

Bank Customer Churn Project

It consists of 10000 observations and 14 variables. Independent variables contain information about customers. Dependent variable refers to customer abandonment status.

Variables:

1. **RowNumber** — corresponds to the record (row) number and has no effect on the output. This column will be removed.
2. **CustomerId** — contains random values and has no effect on customer leaving the bank. This column will be removed.
3. **Surname** — the surname of a customer has no impact on their decision to leave the bank. This column will be removed.
4. **CreditScore** — can have an effect on customer churn, since a customer with a higher credit score is less likely to leave the bank.
5. **Geography** — a customer's location can affect their decision to leave the bank. We'll keep this column.
6. **Gender** — it's interesting to explore whether gender plays a role in a customer leaving the bank. We'll include this column, too.
7. **Age** — this is certainly relevant, since older customers are less likely to leave their bank than younger ones.
8. **Tenure** — refers to the number of years that the customer has been a client of the bank. Normally, older clients are more loyal and less likely to leave a bank.
9. **Balance** — also a very good indicator of customer churn, as people with a higher balance in their accounts are less likely to leave the bank compared to those with lower balances.

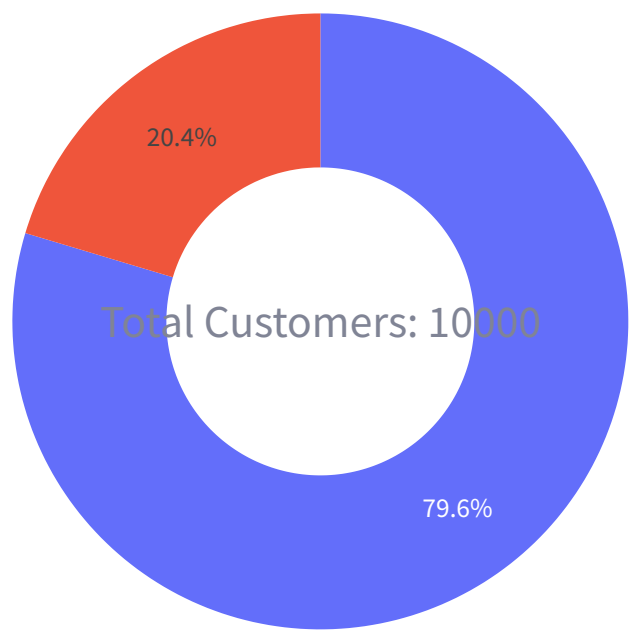
10. **NumOfProducts** — refers to the number of products that a customer has purchased through bank.
11. **HasCrCard** — denotes whether or not a customer has a credit card. This column is also relevant because people with a credit card are less likely to leave the bank. (0=No, 1=Yes)
12. **IsActiveMember** — active customers are less likely to leave the bank, so we'll keep this. (0=No, 1=Yes)
13. **EstimatedSalary** — as with balance, people with lower salaries are more likely to leave the bank compared to those with higher salaries.
14. **Exited** — whether or not the customer left the bank. This is what we have to predict. (0=No, 1=Yes)

Sample Data

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure
0	1	15,634,602	Hargrave	619	France	Female	42	2
1	2	15,647,311	Hill	608	Spain	Female	41	1
2	3	15,619,304	Onio	502	France	Female	42	8
3	4	15,701,354	Boni	699	France	Female	39	1
4	5	15,737,888	Mitchell	850	Spain	Female	43	2

