Name:=JainishBarbhaya Div:=B

RollNo:=4 EnrollmentNo:=230823007

1.Read the dataset and display the 10 records.

```python
import pandas as pd

# Read the dataset
df = pd.read_csv('world_literacy_data_2021_2023.csv')

# Display the first 10 records
print(df.head(10))
```

```
                Country                 GDP Literacy Rate  Life
expectancy  \
0          Afghanistan      $19,101,353,833          9.70%
64.5
1              Albania      $15,278,077,447         55.00%
78.5
2              Algeria     $169,988,236,398         51.40%
76.7
3              Andorra       $3,154,057,987            NaN
NaN
4               Angola      $94,635,415,870          9.30%
60.8
5  Antigua and Barbuda       $1,727,759,259         24.80%
76.9
6            Argentina     $449,663,446,954         90.00%
76.5
7              Armenia      $13,672,802,158         54.60%
74.9
8            Australia   $1,392,680,589,329        113.10%
82.7
9              Austria     $446,314,739,528         85.10%
81.6

    Year
0   2022
1   2022
2   2022
3   2022
4   2022
5   2022
```

```
6  2022
7  2022
8  2022
9  2022
```

2.Print all the column names.

```
# Print all column names
print(df.columns)

Index(['Country', 'GDP', 'Literacy Rate', 'Life expectancy', 'Year'],
dtype='object')
```

3.Print the size(rows, cols) of dataset.

```
# Print the size of the dataset
print(df.shape)

(585, 5)
```

4.Check for the NULL values in dataset and if found replace the NULL values with the average values.

```
# Check for NULL values
print(df.isnull().sum())

# Replace NULL values with average values
df.fillna(df.mean(numeric_only=True), inplace=True)

# Print the updated dataset
print(df.head(10))

Country             0
GDP                 4
Literacy Rate      19
Life expectancy    17
Year                0
dtype: int64
              Country                    GDP Literacy Rate  Life
expectancy  \
0          Afghanistan     $19,101,353,833          9.70%
64.500000
```

```
1              Albania       $15,278,077,447              55.00%
78.500000
2              Algeria      $169,988,236,398              51.40%
76.700000
3              Andorra        $3,154,057,987                 NaN
72.281866
4               Angola       $94,635,415,870               9.30%
60.800000
5   Antigua and Barbuda       $1,727,759,259              24.80%
76.900000
6             Argentina      $449,663,446,954              90.00%
76.500000
7              Armenia       $13,672,802,158              54.60%
74.900000
8            Australia    $1,392,680,589,329             113.10%
82.700000
9              Austria      $446,314,739,528              85.10%
81.600000

    Year
0   2022
1   2022
2   2022
3   2022
4   2022
5   2022
6   2022
7   2022
8   2022
9   2022
```

5.Print the updated dataset.

```
print(df.head(586))
```

```
        Country                        GDP Literacy Rate  Life expectancy
Year
0    Afghanistan        $19,101,353,833              9.70%         64.500000
2022
1        Albania        $15,278,077,447             55.00%         78.500000
2022
2        Algeria       $169,988,236,398             51.40%         76.700000
2022
3        Andorra         $3,154,057,987                NaN         72.281866
2022
4         Angola        $94,635,415,870              9.30%         60.800000
2022
..           ...                    ...                ...               ...
```

```
...
580    Venezuela     $1,185,728,677          2.60%          72.800000
2021
581      Vietnam   $1,119,190,780,753       11.60%          54.300000
2021
582        Yemen     $1,050,992,593         15.70%          83.100000
2021
583       Zambia    $10,220,781,063        136.60%          77.200000
2021
584     Zimbabwe  $1,394,116,310,769       104.60%          81.000000
2021

[585 rows x 5 columns]
```

