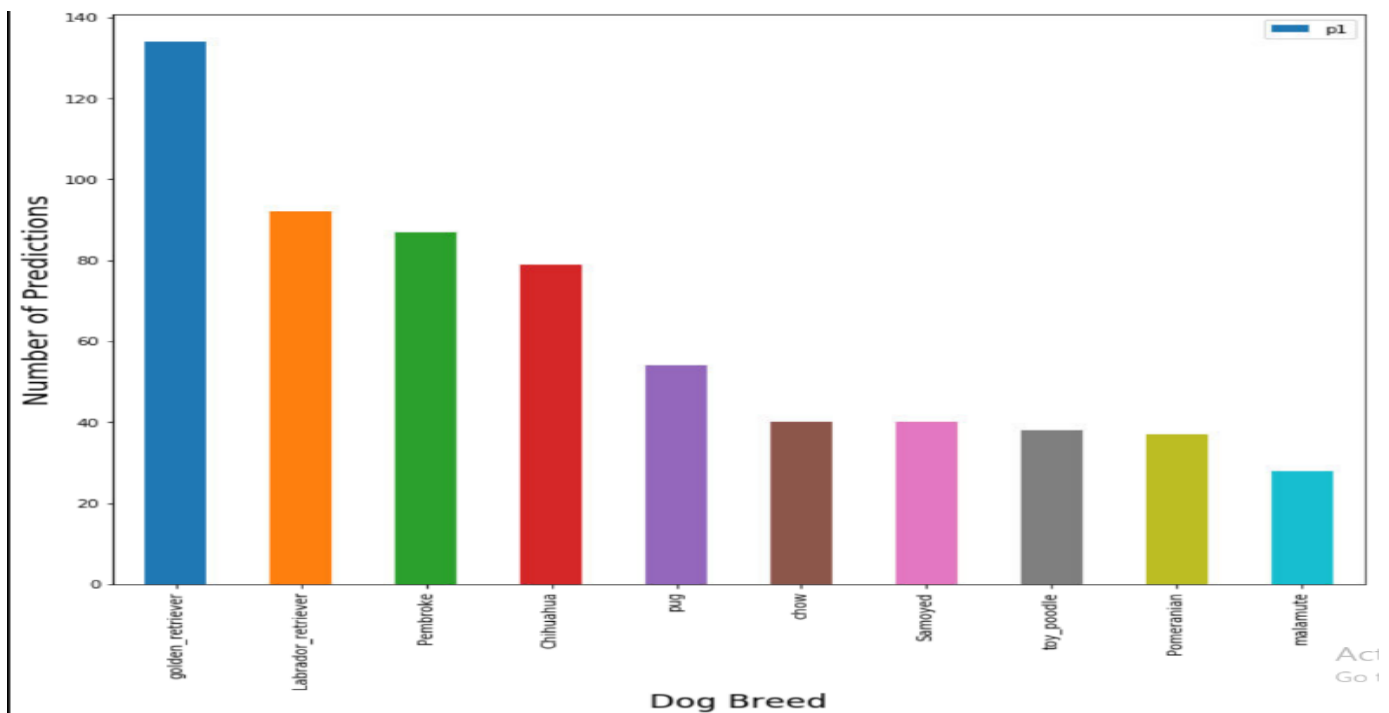# Analyzing and Visualizing WeRateDogs
