



# Linear Regression model using OLS



# Contents

Goals and Requirements .....	3
Automobile price dataset .....	5
Dataset for Clean Missing Data .....	6
Edit Metadata .....	9
Dataset for Split data .....	11
Linear Regression, train model, score model and evaluate model dataset .....	12

# Goals and Requirements

Estimated time to complete lab is 40-45 minutes

## Goals

1. Implement and design a predictive model to predict the price of vehicle based on past data.
2. Approach of using Linear Regression using OLS

## Requirements:

1. Access to an Azure Machine Learning Studio

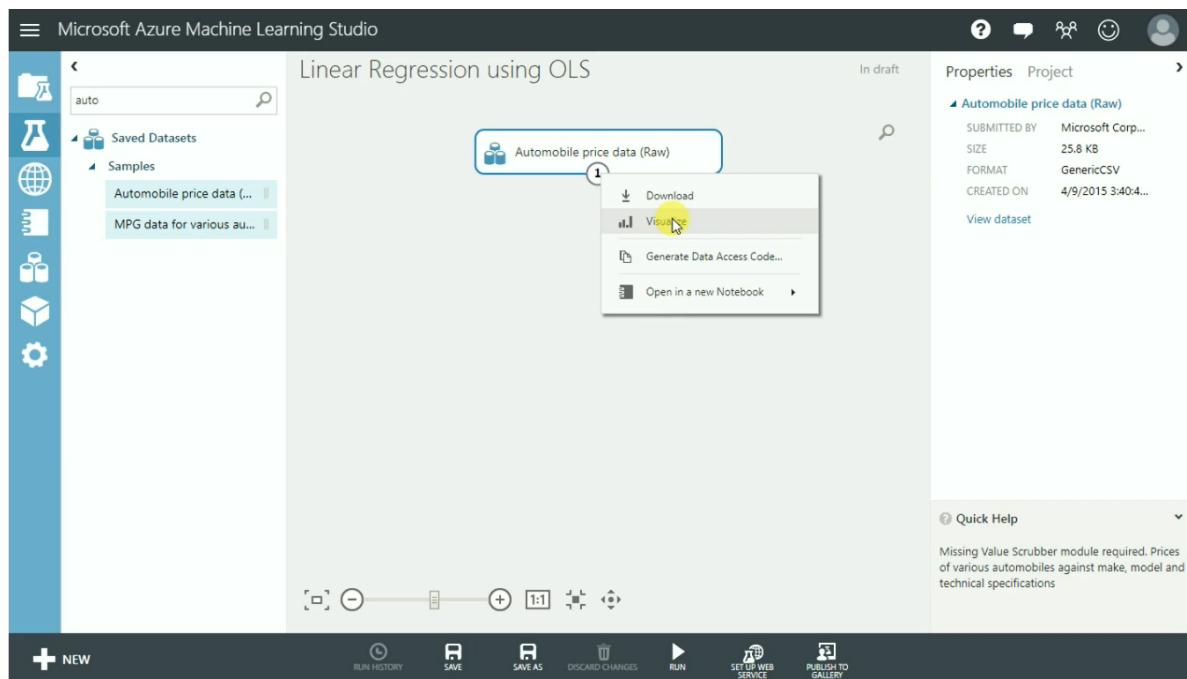
# Build Linear Regression Model

## Business Problem

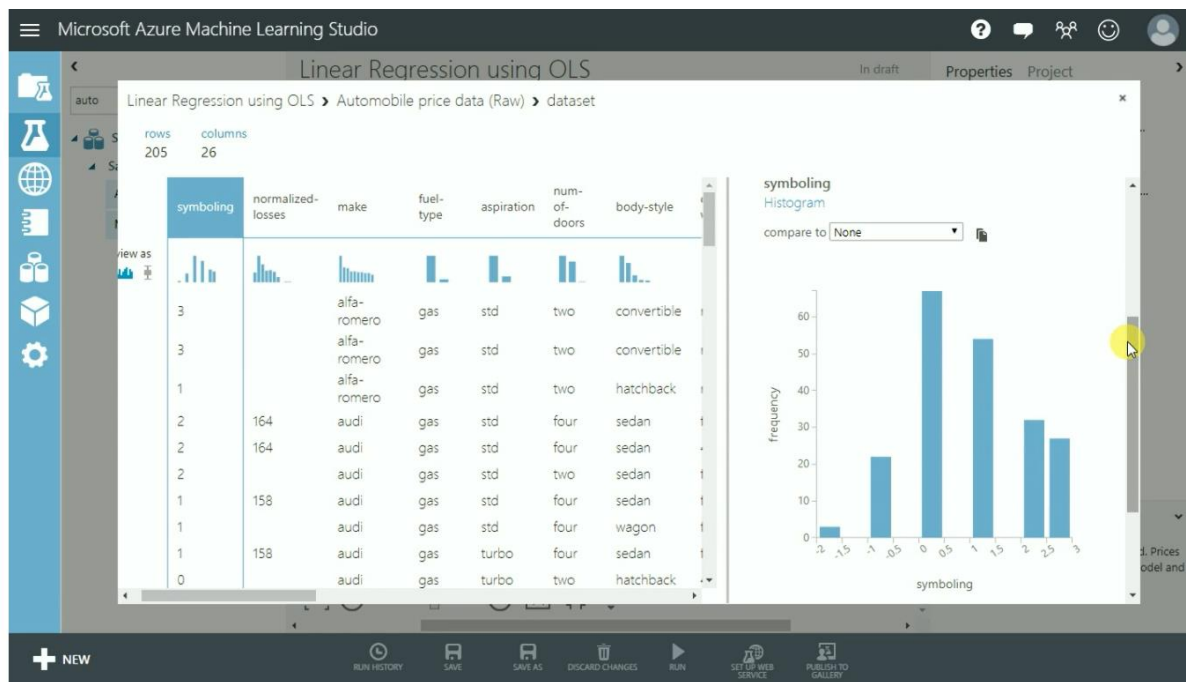
Build a model to predict the price of the vehicle based on the available historic data

