# Store X Analysis Tool

Store X is a small market offering food products to their customers. At times, Store X carries out promotional campaigns. These campaigns are primarily done to boost sales and increase revenues. However, the success of these campaigns is dependent on the number of customers who respond to them; therefore, the identification of these customers is important to the success of the campaign.

To identify these customers, we can use the **Store X Analysis Tool**, which provides certain metrics related to the store, and carries out a machine learning algorithm to predict whether a customer is likely to respond to a campaign or not, given certain consumer information.

Once the tool is opened, the first page that is presented is the home page (shown in **Figure 1**), which provides a brief overview of the insights that the tool generates, in addition to a location where a dataset must be uploaded.

The specific dataset that must be uploaded is one that contains general customer information, information related to sales and purchases, and information related to Store X's previous promotional campaigns.

When the dataset is fully uploaded, the user will be redirected to the main analysis tool, which consists of four pages: the data overview page, the dashboards page, the prediction models page, and the final prediction page.

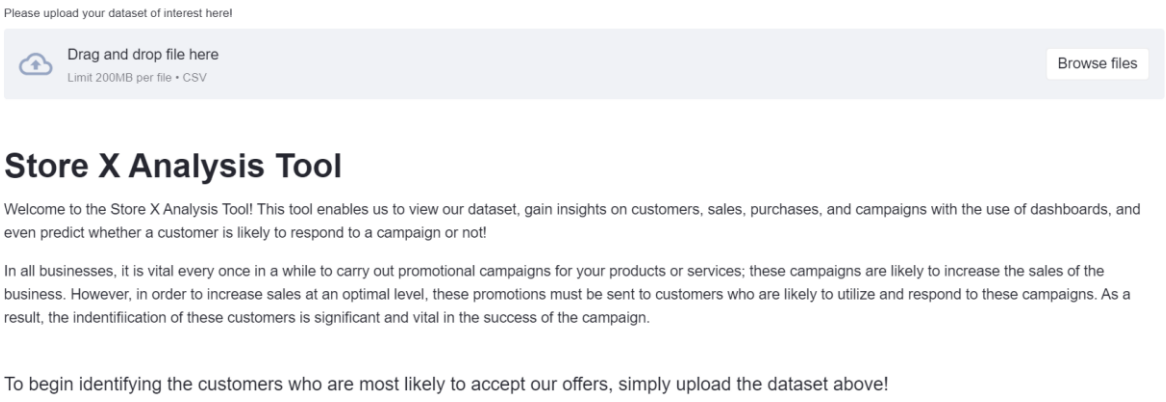
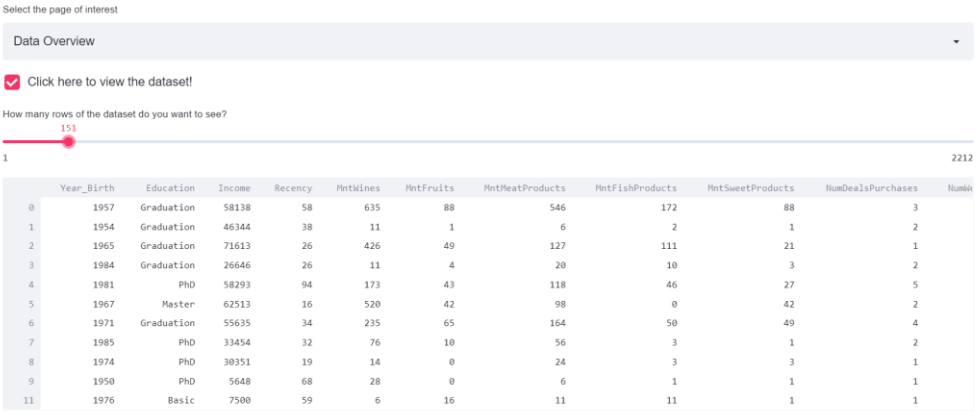



**Figure 1**

***Data overview page***

The data overview page (shown in **Figure 2**) provides a tool that simply enables the user to view the dataset. There is a checkbox that shows the data when the checkbox is enabled. When the checkbox Is enabled, the user can immediately see only the first record of the dataset. The user can decide how many records they want to see using a slider. Each record reflects a different customer; since the dataset has 2,212 records, we have 2,212 customers.




