

**Azure ML Classic Studio**

**Predicting the Flight delays using Regression Model in Azure ML Classic Studio.**

This model (Pipeline) trains a linear regression to predict when mostly the flights are delayed based on data such as Year ,Carrier ,DayofMonth ,DayOfWeek ,OriginAirportID,DestAirportID,CRSDepTime,DepDelay,DepDel15,CRSArrTime,ArrDelay,ArrDel15,Cancelled .

Gallery Link:

https://studio.azureml.net/Home/ViewWorkspaceCached/b0b7eca0449e412fa995253777238f94?#Workspaces/Experiments/Experiment/b0b7eca0449e412fa995253777238f94.f-id.6dd5af9a56554691b0fe7b3af2595045/ViewExperimentsss

FLIGHTS DELAY PREDICTION[19K41A05A2]

**Machine Learning Project Workflow**

1. Import Data

2. Explore Data (Missing values, outliers)

3. Preprocess data (Missing value imputation, outlier

treatment, normalization)

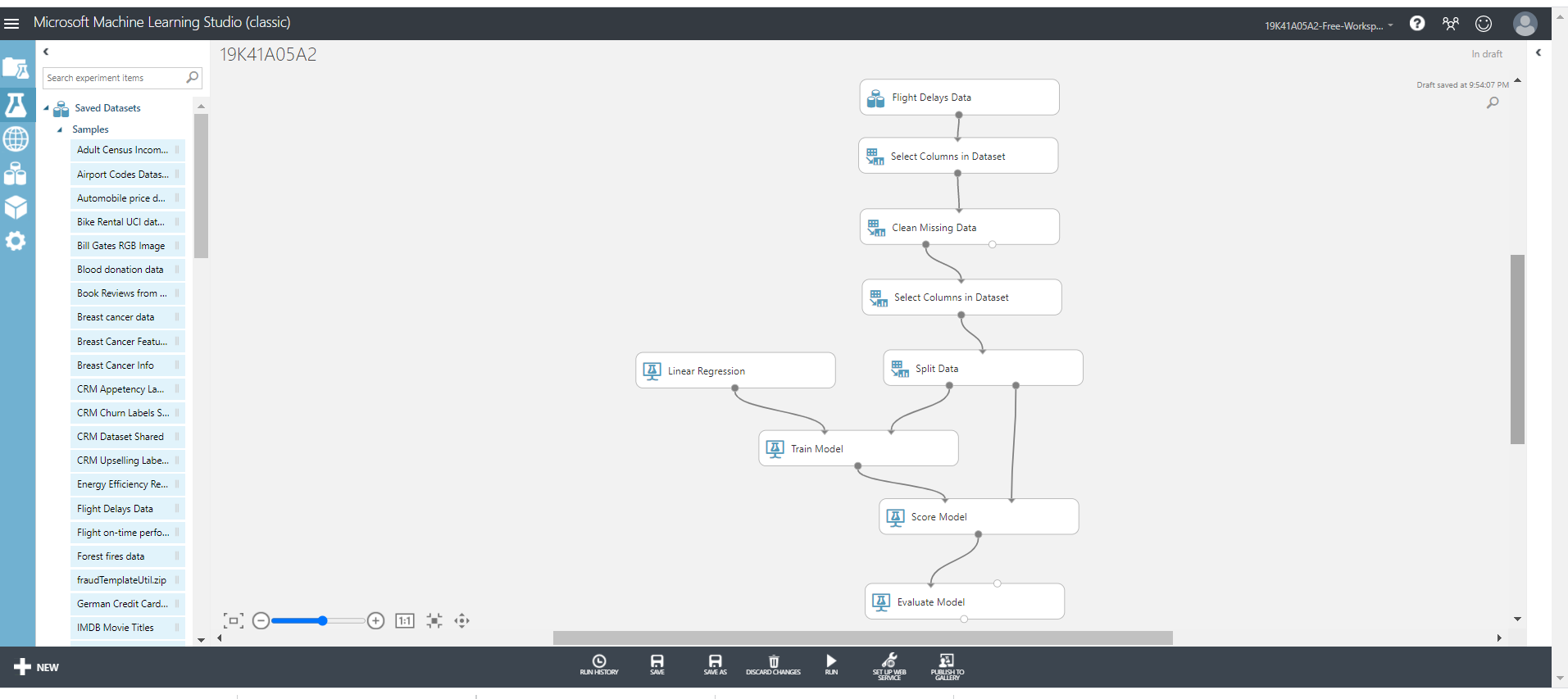
4. Model Selection

5. Model Training

6. Model Testing

7. Model Deployment

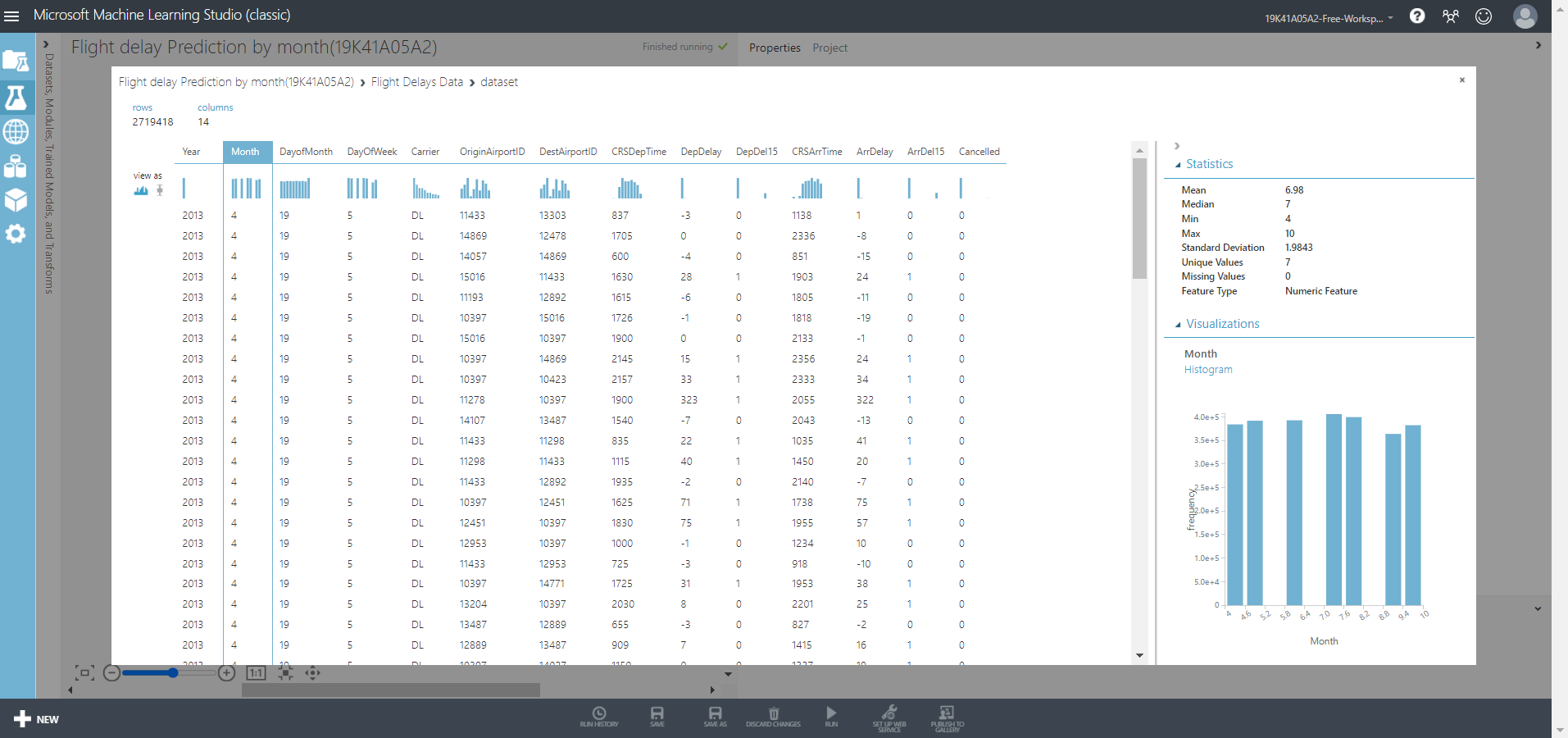
**Workflow**



**Project Workflow**

**Import Data**

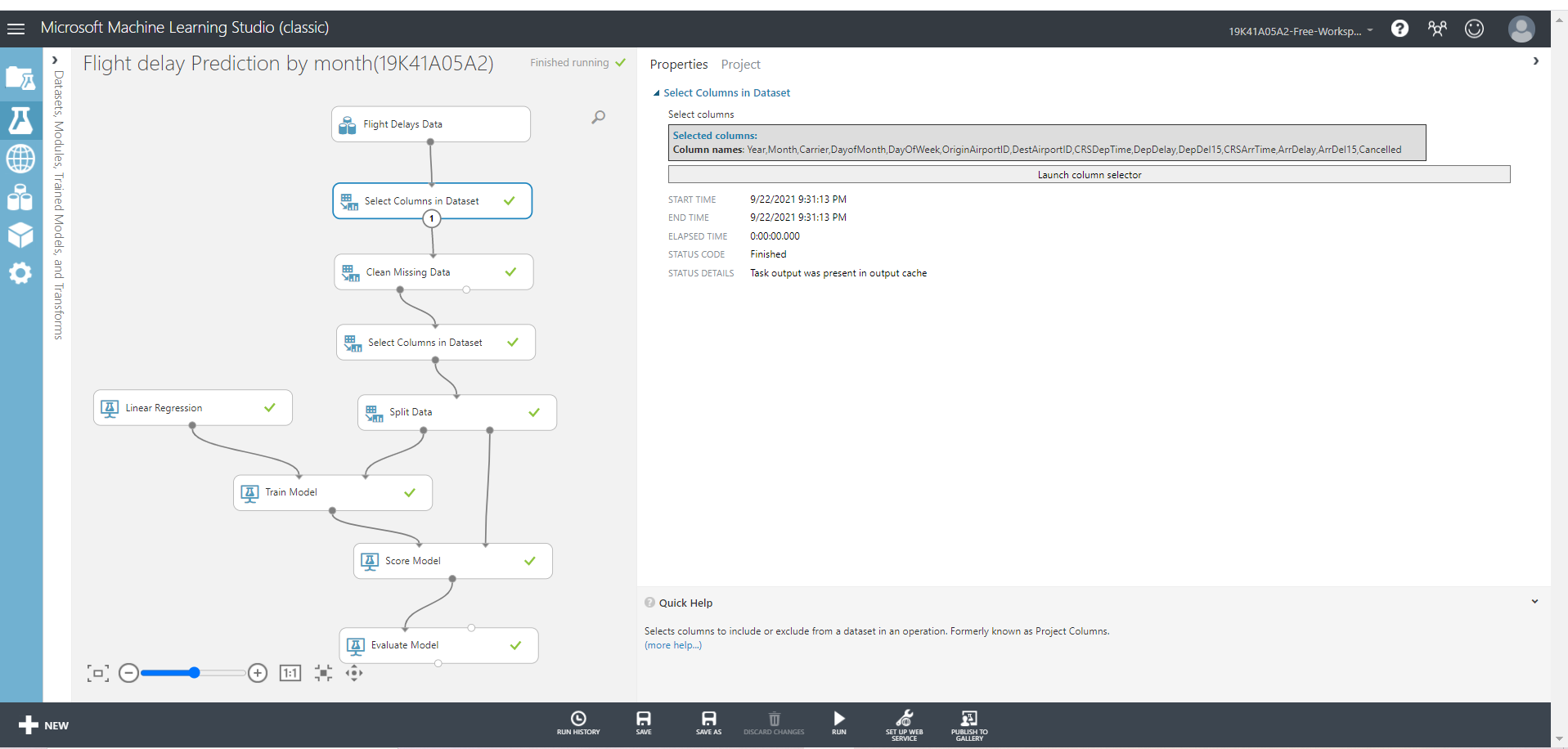
* Importing the RAW dataset which is in CSV format.
* The dataset is pre-available in the Azure ML Classic Studio.