## Automobile price dataset

Search automobile price dataset and visualize the data

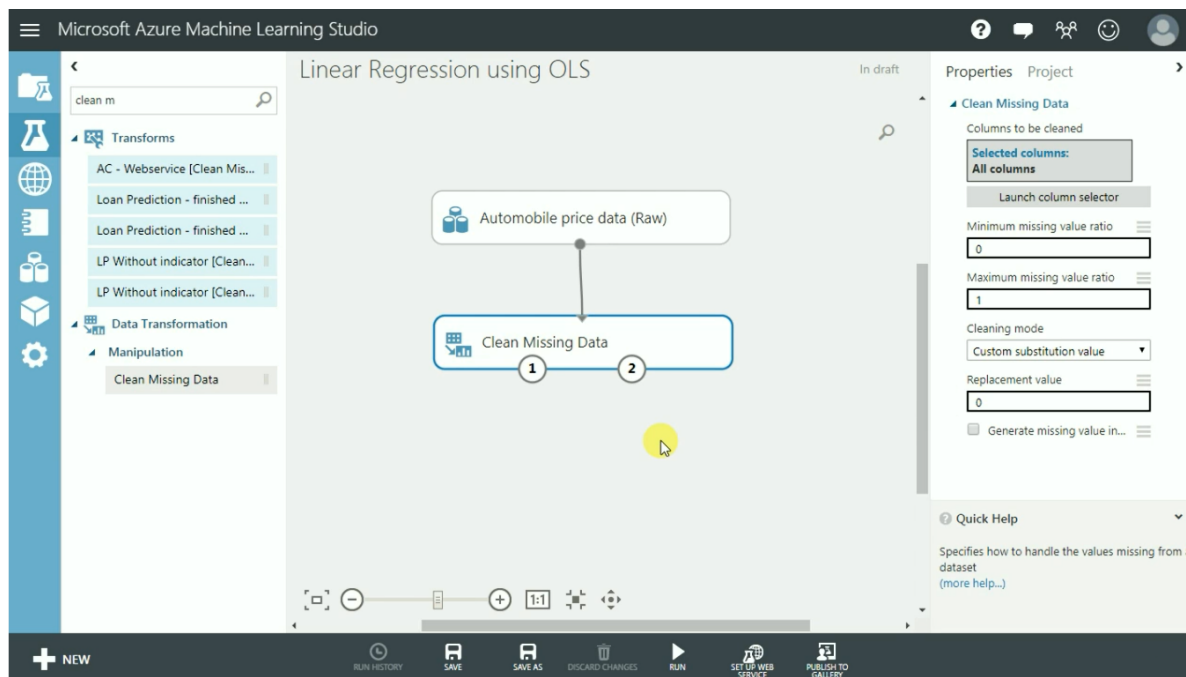


Check the columns involved and their features for knowing missing values and metadata

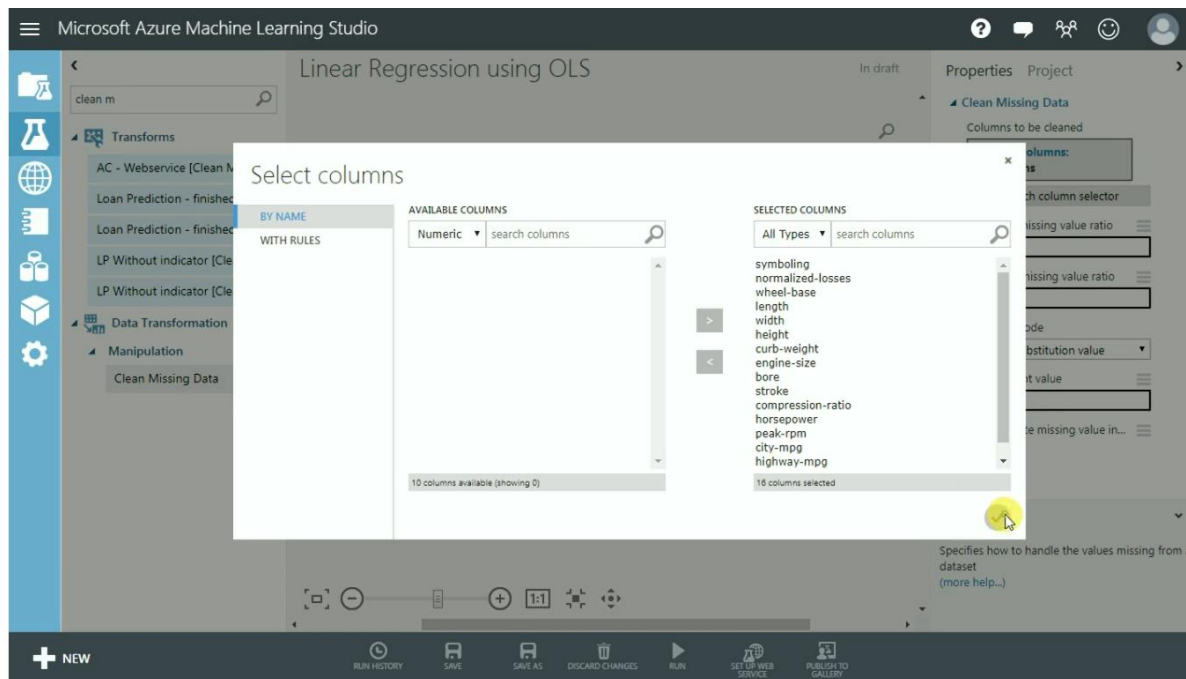


## Dataset for Clean Missing Data

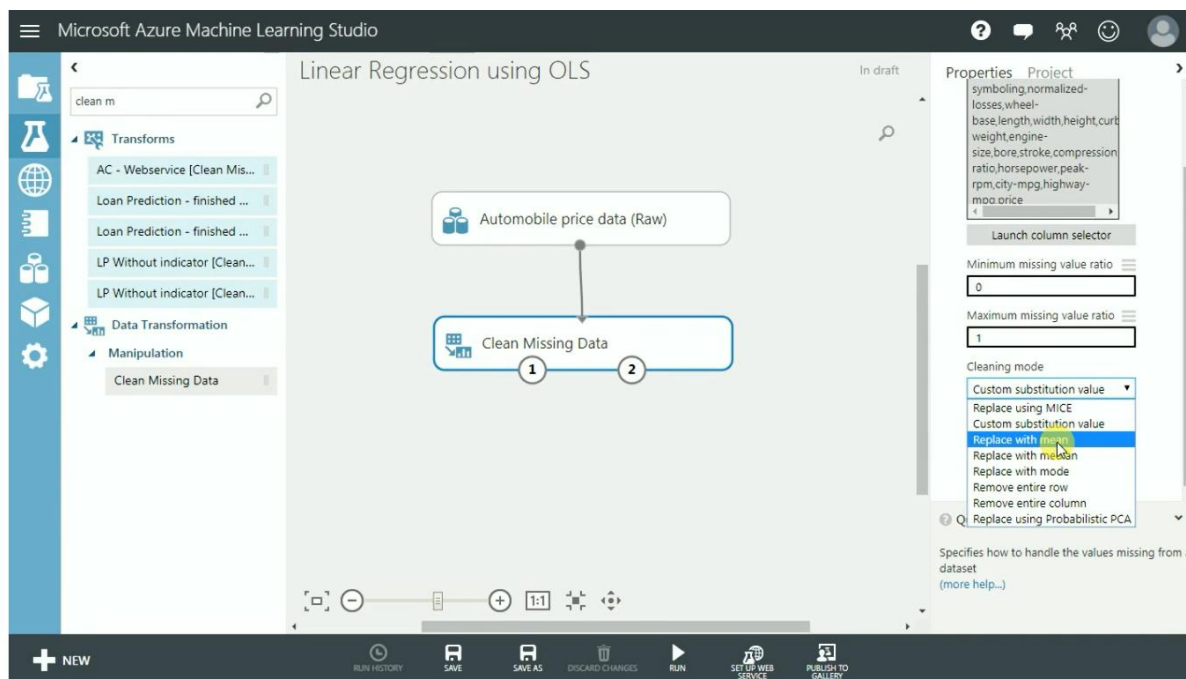
Add clean missing data in canvas and connect the nodes



