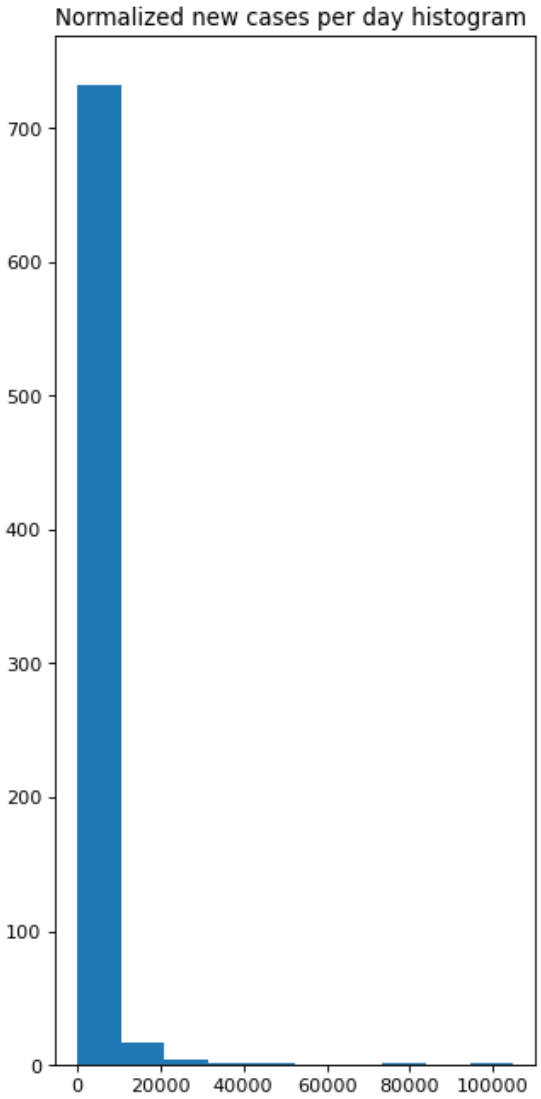
Stage 2:

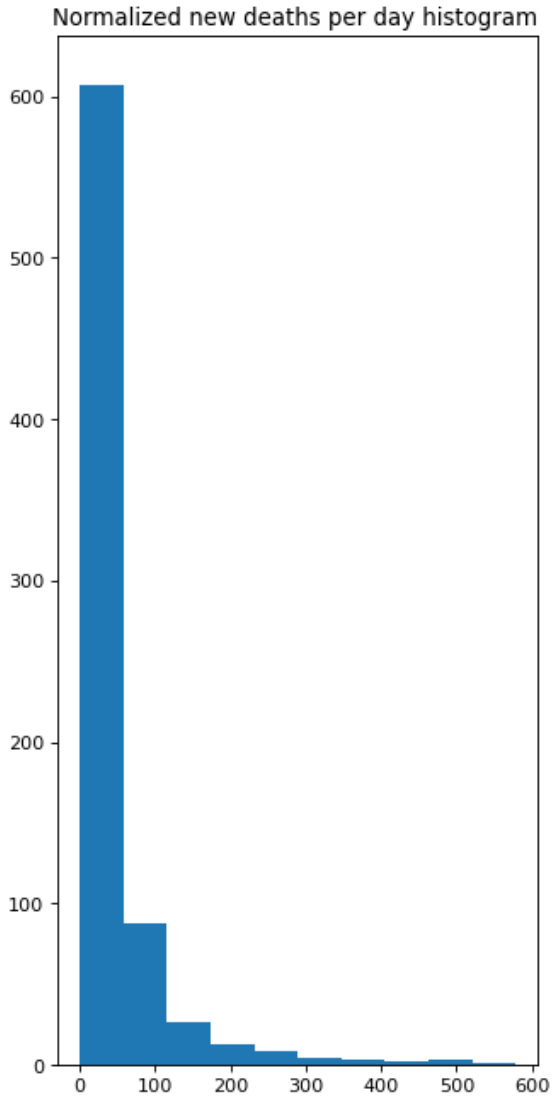
Weekly statistics (mean, median, mode) for number of new cases and deaths across a specific state.

