|  |
| --- |
| **Experiment No. 9**  **Title: Case study: Big data platform / analytics as business need)** |

**Batch: B1** **Roll No.: 1824005** **Experiment No.:9**

**Title: Case study:**

# Regression: Demand Estimation

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Resources needed:**

Microsoft Azure Machine Learning Studio

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Describe the following points with respect to the business under consideration,**

1. **Problem faced by the business**

# Regression: Demand Estimation

This experiment demonstrates the **feature engineering** process for building a **regression** model using bike rental demand prediction as an example. I have demonstrate that effective feature engineering will lead to a more accurate model.

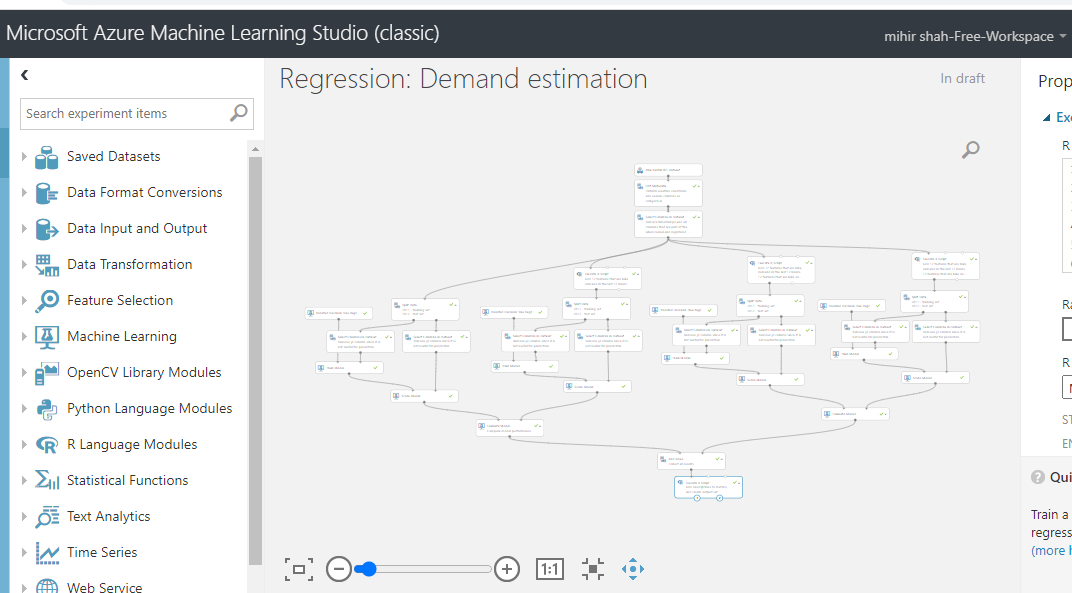
1. **Approach/ Methodology followed by the business**