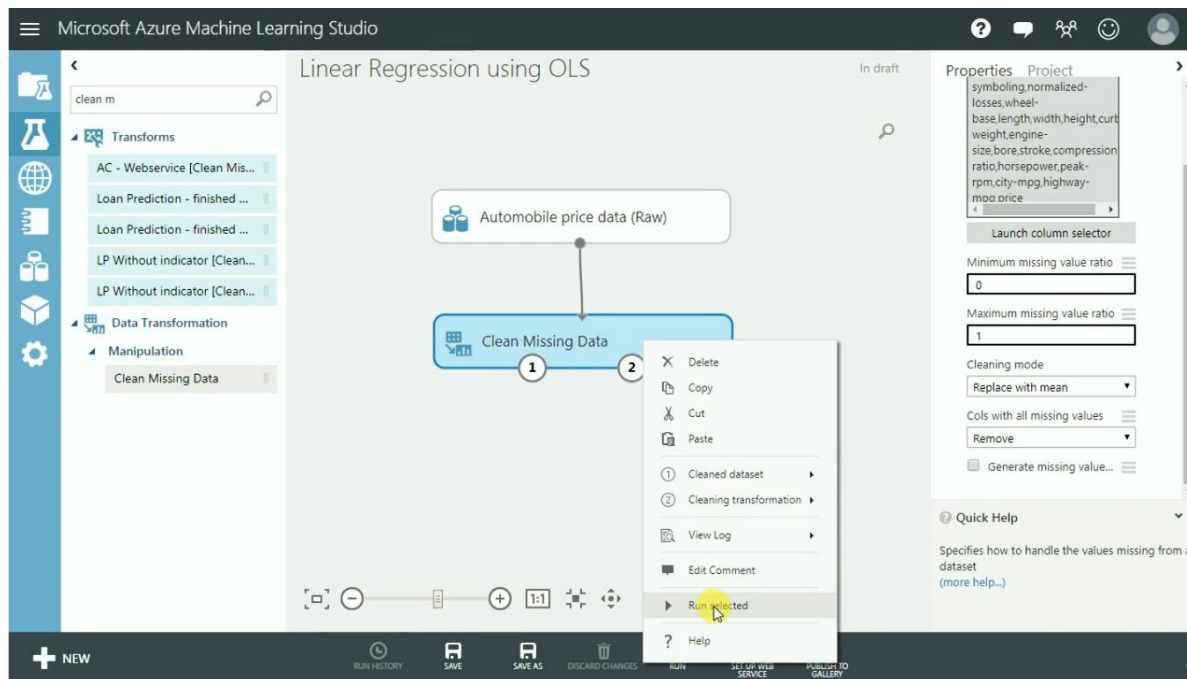
Launch column selector and select all numeric features, press ok



