Key_4

```
import pandas as pd

# Load the dataset
data_url = 'http://archive.ics.uci.edu/dataset/2/adult'
df = pd.read_csv(data_url, header=None)

# Assign column names to the dataframe
df.columns = ['age', 'workclass', 'fnlwgt', 'education', 'education-num', 'marital-
status',
              'occupation', 'relationship', 'race', 'sex', 'capital-gain', 'capital-
loss',
              'hours-per-week', 'native-country', 'salary']

# 1. Count the number of men and women
gender_counts = df['sex'].value_counts()
print("Number of men:", gender_counts['Male'])
print("Number of women:", gender_counts['Female'])

# 2. Calculate the average age of women
average_age_women = df[df['sex'] == 'Female']['age'].mean()
print("Average age of women:", average_age_women)

# 3. Calculate the proportion of German citizens
german_citizens_prop = (df['native-country'] == 'Germany').mean()
print("Proportion of German citizens:", german_citizens_prop)

# 4. Calculate the mean age of those who receive more than 50K per year
mean_age_high_income = df[df['salary'] == '>50K']['age'].mean()
print("Mean age of those who receive more than 50K:", mean_age_high_income)

# 5. Calculate the standard deviation of the age of those who receive more than 50K
per year
std_age_high_income = df[df['salary'] == '>50K']['age'].std()
print("Standard deviation of age of those who receive more than 50K:",
std_age_high_income)

# 4-5. Calculate the mean value and standard deviation of the age of those who
receive less than 50K per year
mean_age_low_income = df[df['salary'] == '<=50K']['age'].mean()
std_age_low_income = df[df['salary'] == '<=50K']['age'].std()
print("Mean age of those who receive less than 50K:", mean_age_low_income)
print("Standard deviation of age of those who receive less than 50K:",
std_age_low_income)

# 6. Check if people who receive more than 50k have at least high school education
education_levels = ['Bachelors', 'Prof-school', 'Assoc-acdm', 'Assoc-voc', 'Masters',
'Doctorate']
high_income_education_check = df[df['salary'] ==
'>50K']['education'].isin(education_levels).all()
print("People who receive more than 50k have at least high school education:",
high_income_education_check)
```