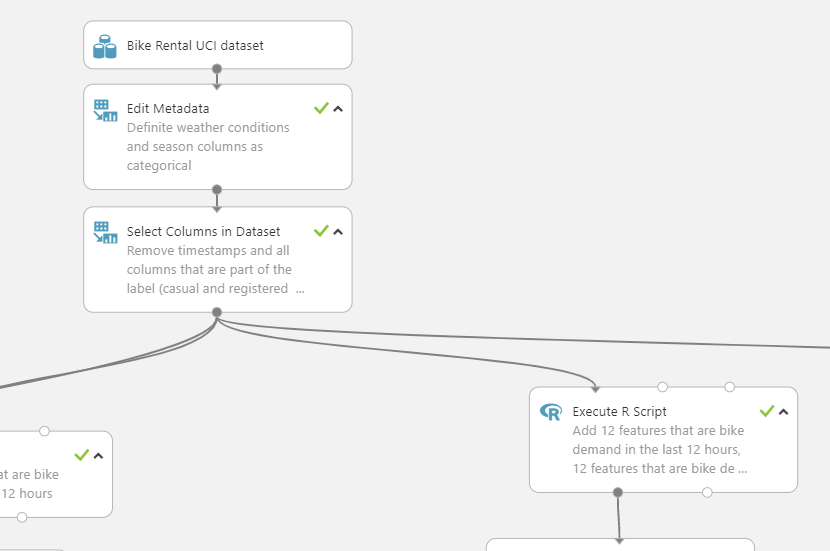
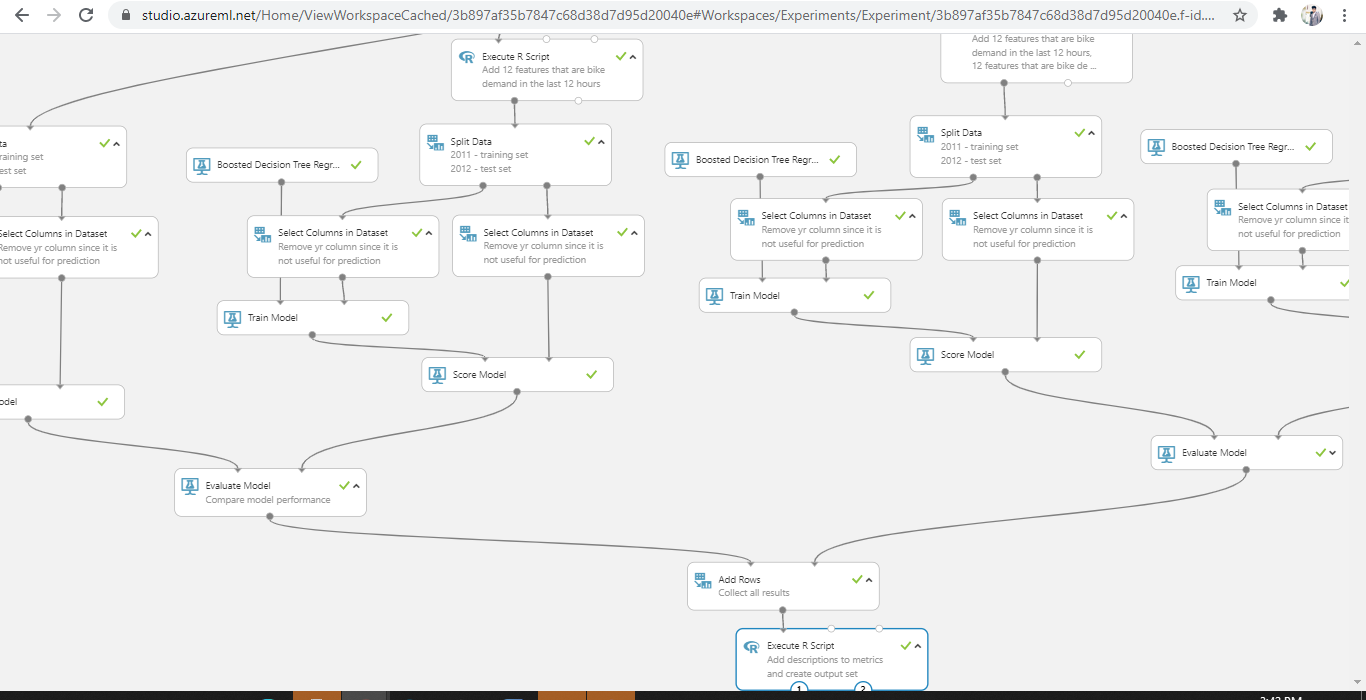
## Data

The Bike Rental UCI dataset is used as the input raw data for this experiment. This dataset is based on real data from the Capital Bikeshare company, which operates a bike rental network in Washington DC in the United States.

The dataset contains 17,379 rows and 17 columns, each row representing the number of bike rentals within a specific hour of a day in the years 2011 or 2012. Weather conditions (such as temperature, humidity, and wind speed) were included in this raw feature set, and the dates were categorized as holiday vs. weekday etc.

The field to predict is "cnt", which contain a count value ranging from 1 to 977, representing the number of bike rentals within a specific hour.



