Names:Liliane Umukunzi
Task 1.5
Data Visualization with Python

Exercise

1. Install the necessary libraries from the requirements.txt file into your corresponding virtual environment. Make sure you start JupyterLab with the extra data rate argument.
2. Start a new notebook in JupyterLab and import all of the libraries you'll need.
3. Load the twentieth-century data that you scraped in Exercise 1.4.
4. Tokenize the words from the text and create a bar chart to plot the 10 most common words.
5. Remove stop words (and punctuation marks if necessary), and rerun the plot. What has changed? Comment on your impressions in a markdown cell under the plot.
6. Create a tags list using TextBlob.
7. Create a list of the Top 10 POS tags for words that appear in the article.
8. Plot this list in a bar chart using the seaborn library. Note: If you were with us for the Data Immersion Program, you've already learned about seaborn in Achievement 4. If Seaborn is new to you, use this introductory resource to walk you through it.
9. Create three bar plots with the top 15 POS labels—one each for nouns, verbs, and adjectives. Analyzing the results—what do you think stands out in these frequencies? Comment on your impressions in a Markdown cell under the plot.
10. Create a dataframe with the countries and the number of times they're mentioned in your text (remember—you'll need to use the countries lookup list if you didn't scrape one in the previous task).
11. Create a plot showing the frequency of the number of times these countries are mentioned in the text. Is there anything in particular you notice in this plot? Describe your findings in a Markdown cell underneath the plot.
12. Save your notebook, download it, push the changes to your GitHub repo, and send it to your mentor for review.

Github repo - https://github.com/liliane774/20th-Century