da420\_homework4

Jason Grahn

1/26/2019

# download from: http://ie.technion.ac.il/serveng/callcenterdata/index.html,   
# http://ie.technion.ac.il/serveng/callcenterdata/FebruaryTxt.ZIP  
  
#make our data frame  
callcenterdata <- read.table(here::here("assignment4/february.txt"),   
 header = TRUE,  
 sep = "") %>%   
 #filter out unnecessary bits  
 filter(server != "PHANTOM", #ghost calls dont count  
 vru\_time >= 0, #vru time can't be negative  
 date < 990208) %>% #just use the one week  
 #make some new variables  
 mutate(wait\_total = vru\_time + q\_time,   
 long\_wait\_binary = if\_else(wait\_total>120, 1, 0), #if it's a long wait, score it 1  
 date = ymd(19000000 + date), #make crappy date format into full date number, then make it ymd  
 day\_of\_week = factor(wday(date), #use date to make day of week factors  
 levels = c(1:7),  
 labels = c("Sunday","Monday","Tuesday",  
 "Wednesday","Thursday","Friday","Saturday")))   
#let's summarize it  
long\_call\_summary <- callcenterdata %>%   
 group\_by(day\_of\_week) %>%   
 summarize(count = n(),  
 long\_count = sum(long\_wait\_binary),  
 percentage\_long = round(long\_count / count \* 100, 2))   
  
#let's plot it  
long\_call\_summary %>%   
 select(day\_of\_week, percentage\_long) %>%   
 ggplot(aes(x = day\_of\_week, y = percentage\_long, label = percentage\_long)) +  
 geom\_col(aes(fill = desc(percentage\_long))) +  
 geom\_label() +  
 theme\_classic() +   
 theme(legend.position = "none") +  
 labs(title= "Days of the week and % of calls with wait above 120 seconds",  
 subtitle = "Sunday and Monday need staffing attention to bring rate lower")

