Objective

The aim of this project is to provide a graphical interpretation of the statistics of Coronavirus all over the world. The data source for this project is the Worldometer website. The data is visualized via graphs, which helps in a better understanding. The spread of Coronavirus is very fast, and as days pass, misinformation about the data is very hard to detect. Therefore, this project is a step in the direction of providing well researched facts and numbers in a form that can be read and interpreted easily.

The Worldometer website

Worldometer is a website run by an international team of developers, researchers, and volunteers with the goal of making world statistics available in a thought-provoking and time relevant format to a wide audience around the world. It is published by a small and independent digital media company based in the United States. For COVID-19, they collect data from official reports, directly from Government's communication channels or indirectly, through local media sources when deemed reliable. In this project, Coronavirus data from the Worldometer website has been taken and analysed.

The website can be found here: https://www.worldometers.info/coronavirus/

Web Scraping

Web Scraping (also termed Screen Scraping, Web Data Extraction, Web Harvesting etc.) is a technique employed to extract large amounts of data from websites whereby the data is extracted and saved to a local or to a database in table (spreadsheet) format. Data displayed by most websites can only be viewed using a web browser. Web Scraping is the technique of automating the process of retrieving data from the website so that instead of manually copying the data, the Web Scraping program performs the same task within a fraction of the time it used to take a human.

BeautifulSoup: Web scraping with Python

BeautifulSoup is a Python library used for parsing documents (i.e. mostly HTML or XML files). Using Requests to obtain the HTML of a page and then parsing whichever information you are looking for with BeautifulSoup from the raw HTML is the quasi-standard web scraping “stack” commonly used by Python programmers. Here, the BeautifulSoup library has been used to scrape information from the Worldometer website and saved in csv files.

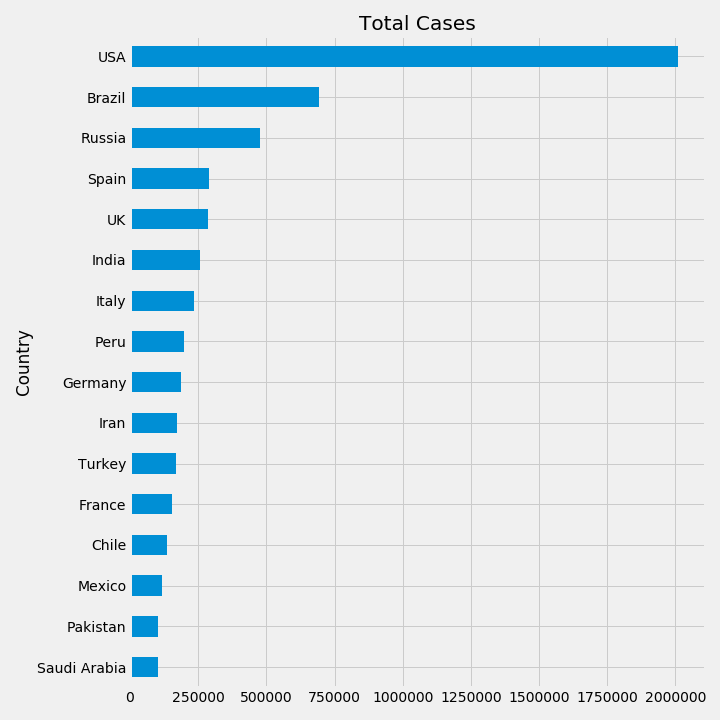
Matplotlib: Visualization with Python

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Here, Matplotlib has been used to plot data that has been scraped from the Worldometer website into four different plots.

Data visualization its importance

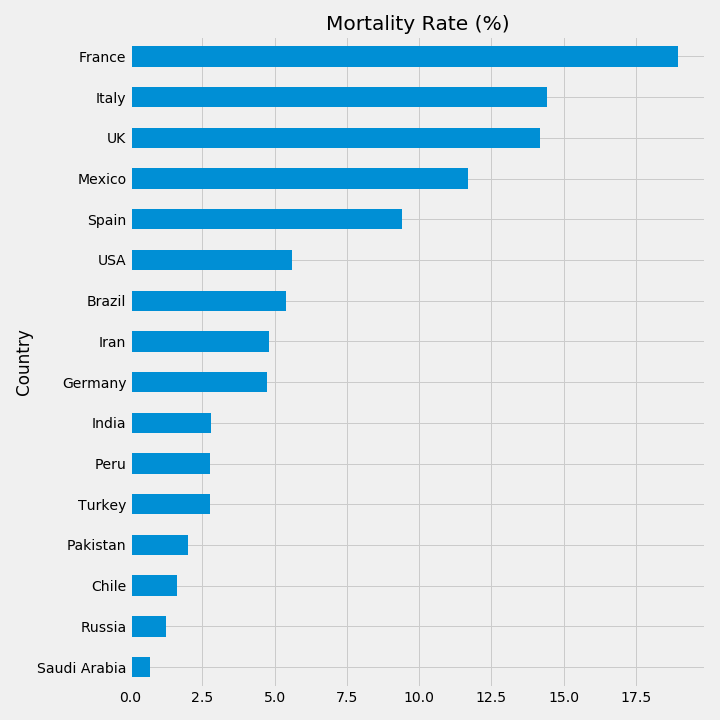
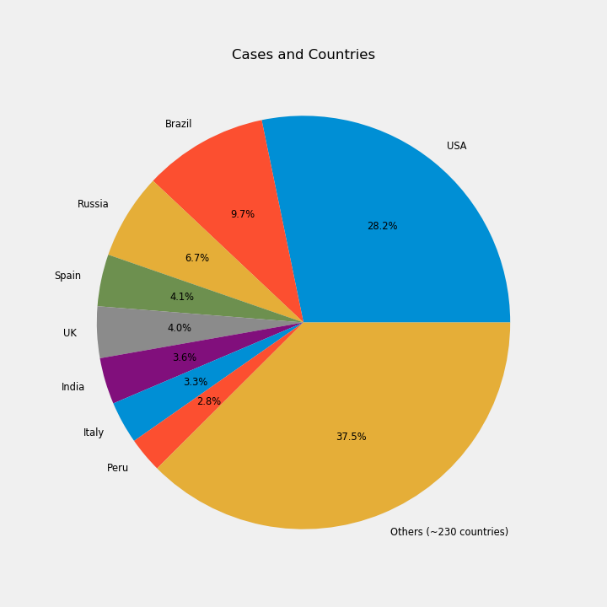
Data visualization is the representation of data or information in a graph, chart, or other visual format. It communicates relationships of the data with images. This is important because it allows trends and patterns to be more easily seen. Data visualization provides a summary of information and makes it easier to identify patterns and trends than looking through thousands of rows on a spreadsheet. It’s the way the human brain works. Since the purpose of data analysis is to gain insights, data is much more valuable when it is visualized. Charts and graphs make communicating data findings easier even if you can identify the patterns without them.

A screenshot of a cell phone

Description automatically generatedThese are the graphs that the program has generated, as of 8th June, 2020.

Total Deaths: countries with highest number of deaths due to Coronavirus recorded, arranged from greatest to lowest number of deaths.

Total Cases: countries with highest number of Coronavirus cases recorded, arranged from greatest to lowest number of total patients.



Total Cases: Percentage of total cases that every major country is contributing.

Mortality Rate (%): Countries with highest mortality rate of Coronavirus patients, arranged from greatest to lowest*.*

*Total Cases: countries with highest number of Coronavirus cases recorded, arranged from greatest to lowest number of total patients.*