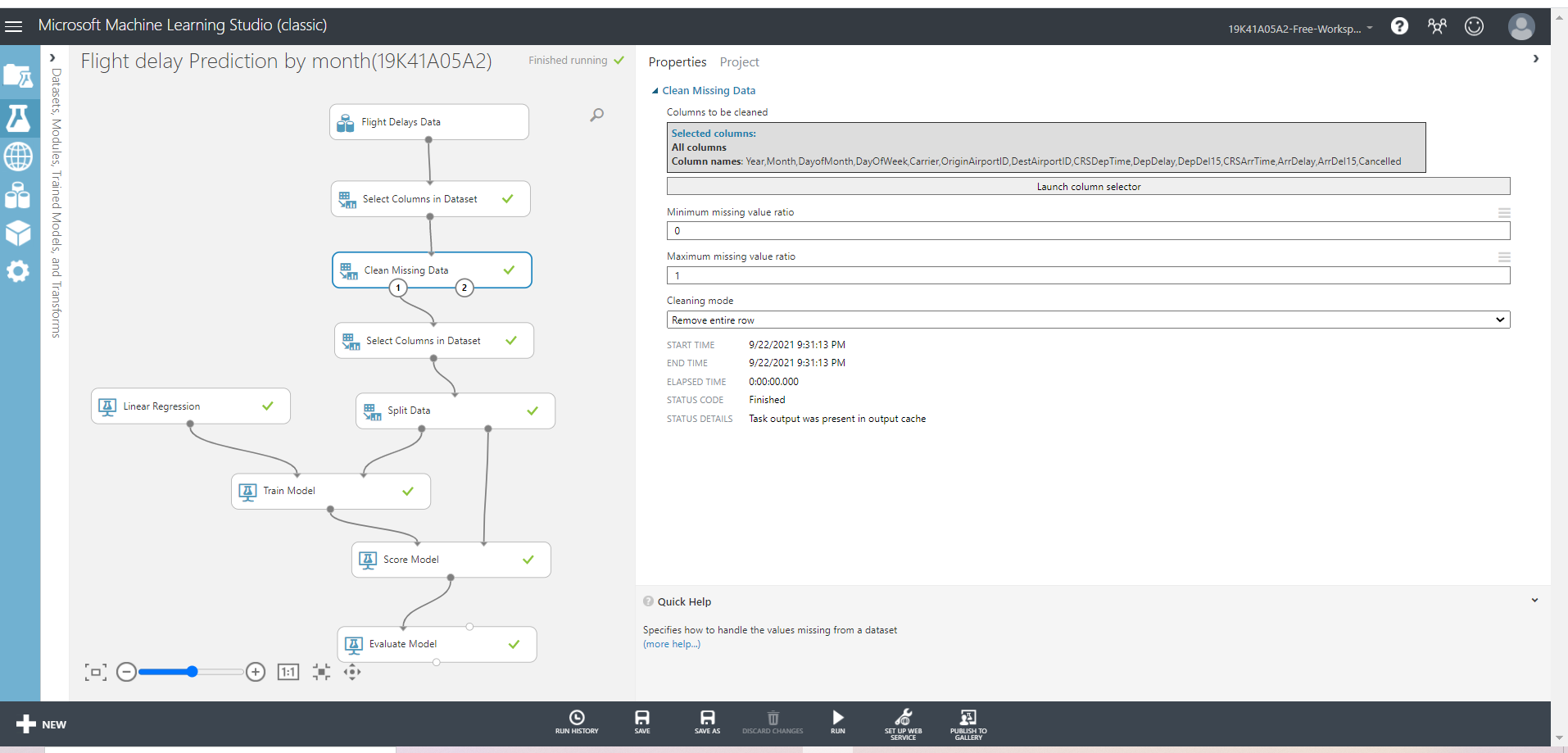


**Flight delays data(CSV format)**

**Explore Data**

* This basically includes data visualization to search for any missing values in the Dataset.
* If any missing values are found, then they needs to be cleaned.
* selecting the required columns and clean the data using the Clean Missing Value module (Just Drag and Drop )

