The Economical Impact of COVID-19

Group Members:
Brandon Nguyen, Calvin Truong, David Hwang, Daniel Kim, Grant Huang, Lei Wang

Background Information

- The Coronavirus also known as COVID-19 an infectious disease that has been circulating worldwide causing a global pandemic.
- Originated within Bats and made its way to wet markets
- Stronger than the common flu or cold
- Fever, shortness of breath
- Created a pause on daily life
    - Unemployment
    - Financial distress
    - Company shutdown

Business Questions

- How has COVID-19 affected retail, workplaces, transit stations, grocery stores & parks across the country?
- Which state is affected the most?
- Which city in California was affected the most?
- Comparison between state with highest confirmed cases to state with lowest confirmed?
- How many recovered vs how many dead?
- How does COVID 19 affects our lives?

Dataset

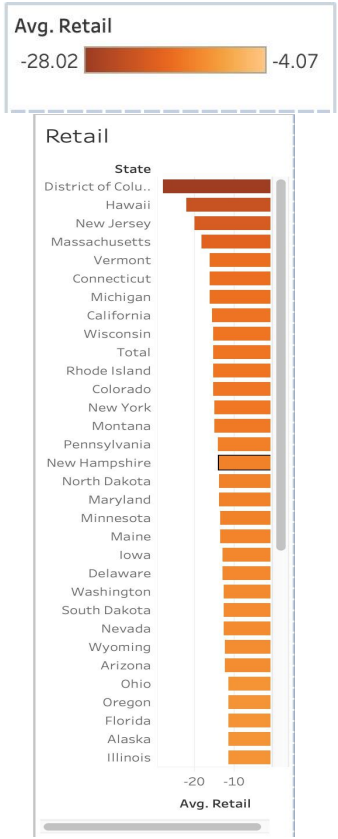
Attribute	Data Type	Example
state	String	Alabama
county	String	Total
date	Date	3/2/2020
retail	Integer	8
grocery_and_pharmacy	Integer	4
parks	Integer	-2
transit_stations	Integer	5
workplaces	Integer	2
residential	Integer	0

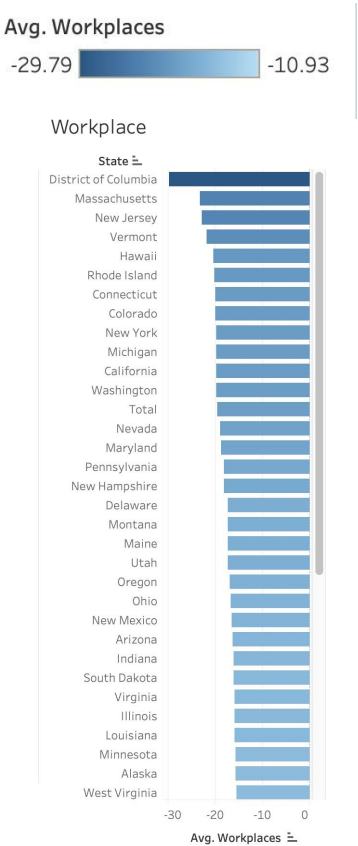
US- Mobility

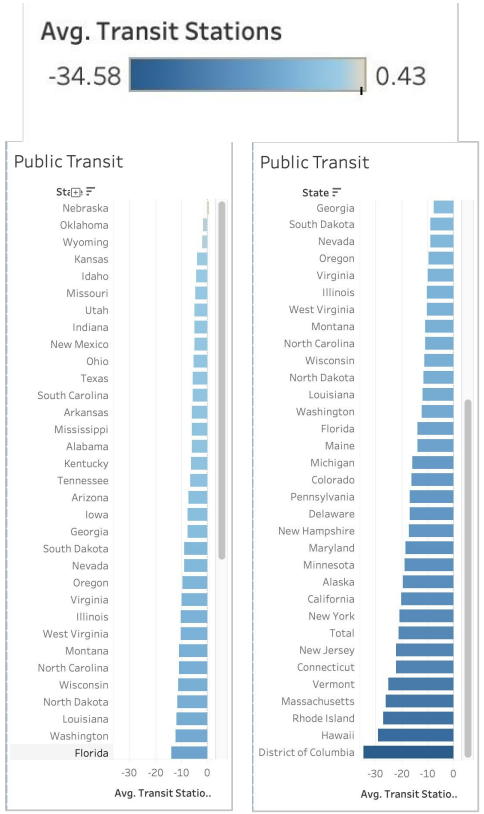
Attribute	Data type	Example
province_state	String	American Samoa
country_region	String	US
lat	Float	-14.271
long	Float	-170.132
combined_key	String	American Samoa, US
date	date	1/22/2020
confirmed	Integer	0

