

Assignment 5: Group Project Interactivity

Team Mask (Group 09)
Paderborn University

June 17, 2020

1 Introduction

At present, where University, Organization, Institution, Restaurant and many more are entirely shut or made online except Hospitals which are working 24*7, yes its **COVID-19**. A pandemic which has shaken the economy of the world. As the battle continues, data sets corresponding to the virus disease are collected around the globe. This concept paper is on the Topic: COVID-19 as part of the Assignment to the subject Module: Interactive Data Visualization which we here summarize about the dataset of COVID-19 disease used, visualizations, and interaction end-user can probably expect.

2 Data Characteristics

The dataset characteristics are described in the Table 2. The dataset *owid-covid-data.csv* is comprised of the dataset sources listed in the Table 1 and is available as comma separated strings. The dataset contains 22,873 records and 33 columns (on June 12, 2020). Also, dataset containing the polygons, used to form a World Map for representing geo spatial mapping information is sourced from Natural Earth (Admin-0 Countries).

European Centre for Disease Prevention and Control	National government reports
Department of Economic and Social Affairs	United Nations
World Population Prospects: The 2019 Revision	Population Division
OECD	Eurostat
World Bank – World Development Indicators, sourced from Food and Agriculture Organization and World Bank estimates	UN Population Division, World Population Prospects, 2017 Revision

Table 1: Sources

3 User and Task

The *owid-covid-19-data* dataset captures information of disease spread and fatalities caused around the globe and also counter measure like testing taken by respective countries.

Tasks for the Visualization can be as follows:

- Visualization of disease information country-wise taking into account of multivariate data like infected cases, fatality count, total cases and total deaths per million.
- Visualization by comparison of the total cases and fatalities reported daily among the continents and top nations.
- Visualization by comparison of multivariate data for each country/countries. For example: total cases vs total deaths.
- Visualization of increasing or decreasing trends of total cases/total deaths of top nations.

Probable target users:

- General public can get to know the statistics of COVID-19 over the globe.

Data	Data Type	Description
iso_code	String	ISO 3166-1 alpha-3 (3 letter country codes)
continent	String	Continent name
location	String	Country name
date	String	Date of observation
total_cases	int64	Total confirmed cases of COVID-19
new_cases	int64	New confirmed cases
total_deaths	int64	Total deaths attributed
new_deaths	int64	New deaths
total_cases_per_million	float64	Total confirmed cases per one million people
new_cases_per_million	float64	New confirmed cases per one million people
total_deaths_per_million	float64	Total deaths per one million people
new_deaths_per_million	float64	New deaths per one million people
total_tests	float64	Total tests conducted
new_tests	float64	New tests conducted
new_tests_smoothed	float64	New tests (7 day smoothed). For countries not reporting data on daily basis.
new_tests_per_thousand	float64	New tests per thousand people
new_tests_smoothed_per_thousand	float64	New tests per thousand people (7 day smoothed)
tests_units	float64	Units used by the location to report the testing data
population	float64	Population in 2020
population_density	float64	Number of people divided by land area, measured in square kilometers, most recent year available.
median_age	float64	Median age of the population, UN projection for 2020.
aged_65_older	float64	Share of the population that is 65 years and older.
aged_70_older	float64	Share of the population that is 70 years and older in 2015.
gdp_per_capita	float64	Gross domestic product at purchasing power parity (constant 2011 international dollars), most recent year available.
extreme_poverty	float64	Share of the population living in extreme poverty, most recent year available since 2010.
cvd_death_rate	float64	Death rate from cardiovascular disease in 2017.
diabetes_prevalence	float64	Diabetes prevalence (% of population aged 20 to 79) in 2017.
female_smokers	float64	Share of women who smoke, most recent year available.
male_smokers	float64	Share of men who smoke, most recent year available.
handwashing_facilities	float64	Share of the population with basic handwashing facilities on premises, most recent year available.
hospital_beds_per_thousand	float64	Hospital beds per 1,000 people, most recent year available since 2010.

Table 2: Data Characteristics

- WHO and independent government bodies of countries can make use of the visualizations for better planning of the medical infrastructure.
- Monetary Institution like: World Bank can estimate the impact of the losses particular to a region and aid respectively.
- Travel and Tourism industry and service providing organization can utilize these visualizations for better trip planning.

4 Visualization Technique

- **Time Visualization:** A Time-Bar-Chart which compares country/continent versus cases/fatalities reported taking into account the time-oriented data(year/month/day), each bar combines time-series information with respect to day/month.
- **Parallel Coordinates:** The visualization of the dataset can be done by grouping the top 10 affected countries together and then compare against each other, the various factors like deaths, active cases, hospital beds and more in order to get an idea of how each country fares in comparison with the others.
- **Geospatial Visualization:** A choropleth world map that differentiates the cases reported/fatalities/case-per-million/deaths-per-million using brightness as a visual variable based on locations and their associated data values.
- **Line-based Technique for Multivariate Data:**

5 Interaction

- **Navigation and Selection:** Medical Institutions/Scientists can use the click and drag, double clicking(zooming) feature, plus and minus keys to zoom to change the level of detail on the choropleth map. And further using the selection feature(hovering) by clicking on a desired region on the choropleth map projects a pop-up of the cases reported, fatality count.
- **Searching and Filtering:** Users can benefit from the search feature by searching with specific keyword such as name of a region and utilize filtering by reducing the complexity of visualization of all COVID-19 infection data by setting constraints (checkboxes to show only fatality count or cases reported) which cannot be expressed using keywords and provide simple organized data showing only the relevant information required for the users.
- **Connection:**Users can much better understand the progression of COVID-19 situation in specific region, if linking one part of the visualization to another. In our case, we can select a specific region on the choropleth map and it should also display equivalent line graph showing only the COVID-19 infection data such as number of cases reported vs. fatality count and elaborate more details on the selected region

....