

Plots

Anisha Dubhashi

11/11/2017

load all data from csv in working directory

```
df <- read_csv("cleandata.csv")

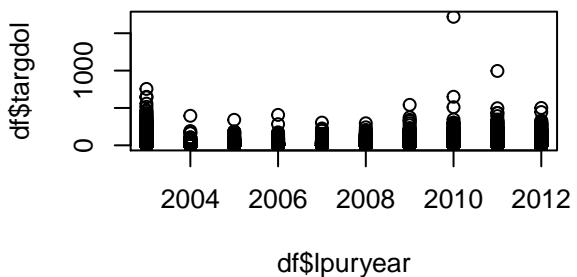
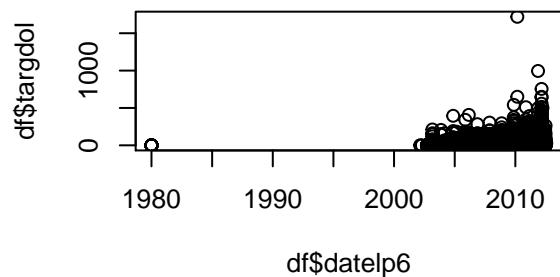
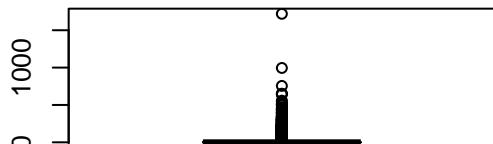
## Warning: Missing column names filled in: 'X1' [1]

## Parsed with column specification:
## cols(
##   .default = col_integer(),
##   targdol = col_double(),
##   datead6 = col_date(format = ""),
##   datelp6 = col_date(format = ""),
##   recentseason = col_character(),
##   consistencycategory = col_character()
## )

## See spec(...) for full column specifications.

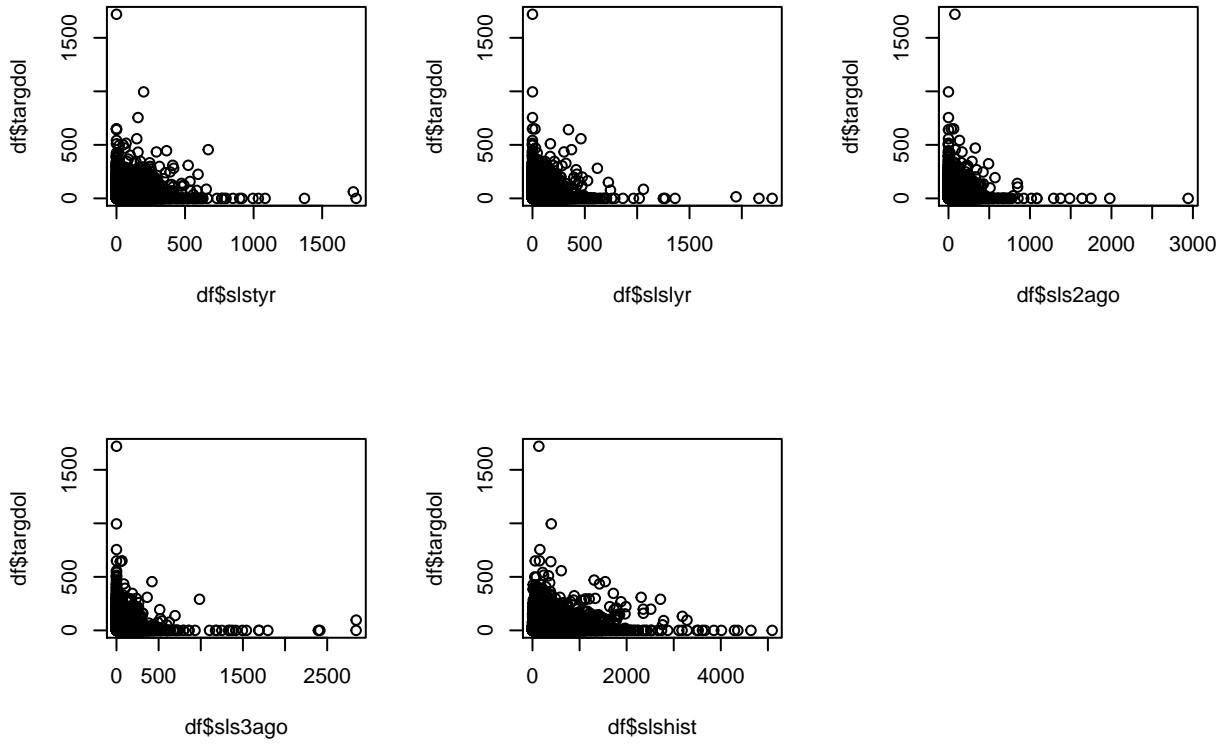
#names(df)

par(mfrow = c(2, 2))
boxplot(df$targdol, outline = TRUE)
plot(df$datead6, df$targdol)
plot(df$datelp6, df$targdol)
plot(df$lpuryear, df$targdol)
```



```
par(mfrow = c(2, 3))
plot(df$slstyr, df$targdol)
plot(df$slslyr, df$targdol)
plot(df$sls2ago, df$targdol)
plot(df$sls3ago, df$targdol)
plot(df$slshist, df$targdol)

par(mfrow = c(2, 3))
```

