install.packages("stringr")

#ADDING LIBRARIES AND DATASET

library(tidyverse)

library(ggplot2)

library(janitor)

library(dplyr)

library(lubridate)

library(stringr)

setwd('E:/DA\_Course/CaseStudy\_Bellabeat/Dataset/Fitabase Data 4.12.16-5.12.16')

daily\_sleep <- read\_csv("E:/DA\_Course/CaseStudy\_Bellabeat/Dataset/Fitabase Data 4.12.16-5.12.16/sleepDay\_merged.csv")

#CLEANING DATA

daily\_sleep<-clean\_names(daily\_sleep)

daily\_sleep<-daily\_sleep %>% separate(sleep\_day,into= c("date","time","am/pm"), sep=" ")

daily\_sleep$sleep\_date\_time <- paste(daily\_sleep$date,daily\_sleep$time,sep=" ")

daily\_sleep <-daily\_sleep %>%

mutate(id=as.character(id),

date=as.Date(date,format="%m/%d/%Y"),

time=format(time,format="%H:%M:%S"),

sleep\_date\_time=as.POSIXct(sleep\_date\_time,format="%m/%d/%Y %H:%M:%S"))

n\_distinct(daily\_sleep$id) # 24 Unique Users, not 30 as mentioned in the case study.

sum(duplicated(daily\_sleep))

daily\_sleep <- distinct(daily\_sleep)

daily\_sleep <- daily\_sleep[, c(1,2,3,8,4,5,6,7)] #Rearranged columns

is.null(daily\_sleep$id) #Checking for nulls

is.null(daily\_sleep$date)

is.null(daily\_sleep$time)

is.null(daily\_sleep$`am/pm`)

is.null(daily\_sleep$total\_sleep\_records)

is.null(daily\_sleep$total\_minutes\_asleep)

is.null(daily\_sleep$sleep\_date\_time)

#ANALYZING DATA

daily\_sleep$time\_taken\_to\_sleep <- daily\_sleep$total\_time\_in\_bed - daily\_sleep$total\_minutes\_asleep

avgtime<-daily\_sleep %>%

group\_by(id) %>%

summarise(average\_time= mean(time\_taken\_to\_sleep))

avgtime$average\_time <- round(avgtime$average\_time,digits = 2)

max(avgtime$average\_time)

min(avgtime$average\_time)

daily\_sleep %>% filter(id=="1844505072")

avgtime %>%

ggplot(aes(x=id,y=average\_time))+ geom\_col(fill="indianred") + theme\_classic()

avgtime$short\_id<-str\_sub(avgtime$id,start=-4) #IDS are not readable in the charts

avgtime$id <- as.character(avgtime$id)

sum(duplicated(avgtime$short\_id)) #Making sure that the ids are unique

avgtime %>%

ggplot(aes(x=short\_id,y=average\_time))+ geom\_col(fill="indianred") +

theme\_classic() +

labs(title="Average Time Taken To Sleep",x="ID",y="Time (In Minutes)") +

geom\_text(aes(label=average\_time),vjust=-0.2,color="black",size=3)

write\_csv(avgtime,file="E:/DA\_Course/CaseStudy\_Bellabeat/Work File/avg\_time.csv")