In [32]:
```
print("mean =",round(al_state_data["Normalized_New_Cases_Per_Day"].mean()))
print("median =",round(al_state_data["Normalized_New_Cases_Per_Day"].median()))
print("variance =",round(al_state_data["Normalized_New_Cases_Per_Day"].var()))
print("skewness =",round(al_state_data["Normalized_New_Cases_Per_Day"].skew()))
print("kurtosis =",round(al_state_data["Normalized_New_Cases_Per_Day"].kurt()))
```

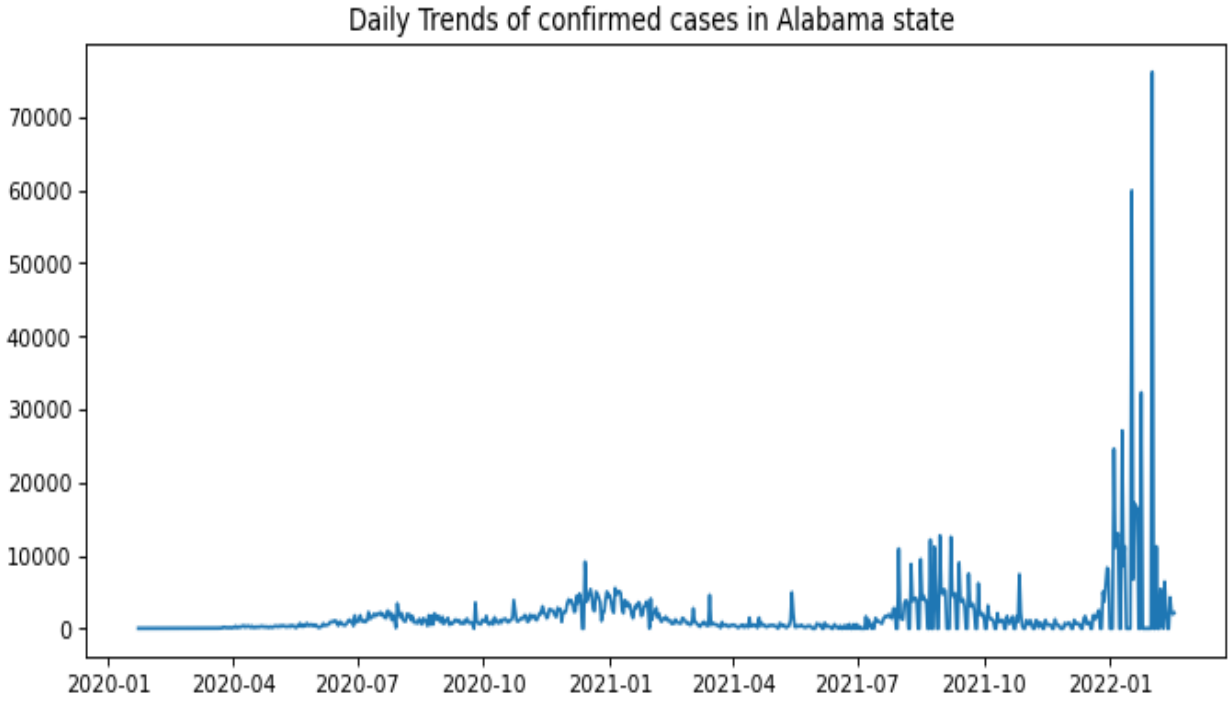
```
mean = 2265
median = 1029
variance = 34249186
skewness = 11
kurtosis = 165
```

