PREDICTIVE ANALYSIS OF THE COVID DELTA VARIANT

1. Are vaccines effective against delta variant?
2. Are we all heading towards another pandemic zone?

There are the questions of the hour. With everyday increase in delta variant cases worldwide, there is panic in the population.

The delta variant (B.1.617.2) of the corona virus (SARS-CoV-2, is contributing to surge in India and now its widespread has been identified and recognized across the world. According to the research analysis of the yale university, the delta variant is 50% more contagious than the alpha variant originated in UK. Data has proved that even in current times, the most of hospitalizations is because of the unvaccinations. According to CDC, fully vaccinated can still spread the delta variant. The research analysis also proves that J&J vaccines are highly effective against the delta variant.

```
In [1]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import plotly.express as px
         import seaborn as sns
         from scipy.stats import ttest_ind
         import statsmodels.api as ss
         import warnings
         warnings.filterwarnings("ignore", category=FutureWarning)
```

```
In [2]:  var = pd.read_csv('covid-variants.csv')
         vacc= pd.read_csv('country_vaccinations.csv')
         latest = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_cc
```

```
In [3]:  #renaming col name in varianst dataset
         var =var.rename(columns={'location':'country'})
         latest =latest.rename(columns={'Country_Region':'country'})
         latest =latest.rename(columns={'US':'United States'})
```

```
In [4]:  # MERGE 1 varinats and vaccines    # LEFT JOIN

         cases = pd.merge(var,vacc, on= "country",how='left')
         cases.dropna()
         # changes the scientific notation to .2 decimal values
         pd.options.display.float_format = '{:.2f}'.format
```

```
In [5]:  deaths=latest[['country','Deaths']].groupby(['country']).sum()
         deaths=deaths.sort_values(ascending=False, by='Deaths').head(20)
         deaths.rename(index={'US':'United States'},inplace =True)
         deaths
```

Out[5]:

country	Deaths
United States	616493
Brazil	561762
India	427371
Mexico	243733
Peru	196818
Russia	160574
United Kingdom	130482
Italy	128187
Colombia	122087
France	112347
Argentina	107213
Indonesia	104010
Iran	93086
Germany	91785
Spain	82006
Poland	75281
South Africa	74352
Ukraine	55792
Turkey	51976
Chile	35880

According to this report, United states ,Brazil, India , Mexico and Peru has recorded highest deaths due to covid.

In [6]:
```python
#MERGE 2
latrep = pd.merge(cases,deaths, on= "country",how='left')

latrep.dropna()
# changes the scientific notation to .2 decimal values
pd.options.display.float_format = '{:.2f}'.format

latrep
```

Out[6]:

	country	date_x	variant	num_sequences	perc_sequences	num_sequences_total	iso_code	date_y	total_vaccinations	people_vacc
0	Angola	2020-12-21	B.1.160	0.00	0.00	93	AGO	2021-03-01	0.00	
1	Angola	2020-12-21	B.1.160	0.00	0.00	93	AGO	2021-03-02	nan	
2	Angola	2020-12-21	B.1.160	0.00	0.00	93	AGO	2021-03-03	nan	
3	Angola	2020-12-21	B.1.160	0.00	0.00	93	AGO	2021-03-04	nan	
4	Angola	2020-12-21	B.1.160	0.00	0.00	93	AGO	2021-03-05	nan	
...	...	...	...	...	...	...	...	...	...	
6972767	Zimbabwe	2021-02-08	non_who	0.00	0.00	37	ZWE	2021-07-31	2413509.00	1645
6972768	Zimbabwe	2021-02-08	non_who	0.00	0.00	37	ZWE	2021-08-01	2433341.00	1654
6972769	Zimbabwe	2021-02-08	non_who	0.00	0.00	37	ZWE	2021-08-02	2473590.00	1674
6972770	Zimbabwe	2021-02-08	non_who	0.00	0.00	37	ZWE	2021-08-03	2540555.00	1707
6972771	Zimbabwe	2021-02-08	non_who	0.00	0.00	37	ZWE	2021-08-04	2604265.00	1740

