Project Title: Cryptocurrency Analysis with Web Scraping and Pandas

Project members : Y. Sachith Nimesh

Project Description: Cryptocurrency has become a popular investment option, and its market is highly volatile. In this project, we will collect data on various cryptocurrencies using web scraping and analyze it to gain insights into market trends. We will use Python, the Pandas library, and web scraping tools like BeautifulSoup to scrape data from websites like CoinMarketCap or CryptoCompare.

The project will include the following steps:

Collecting Data: We will scrape data from websites that provide information on various cryptocurrencies, including market capitalization, price, volume, and circulating supply. We will use web scraping tools like BeautifulSoup to extract data from HTML and XML documents.

Data Cleaning and Preprocessing: We will clean the data by removing any duplicates or errors, handling missing values, and converting data types as needed. We will preprocess the data to make it suitable for analysis using Pandas.

Exploratory Data Analysis: We will perform exploratory data analysis on the collected data to gain insights into market trends. We will use Pandas to explore the relationships between various cryptocurrency metrics, visualize the data using matplotlib, and identify outliers and trends.

Statistical Analysis: We will perform statistical analysis on the data to identify any significant correlations or patterns. We will use Pandas to perform statistical tests like correlation analysis, regression analysis, and hypothesis testing.

Predictive Analysis: We will use the historical data to build predictive models that can forecast future trends in the cryptocurrency market. We will use machine learning algorithms like Linear Regression, Random Forest Regression, and Gradient Boosting Regression to build the models.

Deliverables:

Python code for web scraping and data analysis using Pandas.

A Jupyter notebook documenting the analysis process and findings.

Visualizations of the data using matplotlib.

Predictive models for forecasting future trends in the cryptocurrency market.

Timeline: Week 1: Collect data using web scraping tools and preprocess it using Pandas. Week 2: Perform exploratory data analysis on the collected data. Week 3: Perform statistical analysis and build predictive models. Week 4: Document the analysis process and findings in a Jupyter notebook.

Conclusion: This project will provide an opportunity to gain hands-on experience in web scraping and data analysis using Python and Pandas. It will help understand the dynamics of the cryptocurrency market and build predictive models that can forecast future trends.

# Cryptocurrency Analysis with Web Scraping and Pandas ¶

First, let's import the necessary libraries:

In [13]:
```
import pandas as pd
import requests
from bs4 import BeautifulSoup
```

Next, we'll define a function to scrape data from a website that lists various cryptocurrencies and their market values. Here's an example of such a website: https://coinmarketcap.com/all/views/all/

In [14]:
```
def scrape_cryptocurrencies():
    url = 'https://coinmarketcap.com/all/views/all/'
    res = requests.get(url)
    soup = BeautifulSoup(res.content, 'html.parser')
    table = soup.find('table', {'id': 'currencies-all'})
    rows = table.find_all('tr')

    data = []
    for row in rows:
        cols = row.find_all('td')
        cols = [col.text.strip() for col in cols]
        if len(cols) == 10:
            data.append(cols)

    columns = ['Rank', 'Name', 'Symbol', 'Market Cap', 'Price', 'Circulating Supply', 'Volume (24h)', '% 1h', '% 24h', '% 7d']
    df = pd.DataFrame(data, columns=columns)
    df = df.drop(['Rank'], axis=1)
    return df
```

This function sends an HTTP request to the website, scrapes the data from the HTML using BeautifulSoup, and then converts the data into a Pandas DataFrame. It returns the DataFrame containing the data on various cryptocurrencies.

We can now call this function to get the data and analyze it. For example, we can get the top 10 cryptocurrencies by market cap:

In [15]:
```
df = scrape_cryptocurrencies()
top_10_by_market_cap = df[['Name', 'Market Cap']].head(10)
print(top_10_by_market_cap)
```