Required Assignment 2.3: Data Preparation and Basic EDA

Started: Jun 4 at 7:53pm

Quiz Instructions

🎯 **Learning Outcomes Addressed**

4. Perform data preprocessing operations
5. Prepare data for analysis using the Data-Preproc App

Assignment Instructions

In this assignment, you will practice how to prepare datasets for analysis. Data usually comes in a variety of formats and may contain errors and missing values. In this assignment, you will deal with one such dataset.

Dataset Information:

The dataset is related to direct marketing campaigns of a Portuguese banking institution. These marketing campaigns were based on phone calls. Often, more than one contact to the same client was required to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed.

Refer to the **appendix** ⬇ **(https://participants.eruditus.com/courses/641/files/90380/download?download_frd=1)** for attribute information.

Data Source: bank_assignment.csv ⬇ **(https://participants.eruditus.com/courses/641/files/90352/download?download_frd=1)**

Task: Upload the dataset (bank_assignment.csv) to the Data-Preproc App ⬈ **(https://zerocodelearning.com/) and answer the following questions.**

This is a required assignment and counts towards programme completion.

Question 1	1 pts

Choose the correct separator for properly uploading the bank_assignment dataset in the Data-Preproc App.

Select the most appropriate option.

○ Comma

○ Colon

○ Tab

○ Semicolon

Question 2   1 pts

How many observations (rows) and variables (columns) does the dataset contain?

Type the most appropriate option in the given text field.

[   ] observations/rows

[   ] variables/columns

Question 3   1 pts

What is the overall percentage of missing values?

Select the most appropriate option.

○ 894

○ 1.16

○ 839

○ 2

Question 4   1 pts

How many numeric variables does this dataset contain?

Select the most appropriate option.

○ 17

○ 0

○ 7

○ 10

Question 5

1 pts

What is the percentage of missing values in the age variable? Please approximate the value to two decimal places.

Type the most appropriate option in the given text field.

[   ]

Question 6

1 pts

What is the maximum age of a person in the dataset? Please approximate the value to two decimal places.

Type the most appropriate option in the given text field.

[   ]

Question 7

1 pts

What is the mean balance? Please approximate the value to two decimal places.

Type the most appropriate option in the given text field.

[   ]

Question 8   1 pts

What is the percentage of unemployed customers? Please approximate the value to two decimal places.

Type the most appropriate option in the given text field.

[   ]

Question 9   1 pts

Out of all the customers, 43.4% customers have housing loan?

Select the most appropriate option.

○ True

○ False

Question 10   1 pts

Were most people contacted on cellular phones?

Select the most appropriate option.

○ True

○ False

Question 11　　　　　　　　　　　　　　　　　　1 pts

There are _____ such customers who were contacted between 15 to 20 times before this campaign.

Type the most appropriate option in the given text field.

○ 5

○ 4

○ 2

○ 3

Question 12　　　　　　　　　　　　　　　　　　1 pts

Which of the following month has lowest count as last contact month?

Select the most appropriate option.

○ Jan

○ May

○ Dec

○ Aug

Question 13　　　　　　　　　　　　　　　　　　1 pts

If we considered a complete case strategy for treating missing values, how many observations may we lose?

Select the most appropriate option.

○ 839

○ 880

○ 832

○ 849

Question 14   **1 pts**

Which of the following steps may create new variable?

Select the most appropriate option.

○ Data Imputation

○ All of the above

○ Data Transformation

○ Dummy Encoding

Question 15   **1 pts**

How many new variables will be created when you perform the following steps?

Step 1: Upload data

Step 2: Impute all missing continuous variables with mean and keep original columns

Step 3: Go to Data Transformation Tab and create standardized pdays, duration and balance

Step 4: Go to Dummy Encoding Tab and convert job, marital status and education to one hot encoded variable (remove first dummy & original column)

Select the most appropriate option.

○ 19

○ 30

○ 17

○ 21

Not saved   Submit Quiz