**Skillsets , infrastructure and other impact on the business during implementation**

## Model

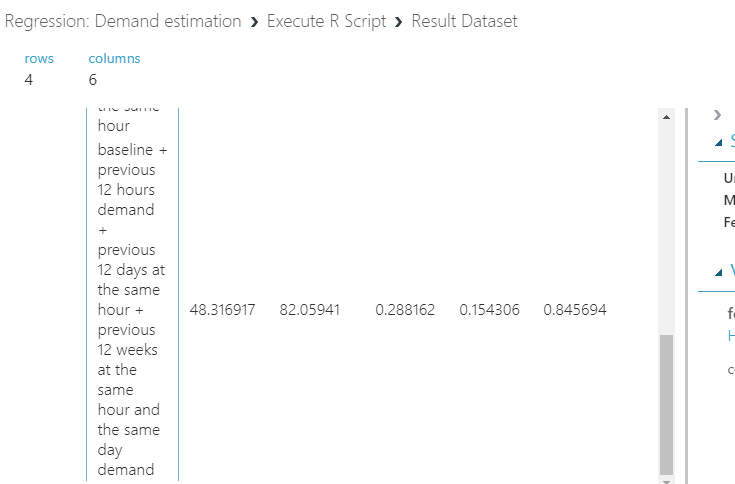
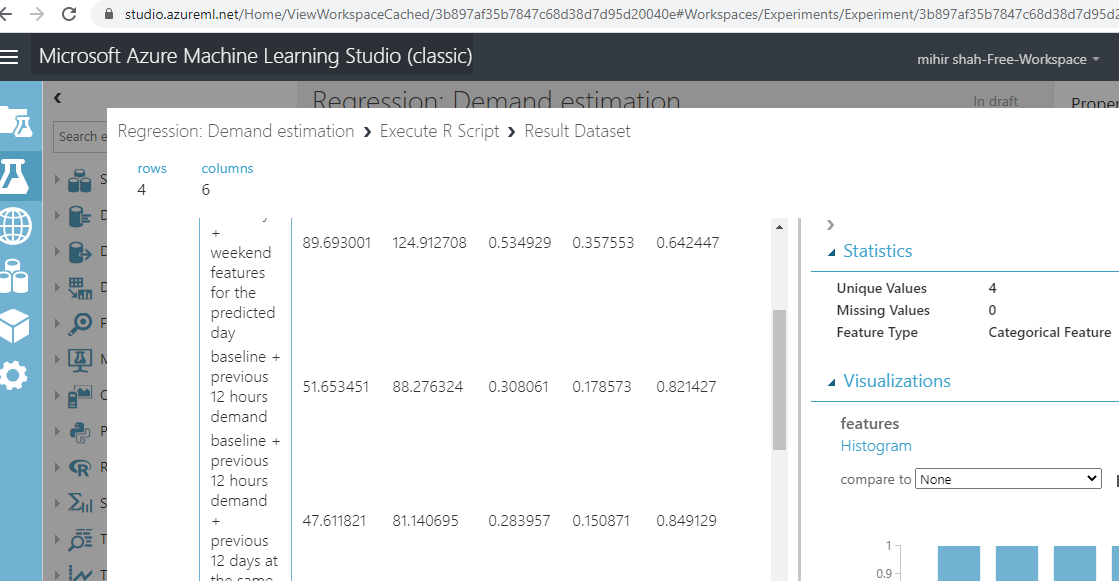
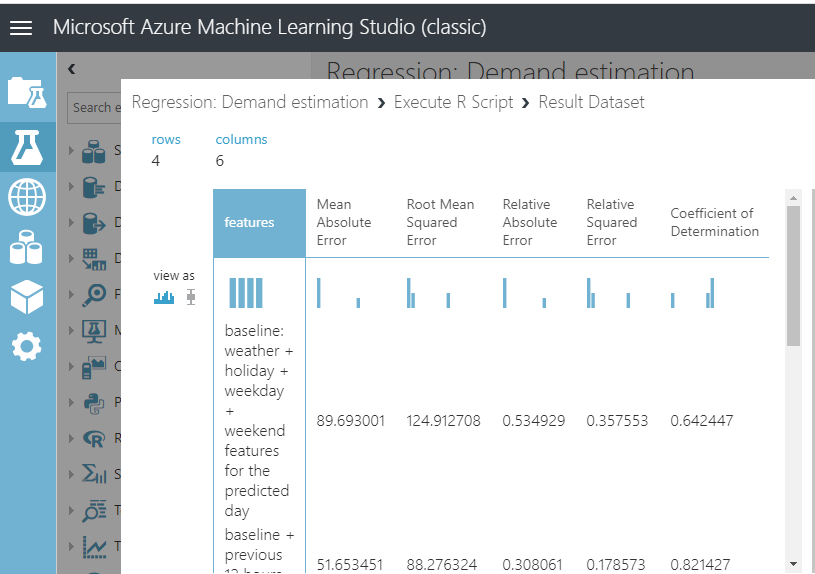
We used a regression model because the label column (number of rentals) contains continuous real numbers.

