Visitor Patterns and Visitation Types of California National Parks

By Cora Johnson and April Morris

22 August 2022

The National Park Service was founded in 1916, spurred on by a worldwide national park movement that began with the establishment of Yellowstone National Park. Today, the National Park Service oversees more than 400 areas in the 50 states, the District of Columbia, American Samoa, Guam, Puerto Rico, Saipan, and the Virgin Islands. The state of California boasts the most national parks, with a total of nine parks: Channel Islands, Death Valley, Joshua Tree, Kings Canyon, Lassen Volcanic, Pinnacles, Redwood, Sequoia, and Yosemite.

The aim of this project is to answer questions about the visitation at national parks in California between the years of 2019 and 2020, including which park saw the most visitors each year, which months were busiest for each park or all parks, which park saw the most overnight visitors, and whether or not southern parks saw more visitors in the winter months due to higher temperatures at southern latitudes. These statistics, as well as the physical location of each park, can provide insight on how social and seasonal changes may affect visitation to different areas of the state.

Our data was pulled from the National Park Service Visitor Use Statistics report page. The format of the data was in reports separated by several different topics. In order to import it into our database, we combed through the reports and entered it into a Microsoft Excel spreadsheet. The spatial aspect of the data also came from the National Park Service in the form of polygon shapefiles. The National Park polygons were imported into ArcGIS and joined to the visitor use statistics data tables.

Database Design

For this database, we had four entities: the California National Parks, which included the park name and its geometry as a polygon shapefile; Yearly Visitors, which included the park name and the number of visitors to each park for 2019 and 2020; Monthly Visitors, which included the park name and the visitors for each month of 2019 and 2020; and the Overnight Visitors, which included the park name and the overnight

stays for each month of 2019 and 2020. The park name was used as the unique identifier and the primary key, and was used to join all of the data together.

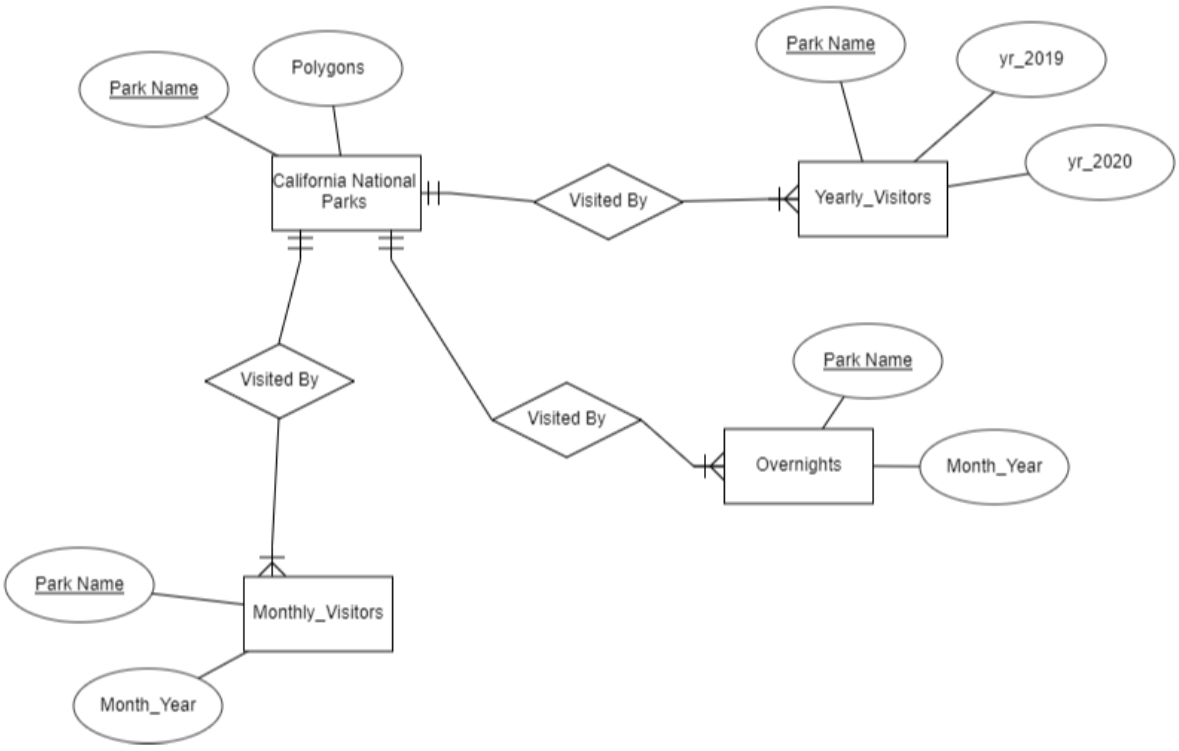

Figure 1: ER Diagram

From our ER diagram, we were able to draft the following relational schema – as all of the data simply relates back to the National Parks entity, it was a fairly straightforward concept.