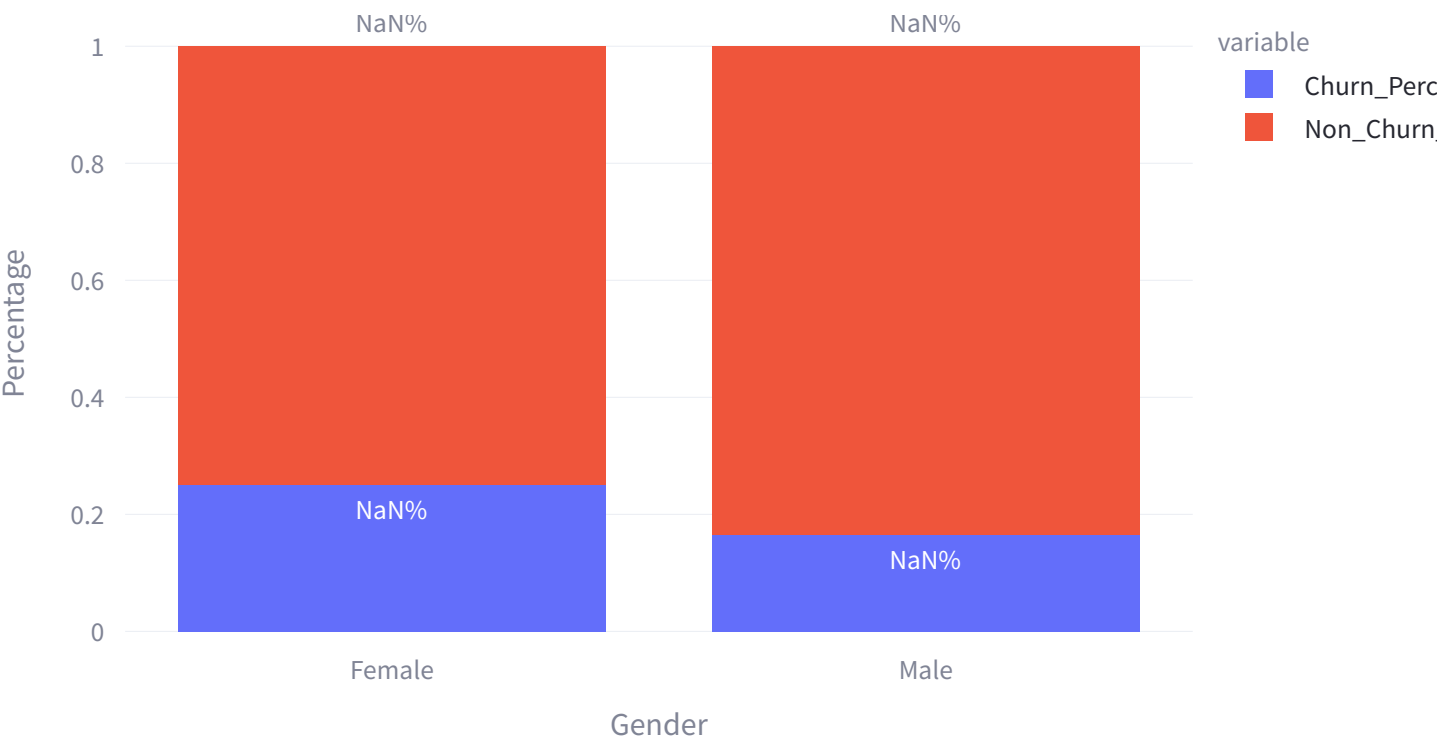
Bank Customer Churn

Churn vs. Non-Churn Customers



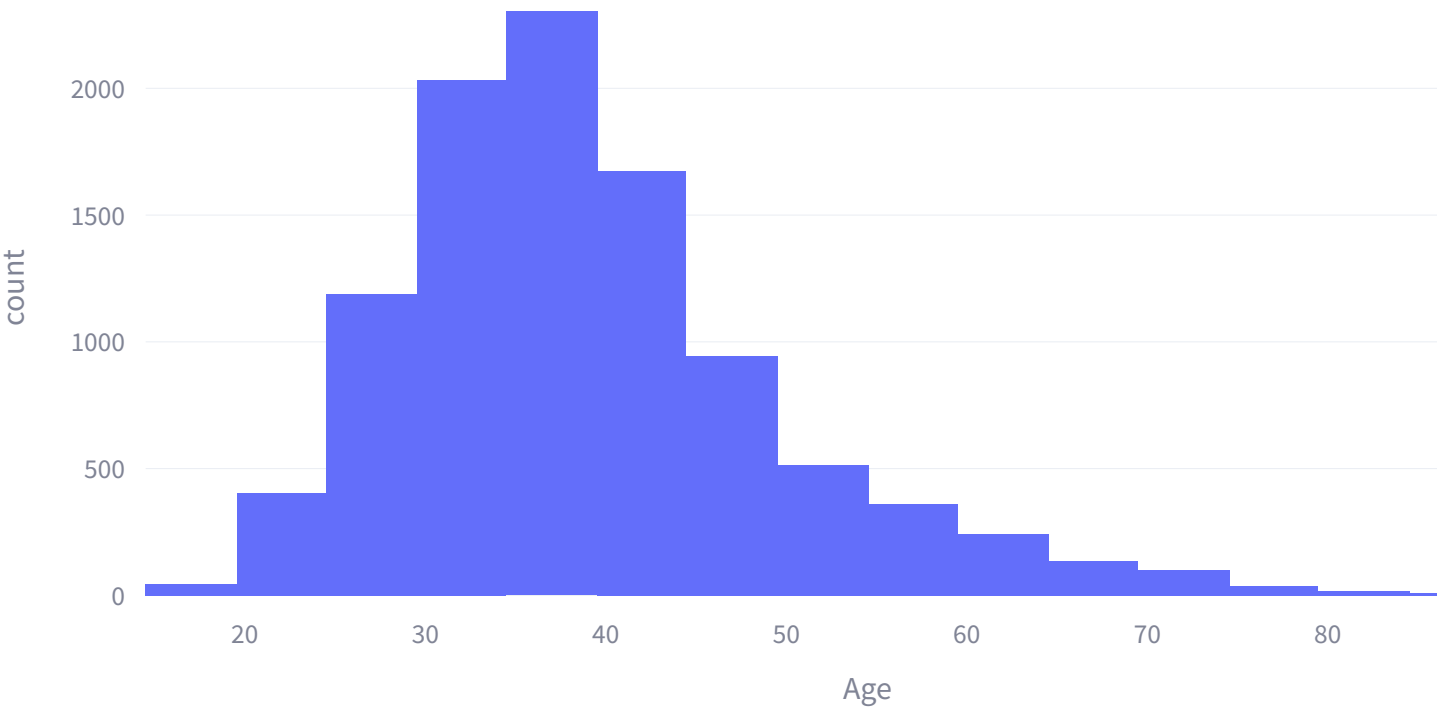
