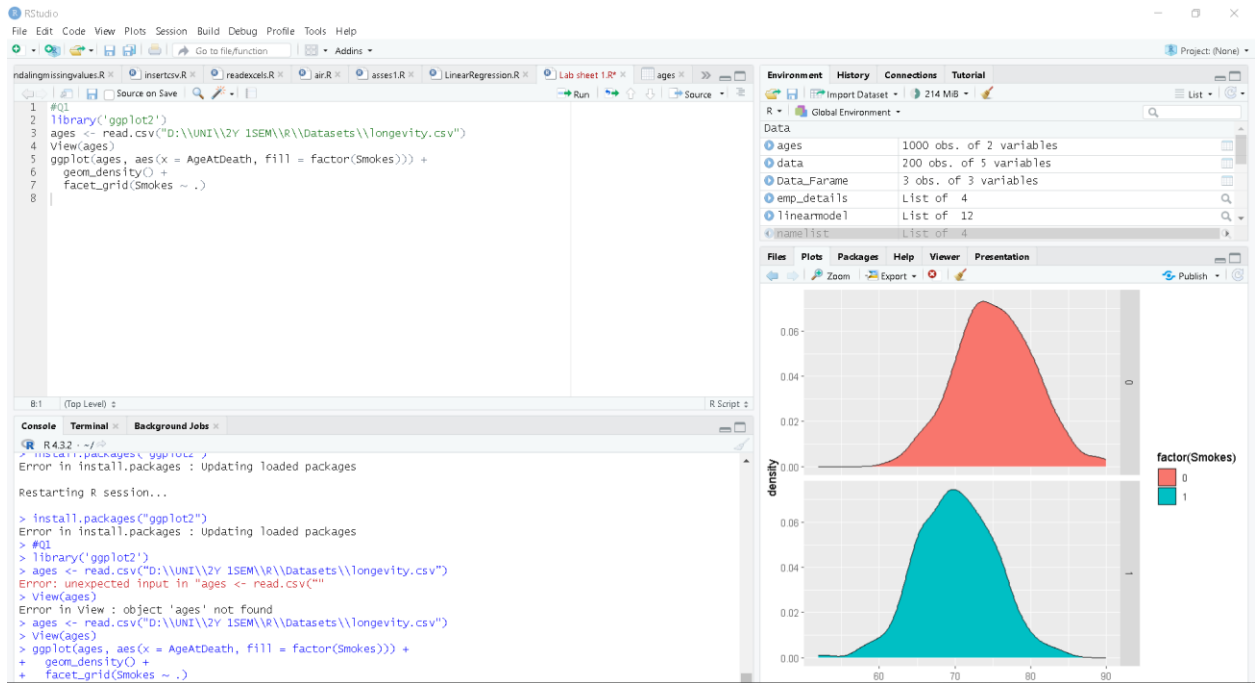


# Lab Sheet 1

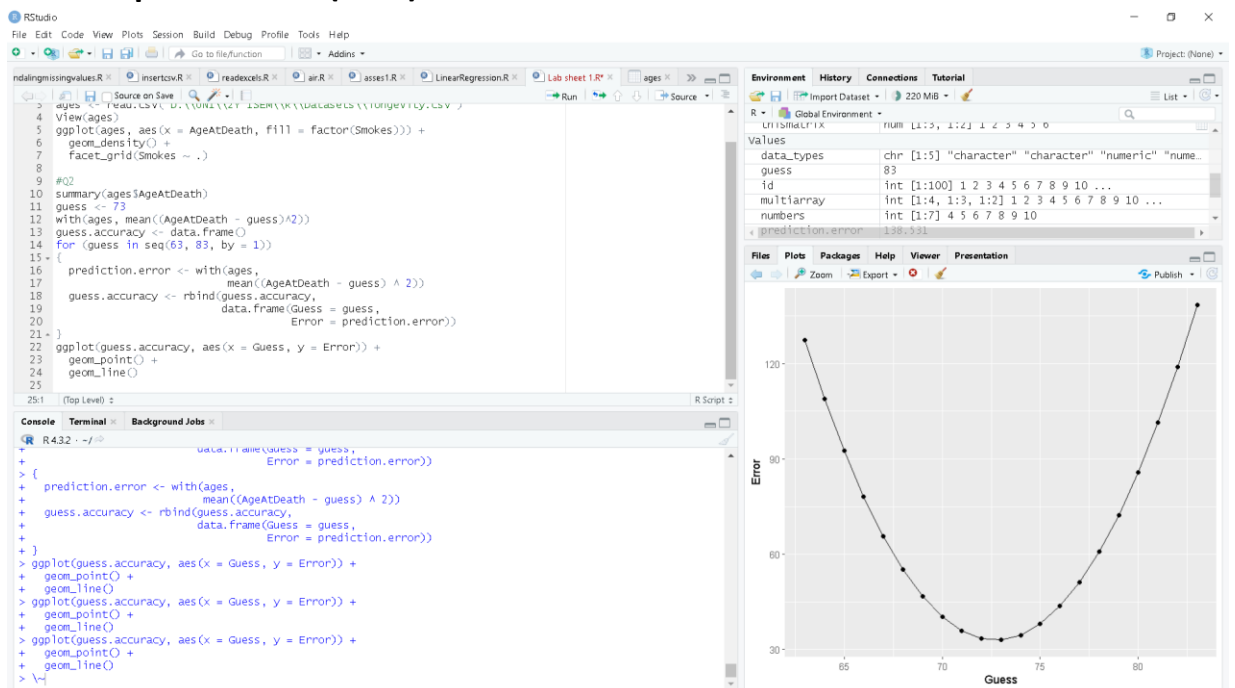
YKA Rathnasiri 27413

## 1. The baseline model



- People