Report of R Tutorial

* Was it straightforward to install R and RStudio?

The installation of R and RStudio was very straightforward. I didn’t have any complications with the installation.

* Was the tutorial useful? Would you recommend it to others?

It was useful to start to know the difference in the coding language to do what we learned in python. I recommend this tutorial to other users that have no idea of the existence of R and RStudio. It is very straightforward and easy to understand. However, there are some typos in the tutorial that could lead to an error if you just copy and paste the code.

* Did you learn anything of potential business value from this analysis?

I learned that the distance of the cars is highly affected by the speed and that the width of the petal of a flower is highly affected by the length of the petal.

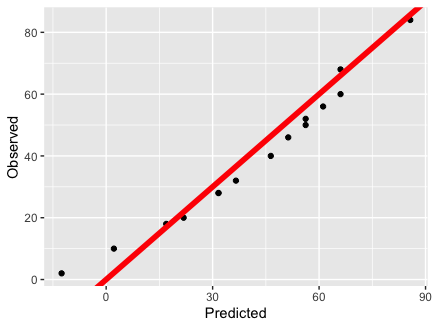
* What are the main lessons you've learned from this experience? What recommendations would you give to other employees who need to get started using R and doing predictive analytics in R instead of WEKA?

The main lessons I learned from this experience are that using RStudio is a lot more user friendly than only using R. Also, it is a lot better to run the code in a script, so that it can be recorded and modified if needed. I also recommend starting a new project and saving it in a known directory with all the data sets are going to be analyzed.

* Can you think of any other things to address that might be of value to management?

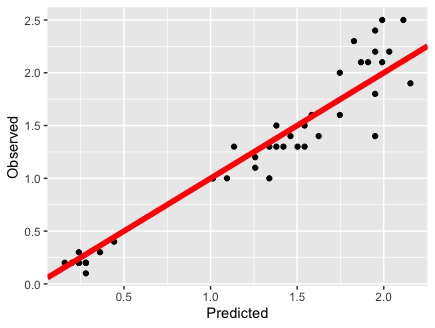
I think it might be important to address the standard packages that someone might need in R to do their analysis.

**Graph 1. The observed distance vs the predicted distance of cars**



The speed of cars is highly significant in predicting the distance. The predictions of shown on **Graph 1** are somewhat close to the observed value.

**Graph 2. The observed petal width vs the predicted petal width of the flowers**



The predictions shown on **Graph 2** are more spread, but there are some predicted values that are very close to the observed ones. Something to take into consideration is the number of values between the two graphs.