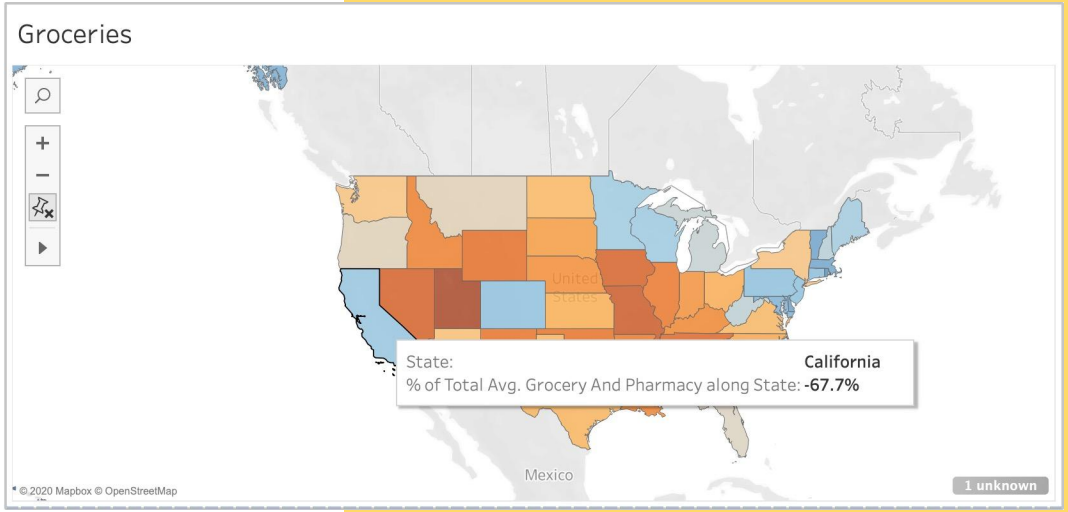
2019 COVID-19 in the US

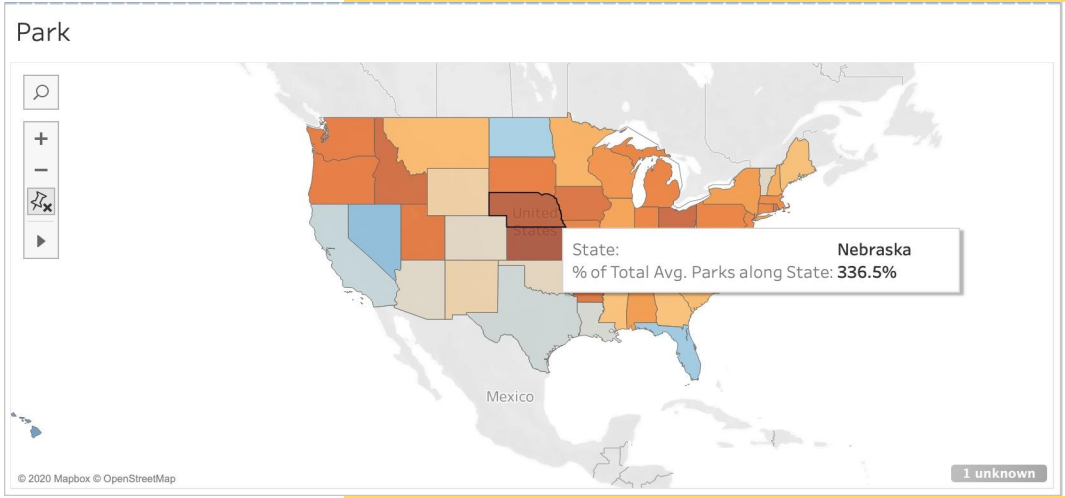
How has COVID-19 affected retail, transit stations, workplaces?



