Helen Liang, Ivy Zhao
Spring 2024, GPH-GU 2022 SAS for Beginners: Data Management and Exploration
School of Global Public Health, New York University, New York, NY, USA

Analyzing Minnesota and Midwest Region COVID-19 Data Trends, June 1-7, 2020

The Coronavirus Disease 2019 (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus that caused a global pandemic in December 2019 and lasted for several years. This analysis focuses on COVID-19 data in the state of Minnesota and consists of datasets on the daily COVID-19 information for the study period from June 1 to June 7, 2020 ("cd0601" – "cd0607") and the day prior to the start of the study period ("cd0531"), both provided by the Center for Systems Science and Engineering at Johns Hopkins University, along with a dataset provided by the U.S. Census indicating region for each U.S. state/territory. Our objective is to conduct a comparative analysis to understand COVID-19 trends, compare outcomes with regional averages, pinpoint disparities, identify potential areas for intervention or support, and make recommendations for policymakers, healthcare providers, and the public to make informed decisions and allocate resources appropriately.

Table 1 presented comparison of COVID-19 outcomes between Minnesota and the Midwest regional average on June 1, 2020. Minnesota reported 255,592 tests (total number of people who have been tested for COVID-19), 25,208 confirmed cases, 3,086 hospitalizations, and 1,060 deaths due to COVID-19. Compared to the Midwest region with 279,553.75 average tests, 29,889.83 average confirmed cases, 2,729.57 average hospitalizations, and 1,564.50 average deaths, Minnesota had a lower number of tests and confirmed cases but a higher number of hospitalizations. Deaths in Minnesota were also significantly lower than the regional average.

Table 1. Comparison of COVID-19 Outcomes Between Minnesota and the Midwest Regional Average on June 1, 2020.

Outcome	Minnesota	Regional Average (Midwest)
Tests	255,592	279,553.75
Confirmed Cases	25,208	29,889.83
Hospitalizations	3,086	2,729.57
Deaths	1,060	1,564.50

We chose to analyze daily trends of new confirmed COVID-19 cases because they directly reflected infection trends over time. Monitoring new confirmed cases is crucial for understanding the virus's transmission dynamics and assists public health administrators in executing timely interventions. Table 2 displayed the number of total and new confirmed COVID-19 cases in Minnesota from June 1 to June 7, 2020, with Figure 1 illustrating daily trends of new confirmed cases. The data highlights daily changes in newly confirmed cases, with a gradual and steady increase from June 2nd to June 5th, then a steep increase with the peak on June 7th (707 new confirmed cases) before the rapid decline on June 8th with 385 new confirmed cases (not shown in Table 2). Despite the decline, the trend on new confirmed COVID-19 cases remains relatively high.

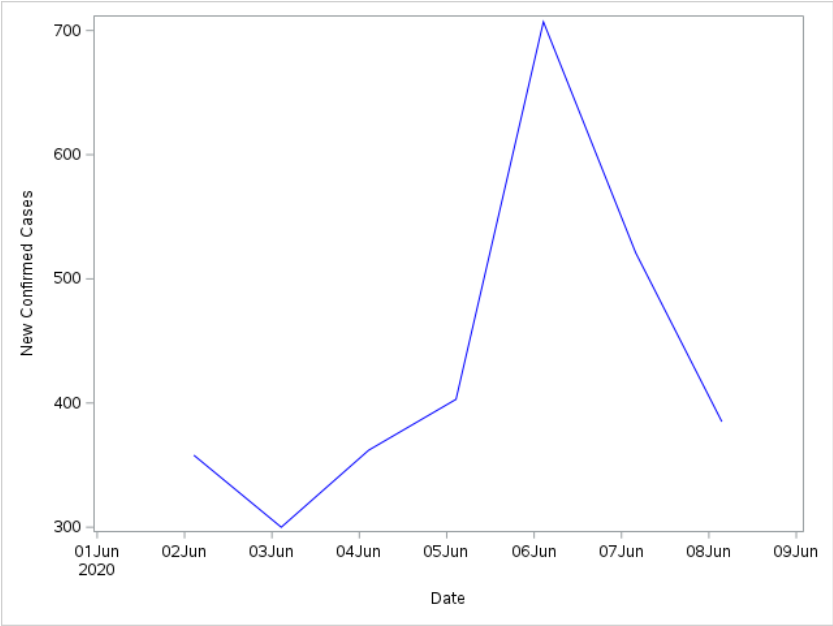
Table 2. Trend of Total and New Confirmed COVID-19 Cases in Minnesota from June 1 to June 7, 2020.