6.Display the top 10 countries having highest GDP year 2023.

```python
# Filter data for the year 2023 and sort by GDP
top_gdp_2023 = df[df['Year'] == 2023].sort_values(by='GDP',
ascending=False).head(10)
print(top_gdp_2023[['Country', 'GDP' , 'Year']])
```

```
                    Country               GDP   Year
387                   Yemen    $96,107,662,398   2023
389                Zimbabwe    $95,503,088,538   2023
385               Venezuela    $94,635,415,870   2023
384                 Vanuatu       $917,058,851   2023
383              Uzbekistan   $909,070,395,161   2023
382                 Uruguay    $88,941,298,258   2023
381           United States    $86,000,000,000   2023
380          United Kingdom       $850,655,017   2023
379    United Arab Emirates    $84,008,783,756   2023
378                 Ukraine       $825,385,185   2023
```

7.Display the top 10 countries having highest Literacy rate in 2023.

```python
# Filter data for the year 2023 and sort by Literacy Rate
top_literacy_2023 = df[df['Year'] == 2023].sort_values(by='Literacy
Rate', ascending=False).head(10)
print(top_literacy_2023[['Country', 'Literacy Rate' , 'Year']])
```

```
          Country Literacy Rate   Year
197       Algeria        90.00%   2023
```

```
283        Kenya           90.00%  2023
195  Afghanistan            9.30%  2023
385    Venezuela            9.30%  2023
360    Sri Lanka            9.30%  2023
352     Slovakia            9.30%  2023
305       Mexico            9.30%  2023
281       Jordan            9.30%  2023
370         Togo            9.00%  2023
315        Nauru            9.00%  2023
```
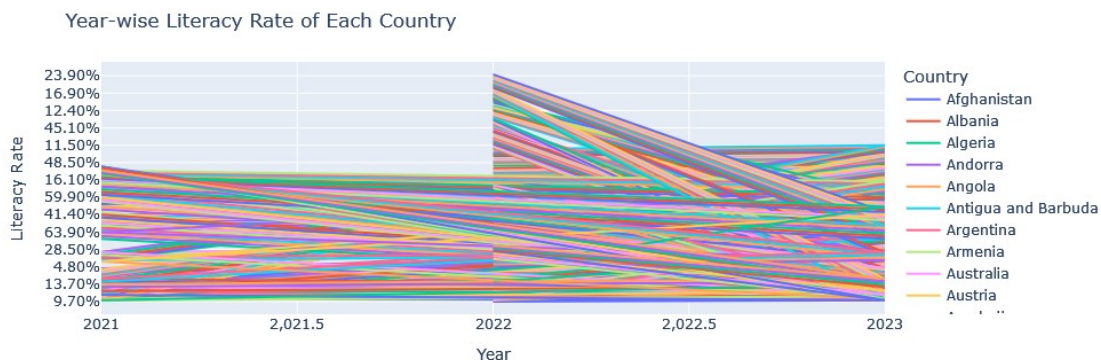
8.Plot the year wise literacy rate of each country.

```python
import pandas as pd
import plotly.express as px

df =  pd.read_csv("world_literacy_data_2021_2023.csv")

# Create the Plotly figure
fig = px.line(df, x='Year', y='Literacy Rate', color='Country')

# Customize the layout
fig.update_layout(
    title='Year-wise Literacy Rate of Each Country',
    xaxis_title='Year',
    yaxis_title='Literacy Rate',
    legend_title='Country'
)

# Show the figure
fig.show()
```