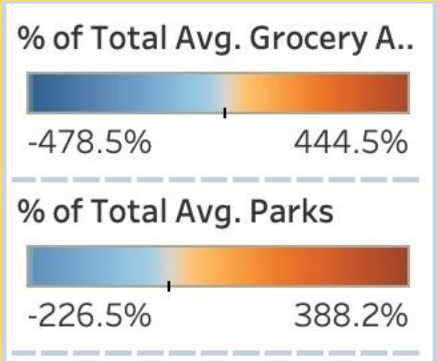




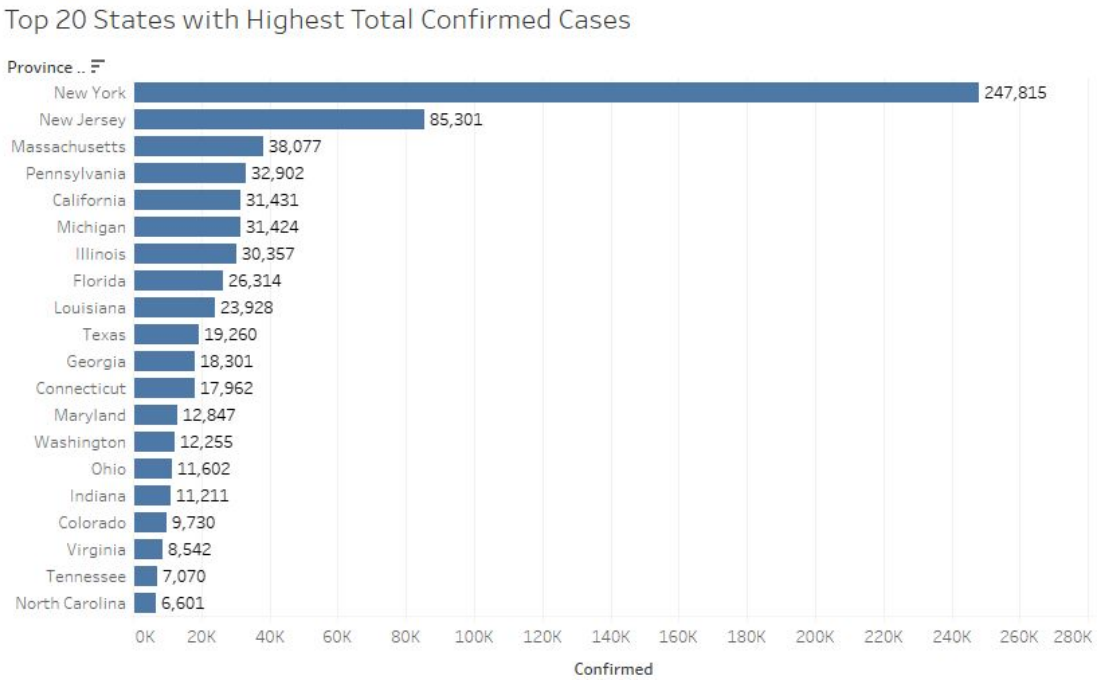




How has COVID-19 affected groceries and parks?



Which US State had the most confirmed cases?

