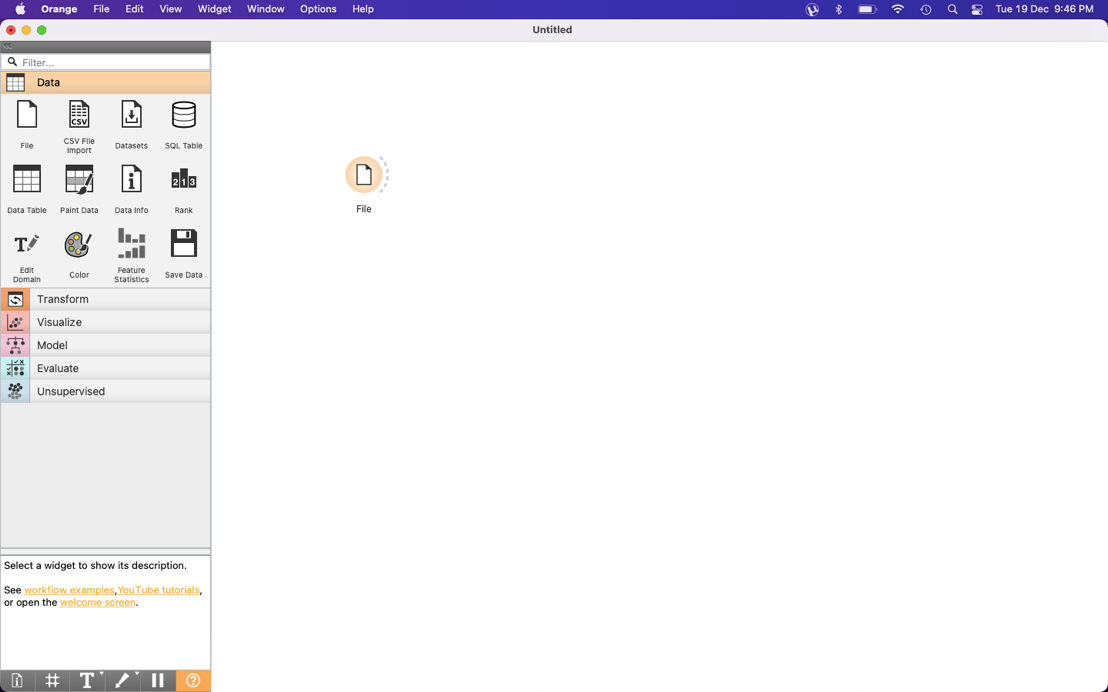
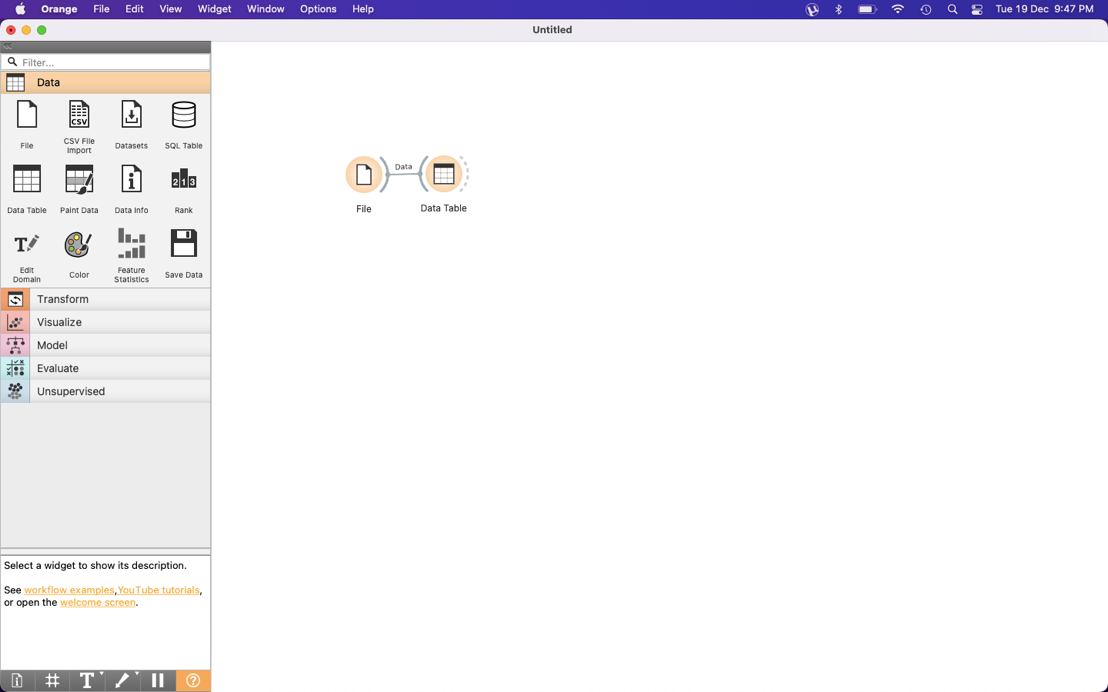
| **CO2** | **Apply pre-processing techniques and explore the using visualization techniques** |
| --- | --- |
| Task 3: | Perform Data Pre-processing, Data Analysis and Visualization for a given dataset. **Platform: Orange, Language: Python** |

