COVID-19 Data Brief

Elizabeth Combs

July 27, 2020

Trends in the United States' COVID-19 Outbreak: Confirmed Cases

At the beginning of 2020, the novel coronavirus COVID-19 spread across the globe, resulting in a pandemic. The virus continued to spread within the United States in June 2020. This data brief describes trends in confirmed cases and hospitalization rates in the United States. The data originates from a study completed in the first week of June from June 1 through June 7, 2020 at the Center for Systems Science and Engineering at Johns Hopkins[1]. The data brief examines the trends in confirmed cases, new confirmed cases per day, and the average hospitalization rate at the national, regional[2], and state or territory levels. Data is available for the 50 U.S. states, six territories, and two cruise ships.

Total Confirmed Cases and New Confirmed Cases in the United States

- Prior to the beginning of the study period on June 1, 2020, the total number of confirmed COVID-19 cases recorded in the U.S. was **1,790,172**.
- The total number of confirmed cases recorded in the U.S. at the end of the study period was **1,943,882**.
- The total increase in confirmed COVID-19 cases during the first week of June, including the first day was **153,710** cases.
- In the U.S., the total daily number of new confirmed cases remained relatively constant throughout the study period, averaging **21,959** with a standard deviation of **3,845**.

Figure 1. New Confirmed Cases in the U.S. by Study Day

Study Day	New Confirmed*
1	20,848
2	20,801
3	19,699
4	21,140
5	29,972
6	23,133
7	18,117

*Daily new confirmed cases were computed using the difference between the previous day and the current day.
**The study period runs from June 1, 2020 through June 7, 2020.

[1]COVID-19 Source: https://github.com/CSSEGISandData
[2] Region Source: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

Hospitalization Rate by Region

- Across all regions in the U.S., the average hospitalization rate declined during the first week of June 2020.
- While the Northeast had the highest average hospitalization rate at **15.4%** of confirmed cases on day one, it declined to **14.2%** by the end of the study period.
- The Midwest had the lowest average hospitalization rate at **11.8%** of confirmed cases on June 1, 2020, which declined to **11.1%** on June 7, 2020.
- The South and West regions had average hospitalizations of **14.9%** and **13.5%** that declined to **14.1%** and **12.8%**, respectively.

Figure 2. Average Hospitalization Rate by U.S. Region

Region*	Average: Day 1	Average: Day 7	Difference
Northeast	15.4%	14.2%	-1.2 pts
South	14.9%	14.1%	-0.8 pts
West	13.5%	12.8%	-0.7 pts
Midwest	11.8%	11.1%	-0.7 pts

*Note: "Other" region which includes U.S. territories and active Cruise Ships are excluded from this analysis due to missing information on hospitalization rate.
** The hospitalization rate represents the total number of people hospitalized per confirmed cases.

Overview of Confirmed Cases at the State-Level

- On June 1, 2020, the state with the highest number of new confirmed cases was Massachusetts with **3,840** cases.
- Meanwhile seven states, territories, and cruise ships had zero new confirmed cases: American Samoa, Diamond Princess, Grand Princess, Idaho, Virgin Islands, Guam, and Rhode Island on June 1, 2020.
- At the end of the study period on Jun 7, 2020, California had the highest number of new confirmed cases with **2,220**.
- Eight locations had zero new cases on the June 7, 2020: American Samoa, Diamond Princess, Grand Princess, Idaho, Virgin Islands, Guam, Rhode Island, and Kentucky.
- Over the entire study period, three locations, American Samoa, Diamond Princess, and Grand Princess, reported no new cases, while

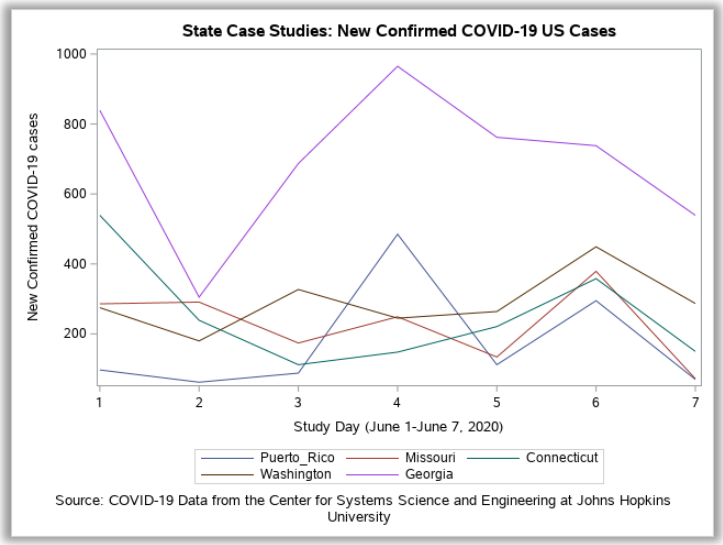
the Virgin Islands reported just **5** new cases in the first week of June.

- California had the highest number of new confirmed cases during the entire study period with **18,664** cases. Texas and Florida also had high numbers of new confirmed cases during the period with **10,756** and **7,775** cases, respectively.

State-Level Case Studies and Comparison

- In this section, one state from each region of the United States is selected to review as case studies.
- States and territories selected by US region are:
  - *Northeast:* Connecticut
  - *South:* Georgia
  - *Midwest:* Missouri
  - *West:* Washington
  - *Other:* Puerto Rico
- The graph on the right indicates that trends are relatively similar across regions, decreasing towards the beginning of the period, increasing in the middle period and spiking on the sixth day followed by another decrease on the seventh day.
- **Georgia** had the highest number of confirmed cases during the entire study period with volatile changes during the first four study days and a decreasing trend during the last three study days, which differs from the other four locations in the final three study days.
- **Connecticut** had a generally decreasing trend over the study period with small increases in cases during days 3-6 with the maximum occurring on the sixth day.
- **Missouri and Washington** followed a similar trend to Connecticut, showing a generally decreasing pattern from days 1 through 5, with a spike occurring during day 6.
- **Puerto Rico** had the lowest number of new confirmed cases during most study days; however, a large increase in new confirmed cases on June 4 brought the territory to the second position during that day.

Figure 3. New Confirmed COVID-19 by Selected Locations



Data Source:

- The data containing the 2019 Novel Coronavirus cases was originally retrieved at https://github.com/CSSEGISandData/COVID-19.
- The repository is operated by the Johns Hopkins University Center for Systems Science and Engineering and supported by ESRI Living Atlas Team and the Johns Hopkins University Applied Physics Lab.
- The codebook is available at https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/README.md.
- The regional source was obtained from https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf, which categorizes U.S. states into four regions.

Authored by:
Elizabeth Combs (eac721)
M.S. in Data Science Candidate

Acknowledgements:
Professor Violet Xu
Professor Emily Goldmann

Code: SAS Code is available on GitHub: https://github.com/lcombs/covid_sas.