Outcomes:

- Working with Python as one of Imperative Programming Languages
- Understanding how Python works and some of the difficulties
- Working with some of Python features
- Working with files in Python
- Solving problems using Python
- Running and testing Python codes

Scoring (465/565):

- (5 points) submission on Git with **10 or more commits**. Your Git history should represent the progress of your work.
- (20 points) Using all of the following Python features:
  - (6 pts) Use of Object-Oriented Programming :
    - Design a logical and practical hierarchy of classes with at least one superclass and a subclass - must be a part of your solution (3 pts).
    - Override an operator, which could be either a logical operator, or a mathematical operator - Does not have to be a part of your solution (2 pts).
    - Override a method - Does not have to be a part of your solution (1 pt).
  - (2 pts) Lambda - must be a part of your solution
  - (2 pts) Map - must be a part of your solution
  - (2 pts) Filter - must be a part of your solution
  - (2 pts) List Comprehension - must be a part of your solution
  - (2 pts) Data Structures (dictionary, set, etc) - must be a part of your solution
  - (2 pts) Variable Positional Argument - must be a part of your solution
  - (2 pts) Yield - Does not have to be a part of your solution

- (12 points) Successful implementation of 3(undergraduate)/4(graduate) problems.
- (3 points) The execution time and style of your code including naming, comments, etc.
- (5 bonus points) Achieve a very good elapsed time.

Requirements:

- On your laptop, add a new folder inside your **CSE465_565** folder, and call it **Homework5**.
- Download **hw5.py**, **zipcodes.txt**, **zips.txt**, **cities.txt**, and **states.txt** from canvas and put them inside the **Homework5** folder.

Instructions:

1. In your project, you are free to create as many classes within the **hw5.py** file or across multiple files as you need. However, ensure that the **hw5.py** file is the only one that contains a __**main**__ method. This specific setup is crucial because your instructor will use the **hw5.py** file to execute and evaluate your work.

2. The programs will process the datafile **zipcodes.txt** which contains over 81K lines of text, one line for each geographic location in the US territories.
   - The first line of the file is the header line that explains each column of the data.
   - Except the first line, all other lines are tab-separated.

3. When **hw5.py** is run, it should create the following output files:
   - **CommonCityNames.txt**: contains all the common city names that appear in all of the states listed in **states.txt**.
     - i. Each city in the file must be unique, no duplicate names allowed.
     - ii. The cities should appear in sorted order.
     - iii. The cities should appear one per line.
   - **LatLon.txt**: for each zip code listed in **zips.txt**, there will be a corresponding line of output. Each output line will list the zip code's latitude and longitude.
     - i. The latitude and longitude must be separated by a space on each line.
     - ii. If a zip code has multiple entries, provide the first one listed in zipcodes.txt.
   - **CityStates.txt**: For each city listed in **cities.txt**, there will be a corresponding line of output. Each output line will list the states containing that city name.
     - i. The states must be separated by a space.
     - ii. The states should appear in sorted order.
     - iii. Each state should be listed only once for each city.
   - **grad.txt**: **(CSE565 only)** Display each of the following pieces of information - one per line
     - i. The two geographically most distant zipcodes - space separated
     - ii. The zip code with the largest population
     - iii. The zip code with the smallest population
     - iv. The zip code with the largest per capita wages
     - v. The zip code with the smallest per capita wages
     - vi. The city with the largest total population (e.g., Brooklyn, NY).

4. In order to get full points you must use features discussed in the class as following:
   - Use of Object-Oriented Programming :
     - i. Design a logical and practical hierarchy of classes with at least one superclass and a subclass. This must be a part of your solution.
     - ii. Override an operator, which could be either a logical operator, or a mathematical operator. This does not have to be a part of your solution.
     - iii. Override a method. This does not have to be a part of your solution.

- Lambda. This must be a part of your solution.
- Map. This must be a part of your solution.
- Filter. This must be a part of your solution.
- List Comprehension. This must be a part of your solution.
- Proper use of Data Structures from python library. This must be a part of your solution.
- Variable Positional Argument. This must be a part of your solution.
- Yield. Does not have to be a part of your solution.

5. Clarity and correctness is the primary consideration when writing your codes.
6. Performance is the secondary concern; however, using obviously wasteful and brute force approaches will result in a loss of points.
7. You can compile and execute your code via Command Line (CMD)/Terminal using the following commands:
   - python3/py/python   hw5.py
   - Just ensure that you are using Python 3.10 or higher. Preferably version 3.12.2
8. Commit and push your work right after any small progress. Your Git should demonstrate the progression of your work.
   - Adding or removing empty lines or spaces in one commit, does not add authenticity to your work.
   - Committing and pushing large volumes of code within a very short time frame, such as minutes or seconds apart, can raise doubts about the work's originality. Producing significant amounts of code this quickly is unrealistic and may indicate that the code is copied from another source. Genuine coding usually requires more development time and deliberate, thoughtful submission of changes.
9. The execution time and style of your code, including indentation and comments, should be acceptable. Brute force approaches which lead to very slow programs, and a lack of clarity will result in a loss of points.

Bonus Points:

To earn 5 bonus points on your assignment, ensure your code achieves a competitive runtime. Your runtime will be compared against others, and if it ranks among the top 15 (with some margin of error), you will receive the full bonus points. At the beginning of your **hw5.py** file, add a special comment indicating whether you want your runtime to be compared for bonus points. Use the following format:

// => I'm competing for BONUS Points <=

- Failing to highlight your participation to achieve bonus points by not adding the required comments at the top of your file will result in 0 bonus points.

By default, these 5 bonus points will be added to your overall Homework grade. However, you have the option to request that these points be applied to either your Quizzes or Exams instead. Please note, these points cannot be applied to Attendance. To make this request, simply email your instructor specifying where you would like your bonus points to be added.

Test your program:

Come up with more tests and test your program comprehensively. Your code will be tested separately with different text files.

Proper testing requires time. Begin by considering various scenarios and delving into the specifics. Eventually, you'll need to come up with a way to validate your solutions using the data at hand. Remember, thorough testing demands quality time devoted at the end to ensure your code functions as intended. So, as you manage your schedule for this assignment, make sure to allocate sufficient time for testing. The consequence of lack of proper testing will result in losing points.

Submission:

Submit the GitHub/GitLab URL of the project **Homework5**.
Inside this folder there should be the following files:
1. **hw5.py** and other complimentary classes.
2. All the input and output text files.