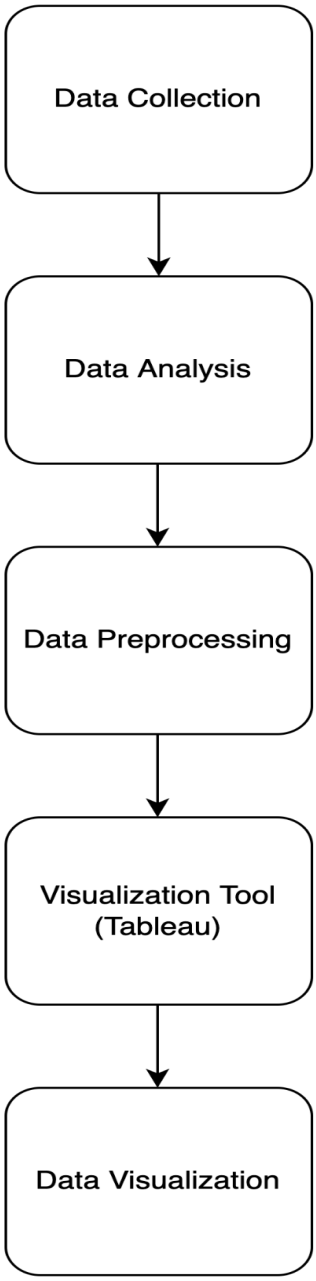
Team Members: Kishan Patel (922243969), Manali Seth (922053012)

Visualization Project - Project Implementation

1) System Architecture



Tools and Technologies

Programming Language: Python
Libraries: Pandas, Numpy
Tool: Visual Studio Code, Tableau

Exploratory visualization dashboards include analyzing data via line chart, pie chart, map etc. We can view confirmed and death cases country-wise globally and state-wise for the USA.

2) Data set Description

COVID19 dataset is collected by Johns Hopkins Center for Systems Science and Engineering (CSSE) and John Hopkins Center for Civic Impact (CCI). It includes datasets based on time-series, US and non-US locations for vaccinations, cases and deaths. The data is collected since January 2020 and is updated on a regular basis, making this a dynamic and scalable dataset.

Field Description:

- Province_State - The name of the State within the USA.
- Country_Region - The name of the Country (USA).
- Last_Update - The most recent date the file was pushed. MM/DD/YYYY HH:mm:ss (24 hour format, in UTC).
- Lat - Latitude.
- Long_ - Longitude.
- Confirmed - Aggregated case count. Counts include confirmed and probable (where reported).
- Deaths - Aggregated death toll. Counts include confirmed and probable (where reported).
- Recovered - Aggregated Recovered case count.
- Active - Aggregated confirmed cases that have not been resolved (Active cases = total cases - total recovered - total deaths).
- FIPS - USA only. Federal Information Processing Standards code that uniquely identifies counties within the USA.
- Admin2: County name. USA only.
- Incident_Rate - cases per 100,000 persons.
- Total_Test_Results - Total number of people who have been tested.
- People_Hospitalized - Total number of people hospitalized.
- Case_Fatality_Ratio - Number recorded deaths * 100/ Number confirmed cases.
- UID - Unique Identifier for each row entry.
- ISO3 - Officially assigned country code identifiers.
- Testing_Rate - Total test results per 100,000 persons. The "total test results" are equal to "Total test results (Positive + Negative)"
- Hospitalization_Rate - USA Hospitalization Rate (%): = Total number hospitalized / Number cases.

- **Confirmed and Death Cases:**

Source:
https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

Description:

time_series_covid19_confirmed_global

```
RangeIndex: 289 entries, 0 to 288
Columns: 1056 entries, Province/State to 12/8/22
dtypes: float64(2), int64(1052), object(2)
```

time_series_covid19_deaths_global

```
RangeIndex: 289 entries, 0 to 288
Columns: 1056 entries, Province/State to 12/8/22
dtypes: float64(2), int64(1052), object(2)
```

time_series_covid19_confirmed_US

```
RangeIndex: 3342 entries, 0 to 3341
Columns: 1063 entries, UID to 12/8/22
dtypes: float64(3), int64(1054), object(6)
```

time_series_covid19_deaths_US

```
RangeIndex: 3342 entries, 0 to 3341
Columns: 1064 entries, UID to 12/8/22
dtypes: float64(3), int64(1055), object(6)
```

