Nikhil Kurmadasu(114361880)

Krishna Teja Reddy Chinnakotla(114507668)

Pranith Kumar Raparthi(113217634)

Varun Doniparti(113270819)

Lakshmi Datta Meghana Siddavatam(114777708)

In [164...
```python
# First we will import the required libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import math
import scipy.stats as st
from scipy.stats import norm,poisson,binom,geom,gamma
from sklearn.preprocessing import StandardScaler
from scipy import stats
import warnings
import seaborn as sns
import plotly.express as px
import plotly.graph_objects as go
import plotly.io as pio
pio.templates.default = "plotly_dark"
warnings.filterwarnings('ignore')
%matplotlib inline
```

In [70]:
```python
df = pd.read_csv('COVID-19_Reported_Patient_Impact_and_Hospital_Capacity_by_State_Timeseri
```

In [71]:
```python
df_x = df.loc[df['state'].isin(['KS', 'IL'])]
df_x['date']= pd.to_datetime(df_x['date'])
df_x_final=df_x.sort_values(['state', 'date'])
df_x_final.reset_index(drop=True, inplace=True)
```

In [73]:
```python
covid_data = df_x_final[['state','date','critical_staffing_shortage_today_yes','inpatient_
covid_data.head()
```

Out[73]:

	state	date	critical_staffing_shortage_today_yes	inpatient_beds	inpatient_beds_used	inpatient_beds_used_covid	
0	IL	2020-02-18	0	36.0	20.0	0.0	
1	IL	2020-02-19	0	86.0	17.0	0.0	
2	IL	2020-02-20	0	86.0	13.0	0.0	
3	IL	2020-02-21	0	86.0	14.0	0.0	
4	IL	2020-02-22	0	86.0	14.0	0.0	

```
In [76]:    df_x_IL = covid_data[covid_data['state'] == 'IL']
```

```
In [141…    df_x_KS = covid_data[covid_data['state'] == 'KS']
```

```
In [95]:    dff = (df_x_IL[(df_x_IL['date']>pd.to_datetime('2020-10-31')) & (df_x_IL['date']<pd.to_dat
```

```
In [96]:    dff
```

Out[96]:

	state	date	critical_staffing_shortage_today_yes	inpatient_beds	inpatient_beds_used	inpatient_beds_used_covid
257	IL	2020-11-01	28	31428.0	20363.0	3472.0
258	IL	2020-11-02	29	30822.0	19976.0	3654.0
259	IL	2020-11-03	29	30721.0	20699.0	3876.0
260	IL	2020-11-04	29	31847.0	21779.0	4105.0
261	IL	2020-11-05	30	31732.0	21943.0	4184.0
262	IL	2020-11-06	29	31647.0	21783.0	4367.0
263	IL	2020-11-07	29	31888.0	21140.0	4716.0
264	IL	2020-11-08	31	31873.0	20622.0	4676.0
265	IL	2020-11-09	35	30967.0	20386.0	4783.0
266	IL	2020-11-10	35	30909.0	21351.0	5170.0
267	IL	2020-11-11	35	32173.0	22815.0	5430.0
268	IL	2020-11-12	38	31707.0	22627.0	5571.0
269	IL	2020-11-13	38	31922.0	22313.0	5391.0
270	IL	2020-11-14	38	31712.0	21600.0	5621.0
271	IL	2020-11-15	39	31767.0	20948.0	5822.0
272	IL	2020-11-16	37	30572.0	20369.0	5898.0
273	IL	2020-11-17	42	30636.0	21065.0	6051.0
274	IL	2020-11-18	41	32056.0	22555.0	6182.0

	state	date	critical_staffing_shortage_today_yes	inpatient_beds	inpatient_beds_used	inpatient_beds_used_covid
275	IL	2020-11-19	39	31954.0	22633.0	6250.0
276	IL	2020-11-20	43	31741.0	22609.0	6319.0
277	IL	2020-11-21	42	31306.0	21644.0	6263.0
278	IL	2020-11-22	39	31322.0	21235.0	6261.0
279	IL	2020-11-23	42	30160.0	20390.0	6223.0
280	IL	2020-11-24	41	30857.0	21495.0	6267.0
281	IL	2020-11-25	43	31727.0	22047.0	6365.0
282	IL	2020-11-26	44	31646.0	20962.0	6135.0
283	IL	2020-11-27	44	31652.0	20548.0	6055.0
284	IL	2020-11-28	44	31571.0	20517.0	5998.0
285	IL	2020-11-29	40	31546.0	20529.0	5995.0
286	IL	2020-11-30	43	30306.0	20051.0	5908.0
287	IL	2020-12-01	45	30122.0	20946.0	5878.0
288	IL	2020-12-02	43	31374.0	22237.0	5769.0
289	IL	2020-12-03	45	31436.0	22192.0	5577.0
290	IL	2020-12-04	44	31445.0	22112.0	5563.0
291	IL	2020-12-05	42	31329.0	21591.0	5391.0
292	IL	2020-12-06	42	31341.0	21089.0	5411.0
293	IL	2020-12-07	40	30207.0	20260.0	5199.0
294	IL	2020-12-08	39	30059.0	21065.0	5293.0
295	IL	2020-12-09	41	32287.0	22916.0	5363.0
296	IL	2020-12-10	35	32248.0	22760.0	5329.0
297	IL	2020-12-11	37	32282.0	22572.0	5247.0