9.Plot the year wise GDP of each country.

```python
import pandas as pd
import plotly.express as px

df =  pd.read_csv("world_literacy_data_2021_2023.csv")

#Plot the year-wise GDP of each country
fig_gdp = px.line(df, x='Year', y='GDP', color='Country',
markers=True)

# Customize the layout
fig.update_layout(
     title='Year-wise GDP of Each Country',
    xaxis_title='Year',
    yaxis_title='GDP',
    legend_title='Country',
    legend=dict(x=1, y=1),  # Position legend outside the plot area
    margin=dict(l=0, r=200, t=40, b=40)  # Adjust margins to fit the
legend
)

# Show the figure
fig.show()
```
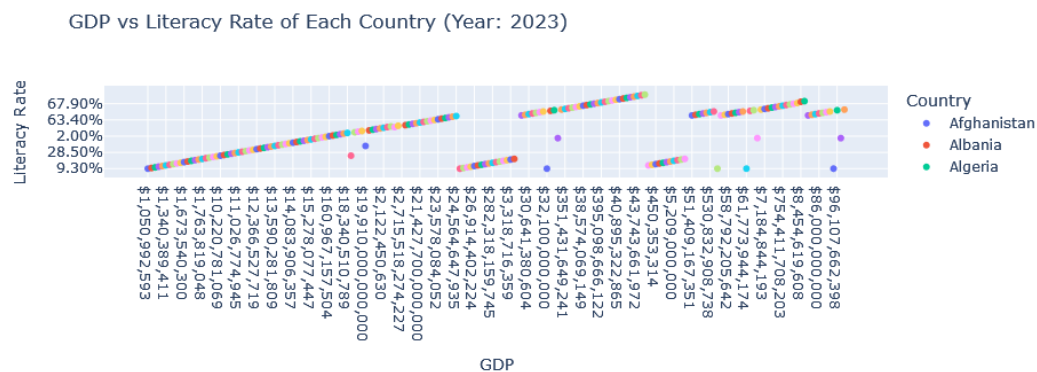

Year-wise GDP of Each Country
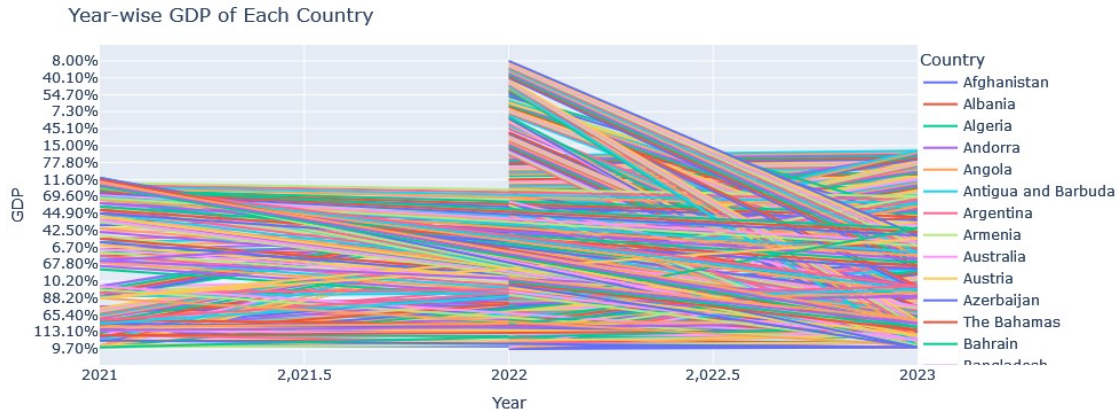
10.Plot the GDP and Literacy rate of the each country (Year : 2023).

```python
import plotly.express as px

# Filter data for the year 2023
df_2023 = df[df['Year'] == 2023]
df = pd.read_csv('world_literacy_data_2021_2023.csv') # iris is a
pandas DataFrame
```

```
fig = px.scatter(df_2023, x="GDP", y="Literacy Rate", color='Country',
                                    title='GDP vs Literacy Rate of Each
Country (Year: 2023)')


fig.update_layout(
    xaxis_title='GDP',
    yaxis_title='Literacy Rate',
    legend_title='Country'
)
fig.show()
```
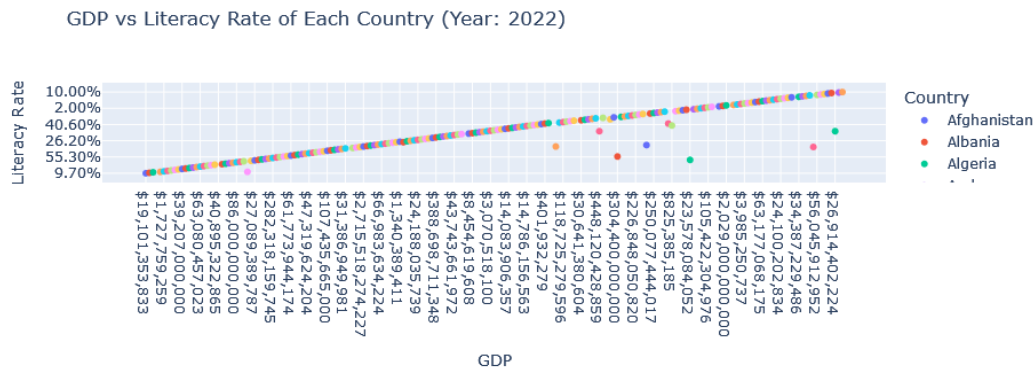

GDP vs Literacy Rate of Each Country (Year: 2023)

11.Plot the GDP and Literacy rate of the each country (Year : 2022).

```
import plotly.express as px

# Filter data for the year 2022
df_2023 = df[df['Year'] == 2022]
df = pd.read_csv('world_literacy_data_2021_2023.csv') # iris is a
pandas DataFrame
fig = px.scatter(df_2023, x="GDP", y="Literacy Rate", color='Country',
                                    title='GDP vs Literacy Rate of Each
Country (Year: 2022)')


fig.update_layout(
    xaxis_title='GDP',
    yaxis_title='Literacy Rate',
    legend_title='Country'
)
fig.show()
```

GDP vs Literacy Rate of Each Country (Year: 2022)

12.Plot the GDP and Literacy rate of the each country (Year : 2021).

```python
import plotly.express as px

# Filter data for the year 2021
df_2023 = df[df['Year'] == 2021]
df = pd.read_csv('world_literacy_data_2021_2023.csv') # iris is a
pandas DataFrame
fig = px.bar(df_2023, x="GDP", y="Literacy Rate", color='Country',
                              title='GDP vs Literacy Rate of Each
Country (Year: 2021)')


fig.update_layout(
    xaxis_title='GDP',
    yaxis_title='Literacy Rate',
    legend_title='Country'
)
fig.show()
```

GDP vs Literacy Rate of Each Country (Year: 2021)
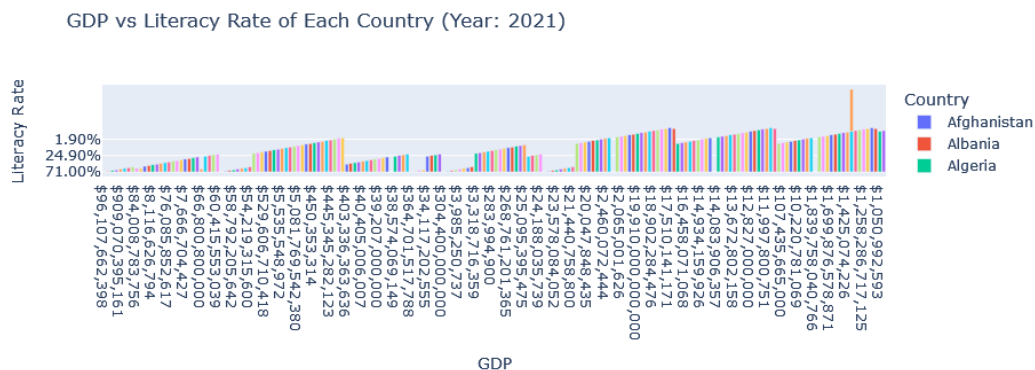
13.Plot the Top 10 countries having highest Literacy Rate of year 2023.

```python
import plotly.express as px

df = pd.read_csv('world_literacy_data_2021_2023.csv')
df_2023 = df[df['Year'] == 2023].sort_values(by='Literacy Rate',
ascending=False).head(10)

fig = px.scatter(df_2023, x="Country", y="Literacy Rate",
color='Country',
                                    title='Top 10 Countries with Highest
Literacy Rate (Year: 2023)' , )

fig.update_layout(
    xaxis_title='Country',
    yaxis_title='Literacy Rate',
    legend_title='Country'
)
fig.show()
```