- **Vaccination Data:**

Source:
https://github.com/govex/COVID-19/tree/master/data_tables/vaccine_data

Description:

time_series_covid19_vaccine_global

```
RangeIndex: 126517 entries, 0 to 126516
Data columns (total 6 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Date                    126517 non-null  object
 1   UID                     125807 non-null  float64
 2   Province_State          0 non-null       float64
 3   Country_Region          126517 non-null  object
 4   Doses_admin             126509 non-null  float64
 5   People_at_least_one_dose  126517 non-null  int64
dtypes: float64(3), int64(1), object(2)
```

time_series_covid19_vaccine_US

```
RangeIndex: 41459 entries, 0 to 41458
Data columns (total 8 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   Date                    41459 non-null  object
 1   UID                     41459 non-null  int64
 2   Province_State          40735 non-null  object
 3   Country_Region          41459 non-null  object
 4   Doses_admin             41459 non-null  int64
 5   People_at_least_one_dose  41459 non-null  int64
 6   People_fully_vaccinated   41459 non-null  int64
 7   Total_additional_doses    41459 non-null  int64
dtypes: int64(5), object(3)
```

- **COVID 19 data**

1) **Source:**
https://raw.githubusercontent.com/CSSEGISandData/COVID-19/web-data/data/cases_country.csv

Description:

```
RangeIndex: 201 entries, 0 to 200
Data columns (total 16 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Country_Region       201 non-null    object
 1   Last_Update          201 non-null    object
 2   Lat                  199 non-null    float64
 3   Long_                199 non-null    float64
 4   Confirmed            201 non-null    int64
 5   Deaths               201 non-null    int64
 6   Recovered            0 non-null      float64
 7   Active               0 non-null      float64
 8   Incident_Rate        196 non-null    float64
 9   People_Tested        0 non-null      float64
 10  People_Hospitalized  0 non-null      float64
 11  Mortality_Rate       201 non-null    float64
 12  UID                  201 non-null    int64
 13  ISO3                 197 non-null    object
 14  Cases_28_Days        201 non-null    int64
 15  Deaths_28_Days       201 non-null    int64
dtypes: float64(8), int64(5), object(3)
```

2) **Source:**
https://raw.githubusercontent.com/CSSEGISandData/COVID-19/web-data/data/cases_time.csv

Description:

```
RangeIndex: 272468 entries, 0 to 272467
Data columns (total 17 columns):
 #   Column               Non-Null Count   Dtype
---  ------               --------------   -----
 0   Country_Region       272468 non-null  object
 1   Last_Update          272468 non-null  datetime64[ns]
 2   Confirmed            272468 non-null  int64
 3   Deaths               272468 non-null  int64
 4   Recovered            0 non-null       float64
 5   Active               0 non-null       float64
 6   Delta_Confirmed      272185 non-null  float64
 7   Delta_Recovered      0 non-null       float64
 8   Incident_Rate        264052 non-null  float64
 9   People_Tested        0 non-null       float64
 10  People_Hospitalized  0 non-null       float64
 11  Province_State       61016 non-null   object
 12  FIPS                 61016 non-null   float64
 13  UID                  272468 non-null  int64
 14  iso3                 268260 non-null  object
 15  Report_Date_String   272468 non-null  object
 16  Delta_Deaths         271972 non-null  float64
dtypes: datetime64[ns](1), float64(9), int64(3), object(4)
```

3) System Description

Countrywise Confirmed Cases and Deaths

A world map shows how different countries worldwide are affected by COVID19. Upon hovering over any region on the world map, the tooltip shows the country name, number of confirmed cases and number of deaths. On the right top, there is a scale showing the range of confirmed cases and deaths. There is a filter option provided to select any country and that will be highlighted on the world map. On the bottom, there are two line charts showing the number of confirmed cases and deaths for the selected country. Selecting a line on line chart shows country name, month and number of cases/deaths.