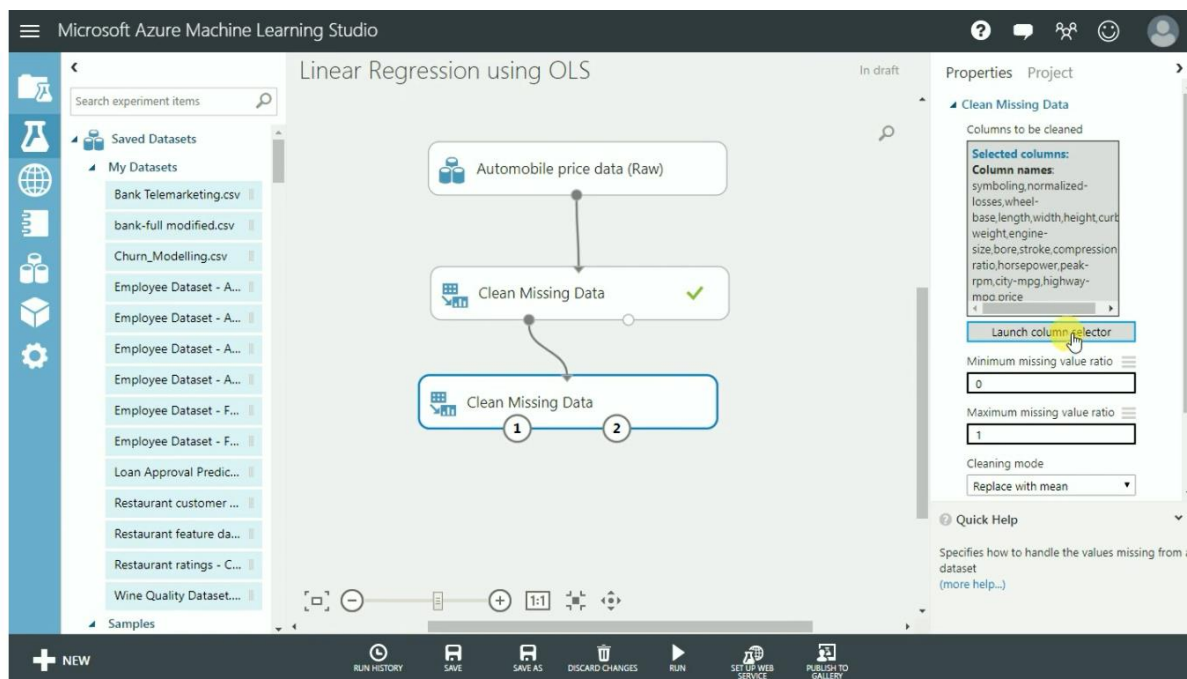
Replace with mean



Run the module

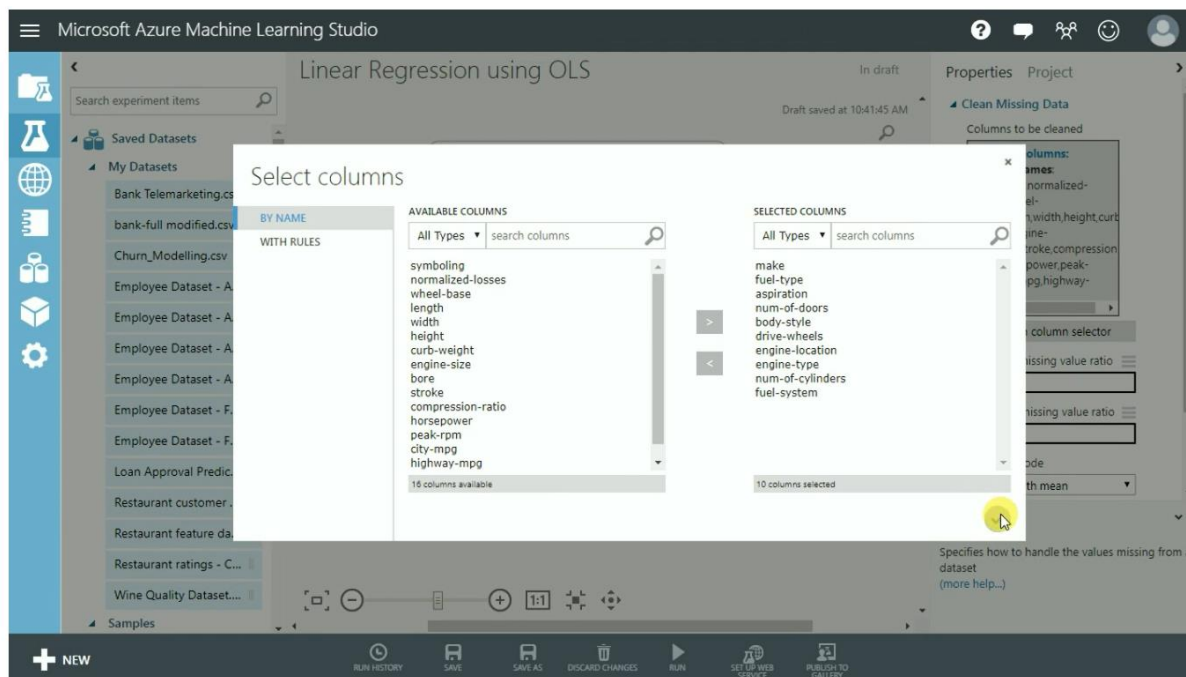


Copy and paste another clean missing data for string features



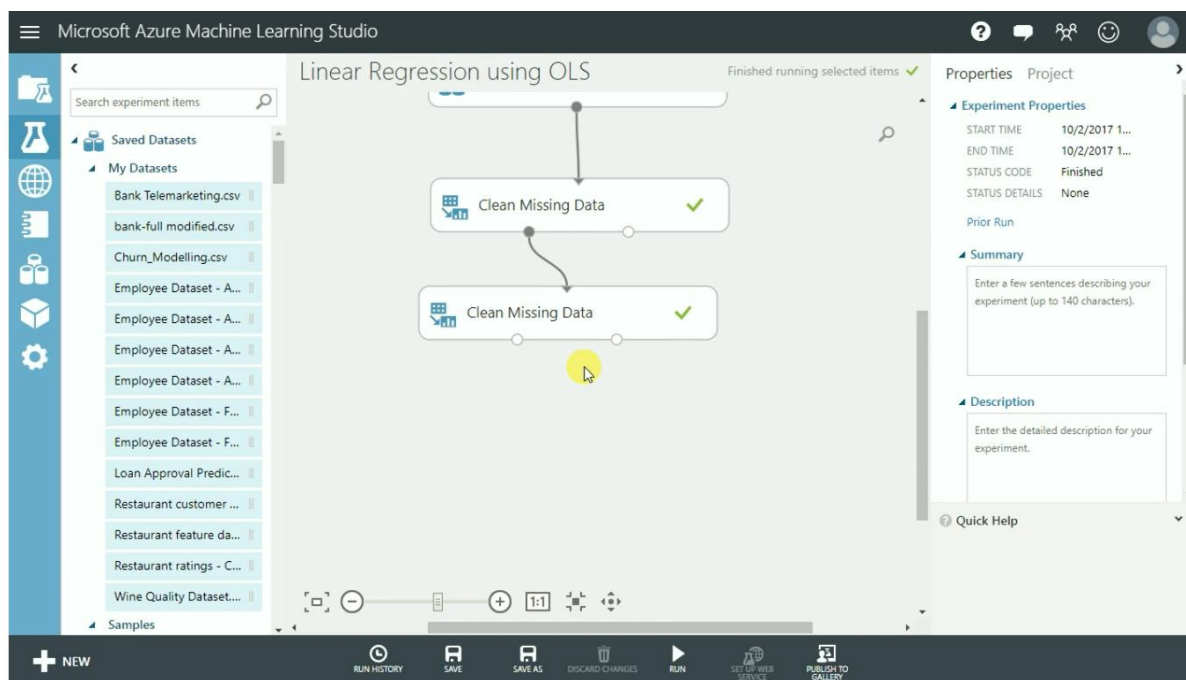


Launch column selector and select string variables and press ok

