**Did you find all the data and was able to access the data you reported in your first report?**

I found all the data, two are available directly with Python scripts, and I need to manually download the other one so that It can preprocess with Python scripts.

Annual Wage Statistics: the websites shows 403 prohibitions for scrape robots, so I tried but cannot access it successfully, so I found an alternative approach to download the mirror data from other websites, then I used Python to gather and collect data from 13 years (2012-2024) into a single CSV result instead of doing it manually. The processed data contains 13 years (rows) and 7 columns.

U.S. Software Developer Salaries: I use kagglehub package to fetch data via API. Different from other Kaggle API approaches, it does not need API keys and it's more convenience. The processed data contains 76 cities (rows) and 4 columns.

Software Engineer Salary: I scrape the websites with request, BeautifulSoup packages, and analyze with regex expressions, I get the data successfully. The processed data contains 12 cities (rows) and 4 columns.

**Did you start implementation of your project**

I finished all the tasks generally, except some polish and details for better readability.

**Did you test your data APIs?**

I test all the data APIs, and it works successfully.

**Technical update**

**Briefly describe here what python packages your plan to use. Do you plan to do a live demo (running code in class) during in class presentations?**

For data preprocessing, I use requests, kagglehub packages for scraping data, use BeautifulSoup for analysis HTML data, use re, os, shutil, regex` standard packages, and pandas package for file I/O and text analysis. For data analysis, I use `matplotlib, seaborn` for plotting data, `pandas, numpy, sklearn, statsmodels` package for calculation.

**What output will your project create (e.g. a data set, notebook with analysis, a web page/service, )?**

The codes can run as a live demo.

The output of my project are: Datasets extracted from webs, and the results and conclusions of the dataset analysis with a notebook.