USA Statewise Confirmed Cases and Deaths

The USA map shows how different states are affected by COVID19. Upon hovering over any region on the USA map, the tooltip shows the state, number of confirmed cases and number of deaths. On the right top, there is a scale showing the range of confirmed cases and deaths. There is a filter option provided to select any country and that will be highlighted on the USA map. On the bottom, there are two line charts showing the number of confirmed cases and deaths for the selected country. Selecting a line on line chart shows state, month and number of cases/deaths.

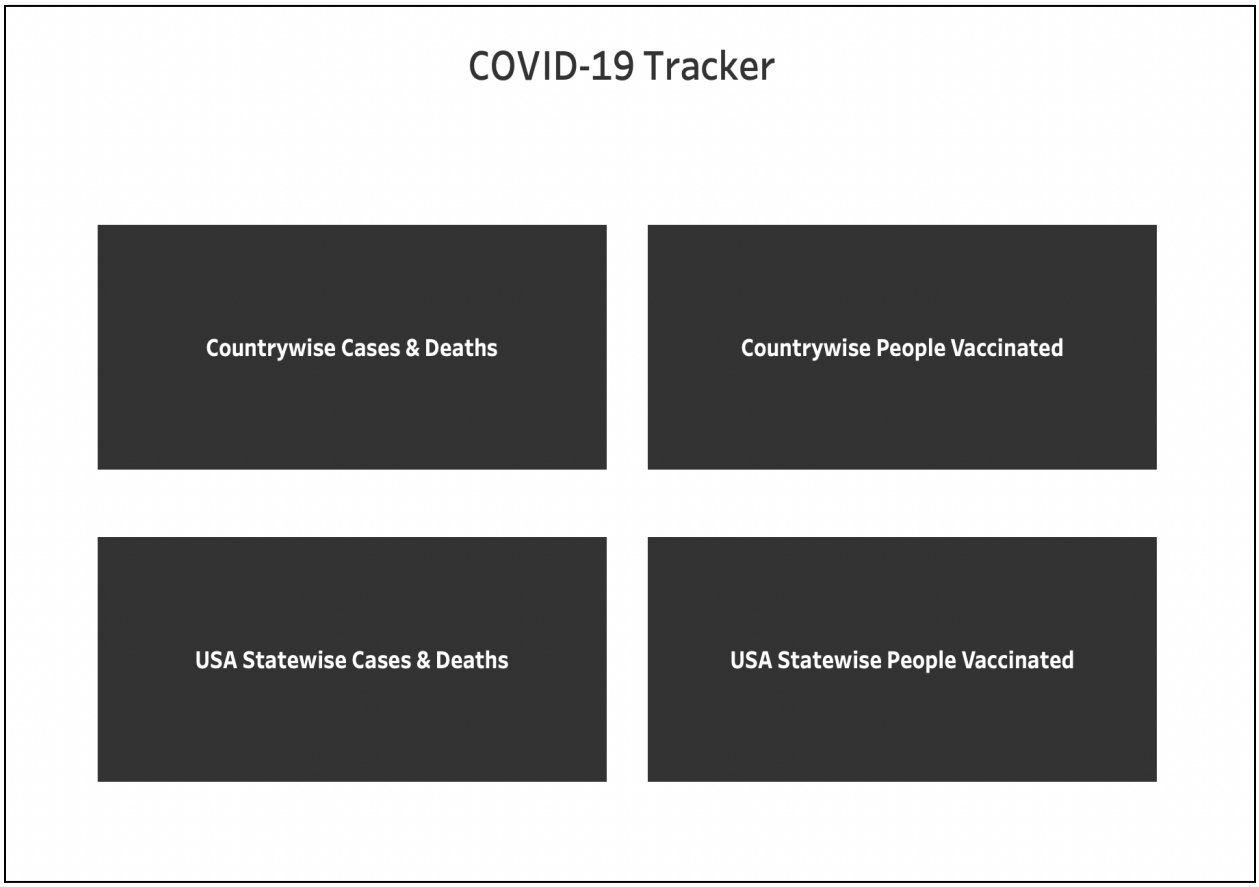
Countrywise People Vaccinated

A world map shows the number of people with at least one dose of vaccination taken. Upon hovering over any region on the world map, the tooltip shows the country name and number of people with at least one dose. On the right top, there is a dropdown to select any country and get its vaccination information. Below the world map is a progression bar showing variations in the number of vaccination doses during different months of the year starting December 2020. Along with the map, there is a pie chart in the bottom right displaying the percentage of population with at least one dose and the percentage of population with no doses. At the bottom, a table shows values for total population, number of people with at least one dose and the number of people with no doses for the selected country.