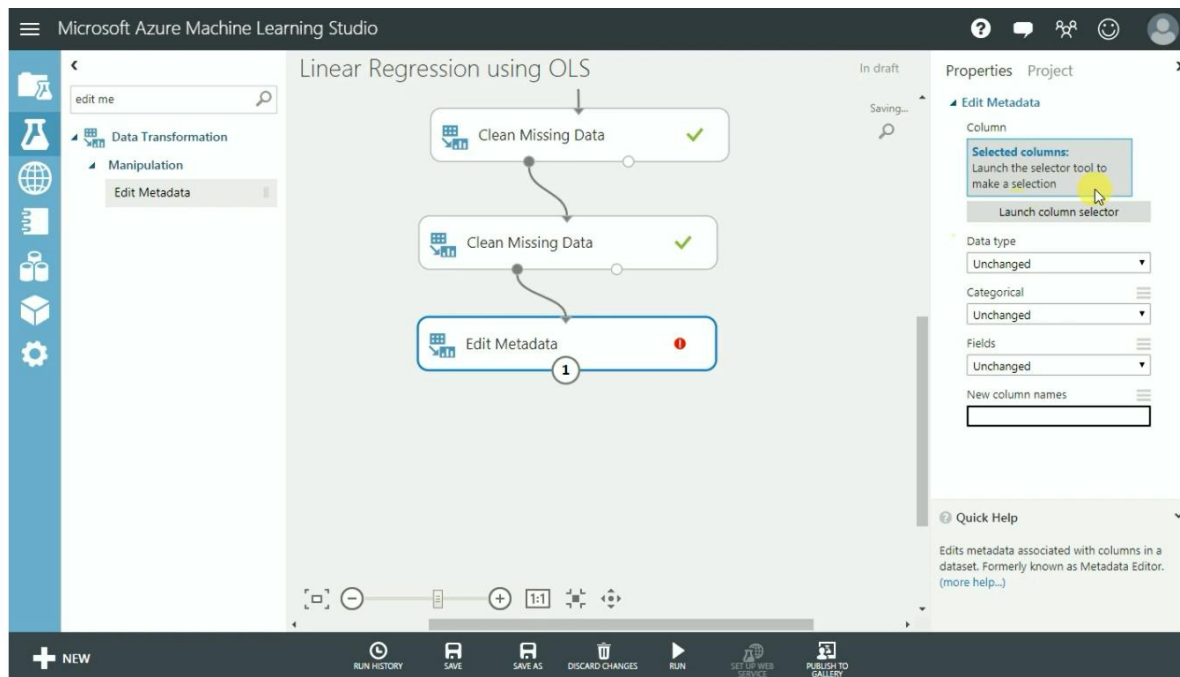


## Edit Metadata

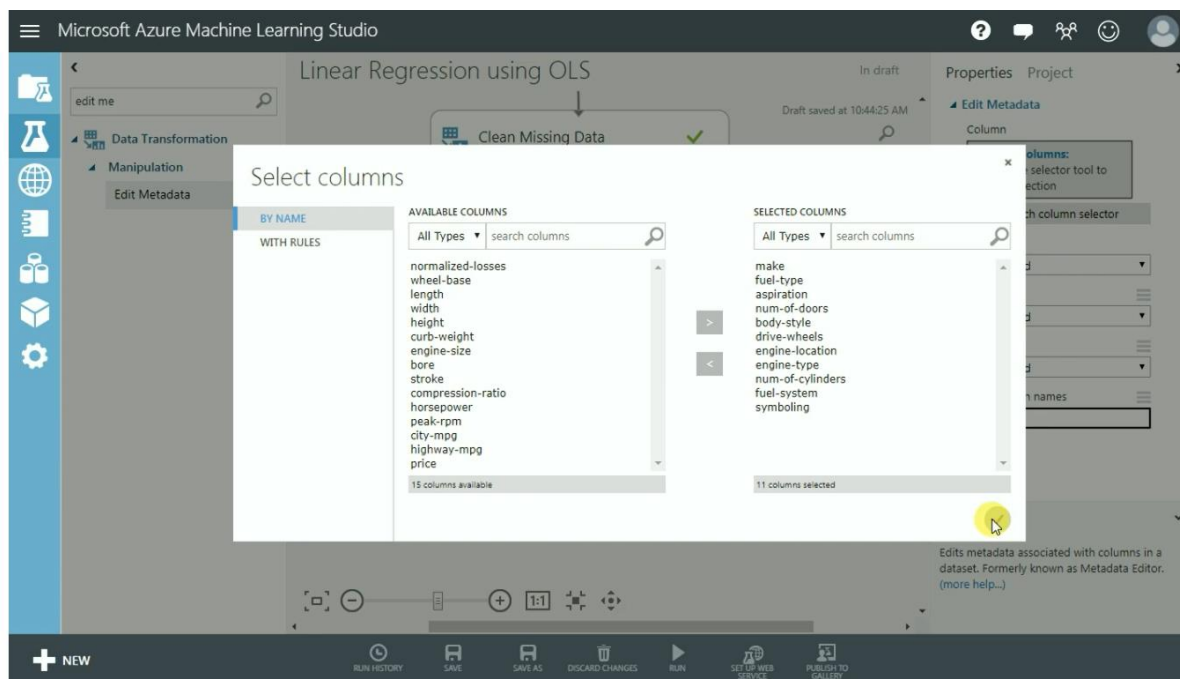
Run the same and visualise if any missing value exists



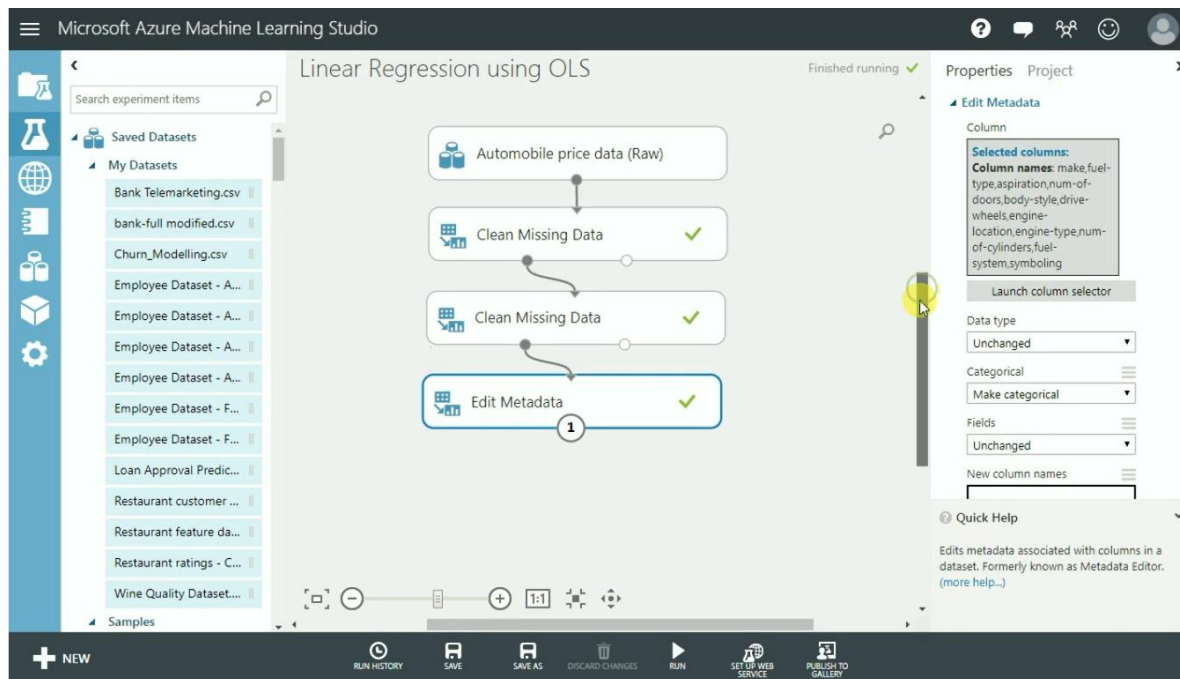
Add metadata and connect the nodes as below



