**Lower difficulty** -> correct =1, wrong =0

1. What is the unit\_name of the most dangerous android?

exemplificative\_vole

1. Calculate the mean volt score per gibbs and chip\_type.



1. Is there a relationship between volt and heat?



1. Build a linear model. The outcome will be kill and the predictors will be speed and agility. Create a table of the parameter estimates from the model
2. Build a linear model. The outcome will be kill and the predictors will be speed, agility, vk, and gibbs. Create a table of the parameter estimates from the model



1. Build a linear model. The outcome will be kill and the predictors will be speed, agility, vk, gibbs, volt, heat, force, and chip\_type. Create a table of the parameter estimates from the model



1. What is the predicted number of kills for an android who runs at a top speed of 35, an agility score of 7, a vk of 400, and has neuromancer as their AI model.



**Higher difficulty** -> correct =3, kind of correct=2 - 1, wrong =0

1. We might be interested in the relationship between chip\_type and agility. We would like a graphical method for examining this. Specifically, you need to create bar plots that displays the count of agility scores corresponding to each chip\_type. Facet by chip\_type.



1. Compare the three linear models that you made. Which one makes better out of sample predictions given the number of effective parameters?



Model 2 makes the best predictions

mod2<-lm(formula= kill~speed + as.numeric(agility)+vk + gibbs, data=d)

1. Create a scatterplot with kill on the y-axis and speed on the x-axis. Also incorporate vk. Facet that by gibbs. Also add a regression line for each faceted plot. Describe what you see.

 

1. We might be particularly interested in the AI models that are being used for the androids, as supposedly, neuromancer, is a particularly vicious one. Specifically, we want an index of the relative strength of evidence from the data about the hypotheses (i.e., the change from prior to posterior odds brought about by the data; i.e., a Bayes factor). Calculate that and summarize what you found.



1. In the first question you needed to find the unit\_name of the most dangerous android. Now, make a function that does that! In other words, make a function that allows you to find the cell that is the name for the corresponding highest value in another column. Find the row index of the maximum value in the specified column. Make sure that you include all function syntax and a test to so that we can see that it works.

find\_highest\_value\_name <- function(df, value\_col, name\_col) {

 max\_index <- which.max(df[[value\_col]])

 # Return the corresponding value in the name column from the same row

 return(df[[name\_col]][max\_index])

}

result <- find\_highest\_value\_name(d,value\_col = "kill", name\_col="unit\_name")

print(result)