Date	Total Confirmed Cases	New Confirmed Cases
6/1/20	24,850	-
6/2/20	25,208	358
6/3/20	25,508	300
6/4/20	25,870	362
6/5/20	26,273	403
6/6/20	27,501	707
6/7/20	27,886	521

Based on the analysis of the COVID-19 data from Minnesota compared to the regional Midwest averages and the observed trends in new confirmed cases, we would recommend increasing testing capacity such as expanding testing centers in Minnesota to help in early detection and isolation of cases to reduce the spread of the virus since both number of tests and confirmed cases at the state-level are lower compared to regional averages. In addition, we should provide targeted public health interventions directed towards high-risk groups such as

the elderly or those with pre-existing conditions as well as immediately and appropriately allocate hospital resources (e.g., beds, medical staff, etc.) for high number of cases and high-risk areas in Minnesota. Furthermore, corresponding policies should be adaptable based on the latest data, including readiness to implement or ease restrictions based on the number of cases, hospitalization rates, etc. along with developing a long-term plan which should include guidelines for potential future outbreaks given the unpredictability of COVID-19.

Figure 1. Line Plot Illustrating Daily Trends in New Confirmed COVID-19 Cases in Minnesota from June 1 to June 7, 2020.



Appendix

```
/*****************************************
PURPOSE: Final Project: COVID-19 Data Analysis
DATA SETS:
    COVID-19 data comes from the Center for Systems Science and Engineering at Johns Hopkins University;
    Region data comes from the U.S. Census;
PROGRAMMER: Helen Liang, Ivy Zhao
DATE UPDATED: 5/7/2020
 *****************************************/


libname mylib "/home/u63668564/GPH-GU-2022 SAS Bootcamp/Final Project";

/* 1. Provide basic information about the data source and the purpose of this data analysis. */

/* This analysis focuses on COVID-19 data in the state of Minnesota during the study period from June 1 to June 7,
Our goal is to understand the COVID-19 trends, compare outcomes with regional averages, and make recommendations to

/* 2a. What was the total number of COVID-19 outcomes of
(a) tests,
(b) confirmed cases,
(c) hospitalizations, and
(d) deaths
recorded in Minnesota at the beginning of the study period (on June 1)? */

data mylib.cd0601;
    set mylib.cd0601;
run;

proc means data=mylib.cd0601 sum;
    where Province_State = "Minnesota";
run;

/* tests = 255,592
   confirmed cases = 25,208
   hospitalizations = 3,086
   deaths = 1,060 */

/* 2b. Now, merge COVID-19 dataset with Census region data. */

data mylib.region;
    set mylib.region;
run;

data mylib.cd0601_merged;
    merge mylib.cd0601
          mylib.region;
run;

/* 2c. What was the average number of COVID-19 outcomes of
(a) tests,
(b) confirmed cases,
(c) hospitalizations, and
(d) deaths
recorded in your state's region at the beginning of the study period (on June 1)? */

proc means data=mylib.cd0601_merged mean;
    where Region = "Midwest";
run;

/* tests = 279,553.75
   confirmed cases = 29,889.83
   hospitalizations = 2,729.57
   deaths = 1,564.50  */

/* 3. Next, merge all COVID-19 datasets provided into one combined dataset. */

data mylib.all_covid;
    set mylib.cd0531
        mylib.cd0601
        mylib.cd0602
        mylib.cd0603
        mylib.cd0604
        mylib.cd0605
        mylib.cd0606
        mylib.cd0607;
run;

/* 4. Describe daily trends for a specific outcome of your choice.
Briefly explain why you chose that one specific outcome to assess daily changes during the study period.
This can be in a graph format but be sure to write a brief description. */
```

```
/* 4a. Compute the number of new COVID-19 outcomes for each day of the study period (June 1-7, 2020)
by subtracting the total number of the same outcome recorded the day before from the total number reported that day
Outcome selected = Confirmed */

/* Sort data by date and filtering for Minnesota */
proc sort data=mylib.all_covid out=mylib.mn_covid;
    by Last_Update;
    where Province_State = "Minnesota";
run;

/* Calculate new daily confirmed cases */
data mylib.mn_new_cases;
    set mylib.mn_covid;
    by Last_Update;
    retain prev_confirmed 0;
    new_confirmed = Confirmed - prev_confirmed;
    prev_confirmed = Confirmed;
    if _n_ = 1 then new_confirmed = .;
    drop prev_confirmed;
run;

/* Print data */
proc print data=mylib.mn_new_cases;
    var Last_Update Confirmed new_confirmed;

/* Plot daily trends in new confirmed cases */
proc sgplot data=mylib.mn_new_cases;
    series x=Last_Update y=new_confirmed / lineattrs=(color=blue);
    xaxis label="Date";
    yaxis label="New Confirmed Cases";
run;

/* For this “new outcome,” you will choose only one outcome from
(a) tests,
(b) confirmed cases,
(c) hospitalizations, or
(d) deaths.

Be sure to explain why you chose that one specific outcome.

How did the total daily number of “new outcome” change over time in your state during the study period (June 1-7, 2
Did these numbers increase, decrease, or pretty much stay the same? */

/* We chose to analyze daily trends of new confirmed COVID-19 cases because they directly reflected infection trend
Monitoring new confirmed cases is crucial for understanding the virus's transmission dynamics and assists public he

Table 2 displayed the number of total and new confirmed COVID-19 cases in Minnesota from June 1 to June 7, 2020, wi
The data highlights daily changes in newly confirmed cases, with a gradual and steady increase from June 2nd to Jun
Despite the decline, the trend on new confirmed COVID-19 cases remains relatively high. */
```