Program1:

To Demonstrate how to utilize the pandas packages to clean,manipulate and anlayze data

**1- import and read a CSV file**

**2-To Generate a basic understanding of a given data.**

1. First 5 rows, last 5 rows of data
2. How to check the basic information of the data
3. How to extract the shape of the data
4. How to know the unique values of the marital status field based on the column
5. Which method is used to make it consistent data( widow and widowed )are two different naming for the same category on the column values
6. How to check for duplicates and null values in the whole dataset

**3- select and filter data based on conditions**

1. How to select a subset of the data points(Birthdate, Education, and Income of every customer) from the data frame.
2. Using loc() and iloc() methods to retrieve the first seven data points.
3. How to filter data using the loc() and isin() methods.(Note: we choose the variable of interest and we select the categories )
4. In our data, that satisfies two conditions such as choosing the customers with an income higher than 75,000 and with a master’s degree(using python operators) and display the output.

**4- apply various data operation tools such as creating new variables or changing data types**

We can apply different operations on the dataset using Pandas such as

1. setting a new index with the variable of our interest using the .set\_index() method
2. sorting the data frame by one of the variable using .sort\_values() with ascending or descending order;
3. creating a new variable which could be the result of a mathematical operation such as sum of other variables
4. changing the datatype of variables into datetime or integer types
5. determining the age based on year of birth
6. creating the week date (calendar week and year) from the purchase date

**5- perform data aggregation using group by and pivot table methods**

After we created new variables, we can further aggregate  and to analyze data by groups,

1. To apply groupby()method to find the mean of income ,recency,number of web and store purchases by educational group.
2. To apply pivot\_table()method to find the aggregated sum of purchases and mean of recency per education and marital status group.