance = 73

The probability distribution is: $\mathrm{N}\left(115, \sqrt{344}\right)$

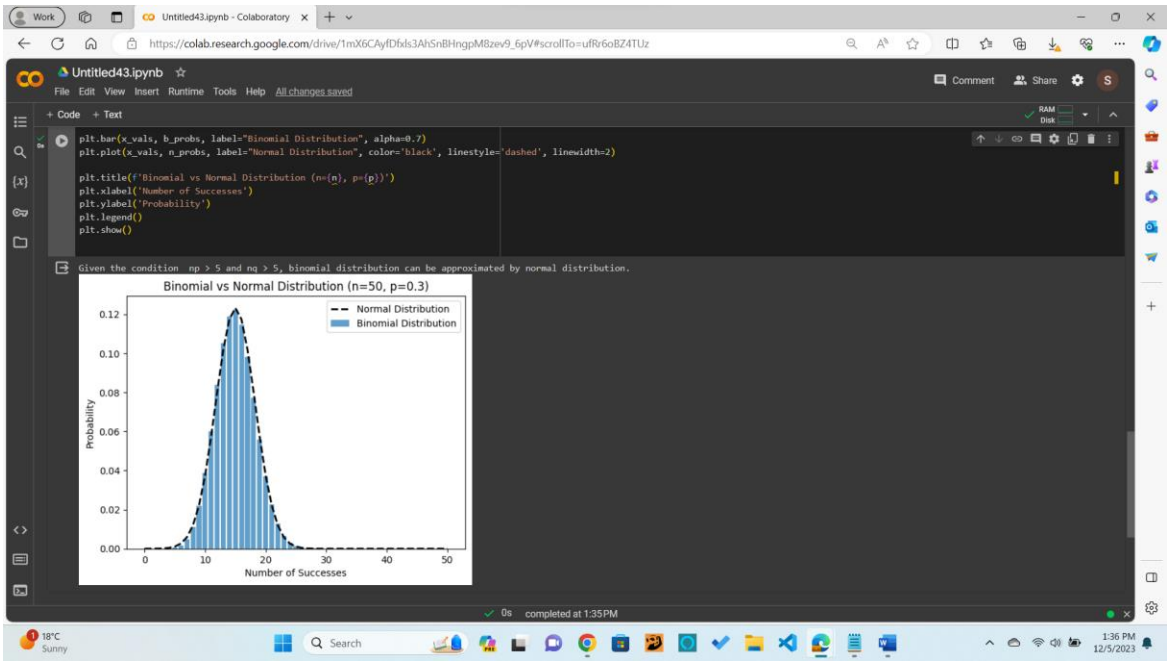
3. For the students in Engineering School, please write Python program to verify the mean $\mu = np$ and standard deviation $\sigma = \sqrt{npq}$ for $p=0.05$ and selecting any n greater than 50 in binomial distribution.
For Business school students, verify in Excel.
→



4. In general, if $np > 5$ and $nq > 5$ in binomial distribution, binomial probabilities can be approximated using the normal distribution. Please select any big enough *n* and *p*'s values to verify in Python program or Excel and plot the histogram.



```
#Question 4.
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import binom, norm

N = 50
P = 0.3


avg_binomial = N * P
std_binomial = np.sqrt(N * P * (1 - P))


if N * P > 5 and N * (1 - P) > 5:
    print("Given the condition  np > 5 and nq > 5, binomial distribution can be approximated by normal distribution.")
else:
    print("Given the  condition np > 5 and nq > 5 , binomial distribution cnnot  be approximated correctly by normal distribution.")


b_exmp = binom.rvs(N, P, size=1000)

x_vals = np.arange(0, N + 1)


b_probs = binom.pmf(x_vals, N, P)


n_probs = norm.pdf(x_vals, avg_binomial, std_binomial)


plt.bar(x_vals, b_probs, label="Binomial Distribution", alpha=0.7)
plt.plot(x_vals, n_probs, label="Normal Distribution", color='black', linestyle='dashed', linewidth=2)

plt.title(f'Binomial vs Normal Distribution (n={n}, p={p})')
plt.xlabel('Number of Successes')
plt.ylabel('Probability')
plt.legend()
```



```
plt.bar(x_vals, b_probs, label="Binomial Distribution", alpha=0.7)
plt.plot(x_vals, n_probs, label="Normal Distribution", color='black', linestyle='dashed', linewidth=2)

plt.title(f'Binomial vs Normal Distribution (n={n}, p={p})')
plt.xlabel('Number of Successes')
plt.ylabel('Probability')
plt.legend()
plt.show()
```

```
Given the condition  np > 5 and nq > 5, binomial distribution can be approximated by normal distribution.
```

