Project Proposal: Classification of Cities and Regions of Saudi Arabia in 2021 in the spread of COVID-19 between regions with high, and low.

Introduction

Because of the widespread distribution of Covid-19 in Saudi Arabia, the degrees of dissemination differ by city or area. It is difficult for the Ministry of Health or people to measure the amount of spread on a daily level. Especially as it relates to scientific or medical studies. I aim at planning a system that would divide Saudi Arabia's regions into high and low spread areas.

My project aims to classify Saudi Arabia's regions as high, or low in terms of the spread of COVID-19 based on the average number of cases per city in each region from January 1st to October 6th, 2021. It is worth noting that, a dataset that I will use from "the King Abdullah Center for Petroleum Studies and Research[1]" website, which included around 13 regions and 102 cities.

Project Question

1- What are the cities most likely to spread to COVID-19?

2- Does urbanization contribute to the spread of COVID-19, as in busy-main cities such as Al Madinah Al Munawwarah city?

3- Do working days or holidays contribute to the spread of COVID-19?

4- Does this study serve doctors or tourists?

5- Do unbusy cities such as Al-Bahah city contribute to the less spread of COVID-19?

Data Description

I will use a dataset from "the King Abdullah Center for Petroleum Studies and Research "website which included around 13 regions and 102 cities. A dataset contains **225.572k** rows and **11** columns.

My contribution to this project is to give a reliable and useful analysis on the classification of level spread cases of COVID-19 around regions and cities between high, and low.

My study has a significant benefit for doctors and tourists in that it contributes to help the Saudi Ministry of Health to take procedures regarding areas or cities where the virus spread is very high. Also, If a tourist plans to travel to one of the regions? It helps him/her make a decision about it.

The sample about COVID-19 patients who live in Saudi Arabia. Also, I will predict as a target (High-low) in spread COVID-19 around Saudi Arabia cities and regions.

Features
1. **DC:** Daily or Cumulative rate of COVID-19 cases.
2. **Indicator:** The status indicator is active, recovered, cases, mortalities in a particular state.
3. **Date:** Date of case.
4. **Day:** Day of case.
5. **City:** A city of COVID-19 case.

6. **Region:** A region of COVID-19 case.
7. **CasesNumber:** Numbers of COVID-19 cases.
8. **AvarageCases:** Average of numbers of COVID-19 cases.
9. **WorkHoliday:** Status on a working day or holiday?
10. **Country:** Saudi Arabia country.

Labels:
11. **SpreedCovid19:** COVID-19 spread case, Is high, or low?

Tool:

I will use python libraries such as Numpy, Pandas, matplotlib, seaborn, sklearn, classification_report, and confusion_matrix.

Furthermore, I will use the python platform is the Jupiter Notebook. In addition, I will apply an MLPClassifier besides; I want to evaluation my model by utilizing some metrics such as precision, recall, and F1-score,..etc.

Dataset Link Source:

In my Github account uploaded CSV file for datasets.

[1]	https://datasource.kapsarc.org/explore/embed/dataset/saudi-arabia-coronavirus-disease-covid-19-situation/table/?disjunctive.daily_cumulative&disjunctive.indicator&disjunctive.event&disjunctive.city_en&disjunctive.region_en&sort=date&dataChart=eyJxdWVyaWVzIjpbeyJjaGFydHMiOlt7InR5cGUiOiJsaW5lIiwiZnVuYyI6IkFWRyIsInlBeGlzIjoiY2FzZV92YWx1ZSIsInNjaWVudGlmaWNEaXNwbGF5Ijp0cnVlLCJjb2xvciI6InJhbmdlLUFjY2VudCJ9XSwieEF4aXMiOiJkYXRlIiwibWF4cG9pbnRzIjoiIiwidGltZXNjYWxlIjoiZGF5Iiwic29ydCI6IiIsInNlcmllc0JyZWFrZG93biI6ImluZGljYXRvciIsImNvbmZpZyI6eyJkYXRhc2V0Ijoic2F1ZGktYXJhYmlhLWNvcm9uYXZpcnVzLWRpc2Vhc2UtY292aWQtMTktc2l0dWF0aW9uIiwib3B0aW9ucyI6eyJkaXNqdW5jdGl2ZS5kYWlseV9jdW11bGF0aXZlIjp0cnVlLCJkaXNqdW5jdGl2ZS5pbmRpY2F0b3IiOnRydWUsImRpc2p1bmN0aXZlLmV2ZW50Ijp0cnVlLCJkaXNqdW5jdGl2ZS5jaXR5X2VuIjp0cnVlLCJkaXNqdW5jdGl2ZS5yZWdpb25fZW4iOnRydWV9fX1dLCJkaXNwbGF5TGVnZW5kIjp0cnVlLCJhbGlnbk1vbnRoIjp0cnVlLCJzaW5nbGVBeGlzIjp0cnVlfQ%3D%3D