Assignment 9 - 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 FB_data.csv consists of a first row containing column names followed by 99903 data lines. The data are a few years old, from 2013. Source: kaggle.com Owned/maintained by Sheena Batra
Columns in each data line:

A.	userid	unique identifier
B.	age	age in 2013
C.	dob_month	date of birth - month
D.	dob_day	date of birth - day
E.	dob_year	date of birth - year
F.	gender	"male" or "female" [or "NA"]
G.	tenure	number of days user has been on facebook
H.	friend_count	how many friends user has
I.	friendships_initiated	how many friendships were requested by the user
J.	likes	how many posts user has liked
K.	likes_received	how many of the user's posts were liked by other users
L.	mobile_likes	portion of column J from using mobile app
M.	mobile_likes_received	portion of column K from using mobile app
N.	www_likes	portion of column J from using FB website
O.	www_likes_received	portion of column K from using FB website

NOTE: You'll need to check for valid data entries in the fields you are using. For example, a quick check of the file revealed 175 "NA" entries for gender.

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.

Besides the program, include a report that contains the following information:

1. What are you trying to show with your chosen analysis?
2. Why are you interested in showing this? Why is this important?
3. How did you conduct the analysis?
4. Describe and discuss the results of your analysis.

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, code, and a screen shot of your output into a zip (compressed) file and submit in Canvas. ***Do not submit the FB_data.csv dataset with your assignment***.

PROGRAMMING GUIDELINES:

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.

Programming can be done in a variety of languages. Programs should employ good programming practices. An example is the use of descriptive variable and function names.

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.