Measure of Central Tendency - Mean, Median and Mode

In [3]:
```python
import numpy as np
import pandas as pd

# Sample data
data = [4, 5, 5, 6, 7, 7, 7, 8, 9, 9]

# Using numpy for mean and median
mean_val = np.mean(data)
median_val = np.median(data)

# Using pandas for mode since numpy does not have a straightforward function for mode
mode_val = pd.Series(data).mode()

# Displaying results
print(f"Mean: {mean_val}")
print(f"Median: {median_val}")
print(f"Mode: {mode_val.iloc[0]}")  # iloc[0] is used to retrieve the first mode if ther

# Note: If the dataset has multiple modes, only the first mode will be printed with the
```

```
Mean: 6.7
Median: 7.0
Mode: 7
```

Hypothesis Testing

In this example, we'll use the scipy library to perform a one-sample t-test. This test determines whether the mean of a sample dataset is statistically different from a known or hypothesized population mean.

Let's assume you have sample data from students' scores in a particular exam, and you want to test if the mean of this sample differs significantly from a hypothesized average score of 50.

Example 1

In [8]:
```python
import numpy as np
from scipy import stats

# Sample exam scores
data = [56, 45, 48, 52, 58, 49, 46, 47, 53, 55]

# Hypothesized population mean
mu = 50

# Perform one-sample t-test
t_statistic, p_value = stats.ttest_1samp(data, mu)

# Display results
print(f"t-statistic: {t_statistic}")
print(f"p-value: {p_value}")

# Interpretation of results
alpha = 0.05  # significance level
if p_value < alpha:
```

```
    print("We reject the null hypothesis: The sample mean is significantly different fro
else:
    print("We fail to reject the null hypothesis: There's no significant difference betw
```

```
t-statistic: 0.6279069767441852
p-value: 0.5456654892990036
We fail to reject the null hypothesis: There's no significant difference between the sam
ple mean and the hypothesized population mean.
```

Example 2

In [11]:
```
import numpy as np
import seaborn as sns
from scipy.stats import ttest_1samp

# Load the iris dataset from seaborn
df = sns.load_dataset('iris')

# Hypothesized population mean for sepal_length (for example, let's take 5.8)
mu_hypothesized = 5.8

# Perform a one-sample t-test
t_stat, p_value = ttest_1samp(df['sepal_length'], mu_hypothesized)

print(f"T-statistic: {t_stat}")
print(f"P-value: {p_value}")

alpha = 0.05
if p_value < alpha:
    print("Reject the null hypothesis: The mean sepal_length of the sample is statistica
else:
    print("Fail to reject the null hypothesis: The data does not provide enough evidence
```

```
T-statistic: 0.6409183514112012
P-value: 0.5225602746220779
Fail to reject the null hypothesis: The data does not provide enough evidence to say the
mean sepal_length is different from the hypothesized mean.
```

In [ ]:
```
!jupyter nbconvert --to webpdf --allow-chromium-download Week7_Lab.ipynb
```