Research of COVID-19 impact and trend in USA using statistical approach

Artsiom Nasledniкau (an62556n@pace.edu, art.nasledniкau@gmail.com)
Computer Science Department, Pace University, New York, NY, USA

Research tends to explore a public dataset "United States COVID-19 Cases and Deaths by State over Time" from Centers for Disease Control and Prevention to visualize impact and trend in the USA on population. The research used computer science methods such as aggregation, pattern recognition, feature classification, data regression, clustering and attribute selection.

A difference-by-day algorithm was developed and applied since raw data from the dataset had linear forms. A new parameter “recovery” was introduced to cross-validate trend results which was also used as the metric to compare results from linear and isotonic models. Both models showed similar results with a little deviation.

As part of the research, data from the dataset have been grouped by US state and each US state has been classified by “mean” feature. The classified result showed average and inflated groups. Also, each US state had been plotted to map with a metric density color. Besides, a k-means cluster was generated for US states.

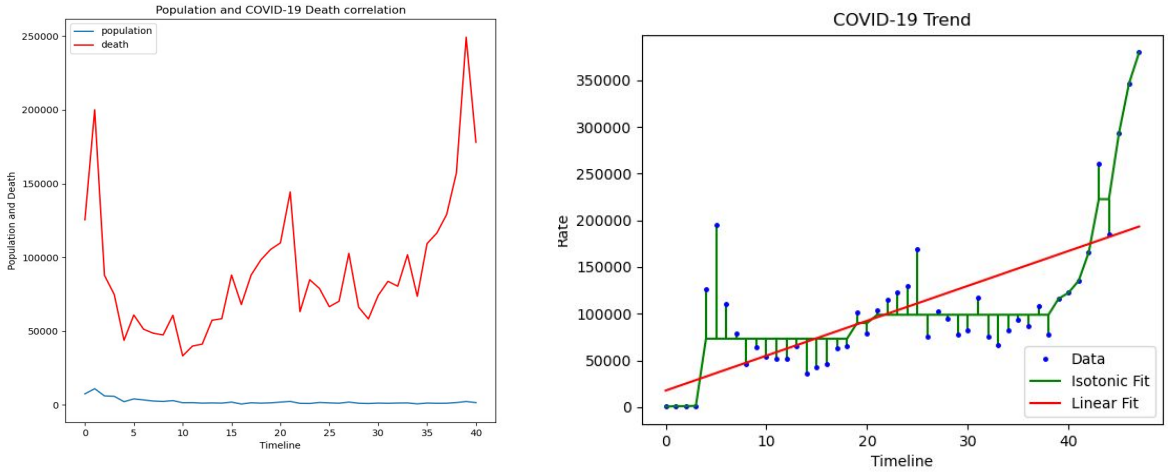

Figure 1a: COVID-19 impact of USA population; Figure 1b: COVID-19 case trend

[1] CDC - United States COVID-19 Cases and Deaths by State over Time, 2020 - https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36/.

[2] Census - U.S. and World Population Clock, 2020 - https://www.census.gov/popclock/.

[3] COVID-19 Pandemic Prediction using Time SeriesForecasting Models, 2020 - https://arxiv.org/pdf/2009.12176.pdf.