JGP Macro Data Scientist Internship Selection Process Test

> **JGP Macro Data Scientist Internship Selection Process Test**

Task

You are a data scientist tasked with analyzing and automating a process for capturing, transforming, and displaying inflation data from the United States.

Instructions:

1. Place the answer to each question below in Python (.py) files.

2. Include a `requirements.txt` file with your code's dependencies.

3. If necessary, describe your ideas in a PDF file that covers all the questions.

4. Store all your test's files and outputs in a GitHub repository, and then send us an email containing the link to the repository.

5. Upon receiving the email, you will have **until 07/30/2024 at 11:59PM** to complete the test and send us your response (step 4).

Question 1

Obtain the U.S. inflation series (CPI) from the BLS (Bureau of Labor Statistics) — https://www.bls.gov/ — via API, process the data, and format it so that the columns correspond to the series and each row represents an observation over time. Save the result in a CSV file. Specifically, obtain the following national-level monthly series:

- CPI All items, seasonally adjusted

- CPI All items, less food and energy, seasonally adjusted

- CPI Gasoline (all types), seasonally adjusted

Example format:

date	series 1	series 2	series 3
1/1/1959	15.16	100.00	100.00
2/1/1959	15.18	100.03	100.03
3/1/1959	15.19	100.10	100.10
4/1/1959	15.22	100.07	100.07
5/1/1959	15.23	100.17	100.17
6/1/1959	15.27	100.63	100.63
7/1/1959	15.30	100.82	100.82
8/1/1959	15.33	100.79	100.79
9/1/1959	15.37	101.40	101.40

Pivot time series format.

Question 2

Using Plotly in Python, develop a chart displaying the **All items, less food and energy, seasonally adjusted** price series with year-over-year percentage variation using the monthly data from 2019 to the present. Keep the frequency monthly in the chart.

Question 3

Describe in words how you would automate the process of extracting the data.

Question 4

Explain how you would relate the price series (All items) with the Gasoline (Gasoline) price series.

BONUS: Implement this solution.

Bonus Question

With the previous questions completed, implement an application using FastAPI to allow requests to the data you stored in the CSV file.