5. In coin tossing experiments, please find the probability of the exact 6 heads from 12 tossing by ONLY using the normal distribution method.

→

$$\mu = np$$
$$\sigma = \sqrt{npq}$$

Using normal distribution to find the probability of exactly 6 heads, we get:

P(X =k) = $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

Substituting the values , we get:

$$\mu = 12 * 0.5$$
$$= 6$$
$$\sigma = \sqrt{12 \times 0.5 \times 0 \cdot 5}$$
$$= \sqrt{3}$$

P(X= 6) = $\frac{1}{\sqrt{3}\sqrt{2\pi}} \exp\left(-\frac{(6-6)^2}{2\times3}\right)$

P(X=6) $\frac{1}{\sqrt{6\pi}}$

Calculating probability , we get:

P(X=6) ~ $\frac{1}{\sqrt{6\pi}}$ ~ 0.1508

The probability of getting exactly 6 heads in 12 coins toss is 0.1508

6. Given that the defective rate of a product of the batteries in a manufacturing company is 6%, 150 batteries are randomly selected from the population. Please find the probability of 12 or more defective ones in them by ONLY using the normal distribution method.

→

p = 0.06

n = 150

$$\mu = NP$$
$$= 150 * 0.06$$
$$Mean = 9$$

Calculating q , we get:

q = 1- p

= 1- 0.06

q= 0.94

Calculating std, we get:

$\sigma = \sqrt{npq}$

$= \sqrt{150 \times 0.06 \times 0 \cdot 94}$

$= \sqrt{8.46} \sim 2.91$

Calculating the z – score, we get:

$Z = \frac{X-\mu}{\sigma}$

$= \frac{12-9}{2 \cdot 91} \sim 1.03$

Using z-score to find P(X<12) using a standard normal distribution :

