

# Chat History

--- User ---

What is the data about?

--- Assistant ---

Based on the data summary provided, the data is about \*\*customer information, likely from a bank or financial institution, with a focus on predicting customer churn or retention.\*\*

Here's what the data includes:

- \* \*\*Size:\*\* The dataset contains \*\*10,000 rows\*\* and \*\*14 columns\*\*.
- \* \*\*Customer Demographics and Activity:\*\*
  - \* \*\*Identifiers:\*\* `RowNumber`, `CustomerId`, `Surname`.
  - \* \*\*Demographics:\*\* `Age` (ranging from 18 to 92 years, with a mean of 38.9), `Gender` (54.57% Male, 45.43% Female), and `Geography` (primarily France (50.14%), Germany (25.09%), and Spain (24.77%)).
  - \* \*\*Financial Behavior:\*\* `CreditScore` (ranging from 350 to 850, mean 650.5), `Tenure` (how long a customer has been with the institution, 0-10 years), `Balance` (customer's account balance, with a significant portion having 0 balance, mean 76,485.89), `NumOfProducts` (number of products the customer uses, 1-4), `HasCrCard` (whether the customer has a credit card, 70.55% do), `IsActiveMember` (whether the customer is an active member, 51.51% are), and `EstimatedSalary` (estimated salary of the customer).
  - \* \*\*Target Variable:\*\* `Exited` (a binary variable indicating whether the customer has exited or churned, with 20.37% of customers having exited).