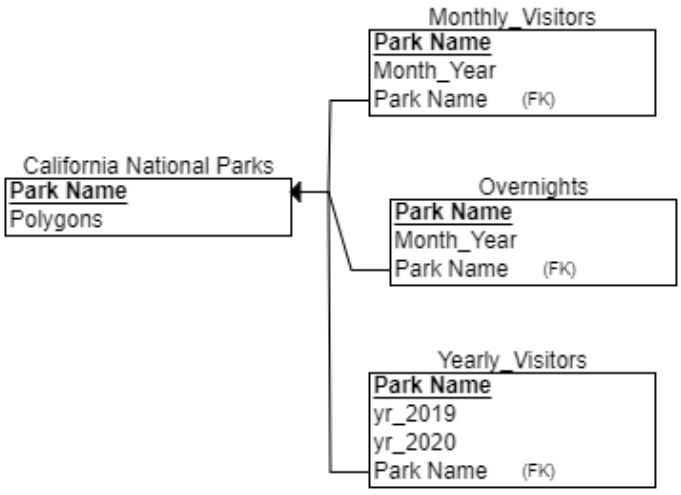

Figure 2: Relational Diagram

Methods

We were able to easily answer our research questions by performing SQL queries. To find the park with the most visitors in 2019 and compare with the number of visitors in 2020, the following query was used:

```
SELECT park_name, yr_2019, yr_2020
FROM yearly_visitors
ORDER by yr_2019 DESC;
```

In order to determine the busiest months for all parks for each year, we used the aggregate average function for each month for all parks:

```
SELECT round(avg(jan_2019)) as Jan19, round(avg(Feb_2019)) as Feb19,
```

```
round(avg(Mar_2019)) as Mar19, round(avg(Apr_2019)) as Apr19,
round(avg(May_2019)) as May19, round(avg(Jun_2019)) as Jun19,
round(avg(Jul_2019)) as Jul19, round(avg(Aug_2019)) as Aug19,

round(avg(Sep_2019)) as Sep19, round(avg(Oct_2019)) as Oct19,
round(avg(Nov_2019)) as Nov19, round(avg(Dec_2019)) as Dec19

FROM monthly_visitors;


SELECT round(avg(jan_2020)) as Jan20, round(avg(Feb_2020)) as Feb20,

round(avg(Mar_2020)) as Mar20, round(avg(Apr_2020)) as Apr20,
round(avg(May_2020)) as May20, round(avg(Jun_2020)) as Jun20,
round(avg(Jul_2020)) as Jul20, round(avg(Aug_2020)) as Aug20,
round(avg(Sep_2020)) as Sep20, round(avg(Oct_2020)) as Oct20,
round(avg(Nov_2020)) as Nov20, round(avg(Dec_2020)) as Dec20

FROM monthly_visitors;
```

To find which months each year saw the most overnight stays, we used a similar query by using the aggregate average function for each month for all parks:

```
SELECT round(avg(jan_2019)) as Jan19, round(avg(Feb_2019)) as Feb19,

round(avg(Mar_2019)) as Mar19, round(avg(Apr_2019)) as Apr19,
round(avg(May_2019)) as May19, round(avg(Jun_2019)) as Jun19,
round(avg(Jul_2019)) as Jul19, round(avg(Aug_2019)) as Aug19,

round(avg(Sep_2019)) as Sep19, round(avg(Oct_2019)) as Oct19,
round(avg(Nov_2019)) as Nov19, round(avg(Dec_2019)) as Dec19

FROM overnights;
```

```
SELECT round(avg(jan_2020)) as Jan20, round(avg(Feb_2020)) as Feb20,
round(avg(Mar_2020)) as Mar20, round(avg(Apr_2020)) as Apr20,
round(avg(May_2020)) as May20, round(avg(Jun_2020)) as Jun20,
round(avg(Jul_2020)) as Jul20, round(avg(Aug_2020)) as Aug20,
round(avg(Sep_2020)) as Sep20, round(avg(Oct_2020)) as Oct20,
round(avg(Nov_2020)) as Nov20, round(avg(Dec_2020)) as Dec20
FROM overnights;
```

