Assignment 7 - Analyze Social Network Data

For this assignment, you will write a program to analyze a social network dataset and print results to the terminal window in both written/tabular and graphical form; and write a brief report describing your analysis and results.

The dataset Pew_Survey.csv consists of a first row containing column names followed by 1372 data lines. The data come from an Internet Core Trends survey conducted by the Pew Research Center in January/February 2019. The column and data field information are included in the document: Pew_Survey_data.pdf on the Canvas site. Be sure to check out this document, as the data may not represent what you think – especially for responses 8 and 9, e.g.

There are many possibilities for analyzing this data, both for business or inter-personal studies. For this assignment, you will design and write a program to analyze the data, and interpret the results. You must include both written/tabular results and also plots or other visualizations of your results. There is a sample plot program included in this module on the Canvas site, but there are many other interesting ways to present data visually.

NOTE
An important part of this assignment -- and part of your grade -- is the purpose and design of the analysis you decide to do. It should not just be a fancy data dump, it must have a point to it. For example, if you were marketing a product, you might use the data to determine your target audience and tailor your marketing campaign accordingly. If you were a politician, you might want to know where to reach your supporters or non-supporters. If you worked for a social service agency, you might be interested in the demographics of users on certain platforms. Be creative and come up with a (non-trivial) idea. I'd like you to have some fun with this. A good approach is to put yourself in the shoes of the designer/decision maker and think about what information you would need in order to be successful in your position.
NOTE

Besides the program, include a report that contains the following *labeled* sections:

PURPOSE: Be sure you address these issues:
   • What are you trying to show with your chosen analysis?
   • Why are you interested in showing this?
   • Why is this important?
   • How would you use your results? *Be specific:* for example, what conclusions are you hoping to be able to draw from analyzing the results, and what would that enable you to do?
You must justify the importance of the study you are doing -- it may not be something trivial.

INPUT:
OUTPUT:
WHAT THE PROGRAM DOES: How did you conduct the analysis?
RESULTS: Describe and discuss the results of your analysis.
ANY ADDITIONAL INFORMATION. (Anything you want to share -- students often talk about special issues that came up and how they addressed them, for instance. You may not have anything you want to include, and that is fine.)

The report should be submitted in .pdf or Word format, in a document **labeled with your last name**, eg: Dugas_HW9_Report.pdf.

Zip your report, program, and a **screen shot of any output** into a zip (compressed) file that is labeled (both inner and zip folders) with your last name, eg: Dugas_HW3.zip (I.E. it should unzip to a folder that also has your last name.) Submit in Canvas. ***Do not submit the Pew_Survey.csv dataset with your assignment***. *Screen shots may be included in your report, but also must be included separately in the zip file.*

PROGRAMMING GUIDELINES:

Do all processing using one program only.

All programming must be in python 3 unless otherwise arranged with the instructor.

If you use external python packages, please note that in your program comments in the run instructions.

Your program should be named: **yourlastname.py**

Programs will be screened for plagiarism. If you "borrow" code, be sure to document the details of the source; otherwise it will be considered plagiarism and result in a zero grade for the assignment. Borrowed code will not count toward your grade, only original code will be considered.

Programs should employ good programming practices. An example is the use of descriptive variable and function names.

Input: Do not prompt for a file name; hard-code the ***Pew_Survey.csv*** name into your program.

Output: All output must have titles and legends that make it clear what is being shown.

Annotation and Comments: ***IMPORTANT***

•	Program header must include your name and assignment information (use comments).
•	Comments must also be used at the beginning of the program to give an overall description of the purpose of the program.
•	Comments must also include detailed running instructions to run in a terminal window.
•	Comments should also be used throughout the code to explain what it is doing. It should be possible to re-create your program based on the comments alone. Poorly commented programs will receive poor grades.