who smoke tend to die earlier than those who don't smoke when you look at the graphs.

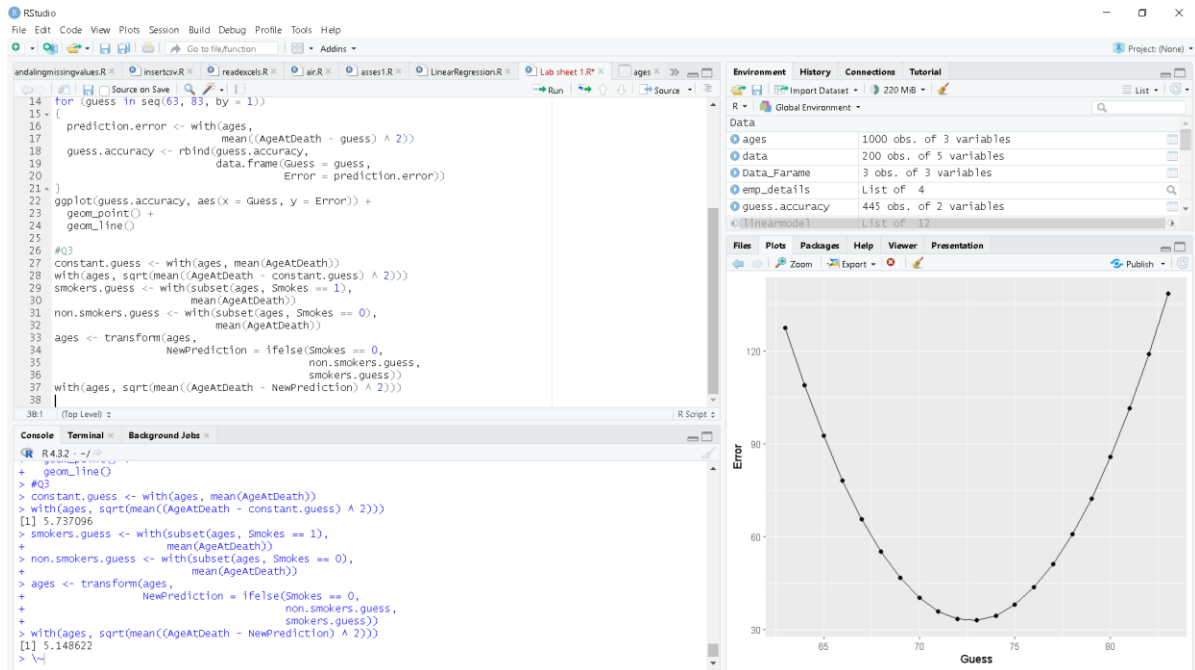
## 2. Mean Squared Error (MSE)