| | Year_Birth | Education | Income | Recency | MntWines | MntFruits | MntMeatProducts | MntFishProducts | MntSweetProducts | NumDealsPurchases | Num |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1957 | Graduation | 58138 | 58 | 635 | 88 | 546 | 172 | 88 | 3 | |
| 1 | 1954 | Graduation | 46344 | 38 | 11 | 1 | 6 | 2 | 1 | 2 | |
| 2 | 1965 | Graduation | 71613 | 26 | 426 | 49 | 127 | 111 | 21 | 1 | |
| 3 | 1984 | Graduation | 26646 | 26 | 11 | 4 | 20 | 10 | 3 | 2 | |
| 4 | 1981 | PhD | 58293 | 94 | 173 | 43 | 118 | 46 | 27 | 5 | |
| 5 | 1967 | Master | 62513 | 16 | 520 | 42 | 98 | 0 | 42 | 2 | |
| 6 | 1971 | Graduation | 55635 | 34 | 235 | 65 | 164 | 50 | 49 | 4 | |
| 7 | 1985 | PhD | 33454 | 32 | 76 | 10 | 56 | 3 | 1 | 2 | |
| 8 | 1974 | PhD | 30351 | 19 | 14 | 0 | 24 | 3 | 3 | 1 | |
| 9 | 1950 | PhD | 5648 | 68 | 28 | 0 | 6 | 1 | 1 | 1 | |
| 11 | 1976 | Basic | 7500 | 59 | 6 | 16 | 11 | 11 | 1 | 1 | |

**Figure 2**

## ***Dashboards page***

The dashboards page (shown in **Figure 3**) provides a tool that enables the user to view dashboards containing insights extracted from the uploaded dataset. There are four dashboards to choose from: the customer overview dashboard, the sales breakdown dashboard, the purchase breakdown & complaint dashboard, and the previous campaign performance dashboard.

### *Customer overview dashboard*

The customer overview dashboard highlights five main customer demographics: year of birth, education level, relationship status, income, and recency. The dashboard answers the following questions:

- What is the age distribution of our customers?
- How old are most of our customers?
- How old is our oldest customer?
- How old is your youngest customer?
- What level of education do most of our customers possess?
- What percentage of our customers are in a relationship?
- What percentage of our customers are not in a relationship?
- What is the median yearly income of our customers?
- How much does our lowest income customer earn?
- How much does our highest income customer earn?
- On average, how many days has it been since the most recent purchase of a customer

### *Sales breakdown dashboard*

The sales breakdown dashboard provides an overview on the sales of Store X. The dashboard answers the following questions:

- What product categories do we sell?
- What are our total sales?
- What product category makes up the highest percentage of sales?
- How much does each product category generate in sales?

*Purchase breakdown & complaints dashboard*

The purchase breakdown & complaints breakdown dashboard provides an overview on the purchases of customers and a brief overview on customer complaints. The dashboard answers the following questions:

- Where do customers purchase their products (online vs. physical)?
- What percentage of customer purchases were carried out in online and physical stores?
- How many customer purchases were carried out in online and physical stores?
- How many units have customers purchased?
- What are the chances of a customer visiting the online store and carrying out a purchase?
- What percentage of purchases were made on a discount?
- What is the percentage of customer complaints?

*Previous campaign performance dashboard*

The previous campaign performance dashboard provides an overview on Store X's previous promotional campaigns. The dashboard answers the following questions:

- What are the success rates of our previous campaigns?
- How many customers have responded to our previous campaigns?

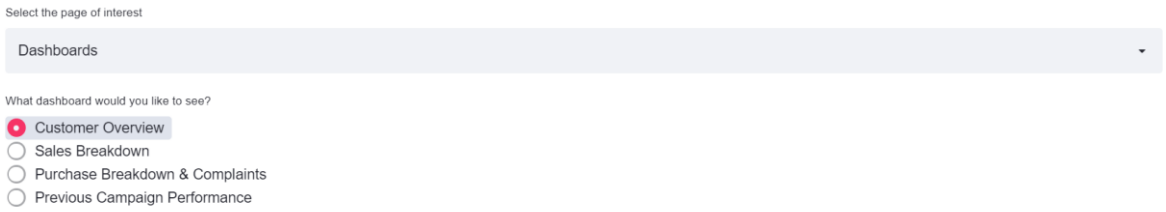



**Figure 3**

***Prediction model page***

The prediction model page (shown in **Figure 4**) provides a tool that enables the user to view specific supervised machine learning algorithms to predict whether a customer is likely to respond to a promotional campaign or not. Four supervised learning models are shown: logistic regression, K-nearest neighbors, decision trees, and random forests. The confusion matrix and the ROC curve are plotted for all models. In addition, the accuracy for all four models is shown.

