**EXPLORATORY DATA ANALYSIS-13**

[**Data-:**](https://drive.google.com/file/d/1xCGpuDQ3mZVaGSUpg__jysMRV7fiJc3u/view?usp=sharing)

**Introduction:** In this assignment, you will perform an in-depth Exploratory Data Analysis (EDA) on the Amazon (AMZN) stock dataset. The dataset contains the following columns: date, open, high, low, close, volume, adjusted\_close, change\_percent, avg\_vol\_20d.

**Instructions:**

1. Download the Amazon (AMZN) stock dataset provided to you.
2. Import the necessary libraries, including pandas, matplotlib, and any other libraries you may need.
3. Load the dataset into a pandas DataFrame.
4. Perform the following data preprocessing steps:

a. Check for missing values and handle them appropriately.

b. Convert the 'date' column to a datetime data type.

c. Set the 'date' column as the index of the DataFrame.

1. Explore the dataset using pandas functions, and answer the following questions:

**Data Exploration Questions:**

1. What is the range of dates included in this dataset?
2. What is the average closing price of Amazon (AMZN) stock over the entire period?
3. What was the highest closing price, and on which date did it occur?
4. Calculate the 20-day rolling average of the 'close' column and create a new column for it.
5. What is the correlation between the 'close' price and the 'volume' of Amazon stock?
6. Create a new column 'price\_diff' that represents the daily price change (close - open).
7. How many days had a positive price change ('price\_diff' > 0) and a negative price change ('price\_diff' < 0)?
8. Calculate the average daily price change for days with positive price change and negative price change.
9. Plot a line chart showing the daily closing price of Amazon (AMZN) stock over time.
10. Plot a histogram of the daily volume of Amazon (AMZN) stock.
11. Perform more advanced EDA with pandas and visualization techniques:

a. Use groupby to calculate the average adjusted close price for each year in the dataset.

b. Create a line chart to visualize the annual trend in average adjusted close price.

c. Create a box plot to visualize the distribution of daily price changes ('price\_diff').

1. Based on your analysis, provide insights and observations about the Amazon (AMZN) stock data.

**Submission Instructions:**

* Create a Jupyter Notebook to complete the assignment.
* Include comments and explanations for each step in your code.
* Provide answers to the questions in the form of comments or markdown cells.
* Include appropriate data visualizations with titles and labels.