For our final research question we were interested to see whether parks located in the southern part of the state saw more visitors in winter months rather than in summer months due to climatic variations between the northern and southern parts of California. To determine which parks were located further south, we used the centroid of the park polygons. If the latitude of the centroid fell below 37 degrees, they were identified as southern parks. The parks we ended up with were Channel Islands, Death Valley, Joshua Tree, Pinnacles, Sequoia, and Kings Canyon.

```
SELECT park_name, ST_Y(ST_Centroid(geom)) as Latitude
FROM monthly_visitors
WHERE ST_Y(ST_Centroid(geom)) < 37;
```

To find monthly visitors to only southern parks, we combined the southern parks query with the aggregate average function:

```
SELECT
```

```
round(avg(jan_2019)) as Jan19, round(avg(feb_2019)) as Feb19,
round(avg(mar_2019)) as Mar19, round(avg(apr_2019)) as Apr19,
round(avg(may_2019)) as May19, round(avg(jun_2019)) as Jun19,
round(avg(jul_2019)) as Jul19, round(avg(aug_2019)) as Aug19,
round(avg(sep_2019)) as Sep19, round(avg(oct_2019)) as Oct19,
round(avg(nov_2019)) as Nov19, round(avg(dec_2019)) as Dec19
FROM monthly_visitors
WHERE ST_Y(ST_Centroid(geom)) < 37;

SELECT
round(avg(jan_2020)) as Jan20, round(avg(Feb_2020)) as Feb20,
round(avg(Mar_2020)) as Mar20, round(avg(Apr_2020)) as Apr20,
round(avg(May_2020)) as May20, round(avg(Jun_2020)) as Jun20,
round(avg(Jul_2020)) as Jul20, round(avg(Aug_2020)) as Aug20,
round(avg(Sep_2020)) as Sep20, round(avg(Oct_2020)) as Oct20,
round(avg(Nov_2020)) as Nov20, round(avg(Dec_2020)) as Dec20
FROM monthly_visitors
WHERE ST_Y(ST_Centroid(geom)) < 37;
```

Results

Our results were not entirely what we expected. When it came to comparing the number of visitors to each park in 2019 and 2020, we were expecting the most popular park in 2019 would still be the most popular park in 2020, but that was not the case. Figure 3 shows the result of our yearly visitor queries in bar graph form.

