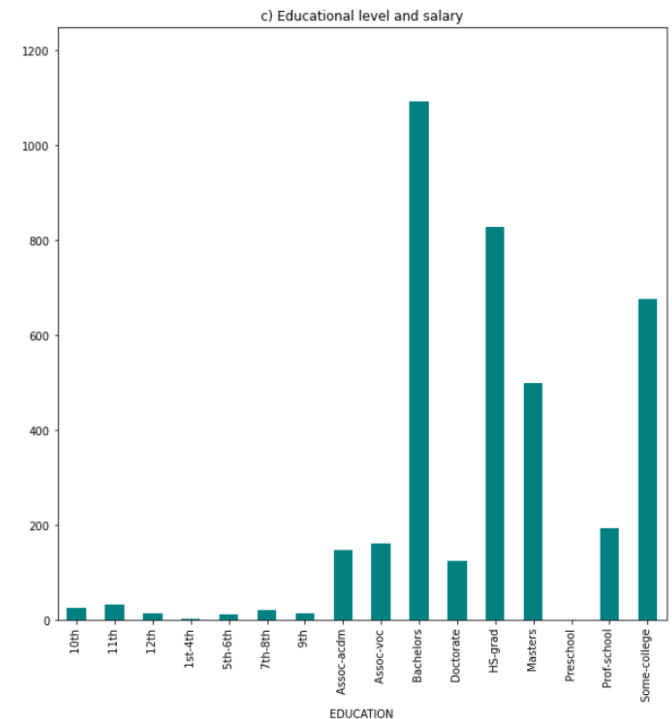
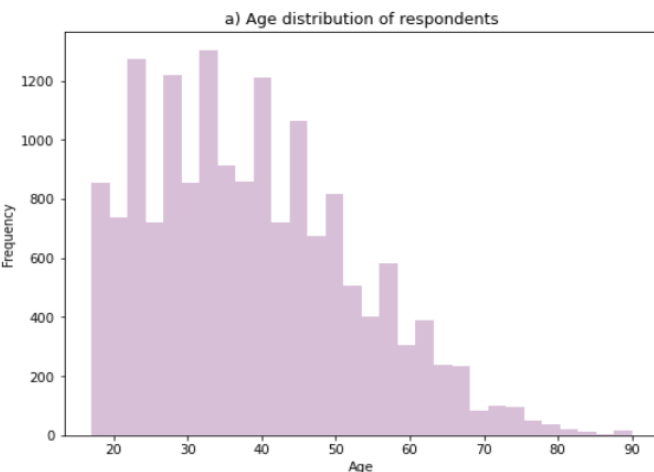


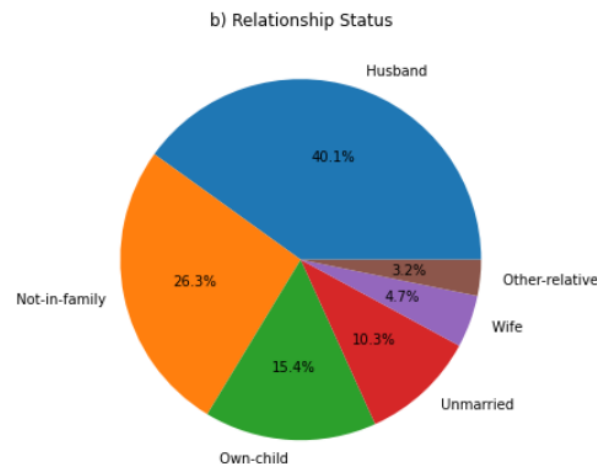
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
a = pd.read_csv("census_income_dataset.csv")
a = a[~a["AGE"].isnull()]
a["AGE"].plot.hist(bins=30,figsize=(8,6),color='thistle',title=("a) Age distribution of respondents"))
plt.xlabel("Age");plt.show();
b=a["RELATIONSHIP"].value_counts()
b.plot(kind='pie',autopct='%1.1f%%',title=("b) Relationship Status"),figsize=(8,6))
plt.ylabel('');plt.show()
a.loc[a["SALARY"] == " <=50K", "SALARY"] = 0
a.loc[a["SALARY"] == " >50K", "SALARY"] = 1
from scipy.stats import pearsonr
corr, p_value = pearsonr(a["EDUCATION-NUM"], a["SALARY"])
a_groups = a.groupby(['EDUCATION-NUM'])["SALARY"].sum()
a_groups.plot(kind='bar',figsize=(10,10),ylim=(0,1250),color='teal')
plt.title("c) Educational level and salary")
plt.show()
```



- This bar plot provide the correlation between the educational level and salary obtained.
- Also, in this graph we have used pearson's correlation in order to find out the relate the EDUCATION_NUM and SALARY.
- It appears that those who successfully complete their bachelor's degrees earn more than \$50,000, whereas younger children (aged 1-4) get paid considerably less.



- The above histogram plot perfectly shows how the age of respondents are distributed in the census data.
- 75% of the total population of young people is, on average, between the ages of 25 and 35.



- This pie chart provides a well informative data with exact percentage values.
- The typical ratio between married and single people is four to one in favor of husbands.About one-third of total population is not in family.

NAME: ABHIK SARKAR
 MATRIKELNUMMER: 23149662
 IDM Id : lo88xide
 EMAIL: abhik.sarkar@fau.de