



# Multi Class Logistic Regression with Azure ML



# Contents

## Table of Contents

Project Expectation .....	4
Dataset Selection .....	5
SQL Transformation .....	6
Query Script .....	8
Select Columns.....	10
Split Data.....	11
Multi Class Logistic Regression.....	12
Score & Evaluate Model.....	14
Visualize .....	15
Result .....	15

# Goals and Requirements

**Estimated time to complete lab is 15-20 minutes.**

## **Goals**

1. Predict the Wine Quality using the dataset provided.
2. Multi Class Logistic Regression Model Implementation

## **Requirements**

1. Access to an Azure subscription (with subscription administrator permissions).

# Multi Class Logistic Regression

**Project Expectation:** To Build a Predictive Model for Wine Quality

## Wine Quality Prediction

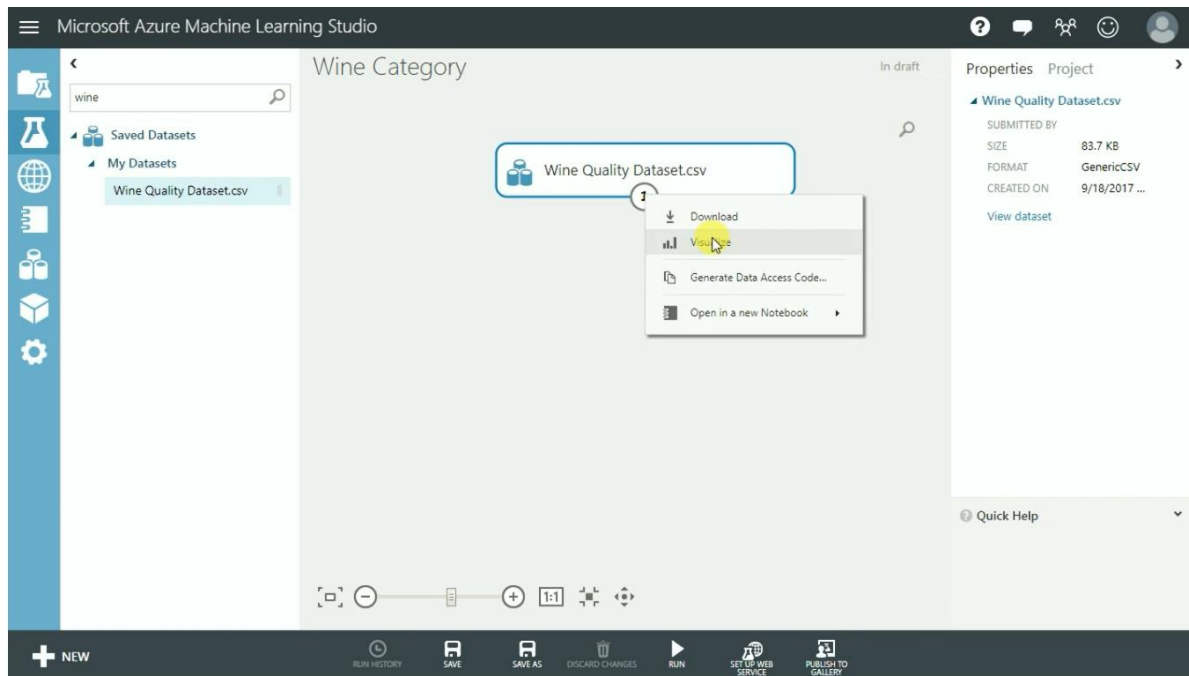


- Fixed and Volatile Acidity
- Citric acid
- Residual sugar
- Chlorides
- Free and Total Sulfur dioxide
- Density
- pH
- Sulphates
- Alcohol Content

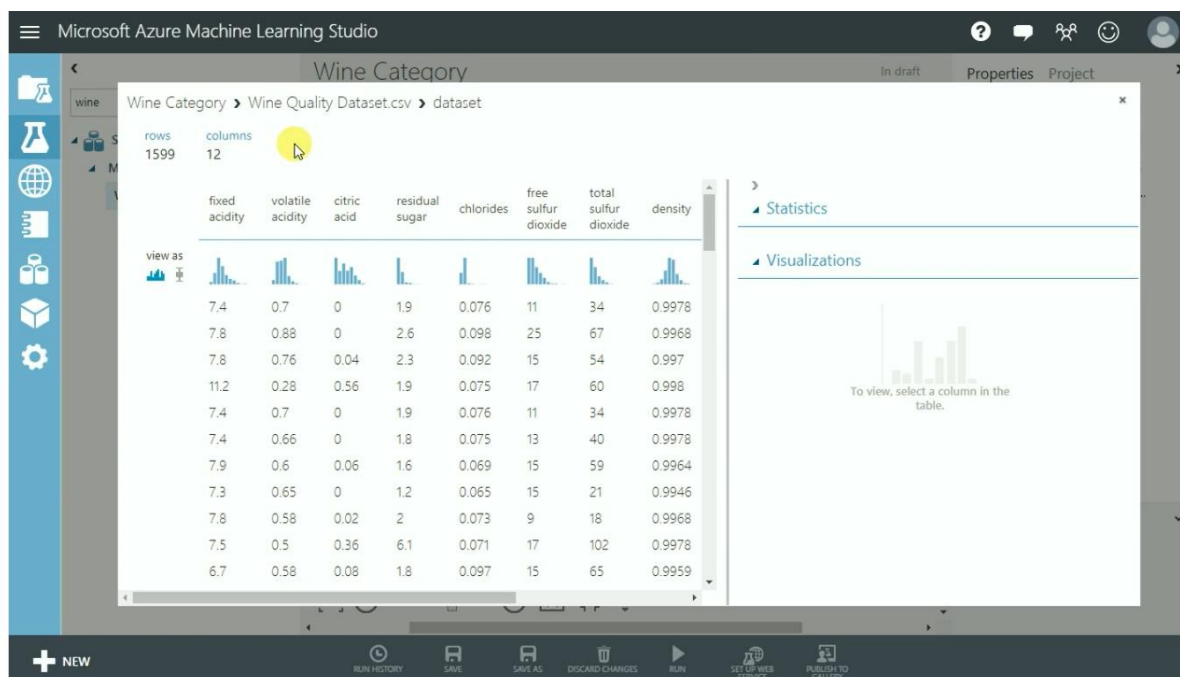
P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis.  
Modeling wine preferences by data mining from physicochemical properties. In Decision Support Systems, Elsevier, 47(4):547-553, 2009

