


109 lines (83 sloc) 5.03 KB

Raw

Blame



```
1 library(readr)
2 newdata <- read_csv("~/Documents/GitHub/WebMining-OFD/newdata.csv")
3
4 newdata$LogComments <- log(newdata$`Number of Comments`)
5 newdata$LogComments[newdata$LogComments=="-Inf"] <- NA
6 newdata$TimeDif <- newdata$`Expected Delivery Time` - newdata$`Time(min)`
7
8 summary(newdata$LogComments)
9 summary(newdata$TimeDif)
10
11 library(tidyr)
12 DataMoment <- newdata %>% gather(`Typical Traffic Afternoon`, `Typical Traffic Noon`, `Typical Traffic Morning`, key = "Typical Traffic", value = "Cases")
13
14 library(dplyr)
15 DataMoments <- data.frame(newdata[1], newdata[5:8], newdata[11:14], newdata[20:21])
16 Morning <- DataMoments %>% filter(Moment == "Morning")
17 Noon <- DataMoments %>% filter(Moment == "Noon")
18 Afternoon <- DataMoments %>% filter(Moment == "Afternoon")
19
20
21 summary(Morning$Number.of.Comments)
22 summary(Noon$Number.of.Comments)
23 summary(Afternoon$Number.of.Comments)
24
25 summary(Morning$TimeDif)
26 summary(Noon$TimeDif)
27 summary(Afternoon$TimeDif)
```



Cada línea tiene una sintáxis (a veces algunas son largas)
En total son 109, pero trabajaremos las primeras 52.