

ASSIGNMENT 1

REPORT ON COVID-19 DATA ANALYSIS

April 1, 2021

1 Analysis and conclusion

This report suggests that we have a district,state wise analysis for cases,z-scores,standard deviations,means for every district according to its neighbors and state.We had to undergo a lot of data preprocessing in order to match the district names as Covid-19 portal's districts .After going through the data,we had to add some districts manually as per their COVID-19 portal district names.Some samller districts are merged into a larger one,plus the state which have unknown entries are combined and named as unknown with 'state name'Furthermore,doing careful analysis of data,we present the following conclusions.

From the analysis of the COVID data in the span of March 15,2020 to September 5,2020, we can conclude that the hike in the number of cases is typically observed from the week 7.Prominently in the week 1-6, i.e. from March 15,2020 to April 26,2020 there are only cases in Delhi,Chandigarh and Telangana.

The top 5 hot-spot districts according to the neighborhood z-score are Bangalore(urban),Assam,Bhopal,Surat,Delhi.Also the top 5 cold-spot according to neighborhood z-score are Upper Dibang,Kairakal,Unokoti(Tripura),Mahe,Sheohar.

The top 5 hot-spot districts according to the state z-score are Puducherry,Bangalore(urban),Chennai,Raipur,the top 5 cold-spot according to state z-score are Diu,Krishna(Andra Pradesh),Lahaul and Spiti,Vizianagaram,Wayanad.