Given that the number of features is relatively small (less than 100) and these features are not sparse, the decision boundary is very likely to be nonlinear. Based on these observations, we decided to use the **Boosted Decision Tree Regression** algorithm for the experiment.

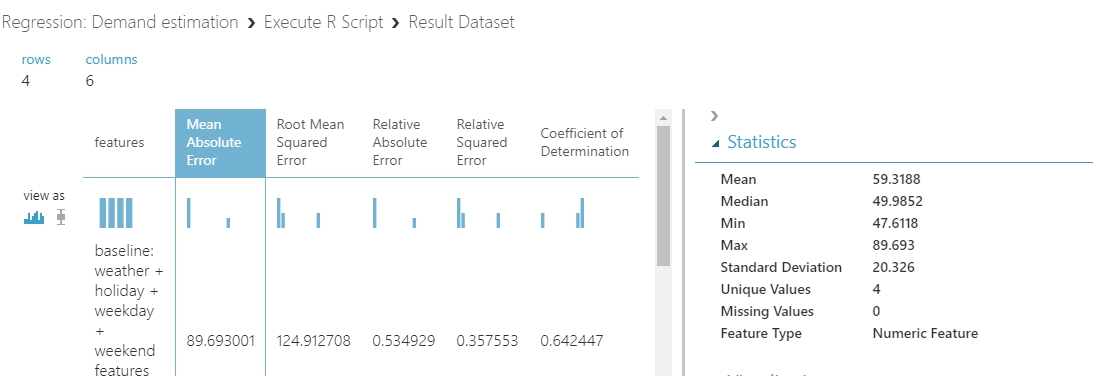
## Running the Experiment

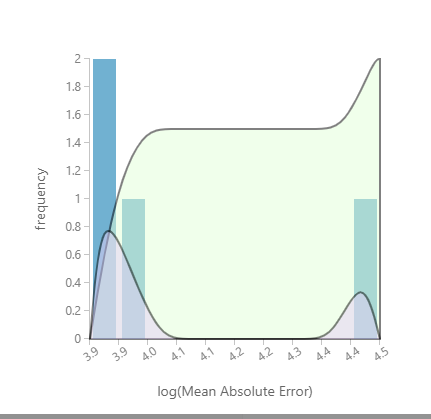
Overall, the experiment had five major steps:

* Step 1: Get data
* Step 2: Data pre-processing
* Step 3: Feature engineering
* Step 4: Train the model
* Step 5: Test, evaluate, and compare the model

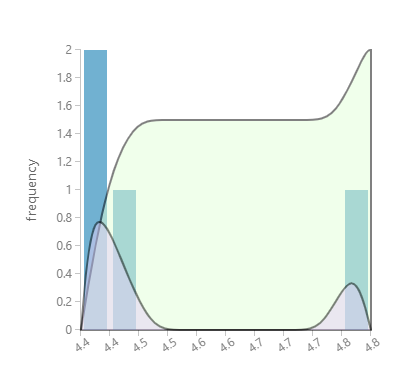
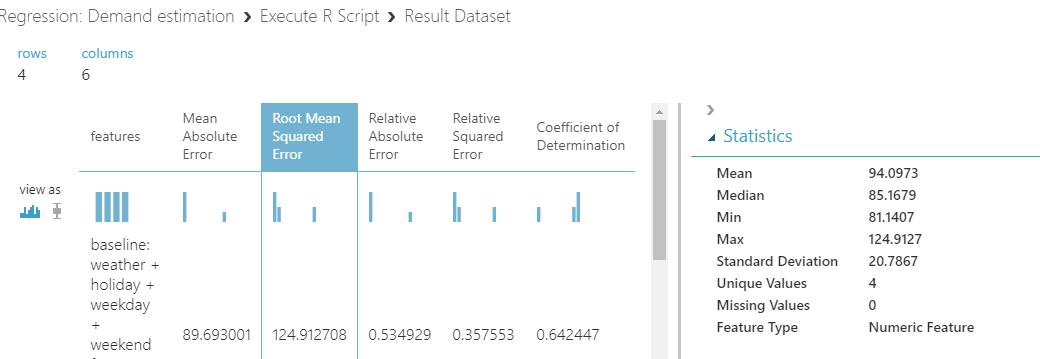