Bank Customer Churn vs Age

Churn Percentage by Gender

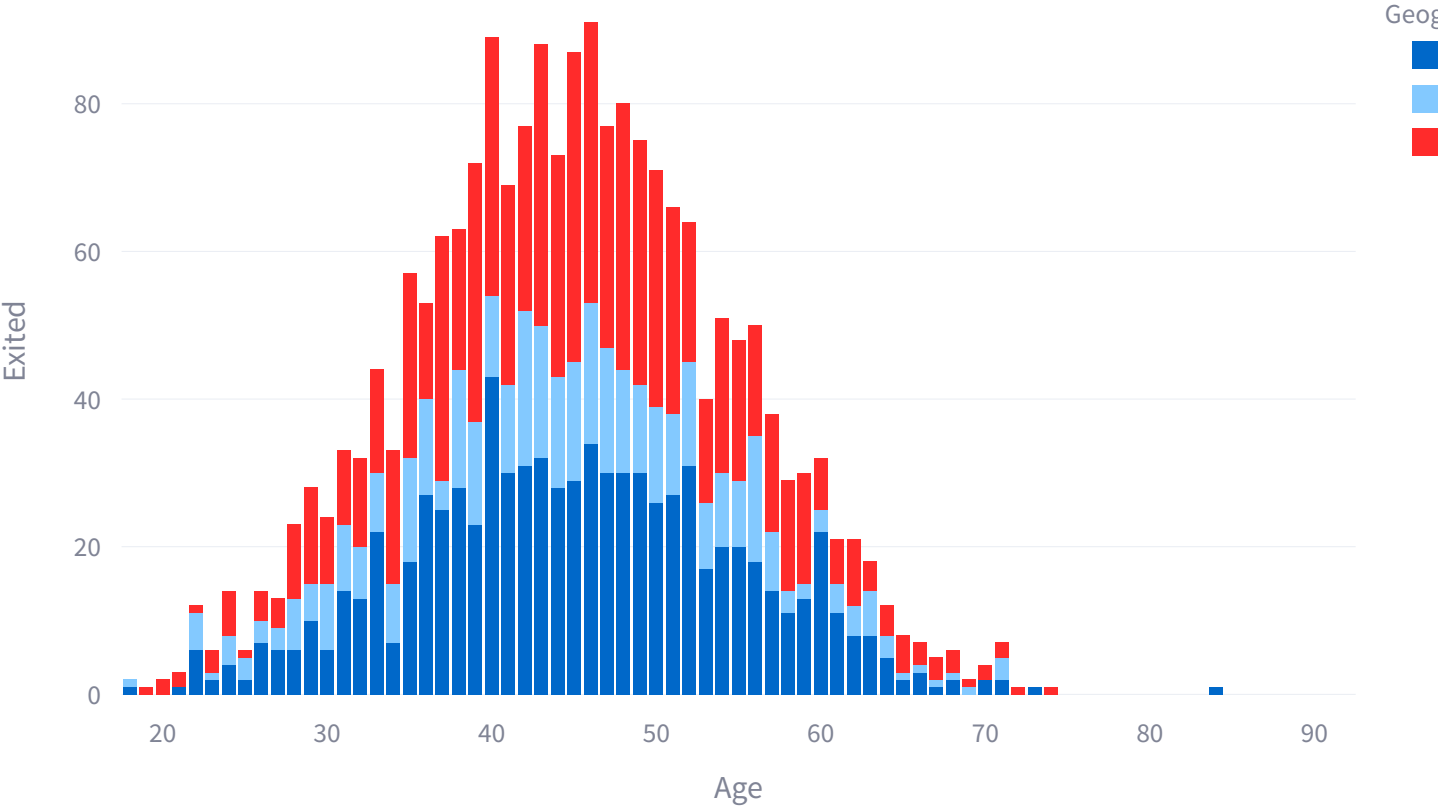