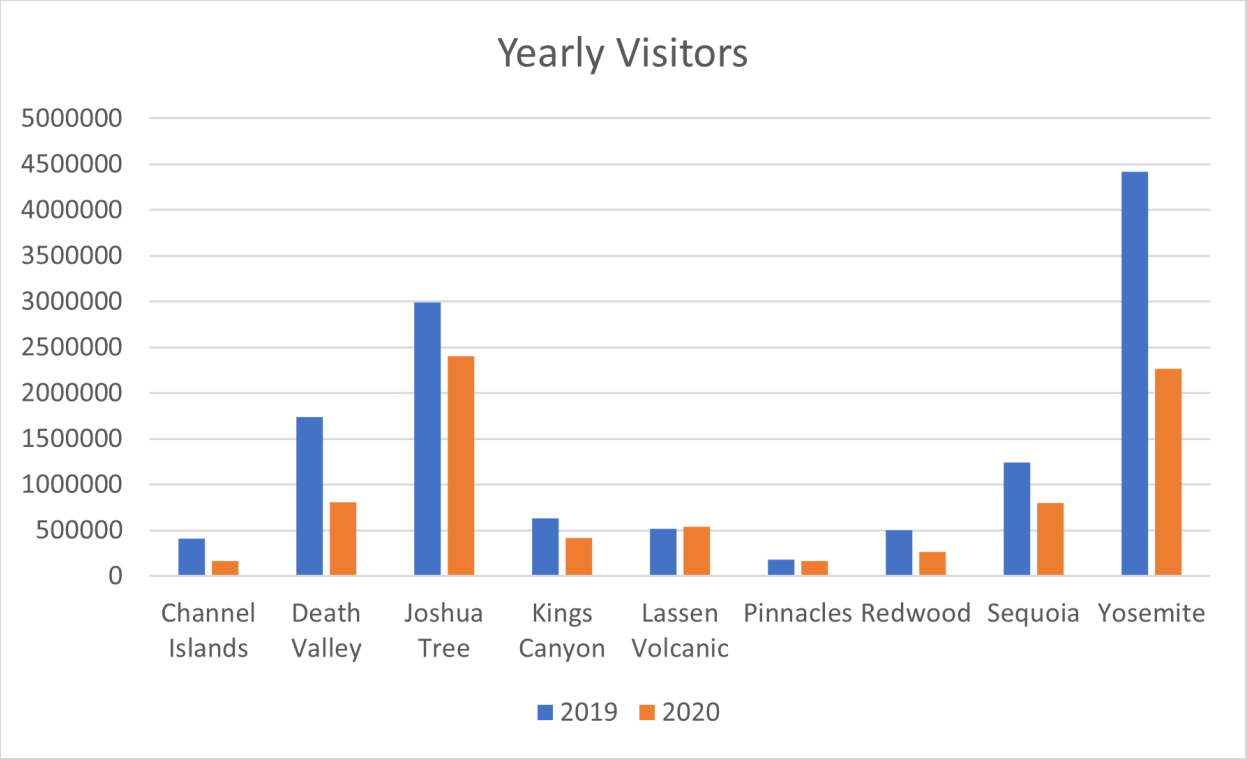

Figure 3: Yearly Visitors

In 2019, Yosemite had the most visitors by far with 4,422,861 visitors. However in 2020, Joshua Tree had slightly more visitors than Yosemite, 2,399,542 and 2,268,313 respectively. We anticipated slightly fewer visitors overall in 2020 than in 2019 due to the coronavirus pandemic, and while that was the case for some parks, other parks saw drastically fewer visitors like Yosemite and Death Valley.

We had less concrete expectations regarding the busiest months for all parks. We anticipated that the summer months would be busy due to summer breaks from

schools. We also expected a decline in visitors around the initial pandemic stay at home order in March of 2020 and a general decline in 2020 due to restricted access to national parks related to the pandemic. Our assumptions were largely correct. July was the busiest month for all parks for both years, with an average of 170,097 visitors for all parks in 2019 and 131,998 in 2020. August was a close second for the busiest month. There was a dramatic dip in visitation in April 2020, which slowly climbed to peak in July.The average visitation to all parks in April 2019 decreased by 97.64% in April 2020. The results of our queries for average monthly recreation visits are visualized in figure 4 and figure 5.