# Lab Sheet 1

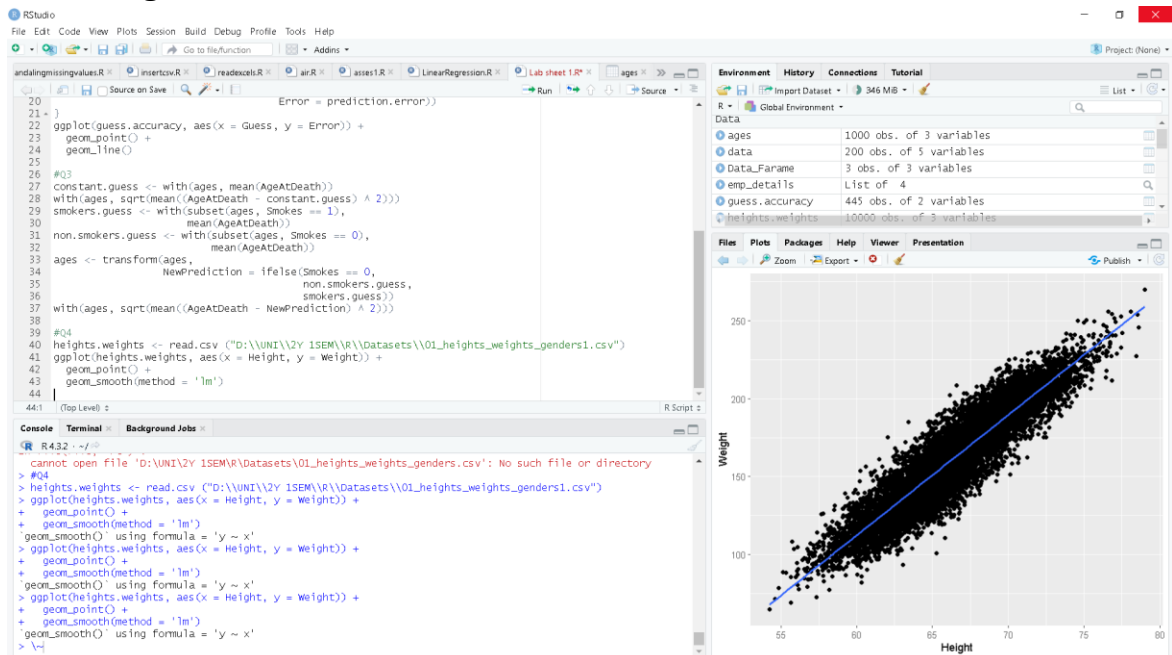
YKA Rathnasiri 27413

## 3. Root Mean Squared Error (RMSE)



- Without considering smoking data, our prediction error (RMSE) is 5.737096. When we include smoking data, the error improves to 5.148622. This means that the prediction is more accurate when we take smoking into account, making it a better estimate than just using age at death alone.

