

CSE 110

[Home](#)[W1](#)[W2](#)[W3](#)[W4](#)[W5](#)[W6](#)[W7](#)

Project 06: Data Analysis

Regarding Milestones:

A milestone or milepost is a marker placed along a highway to tell you how far you have come, or to indicate your progress toward your destination. In software development projects, a Project Milestone marks a specific point along a project timeline.

To help you make progress toward finishing this project, you will complete part of the program during the middle of the week and submit a "Milestone Submission." Then, by the end of the week, you will complete the program and submit the finished version.

You should read over the complete project description first. Then, at the bottom of this page, you will see which features are required for the milestone and which are required for the overall project.

Overview

At this point in the course, you have learned how to use the major building blocks of programming, including variables, if-statements, loops, lists, and now files. Equipped with these tools, you can use them to solve real-world problems.

For this assignment you will write a program to analyze a dataset containing information about life expectancies over the years throughout the countries of the world.

Project Description

The dataset you will be using comes from [OurWorldInData.org](https://ourworldindata.org) from an article on the Spanish Flu. The first graph on that page shows the life expectancies over the years for various countries.

You can download the dataset directly here: [life-expectancy.csv](#). This is a .csv (Comma Separated Values) file that contains the data you'll need with each column separated by commas. There are roughly 19,000 rows in this dataset.

This dataset is licensed under the [Creative Commons BY license](#), you may also read the [Life Expectancy Data License](#).

Your task is to write a program to help analyze this large amount of data.

Assignment

Download the dataset and write a Python program to analyze it to answer the following questions:

1. What is the year and country that has the lowest life expectancy in the dataset?
2. What is the year and country that has the highest life expectancy in the dataset?
3. Allow the user to type in a year, then, find the average life expectancy for that year. Then find the country with the minimum and the one with the maximum life expectancies for that year.

A sample run could look as follows:

```
Enter the year of interest: 1959
```

```
The overall max life expectancy is: 86.751 from Monaco in 2019
```

```
The overall min life expectancy is: 17.76 from Iceland in 1882
```

```
For the year 1959:
```

```
The average life expectancy across all countries was 54.95
```

```
The max life expectancy was in Norway with 73.49
```

```
The min life expectancy was in Mali with 28.077
```

Milestone Requirements

By the middle of the week, to help make sure you are on track to finish the assignment, you need to complete the following:

1. Download the dataset
2. Load the dataset in your Python program
3. Iterate through the data line by line
4. Split each line into parts

- Find the lowest value for life expectancy and the highest value for life expectancy in the dataset, and display both values. (Note that at this point, you just need the value for this, not the year and the country for that value.)

Final Requirements

Finish the program by getting and displaying the answers to the questions above and adding the required functionality.

External Libraries

This assignment is designed to help you practice all the principles we have been learning during this course. To help achieve this goal, **you may not use any external libraries on this assignment** such as Pandas or CSV Reader. Instead, please follow the examples in the learning activities and the team activity.

Showing Creativity and Exceeding Requirements

You can show creativity and exceed the core requirements by adding any kind of data exploration or additional features. For example, you could:

- Identify the year and country that has the largest drop from one year to the next.
- Allow the user to type in a country, then show the minimum, maximum, and average life expectancy for that country.
- Look for interesting anomalies or patterns in the data.
- Anything else you can think of!

Important: In order to receive credit for showing creativity, you must include a comment at the top of the program that describes in 1-2 sentences what you have added.

Submission

There is an assignment submission in Canvas for both the Milestone and the Final Submission. Make sure to:

- [Return to Canvas](#) to submit:

1. The milestone submission (midweek)
2. The finished project (end of week)

Up Next

- Ponder: [Week 06 Reflection](#)

Other Links:

- Return to: [Week Overview](#) | [Course Home](#)

Copyright © Brigham Young University-Idaho | All rights reserved