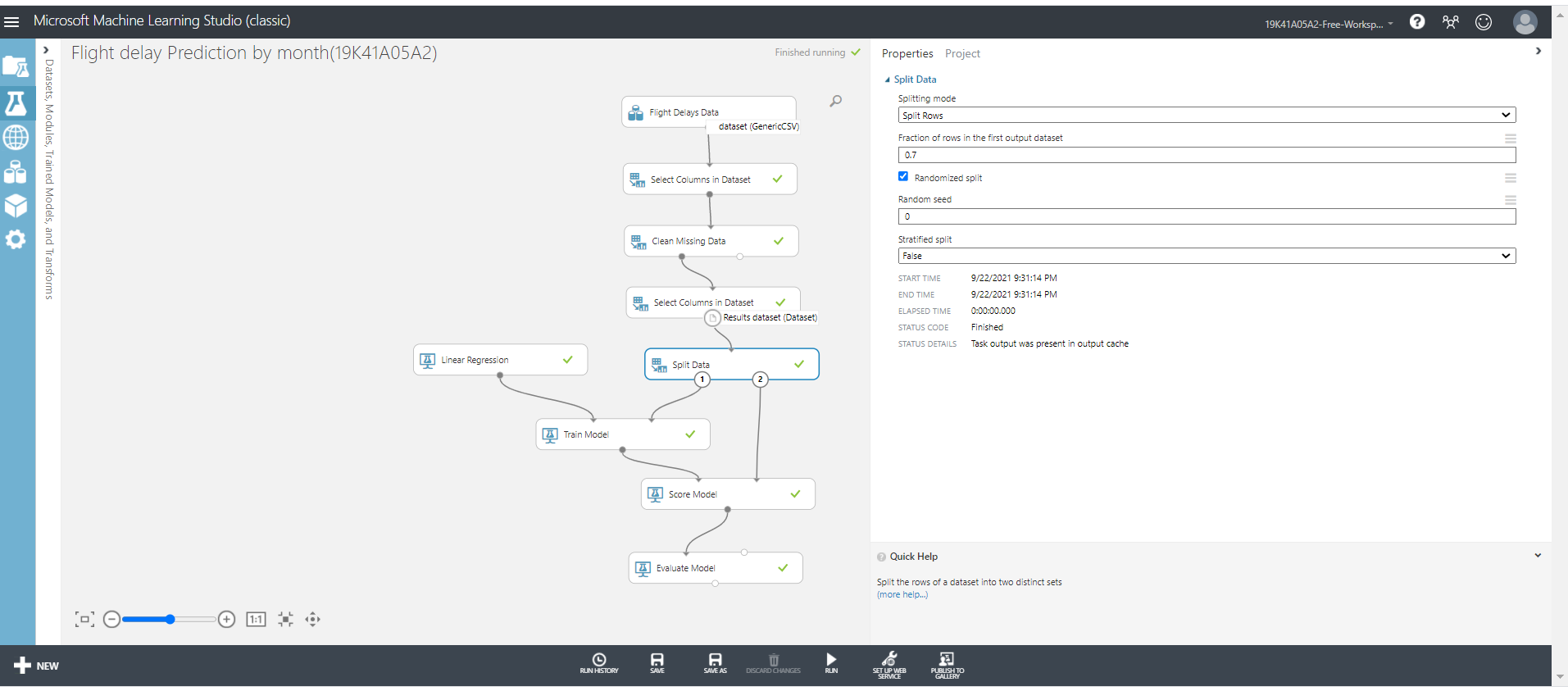




**Data Cleaning**

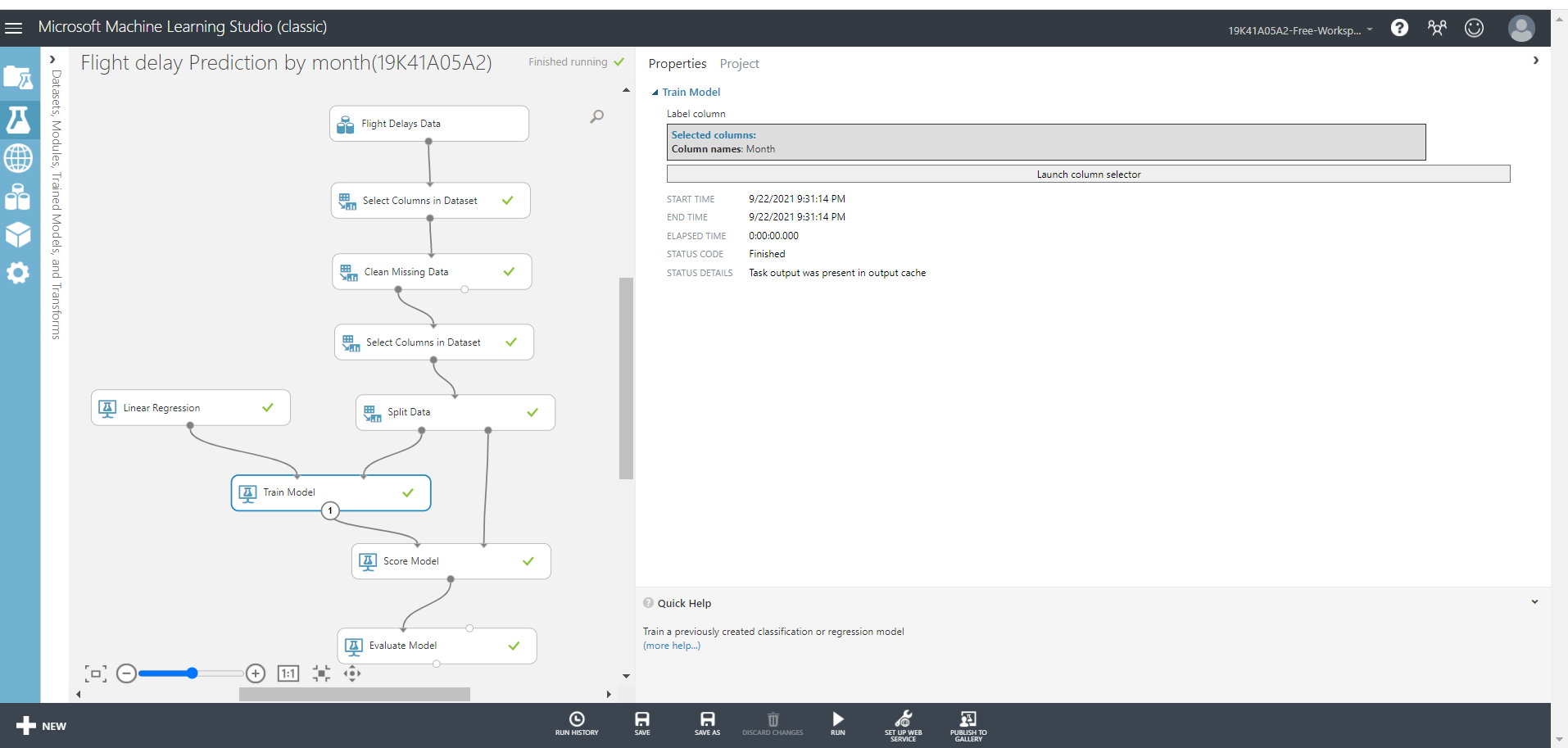
**Split Data**

* Use the Split Data module to randomly divide the input data so that the training dataset contains 70% of the original data and the testing dataset contains 30% of the original data.



