# Usability and Difficulties of Introduction to R in the world of Data Science
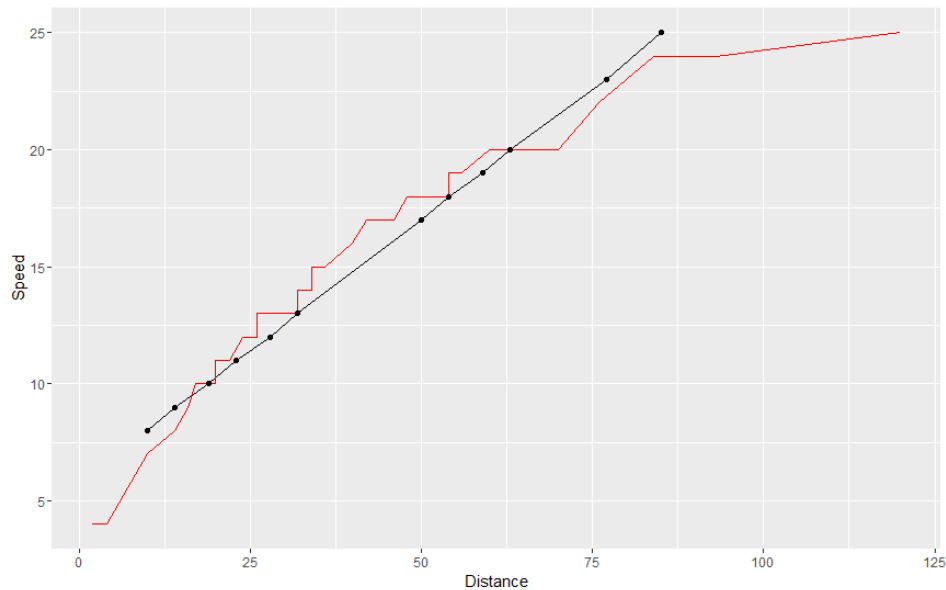
**By**

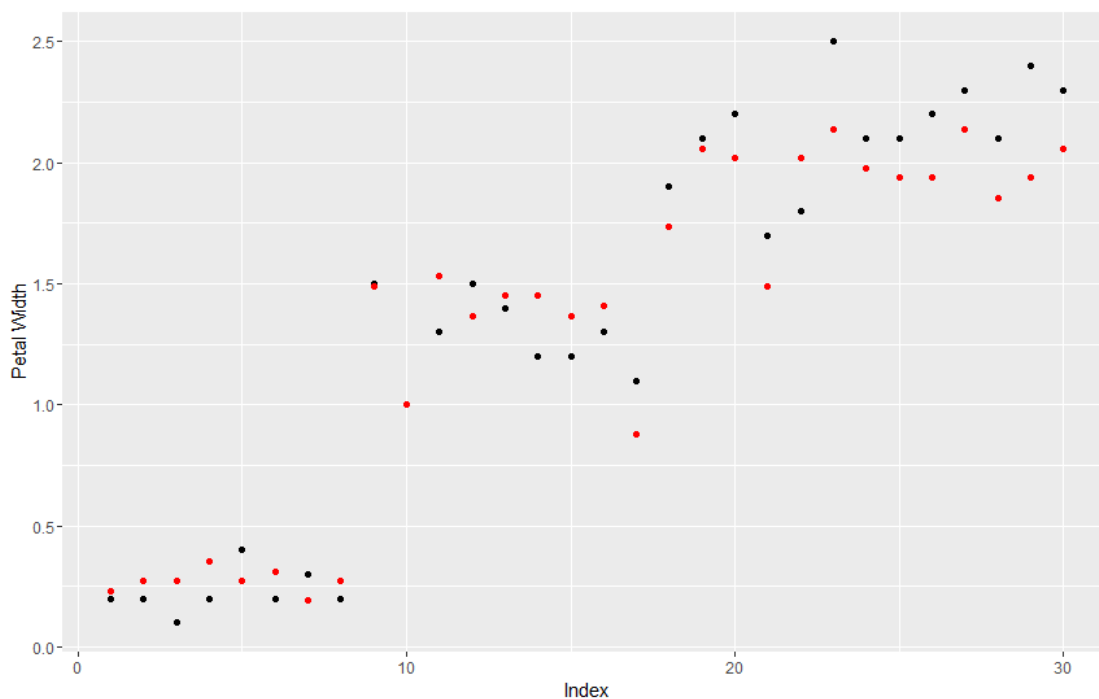**Ruthesan Balasubramaniam**

**ruthesan@gmail.com**

# Distance Prediction of Cars:

The cars distance prediction is done with linear regression algorithm in R with the help of R studio IDE.



The Red line is illustration of the collected speed vs distance profile. The black line is the predicted distance value corresponding to the speed. The data for this prediction is a subset of a red line. This prediction has errors because of the model we used. The speed vs distance is a non-linear data, but we used linear regression to simplify the process.

# Prediction of Petal Width:

The iris data is split into 80/20. The above diagram is a validation where red points indicates the predictions and the black point are the real values. This 20 percent data is predicted using linear regression model based on the petal length.

**Error Overcoming in R:**

Due to prior knowledge in R, I always read the error for unknown object, missing symbols and typo errors. If I could not identify it then I will copy paste the error in google search or stack overflow for the solutions.

**Installation Difficulties:**

I already installed R and R studio in my system several times, so it is not a difficult task. Sometimes few packages made troubles while install but I found solutions in google and installed through CMD.

**Feedback About the Tutorials:**

The tutorials provided in this task is very limited for me. Maybe because of my previous knowledge in this tool.

I will recommitment R for people who using RapidMiner because of its packages and customizations. Especially the plotting options.