**DAT 205 Meeting Minutes Feb 25, 2021 at 2000hrs ET**

**Topic:** Team meeting – Project proposal and dataset selection

**Participants:**

Bhavika Patil

Cindy Guo

Dennis Hung

**Discussion**

* Dataset options
  + Option 1: NBA API
    - <https://www.playingnumbers.com/2019/12/how-to-get-nba-data-using-the-nba_api-python-module-beginner/>
  + Option 2: NYC Taxis
    - <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>
    - 20 features
    - Approximately 72GB of data from 2009 to 2020 June
    - New York City taxis only
  + Investigate dataset via possible API’s
    - <https://sports-statistics.com/sports-data/sports-data-sets-for-data-modeling-visualization-predictions-machine-learning/>
      * Sports Datasets for Data Modeling, Data-Vis, Predictions, Machine-Learning
    - <https://pypi.org/project/python-nasa-api/>
      * python-nasa-api Simple NASA Api Wrapper written in Python!pypi.org
    - <https://www.springboard.com/library/data-science/top-apis-for-data-scientists/>
      * 22 APIs Every Data Scientist Should Know. APIs are the essential building blocks for data science. They provide key data sources and enable data integration and visualization. In this blog, we look at the most important APIs and how to lev...www.springboard.com
    - <https://www.toptal.com/python/twitter-data-mining-using-python>
      * Twitter Data Mining: A Guide to Big Data Analytics Using Python. We will take a look at how to use Tweepy to gather data from Twitter using Python, followed by various usage examples.www.toptal.com

Timelines

|  |  |
| --- | --- |
| **Date** | **Milestone** |
| Sun, Feb 28 | Decide on dataset |
| Tue, Mar 2 | Finalized Initial proposal |
| Thu, Mar 4 | Submit initial proposal |

**Action Item**

Group:

* Investigate further into other possible datasets by Feb 28