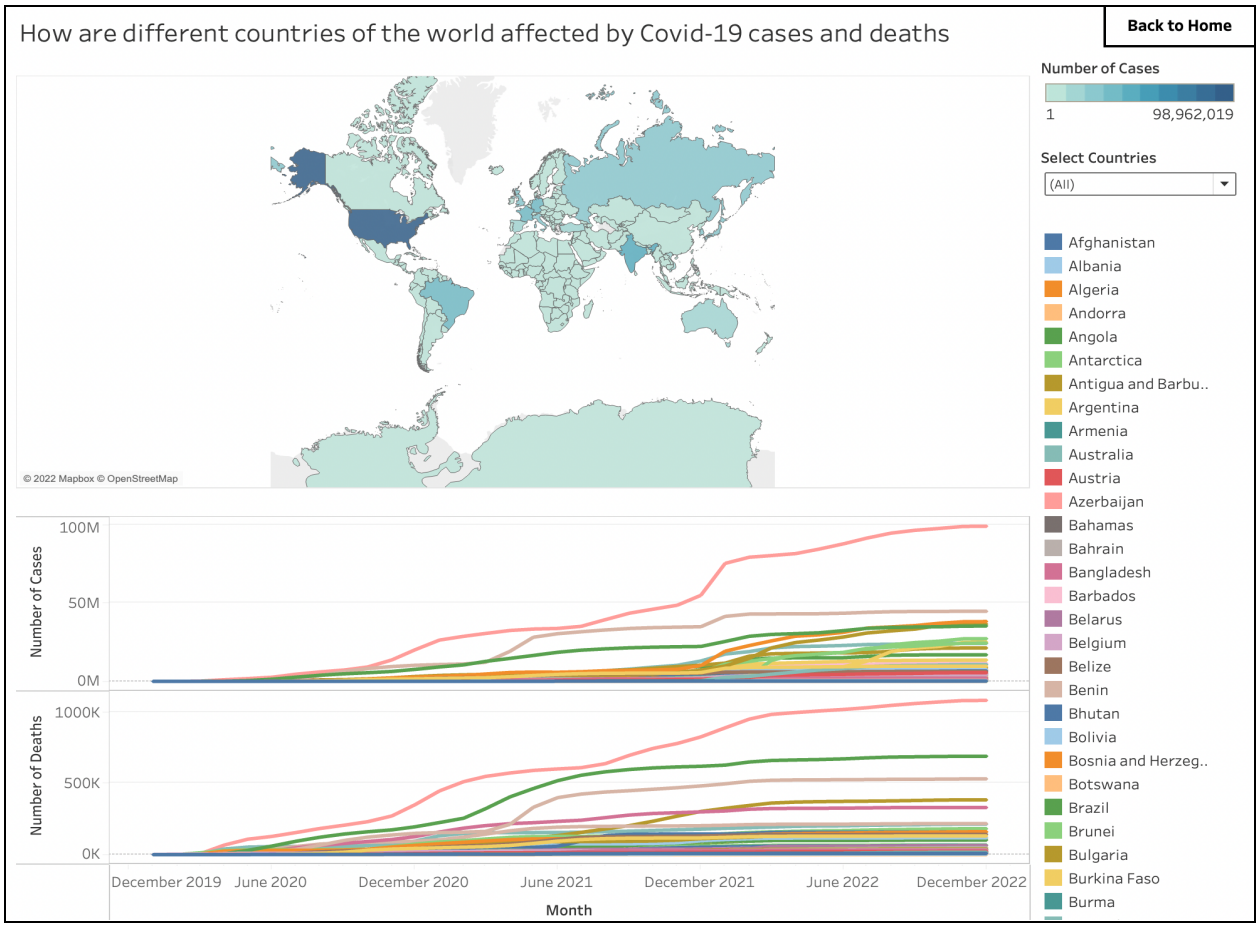
USA Statewise People Vaccinated

The USA map shows the number of people with at least one dose of vaccination taken. Upon hovering over any region on the USA map, the tooltip shows the state, number of people with at least one dose and number of people fully vaccinated. On the right top, there is a dropdown to select any USA state and get its vaccination related information. Below the USA map is a progression bar showing variations in the number of vaccination doses during different months of the year starting December 2020. Along with the map, there is a pie chart in the bottom right displaying the percentage of population not vaccinated, percentage of population fully vaccinated, percent of population partially vaccinated. At the bottom, a table shows values for total population, number of people fully vaccinated, number of people partially vaccinated and number of people not vaccinated for the selected state.