Now launch column selector and select string, include symboling and click ok

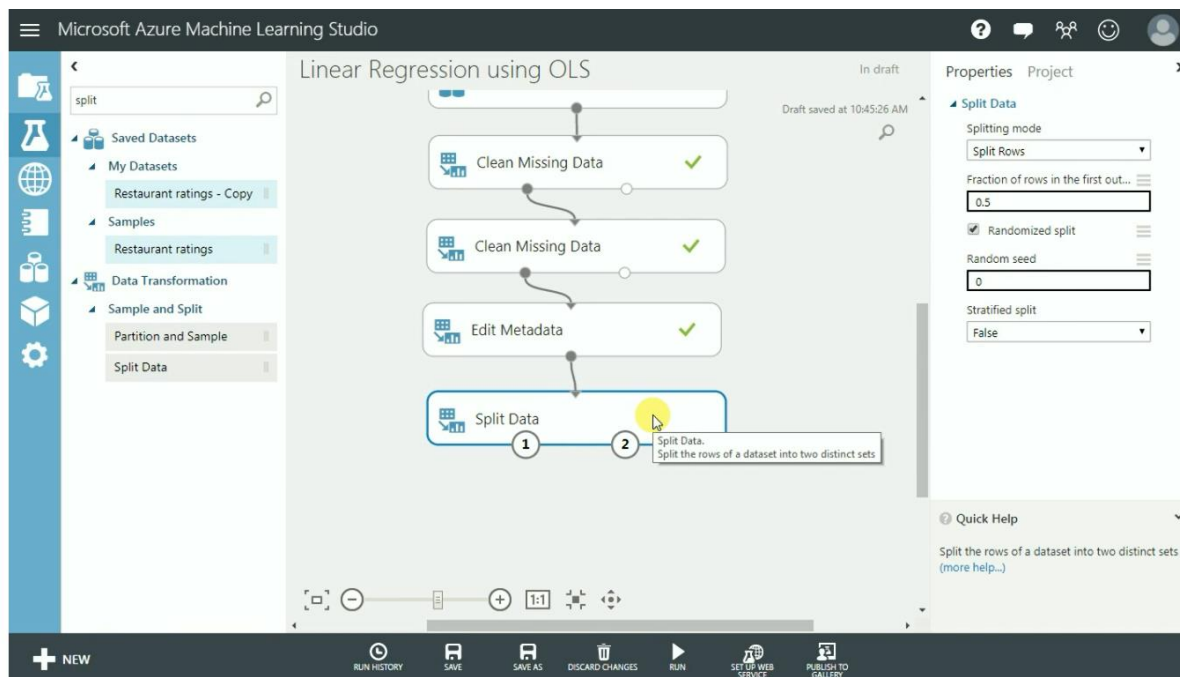


Run the module before application of split data

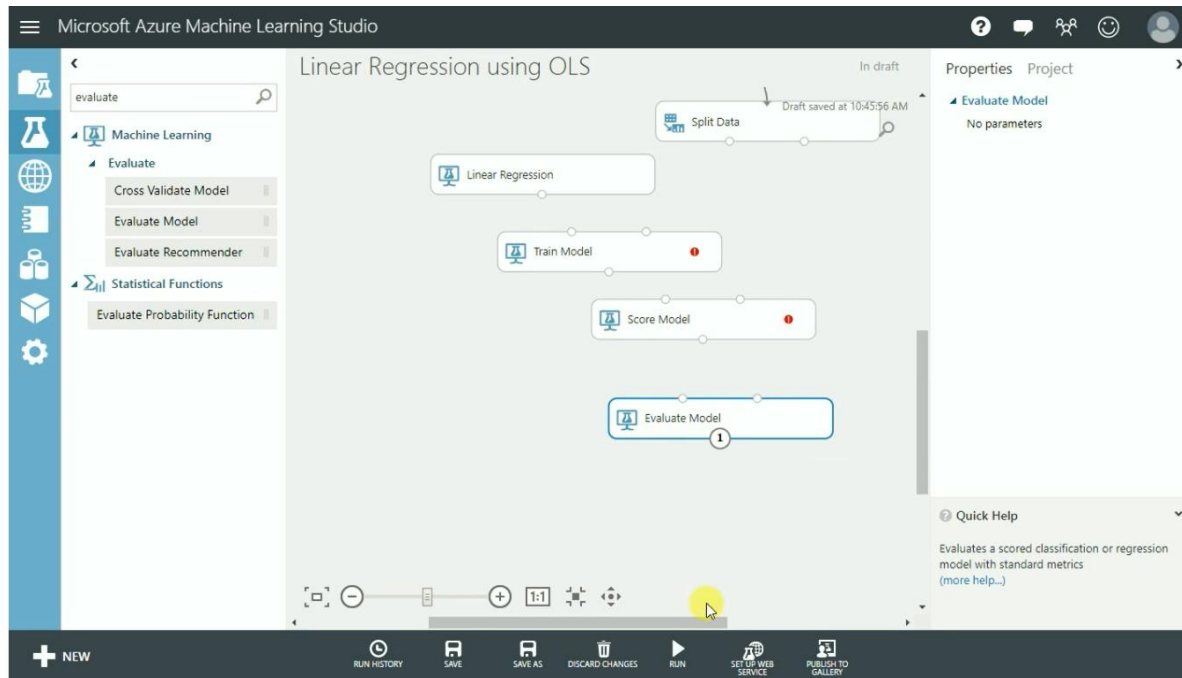


## Dataset for Split data

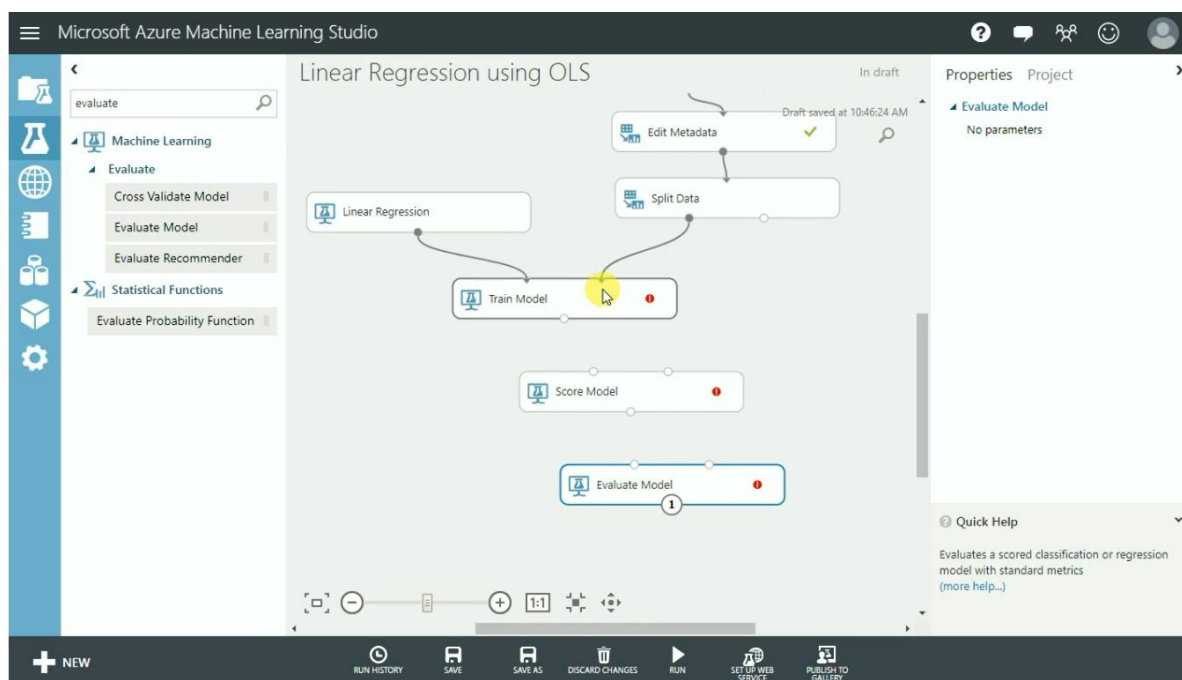
Add split data and connect the nodes