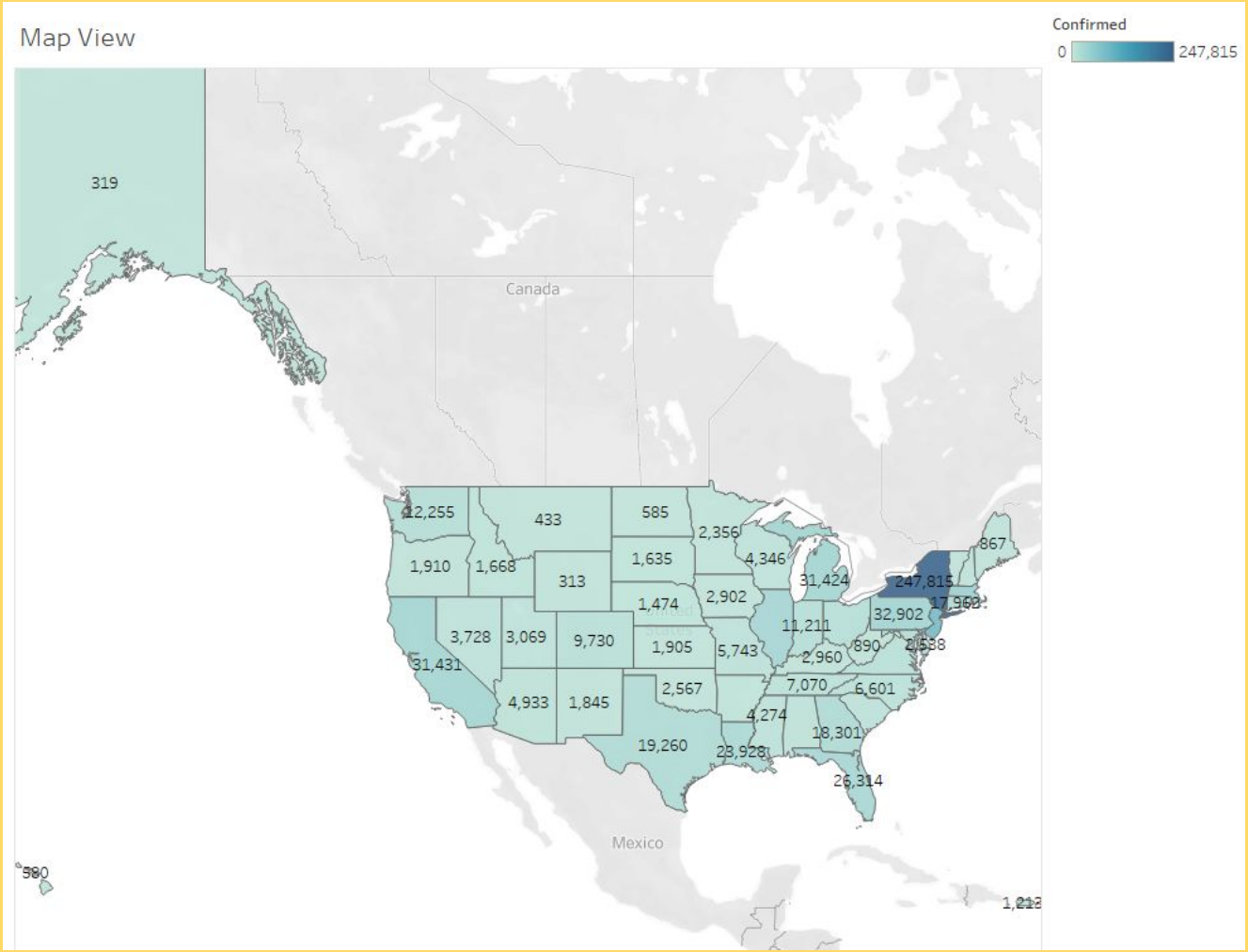


Confirmed

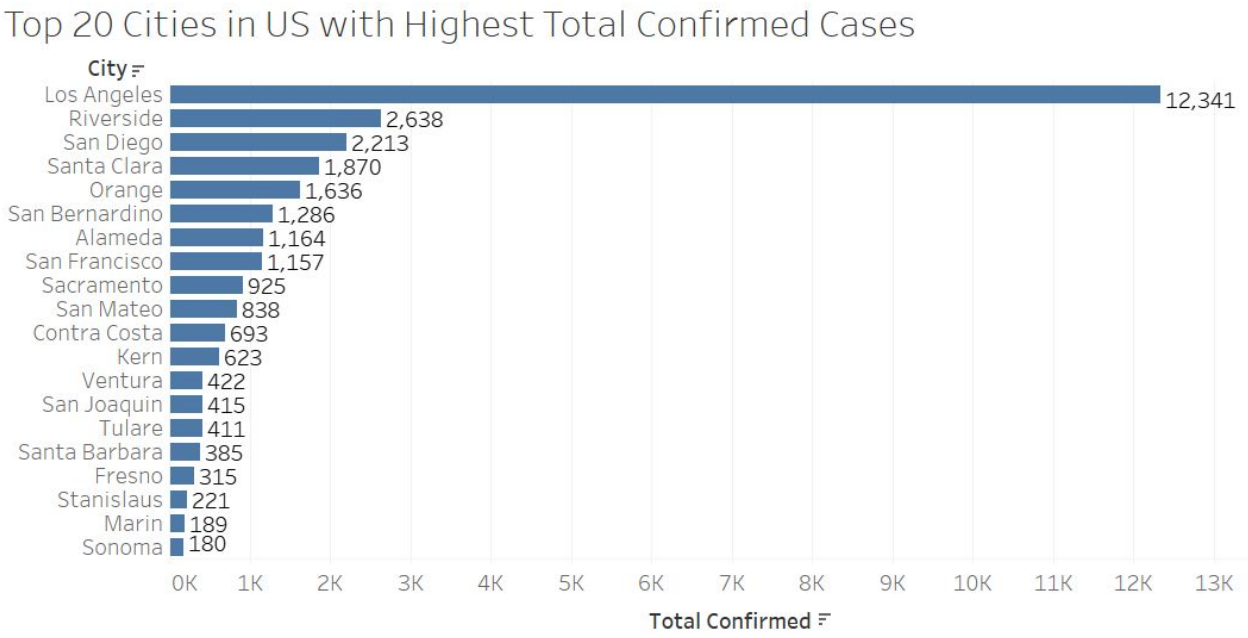
7K 248K

Province ..	
New York	36.29%
New Jersey	12.49%
Massachusetts	5.58%
Pennsylvania	4.82%
California	4.60%
Michigan	4.60%
Illinois	4.45%
Florida	3.85%
Louisiana	3.50%
Texas	2.82%
Georgia	2.68%
Connecticut	2.63%
Maryland	1.88%
Washington	1.79%
Ohio	1.70%
Indiana	1.64%
Colorado	1.42%
Virginia	1.25%
Tennessee	1.04%
North Carolina	0.97%

Data set is from Jan 22 - April 22


Map View
Confirmed
0
247,815
319
Canada
12,255
433
585
2,356
867
1,910
1,668
313
1,635
4,346
247,815
1,474
2,902
31,424
17,962
32,902
3,728
3,069
9,730
1,905
5,743
11,211
890
2,538
31,431
2,960
7,070
6,601
4,933
1,845
2,567
4,274
18,301
19,260
23,928
26,314
Mexico
980
1,212

Which city was affected the most in California?