## Dataset Selection

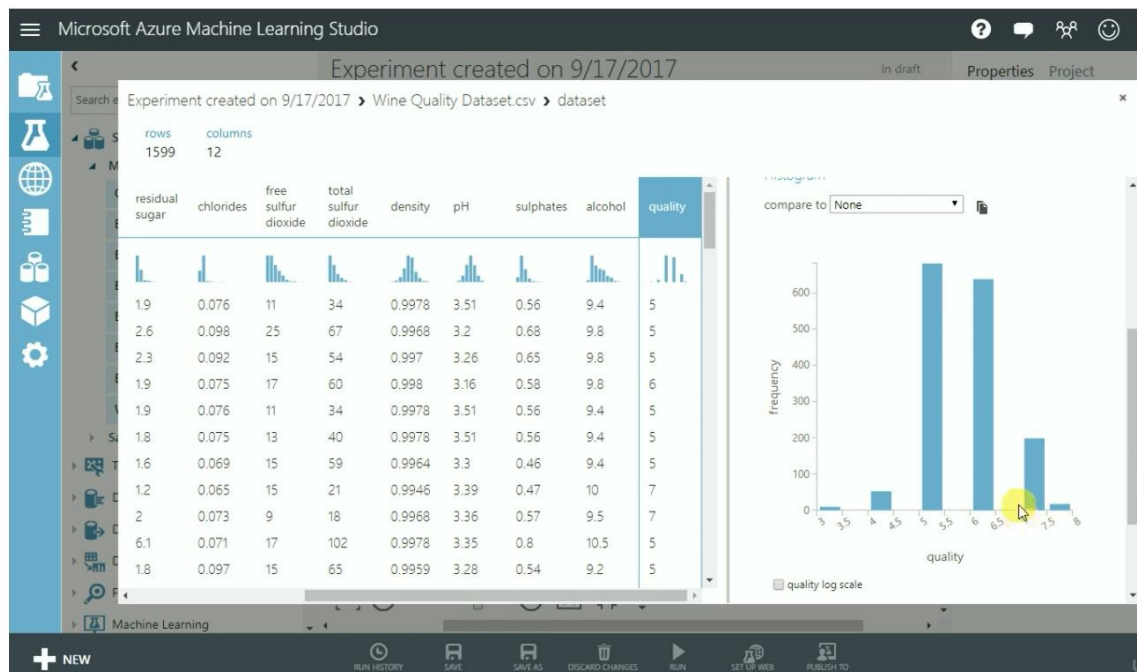
Select the dataset and drop in canvas and visualize the same



Visualize the data

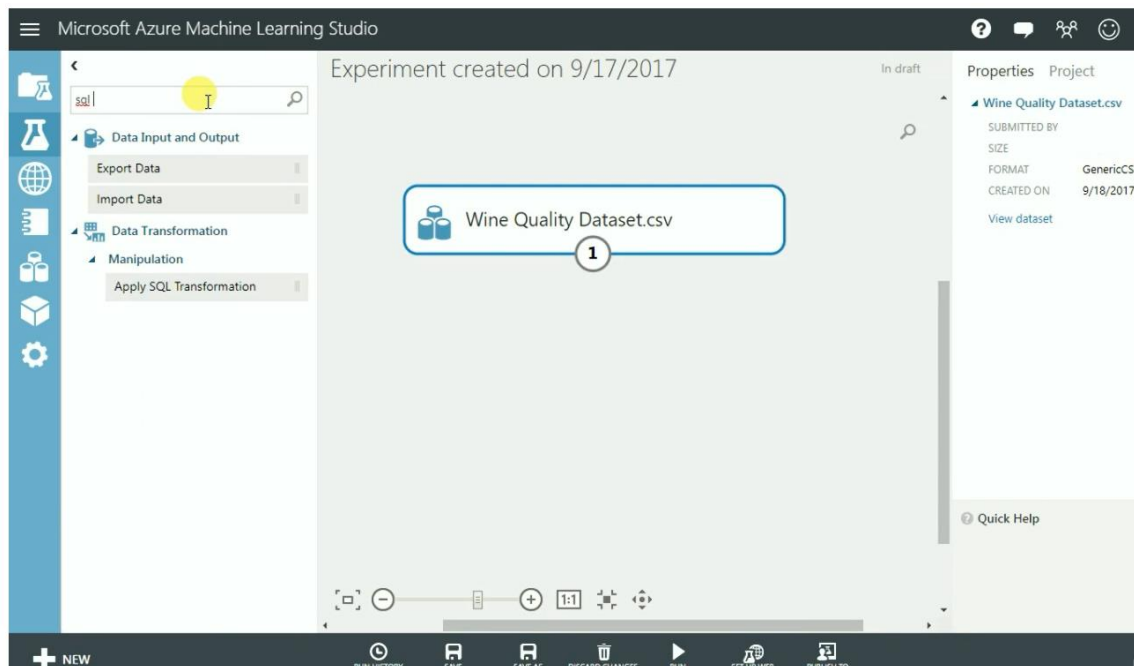


Considering 3 and 4 as low, 5 and 6 as medium, 7 and 8 as high from graph

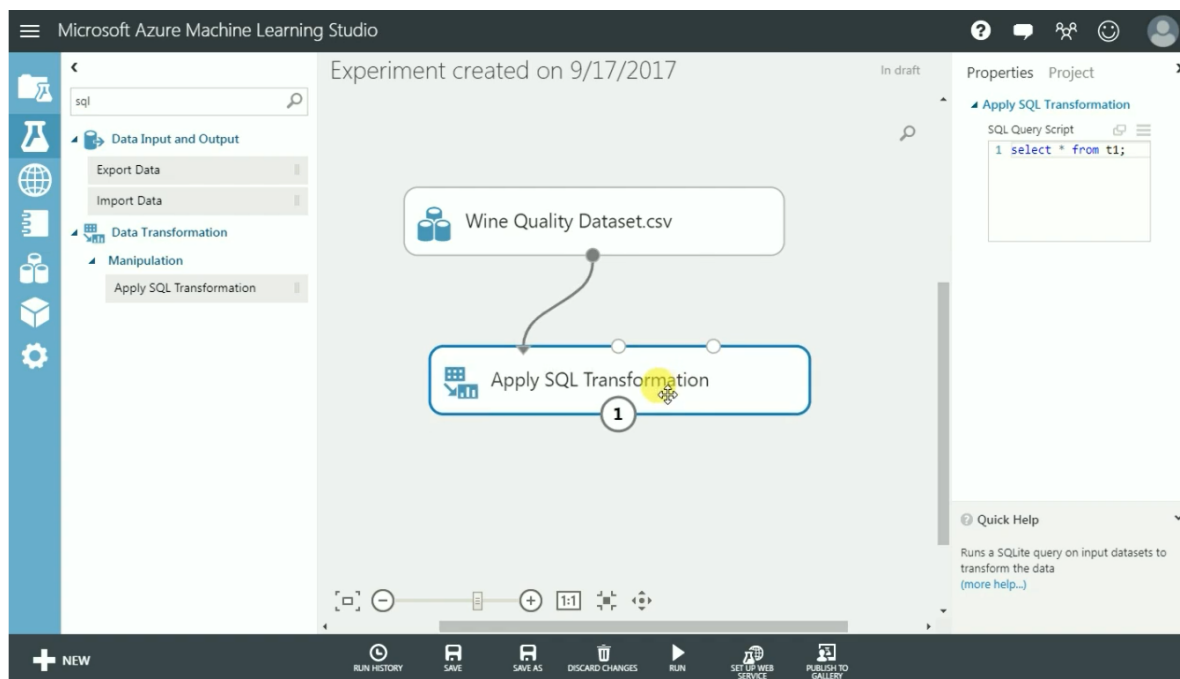
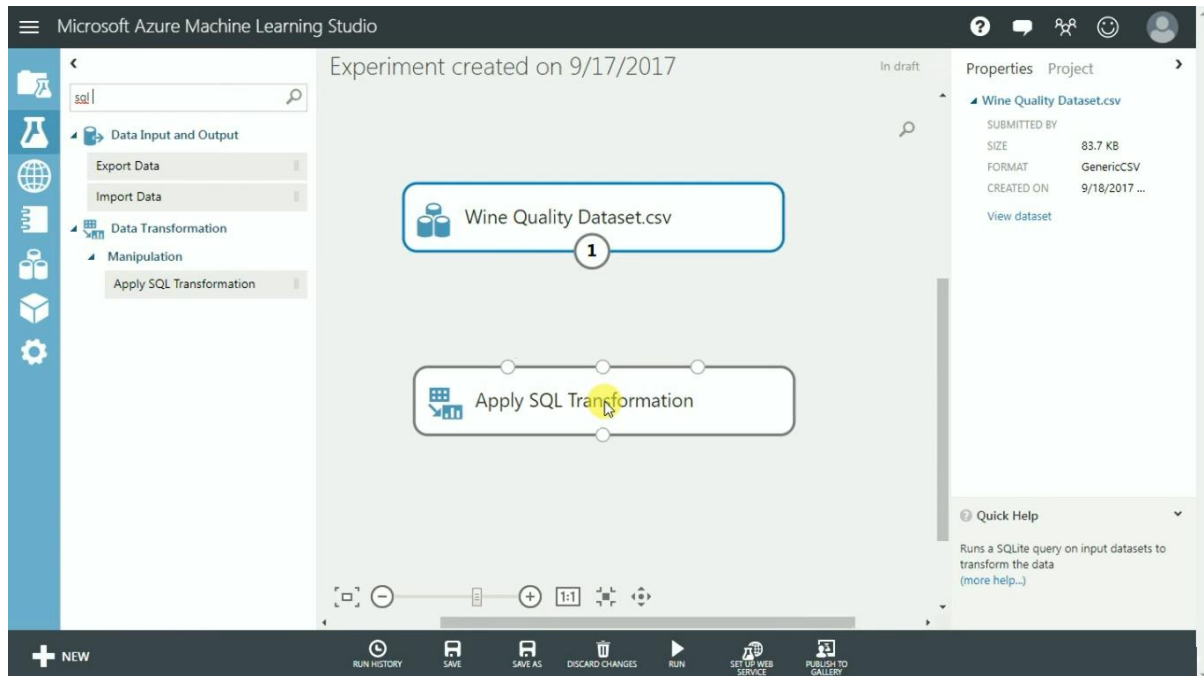