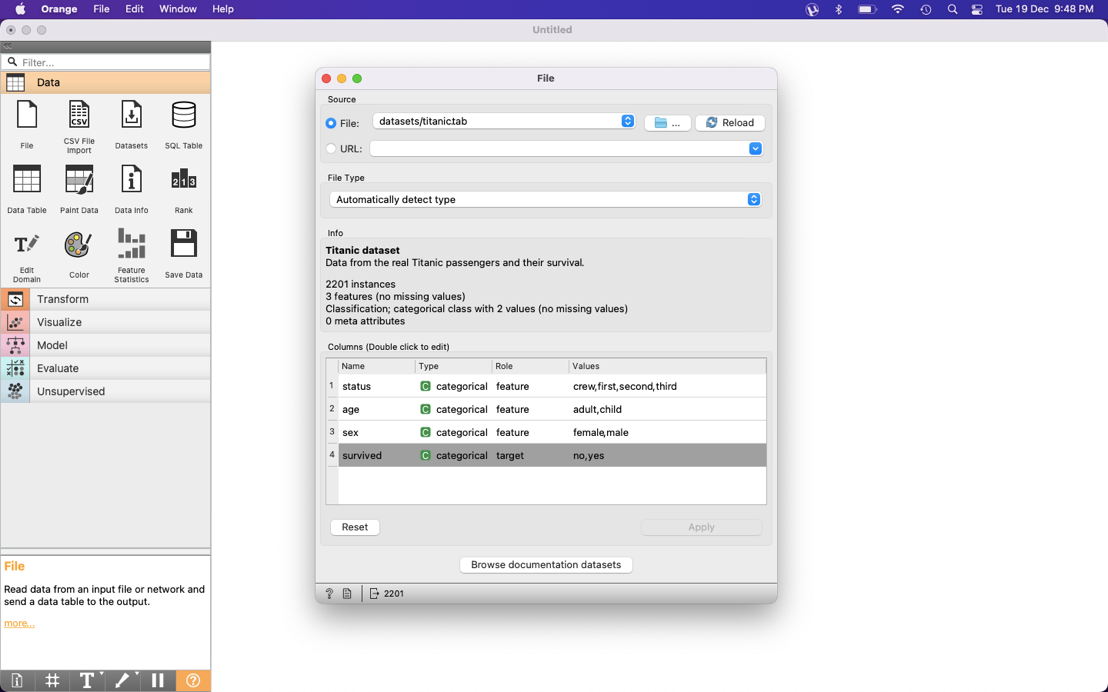
Step 1: Load the Dataset -> Drag and drop the "File" widget onto the canvas



Step 2: Connect the "File" widget to the "Data Table" widget.



Step3: Load your customer dataset using the "Browse" button in the "File" widget.   
 ( titanic dataset loaded here )