Age Distribution

Age Distribution

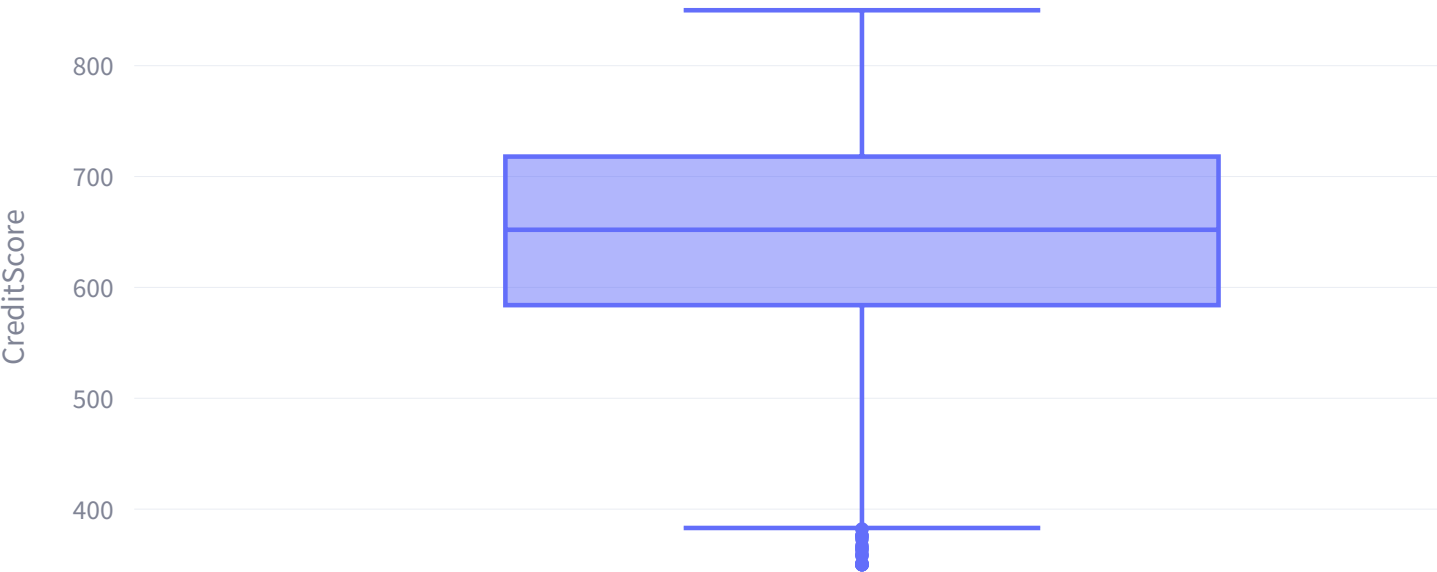


Geography vs age

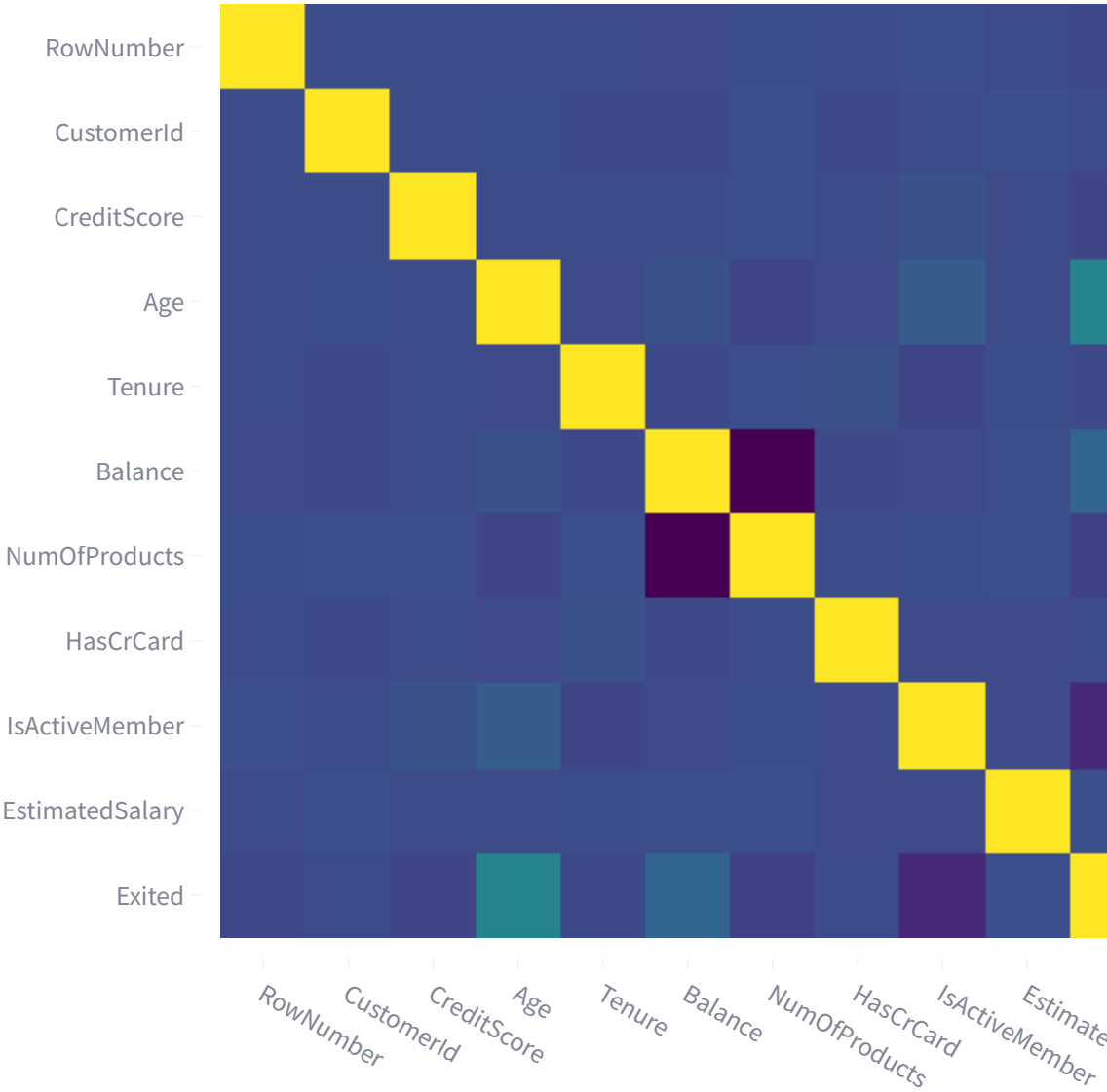


Credit Score Distribution

Credit Score Distribution



Correlation Matrix Heatmap



Made with Streamlit