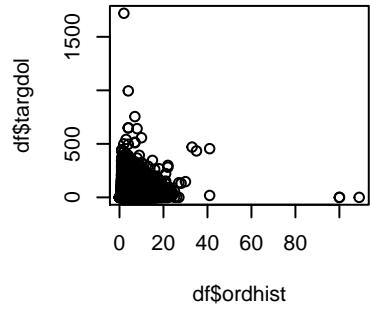
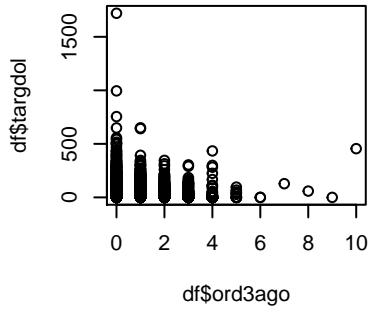
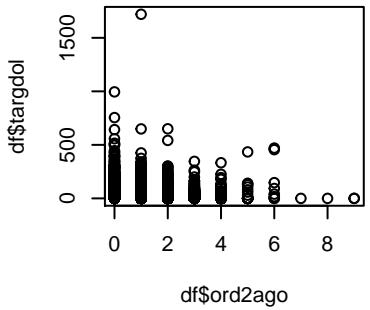
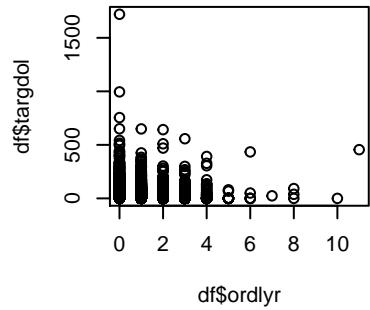
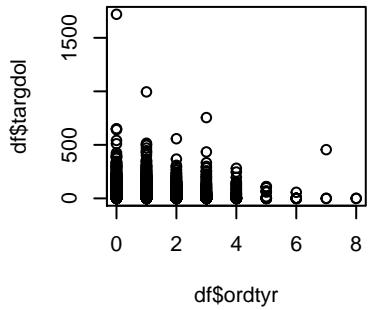


```

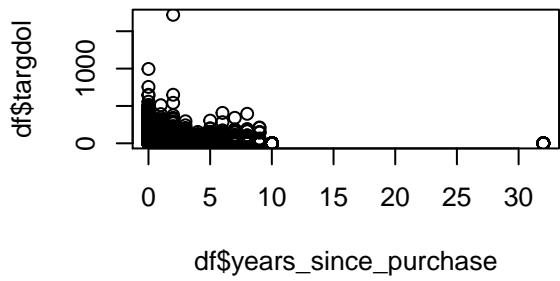
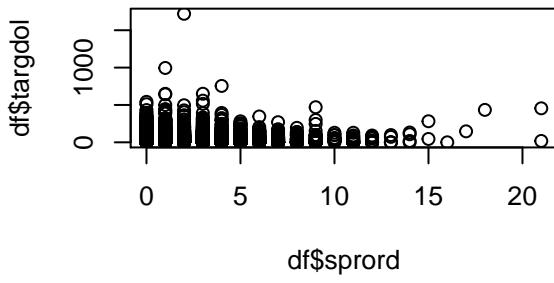
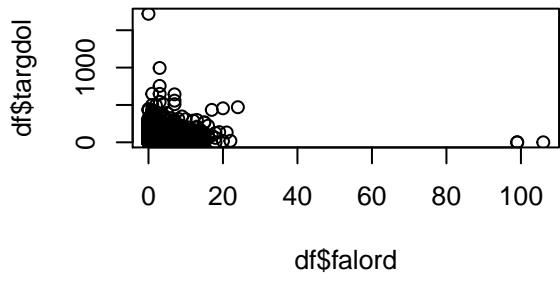
plot(df$ordtyr, df$targdol)
plot(df$ordlyr, df$targdol)
plot(df$ord2ago, df$targdol)
plot(df$ord3ago, df$targdol)
plot(df$ordhist, df$targdol)

par(mfrow = c(2, 2))

```



```
plot(df$falord, df$targdol)
plot(df$sprord, df$targdol)
plot(df$years_since_purchase, df$targdol)
```



```
df[df$targdol == outlier(df$targdol),]
```

```
## # A tibble: 1 x 22
##       X1 targdol    datead6    datelp6 lpuryear slstyr slslyr sls2ago
##   <int>   <dbl>     <date>     <date>     <int>   <int>   <int>
## 1 41618    1720 2008-03-01 2010-03-01     2010      0      0     79
## # ... with 14 more variables: sls3ago <int>, slshist <int>, ordtyr <int>,
## #   ordlyr <int>, ord2ago <int>, ord3ago <int>, ordhist <int>,
## #   falord <int>, sprord <int>, train <int>, recentseason <chr>,
## #   recency_year <int>, years_since_purchase <int>,
## #   consistencycategory <chr>
```

```
df[df$ordhist >= outlier(df$ordhist),]
```

```
## # A tibble: 1 x 22
##       X1 targdol    datead6    datelp6 lpuryear slstyr slslyr sls2ago
##   <int>   <dbl>     <date>     <date>     <int>   <int>   <int>
## 1 79281     0 1993-09-01 2011-03-11     2011      0     17     29
## # ... with 14 more variables: sls3ago <int>, slshist <int>, ordtyr <int>,
## #   ordlyr <int>, ord2ago <int>, ord3ago <int>, ordhist <int>,
## #   falord <int>, sprord <int>, train <int>, recentseason <chr>,
## #   recency_year <int>, years_since_purchase <int>,
## #   consistencycategory <chr>
```