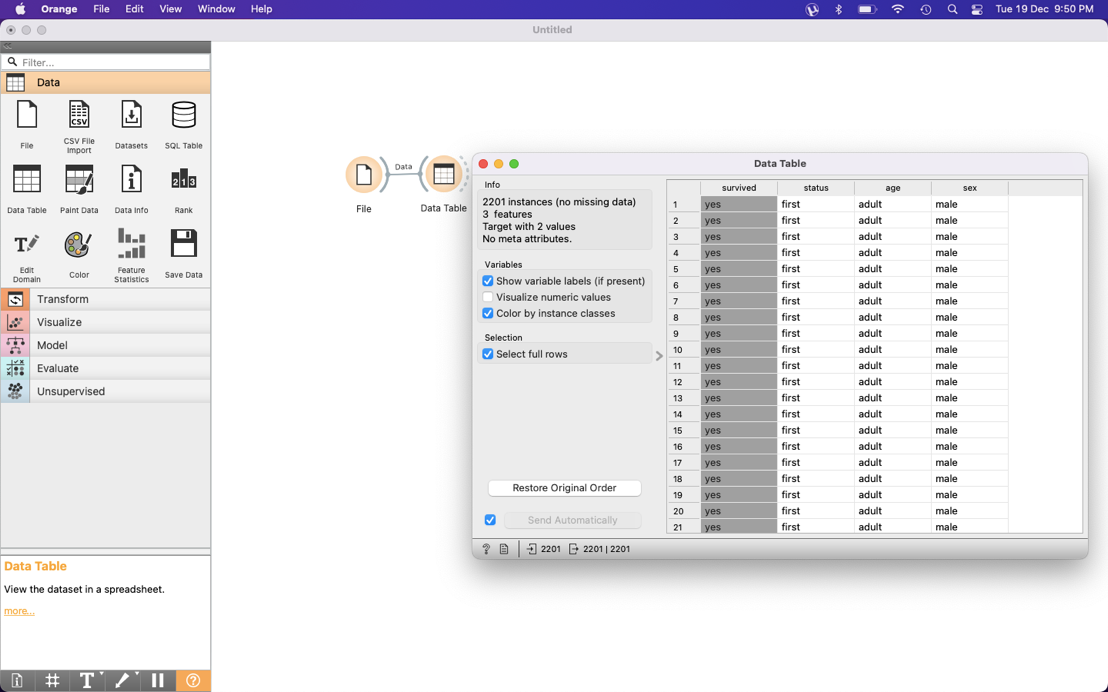


Step 4: Data Pre-processing

Drag and drop the "Data Table" widget onto the canvas.

Connect the "Data Table" widget to the "File" widget.

Use the "Data Table" widget to inspect your dataset. Address any missing values or outliers.

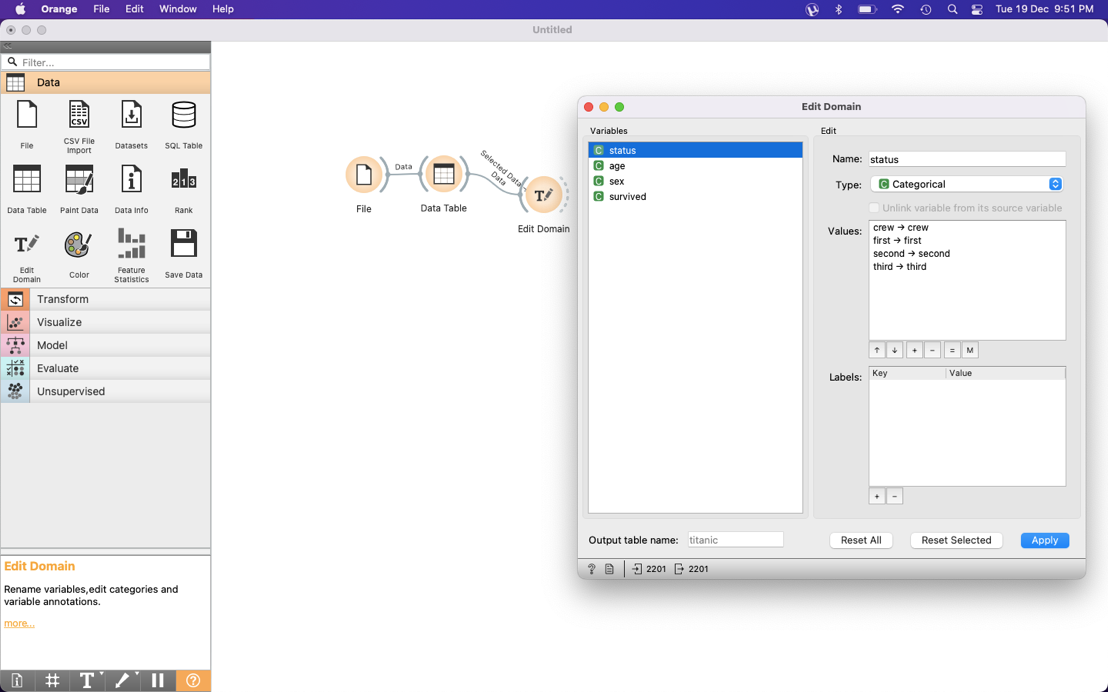


Step 5: Encode Categorical Variables

Drag and drop the "Edit Domain" widget onto the canvas.