**Data Splitting**

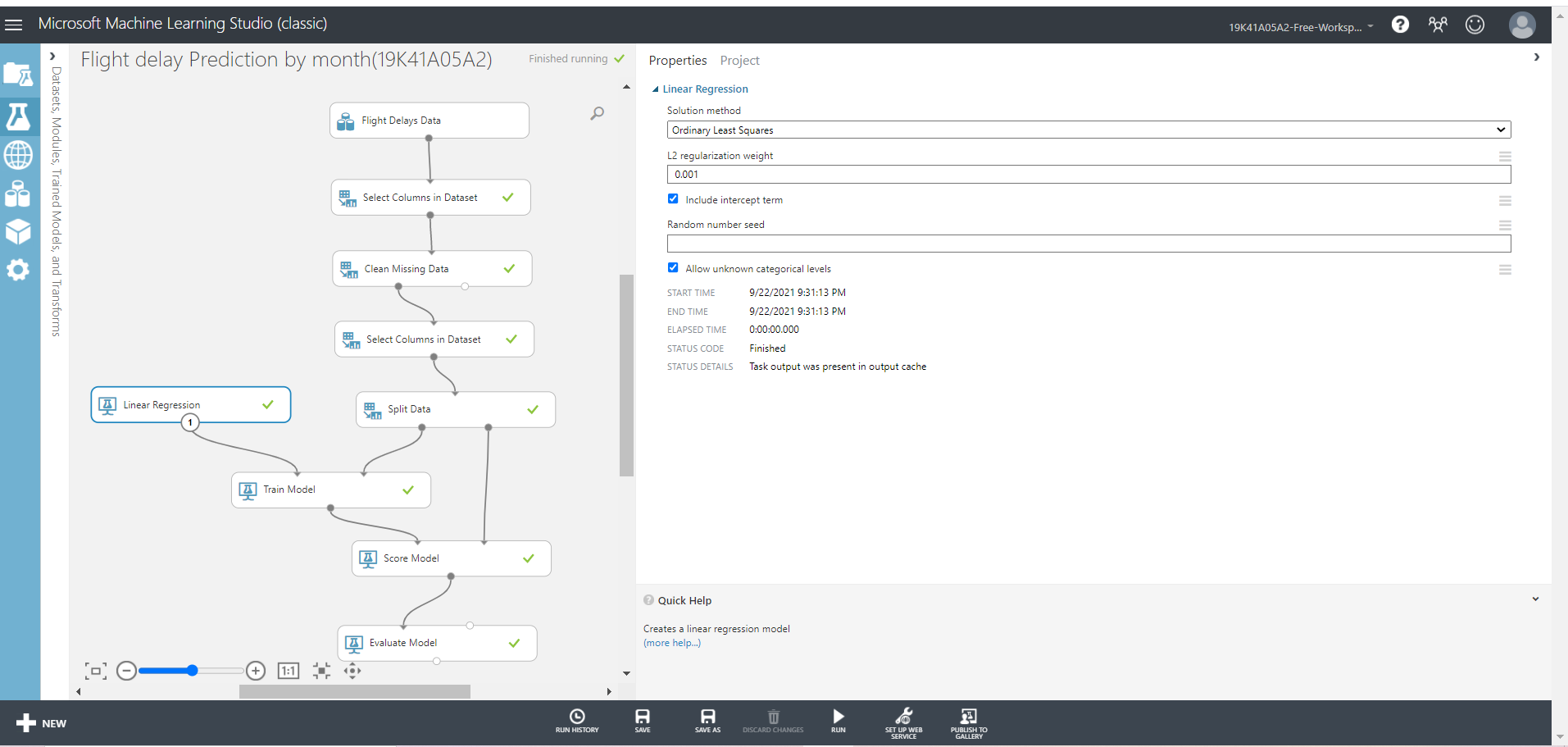
**Model Training and Algorithm**



**Model Training**

**Using Linear regression to train the model**

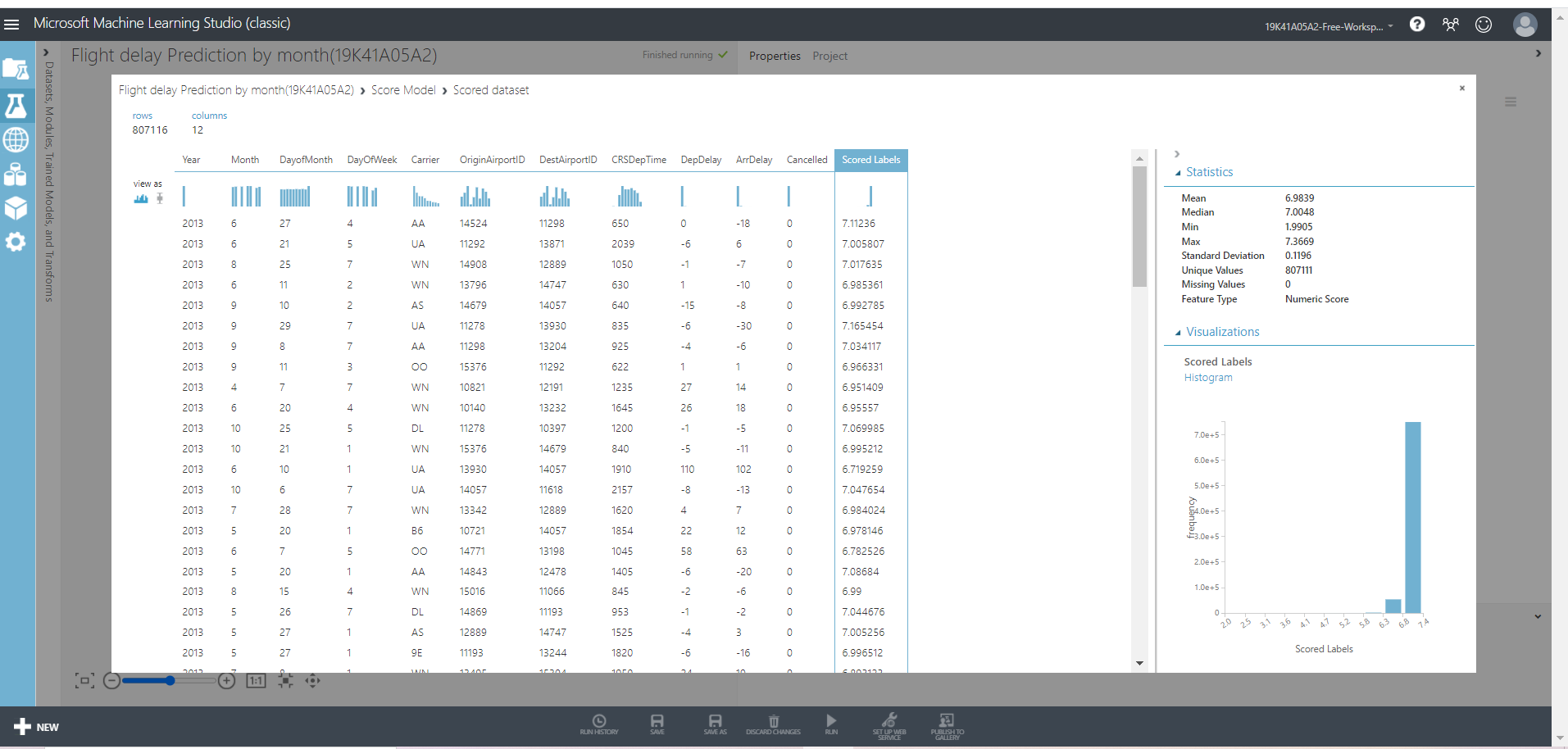
* Since the goal of this sample is to predict flight delays, and because the label column is continuous data, a regression model can be a good choice.We use Linear Regression for this pipeline.