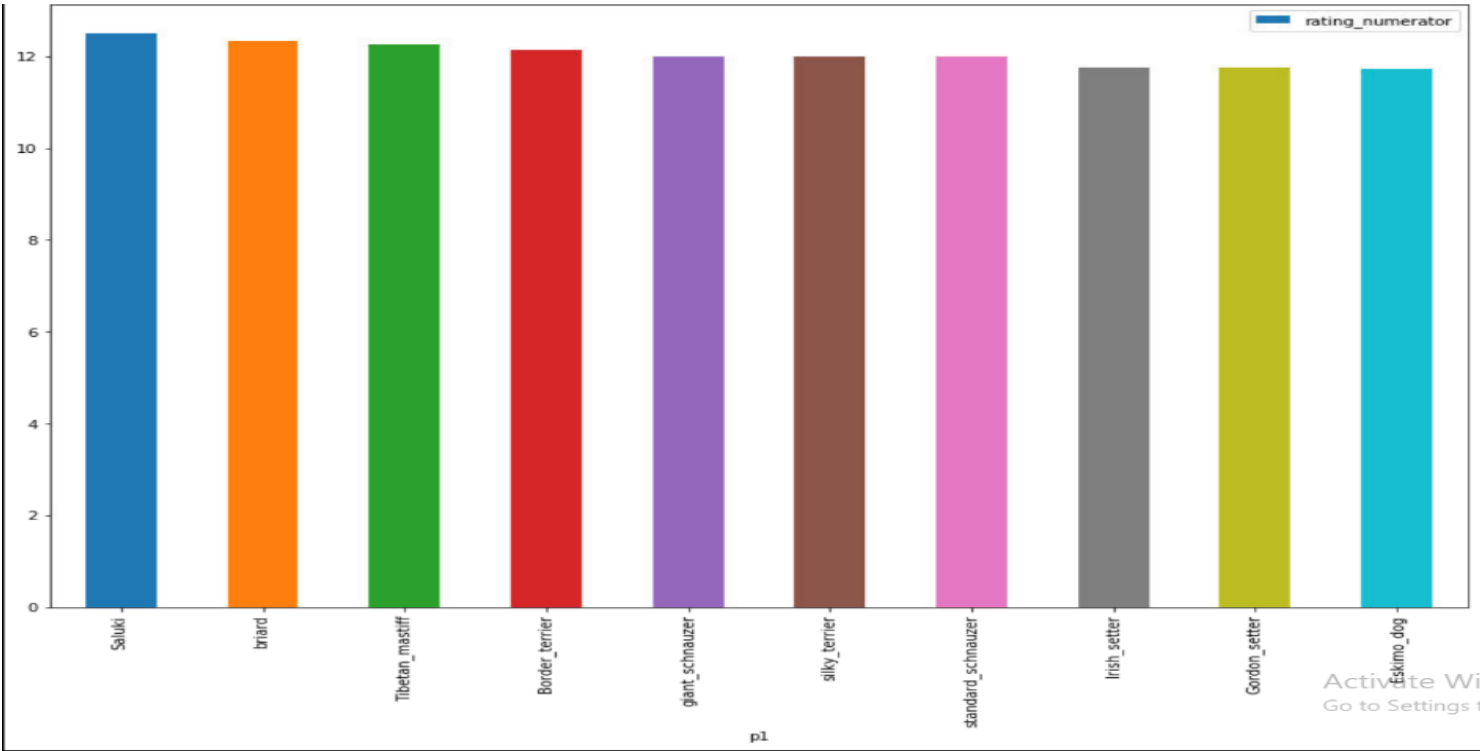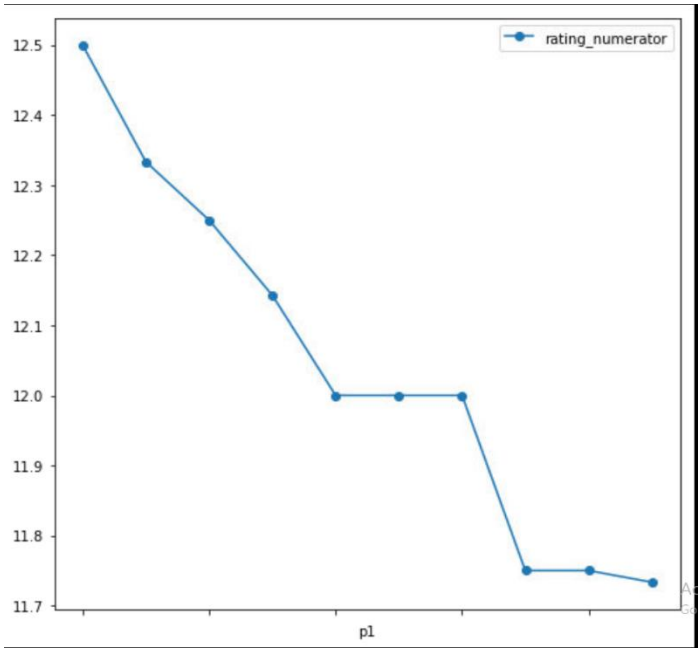
By Mohit Patil

