```
#load excel file
import pandas as pd
df_total = pd.read_excel('/content/total 5years.xlsx')
df_total
```

	Country Name	Country Code	Series Name	Series Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	India	IND	Population, total	SP.POP.TOTL	1402617695	1414203896	1425423212	1438069596	..

```
#load excel file
import pandas as pd
df_TF= pd.read_excel('/content/total female.xlsx')
df_TF
```

	Series Name	Series Code	Country Name	Country Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	Population, female	SP.POP.TOTL.FE.IN	India	IND	678442228	684279793	689891756	696186332	..

```
#load total male
import pandas as pd
df_TM= pd.read_excel('/content/total male.xlsx')
df_TM
```

	Series Name	Series Code	Country Name	Country Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	Population, male	SP.POP.TOTL.MA.IN	India	IND	724175467	729924103	735531456	741883264	..

```
#load male%
import pandas as pd
df_M_per= pd.read_excel('/content/male %5years.xlsx')
df_M_per
```

	Series Name	Series Code	Country Name	Country Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	Population, male (% of total population)	SP.POP.TOTL.MA.ZS	India	IND	51.630282	51.613781	51.600918	51.588829	..

```
#load percentange female
import pandas as pd
df_F_per= pd.read_excel('/content/female%5years.xlsx')
df_F_per
```

	Country Name	Country Code	Series Name	Series Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	India	IND	Population, female (% of total population)	SP.POP.TOTL.FE.ZS	48.369718	48.386219	48.399082	48.411171	..

```
#to combine 5 dataset into a single
df= pd.concat([df_total,df_TF,df_TM,df_M_per,df_F_per])
df
```

	Country Name	Country Code	Series Name	Series Code	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]	2024 [YR2024]
0	India	IND	Population, total	SP.POP.TOTL	1.402618e+09	1.414204e+09	1.425423e+09	1.438070e+09	..
0	India	IND	Population, female	SP.POP.TOTL.FE.IN	6.784422e+08	6.842798e+08	6.898918e+08	6.961863e+08	..
0	India	IND	Population, male	SP.POP.TOTL.MA.IN	7.241755e+08	7.299241e+08	7.355315e+08	7.418833e+08	..
0	India	IND	Population, male (% of total population)	SP.POP.TOTL.MA.ZS	5.163028e+01	5.161378e+01	5.160092e+01	5.158883e+01	..
			Population,						

Next steps: Generate code with df View recommended plots New interactive sheet

```
#remove series code and 2024 year column,country code,name and series name

df.drop(['Series Code','2024 [YR2024]','Country Name','Country Code'],axis=1,inplace=True, errors='ignore')
df
```

	Series Name	2020 [YR2020]	2021 [YR2021]	2022 [YR2022]	2023 [YR2023]
0	Population, total	1.402618e+09	1.414204e+09	1.425423e+09	1.438070e+09
0	Population, female	6.784422e+08	6.842798e+08	6.898918e+08	6.961863e+08
0	Population, male	7.241755e+08	7.299241e+08	7.355315e+08	7.418833e+08
0	Population, male (% of total population)	5.163028e+01	5.161378e+01	5.160092e+01	5.158883e+01
0	Population, female (% of total population)	4.836972e+01	4.838622e+01	4.839908e+01	4.841117e+01

Next steps: Generate code with df View recommended plots New interactive sheet

```
#change column name
df.columns=['Series Name','2020','2021','2022','2023']


df=df.transpose()


#to set series name as column name
df.columns=df.iloc[0]
df
```

Series Name	Population, total	Population, female	Population, male	Population, male (% of total population)	Population, female (% of total population)
Series Name	Population, total	Population, female	Population, male	Population, male (% of total population)	Population, female (% of total population)
2020	1402617695.0	678442228.0	724175467.0	51.630282	48.369718
2021	1414203896.0	684279793.0	729924103.0	51.613781	48.386219
2022	1425423212.0	689891756.0	735531456.0	51.600918	48.399082
2023	1438069596.0	696186332.0	741883264.0	51.588829	48.411171

Next steps: Generate code with df View recommended plots New interactive sheet

```
#to remove Series name
df=df.iloc[1:]
df
```

Series Name	Population, total	Population, female	Population, male	Population, male (% of total population)	Population, female (% of total population)
2020	1402617695.0	678442228.0	724175467.0	51.630282	48.369718
2021	1414203896.0	684279793.0	729924103.0	51.613781	48.386219
2022	1425423212.0	689891756.0	735531456.0	51.600918	48.399082
2023	1438069596.0	696186332.0	741883264.0	51.588829	48.411171

Next steps: Generate code with df View recommended plots New interactive sheet

```
df.head()
```

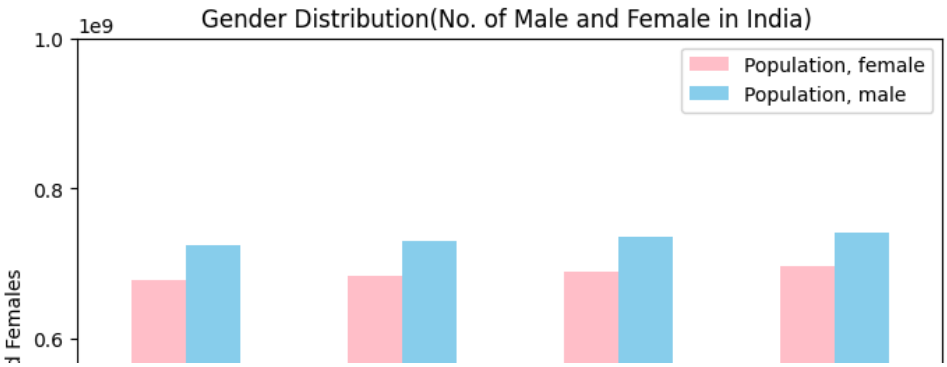
Series Name	Population, total	Population, female	Population, male	Population, male (% of total population)	Population, female (% of total population)
2020	1402617695.0	678442228.0	724175467.0	51.630282	48.369718
2021	1414203896.0	684279793.0	729924103.0	51.613781	48.386219
2022	1425423212.0	689891756.0	735531456.0	51.600918	48.399082
2023	1438069596.0	696186332.0	741883264.0	51.588829	48.411171

Next steps: Generate code with df View recommended plots New interactive sheet

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 5 entries, 0 to 0
Data columns (total 5 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Series Name    5 non-null      object
 1   2020 [YR2020]  5 non-null      float64
 2   2021 [YR2021]  5 non-null      float64
 3   2022 [YR2022]  5 non-null      float64
 4   2023 [YR2023]  5 non-null      float64
dtypes: float64(4), object(1)
memory usage: 412.0+ bytes
```

```
#plot bar graph of male female population for 4 years
import pandas as pd
import matplotlib.pyplot as plt
#extract total population, no. of male,female population
df_new=df.iloc[:,1:3]
df_new
df_new.plot(kind='bar',figsize=(8,7), color=['pink', 'skyblue'])
plt.xlabel('Year')
plt.ylabel('Number of Males and Females')
plt.legend()
plt.ylim(0,1000000000)
plt.title('Gender Distribution(No. of Male and Female in India) ')
plt.show()
```




```
#to draw gender distribution bar graph using number of male and female
import pandas as pd
import matplotlib.pyplot as plt
#extract total population, no. of male,female population
df_new=df.iloc[:,3:5]
df_new
df_new.plot(kind='bar',figsize=(8, 6),width=.4, color=['pink', 'skyblue'])
plt.xlabel('Year')
plt.ylabel('% of Males and Females')
plt.ylim(0,100)
plt.title('Gender Distribution in % (India)')
plt.show()
```

Start coding or generate with AI.

Double-click (or enter) to edit