**COVID-19 GLOBAL REPORT PROJECT PROPOSAL**

Group Members

Ege Öngül 29037

Elif Ilgın Arat 28818

Zeynep Kurtuluş 29045

Sezin Tekin 28884

1. **Motivation**

In this project, we are going to work on a dataset that consists of the cumulative count of confirmed death and recovered cases of COVID-19 from different countries. We have decided to conduct this project because in the past few years, COVID-19 had significantly affected the agenda of the entire world. With this project, we are not only, aiming to analyze the data both visually and statistically but also investigate a pattern between COVID-19 and geographical region to be able to make future prediction.

1. **Significance of the project**

The project that is going to be conducted will have a significant importance in the sense that the results of our analyses will provide information about mainly how the global pandemic has affected and continues to affect the globe. Additionally, we will obtain the comparisons about how gender, climate, age, and difference in geographical regions influence course of COVID-19 cases in terms of death, recovery and also how the amount of confirmed COVID-19 cases are affected by these factors. This project will be influential and useful in the sense that, the results of this research may give rise to the discovery of unknown underlaying factors behind COVID-19 cases. Human race will benefit for from the discovery of these underlaying factors for sure, since once detected the effects and contagion can be minimized.

1. **Description of the Project**

To begin with, we will analyze the relation between age and confirmed death cases and after we will use tools in programming languages such as pandas and matplotlib to visualize and interpret data. These visualizations will show that, whether there exists a correlation between age and death ratios, and if so, the rate of the correlation. Moving on, we will examine how the spread of the virus is affected with respect to climate. So, we will look at the ratios of cases in colder and warmer regions. After that we are planning to investigate biological factors such as gender, to decide whether different genders have more resistance to the virus in terms of immunity. Lastly, different geographical regions will be examined to detect if there is a connection between COVID-19 cases and development levels of countries.

1. **Expected Output**

Clearly, when the data visualizations conducted from the given data set are examined, we will reach to a general idea about how different parameters affect the number of confirmed COVID-19 cases and death amounts. Expected results of this project are, age and number of deaths will be positively correlated, climate and death amounts will not provide a significant pattern, so outliers are expected in this category. Finally, men are expected to be less immune to virus compared to women.

1. **Plan Table**

|  |  |
| --- | --- |
| **Weeks** | **Explanation** |
| 5 | Project Group Formation |
| 6 | Project Proposal |
| 7-8 | Gathering Information & Analyzing Data |
| 9 | Progress Report |
| 10-11 | Data Visualization |
| 12 | Analyzing Visualized Data |
| 13 | Interpreting the Data & Reaching to Conclusions |
| 14 | Final Report & Presentation |