Introduction

Analyzing and Visualizing the dataset is the most significant process in terms of getting the insights and solving business problems. The output from this stage forms the base for answering the questions asked by the organization. Though our primary focus for the project was Data Wrangling following is analysis and visualization for some of the common questions that can be asked to the dataset.

Question 1 : Which are the most commonly predicted dog breeds?

The above visualization clearly states that the golden_retriever was the most predicted dog. It is followed by Labrador_retriever and Pembroke.The golden retriever had approximately 140 predictions. The data can be helpful in predicting the accuracy of the algorithm used to predict the dogs.

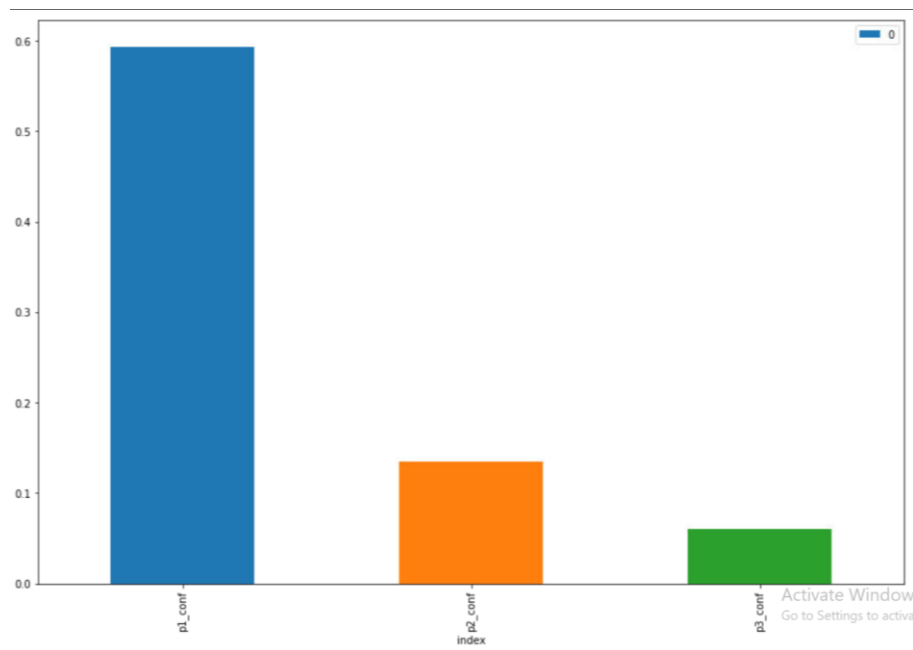# Question 2 : How does the highest average rating for predicted dogs vary?

The highest average rating is quite steep and hence not many dogs get a high rating of

12.5.The highest rated dog is Saluki which is 12.5 and It is followed by the briard and Tibetan mastiff.

We can conclude from the visualization that the dog breed 'Saluki' is the most rated dog.

Question 3 : How does the confidence interval for the first,second and third prediction vary?

As we move to the second and the third prediction the confidence intervals decrease in an irregular fashion. The first prediction has the highest confident interval.The first prediction seems to be the most confident as long as the image prediction algorithm is concerned. The prediction confidence decreases after the first prediction for the further predictions.