6972772 rows × 21 columns

In [7]:
```python
varinfo=cases[['variant','num_sequences']].groupby(['variant']).sum()
varinfo = varinfo.sort_values(ascending=False, by='num_sequences')
varinfo
```

Out[7]:

variant	num_sequences
Alpha	226034126.00
non_who	177008884.00
others	114766394.00
Delta	69607923.00
B.1.177	38013279.00
Epsilon	13229537.00
Gamma	13021244.00
Iota	8729201.00
B.1.160	6758627.00
Beta	6572828.00
B.1.258	5954672.00
B.1.1.519	4762561.00
B.1.221	3284763.00
Eta	1548546.00
S:677H.Robin1	1330211.00
Kappa	1197171.00
S:677P.Pelican	1009929.00
Lambda	753353.00
B.1.621	389389.00
B.1.1.277	262340.00
B.1.367	198638.00

B.1.620	173605.00
B.1.1.302	104476.00

In [8]:
```python
delta=latrep[latrep['variant']=='Delta'].groupby(['country','variant']).sum()
delta =delta.sort_values(ascending=False, by='num_sequences').head(10)
delta=delta.reset_index()
delta=delta[['country','variant','num_sequences','num_sequences_total','total_vaccinations','people_fully_vaccina
delta
```

Out[8]:

	country	variant	num_sequences	num_sequences_total	total_vaccinations	people_fully_vaccinated	Deaths
0	United Kingdom	Delta	42138790.00	136900660	314317394028.00	104709971553.00	1011887910.00
1	United States	Delta	7867982.00	147972243	1379373457210.00	600503986198.00	4800014498.00
2	Denmark	Delta	3116883.00	30192162	15578339733.00	6612218646.00	0.00
3	India	Delta	2990799.00	7753991	988950434240.00	189154884864.00	2776202016.00
4	Germany	Delta	1474070.00	31715268	247436307028.00	90888946234.00	689672490.00
5	Italy	Delta	1054500.00	8761230	161287848017.00	59589396104.00	882182934.00
6	Sweden	Delta	1012542.00	17646780	5150712256.00	1830978816.00	0.00
7	Netherlands	Delta	913716.00	9196728	7268574060.00	2580247624.00	0.00
8	Spain	Delta	911640.00	9063789	98706961980.00	39321692025.00	576420174.00
9	France	Delta	853281.00	10767562	169374781437.00	60589412921.00	769689297.00

Is there any ind of association between the delta varaint number seuqnces and total vaccinataions in top 10 listed countries?

In [9]:
```python
delta['num_sequences'].corr(delta['total_vaccinations'])
```

Out[9]: 0.12600079283189067

Is there any evidence that shows that delta varaint cases is decreased with total vaccinations and deaths?

In [10]:
```python
y=delta['num_sequences']
x=delta['total_vaccinations']
z=delta['Deaths']
model=ss.OLS(x,y+z).fit()
prediction=model.predict(x)
model.summary()
```

```
C:\ProgramData\Anaconda3\lib\site-packages\scipy\stats\stats.py:1603: UserWarning: kurtosistest only valid for n>
=20 ... continuing anyway, n=10
  warnings.warn("kurtosistest only valid for n>=20 ... continuing "
```
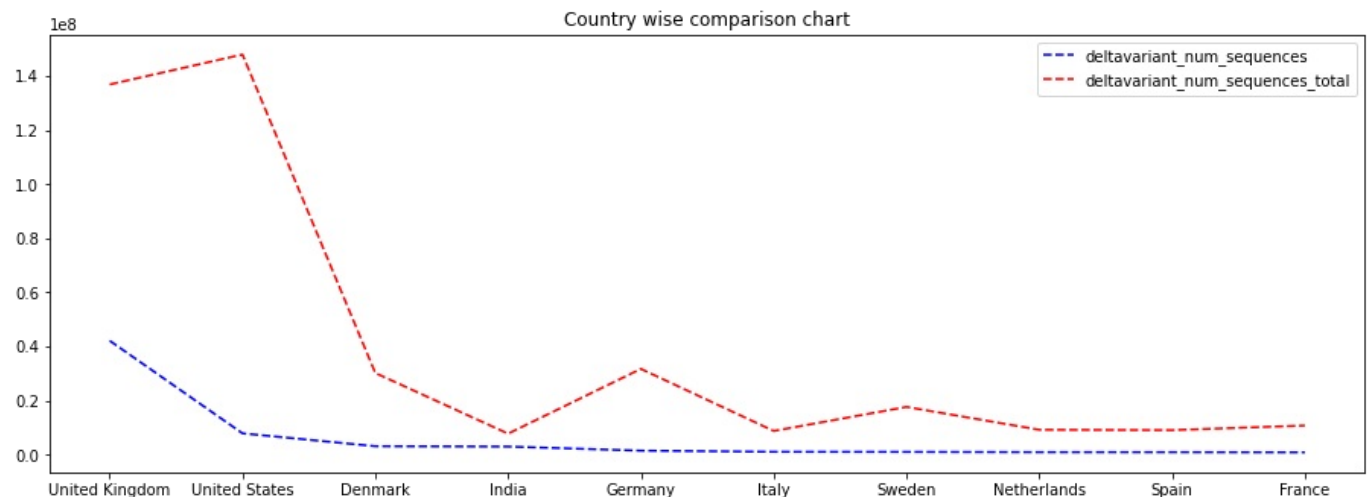
Out[10]:

OLS Regression Results

Dep. Variable:	total_vaccinations	**R-squared (uncentered):**	0.984
Model:	OLS	**Adj. R-squared (uncentered):**	0.982
Method:	Least Squares	**F-statistic:**	549.5
Date:	Mon, 09 Aug 2021	**Prob (F-statistic):**	2.23e-09
Time:	02:20:26	**Log-Likelihood:**	-264.01
No. Observations:	10	**AIC:**	530.0
Df Residuals:	9	**BIC:**	530.3
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	299.1808	12.763	23.441	0.000	270.309	328.053

Omnibus:	4.331	**Durbin-Watson:**	1.635
Prob(Omnibus):	0.115	**Jarque-Bera (JB):**	1.468
Skew:	0.902	**Prob(JB):**	0.480
Kurtosis:	3.520	**Cond. No.**	1.00

Notes:

[1] R² is computed without centering (uncentered) since the model does not contain a constant.

[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

DATA VISUALIZATION

In [11]:
```python
fig=plt.figure()
ax=fig.add_axes([0,0,2,1])
plt.plot(delta['country'],delta['num_sequences'],'b--',label='deltavariant_num_sequences')
plt.plot(delta['country'],delta['num_sequences_total'],'r--',label='deltavariant_num_sequences_total')
#plt.plot(delta['country'],delta['total_vaccinations'],'o--',label='total_vaccinations')
#plt.plot(delta['country'],delta['people_fully_vaccinated'],'g--',label='people_fully_vaccinated')
#plt.plot(delta['country'],delta['Deaths'],'y--',label='Deaths')

plt.legend(loc='upper right')
plt.title ("Country wise comparison chart")
plt.show()
```


In [12]:
```python
fig=plt.figure()
ax=fig.add_axes([0,0,2,1])
plt.plot(delta['country'],delta['num_sequences'],'b--',label='deltavariant_num_sequences')
plt.plot(delta['country'],delta['num_sequences_total'],'r--',label='deltavariant_num_sequences_total')
plt.plot(delta['country'],delta['total_vaccinations'],'o--',label='total_vaccinations')
plt.plot(delta['country'],delta['people_fully_vaccinated'],'g--',label='people_fully_vaccinated')
plt.plot(delta['country'],delta['Deaths'],'y--',label='Deaths')

plt.legend(loc='upper right')
plt.title ("Country wise comparison chart")
plt.show()
```
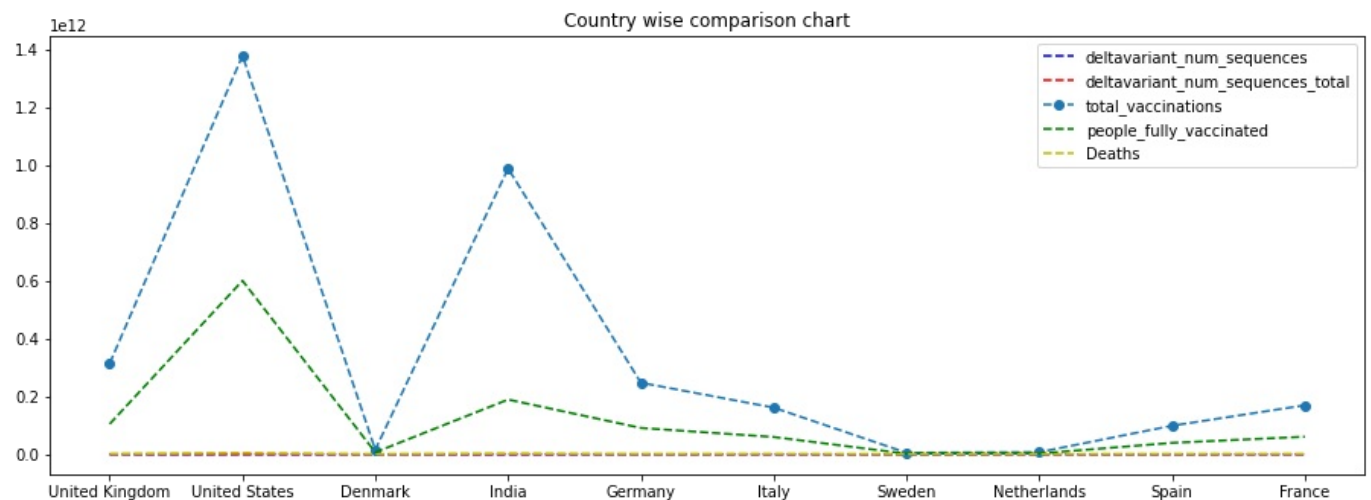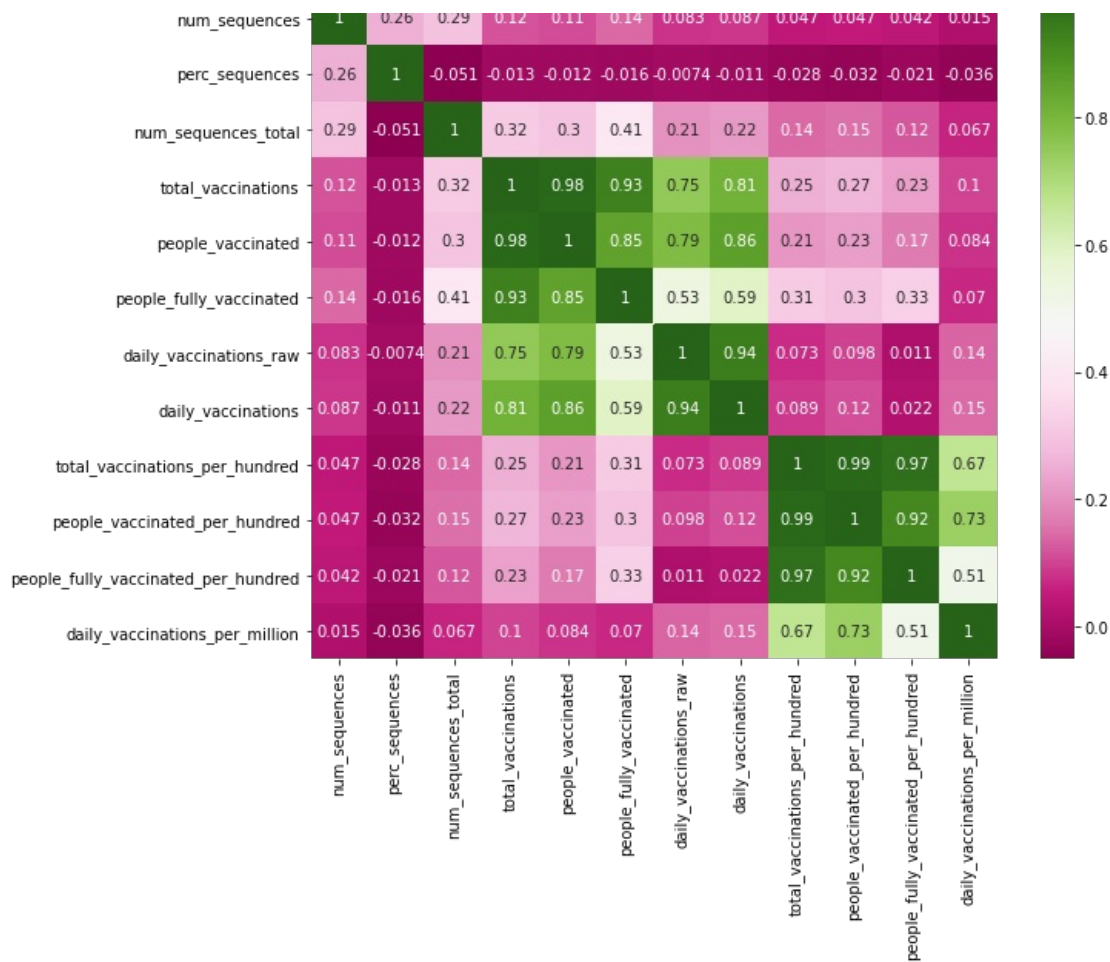

