Exercise: Hangman

Background

For this exercise we will use plotting to examine factors that may correlate with life expectancy.

The file `life_expectancy.csv` comes from https://www.kaggle.com/kumarajarshi/life-expectancy-who (column headings have been cleaned up to remove leading/trailing whitespace and to have consistent casing). It contains annual data on life expectancy and several other factors for several countries, spanning the years 2000-2015. It contains the following columns:

- Country: the name of the country
- Year: the year in which the data were obtained
- Status: "Developing" or "Developed"
- Life expectancy: the average lifespan of a resident of the country
- Adult mortality: probability that residents of this country will die between age 15 and age 60, expressed in terms of 1000 residents
- Infant deaths: number of annual infant deaths per 1000 people
- Alcohol: alcohol consumption per capita among residents 15 and older, expressed in liters of pure alcohol
- Percentage expenditure: expenditure on health as a percentage of Gross Domestic Product per capita
- Hepatitis B: percentage of 1-year-olds immunized against Hepatitis B
- Measles: number of measles cases reported per 1000 people
- BMI: average body mass index of the population
- Under-five deaths: number of deaths of children under the age of five per 1000 people
- Polio: percentage of 1-year-olds immunized against polio
- Total expenditure: government expenditure on health as a percentage of total government expenditure
- Diphtheria: percentage of 1-year-olds immunized against diptheria, tetanus, and pertussis
- HIV/AIDS: deaths due to HIV/AIDS (0-4 years) per 1000 live births
- GDP: gross domestic product per capita
- Population: population of the country
- Thinness 10-19 years: prevalence of thinness among people ages 10 to 19 as a percentage; a measure of malnourishment in adolescents
- Thinness 5-9 years: prevalence of thinness among people ages 5 to 9 as a percentage; a measure of malnourishment in children
- Income/composition of resources: Human Development Index in terms of income/composition of resources (index ranging from 0 to 1)
- Schooling: number of years of schooling

Template

A template script, `expectancy.py`, is provided. The script contains a function `parse_args()` which handles command-line arguments. It also contains import statements and an `if __name__ == "__main__":` statement to manage overall execution of the program.

Problem statement

Write a function `plot_life_expectancy` that will create a joint plot with 2015 life expectancy as the x axis and a factor of the user's choosing as the y axis.

Instructions

Please use the template to implement your solution to this exercise. Make sure your program is called `expectancy.py`.

`plot_life_expectancy()` function

Write a function called `plot_life_expectancy()`. Your function should take two arguments:

- a string containing the path to a CSV file containing life expectancy data
- a string containing the name of a column in the CSV file that has numeric values

Your function should read the CSV file into a Pandas DataFrame (do not hard-code the path to the file anywhere in your script!). It should create a Seaborn joint plot **using only the rows from the DataFrame where the year is 2015**. Use life expectancy as the x axis and the column specified by the second parameter of the function as the y axis. Use the `"Status"` column as the value for the `hue` argument.

Hint: When you are not using Jupyter, one extra step is required to display a pyplot or seaborn plot.

Docstrings

Please write a docstring for your `plot_life_expectancy()` function. For information on how to write docstrings, see https://umd.instructure.com/courses/1299872/pages/docstrings.

Running your code

To run your program within the VS Code built-in terminal, first make sure you have opened (in VS Code) the directory where your program is saved. If necessary, you can go to the VS Code File menu and select "Open…" on macOS or "Open Folder…" on Windows, and navigate to the directory where your program is.

Then, open the VS Code built-in terminal. Type `python3` (on macOS) or `python` (on Windows) followed by a space and the name of your program. Include the path to `life_expectancy.csv` and a column name from this CSV file as command-line arguments. Here is an example:

```
python3 expectancy.py life_expectancy.csv "Adult mortality"
```

Academic integrity

This assignment is to be done by you individually or with a partner or partners as assigned in class. Outside help of any kind (including, but not limited to, help from the internet, tutors, or classmates other than your assigned partner) is not allowed. Disseminating these instructions in whole or in part without written permission of the instructor is considered an infraction of academic integrity. Posting the instructions, or any part thereof, on the internet is considered dissemination and is strictly prohibited.