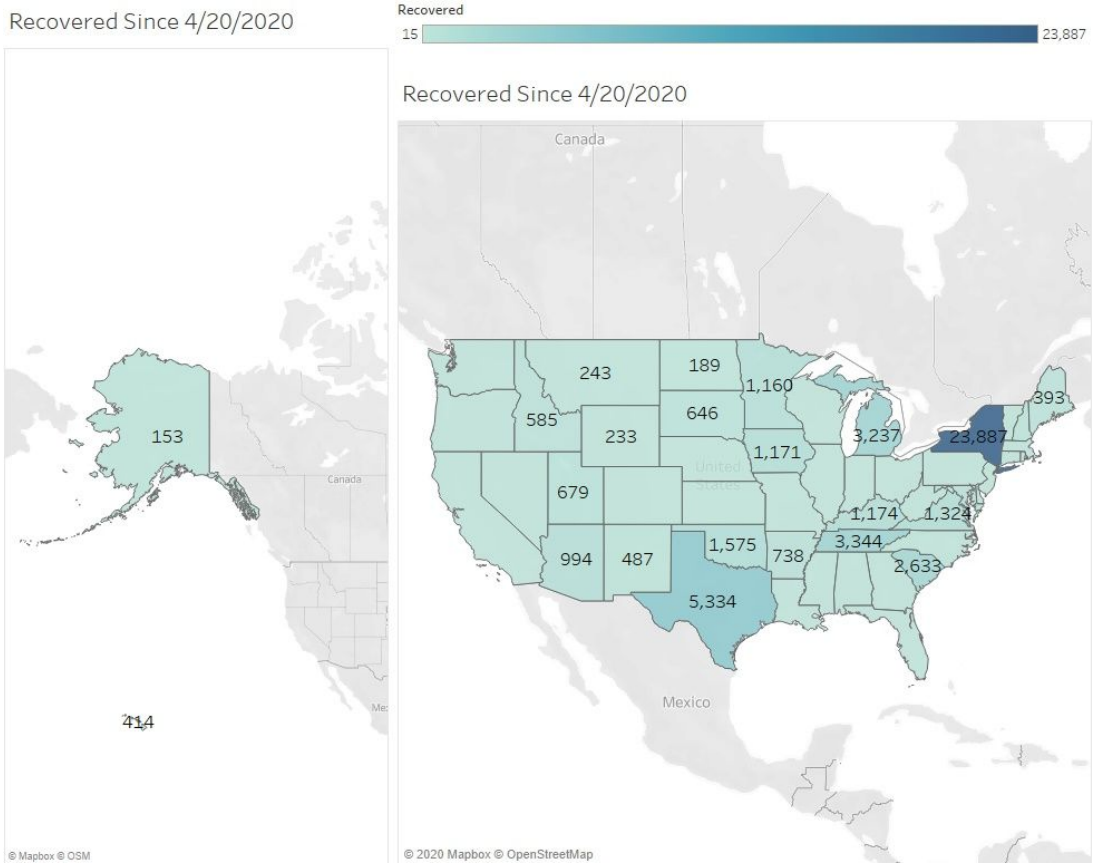
Total Confirmed

180 12,341

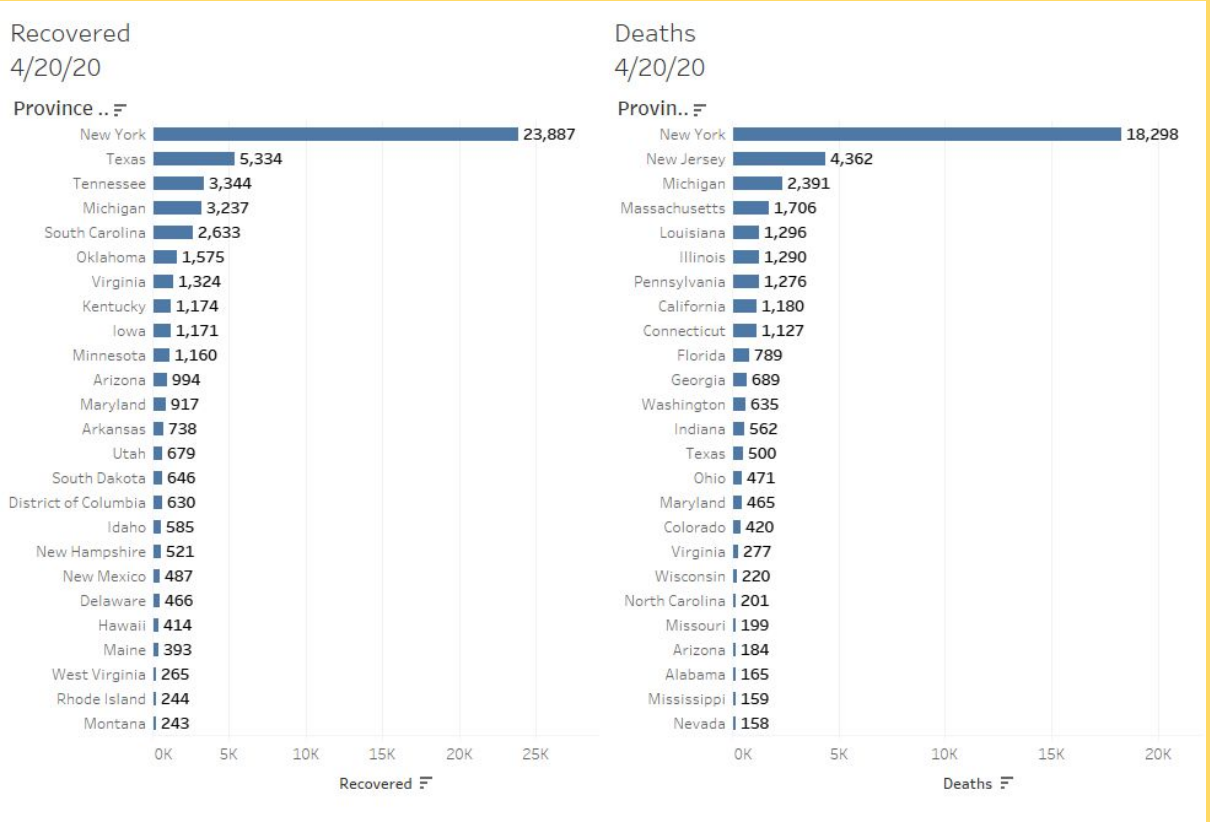
City	
Los Angeles	41.24%
Riverside	8.82%
San Diego	7.40%
Santa Clara	6.25%
Orange	5.47%
San Bernardino	4.30%
Alameda	3.89%
San Francisco	3.87%
Sacramento	3.09%
San Mateo	2.80%
Contra Costa	2.32%
Kern	2.08%
Ventura	1.41%
San Joaquin	1.39%
Tulare	1.37%
Santa Barbara	1.29%
Fresno	1.05%
Stanislaus	0.74%
Marin	0.63%
Sonoma	0.60%