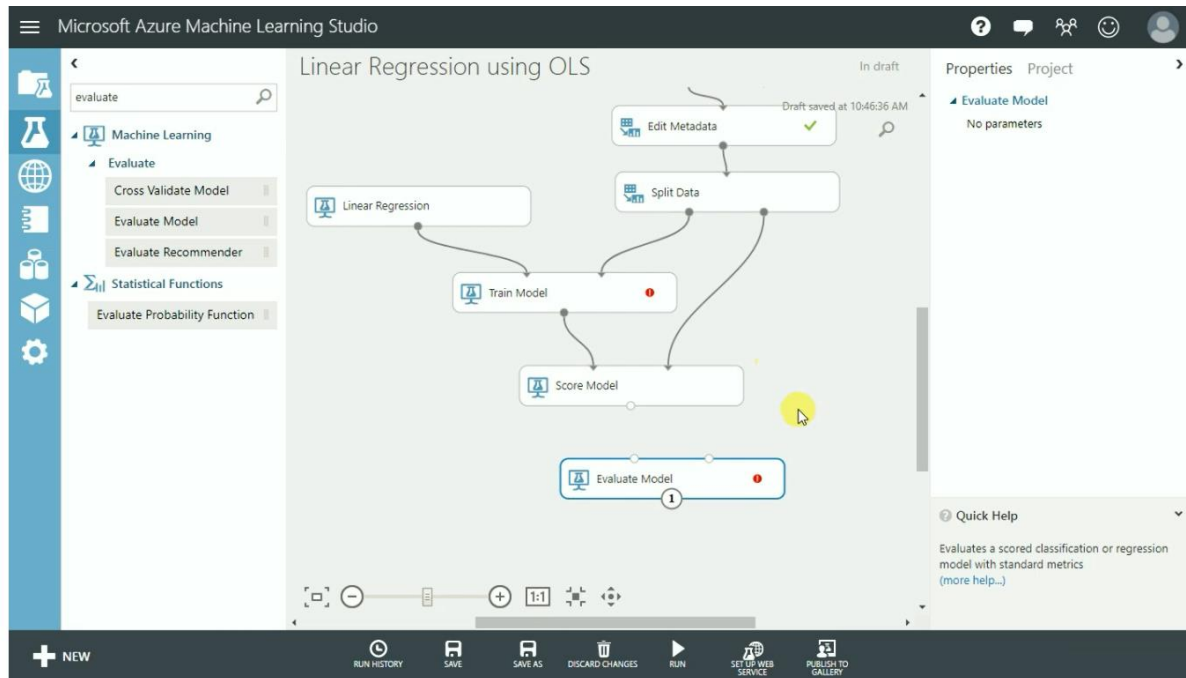
## Linear Regression, train model, score model and evaluate model dataset



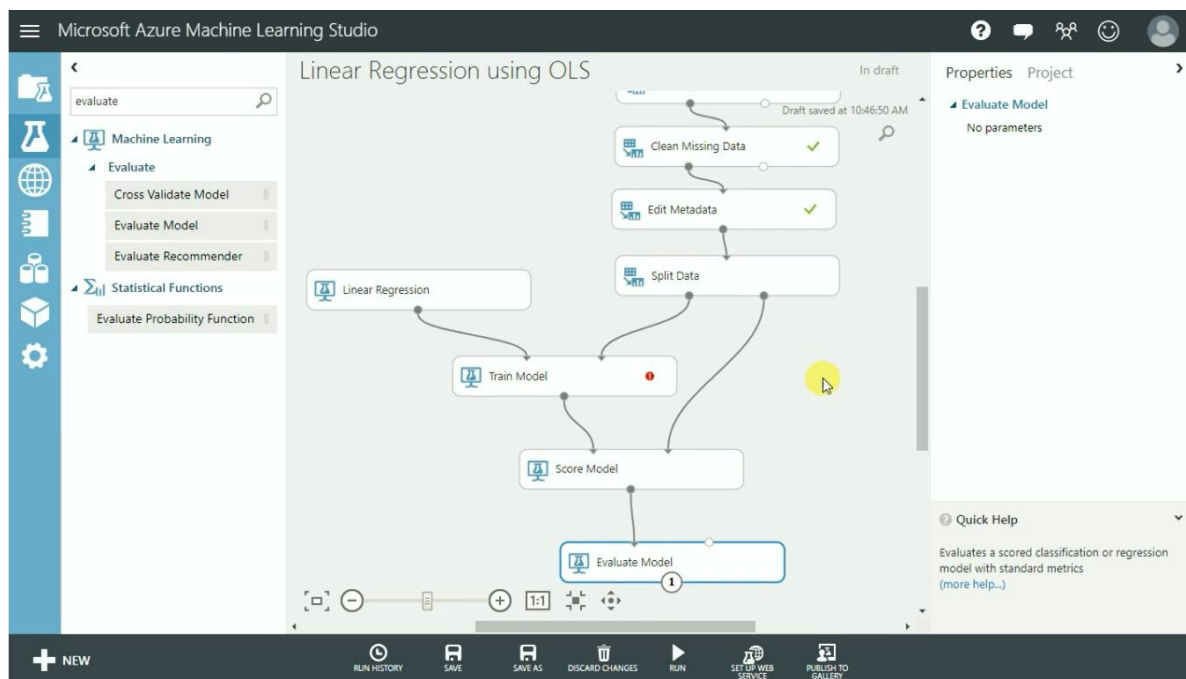
Connect Linear regression node and split data node to train model as shown



Connect train model node 1 and split data node 2 to score model nodes respectively