Connect the "Edit Domain" widget to the "Data Table" widget.

In the "Edit Domain" widget, select the categorical variables and choose the appropriate encoding method.

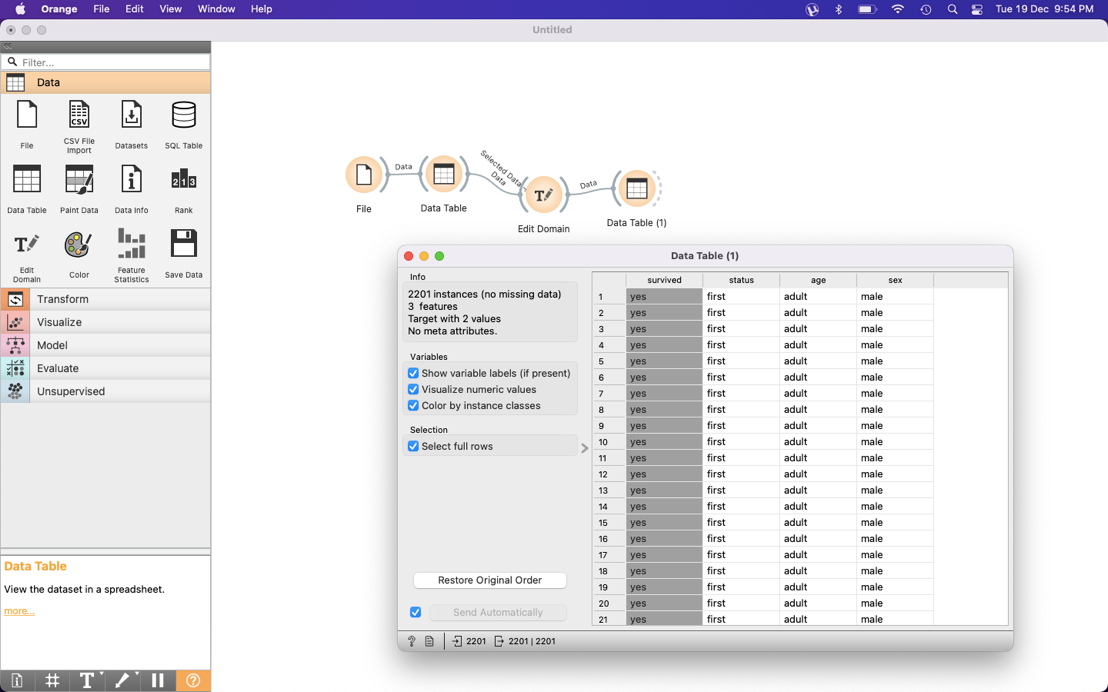


Step 6: Data Analysis

Drag and drop the "Data Table" widget onto the canvas.

Connect the "Data Table" widget to the "Edit Domain" widget.

Use the "Data Table" widget to explore basic statistics and demographics of your customers.