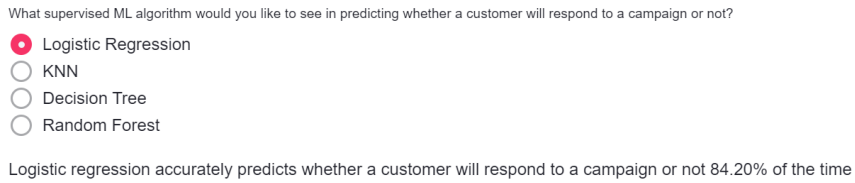



**Figure 4**

***Final prediction page***

The final prediction page (shown in **Figure 5**) provides a tool that enables the user to predict whether a customer is likely to respond to a promotional campaign or not, given certain customer information such as their year of birth, total amount spent, total purchases, etc. The random forest classifier was used to predict response.

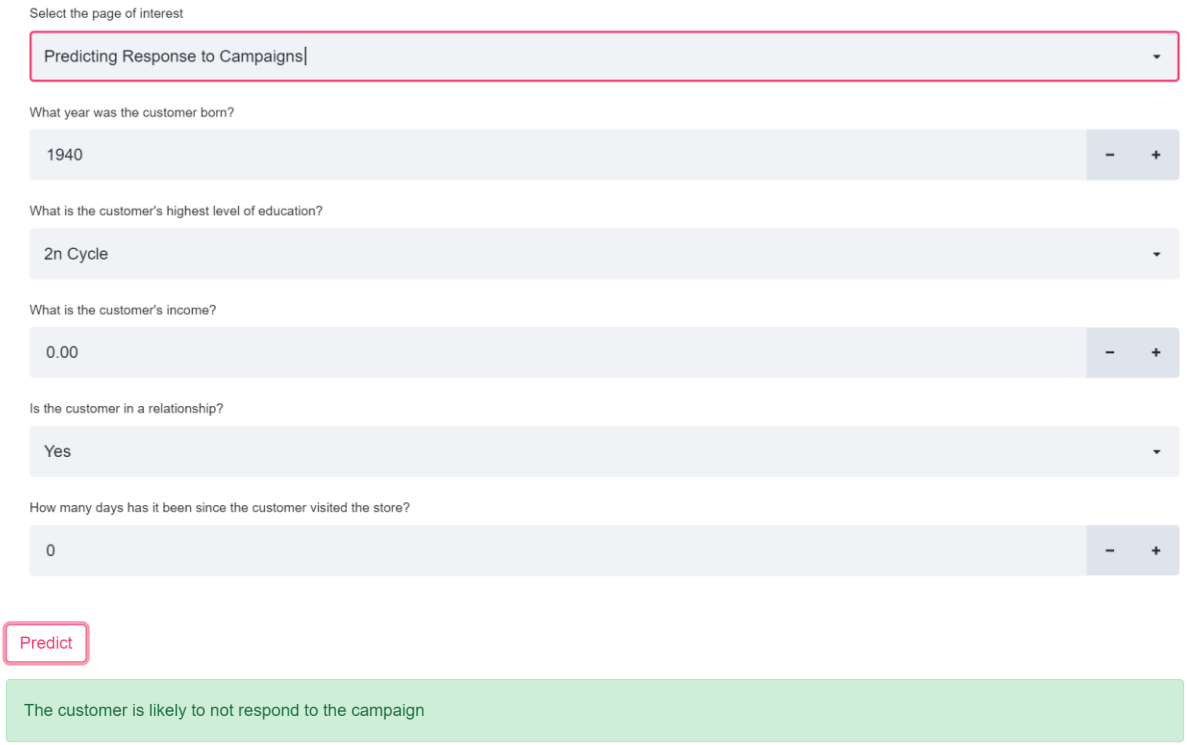



**Figure 5**