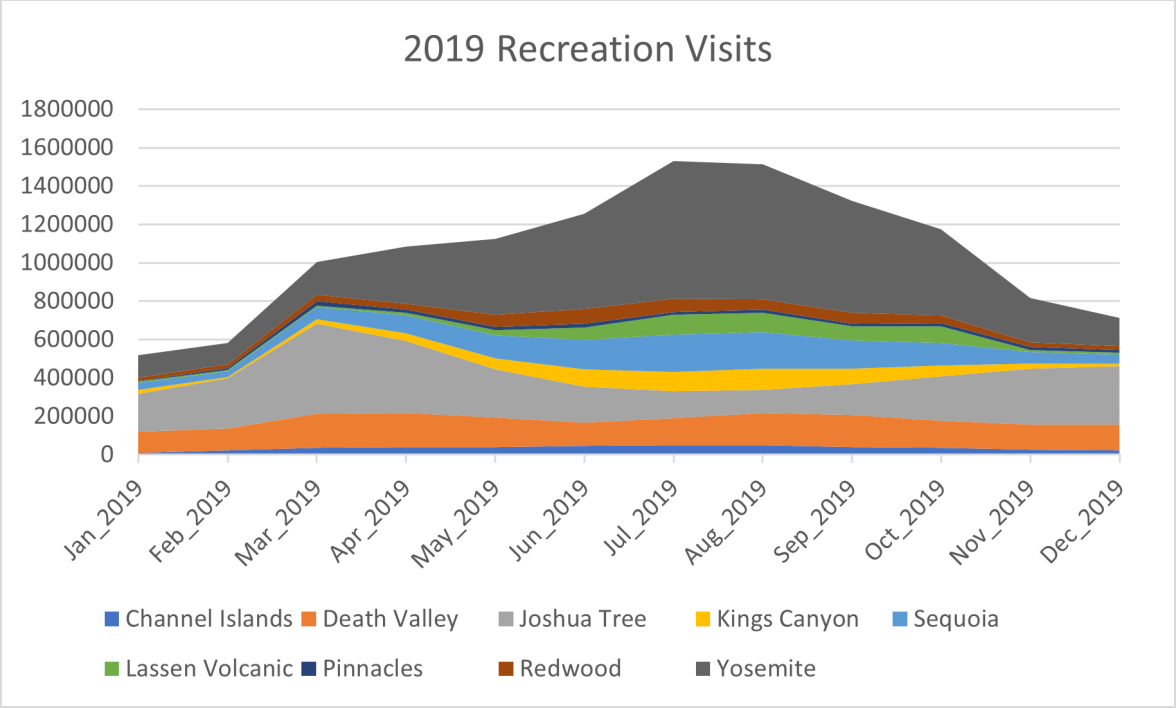

Figure 4: 2019 Recreation Visits

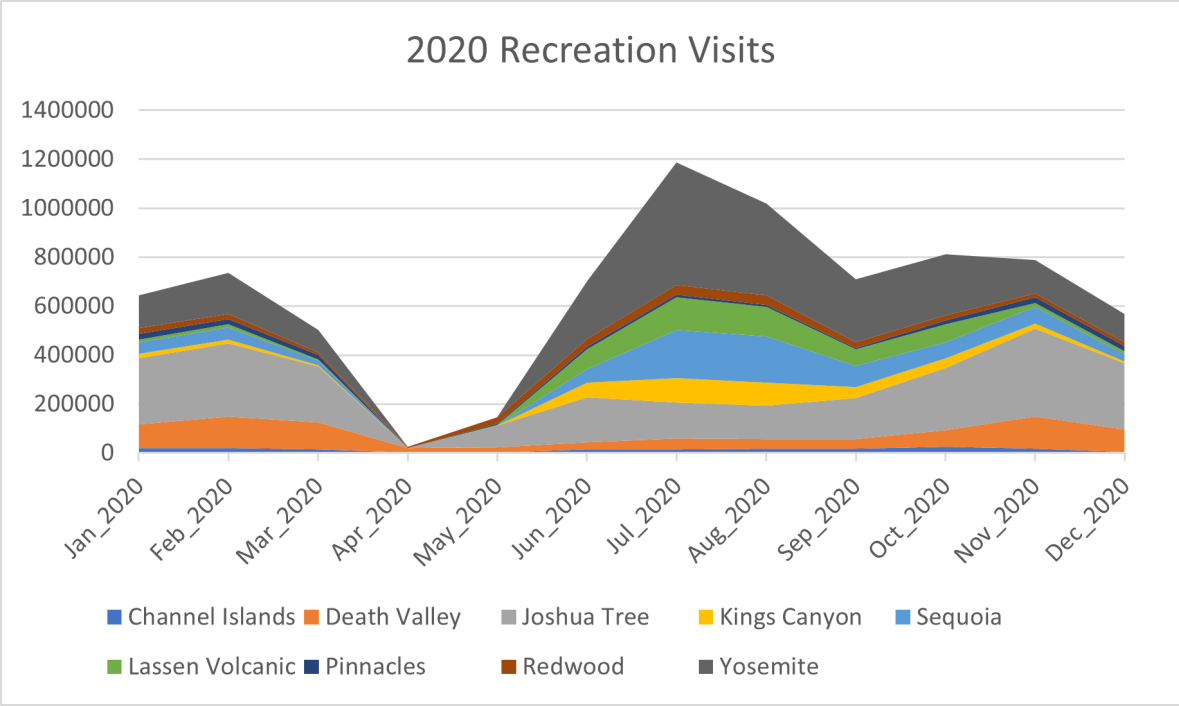

Figure 5: 2020 Recreation Visits

The expectations we had for overnight stays were roughly the same as what we expected for the monthly visits. We anticipated a busy summer and a slow down in the spring of 2020. The results matched our hypothesis and were largely similar to the recreational visits results, with the exception of August being the busiest month for overnight stays both years instead of July. The results of our queries for average monthly overnight visits are visualized in figure 6 and 7.

