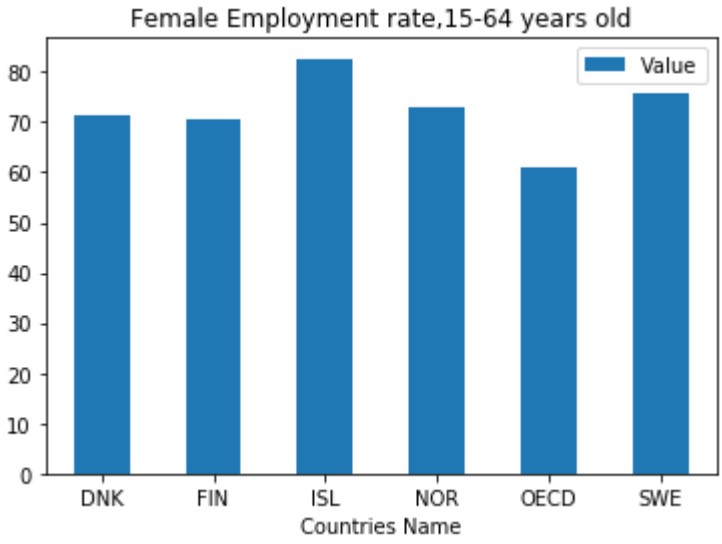
1. **Analyze the visualization and point out possible problem areas and misuses of basic visualization principles. Put your findings in written form.**

The chart is pie chart which is very bad at translating angles towards values and it is not good for categorical and numerical relationship for big data Here colors are the sequential colors-light to deep colors which is used for lowest to highest or highest to lowest sequential data. It is not easy to read because the values of percentages and slices of the graph are quite confusing. Here the categories are more than four which is not good to use pie chart for this data. It is easy to compare to use bar chart for this data rather than pie chart becausea bar plot shows the relationship between a numeric and a categorical variable where each entity of the categorical variable is represented as a bar.

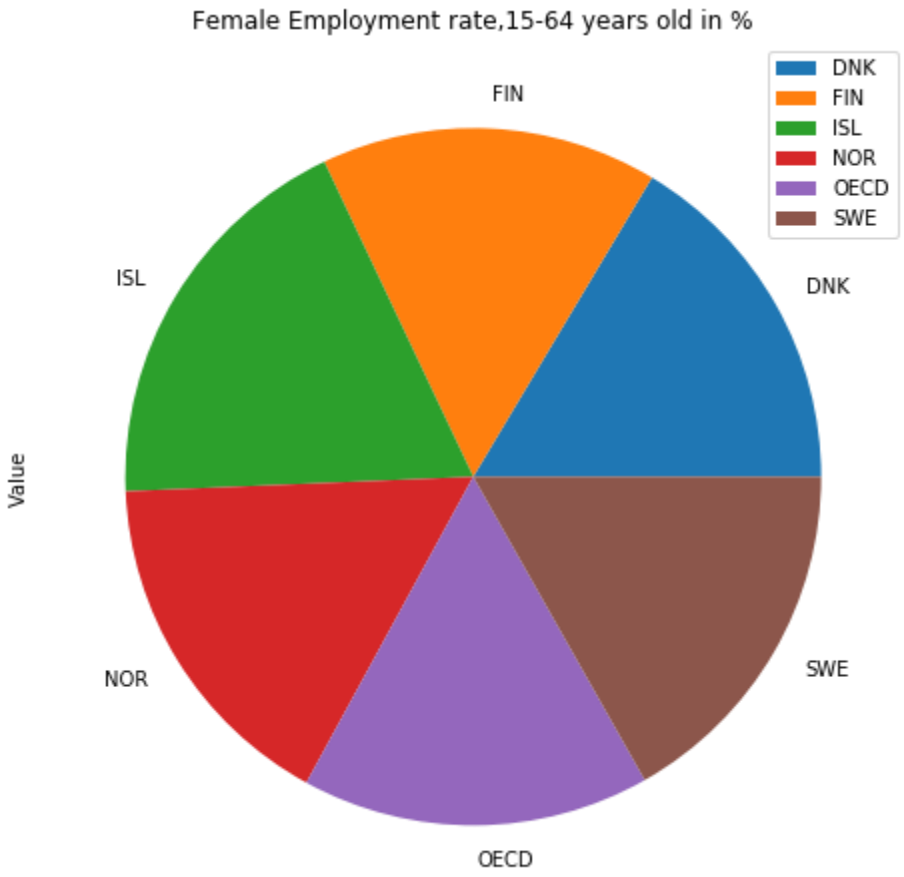
2. Create an improved version of the visualization. Include at least one more data set When creating your visualization to provide more context to the existing data.