## SQL Transformation

Search for Apply SQL transformation

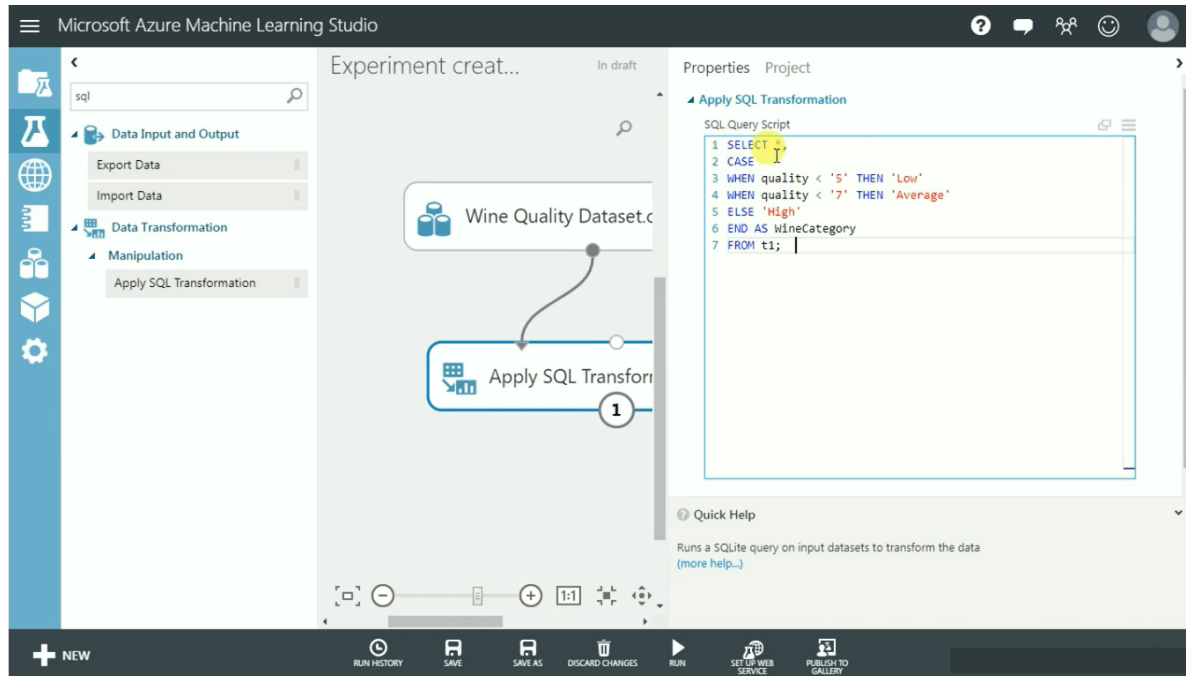


## Connect the nodes

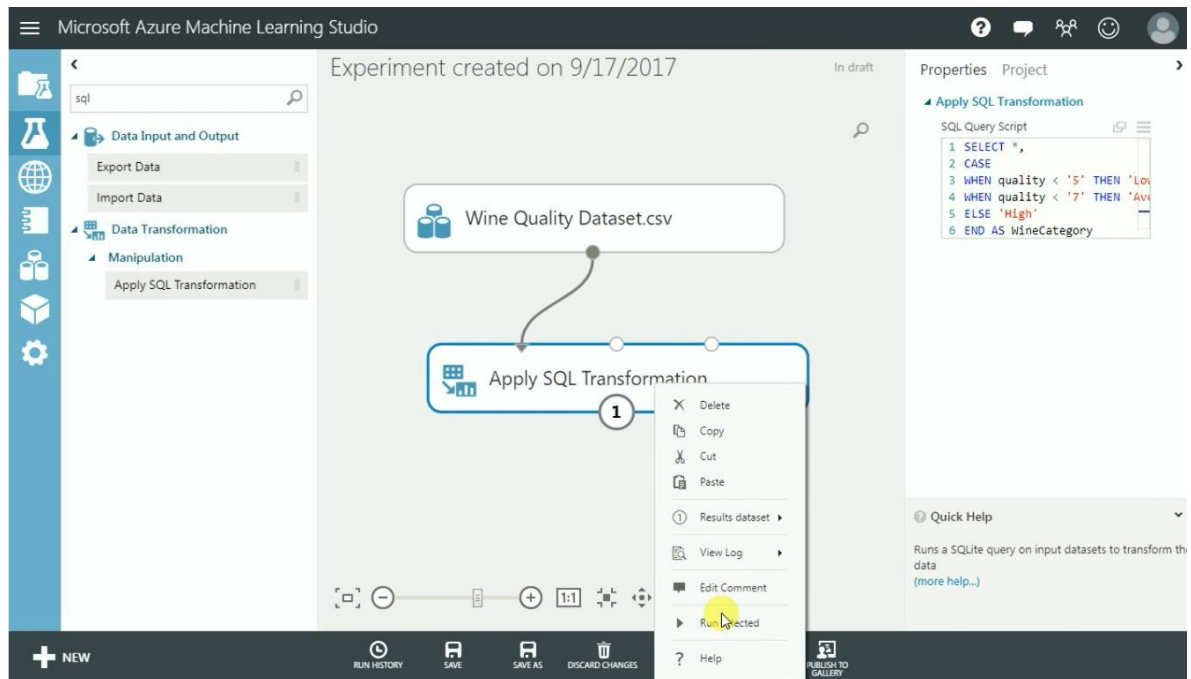


## Query Script

Enter the query script in the right to categorize

