

CPE 101, Project Assignment-2 : Education and Poverty Analysis

Please note that this project assignment must be attempted and submitted by each student individually. Only the Python program file needs to be submitted through Canvas. The deadline to submit this assignment is May 13th, 11:59 PM.

Your program will create an application for examining the influence of race on police shootings using real-world dataset. Based on the findings, you are required to report/answer questions and state your opinion.

You will be practicing the following concepts:

- variables
- integers
- if-else
- functions
 - Loops

Project Description

Background:

You are analyzing datasets containing statistics for each county in the United States. These statistics include information such as median household income in each county, education level, etc. In the following tasks, you will use what you have learned thus far to perform some data analysis and uncover useful insights.

Note:

There are several text files you will be working with on this project. Don't worry, you will not actually be writing code to read data from these files. Instead, you are provided helper code that conveniently reads the data from each file. This helper code is organized by labels (ex. Helper Code 1, Helper Code 2) and is

located at the beginning of “Project2.py”. Do NOT modify/delete this helper code, but feel free to take a look at it.

Task 1:

Question 1.1: What is the national percentage of people who have a Bachelor's degree? Work on the function “percent_bachelors”. It takes two parameters, “bach_data” and “pop_data,” which contain the percent of bachelor’s graduates and total population for each county, respectively. This function should **return** the final percentage you calculate to the calling function.

Expected return value: 28.920803001455276

Print the returned value with an appropriate string/prompt using print statement

Note: Take a look at the helper code that uses the files “Bachelors.txt” (Helper Code 1) and “Population.txt” (Helper Code 2). This helper code puts the percentage of bachelor’s graduates for each county in the list “bach_data”, as well as the population of each county into the list “pop_data”.

Note: The bachelor’s graduation data is in percentage values, not decimals (i.e. for 20%, the value in the list bach_data would be 20, not 0.20).

If you are confused about how to utilize the helper code, take a look at this example.

Example:

This function prints each county’s bachelor’s graduation percentage and county population.

```
def interesting_data (bach_data,population_data):  
    for i in range(len(bach_data)):  
        print("County",i,":",bach_data[i],population_data[i])
```

Calling this function would look like this:

```
interesting_data(bach_data, pop_data)
```

The first few lines of output from calling this function would look like this:

County 0: 20.9 55395

County 1: 27.7 200111

County 2: 13.4 26887

Question 1.2: What percent of counties have an average income greater than \$50,000? Work on the function “income_above_50k”. It takes one parameter, “avg_income,” which contains the median income for each county, respectively. This function should **return** the final percentage.

Expected return value: 28.47597836461979

Print the above returned value with an appropriate string/prompt using print statement

Note: Take a look at the helper code that uses the files “Income.txt” (Helper Code 3). This helper code puts the average income for each county into the list called “avg_income”.

Task 2: Find and **return** the total number of counties with a percentage of high school graduates strictly below a given threshold value. Use the function high_school_below_threshold that takes as input the percent of high school graduates for each county (“high_school_data”), and a threshold value.

Expected return value when threshold = 60: 9

Print the above returned value with an appropriate string/prompt statement

Note: Take a look at the helper code that uses the files “HighSchool.txt” (Helper Code 4). This helper code puts the high school graduation rate for each county into the list called “high_school_info”.

Task 3: Work on the function `bachelors_above_threshold`, that takes as inputs the bachelors data (“bach_data”) and a number as a threshold. The function must return the total number of counties with a percentage of bachelors strictly greater than the threshold percentage.

Expected return value when threshold = 60: 4

Print the above returned value with an appropriate string/prompt message

Note: Similar to task 1, question 1.1, use helper code 1 to get the bachelor’s data in a list.

Task 4: Work on the function `below_poverty_total`, that takes input about the percentage of people below poverty level for each county (`poverty_data`) as well as the county population (`pop_data`). The function must return the total number of people below the poverty level across all counties. Do not round!

Expected return value: 48996488.47399998

Note: Take a look at the helper code that uses the file “Poverty.txt” (Helper Code 5). This helper code puts the percentage of people below the poverty level for each county into the list called “poverty_info”. Additionally, use helper code 2 to access the population data.

Task 5: Write a function named `percent_below_poverty`. The function is given the same arguments from task 4, and it must return the total percentage of the national population below the poverty level. This function must call the function you wrote in task 4.

Expected return value: 15.366286413307403

Print the above returned value with an appropriate string/prompt statement

Task 6: Write a function called `education_vs_poverty` that takes as input the percentage of bachelors graduates for each county as well as the percentage of people below the poverty level for each county. Find the 4 national average poverty levels for counties with bachelors graduation rates $<10\%$, $\geq 10\%$ and $<20\%$, $\geq 20\%$ and $<30\%$, and $\geq 30\%$. Return this information in a string in the following format:

Average Poverty Levels:

Bachelors $<10\%$: w%

Bachelors $\geq 10\%$ and $<20\%$: x%

Bachelors $\geq 20\%$ and $<30\%$: y%

Bachelors $\geq 30\%$: z%

Expected return value:

Average Poverty Levels:

Bachelors $<10\%$: 23.986627906976732%

Bachelors $\geq 10\%$ and $<20\%$: 17.986990508096028%

Bachelors $\geq 20\%$ and $<30\%$: 14.181987577639747%

Bachelors $\geq 30\%$: 12.437333333333338%

Based on your results in task 6, observe the relationship between education and poverty level in this dataset. That is, what do you notice about the poverty level as the percentage of bachelor's graduates increases? Write your response. Submit this response as a comment for the project 2 Canvas assignment.