COVID New Case Data

Source: <https://github.com/owid/covid-19-data/blob/master/public/data/owid-covid-data.csv>

Code we use to clean the data:

df = pd.read\_csv('https://raw.githubusercontent.com/owid/covid-19-data/master/public/data/owid-covid-data.csv')

# to clean the data into only containing the United States COVID data

us\_df = df.query("location == 'United States'")

# to check whether there are any missing data in the new\_cases column

us\_df[us\_df["new\_cases"].isna()]

## the result of this shows that there are no missing data in the new\_cases column until 2023-03-26 which is well beyond our scope of analysis

## so the missing data in the new\_cases column is not a problem

## we can use this data to do our analysis

Economic factors

Selected these base on:  
<https://www.investopedia.com/articles/personal-finance/020215/top-ten-us-economic-indicators.asp>

GDP

Source: <https://fred.stlouisfed.org/series/GDP>

Unemployment rate

Source: <https://fred.stlouisfed.org/series/UNRATE>

Industrial Production rate

Source: <https://fred.stlouisfed.org/series/INDPRO>

Fed fund rate

Source: <https://fred.stlouisfed.org/series/FEDFUNDS>

Inflation rate

Source: <https://fred.stlouisfed.org/series/CPIAUCSL>

Total Construction Spending (VPN Required)

Source: <https://www.census.gov/construction/c30/prexcel.html&c_year=2021>

Code we use to clean the data:

result\_df = pd.DataFrame(columns=["month", "Total\_Construction\_Value"])

for year in [20, 21]:

for month in range(1, 13):

link = "https://www.census.gov/construction/c30/xls/pr20{}{}.xlsx".format(str(year), str(month).zfill(2))

print(link)

temp\_df = pd.read\_excel(link)

total\_construction\_val = temp\_df[temp\_df["Table 1. Value of Construction Put in Place in the United States, Seasonally Adjusted Annual Rate"] == "Total Construction"].iloc[0,1]

result\_df = pd.concat([result\_df, pd.DataFrame({"month":["20"+ str(year) + str(month).zfill(2)], "Total\_Construction\_Value":[total\_construction\_val]})]).reset\_index(drop=True)

result\_df

Manufacturing Total

Source: <https://fred.stlouisfed.org/series/AMTMNO>

Total Home Sales (VPN Required)

Source: N/A