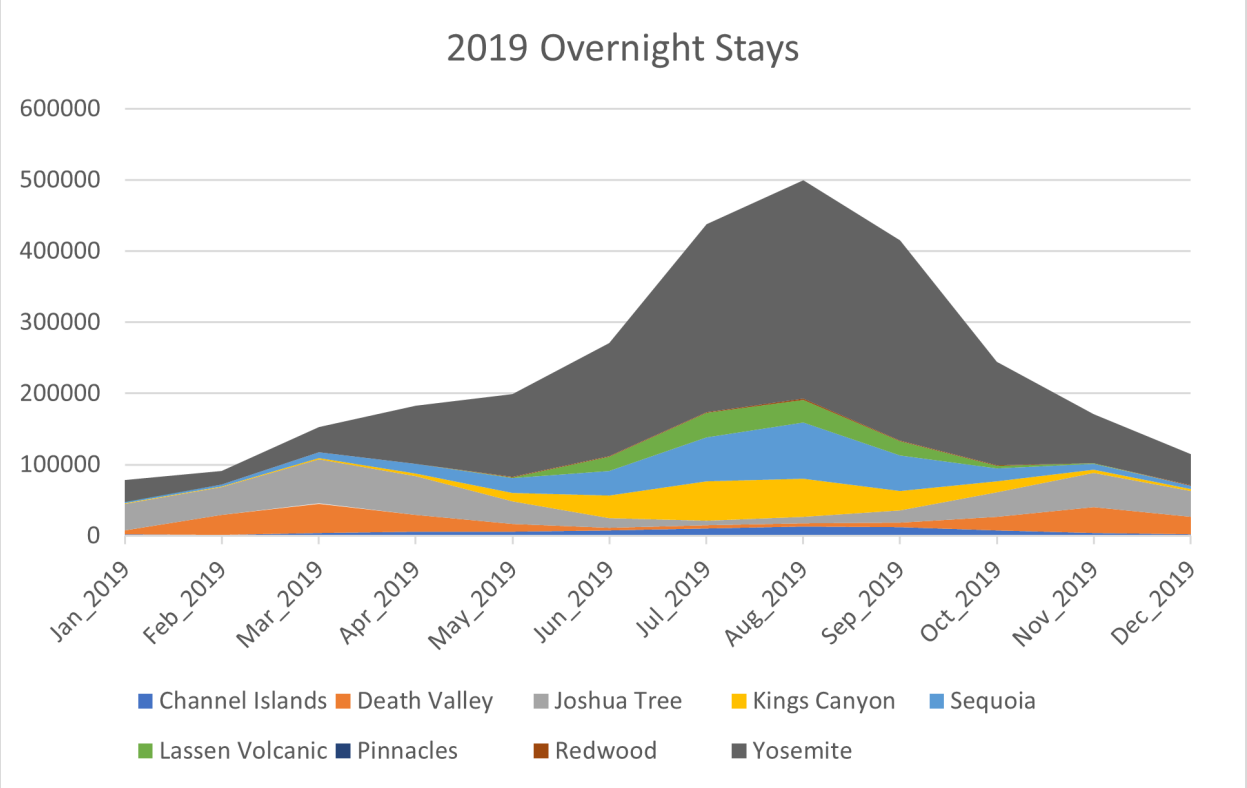

Figure 6: 2019 Overnight Stays

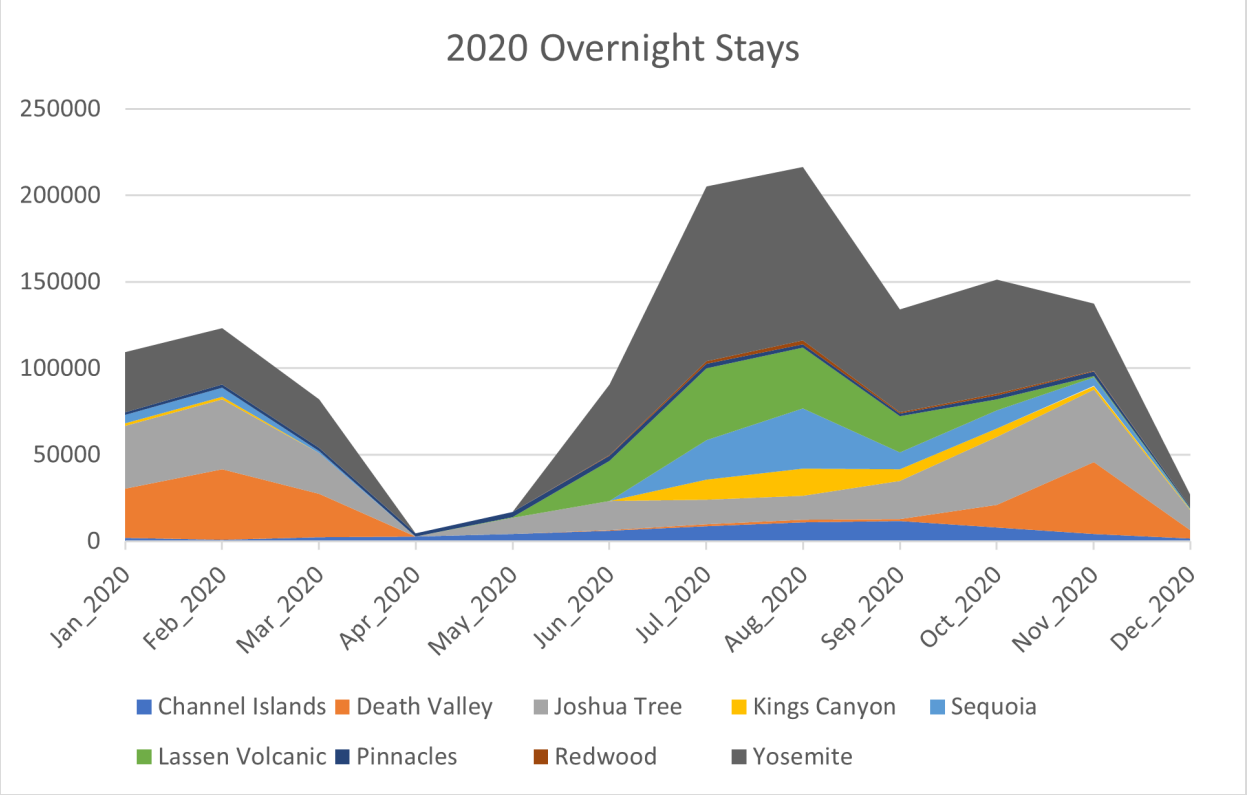

Figure 7: 2020 Overnight Stays

Our query to find California parks with centroid latitudes south of 37° N yielded 6 of the 9 California national parks: Channel Islands, Death Valley, Joshua Tree, Pinnacles, Sequoia, and Kings Canyon. The result of this spatial query is visualized in a map in figure 8.