**Linear regression**

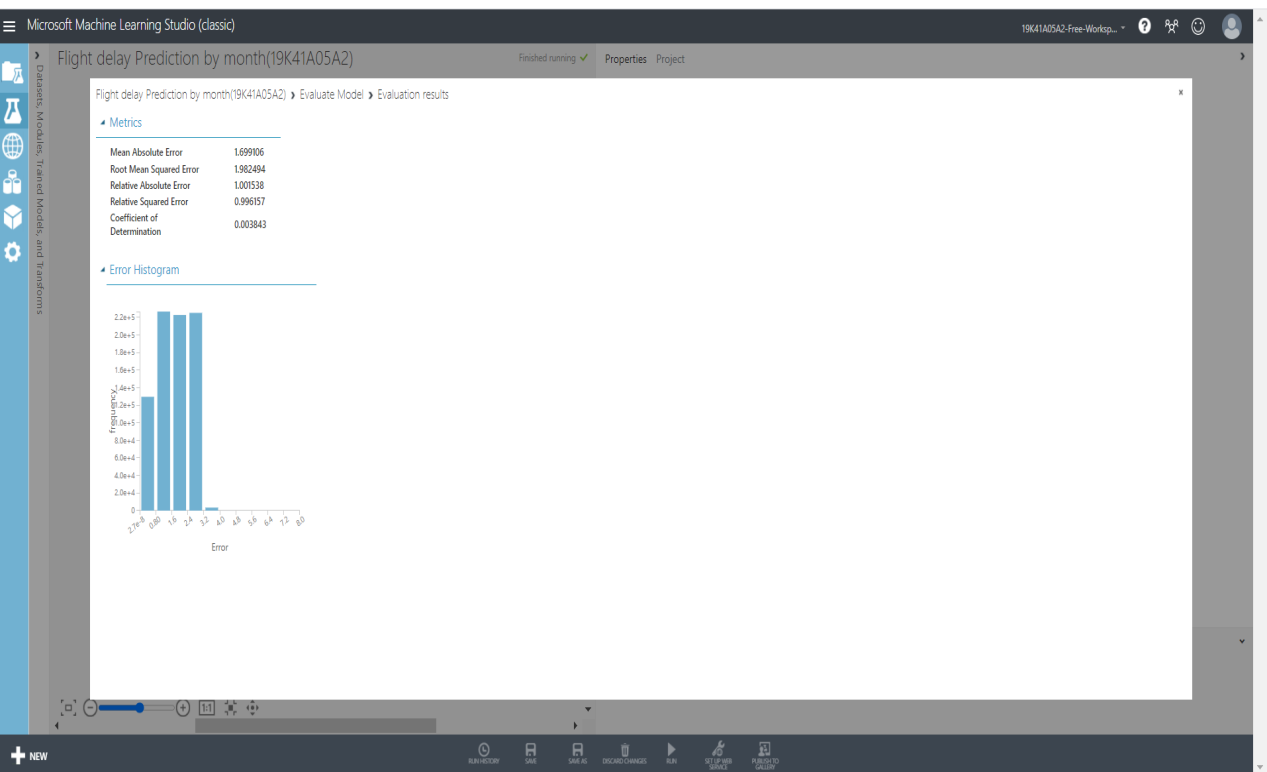
**Score Model and Evaluate Model**

* After the model is trained, we can use the Score Model and Evaluate Model modules to generate predicted results and evaluate the models.



**Score Labels**

**Evaluation Results**

****

**Model Evaluation Results**