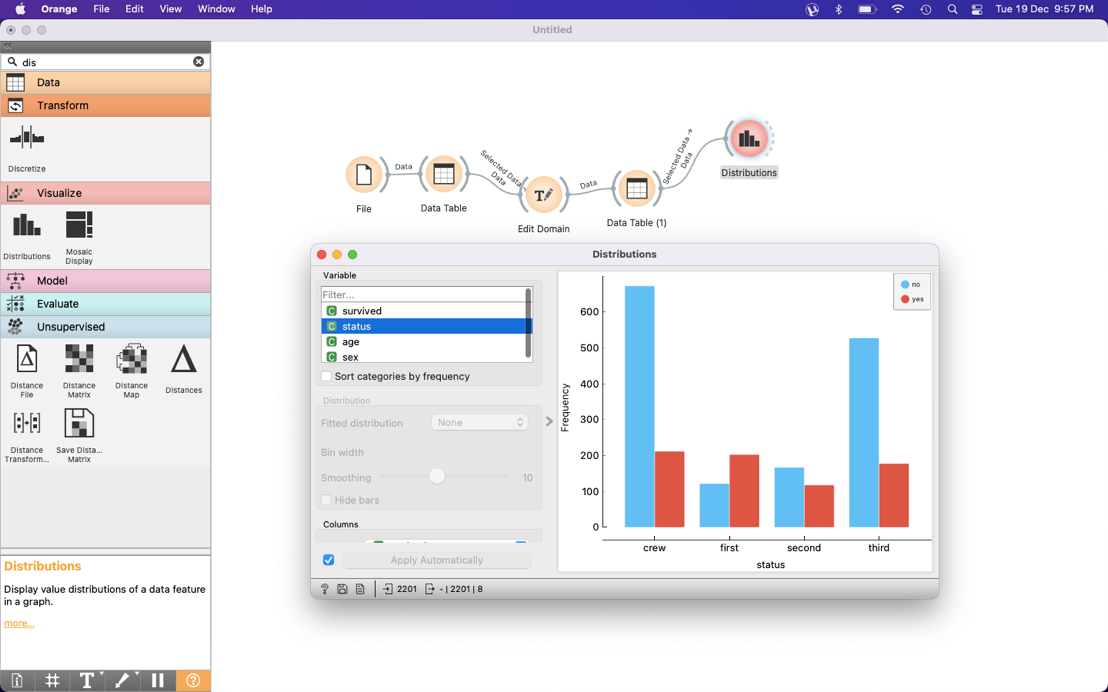


Step 7: Data Visualization

Drag and drop the "Distributions" widget onto the canvas.

Connect the "Distributions" widget to the "Edit Domain" widget.

In the "Distributions" widget, select variables like age, income, and spending to plot histograms and visualize distributions.

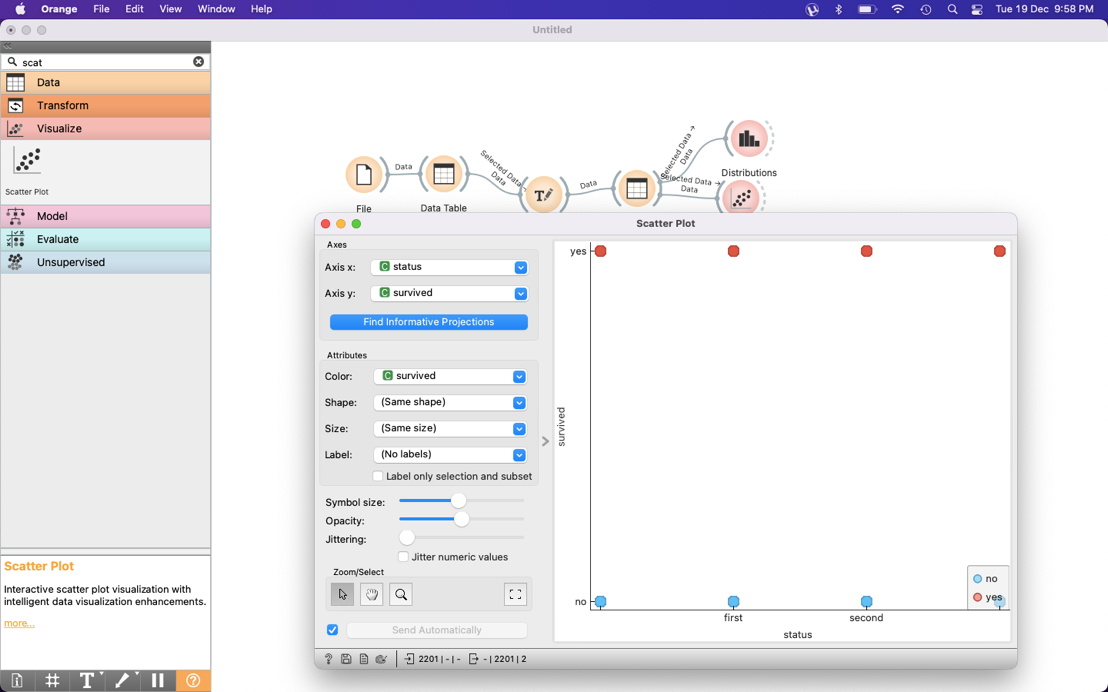


For visualizing relationships between income and spending, you can use the "Scatter Plot" widget.

Drag and drop the "Scatter Plot" widget onto the canvas.

Connect the "Scatter Plot" widget to the "Edit Domain" widget.

Choose "Income" as the X-axis variable and "Spending" as the Y-axis variable.



Final Design:

