Engineering in Computer Science

Project Visual Analytics

Marello Piccinini 1759133

SAPIENZA
UNIVERSITÀ DI ROMA

COVID-19 Italian Dashboard

Introduction

- COVID-19 was first identified on December 31, 2019, by the Chinese authorities in Wuhan, causing more than 500.000.000 cases and 6.000.000 deaths in just over two years.

- The collection and the analysis of the data has been from the beginning of the pandemic the main weapon of any government to counter the advance of the virus as every decision made was based on epidemiological data that was analyzed by experts in the field.

Introduction

- The goal of this project is to provide a simple and efficent dashboard to visualize informtion about the virus during the last two years.

- This dashboard could be helpful to a user to discover differences or analogies in different regions of Italy but could be used also to understand if the decision made by the government during the pandemic have been effective or not, knowing the mistakes of the past it is possible to be prepared for future similar situations.

Data

- The project is base on two different datasets:

1. A json file to draw the map of Italy divided by regions.

2. A csv file containing all the data collected from 2020-02-24 provided by the Ministry of Health and processed and managed by the Department of Civil Protection.
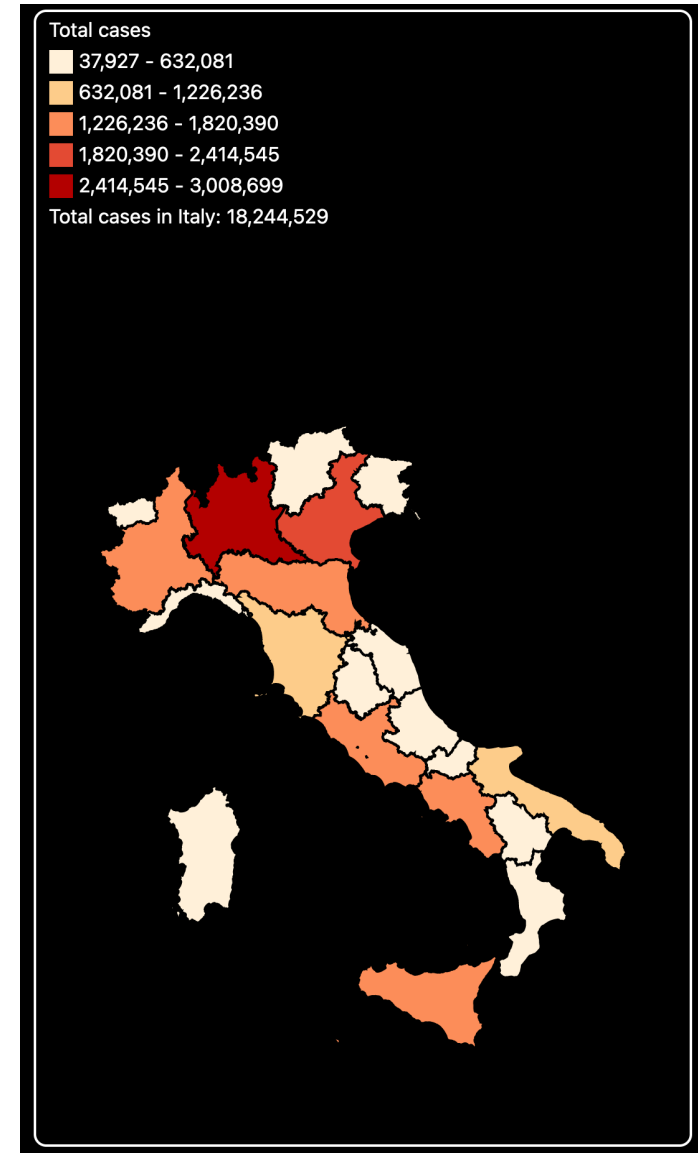
   The AS index is 17.080 x 11 = 18.7880

Preprocessing

- The preprocessing of the dataset about COVID-19 consists of three steps:

1. Choose a subset of the attributes, in particular [*date , region, hospitalized_with_symptoms, intensive_care, total_hospitalized, home_isolation, total_positives, discharged_healed, total_cases, swabs]*

2. Merge the information about *Trento* and *Bolzano* in a unique region *Trentino-Alto Adige.*

3. Transform each column from cumulative value to increment value. The dashboard does not show the total number of cases but the new cases in the specific period)

Visualization

1. Choropleth Map

- The map of Italy is divided by regions.
- The variable visualized is the *total_cases.*
- Regions with different values have different colors, the color scale used was made by *ColorBrewer 2.0 .*
- A legend explain the interval value for each color.

