DoctorsVisits-Analysis

Using python in Anaconda Navigator

Import libraries

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

Read the dataset

```
df = pd.read_csv("DoctorVisits.csv") df.head(15)
```

Display complete information about the columns of the dataset such as column name,count,Data type and overoll memory usage

```
df.info()
```

Find out the total no: of people based on their count of illeness

```
df["illness"].value_counts()
df["gender"].value_counts()
```

visualize and analyse the maximum,minimum and medium income

```
y = list(df.income)
plt.boxplot(y)
plt.show()
```

Find out the no of days of reduced activity of male and female seperately due to illness

```
df.groupby(['gender', 'reduced']).mean()
```

visualize is their is any missing values in the dataset based on a heat map

```
#missing values
sns.heatmap(df.isnull(),cbar=False,cmap='viridis')
```

Find out the correlation between variables in the given dataset correlation between different variables

```
plt.figure(figsize=(10,10))
sns.heatmap(df.corr(),cbar=True,annot=True,cmap='Blues')
```

Analyse how the income of a patient affects the no of visits to the hospital

```
#relation between income and visits
plt.figure(figsize=(10,10))
plt.scatter(x='income',y='visits',data=df)
plt.xlabel('income')
plt.ylabel('visits')
```

count and visualize the number of males and females affected by illness

```
sns.histplot(df.gender,bins=2)
```

visualize the percentage of people getting govt health insurance due to low income,due to old age and also the percentage of the people having private health insurance

```
# % of people getting Insurance due to low income
label=['yes','no']
Y = df[df['freepoor']=='yes']
N = df[df['freepoor']=='no']
x = [Y.shape[0],N.shape[0]]
plt.figure(figsize=(5,5))
plt.pie(x,labels=label)
plt.title("% of people getting govt health Insurance due to low income ")
plt.show()
```

```
# % of people having private Insurance
Y = df[df['private']=='yes']
N = df[df['private']=='no']
x = [Y.shape[0],N.shape[0]]
plt.figure(figsize=(5,5))
plt.pie(x,labels=label)
plt.title("% of people having private health Insurance ")
plt.show()
# % of people getting govt Insurance due to old age,disability or veteran status
Y = df[df['freerepat']=='yes']
N = df[df['freerepat']=='no']
x = [Y.shape[0],N.shape[0]]
plt.figure(figsize=(5,5))
plt.pie(x,labels=label)
plt.title("% of people getting govt health Insurance due to old age, disability or veteran status ")
plt.show()
```

Plot a horizantal bar chart to analyse the reduced days of activity due to illness based on gender

```
db= df.groupby('gender')['reduced'].sum().to_frame().reset_index()
#creating the bar chart
plt.barh(db['gender'],db['reduced'],color = ['cornflowerblue','lightseagreen'])
#Adding the aesthetics
plt.title('Bar Chart')
plt.xlabel('reduced activity')
plt.ylabel('gender')
#Show the plot
plt.show()
```