In [13]:
```python
# correlation
corr = cases.corr()
plt.figure(figsize=(10,8))
sns.heatmap(corr,cmap ="PiYG",annot=True)
```

Out[13]: <AxesSubplot:>

	num_sequences	perc_sequences	num_sequences_total	total_vaccinations	people_vaccinated	people_fully_vaccinated	daily_vaccinations_raw	daily_vaccinations	total_vaccinations_per_hundred	people_vaccinated_per_hundred	people_fully_vaccinated_per_hundred	daily_vaccinations_per_million
num_sequences	1	0.26	0.29	0.12	0.11	0.14	0.083	0.087	0.047	0.047	0.042	0.015
perc_sequences	0.26	1	-0.051	-0.013	-0.012	-0.016	-0.0074	-0.011	-0.028	-0.032	-0.021	-0.036
num_sequences_total	0.29	-0.051	1	0.32	0.3	0.41	0.21	0.22	0.14	0.15	0.12	0.067
total_vaccinations	0.12	-0.013	0.32	1	0.98	0.93	0.75	0.81	0.25	0.27	0.23	0.1
people_vaccinated	0.11	-0.012	0.3	0.98	1	0.85	0.79	0.86	0.21	0.23	0.17	0.084
people_fully_vaccinated	0.14	-0.016	0.41	0.93	0.85	1	0.53	0.59	0.31	0.3	0.33	0.07
daily_vaccinations_raw	0.083	-0.0074	0.21	0.75	0.79	0.53	1	0.94	0.073	0.098	0.011	0.14
daily_vaccinations	0.087	-0.011	0.22	0.81	0.86	0.59	0.94	1	0.089	0.12	0.022	0.15
total_vaccinations_per_hundred	0.047	-0.028	0.14	0.25	0.21	0.31	0.073	0.089	1	0.99	0.97	0.67
people_vaccinated_per_hundred	0.047	-0.032	0.15	0.27	0.23	0.3	0.098	0.12	0.99	1	0.92	0.73
people_fully_vaccinated_per_hundred	0.042	-0.021	0.12	0.23	0.17	0.33	0.011	0.022	0.97	0.92	1	0.51
daily_vaccinations_per_million	0.015	-0.036	0.067	0.1	0.084	0.07	0.14	0.15	0.67	0.73	0.51	1

In [14]:
```python
delta=latrep[latrep['variant']=='Delta'].groupby(['country','variant']).sum()
delta =delta.sort_values(ascending=False, by='num_sequences_total')
delta=delta.reset_index()
```

In [15]:
```python
fig = px.choropleth(delta, locations="country",
                    color="num_sequences_total",
                    locationmode='country names',
                    color_continuous_scale=px.colors.sequential.Plasma,
                    title="Delta Cases in the world in last two weeks",
                    height=600)
fig.show()
```

Delta Cases in the world in last two weeks