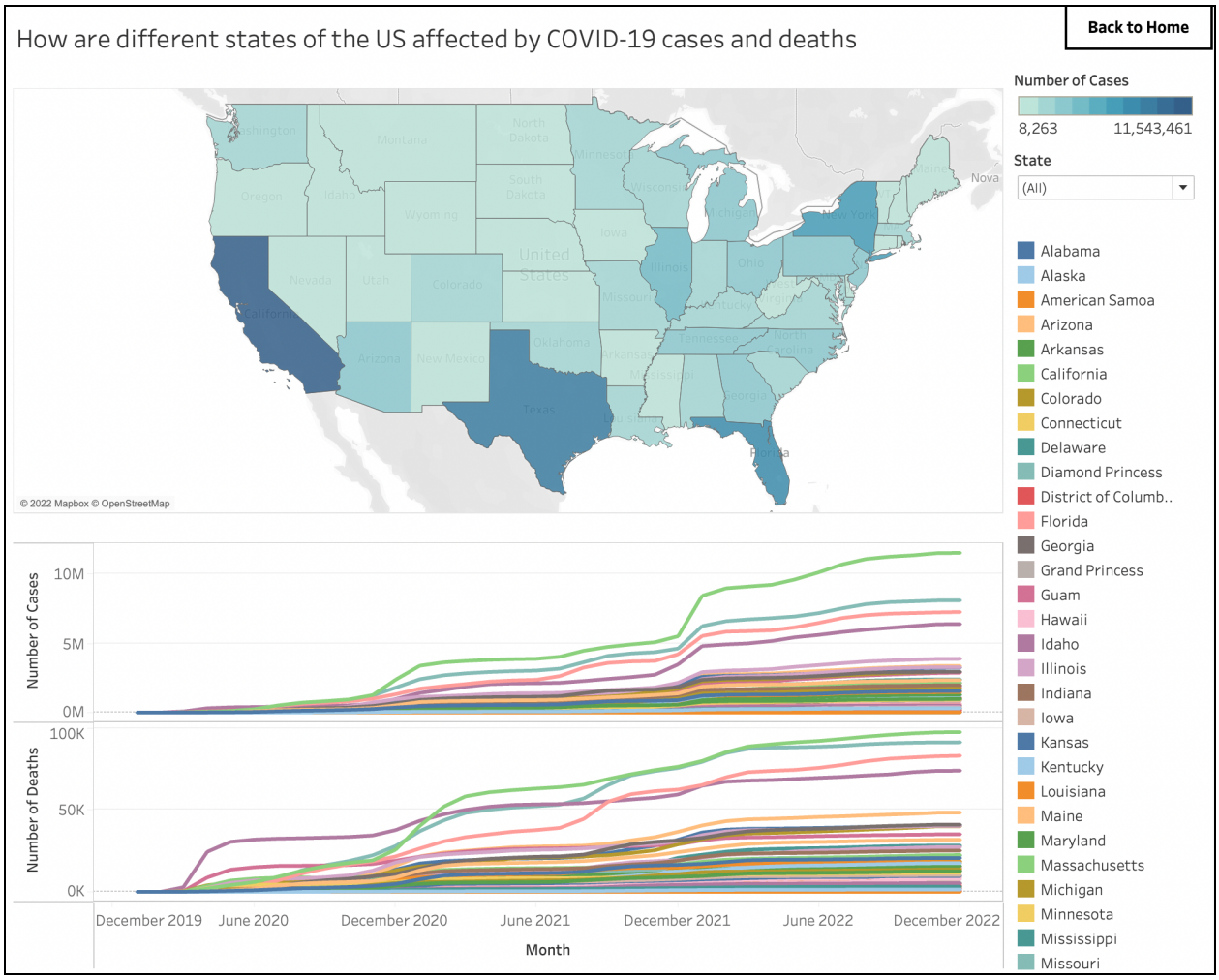
4) Screenshots



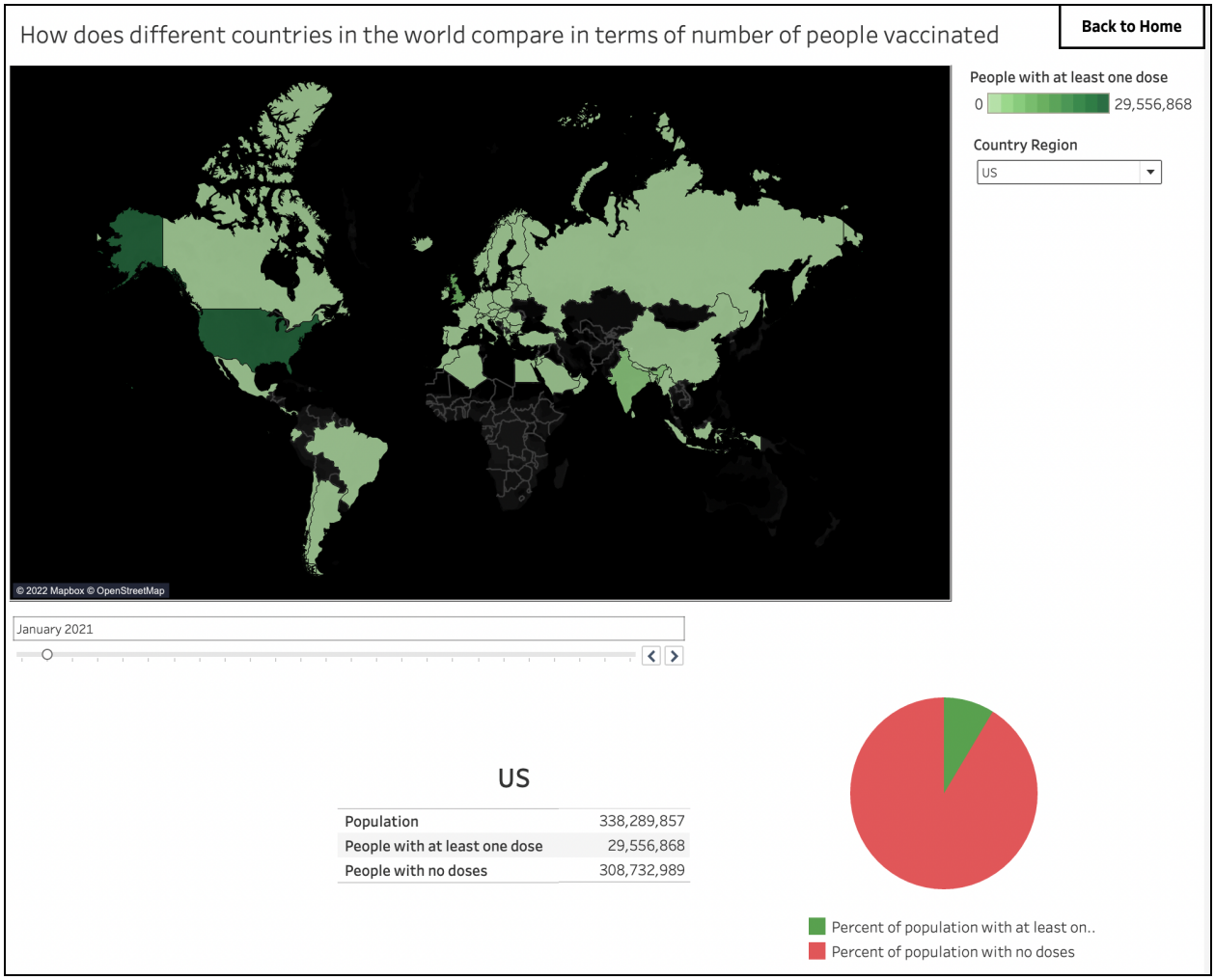