$P(X \geq 12) = 1 - P(X < 12)$

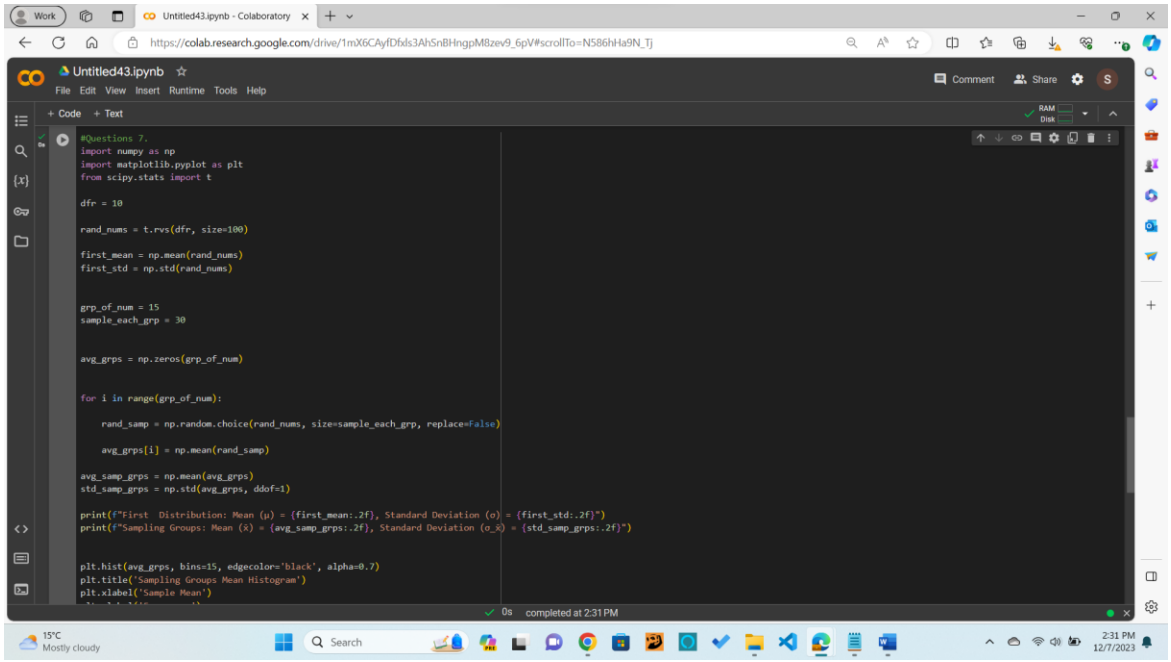
$= 1 - 0.8485$

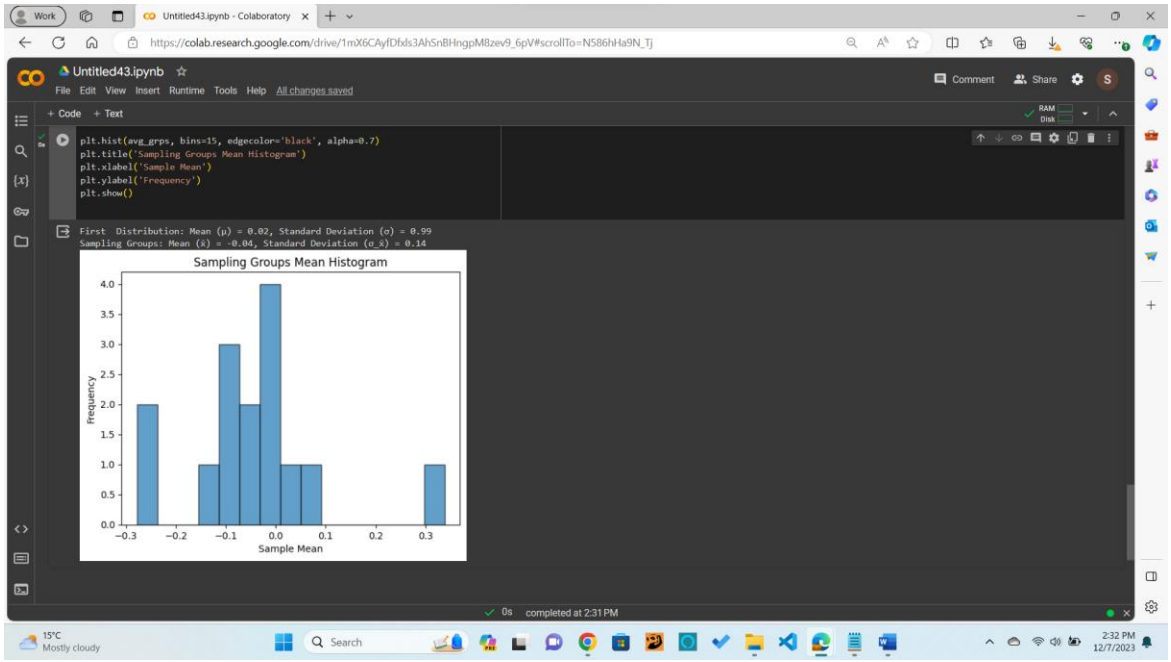
$P(X \geq 12) \sim 0.1515$

The probability of having 12 or more defective batteries out of 150 is 0.1515 or 15.15%.

7. For the students in Engineering School, please write a Python program by calling functions in the following link to create *100* random numbers in T distribution with *df =10* (degree of freedom) and calculate the mean μ and standard deviation σ. After that, the *30* samples will be randomly selected from these random numbers in each sampling group. A total of 15 sampling groups should be created. Based on Central Limit Theorem (CLT), the mean value $\bar{\bar{x}}$ in total 15 sampling group is roughly the mean μ of *100* random numbers and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Please verify it and plot the histogram, which should be normal distribution. For Business school students, complete the above process in Excel.

➔





Python function link:
https://www.statology.org/t-distribution-python/

Excel function link:
https://support.microsoft.com/en-us/office/t-dist-function-4329459f-ae91-48c2-bba8-1ead1c6c21b2