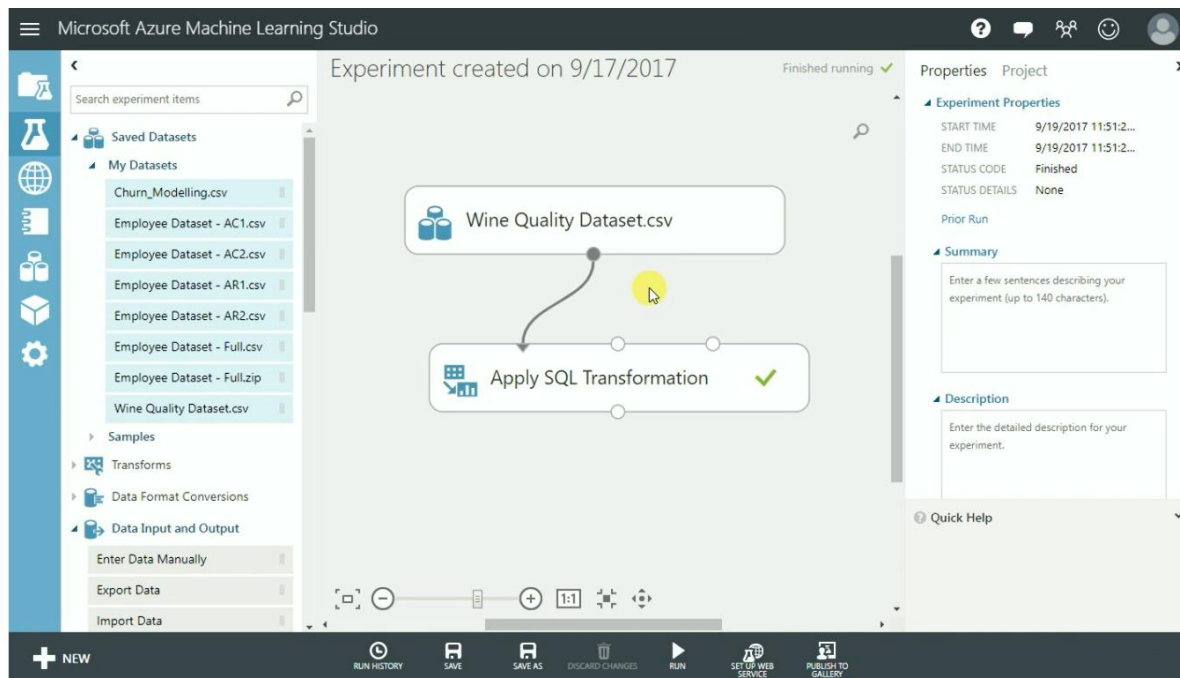


Run for output

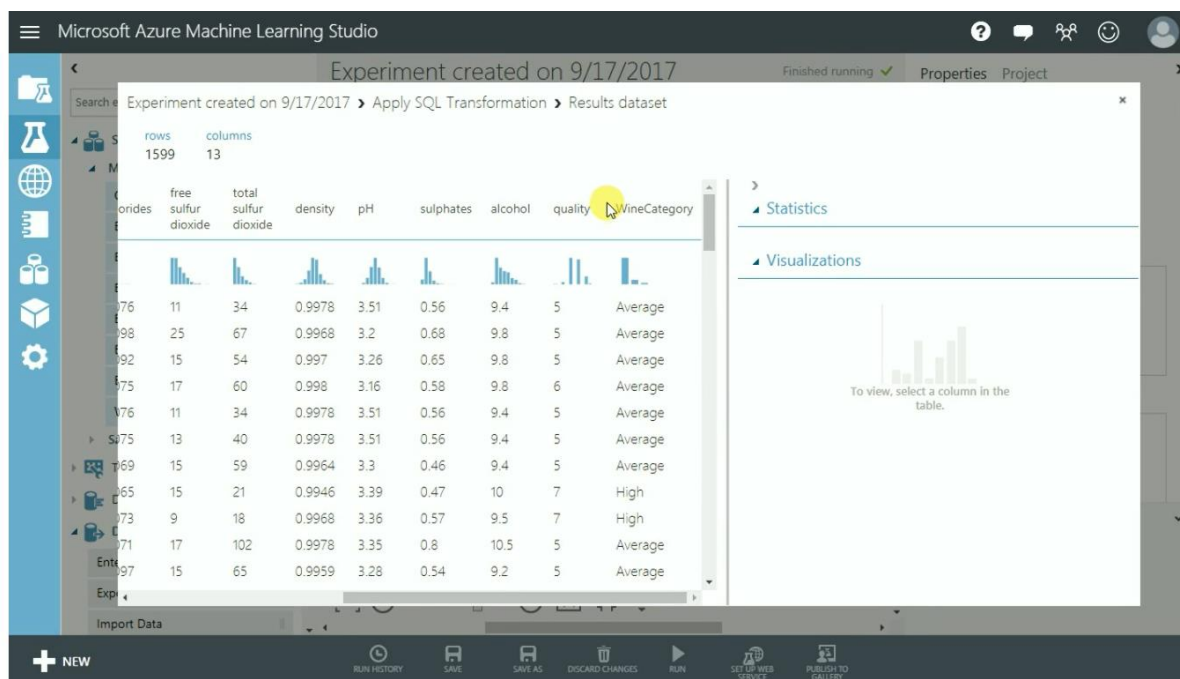




Right click and visualize for result

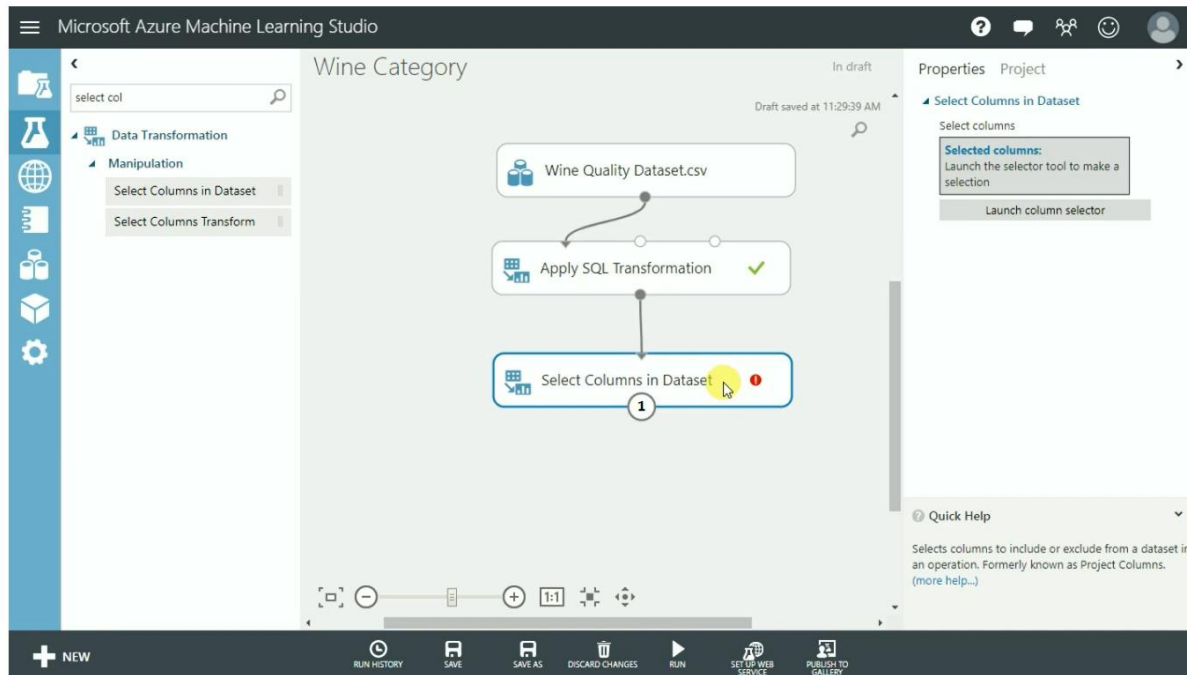


Result

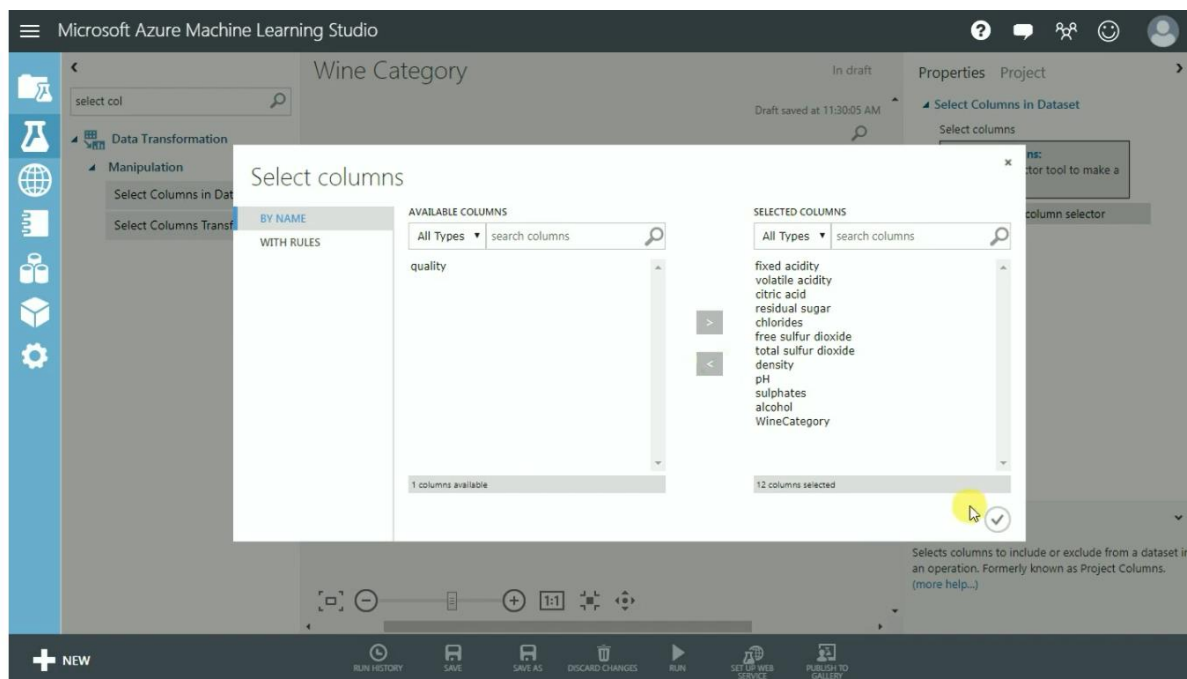


