

Answers 5.4

Step 2. Data quality assessment & descriptive analysis

Following steps have been taken to assess the quality of Pig E Bank's client data set.

- a) Data consistency check: typos and inconsistent entries have been corrected
- b) Duplicate & missing data check: no duplicate entry was found. Some missing values <figure 1> were detected but remained intact due to its small volume.
- c) Descriptive analysis <figure 2> was carried out.

Attribute	Missing Values
Row Number	0
Customer ID	0
Last Name	1
Credit Score	3
Country	0
Gender	1
Age	1
Tenure	0
Balance	0
NumOfProducts	0
HasCrCard?	0
IsActiveMember	0
Estimated Salary	2
ExitedFromBank?	0

<Figure 1. Missing value check result>

Attribute	Count	Mode	Mean	Std	Min	25%	50%	75%	Max
Row Number	991								
Customer ID	991								
Last Name	990	-	-	-	-	-	-	-	-
Credit Score	988		648.51	98.2	376	579.75	653.5	721	850
Country	991	FR	-	-	-	-	-	-	-
Gender	990	Male	-	-	-	-	-	-	-
Age	990		38.68	10.96	2	32	37	44	82
Tenure	991		5.07	2.93	0	2	5	8	10
Balance	991		78,002.72	62,799.24	0	0	98,668.18	129,461.85	213,146.20
NumOfProducts	991		1.52	0.57	1	1	1	2	4
HasCrCard?	991		0.71	0.46	0	0	1	1	1
IsActiveMember	991		0.51	0.5	0	0	1	1	1
Estimated Salary	989		98,574.54	57,224.72	371.05	48,559.19	98,368.24	146,326.45	199,725.39
ExitedFromBank?	991		0.21	0.4	0	0	0	0	1

<Figure 2. Descriptive analysis>

Step 3.

Key observations of the data set

- **Credit Score:** The average credit score is 648.51, with scores ranging from 376 to 850.
- **Age:** The average age is 38.68 years, with the youngest customer being 2 and the oldest 82.
- **Balance:** The average account balance is \$78,002.72, with balances ranging from \$0 to \$213,146.20.
- **NumOfProducts:** Most customers have between 1 and 2 products, with an average of 1.52 products per customer.
- **Has Credit Card:** Approximately 70.64% of customers have a credit card.
- **IsActiveMember:** About 50.76% of customers are active members.
- **Estimated Salary:** The average estimated salary is \$98,574.54, with a wide range from \$371.05 to \$199,725.39.
- **ExitedFromBank?:** 20.58% of customers have exited the bank.

Common characteristics of customers who exited from bank

Out of 204 customers who exited from bank,

- About 87% of them had balance below \$150,000.
- About 85% of them had credit score between 500 and 700.
- About 70% of them had (at least) one credit card.
- About 70% of them had inactive member status.
- About 65% of them were in their 30s and 40s.
- About half of them earned expected salary of \$100,000 or less.

Step 4. Decision Tree

Which customers will leave us?

