Data

**Read and show all data used.**

Read and show data Covid-19 cases per district.

*# 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 dataset*  
df\_cases



Read and show the top 5 data rows from Covid-19 cases per district.

df\_cases.head()



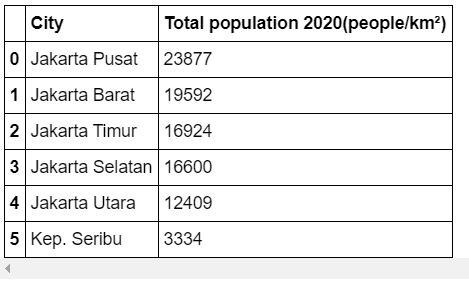
Read and show the bottom 5 data rows from Covid-19 cases per district

df\_cases.tail()



Read and show the total population data in DKI Jakarta 2020.

**import** **pandas** **as** **pd**  
*# Read in the data total population in DKI Jakarta 2020*  
df\_population = pd.read\_csv("https://raw.githubusercontent.com/cahyati/Coursera\_Capstone/master/population2020\_DKI\_Jakarta.csv")  
*# View the top rows of the dataset*  
df\_population

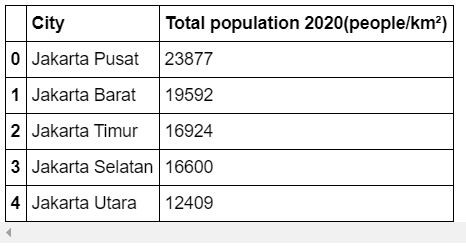


Total population in Jakarta.

df\_population.info()  
  
*# Get the number of total / confirmed POSITIVE cases in Jakarta per 28 May 2020*  
print ("Total Polulation :", df\_population['Total population 2020(people/km²)'].sum())<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 6 entries, 0 to 5  
Data 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+ bytes  
Total Polulation : 92736

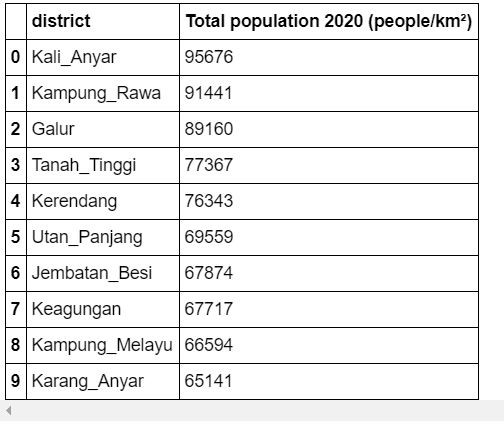
Read and show the top 5 data rows from total population in DKI Jakarta, 2020.

df\_population.head()



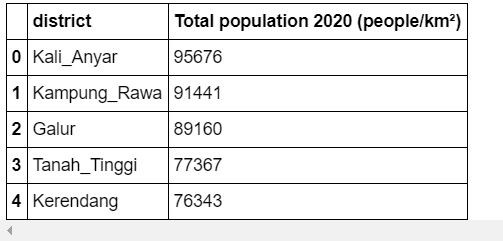
Read and show the data from 10 districts most pupulated in DKI Jakarta, 2020.

*# Read in the data 10 most pupulation in DKI Jakarta 2020 per district*  
df\_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 dataset*  
df\_most\_population



Read and show the top 5 data rows from 10 most populated areas in DKI Jakarta, 2020 per district.

df\_most\_population.head()



According 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:

1. RSPI Sulianti Saroso, Jakarta Utara
2. RSUP Persahabatan, Jakarta Timur
3. RSPAD Gatot Soebroto, Jakarta Pusat
4. RSUP Fatmawati, Jakarta Selatan
5. RSU Bhayangkara, Jakarta Timur
6. RSAL Mintohardjo, Jakarta Pusat
7. RSUD Cengkareng, Jakarta Barat
8. RSUD Pasar Minggu, Jakarta Selatan
9. RSKD Duren Sawit, Jakarta Timur
10. RS Pelni, Jakarta Barat
11. RSUD Tarakan, Jakarta Pusat
12. RSUD Koja, Jakarta Utara
13. RSU Pertamina Jaya, Jakarta Pusat

Construct 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-19*  
df\_hospital = pd.read\_csv("https://raw.githubusercontent.com/cahyati/Coursera\_Capstone/master/Hospital**%20f**or%20treatment**%20c**ovid-19.csv")  
*# View the top rows of the dataset*  
df\_hospital



Read and show the top 5 data rows from Hospital data providing treatment Covid-19.

df\_hospital.head()



This 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.