- The date range is from Jan 22 - April 19
- Los Angeles has the most total confirmed cases in the California

Number of Patients Recovered by Region

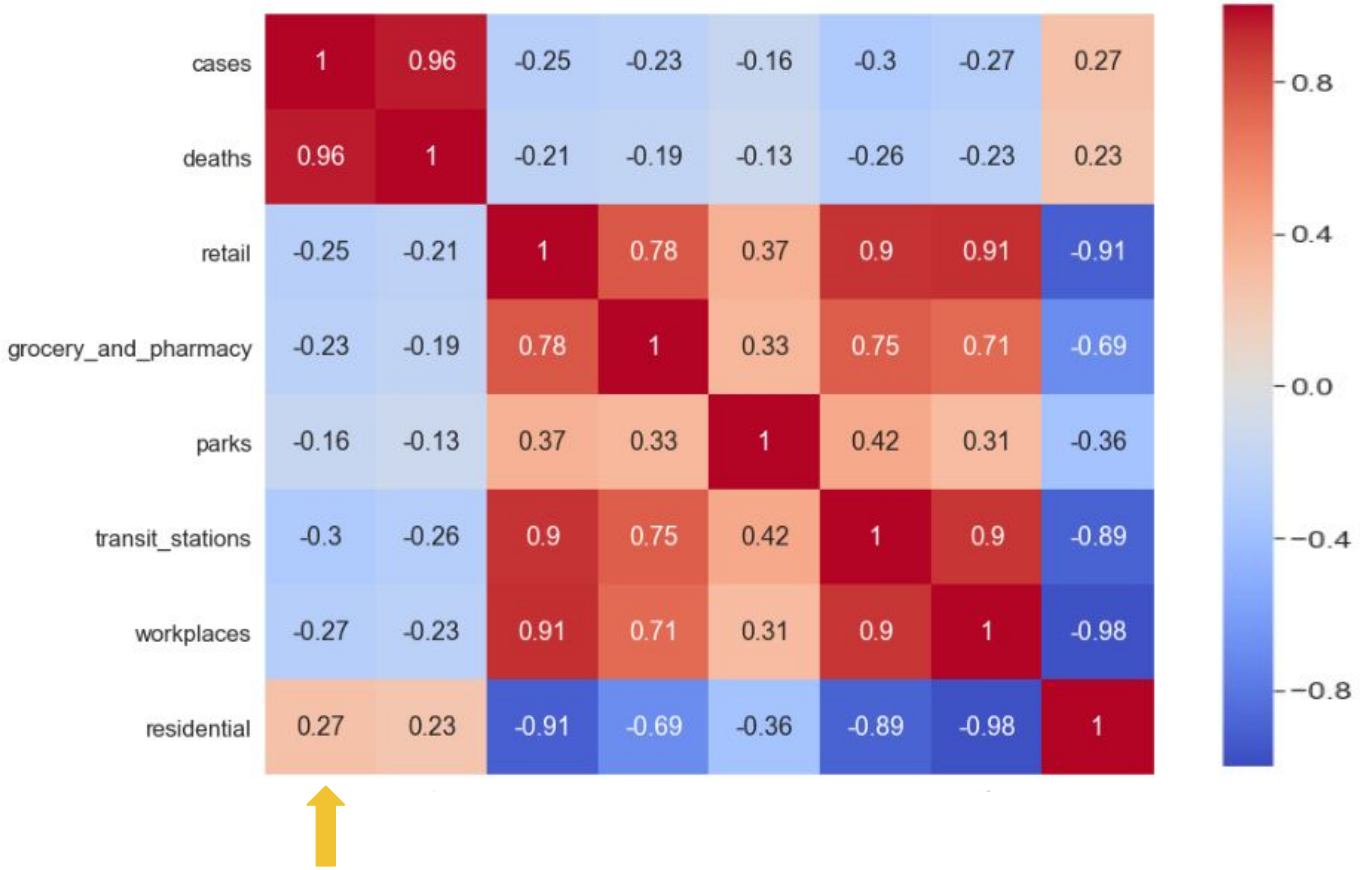


Recovered Patients vs Number of Death in Each State



Correlation Matrix

Negative-- Blue
Positive-- Red



How does COVID-19 Affect Our Lives?

Mobility $_{\text{Groc\&Phar}}$ = -0.002 **cases** + 0.0074 **deaths** + 0.4146 retail + 0.016 parks + 0.2164 transit_station + 0.2159 workplace + 0.8593 residential + 9.7266

R Squared = 0.633

OLS Regression Results

Dep. Variable:	grocery_and_pharmacy	R-squared:	0.633
Model:	OLS	Adj. R-squared:	0.632
Method:	Least Squares	F-statistic:	402.6
Date:	Mon, 04 May 2020	Prob (F-statistic):	0.00
Time:	21:25:41	Log-Likelihood:	-5923.5
No. Observations:	1640	AIC:	1.186e+04
Df Residuals:	1632	BIC:	1.191e+04
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	9.7266	0.399	24.395	0.000	8.945	10.509
cases	-0.0002	0.000	-1.855	0.064	-0.000	1.29e-05
deaths	0.0074	0.005	1.543	0.123	-0.002	0.017
retail	0.4146	0.027	15.080	0.000	0.361	0.468
parks	0.0160	0.007	2.240	0.025	0.002	0.030
transit_stations	0.2164	0.026	8.425	0.000	0.166	0.267
workplaces	0.2159	0.055	3.922	0.000	0.108	0.324
residential	0.8593	0.127	6.771	0.000	0.610	1.108

How Does Mobility Affect COVID-19?

Deaths # = 1.2297 retail - 0.3252 groc_phar - 0.0851 parks - 2.0323 transit_statation + 0.2359 workplace +2.5239 residential

R Squared = 0.071

OLS Regression Results

Dep. Variable:	deaths	R-squared:	0.071
Model:	OLS	Adj. R-squared:	0.067
Method:	Least Squares	F-statistic:	20.74
Date:	Mon, 04 May 2020	Prob (F-statistic):	1.57e-23
Time:	21:15:55	Log-Likelihood:	-10599.
No. Observations:	1640	AIC:	2.121e+04
Df Residuals:	1633	BIC:	2.125e+04
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-9.0669	8.054	-1.126	0.260	-24.864	6.730
retail	1.2297	0.506	2.430	0.015	0.237	2.222
grocery_and_pharmacy	-0.3252	0.428	-0.760	0.447	-1.164	0.514
parks	-0.0851	0.123	-0.689	0.491	-0.327	0.157
transit_stations	-2.0323	0.449	-4.527	0.000	-2.913	-1.152
workplaces	0.2359	0.955	0.247	0.805	-1.638	2.110
residential	2.5239	2.221	1.136	0.256	-1.833	6.881

References

Data Source:

- https://www.kaggle.com/roche-data-science-coalition/uncover
- https://www.kaggle.com/arghadeep/covid19-community-mobility-dataset