Data Processing and Visualization Report

1. Dataset Chosen

Title: Unemployment by sex and age – monthly data (une_rt_m)

Portal: Eurostat Open Data

Link: https://ec.europa.eu/eurostat/api/dissemination/statistics/1.0/data/une_rt_m

Reason for choosing: This dataset is clean, openly accessible, updated regularly, and includes monthly unemployment data for Lithuania. It enables easy computation of yearly averages.

2. Access Method

- Accessed using Eurostat JSON API.

- Parameters: geo=LTU, sex=T, age=Y15-74, s_adj=SA, unit=PC_ACT.

- Data retrieved with Python requests and parsed as JSON.

3. Data Cleaning & Transformations

1. Extracted monthly unemployment values.

2. Filtered years 2015–2024.

3. Converted date strings to datetime.

4. Built DataFrame.

5. Resampled to annual averages.

6. Rounded values.

7. Saved as processed.json.

8. Placed in React /public folder.

4. Final Output (Visualization)

- Interactive chart using React + Recharts.

- Displays annual unemployment trend.

- Includes table and interpretation.

Interpretation:

Lithuania's unemployment rate shows improvement over the last decade, with a downward trend despite minor fluctuations.

5. Obstacles & Solutions

1. Incorrect API codes → fixed using metadata.

2. Complex SDMX structure → mapped dimensions correctly.

3. Partial year data → filtered invalid years.

4. React integration → ensured JSON is in /public.

6. Licensing

Eurostat data licensed under CC BY 4.0.

7. Reproducibility

- Python runs with no manual steps.

- Running main.py updates processed.json.

- React loads visualization automatically.

8. Conclusion

The project meets all assignment requirements: data fetching, cleaning, indicator computation, and visualization.