MCIS 540 - Project

Project Overview

The goal of this project is to gain more practice with Python programming and using Python to perform data analysis to solve real-world problems.

Background

In this project, we will review poverty data collected by the 2016 U.S. census for each county in Pennsylvania. We will do some statistics using this data.

The data comes from the U.S. census of 2016. It is located at https://www.census.gov/data/datasets/2016/demo/saipe/2016-state-and-county.html and is available in a number of formats. We will work with the txt version. The web site has description about the format of the text file, indicating the columns the data occurs in.
A few notes about the data:

- The first line has data for the entire state. We are only interested in county data—the rest of the lines in the file.
- The data we care about for each **county** are:
    - count of people age 0-17 (i.e., children) in poverty (position 50-57)
    - percentage of people age 0-17 (i.e., children) in poverty (position 77-80)
    - median household income (position 134-139)
    - county name (position 194-238)

Data Analysis

Using this data to:

1. Find the county with the highest percentage of children in poverty and **print** its name, percentage, count of children in poverty and the median household income.
2. Find the county with the lowest percentage of children in poverty and **print** its name, percentage, count of children in poverty and the median household income.
3. Loop to continually prompt for a county and **print** its name, percentage, count of children in poverty and the median household income. The entered value should not be case sensitive and only the county name is used, i.e. entering “chester” or “Chester” gets the data for Chester County (that is, you don’t search using “Chester County”). Re-prompt if county is not found. Type ‘q’ or ‘quit’ in any case to quit the loop.

Program specifications

Required functions:

a. Each print specified above will print the same data so create a `print_data` function. This function will be used by the next functions.
b. Create a function for the highest percentage data: `print_highest_data`
c. Create a function for the lowest percentage data: `print_lowest_data`
d. Create a function for the county data: `print_county_data`

Deliverable

The deliverable for this project is a Jupyter Notebook file which should be named as MCIS540_ [your_ name]_project.ipynb. Your submission should include your code and your write-up (such as introduction, purpose, program/functions description, results in Data Analysis section, and conclusion/discussion).