	state	date	critical_staffing_shortage_today_yes	inpatient_beds	inpatient_beds_used	inpatient_beds_used_covid
298	IL	2020-12-12	37	32216.0	21970.0	5212.0
299	IL	2020-12-13	38	31730.0	21598.0	5136.0
300	IL	2020-12-14	40	29953.0	20442.0	5051.0
301	IL	2020-12-15	39	30110.0	21377.0	4988.0
302	IL	2020-12-16	36	31953.0	23000.0	4972.0
303	IL	2020-12-17	36	32066.0	23013.0	4852.0
304	IL	2020-12-18	36	32063.0	22957.0	4794.0
305	IL	2020-12-19	36	32032.0	22215.0	4631.0
306	IL	2020-12-20	34	32200.0	21482.0	4659.0
307	IL	2020-12-21	36	30201.0	19933.0	4587.0
308	IL	2020-12-22	34	30133.0	20906.0	4635.0
309	IL	2020-12-23	33	32398.0	22496.0	4660.0
310	IL	2020-12-24	32	32397.0	21356.0	4452.0
311	IL	2020-12-25	33	32345.0	19384.0	4176.0
312	IL	2020-12-26	33	32371.0	19359.0	4187.0
313	IL	2020-12-27	35	32378.0	20091.0	4356.0

```
In [98]: dfc = (df_case_IL[(df_case_IL['submission_date']>pd.to_datetime('2020-10-31')) & (df_case_
```

```
In [99]: dfc
```

Out[99]:

	state	submission_date	tot_cases	tot_death	per_day_cases	per_day_deaths
284	IL	2020-11-01	423539	10075	6980	35
285	IL	2020-11-02	429761	10093	6222	18
286	IL	2020-11-03	436277	10161	6516	68
287	IL	2020-11-04	443815	10216	7538	55
288	IL	2020-11-05	453750	10313	9935	97
289	IL	2020-11-06	465540	10397	11790	84

	state	submission_date	tot_cases	tot_death	per_day_cases	per_day_deaths
290	IL	2020-11-07	477978	10488	12438	91
291	IL	2020-11-08	487987	10538	10009	50
292	IL	2020-11-09	498560	10563	10573	25
293	IL	2020-11-10	511183	10645	12623	82
294	IL	2020-11-11	523840	10798	12657	153
295	IL	2020-11-12	536542	10846	12702	48
296	IL	2020-11-13	551957	10891	15415	45
297	IL	2020-11-14	562985	11088	11028	197
298	IL	2020-11-15	573616	11162	10631	74
299	IL	2020-11-16	585248	11204	11632	42
300	IL	2020-11-17	597849	11317	12601	113
301	IL	2020-11-18	606771	11468	8922	151
302	IL	2020-11-19	621383	11648	14612	180
303	IL	2020-11-20	634395	11795	13012	147
304	IL	2020-11-21	646286	11952	11891	157
305	IL	2020-11-22	656298	12051	10012	99
306	IL	2020-11-23	664620	12112	8322	61
307	IL	2020-11-24	674089	12262	9469	150
308	IL	2020-11-25	685467	12440	11378	178
309	IL	2020-11-26	697489	12596	12022	156
310	IL	2020-11-27	705063	12686	7574	90
311	IL	2020-11-28	712936	12838	7873	152
312	IL	2020-11-29	720114	12882	7178	44
313	IL	2020-11-30	726304	12985	6190	103
314	IL	2020-12-01	738846	13131	12542	146
315	IL	2020-12-02	748603	13397	9757	266
316	IL	2020-12-03	759562	13625	10959	228
317	IL	2020-12-04	770088	13782	10526	157
318	IL	2020-12-05	779975	14017	9887	235
319	IL	2020-12-06	787573	14116	7598	99
320	IL	2020-12-07	796264	14216	8691	100
321	IL	2020-12-08	804174	14384	7910	168
322	IL	2020-12-09	812430	14613	8256	229
323	IL	2020-12-10	823531	14845	11101	232
324	IL	2020-12-11	832951	15067	9420	222
325	IL	2020-12-12	841688	15232	8737	165

	state	submission_date	tot_cases	tot_death	per_day_cases	per_day_deaths
326	IL	2020-12-13	848904	15339	7216	107
327	IL	2020-12-14	856118	15455	7214	116
328	IL	2020-12-15	863477	15587	7359	132
329	IL	2020-12-16	870600	15777	7123	190
330	IL	2020-12-17	879428	15985	8828	208
331	IL	2020-12-18	886805	16206	7377	221
332	IL	2020-12-19	894367	16326	7562	120
333	IL	2020-12-20	900370	16407	6003	81
334	IL	2020-12-21	905069	16527	4699	120
335	IL	2020-12-22	911308	16671	6239	144
336	IL	2020-12-23	918070	16842	6762	171
337	IL	2020-12-24	925107	16960	7037	118
338	IL	2020-12-25	930849	17154	5742	194
339	IL	2020-12-26	934142	17224	3293	70
340	IL	2020-12-27	937909	17336	3767	112

In [133…
```python
# helper function to compute correlation
def computeCorrelation(x,y):
    x_mean = np.mean(x)
    y_mean = np.mean(y)
    print(x_mean)
    print(y_mean)
    xy = 0
    xi_x = 0
    yi_y = 0

    for i in range(len(x)):
        xy += ((x[i]-x_mean) * (y[i] - y_mean))
        xi_x += np.square(x[i] - x_mean)
        yi_y += np.square(y[i] - y_mean)

    return xy/((np.sqrt(xi_x * yi_y)))
```

In [121…
```python
x = dff['critical_staffing_shortage_today_yes']
y = dfc['per_day_cases']
correlation = computeCorrelation(x,y)
#validate it with corr

print("Our Function Correlation: %1.3f " % (correlation))
```

```
37.49122807017544
9146.491228070176
Our Function Correlation: 0.843
```

In [131…
```python
x.iloc[0]
```

Out[131…
```
3472.0
```

```
In [134... x = np.array(dff['inpatient_beds_used_covid'])
         y = np.array(dfc['per_day_cases'])
         correlation = computeCorrelation(x,y)
         #validate it with corr

         print("Our Function Correlation: %1.3f " % (correlation))

         5234.666666666667
         9146.491228070176
         Our Function Correlation: 0.529
```

```
In [135... df_prewave= (df_case_IL[(df_case_KS['submission_date']>pd.to_datetime('2021-05-24')) & (di
```

```
In [136... df_postwave= (df_case_IL[(df_case_KS['submission_date']>pd.to_datetime('2021-07-30')) & (c
```

```
In [137... def permutation_test(X, Y, n, threshold):
             T_obs = abs(np.mean(X) - np.mean(Y))
             print(T_obs, np.mean(X), np.mean(Y))
             xy = np.append(X,Y)
         #     xy.info()
             p_value = 0.0
             for i in range(n):
                 permutation = np.random.permutation(xy)
                 X1 = permutation[:len(X)]
                 Y1 = permutation[len(X):]
                 Ti = abs(np.mean(X1) - np.mean(Y1))
                 if(Ti > T_obs):
                     p_value += 1.0
         #             print(p_value, T_obs, Ti)
             p_value = p_value/n
             print("The p-value is: ", p_value)
             if(p_value <= threshold):
                 print("==> Reject the Null Hypothesis")
             else:
                 print("==> Accept the Null Hypothesis")
             return
```

```
In [140... # PERMUTATION TEST: Hypotheses and Results
         print("-----------------------------------------------------------------------------------
         print("H0: For Oct'21 to Dec'21, the distribution of #deaths due to COVID is similar in II
         permutation_test(np.array(df_postwave['per_day_deaths']),np.array(df_prewave['per_day_deat
         print(np.mean(np.array(df_postwave['per_day_deaths'])))
         print(np.mean(np.array(df_prewave['per_day_deaths'])))
         print("-----------------------------------------------------------------------------------
         print("H0: For Oct'21 to Dec'21, the distribution of #cases due to COVID is similar in IL
         permutation_test(np.array(df_postwave['per_day_cases']),np.array(df_prewave['per_day_cases
         print("-----------------------------------------------------------------------------------

         -----------------------------------------------------------------------------------
         -----------
         H0: For Oct'21 to Dec'21, the distribution of #deaths due to COVID is similar in IL and KS
         2.0333333333333314 17.866666666666667 19.9
         The p-value is:  0.506
         ==> Accept the Null Hypothesis
         17.866666666666667
         19.9
         -----------------------------------------------------------------------------------
         -----------
         H0: For Oct'21 to Dec'21, the distribution of #cases due to COVID is similar in IL and KS
         2721.7666666666664 3170.133333333333 448.3666666666667
         The p-value is:  0.0
```

```
       ==> Reject the Null Hypothesis
       --------------------------------------------------------------------------------
       -----------
```

In [ ]:

```
       ==> Reject the Null Hypothesis
       --------------------------------------------------------------------------------
       -----------
```