

Vallum Assignment

"1. What is equity investing, and how would you define it?"

Investing in company stocks to earn returns through dividends and capital appreciation. It involves owning shares of publicly traded companies.

2. Can you explain the GARP style of investing?

Growth at a Reasonable Price (GARP) combines growth and value investing, focusing on companies with steady growth and reasonable valuations.

Key metric: PEG ratio (P/E ratio divided by earnings growth rate).

3. When analyzing companies listed on BSE & NSE, how do you differentiate between GARP style opportunities, growth-only, and value-only companies?

- **GARP:** Moderate growth and reasonable valuations (PEG around 1).
- **Growth-Only:** High growth, high P/E and PEG ratios.
- **Value-Only:** Undervalued with low P/E and P/B ratios, but slower growth.

4. How comfortable are you with Python? Please provide details about your knowledge and practical application level.

I am proficient in Python with extensive experience in web scraping, data cleaning and data analysis. I use libraries such as:

- **Pandas** for data manipulation and analysis.
- **NumPy** for numerical computations.
- **Matplotlib** and **Seaborn** for data visualization.
- **BeautifulSoup** and **Selenium** for web scraping.
- **Jupyter Notebooks** for interactive coding and visualization.

I have created a project called "Cars24 Data Analysis" where web scraping and cleaning has been done with the help of Python and its libraries.

Kindly go through the repository link for more details about project:

[**https://github.com/gauravsyadav/Cars24_Data_Analysis**](https://github.com/gauravsyadav/Cars24_Data_Analysis)

5. Using Python, how would you determine if a company listed on BSE & NSE follows the GARP style, considering the available data for approximately 6000 companies?

- **Data Collection:** Use web scraping to collect financial data including earnings growth, P/E ratios, and other relevant metrics.
- **Data Cleaning:** Handle missing values and outliers using Pandas.
- **Calculation:** Compute PEG ratios and filter companies with PEG ratios around 1.
- **Analysis:** Use Pandas for filtering and NumPy for calculations. Plot distributions and trends using Matplotlib/Seaborn.

6. Based on your knowledge, what insights can you derive and showcase about the following stocks: SBIN, Adani Enterprises, HUL, Tata Steels, Moil?

- **SBIN (State Bank of India):** A leading public sector bank in India, focusing on retail banking and digital transformation.
- **Adani Enterprises:** A conglomerate with interests in diverse sectors such as energy, resources, logistics, agribusiness, and more.
- **HUL (Hindustan Unilever Limited):** A major FMCG company known for its strong portfolio of consumer goods.
- **Tata Steels:** One of the largest steel manufacturing companies, involved in the production and distribution of steel.
- **Moil (Manganese Ore India Limited):** India's largest producer of manganese ore, catering to the needs of the steel industry.

7. Are you familiar with web scraping techniques?

Yes, I am familiar with web scraping. Techniques include:

- **BeautifulSoup:** For parsing HTML and XML documents.
- **Selenium:** For automating web browsers to scrape dynamic content.
- **Process:** Identify target URLs, request data, parse HTML, extract relevant information, and store in a structured format like CSV or a database.

8. If tasked with extracting the number of NRIs across PMSs from SEBI's monthly reports for June '23, Sep '23, Dec '23, and Mar '24, how would you approach this task in terms of process, time, output file, and data accuracy?

- **Process:**
 1. Identify the URLs or sources for SEBI's monthly reports, write a script to parse the relevant HTML files and extract the texts.
 2. Clean and structure the data using Python.
 3. Validate the extracted data against the source.
- **Time:** Depending on the complexity and volume, it may take a few hours to a day to complete.
- **Output File:** Store the extracted data in a structured format such as CSV or Excel.
- **Data Accuracy:** Ensure accuracy by cross-referencing extracted data with original reports and handling errors.

9. What configuration of devices do you believe is necessary to perform these tasks on a daily basis?"

- **Hardware:** A latest laptop or desktop with at least an Intel i5 or AMD Ryzen 3 or equivalent processor, 8GB RAM or more and SSD storage for faster performance.
- **Software:**
 - Python environment with Jupyter Notebook or PyCharm for coding and analysis.
 - Essential libraries like Pandas, NumPy, Matplotlib, Seaborn for data manipulation and visualization.
 - Tools such as BeautifulSoup, Scrapy, and Selenium for web scraping tasks.
- **Internet:** Reliable and fast internet connection.