## Select Columns

Search for select columns in dataset and connect with existing dataset

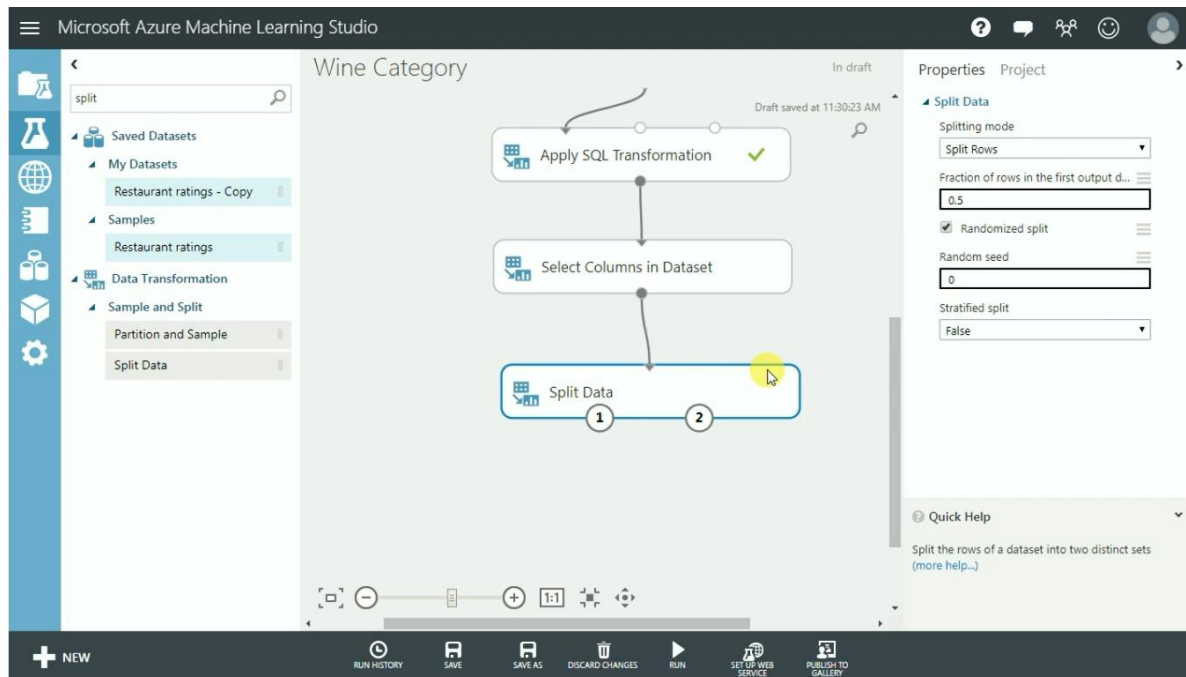


Launch column selector and select all types except quality and click ok

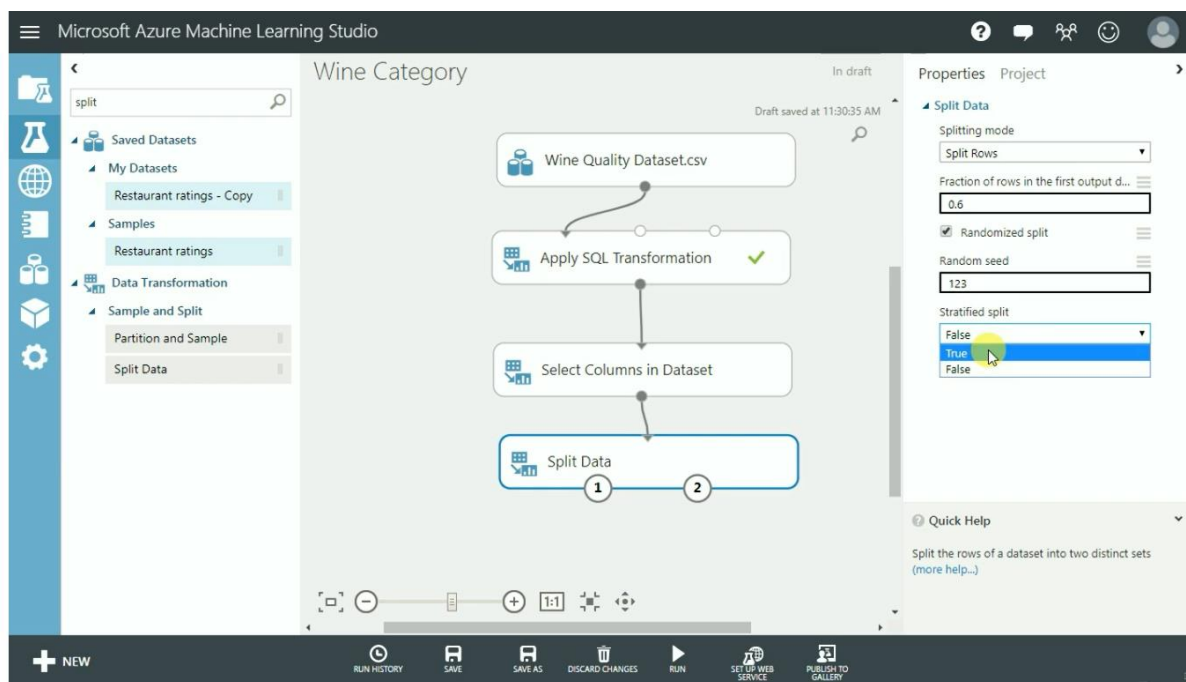


## Split Data

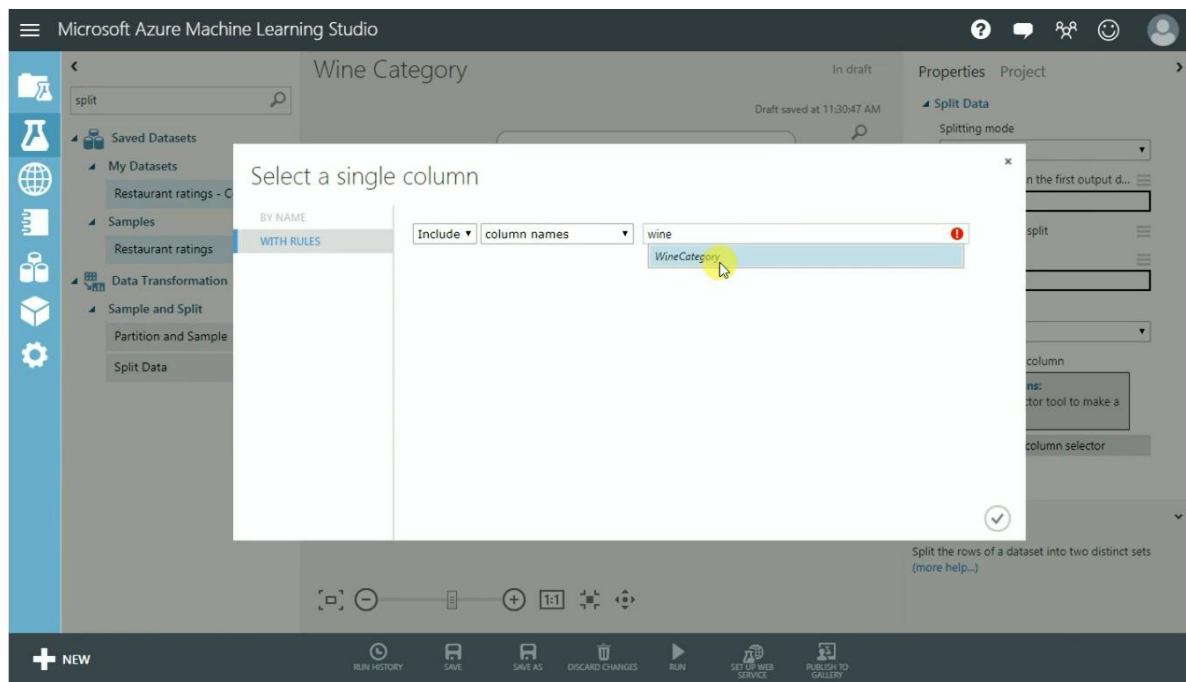