Mean Absolute Error Features

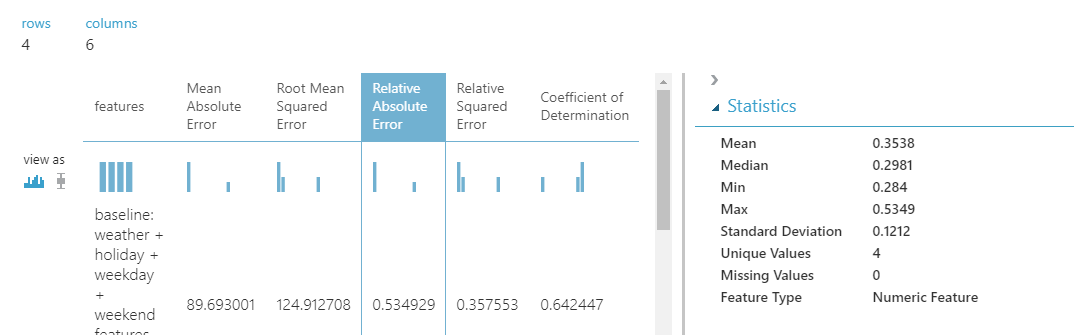


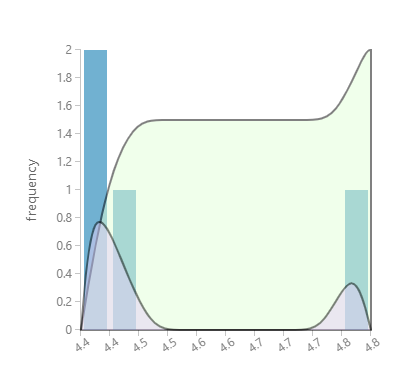


Root Mean Square Features

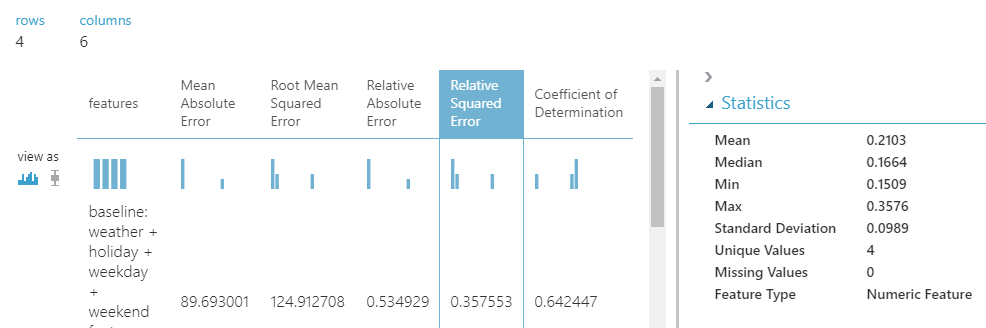


Realitive Absolute error

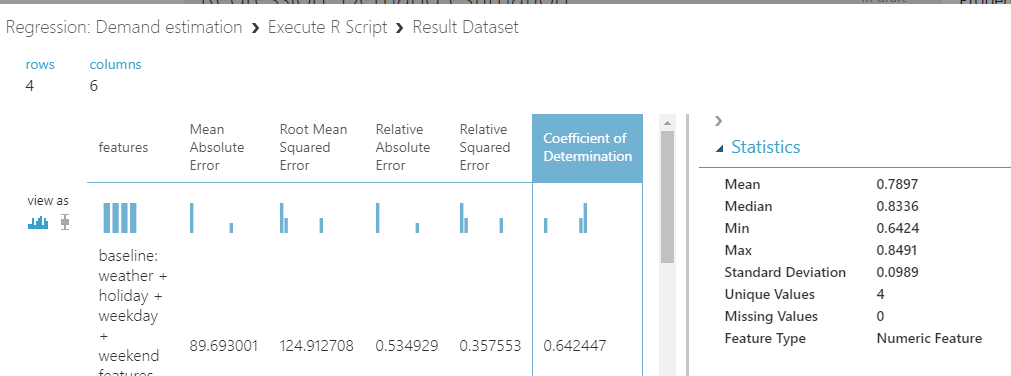