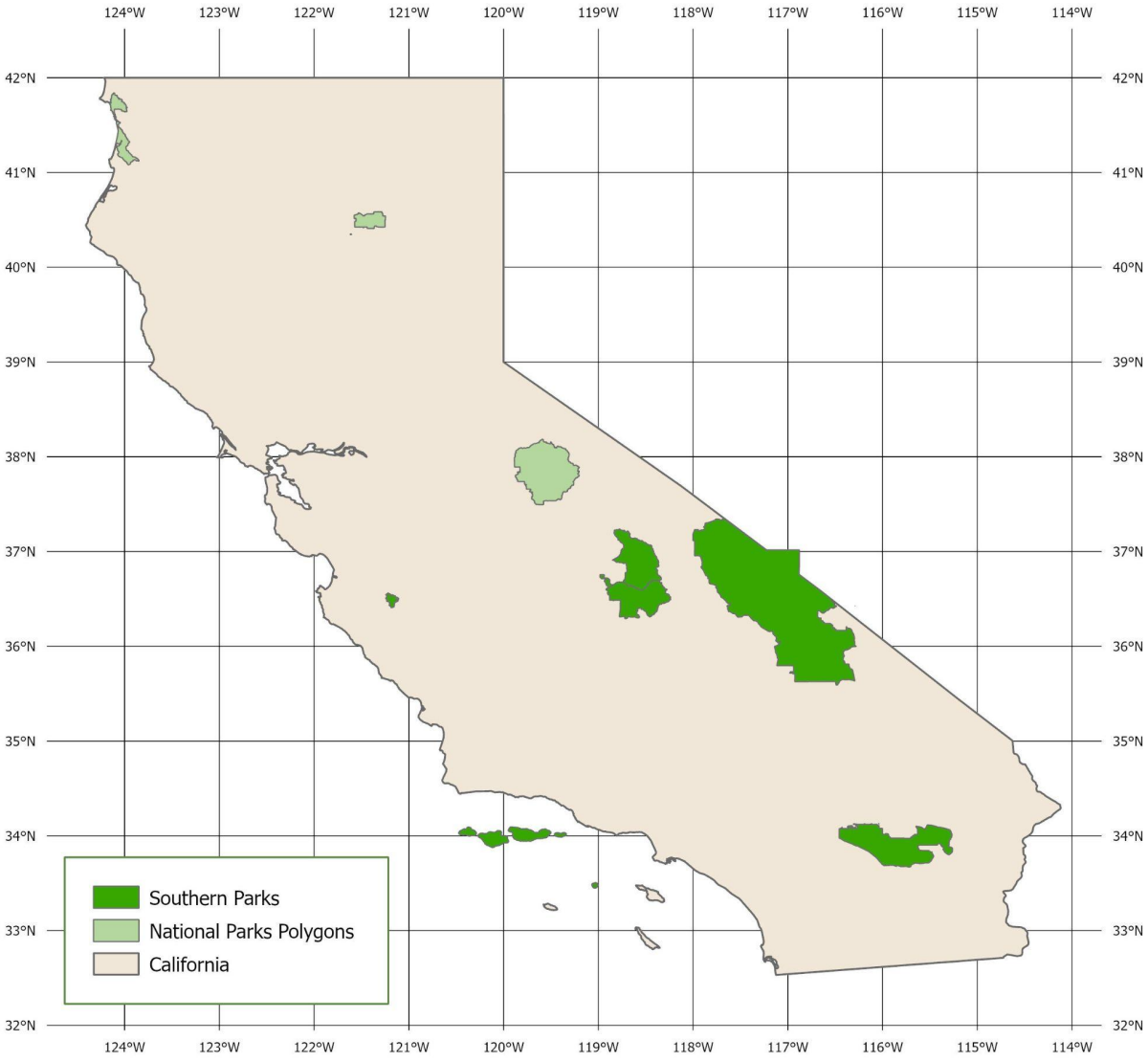

Figure 8: Southern parks with centroids south of 37° N

When it came to looking at visitation to southern parks during the summer, we expected fewer visitors than the data actually showed. However, we were correct in that summer was not the most popular time to visit these parks – it was actually spring that saw the most visitors, at least in 2019. March was the busiest month for the southern parks in 2019 (132,130 visitors), followed closely by April (126,691 visitors). No other month of the year in 2019 had an average number of visitors above 110,000.

In 2020, the results were much different. While there were many visitors in March (65,707 on average), it was still 50.3% less than in March of 2019. There was a 97.3% decline in average number of visitors in April (3,394 visitors). Due to the pandemic, the previous year's busiest month was this 2020's slowest. The busiest month in 2020 was November (103,140 visitors), followed by February (89,176 visitors). While our initial expectations were still correct that summer was not the busiest season, the pandemic dramatically altered the pattern from the previous year.

Discussion/Conclusions

There are many factors involved when it comes to how and when people choose to visit national parks. From the weather, school obligations, and unprecedented pandemics, this project has shown that the national parks continue to be a beloved destination. While we were able to answer our research questions and establish a rudimentary understanding of visitation habits for these parks, many more questions arose during this process.

One interesting inquiry would be to delve further into the climate aspect – our method of locating southern parks by latitude, where we inferred it was warmer, was simple and effective but the unique topography of the state makes for a more complicated climate. While Kings Canyon and Sequoia National Parks are at a similar latitude to Death Valley National Park, their higher elevation creates a climate more similar to northern California parks.This idea could be explored further with a more robust dataset and a larger sample size of years. Monthly average tables could be added to the database for climate variables such as temperature and precipitation. An attribute for elevation of each park could also be added. Such information could be further used to study how climate change affects visitation. Extreme heat and drought in California have led to a higher occurence of wildfire in the past few years causing another obstacle to park visitation. Whether rising temperatures simply become unbearable or other hazards and damages are inflicted on these parks, visitation patterns will likely reflect changes in climate.

References

National Park Service. *Visitor Use Statistics.* Integrated Resource Management Applications Portal. Retrieved August 5, 2022 from https://irma.nps.gov/STATS/

National Park Service. 2022, June 30. *Administrative Boundaries of National Park System Units.* Integrated Resource Management Applications Data Store. https://irma.nps.gov/DataStore/Reference/Profile/2224545?Inv=True

Natural Earth. *Admin 1 - States, Provinces.* 1:10 M Cultural Vectors. https://www.naturalearthdata.com/downloads/10m-cultural-vectors/