Battle of Neighbourhood Week-2

A few Identified factors that influence our decision are:Covid-19 cases per district “Riwayat File Covid-19 DKI Jakarta”Total population in DKI Jakarta 2020 statistik.jakarta.go.id10 most population in DKI Jakarta 2020 per district statistik.jakarta.go.idHospital for treatment covid-19 megapolitan.kompas.comThe following data sources are needed to extract/generate the required information:Processed covid-19 positive case data collection on 28 May 2020 At 09.00.The distribution of mask sales based on the population in the DKI Jakarta area.The distribution of mask sales based on 5 districts with the most densely populated populations.New datasets (to be created) from Hospital table that contains city, district, along with their latitudes and longitudes.

Read in the data Covid-19 cases per district (28 May,2020)df_cases = pd.read_csv("https://raw.githubusercontent.com/cahyati/Coursera_Capstone/master/Standar%20Kelurahan%20Data%20Corona%20(28%20MEI%202020%20Pukul%2009.00).csv")# View the top rows of the datasetdf_cases

df_cases.head() import pandas as pd# Read in the data total population in DKI Jakarta 2020df_population = pd.read_csv("https://raw.githubusercontent.com/cahyati/Coursera_Capstone/master/population2020_DKI_Jakarta.csv")# View the top rows of the datasetdf_population

df_cases.tail() Total population in Jakarta.df_population.info()# Get the number of total / confirmed POSITIVE cases in Jakarta per 28 May 2020print ("Total Polulation :", df_population['Total population 2020(people/km²)'].sum())<class 'pandas.core.frame.DataFrame'>RangeIndex: 6 entries, 0 to 5Data columns (total 2 columns): # Column Non-Null Count Dtype --- ------ -------------- ----- 0 City 6 non-null object 1 Total population 2020(people/km²) 6 non-null int64 dtypes: int64(1), object(1)memory usage: 224.0+ bytesTotal Polulation : 92736Read and show the top 5 data rows from total population in DKI Jakarta, 2020.df_population.head()Image for postRead and show the data from 10 districts most pupulated in DKI Jakarta, 2020.# Read in the data 10 most pupulation in DKI Jakarta 2020 per districtdf_most_population = pd.read_csv("https://raw.githubusercontent.com/cahyati/Coursera_Capstone/master/10_kelurahan%20terpadat_DKI_Jakarta.csv")# View the top rows of the datasetdf_most_populationImage for postRead and show the top 5 data rows from 10 most populated areas in DKI Jakarta, 2020 per district.df_most_population.head()Image for postAccording to the information update from Kompas.com (megapolitan.kompas.com), the following hospitals are the existing reference hospitals for Covid-19 testing in Jakarta area:RSPI Sulianti Saroso, Jakarta UtaraRSUP Persahabatan, Jakarta TimurRSPAD Gatot Soebroto, Jakarta PusatRSUP Fatmawati, Jakarta SelatanRSU Bhayangkara, Jakarta TimurRSAL Mintohardjo, Jakarta PusatRSUD Cengkareng, Jakarta BaratRSUD Pasar Minggu, Jakarta SelatanRSKD Duren Sawit, Jakarta TimurRS Pelni, Jakarta BaratRSUD Tarakan, Jakarta PusatRSUD Koja, Jakarta UtaraRSU Pertamina Jaya, Jakarta PusatConstruct a Pandas data frame for subsequent data analysis.Read and show Hospital data that provide treatment Covid-19.# Read in the data Hospital for treatment covid-19df_hospital = pd.read_csv("https://raw.githubusercontent.com/cahyati/Coursera_Capstone/master/Hospital%20for%20treatment%20covid-19.csv")# View the top rows of the datasetdf_hospitalImage for postRead and show the top 5 data rows from Hospital data providing treatment Covid-19.df_hospital.head()Image for postThis sums up our data mining and data exploration section. In the following METHODOLOGY section, we will describe the process of how to do a ‘Visual’ approach to better understand our data using data science and data analytics tool kits.