Assignment-3

Download the dataset (in csv format) from the following link:

https://www.kaggle.com/jahias/microsoft-adventure-works-cycles-customer-data

The Marketing department of Adventure Works Cycles wants to increase sales by targeting specific customers for a mailing campaign. The company's database contains a list of past customers and a list of potential new customers. By investigating the attributes of previous bike buyers, the company hopes to discover patterns that they can then apply to potential customers. They hope to use the discovered patterns to predict which potential customers are most likely to purchase a bike from Adventure Works Cycles.

Q.1 Based on Feature Selection, Cleaning, and Preprocessing to Construct an Input from Data Source

(a) Examine the values of each attribute and select a set of attributes only that would affect to predict future bike buyers to create your input for data mining algorithms. Remove all the unnecessary attributes. (Select features just by analysis).

(b) Create a new Data Frame with the selected attributes only.

(c) Determine a Data value type (Discrete, or Continuous, then Nominal, Ordinal, Interval, Ratio) of each attribute in your selection to identify preprocessing tasks to create input for your data mining.

Q.2 Is there a significant difference in the average income of customers who purchased a bike vs. those who did not?

Q.3 What percentage of customers in each commute distance category purchased a bike?

Q.4 Does age distribution differ between bike buyers and non-buyers?

Q.5 Is the proportion of bike buyers different across marital status categories?