

5. Web Scrapping and Horizontal Bar Plots

Prepared by Prajwal Singh

Goal

- Reproduce a recent bar plot, "Killer Quakes", from Sept 10, TOI (see pictures at the end for reference). The source of the data file has been mentioned to be the following - <https://ourworldindata.org>
- The goal is to acquire the data from the website, plot an equivalent bar chart from the acquired data and verify the match (and mismatch if any).

Details

- The plot uses data from two different sets of data available at the following website - <https://ourworldindata.org>.
- One dataset is available as a page table in the following URL - <https://ourworldindata.org/the-worlds-deadliest-earthquakes>.
- The other set is available from a natural disaster dataset available in the **OurWorldInData** servers.

Data Acquisition

- Use the pandas read_html method to Web Scrape data from the following website <https://ourworldindata.org/the-worlds-deadliest-earthquakes>.
- See examples from https://pandas.pydata.org/docs/user_guide/io.html#html.
 - Note that read_html returns an array of tables.
- Acquire the natural disaster dataset using owid-catalog API:
 - See: <https://pypi.org/project/owid-catalog/>
 - See: <https://github.com/owid/owid-catalog-py>

You must install owid-catalog and import catalog as "from owid import catalog" before use.

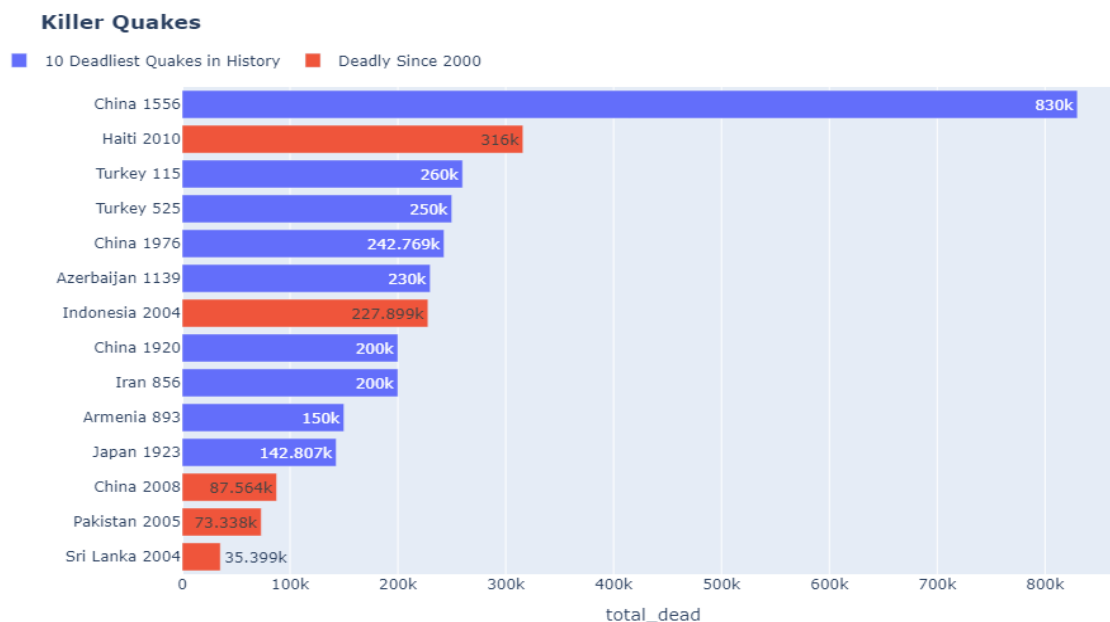
Data Preparation and Cleaning Component

- From the owid catalog acquired data choose only those rows that correspond to earthquakes and are for years 2020 or higher.
- Remove data rows to "World", "European Union (27)" and Continents. (you may list the continents manually or use a python API such as a-world-of-countries to get_continent_list. See <https://pypi.org/project/a-world-of-countries>).
- Additionally remove the names from the following list: ['Upper-middle-income countries', 'Lower-middle-income countries', 'High-income countries', 'Low-income countries'].
- Sort the resulting data frame in decreasing order, and pick only 5 rows from the data frame to concatenate with the scraped dataset. Before concatenating make sure that both the data frames have the same column names.
- To avoid duplication of rows, use the pandas data frame method `drop_duplicates(subset=["country", "year"])`.

Data Visualization

- Use bar plot method of Plotly express or Seaborn to plot horizontal bar plots.
- Note that you need to set the color (for Plotly), hue for (seaborn) to assign different hues to distinguish between bars before 2020 and the others.

Sample Image

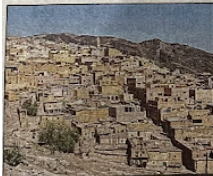


Moroccan homes of masonry & stone likely to raise quake toll

Mountain Villages Hard To Reach; 12th Century Mosque Damaged

Moulay Brahim/Marrakech: A rare, powerful earthquake struck Morocco on Friday night, sending people racing from their beds into the darkened streets and toppling buildings in mountainous villages and ancient cities not built to withstand such force. Updated interior ministry figures on Saturday showed the quake killed at least 1,037 people, the vast majority in Al-Haouz, the epicentre, and Taroudant provinces.

Marrakech's famous Koutoubia Mosque, built in the 12th century, was damaged, but the extent was not immediately clear. Its 69-metre minaret is known as the "roof of Marrakech." Moroccans posted videos showing damage to parts of the famous red walls that surround the old city, a



The Moulay Brahim village, near the epicentre of the earthquake

Unesco World Heritage site.

The epicentre of tremor was near the town of Ighil in

KILLER QUAKE

Ten deadliest quakes in history		Five deadliest since 2000	
	Year		Approximate toll
Shaanxi China	1556		8,30,000
Port-au-Prince Haiti	2010		3,16,000
Anatakya (Antioch) Turkey	115		2,60,000
Anatakya (Antioch) Turkey	525		2,50,000
Tangshan China	1976		2,43,000
Gyzndzha Azerbaijan	1139		2,30,000
Indian Ocean tsunami	2004		2,28,000
Damghan Iran	856		2,00,000
Gansu China	1920		2,00,000
Dvin Armenia	693		1,50,000
Tokyo Japan	1923		1,43,000
POK & NWFP Pakistan	2005		87,350
Sichuan China	2008		69,200
Central Turkey	2023		59,260



Sources: ourworldindata.org

the High Atlas, was the worst hit of any *Reuters* journalists saw in rural areas south of Marrakech where officials said most of the more than 1,000 deaths had occurred.

The quake brought down walls made from stone and masonry not designed to withstand quakes, covering whole communities with rubble and leaving residents picking their way precariously through remains. "The problem is that where destructive earthquakes are rare, buildings are simply not constructed robustly enough to cope with strong ground shaking, so many collapse, resulting in high casualties," said Bill McGuire, professor emeritus of geophysical and climate hazards at University College London. "I would expect the death toll to climb into the thousands once more is known. As with any big quake, aftershocks are likely, which will lead to further casualties and hinder search efforts."

Strong tremors were also felt in the coastal cities of Rabat, Casablanca and Essaouira. "I was nearly asleep when I heard the doors and the shutters banging," said Ghannou Na'iem, a Casablanca resident in her 80s who was visiting Marrakech when the quake hit. In the mountain village of Moulay Brahim near the

Algeria suspends ban, opens airspace

Algiers: Algeria will allow aircraft carrying humanitarian aid to