# **Part 1. Answer only 1 Question** (If you solve more the top 1 mark will be awarded) **Question 1**

Data description

This random sample of dataset provide here.

https://archive.ics.uci.edu/ml/datasets/adult

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 48842 | **Area:** | Social |
| **Attribute Characteristics:** | Categorical, Integer | **Number of Attributes:** | 14 | **Date Donated** | 1996-05-01 |
| **Associated Tasks:** | Classification | **Missing Values?** | Yes | **Number of Web Hits:** | 523007 |

# Data Set Information:

Extraction was done by Barry Becker from the 1994 Census database. A set of reasonably clean records was extracted using the following conditions: ((AAGE>16) && (AGI>100) && (AFNLWGT>1) && (HRSWK>0))  
  
Prediction task is to determine whether a person makes over 50K a year.

# Attribute Information:

Listing of attributes:  
  
>50K, <=50K.  
  
age: continuous.  
workclass: Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.  
fnlwgt: continuous.  
education: Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool.  
education-num: continuous.  
marital-status: Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.  
occupation: Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces.  
relationship: Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.  
race: White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black.  
sex: Female, Male.  
capital-gain: continuous.  
capital-loss: continuous.  
hours-per-week: continuous.  
native-country: United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinadad&Tobago, Peru, Hong, Holand-Netherlands.

# Your task is:

1. Define the missing values and update it with the mean for numerical values and the highest frequency for categorical values or delete it.
2. Split the cleaned data to two datasets where (<=50K & >50k) and get the flowing foreach category
   1. get the max, min, and the mean age.
   2. get the highest and lowest frequency for education level.
   3. get the max, min, and mean hours per week.
3. Is male’s income >50k?
4. Is the most frequency for male’s income in >50k?
5. Is the most frequency for Husband’s income in >50k?

# Question 2

**Write a class named Employee that holds the following data about an employee in attributes: name, ID number, department, and job title.**

Once you have written the class, write a program that creates three Employee objects to hold the following data:

Name: Ayman Nabil, Ahmed Saeed, Esraa Mohamed

ID Number: 47899, 39119, 81774

Department: Accounting, IT, Manufacturing

Job Title: Vice President, Programmer, Engineer

The program should store this data in the three objects, then display the data for each employee on the screen. Finally write the data to a csv file.

**Then Create a program that reads Employee data from file employees.csv then stores Employee objects in a dictionary.**

Use the employee ID number as the key. The program should present a menu that lets the user perform the following actions:

1. Look up an employee in the dictionary
2. Add a new employee to the dictionary
3. Change an existing employee’s name, department, and job title in the dictionary
4. Delete an employee from the dictionary
5. Quit the program

When the program ends, it should save the dictionary data to a file. Each time the program starts, it should try to load the pickled dictionary from the file. If the file does not exist, the program should start with an empty dictionary.

# **Question** **3**

Write a Person class with an instance variable, and a constructor that takes an integer, as a parameter. The constructor must confirm the argument passed as is not negative; if a negative argument is passed, the constructor should set to 0 and print Age is not valid, setting age to 0. In addition, you must write the following instance methods:

1. *yearPasses ()* should increase the instance variable by 1.
2. *greetings ()* should perform the following conditional actions:
   * If less than 12, print You are young.
   * If more than or equal 12 and less than 18, print You are a teenager.
   * Otherwise, print You are old and wise.

Write a main function to test your class. Once the age is entered you should print the greeting message, then ask the user how many years passed. Use *yearPasses ()* method to add those years to the person age. Then print the greetings one more time.

***Test Case 0*:**   
**Input** 0,3

**Output** prints the "Age is not valid..." message followed by the young message. Three years pass and, so we print the young message.

***Test Case 1:***

**Input** 11, 1  
**Output** code should print that the person is young. 1 year pass and, so we print that the person is now a teenager.

***Test Case 2:***

**Input** 18, -1, 1  
**Output** code should print that the person is a old and wise. Then ask for input again as negative numbers are not allowed, 1 year pass and, so we print that the person is old and wise.

***Test Case 3:***   
Think of test case of you choice here, write the case in the comments before the code that executes it.