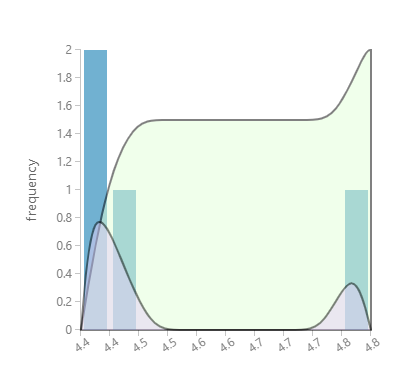


Relative square error



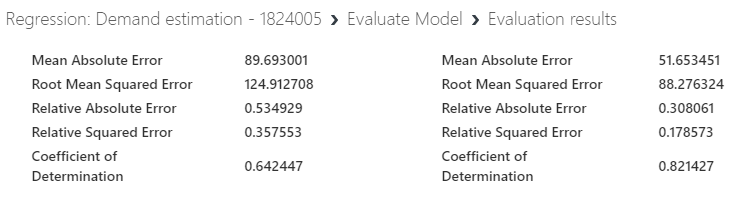
Coefficient of determination

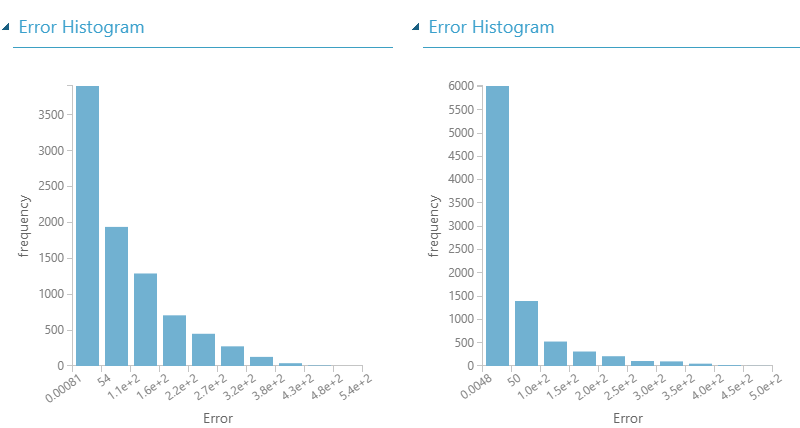




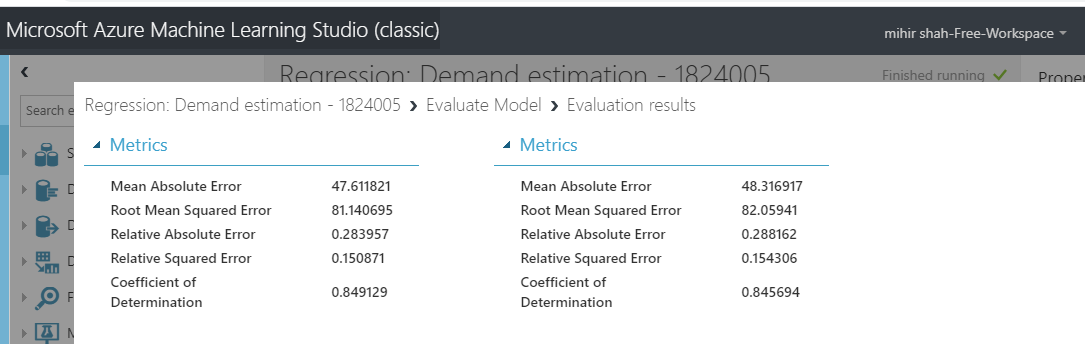
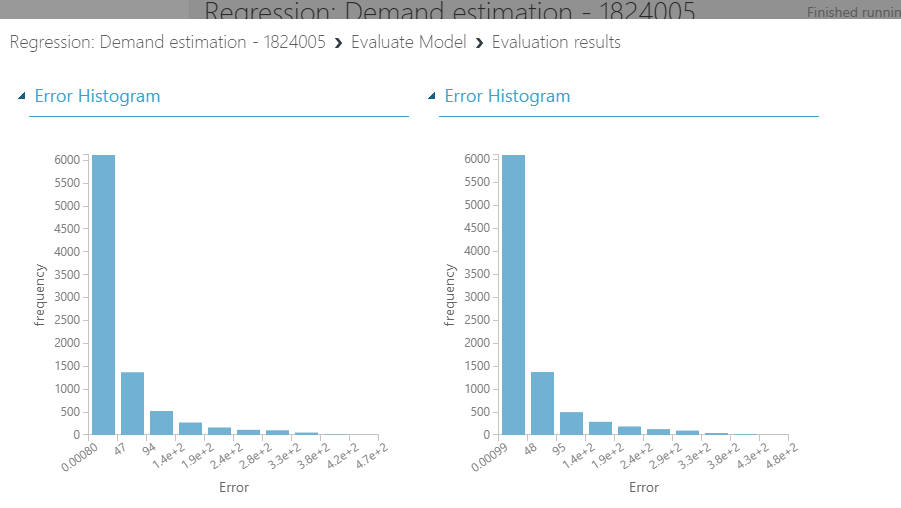
Evaluatiom Results

