# L01_Exploring_Data

September 19, 2025

[20]:
```
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as mdates
```

[11]:
```
df_lab01_clean_data = pd.read_csv('lab01_clean_data.csv', sep = ',')
df_lab01_clean_data.head()
```

[11]:

| | Unnamed: 0 | date | time | abbreviation_canton_and_fl | ncumul_tested | ncumul_conf | new_hosp | current_hosp | current_icu | current_vent | ncumul_released | ncumul_deceased | source | current_isolated | current_quarantined | current_quarantined_riskareatravel |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 2020-02-01 | 00:00 | GE | 109.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | infocovid.smc.unige.ch | 1.0 | 0.0 | 0.0 |
| 1 | 1 | 2020-02-02 | 00:00 | GE | 115.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | infocovid.smc.unige.ch | 1.0 | 0.0 | 0.0 |
| 2 | 2 | 2020-02-03 | 00:00 | GE | 136.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | infocovid.smc.unige.ch | 1.0 | 0.0 | 0.0 |
| 3 | 3 | 2020-02-04 | 00:00 | GE | 156.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | infocovid.smc.unige.ch | 1.0 | 0.0 | 0.0 |
| 4 | 4 | 2020-02-05 | 00:00 | GE | 169.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | infocovid.smc.unige.ch | 1.0 | 0.0 | 0.0 |

## 1 What was your initial question or idea?

My idea was to make a plot to show how many people are currently quarantined over time. The question then would be: How many people are quarantined at specific points in time, and how does this number change over time?

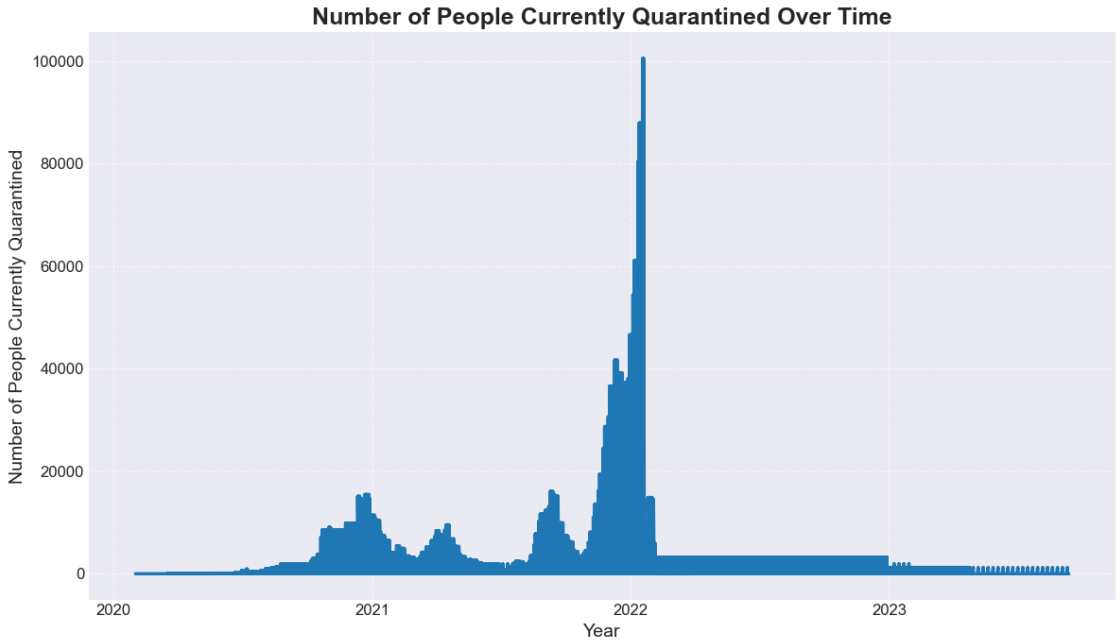
[24]:
```
# Line-Graph showing how many people are currently (current_quarantined)␣
 ↪quarantined over time

df_lab01_clean_data['date'] = pd.to_datetime(df_lab01_clean_data['date'])

plt.style.use('seaborn-v0_8-darkgrid')
plt.figure(figsize=(12,7))
plt.plot(
    df_lab01_clean_data['date'],
    df_lab01_clean_data['current_quarantined'],
    color='#1f77b4',
    linewidth=2.5
)
plt.title('Number of People Currently Quarantined Over Time', fontsize=18,␣
 ↪fontweight='bold')
plt.xlabel('Year', fontsize=14)
plt.ylabel('Number of People Currently Quarantined', fontsize=14)

ax = plt.gca()
ax.xaxis.set_major_locator(mdates.YearLocator())
ax.xaxis.set_major_formatter(mdates.DateFormatter('%Y'))

plt.xticks(fontsize=12)
plt.yticks(fontsize=12)
plt.grid(True, which='major', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```




## 2 How did you proceed to arrive at an answer?

I'm creating a line graph to show how many people are currently quarantined over time. I'm using the 'seaborn-v0_8-darkgrid' style, setting the date column to datetime, and then plotting 'current_quarantined' against 'date'. I've added a title, axis labels, and formatted the x-axis to show years. Finally, I'm displaying the plot.

## 3 What are your results?

The result is a plot that shows the trend of how many people are currently quarantined over time. The graph indicates fluctuations in the number of quarantined individuals, with noticeable peaks and troughs. It is noteworthy that the data from 2023 onwards was no longer maintained or collected differently, as the graph shows gaps.