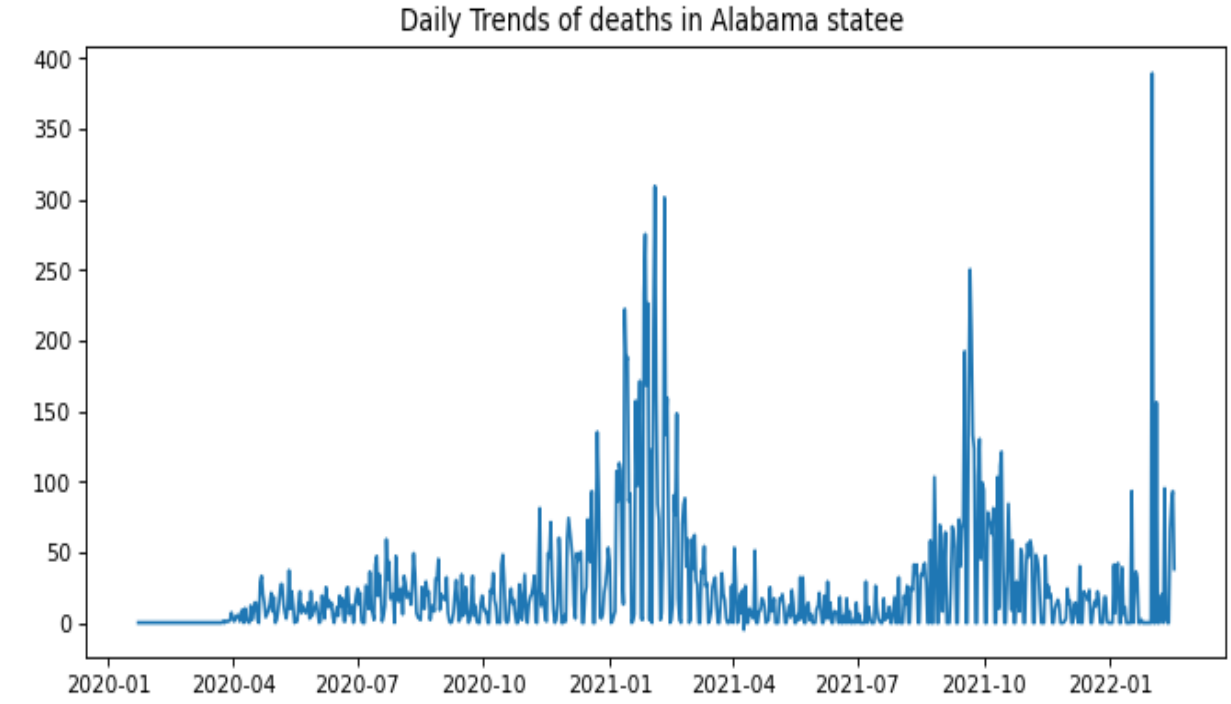
Histogram is skewed to the left, considering that Poisson distribution would be a good fit for distribution



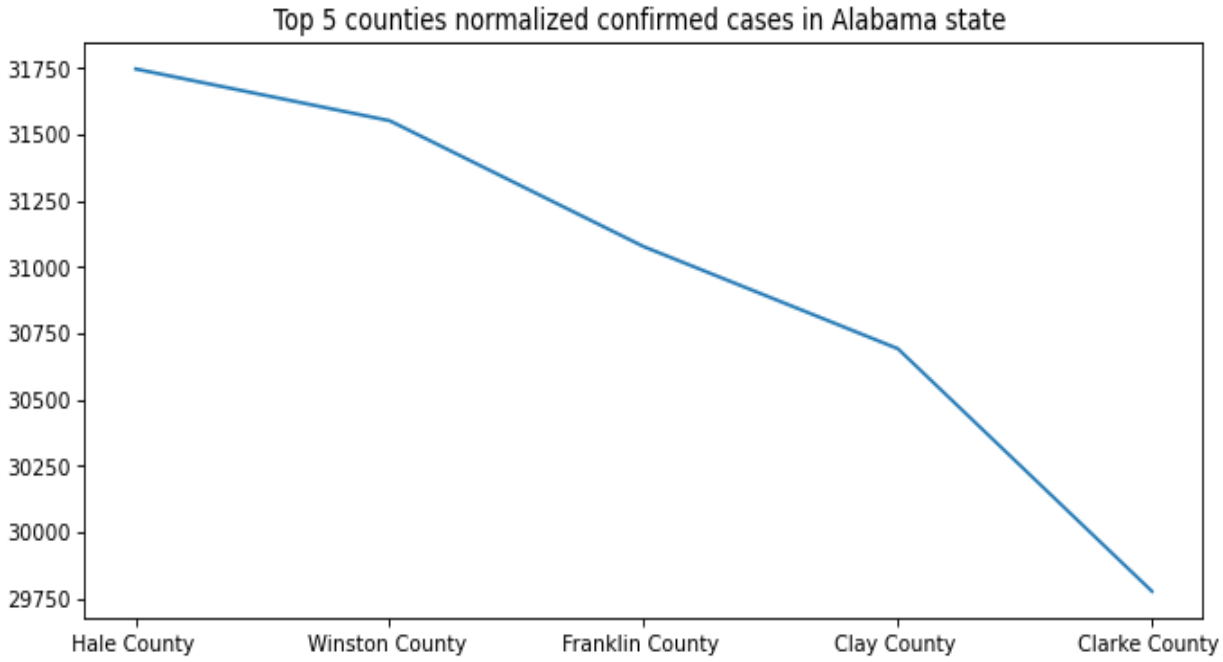


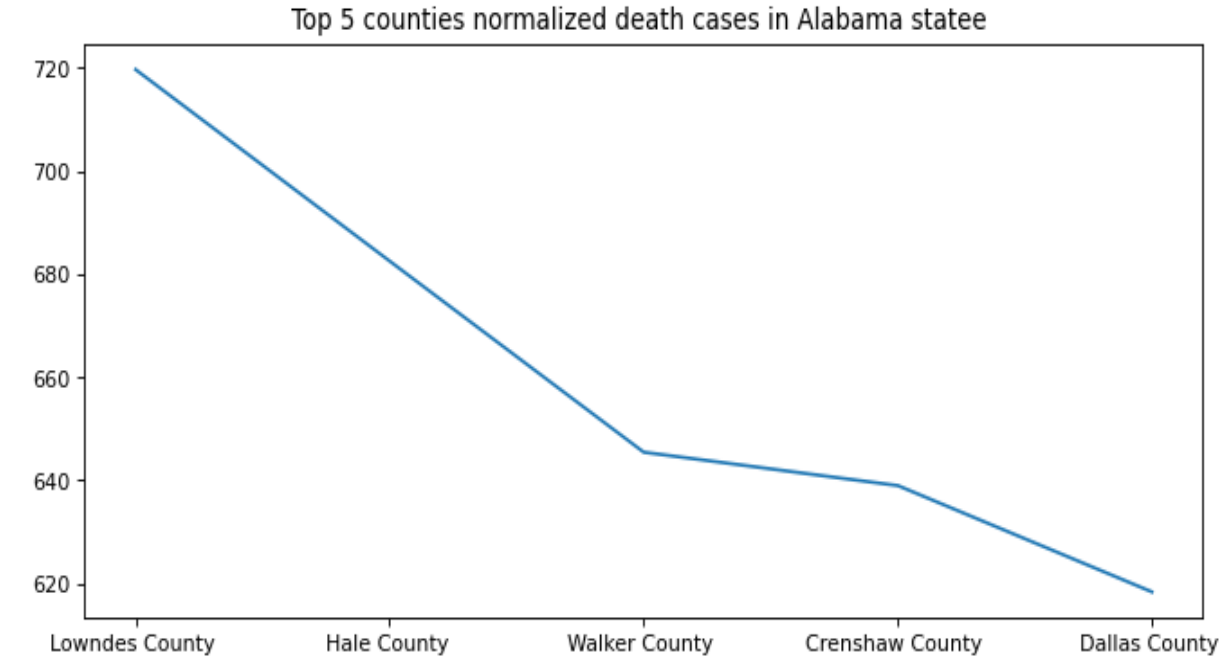
Daily Trends of confirmed cases in Alabama state



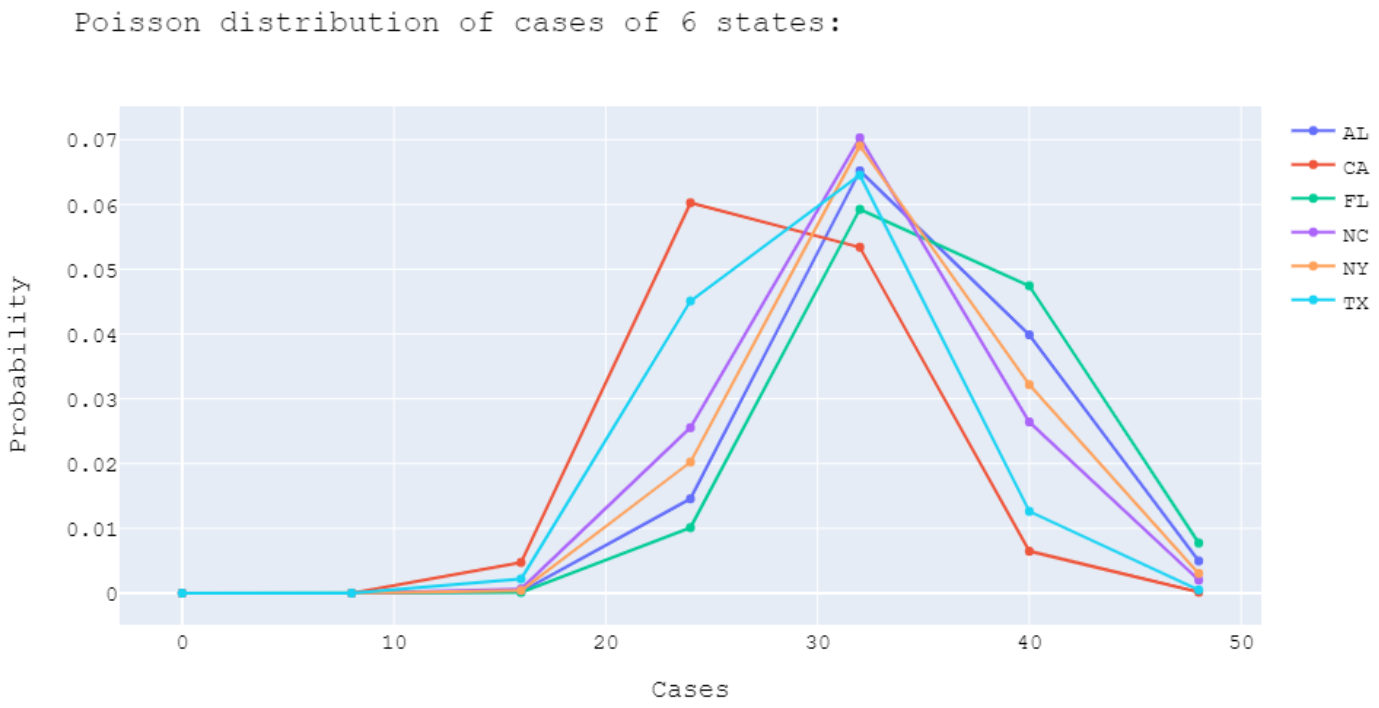


Top 5 counties with most effective cases in Alabama state

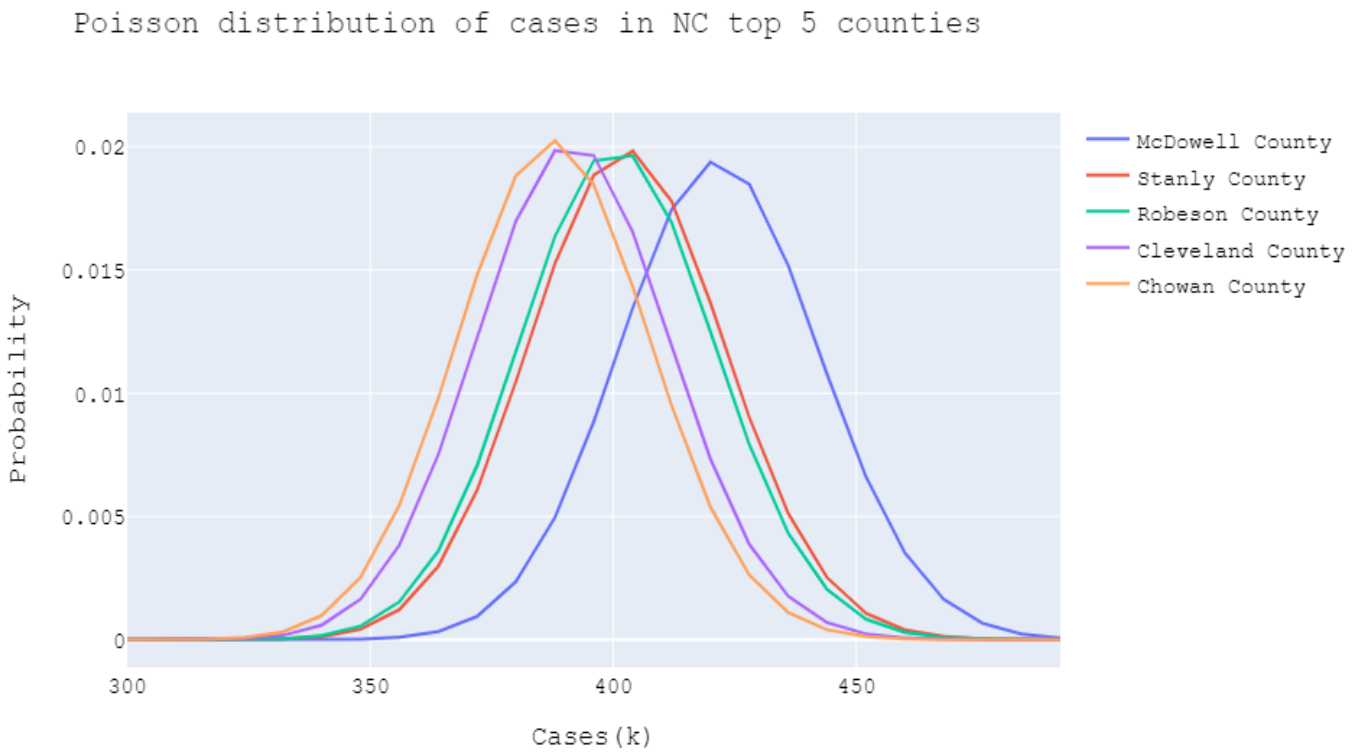




Poisson distribution plot of confirmed cases:



Poisson distribution plot of confirmed cases in NC state counties:



Correlation Matrix of Merged covid hospital dataset County grouped

```
merged_enrichment_grouped.corr()
```

51]:

	countyFIPS	population	Cases	Death	total_beds_7_day_avg	inpatient_beds_used_7_day_
countyFIPS	1.000000	-0.074984	-0.085329	-0.071868	-0.004485	-0.011
population	-0.074984	1.000000	0.900709	0.851136	0.388930	0.396
Cases	-0.085329	0.900709	1.000000	0.937894	0.375275	0.380
Death	-0.071868	0.851136	0.937894	1.000000	0.382275	0.375
total_beds_7_day_avg	-0.004485	0.388930	0.375275	0.382275	1.000000	0.970
inpatient_beds_used_7_day_avg	-0.011672	0.396266	0.380709	0.375504	0.970789	1.000
total_adult_patients_hospitalized_confirmed_covid_7_day_avg	-0.027966	0.370390	0.369202	0.361198	0.871493	0.881
al_pediatric_patients_hospitalized_confirmed_covid_7_day_avg	0.000793	0.303678	0.324913	0.305741	0.417846	0.446
inpatient_beds_7_day_avg	-0.009785	0.392277	0.380251	0.379581	0.985790	0.982
total_icu_beds_7_day_avg	-0.019979	0.393487	0.389868	0.374942	0.917246	0.931
icu_beds_used_7_day_avg	-0.017726	0.383396	0.376687	0.352280	0.897484	0.931
staffed_icu_adult_patients_confirmed_covid_7_day_avg	-0.026838	0.316027	0.318538	0.292967	0.799105	0.820

Observation:

- County population vs Cases and Death correlation coefficient are 0.9 abd 0.85 which indicates that counties with more population having more number of covid Cases and deaths.
- County population vs total_beds_7_day_avg coefficient is 0.38 which means counties population and total hospital beds doesn't have much dependency.

Correlation Matrix of Merged covid hospital dataset State grouped

```
merged_enrichment_grouped_State.corr()
```

[56]:

	population	Cases	Death	total_beds_7_day_avg	inpatient_beds_used_7_day_avg	total
population	1.000000	0.450430	0.398566	0.337262	0.368567	
Cases	0.450430	1.000000	0.926858	0.639839	0.671495	
Death	0.398566	0.926858	1.000000	0.647885	0.686301	
total_beds_7_day_avg	0.337262	0.639839	0.647885	1.000000	0.986332	
inpatient_beds_used_7_day_avg	0.368567	0.671495	0.686301	0.986332	1.000000	
total_adult_patients_hospitalized_confirmed_covid_7_day_avg	0.338357	0.633503	0.607707	0.892260	0.858015	
total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg	0.275515	0.354311	0.365426	0.595936	0.640229	
inpatient_beds_7_day_avg	0.335702	0.654567	0.665840	0.994717	0.989135	
total_icu_beds_7_day_avg	0.434705	0.643058	0.668117	0.927515	0.946878	
icu_beds_used_7_day_avg	0.395060	0.605813	0.598783	0.902276	0.931068	
staffed_icu_adult_patients_confirmed_covid_7_day_avg	0.290703	0.549911	0.518089	0.717839	0.692699	

Observation:

- As we can see correlation coefficient between population vs Cases and deaths is 0.45 and 0.4 which means as population increases there is significant increase in cases and deaths.
- Cases vs total_icu_beds_7_day_avg correlation coefficient is 0.64 which mean most of the people are admitting to hospital.
- Cases vs Death correlation coefficient is 0.92 which means state with higher number of cases having higher number of deaths.