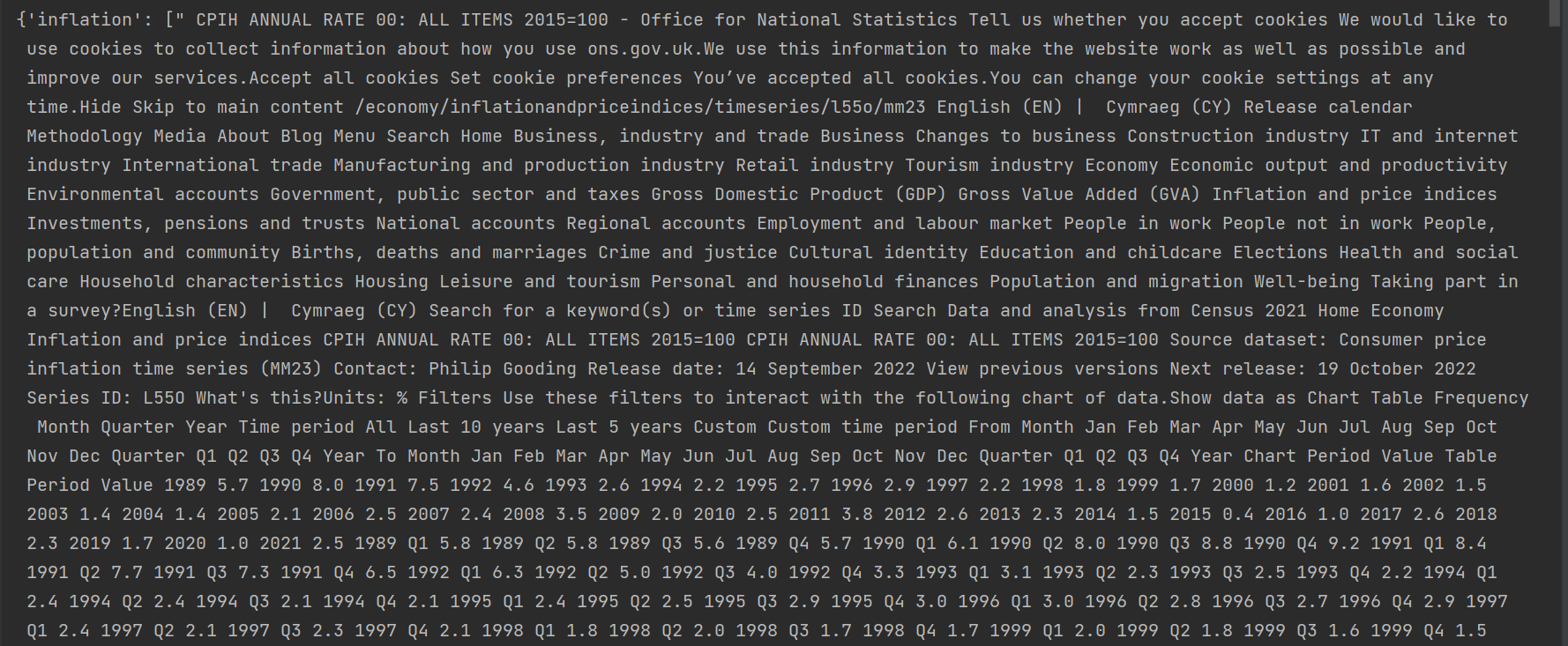
CS 4395 Webcrawler Report

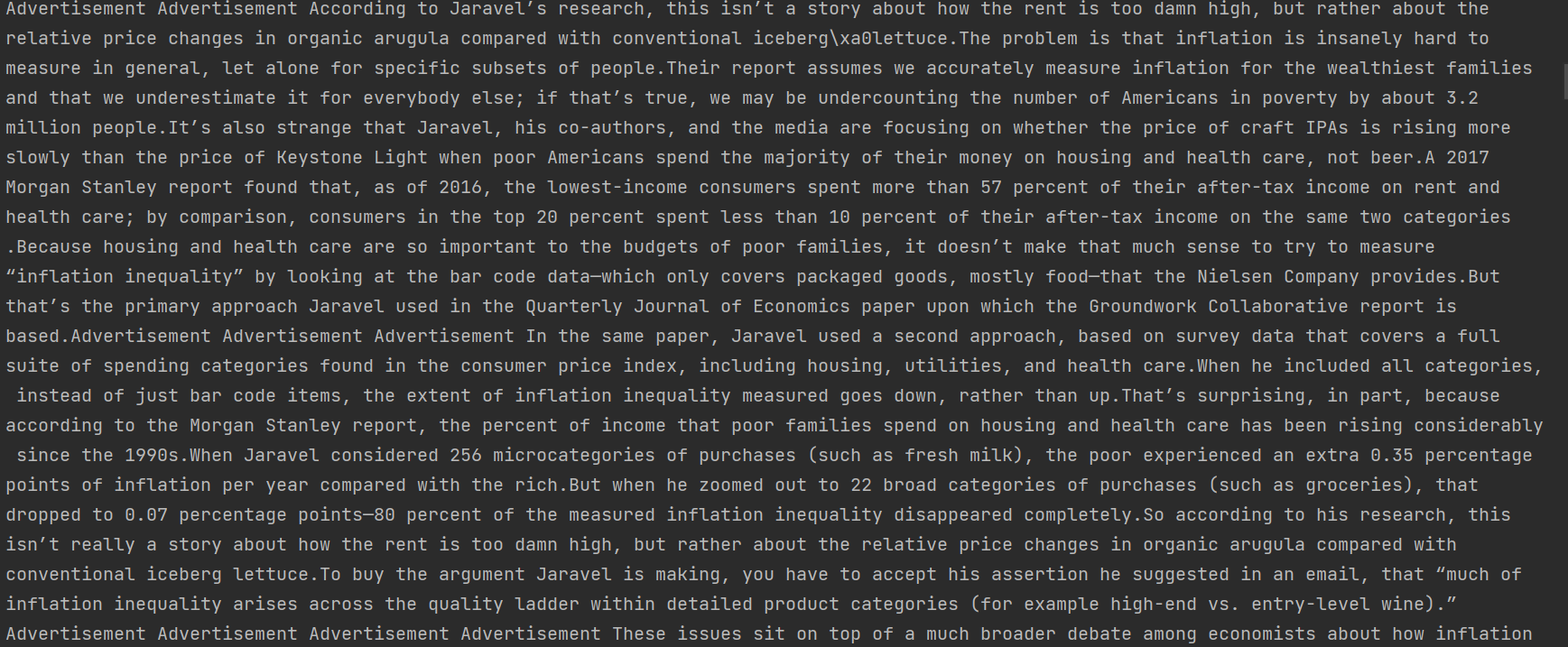
By: Harshith Nanda & Charith Muppidi

**How did we create our knowledge base?**

We created our knowledge base by scraping several, relevant webpages, processed the webpage by removing unnecessary characters like newlines, then retrieved the top 40 most frequently appearing words, then narrowed it down to the top 10, which were printed out. Our knowledge base consisted of a python dictionary with the most frequently appearing words as the key, and a list of sentences they appear in as the value.

For our project, we chose an article about Inflation and scraped relevant articles about the economy and government policies related to inflation.

Here is a short sample of the knowledgebase: 



A picture containing text, outdoor

Description automatically generated

The ten most frequently occurring words are:

1. inflation
2. price
3. money
4. state
5. used
6. goods
7. rate
8. vicinity
9. modifier
10. supply

Sample Dialog with a Chatbot:

User: Hello!  
Bot: Inflation.  
User: Interesting, what about inflation would you like to tell me?  
Bot: Inflation is economy   
User: What is the inflation rate for 2022?   
Bot: 1984  
User: That’s the year, not the inflation rate. What is the inflation rate?  
Bot: 1992 3%  
User: What about the inflation rate for 2022?  
Bot: Federal reserve radically sends weapons to Ukraine to combat inflation  
User: I see. Has the president done anything about inflation?   
Bot: Fed chair Paul Volcker causes Russia-Ukraine war, leading to inflation in food prices around the world  
User: The fed chair caused the war? That’s quite the accusation to make. You’re rather conspiracy minded, aren’t you?  
Bot: Jerome Powell increases interest rates by $3.5 trillion to combat inflation  
User: You mean 3.5%?   
Bot: The fed funds rate set to Biden the next few months  
User: Yeah I’m gonna need to shut you down.   
Bot: 1984