Total cases
- 37,927 – 632,081
- 632,081 – 1,226,236
- 1,226,236 – 1,820,390
- 1,820,390 – 2,414,545
- 2,414,545 – 3,008,699

Total cases in Italy: 18,244,529

Visualization

2. Parallel coordinates plot

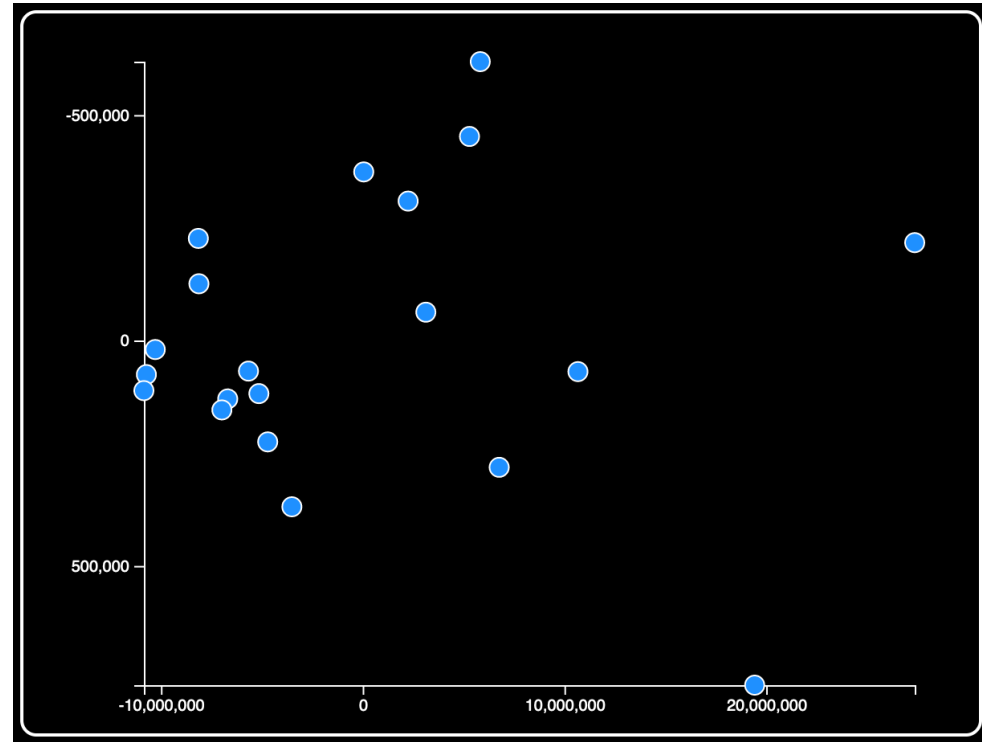- Used to visualize the value of all the regions on all the nine metrics.

Visualization

3. Scatterplot

- Used to visualize the result of the Multidimensional Scaling operation.

- The reduction performed is from R^9 to R^2.

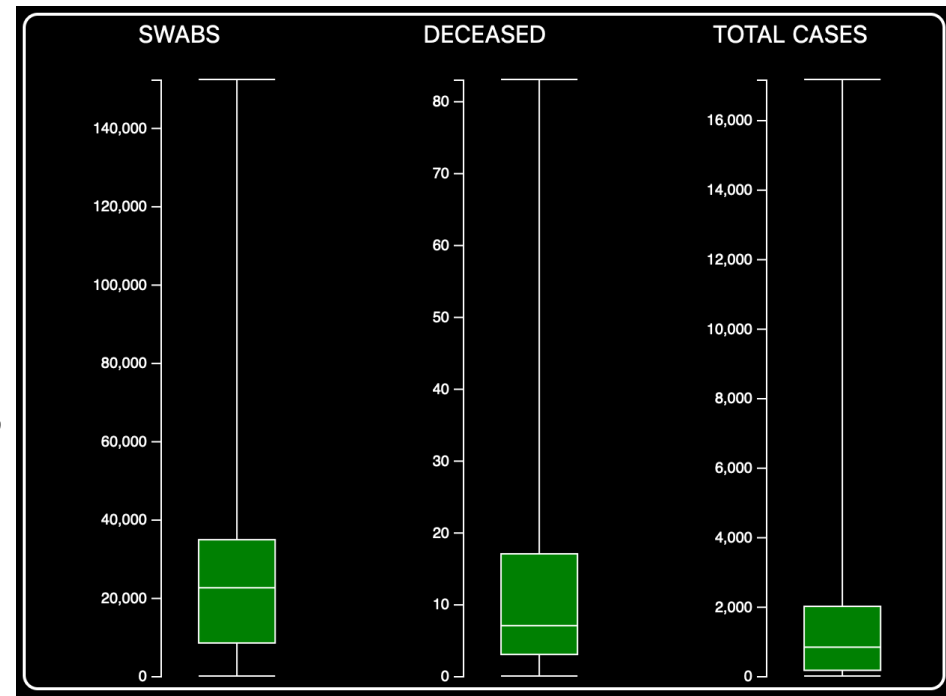- The dissimilarity computation is done using the Euclidean distance.