Now connect score model node to evaluate model



Change the parameters in linear regression as shown

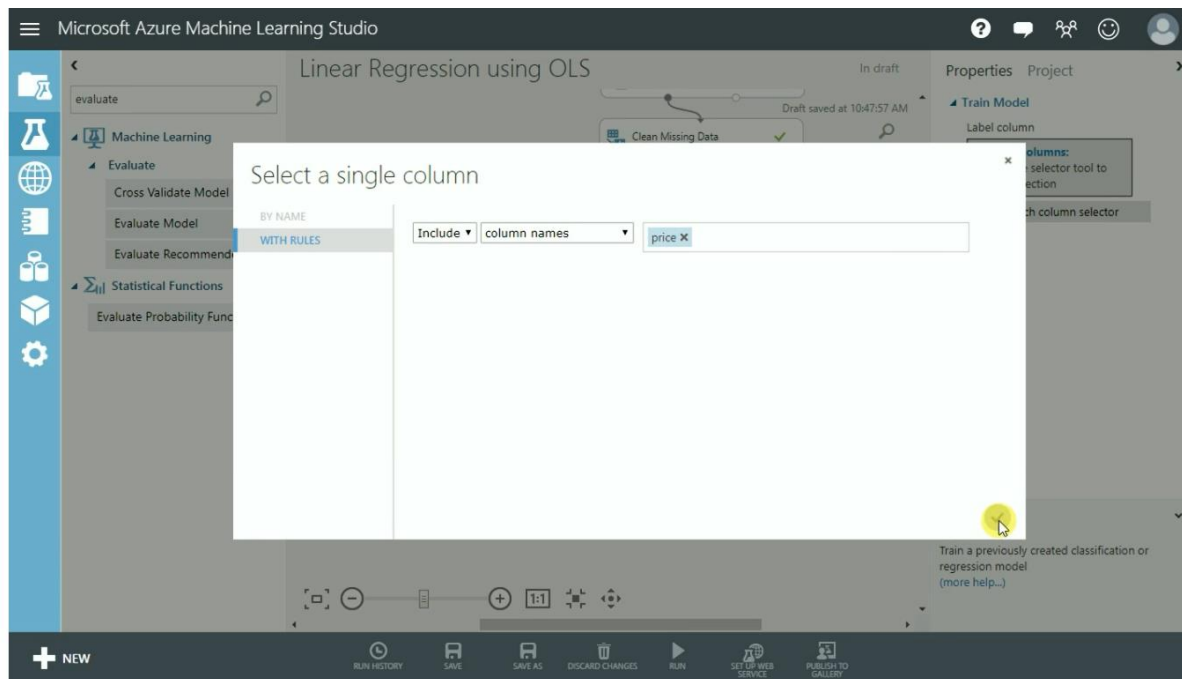
The screenshot shows the Microsoft Azure Machine Learning Studio interface. The central workspace displays a workflow titled "Linear Regression using OLS" in draft mode. The workflow consists of the following modules: "Clean Missing Data" (twice), "Edit Metadata", "Split Data", "Linear Regression", "Train Model", "Score Model", and "Evaluate Model". The "Linear Regression" module is selected, and its properties are visible on the right. The "Properties" pane for the "Linear Regression" module shows the following settings: "Solution method" is set to "Ordinary Least Squares", "L2 regularization weight" is set to "0.001", "Include intercept term" is checked, "Random number seed" is set to "123", and "Allow unknown categories" is checked. The "Quick Help" section at the bottom of the properties pane states: "Creates a linear regression model (more help...)". The bottom toolbar includes buttons for "NEW", "RUN HISTORY", "SAVE", "SAVE AS", "DISCARD CHANGES", "RUN", "SET UP WEB SERVICE", and "PUBLISH TO GALLERY".

Add label column to train model before running the module

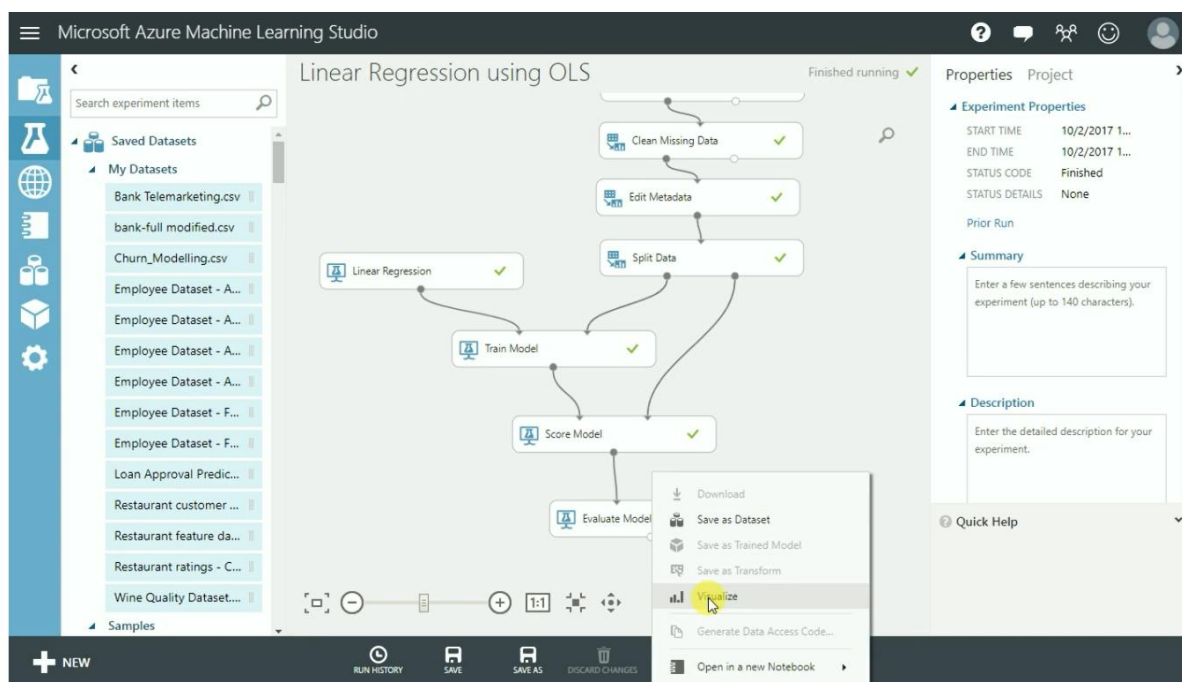
The screenshot shows the Microsoft Azure Machine Learning Studio interface. The central workspace displays a workflow titled "Linear Regression using OLS" in draft mode. The workflow consists of the following modules: "Clean Missing Data" (twice), "Edit Metadata", "Split Data", "Linear Regression", "Train Model", "Score Model", and "Evaluate Model". The "Train Model" module is selected, and its properties are visible on the right. The "Properties" pane for the "Train Model" module shows the "Label column" property, which is currently set to "Selected columns: Launch the selector tool to make a selection". A yellow circle highlights the "Launch column selector" button. The "Quick Help" section at the bottom of the properties pane states: "Train a previously created classification or regression model (more help...)". The bottom toolbar includes buttons for "NEW", "RUN HISTORY", "SAVE", "SAVE AS", "DISCARD CHANGES", "RUN", "SET UP WEB SERVICE", and "PUBLISH TO GALLERY".



Add price column and click ok



Run and visualize the evaluate model now



Observe the result

