



Shastra Biogen 2023

Computational Epidemiology of SARS-Cov-2 (COVID-19)

Problem Statement 2:

BACKGROUND

Covid 19 struck but unexpectedly lasted for more than two years. It had a significant impact on the social and economic worlds. The spread of **COVID-19** threatens public health and dramatically affects the global economy. Labor displacement, business closures, and stock crashes are just some of the impacts of this global lockdown during the pandemic. Vaccines worked, but the production was short so they couldn't be distributed. However, proper prediction of the risk of infections based on various factors helped to predict the verdict population. From the **given data**, considering the possible factors, you must make a **model prediction** regarding the topic discussed. Note You, as a participant, can't add your own characteristics data sets using the web or any other sources. Report your readings and outcomes in the form of percentages and their distribution. Arrive /Mention a proper conclusion based on your tasks. You can use software, e.g., Excel, or even go manually. The more accuracy better, the more points. This data analysis is constructive even in the future, for accumulating the necessary resources smartly. The following data set we provide :

PROBLEM

Part 1

For our case, we need to apply the regression analysis to the attached dataset below(in txt/CSV format). Prepare a regression model based on the '**DataSet_Test**' File. This is common to all participants. Further for the other file named '**DataSet_Participant**'. This is unique to you and contains similar data only without the actual COVID test result. Using the analytical model prepared, you have to predict the result for those patients and complete the columns

Part 2

Similarly prepare another analytical model (on the same Dataset, check Columns AH to AO {consider all of them} based on Covid Symptoms). Thus this will be our non-clinical way to

predict Covid infection. Using the model prepared, you have to predict the result for those patients in the other set and thus fill in that column.

You can use any method (any language, or software) to analyze the data and perform the necessary operation. The final results should be in the tabular format submitted through a Google sheet file (two separate files for part 1 and part 2). The method used should also be presented so that we can run the model as well as check its authenticity.

DataSet_Test link

https://drive.google.com/file/d/123VNKt9HkCkP3BOyqzBNQdl1jKoKSX5G/view?usp=share_link

The individual dataset will be mailed separately.