1. A random variable is a function that assigns a numerical value to each outcome of a random experiment .It translates outcomes from a sample space into numbers.
   There are two main types: discrete-countable values and continuous -any value in an interval.Random variables allow us to calculate probabilities, expectations, and other statistical measures.
2. There are two main types of random variables:

   Discrete random variable – Takes countable values (e.g., 0, 1, 2, 3).

   Continuous random variable – Takes any value within a given interval (e.g., real numbers between 0 and 1).

3. **Discrete distribution** deals with random variables that take **countable values** (like 0, 1, 2, 3).
It is described by a **probability mass function (PMF)**, and probabilities are found by summing values.
Each specific value can have a positive probability.

Continuous distribution deals with random variables that take **uncountable values** (any value in an interval).
It is described by a **probability density function (PDF)**, and probabilities are found by integration.
The probability of any single exact value is zero.

4. The **binomial distribution** is a discrete probability distribution that models the number of successes in a fixed number of independent trials.
Each trial has only two possible outcomes (success or failure) and the same probability of success.
It is commonly used to calculate the probability of getting a certain number of successes in experiments like coin tosses.
It is defined by two parameters: n(number of trials) and p(probability of success).

5. The **standard normal distribution** is a continuous probability distribution with mean 0 and standard deviation 1.
It is a special case of the normal distribution and is symmetric about zero.
It is used to convert normal random variables into **z-scores** for easier probability calculations.
It is important because many real-world phenomena approximately follow it, and statistical tests rely on it.

6. The **Central Limit Theorem (CLT)** states that the sampling distribution of the sample mean approaches a normal distribution as the sample size becomes large, regardless of the population's original distribution.
This holds as long as the samples are independent and identically distributed with a finite mean and variance.
It explains why the normal distribution appears so frequently in statistics.
The CLT is critical because it allows us to make inferences about population parameters using normal-based methods, even when the population is not normally distributed.

7. Confidence intervals provide a range of values that likely contains the true population parameter (such as a mean or proportion).
They show the precision and reliability of an estimate, not just a single value.
A higher confidence level (e.g., 95%) means greater certainty but usually a wider interval.
They are significant because they help in making informed decisions and assessing statistical uncertainty.

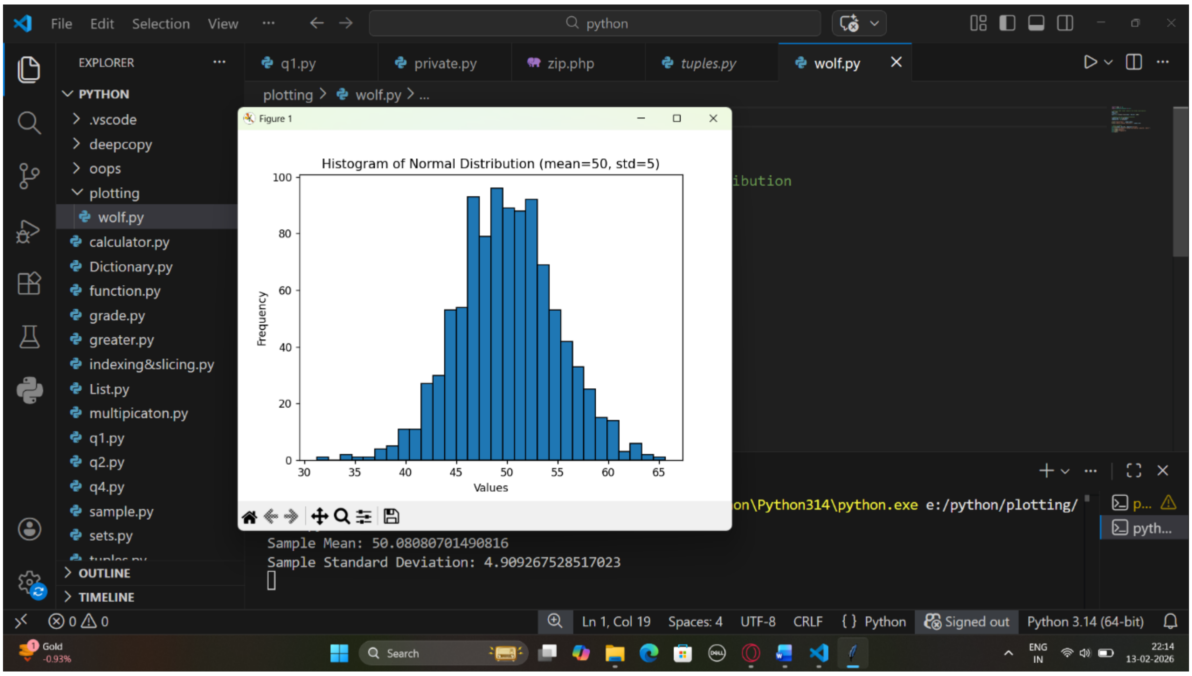
8. The **expected value** is the long-run average value of a random variable over many repeated trials.
It is calculated as the weighted average of all possible values, using their probabilities.
For a discrete distribution, $E(X) = \sum x\, P(x)$; for a continuous distribution, it is found using integration.
It represents the theoretical mean of the probability distribution.