Search for split data and connect with select column dataset



Change the parameters as shown and stratified split as true



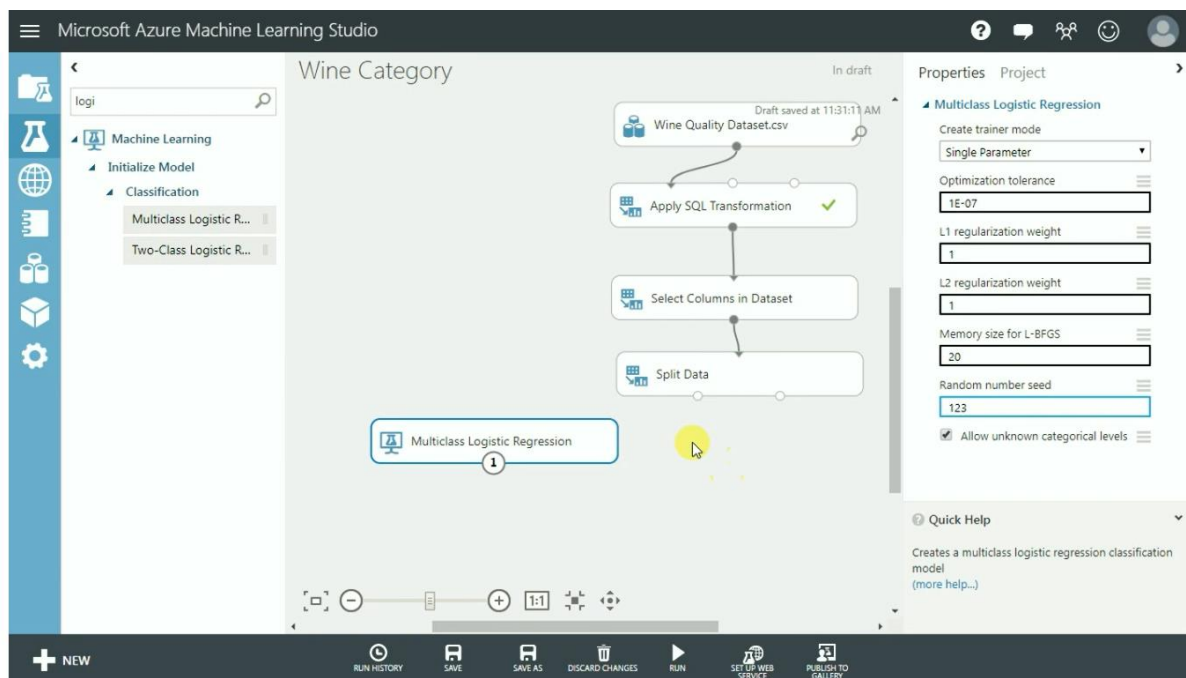
Launch column selector and change to wine category and click ok



