Python Script

Assignment 6.2: Histogram, Boxplot, Bullet Chart, Parallel Coordinate

DSC640

Taniya Adhikari

In [48]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
from numerize import numerize
import matplotlib.ticker as ticker
from matplotlib.ticker import FuncFormatter
import plotly.express as px
import plotly.graph_objects as go
```

In [49]:
```
birthrates = pd.read_csv('birth-rate.csv')
birthrates.head(3)
```

Out[49]:

	Country	1960	1961	1962	1963	1964	1965	1966	1967	1968	...	1999	2000	2001	2002	2003	2004	200
0	Aruba	36.400	35.179	33.863	32.459	30.994	29.513	28.069	26.721	25.518	...	15.024	14.528	14.041	13.579	13.153	12.772	12.44
1	Afghanistan	52.201	52.206	52.208	52.204	52.192	52.168	52.130	52.076	52.006	...	51.229	50.903	50.486	49.984	49.416	48.803	48.17
2	Angola	54.432	54.394	54.317	54.199	54.040	53.836	53.585	53.296	52.984	...	48.662	48.355	48.005	47.545	46.936	46.184	45.33

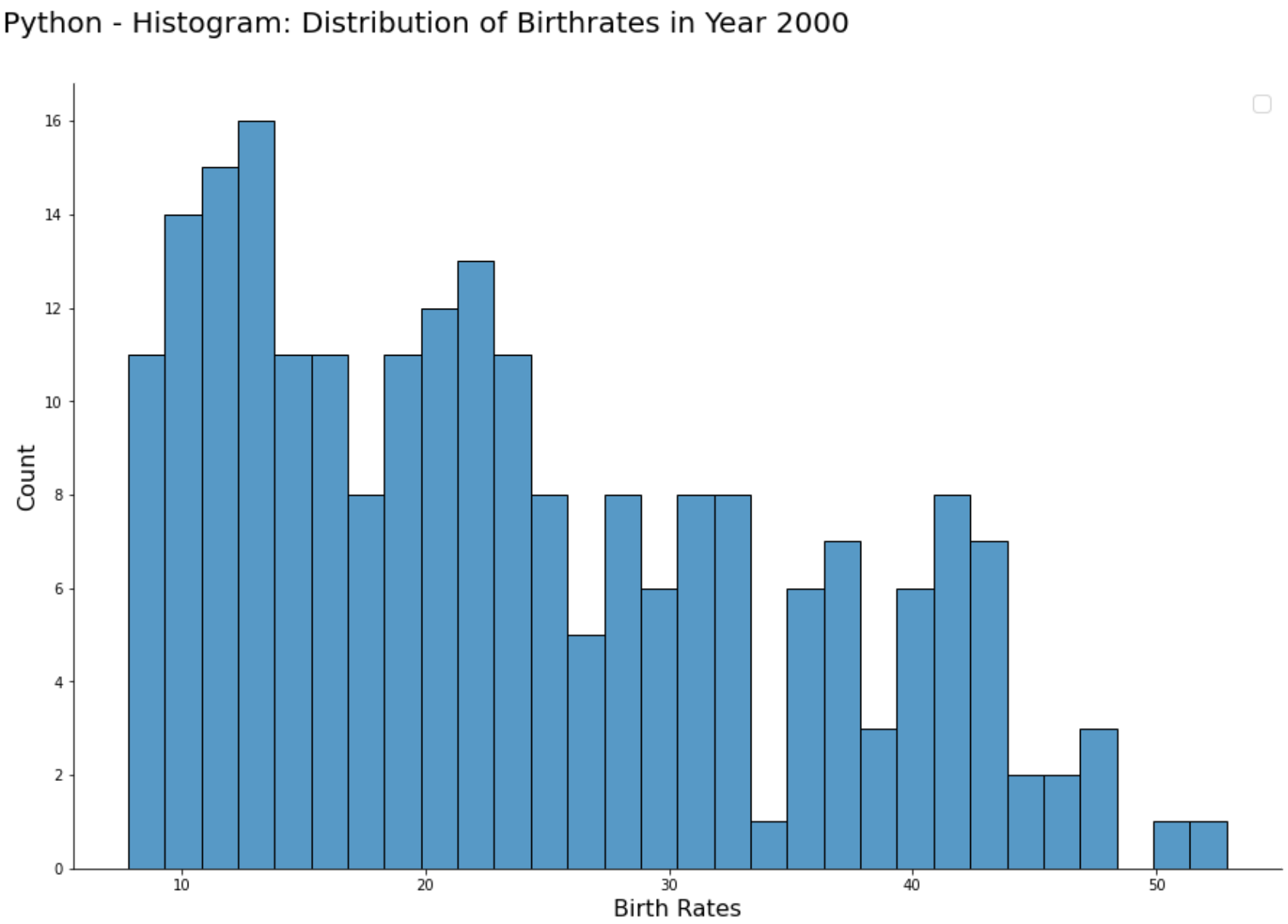
3 rows × 50 columns

Python Histogram

In [50]:
```
plt.rcParams['figure.figsize'] = [15,10]
fig, ax = plt.subplots()
```

```
sns.histplot(data=birthrates, x="2000", bins=30)


plt.suptitle("Python - Histogram: Distribution of Birthrates in Year 2000",
             size=20, x=0.08, y=.95,horizontalalignment='left', verticalalignment='top')

plt.ylabel('Count', size=16)
plt.xlabel('Birth Rates', size=16)

plt.legend(fontsize=15)

right_side = ax.spines["right"]
right_side.set_visible(False)
top = ax.spines["top"]
top.set_visible(False)

plt.show()
```

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when
legend() is called with no argument.
```



Python - Box Plot

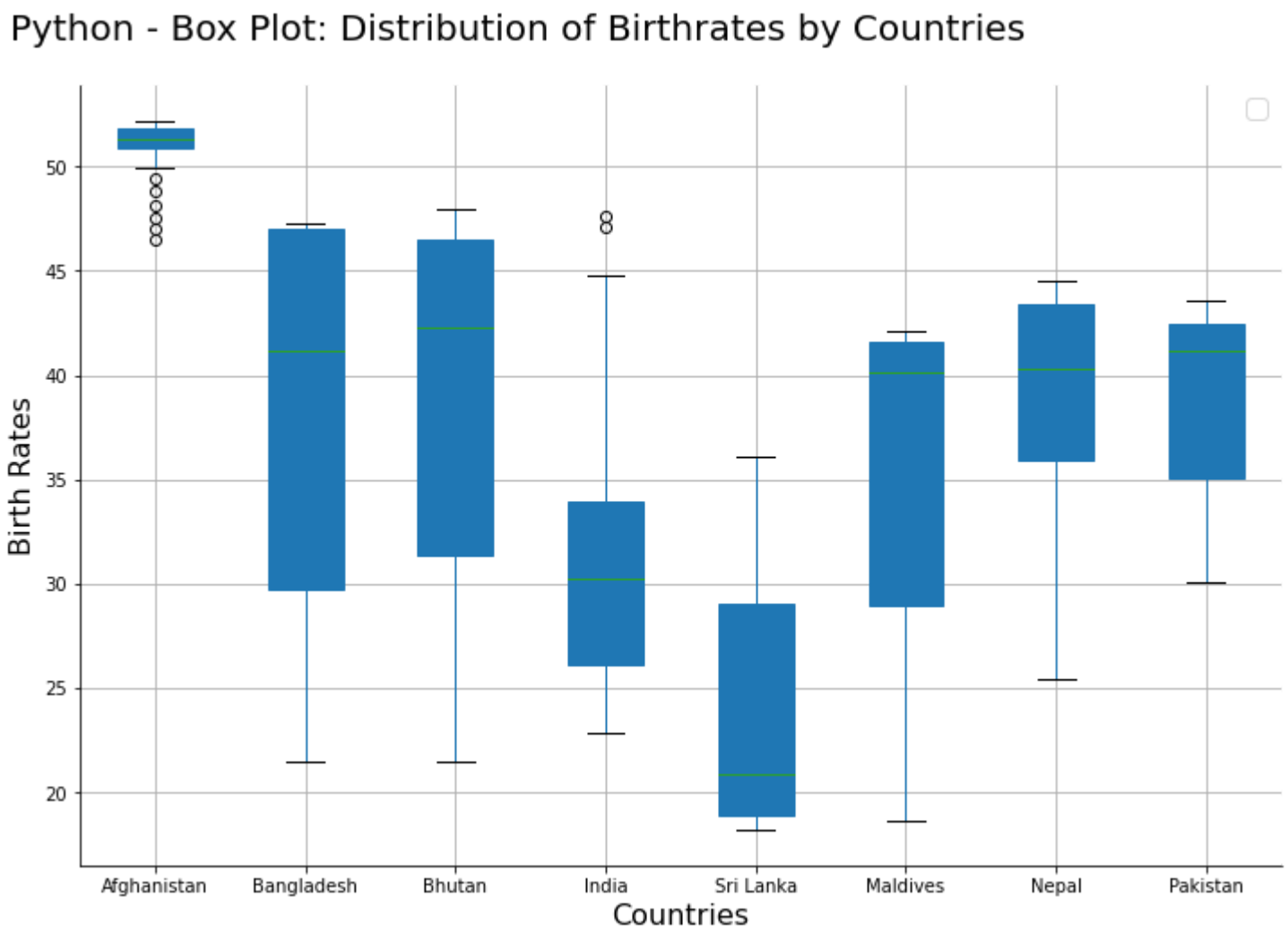
In [51]:
```
southasian = ['India', 'Pakistan', 'Bangladesh', 'Nepal', 'Bhutan', 'Maldives', 'Afghanistan', 'Sri Lanka']
```

In [53]:
```
bd = birthrates.set_index('Country')
```

In [54]:
```
Filter_df  = bd[bd.index.isin(southasian)]
```

In [55]:
```
plt.rcParams['figure.figsize'] = [12,8]

ax = Filter_df.T.boxplot(patch_artist=True)
plt.suptitle("Python - Box Plot: Distribution of Birthrates by Countries",
             size=20, x=0.08, y=.95,horizontalalignment='left', verticalalignment='top')

plt.ylabel('Birth Rates', size=16)
plt.xlabel('Countries', size=16)

plt.legend(fontsize=15)

right_side = ax.spines["right"]
right_side.set_visible(False)
top = ax.spines["top"]
top.set_visible(False)
```

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when
legend() is called with no argument.
```



Python - Bullet Chart

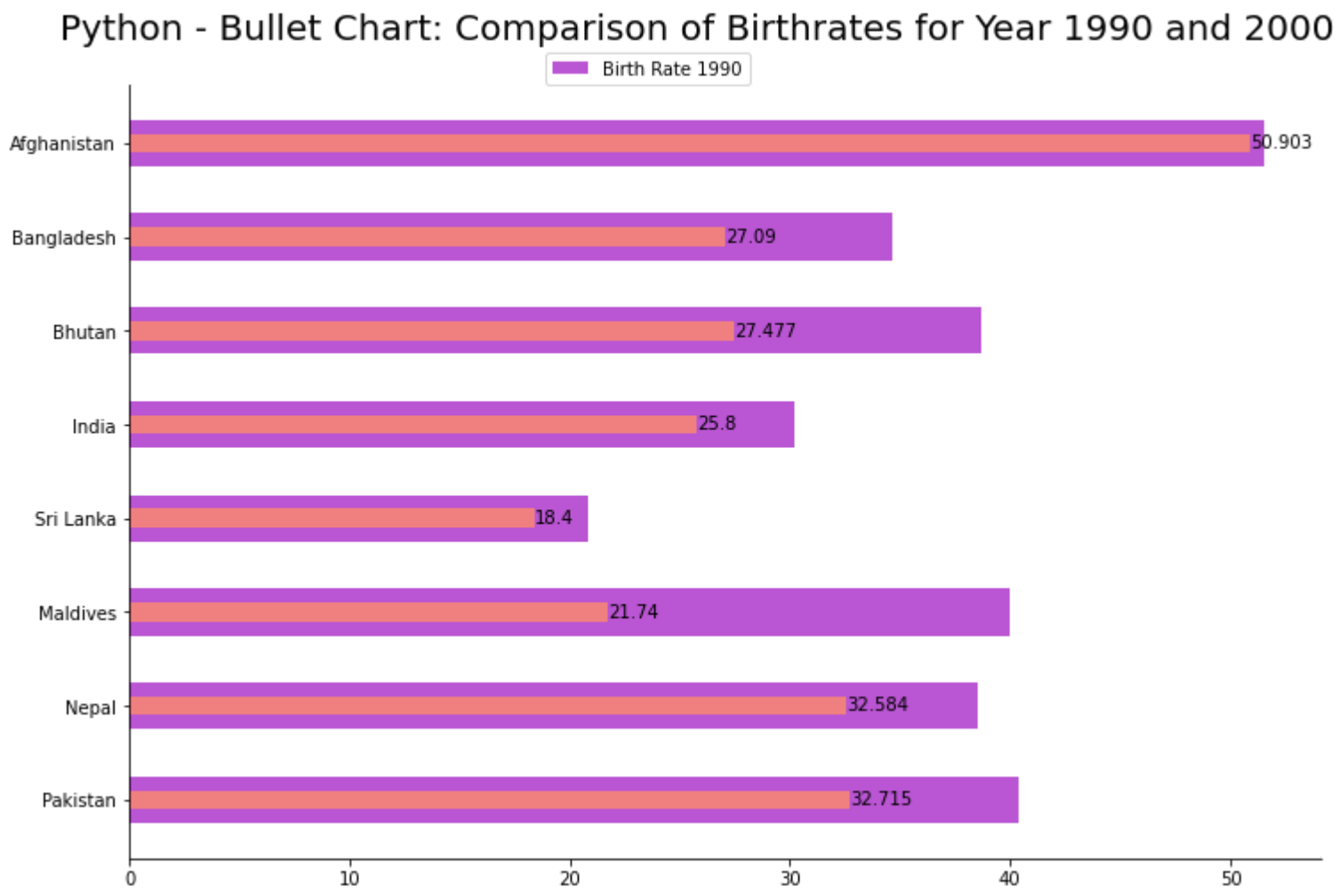
In [74]:

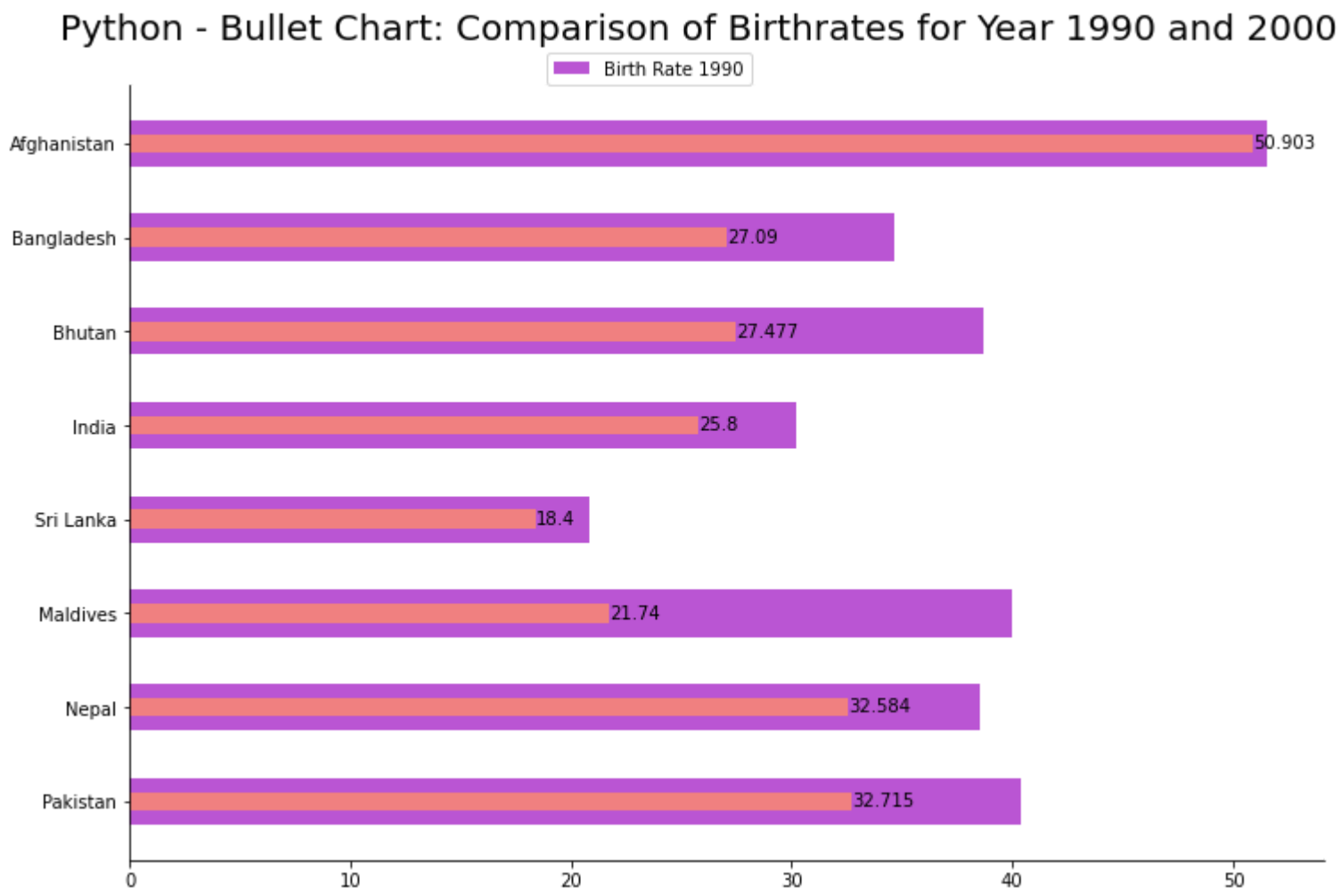
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Function: bullet_chart, creates a horizontal bar graph representing a bullet chart
# Inputs: (1) dataframe with 4 rows - the name of the column ("business"), "base goal" or the quota, "stretch goal" or
# Output: matplotlib image representing a bullet chart
def bullet_chart(df,color_code=False):
    y_pos = np.arange(len(df.index))
```

```
  #assign coloring
    df["col"]="indigo"
    if (color_code==True):
        for i in y_pos:
            if(df["2000"][i]>=df["1990"][i]):
                df["col"][i]="gold"
            else:
                df["col"][i]="lightcoral"

    #Initialize plot
    fig, ax = plt.subplots()
    ax.barh(y_pos, df["1990"], height=0.5, align='center', color='mediumorchid', label = "Birth Rate 1990")
    ax.barh(y_pos, df["2000"], height=0.2, align='center',color=df["col"])
    ax.set_yticklabels(df.index)
    ax.set_yticks(y_pos)
    ax.invert_yaxis()

  #add data labels
    for i in y_pos:
        ax.text(df["2000"][i], i+0.05, df["2000"][i])

      #add legend and format borders
    plt.legend(loc=(0.35,1.0))
    plt.suptitle("Python - Bullet Chart: Comparison of Birthrates for Year 1990 and 2000",
             size=20, x=0.08, y=.95,horizontalalignment='left', verticalalignment='top')

    right_side = ax.spines["right"]
    right_side.set_visible(False)
    top = ax.spines["top"]
    top.set_visible(False)

    return fig

# Use Sisense for Cloud Data Teams to visualize a dataframe or an image by passing data to periscope.output()
bullet_chart(bullet, color_code=True)
```

```
C:\Users\bibek\AppData\Local\Temp\ipykernel_4264\87858703.py:12: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a
-view-versus-a-copy
  df["col"]="Blue"
C:\Users\bibek\AppData\Local\Temp\ipykernel_4264\87858703.py:18: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a
-view-versus-a-copy
  df["col"][i]="lightcoral"
C:\Users\bibek\AppData\Local\Temp\ipykernel_4264\87858703.py:24: UserWarning: FixedFormatter should only be used togeth
er with FixedLocator
  ax.set_yticklabels(df.index)
```

Out[74]:


Python - Bullet Chart: Comparison of Birthrates for Year 1990 and 2000
Birth Rate 1990
Afghanistan
Bangladesh
Bhutan
India
Sri Lanka
Maldives
Nepal
Pakistan
50.903
27.09
27.477
25.8
18.4
21.74
32.584
32.715
0
10
20
30
40
50

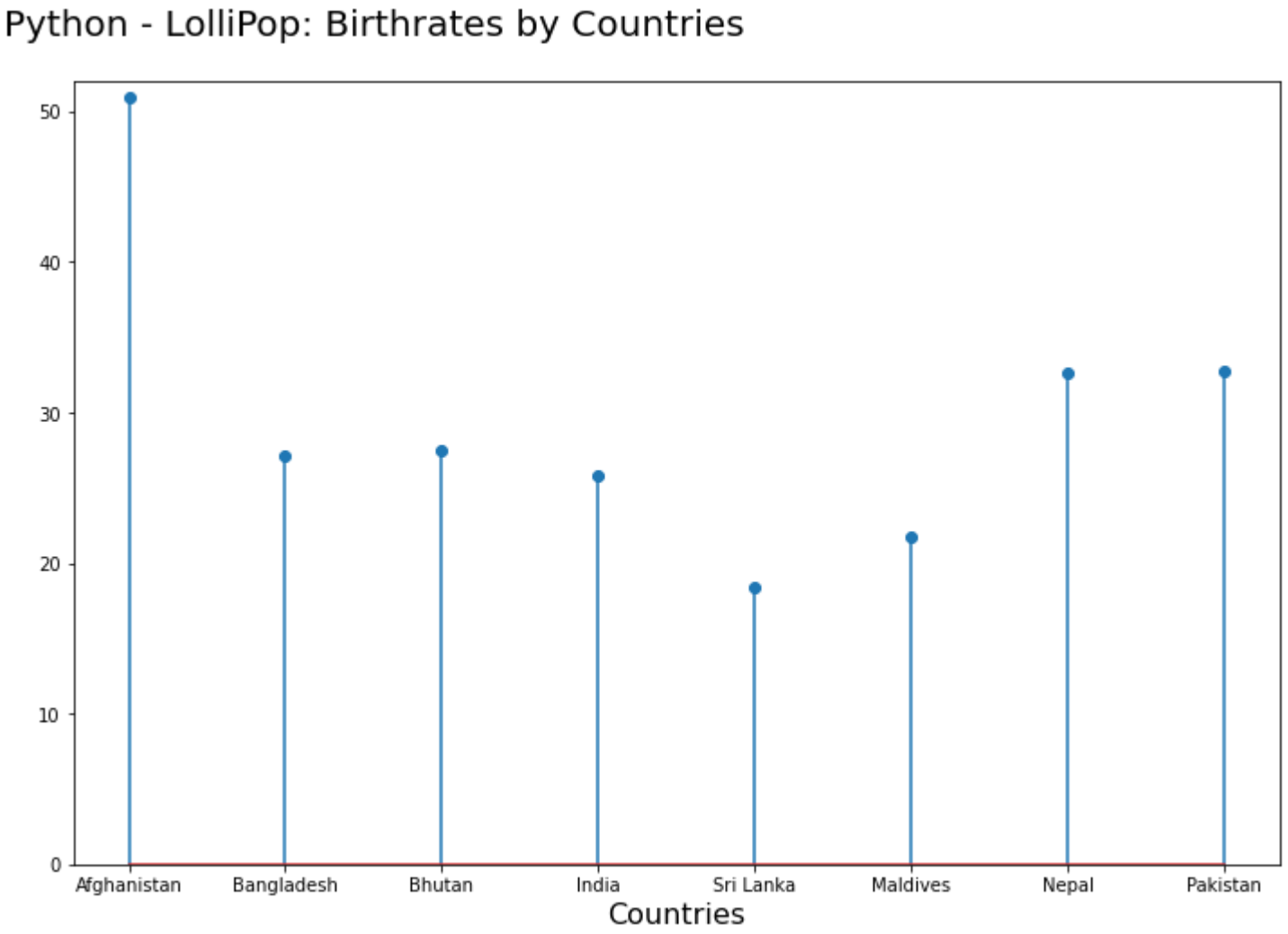


Python - Lollipop Plot:

In [75]:

```
plt.rcParams['figure.figsize'] = [12,8]

# stem function
ax = plt.stem(bullet.index, bullet['2000'])
plt.ylim(0, 52)
plt.suptitle("Python - LolliPop: Birthrates by Countries",
             size=20, x=0.08, y=.95,horizontalalignment='left', verticalalignment='top')

plt.xlabel('Countries', size=16)
```

Out[75]:
```
Text(0.5, 0, 'Countries')
```



In [ ]: