

## Abstract

Due to Reddit's current regulations on data scraping from the forum, there are limits on the number of scrape and the timeframe allowed. Therefore, if you, the evaluating professor, use my code to directly fetch data, the data will not correspond to the exact moment when I conducted my research, leading to discrepancies between the output data and the content in my Business Project. To this end, I have provided the data as originally captured in its CSV file format, to facilitate the smooth operation of my code.

### 1. Folder Links (Students provide two options, professors may use either one)

- Github:

<https://github.com/peiyuntsai/BEMM466-Business-Project.git>

- Onedrive:

[https://universityofexeteruk-my.sharepoint.com/:f:/g/personal/pt493\\_exeter\\_ac\\_uk/Ep8n5ycukWRHkbA3t1ne\\_VsBr7kHONDoxkL1GGdwfmQbAg?e=6aXgjW](https://universityofexeteruk-my.sharepoint.com/:f:/g/personal/pt493_exeter_ac_uk/Ep8n5ycukWRHkbA3t1ne_VsBr7kHONDoxkL1GGdwfmQbAg?e=6aXgjW)

### 2. Code

- Web scraping and merging of two subreddit datasets
  - File name: Reddit Scrape.ipynb
- NLP data analysis
  - File name: Reddit EV Analysis.ipynb

### 3. Data

- Original data (from May 8, 2024, to July 17, 2024)
  - (1). File name: relectricvehiclesdata.csv
  - (2). File name: relectriccarsdata.csv
- Merged original data (subsequent data analysis will use this file)
  - File name: merged\_reddit\_posts.csv

4. Official Reddit Website: <https://www.reddit.com/>