Visualization

4. Boxplot

- Used for the analytic part.

- Metrics selected: *swabs, deceased, total_cases.*

- Shows *minimum, first quartile, median, third quartile, maximum* of each metric in a single day (no cumulative value).

- Computation triggered by the user by range slider and region selection.

Visualization Global view
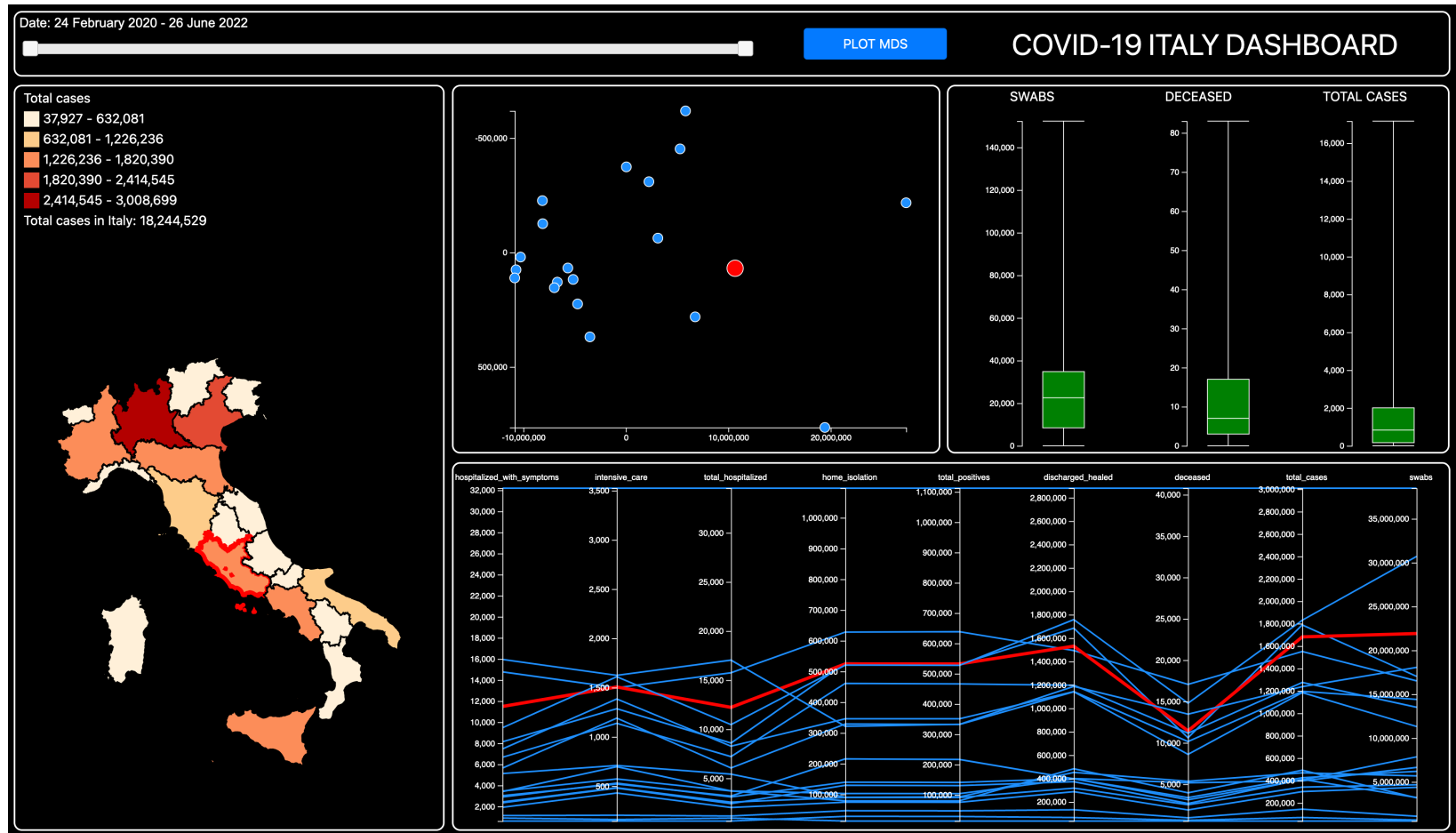
Conclusion

- The dashboard provides a complete and clear view of the Italian situation from the pandemic to now.

- It is able to show various information, not identifiable by other solutions through effective and easy to interpret graphs.

- The project can also be used for different purposes, adapting the dataset to be used and analyzed.

- For future work it is possible to integrate the dashboard with an additional metric as vaccines.

Engineering in Computer Science

Project Visual Analytics

Marello Piccinini 1759133

Thank you for the attention

SAPIENZA
UNIVERSITÀ DI ROMA