9. import numpy as np

```
import matplotlib.pyplot as plt

# Generate 1000 random numbers from normal distribution

mean = 50

std_dev = 5

data = np.random.normal(mean, std_dev, 1000)

sample_mean = np.mean(data)

sample_std = np.std(data)

print("Sample Mean:", sample_mean)

print("Sample Standard Deviation:", sample_std)

# Draw histogram

plt.hist(data, bins=30, edgecolor='black')

plt.title("Histogram of Normal Distribution (mean=50, std=5)")
```

plt.xlabel("Values")

plt.ylabel("Frequency")



plt.show()

10. **Steps to Estimate Average Sales with 95% Confidence Interval**

1. **Calculate the sample mean**

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2. **Calculate the sample standard deviation (s)**

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

3. **Compute Standard Error (SE)**

$$SE = \frac{s}{\sqrt{n}}$$

4. **Use t-distribution for 95% CI (because n is small)**

$$CI = \bar{x} \pm t_{\alpha/2,n-1} \cdot SE$$

Where:

- $t_{\alpha/2,n-1}$ is the **critical t-value**
- $\alpha = 0.05$

daily_sales = [220, 245, 210, 265, 230, 250, 260, 275, 240, 255,
235, 260, 245, 250, 225, 270, 265, 255, 250, 260]

Step 1: Sample mean

$$\bar{x} = \frac{\sum x_i}{n} = \frac{4965}{20} = 248.25$$

Step 2: Sample standard deviation

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \approx 17.26$$

Step 3: Standard error

$$SE = \frac{s}{\sqrt{n}} \approx \frac{17.26}{4.472} \approx 3.86$$

Step 4: t-critical value (95% CI, df = 19)

$$t_{0.975,19} \approx 2.093$$

Step 5: Margin of error

$$ME = t \cdot SE \approx 2.093 \cdot 3.86 \approx 8.08$$

Step 6: 95% Confidence Interval

$$CI = \bar{x} \pm ME = 248.25 \pm 8.08 = (240.17, 256.33)$$

☑ **Final Answer:**

- **Mean daily sales:** 248.25

- **95% Confidence Interval:** (240.17, 256.33)

Python Code to Compute Mean & 95% CI

```
import numpy as np
import scipy.stats as stats

daily_sales = [220, 245, 210, 265, 230, 250, 260, 275, 240, 255,
               235, 260, 245, 250, 225, 270, 265, 255, 250, 260]

# Sample size
n = len(daily_sales)

# Sample mean
mean_sales = np.mean(daily_sales)

# Sample standard deviation (ddof=1 for sample sd)
std_sales = np.std(daily_sales, ddof=1)

# Standard Error
se = std_sales / np.sqrt(n)

# t-critical value for 95% CI
t_crit = stats.t.ppf(0.975, df=n-1)

# Margin of Error
margin = t_crit * se

# Confidence Interval
```

ci_lower = mean_sales - margin

ci_upper = mean_sales + margin

mean_sales, (ci_lower, ci_upper)