Home Page



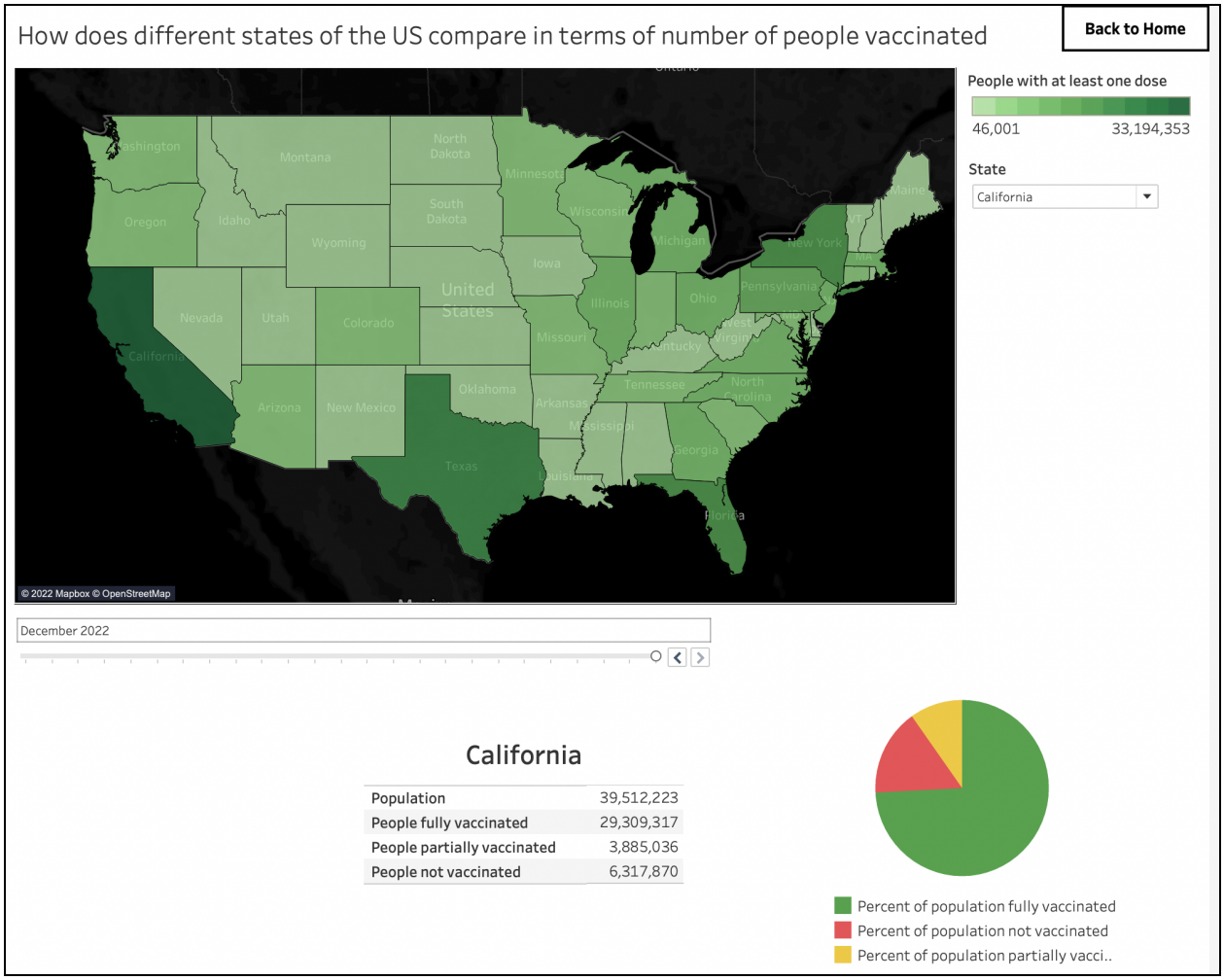
Countrywise Cases and Deaths



USA Statewise Cases and Deaths



Countrywise People Vaccinated



USA Statewise People Vaccinated