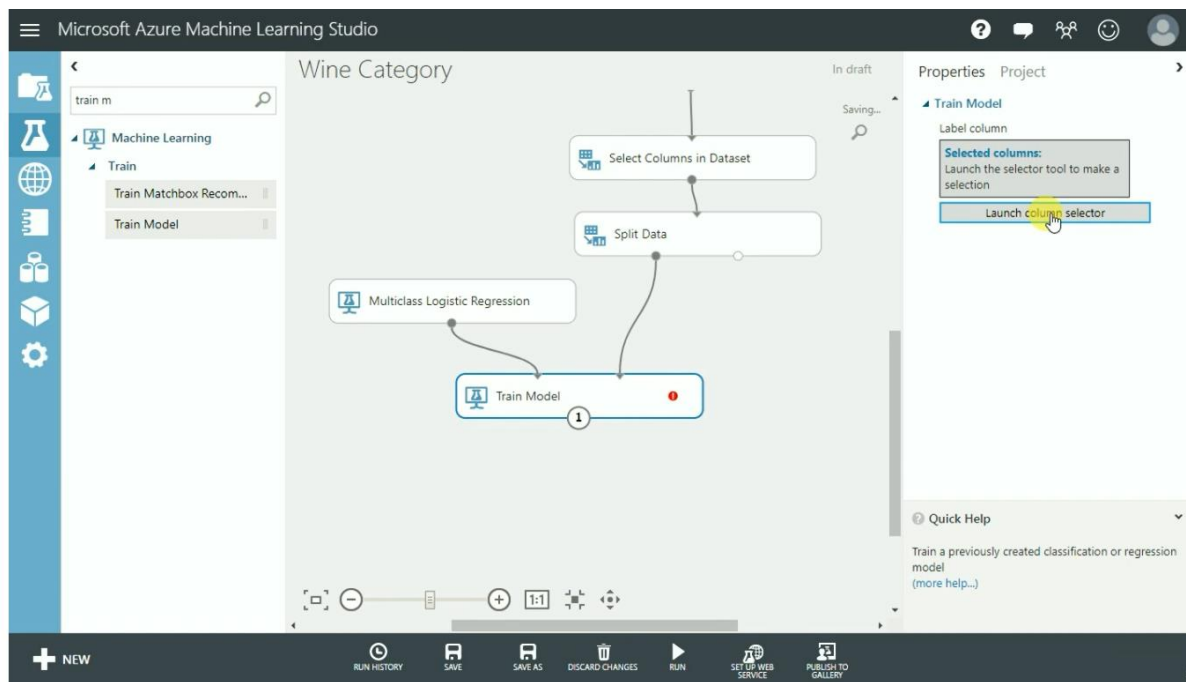
## Multi Class Logistic Regression

Search for multiclass logistic regression and change random number seed

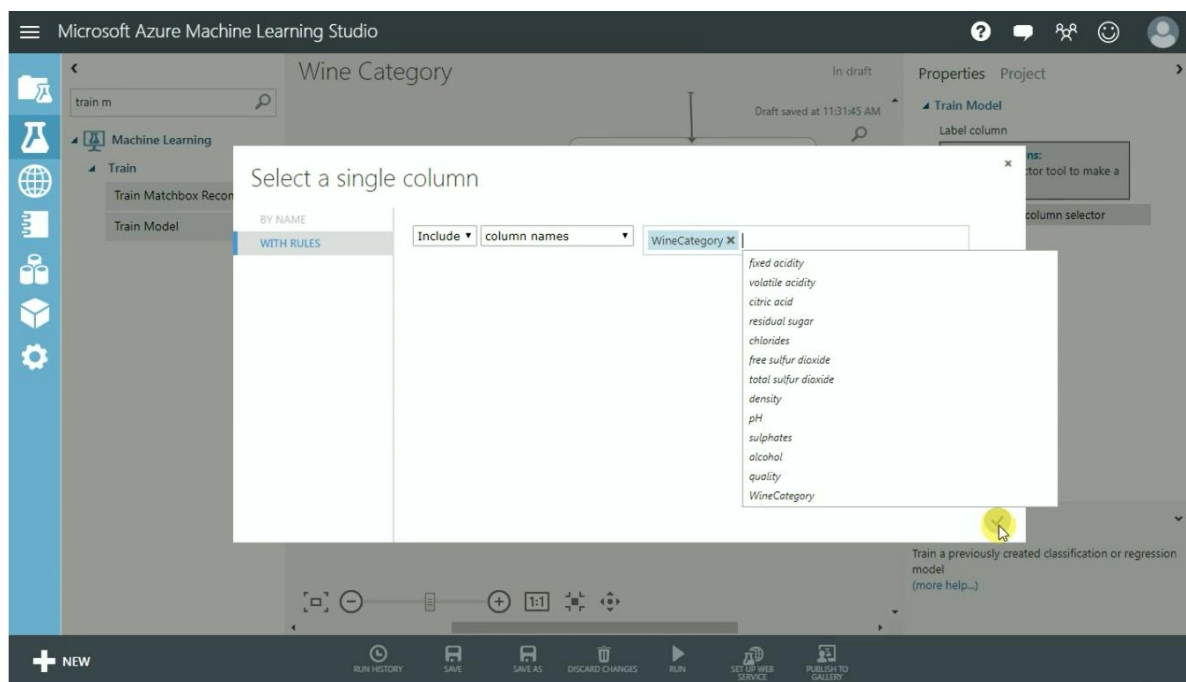
Parameter as 123



Add train model in canvas and connect nodes as shown



Launch column selector and select wine category and click ok



Drag and drop the Score model now and connect node1 with Train model

The screenshot displays the Microsoft Azure Machine Learning Studio environment. The main workspace shows a workflow diagram for training and scoring a model. The workflow consists of the following steps:

- Select Columns in Dataset**: The first step in the workflow.
- Split Data**: The second step, which is highlighted with a yellow circle. It has a dropdown menu open, showing options like "Train and Test Split" and "Random Split".
- Multiclass Logistic Regression**: A model component connected to the "Split Data" step.
- Train Model**: A step that receives input from both the "Split Data" step and the "Multiclass Logistic Regression" model.
- Score Model**: The final step in the workflow, which receives input from the "Train Model" step. It is highlighted with a yellow circle and has a red error icon.