First I downloaded the dataset for women perspective from the given link in the assignment. I, however, run codes in the python kernel to create the improved version of the visualization .Codes are given below:

```
>df=pd.read_csv("~/Desktop/women.csv")
>dfi=df[df['LOCATION'].isin(['DNK','FIN','ISL','NOR','SWE','OECD'])]
>dfe=dfi[['LOCATION', 'Value']]
>dfe=dfe.groupby("LOCATION").mean()
>dfe.plot.bar(rot=0, subplots=True)
>plt.title('Female Employment rate,15-64 years old')
>plt.xlabel('Countries Name')
>#plt.legend("Female Employment rate")
```



```
dfv.plot(kind='pie', subplots=True, figsize=(16,8))
plt.title('Female Employment rate,15-64 years old in %')
```

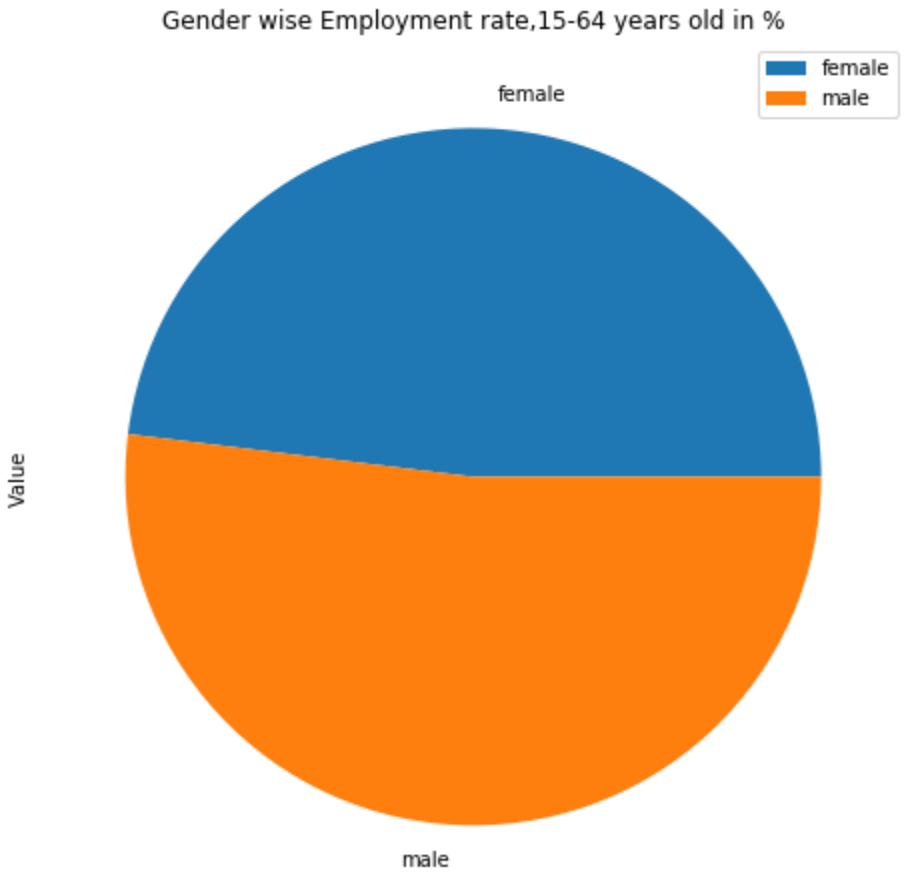


I downloaded the dataset for men perspective from the given link in the assignment and merge this dataset with women dataset. Moreover, I run codes in the python kernel to provide more contexts to the existing data**.** .Codes is given below:

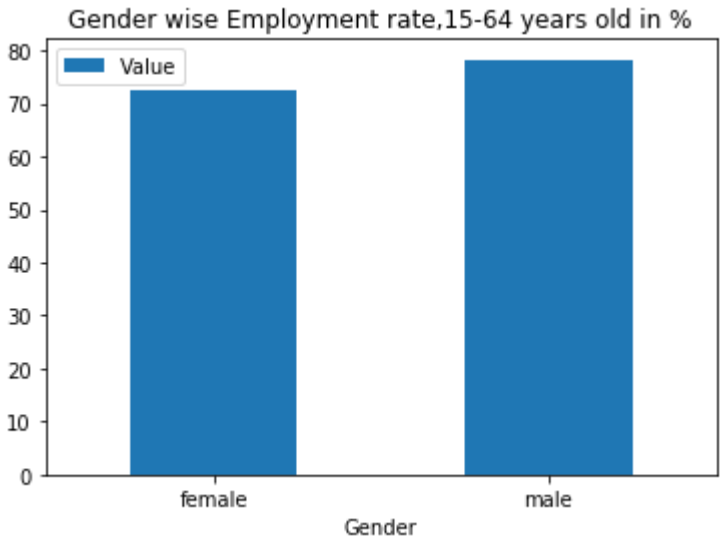
```
df=pd.read_csv("~/Desktop/women.csv")
dfm=pd.read_csv("~/Desktop/men.csv")
dfm["Sex"] = "male"
df["Sex"] = "female"
menWomen = pd.concat([ df, dfm])
dfsex=menWomen[menWomen['LOCATION'].isin(['DNK','FIN','ISL','NOR','SWE','OECD'])]
```

```
dfg=dfsex[['LOCATION', 'Value']]
dfl=dfg.groupby("LOCATION").mean()
dfsub=dfsub.drop(["Flag Codes"],axis=1)

dfsub.plot(kind='pie', subplots=True, figsize=(16,8))
plt.title('Gender wise Employment rate,15-64 years old in %')
```




```
dfsub.plot.bar(rot=0)
plt.title('Gender wise Employment rate,15-64 years old in %')
plt.xlabel('Gender')
```


Gender wise Employment rate,15-64 years old in %
Value
80
70
60
50
40
30
20
10
0
female
male
Gender