DS4300 Project Proposal

Names: Srishti Kundu, Aanay Anandpara, Rosan Wang, Eirean Co, Wai Dzing Soo
NoSQL: MongoDB, Python
Data Source:
1. Twitter API
   a. A developer account from Twitter provides access tokens which fulfills the authentication requirements needed for us to make queries with Python.
   b. The Tweepy package allows us to make calls to the Twitter API such as searching for tweets based on keywords, and more.
   c. Using the output from Tweepy/Twitter API, we can populate our mongoDB using PyMongo.

2. Inflation indicators from the U.S. Bureau of Labor Statistics
   a. Inflation is the change over time in the prices of goods and services. To track inflation, the U.S. The Bureau of Labor Statistics (BLS) calculates and publishes the consumer price index (CPI) on their website. The CPI captures the price of a "market basket" of goods and services, i.e. expenses of average American families and individuals
   b. We intend to use the most general CPI, which includes all items in the CPI market basket averaged across all U.S. cities. We can use this data to compute inflation rates as the percent change in the CPI between time periods.

Motivation: According to the Federal Reserve, inflation expectations play a significant role in determining inflation.[1] Consumer sentiment influences investment decisions, which in turn influences economic conditions and outcomes. For example, if consumers are optimistic about the future, they are more likely to invest and consume, which boosts the economy. In many ways, economic conditions can be a self-fulfilling prophecy. If consumers believe recession is coming, they will reduce their spending which in turn will eventually lead to recession-like conditions.

Therefore, analyzing the sentiments of consumers can provide a lot of valuable information on predicting economic conditions. By analyzing this relationship, the information can be useful for policymakers and businesses to make informed decisions regarding economic policies, investments, and business strategies.

Description:
- Perform sentiment analysis on tweets about the economy. Potential libraries include VADER (relevant for social media), NLTK and TextBlob
- Examine the relationship between sentiment and inflation indicator. Is there a correlation between economic sentiment and economic indicators like inflation?

---

[1] https://www.federalreserve.gov/monetarypolicy/files/FOMC20091201memo05.pdf