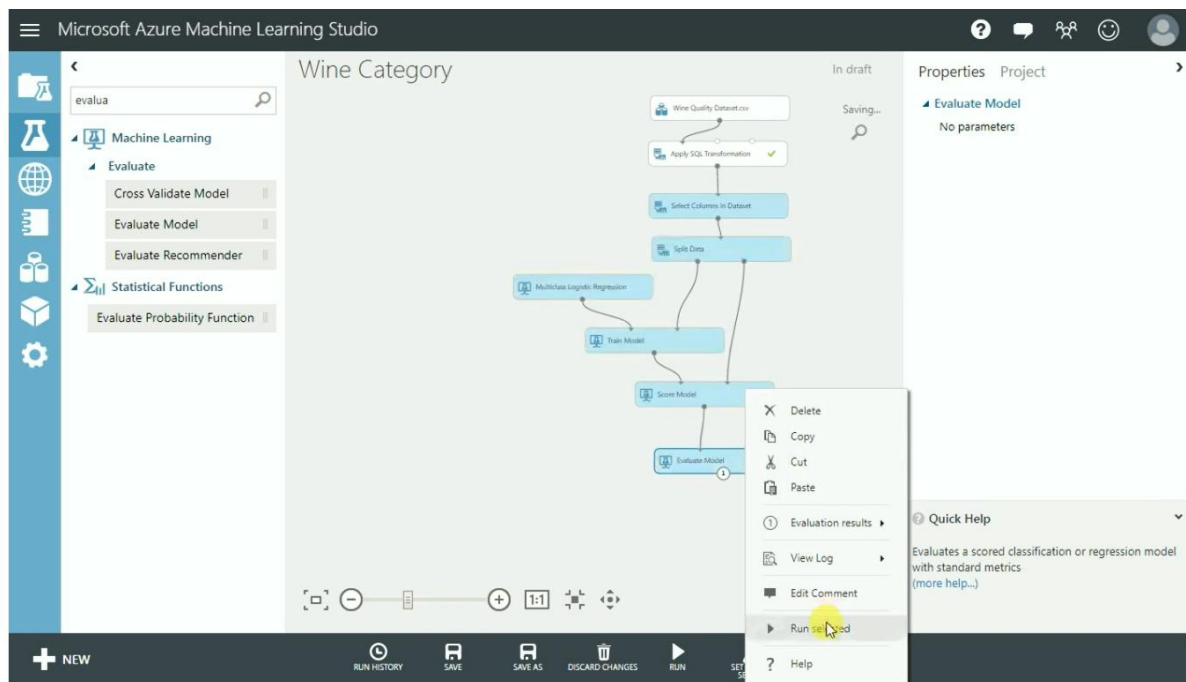
The interface includes a sidebar on the left with a search bar and a list of components under the "Machine Learning" and "Score" sections. The "Score" section lists "Score Matchbox Recom..." and "Score Model". The right sidebar shows the "Properties" panel for the selected "Score Model" component, with the "Append score columns to output" checkbox checked. The bottom of the interface features a toolbar with icons for zooming, saving, and running the workflow.

The screenshot displays the Microsoft Azure Machine Learning Studio interface. The main workspace is titled "Wine Category" and shows a workflow diagram. The workflow consists of the following steps:

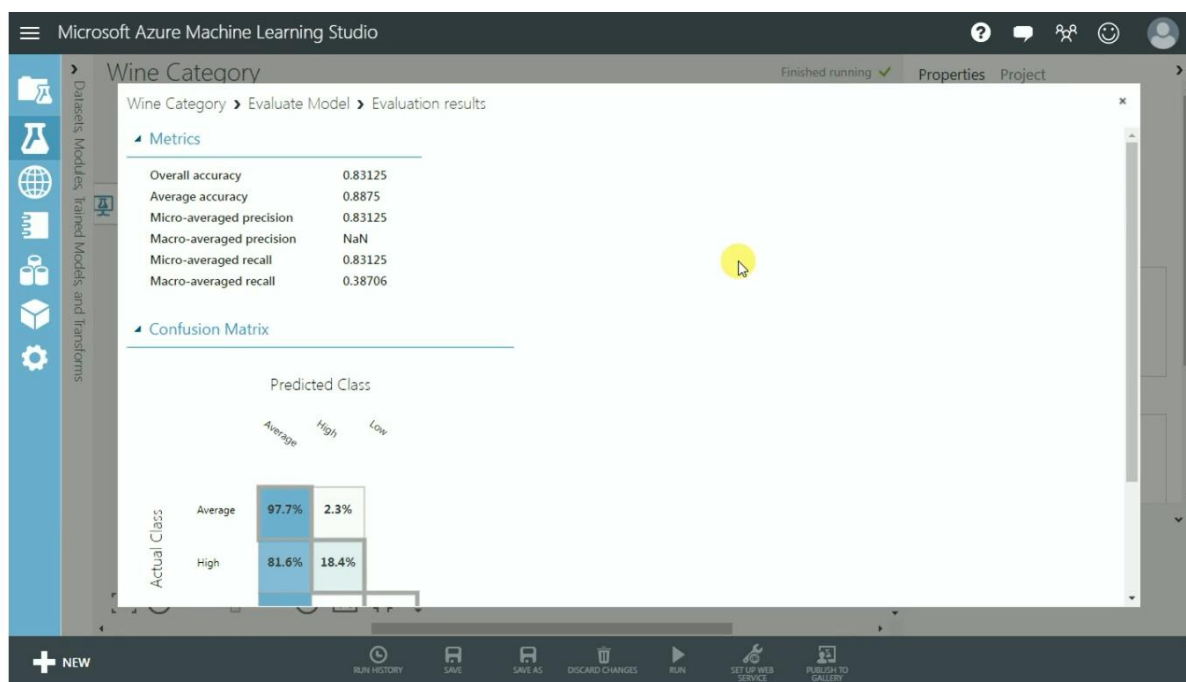
- Select Columns in Dataset**: The first step in the workflow.
- Split Data**: The second step, which branches into two paths.
- Multiclass Logistic Regression**: A step that receives input from the "Split Data" step.
- Train Model**: A step that receives input from the "Multiclass Logistic Regression" step and the "Split Data" step.
- Score Model**: The final step, which receives input from the "Train Model" step and the "Split Data" step. It is highlighted with a blue border and a "1" in a circle.

The left sidebar shows the "Machine Learning" section with the "Score" option selected. The right sidebar shows the "Properties" panel for the "Score Model" step, with the "Append score columns to output" checkbox checked. The top bar indicates "In draft" and "Draft saved at 11:32:18 AM".

14



Visualize the result



Result

As the result high and low quality is not predicted as expected due to Low observation available