Model1





Model2

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Questions:**

Discuss the tangible and intangible benefits the business has observed after the implementation.

The tangible and intangible benefits the business has observed after the implementation used to oversee and strategically control specific endeavors. It takes time and money to manage a project, and the results of project management are difficult to separate from the results of the project as a whole. However, project management itself lends both tangible and intangible benefits to an organization.

## Budget Management

One of the major tangible benefits of project management is budgetary savings. Project managers control budgets and make decisions about how best to allocate resources in the process of working toward a project's objectives. In the end, the difference between coming in under budget or having an overrun is a function of workforce efficiency and project management. The money that project managers save their businesses increases the project's return on investment and remains available for future projects.

## Managing Risk

Rick management is much more difficult to quantify than cost savings. Project managers can anticipate risk and guide a project to best avoid it. While there are some financial metrics for placing a tangible value on risk, project management also has the intangible benefit of taking on risks only when they are necessary or worthwhile in the first place. A skilled project manager can lead a project team toward its objectives but remain flexible to react to changing rick factors.

## Meeting Deadlines

Project management is also concerned with managing time and meeting deadlines. Completing a project on time can be as important as completing it on budget. Projects that take too long to complete cost money in terms of overtime wages, lost productivity and time spent modifying schedules and timelines. Project management's impact on meeting deadline is a tangible benefit when the costs of late completion are known. For example, if a business spends $100,000 each day operating a factory to meet a production quota, the tangible benefit of project management that competes the project on time is $500,000 for every week of estimated late delivery.

## Teamwork

Another intangible benefit of project management is its effect on teamwork within an organization. One of the key areas project management deals with is the allocation of human resources. Successful project management gets the most out of each worker and fosters an environment of cooperation and mutual responsibility that can remain long after the project is completed. The intangible benefit of a productive, collaborative workforce is a fundamental part of a successful business's organizational culture despite the impossibility of assigning a dollar value to it

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Outcomes: Realise adequate perspectives of big data analytics in various applications**

**Conclusion: (Conclusion to be based on the objectives and outcomes achieved)**

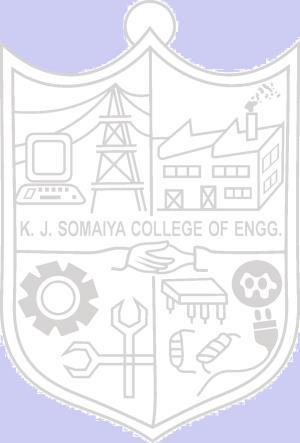
In the given experiment I have demonstrates the **feature engineering** process for building a **regression** model using bike rental demand prediction as an example. I have demonstrate that effective feature engineering will lead to a more accurate model.

**Grade: AA / AB / BB / BC / CC / CD /DD**

**Signature of faculty in-charge with date**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**References:**



**Books/ Journals/ Websites:**