# Data Wrangling #2 – Titanic

# 0. Load data in Studio

titanic <- read.csv("titanic\_original.csv")

titanic <- tbl\_df(titanic)

#1. Update port of embarkation

titanic$embarked[which(titanic$embarked == "")] <- "S"

# 2. Update missing values of Age column

titanic$age[is.na(titanic$age)] <- mean(titanic$age, na.rm = TRUE)

# 3. Fill missing values of life boat with NA

titanic$boat[titanic$boat == ""] <- NA

# 4. Created new column based on cabin entry. 1 means yes and 0 means no

titanic$has\_cabin\_number <- ifelse(titanic$cabin == "",0,1)

#. Save results in titanic\_clean.csv file

write.csv(titanic,"titanic\_clean.csv")