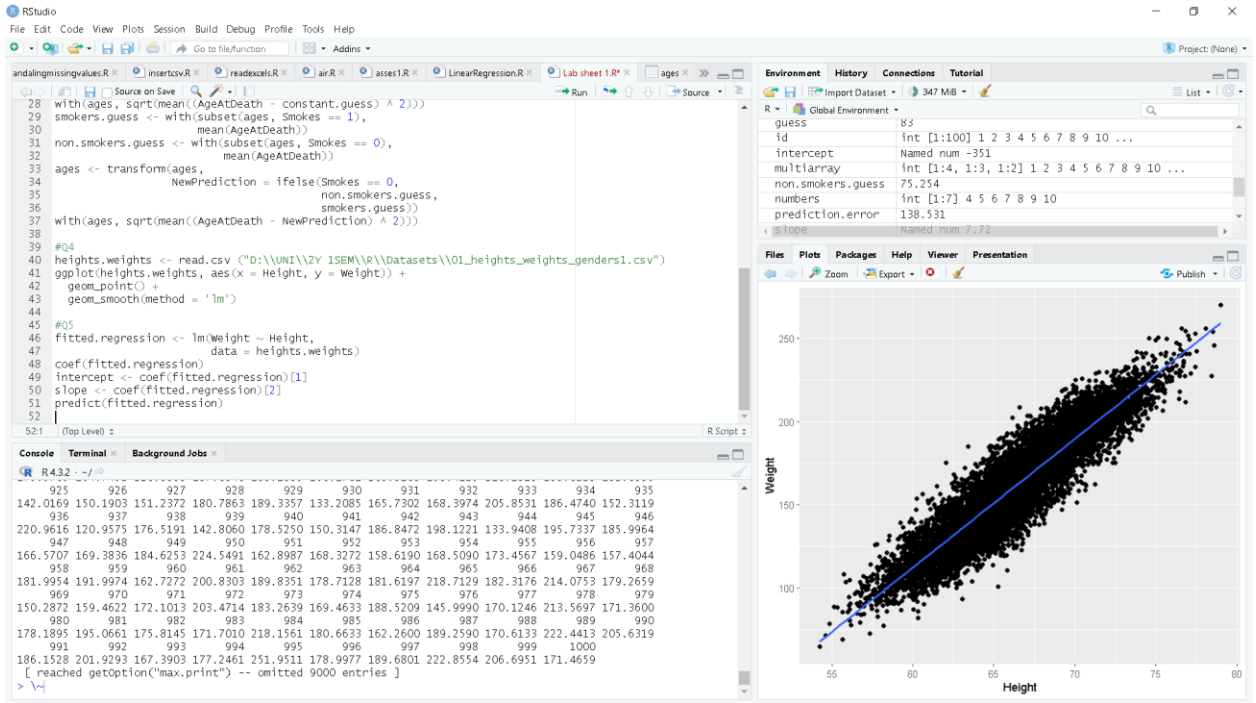
## 4. Linear regression in a nut shell



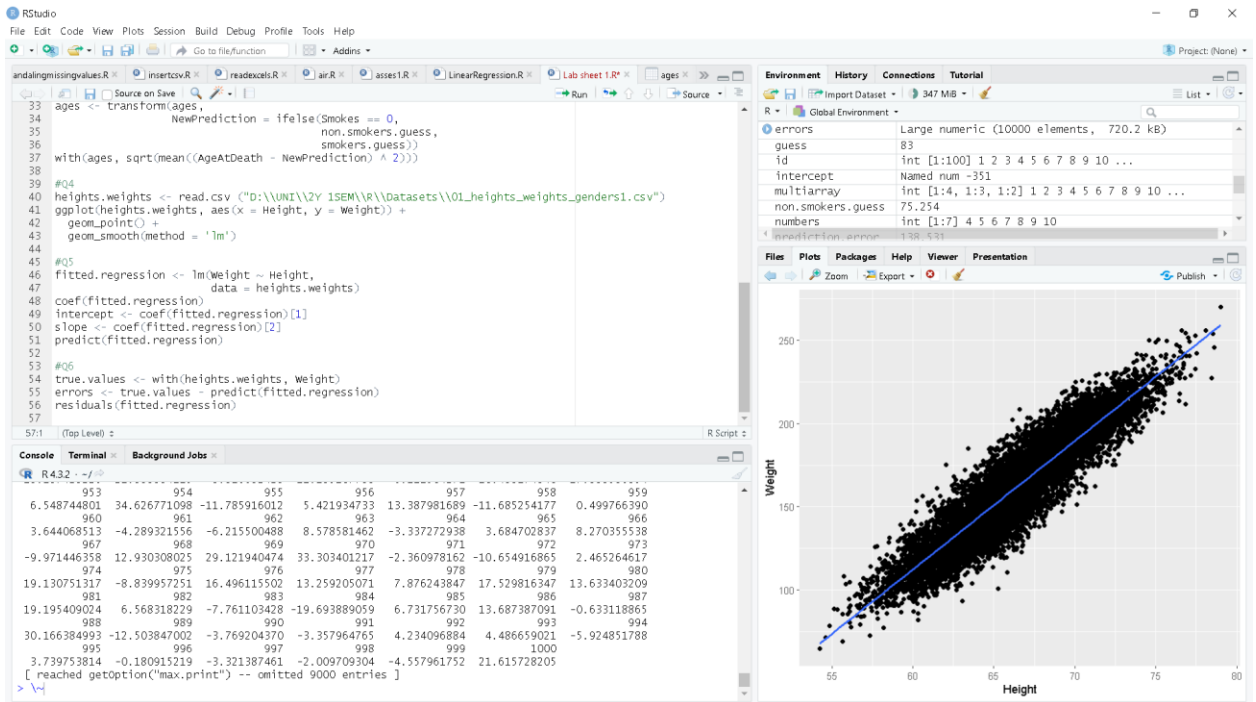
# Lab Sheet 1

YKA Rathnasiri 27413

## 5. Coefficients



## 6. Residuals



# Lab Sheet 1

YKA Rathnasiri 27413

## 7. R-squared

