[redacted]

Brian Hogan

[redacted]

[redacted]BRIAN+HOGAN+CORRELATION+HEATMAP+PREDIC

```
#LA Tickets for 2018 - Applied Data Science Final Project
library(gdata,warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'gdata' was built under R version 3.5.2
```

```
## gdata: read.xls support for 'XLS' (Excel 97-2004) files ENABLED.
```

```
##
```

```
## gdata: read.xls support for 'XLSX' (Excel 2007+) files ENABLED.
```

```
df0 <- read.csv("C:/Users/Brian P Hogan Jr/Desktop/Data/streetsweeping-citations-2018-clean.csv")
apply(apply(df0,2,is.na),2,sum)  #confirm no NAs....
```

```
##                     X         ticket.number            issue.date
##                     0                     0                     0
##            issue.year           issue.month             issue.day
##                     0                     0                     0
##         issue.weekday            issue.time        issue.time.bin
##                     0                     0                     0
##             agency.id              meter.id              route.id
##                     0                     0                     0
##         issue.address     issue.address.lat     issue.address.lon
##                     0                     0                     0
##          violation.id        violation.desc    violation.fine.amt
##                     0                     0                     0
##     plate.expire.date     plate.expire.year    plate.expire.month
##                     0                     0                     0
##     plate.expire.flag           plate.state               car.make
##                     0                     0                     0
## car.make.import.flag         car.bodystyle             car.color
##                     0                     0                     0
```

```
# str(df0)
#colnames(df0)
# head(df0,1)
df1 <- data.frame(df0) ##594546 obs. of  27 variables:
#Notes.1 ==> violation.fine all $73; all street cleaning; meter.id=dont use all same meter
#Note.2==> when merge dataframe columns are reordered....
#DATA TRANSFORMATION - add numeric codes for scatter+melt & correlations
    # f.make.numeric.id <-function(v.name)  ##FIX - v.name
```

```
    #     # { df.unique.temp <- data.frame(unique(df1$v.name))
    #     # level.rows <- c(1:nrow(df.unique.temp))  #fill in blanks to add a group size variable
    #     # df.unique.temp <-cbind(df.unique.temp ,level.rows)
    #     # colnames(df.unique.temp) <- c("v.name",paste("v.name",".2"))
    #     # df2 <- merge(x=df1, y=df.unique.temp, by ="v.name", all.x=TRUE)
    #     # remove(df.unique.temp)  }
#NUMBERIC IDs -------------FOR correlation & scatterplots--------------------------------
c.names <- colnames(df1)
c.names.2 <- gsub("\\.","",c.names)  #remove dots for sqldf
c.names.2
```

```
##  [1] "X"                 "ticketnumber"      "issuedate"
##  [4] "issueyear"         "issuemonth"        "issueday"
##  [7] "issueweekday"      "issuetime"         "issuetimebin"
## [10] "agencyid"          "meterid"           "routeid"
## [13] "issueaddress"      "issueaddresslat"   "issueaddresslon"
## [16] "violationid"       "violationdesc"     "violationfineamt"
## [19] "plateexpiredate"   "plateexpireyear"   "plateexpiremonth"
## [22] "plateexpireflag"   "platestate"        "carmake"
## [25] "carmakeimportflag" "carbodystyle"      "carcolor"
```

```
colnames(df1)<-c.names.2
remove(c.names.2)
#route.ID  ==>levels=674
df.temp <- data.frame(unique(df1$routeid))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
colnames(df.temp) <- c("routeid",paste("routeid2"))
df1 <- merge(x=df1, y=df.temp, by ="routeid", all.x=FALSE)
#plate.state  ==> 71levels=73
df.temp <- data.frame(unique(df1$platestate))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
colnames(df.temp) <- c("platestate",paste("platestate2"))
df1 <- merge(x=df1, y=df.temp, by ="platestate", all.x=FALSE)
#car.make ==> levels=62
df.temp <- data.frame(unique(df1$carmake))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
colnames(df.temp) <- c("carmake",paste("carmake2"))
df.carid.master <-df.temp
df1 <- merge(x=df1, y=df.temp, by ="carmake", all.x=FALSE)
write.csv(df.carid.master, "df.carid.master.csv") #spped up normalization
#car.bodystyle  ==> levels=12
df.temp <- data.frame(unique(df1$carbodystyle))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
colnames(df.temp) <- c("carbodystyle",paste("carbodystyle2"))
df1 <- merge(x=df1, y=df.temp, by ="carbodystyle", all.x=TRUE)
#car.color  ==>levels=16
df.temp <- data.frame(unique(df1$carcolor))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
```

```
colnames(df.temp) <- c("carcolor",paste("carcolor2"))
df1 <- merge(x=df1, y=df.temp, by ="carcolor", all.x=TRUE)
#issue.address ==>levels=292164
df.temp <- data.frame(unique(df1$issueaddress))
level.rows <- c(1:nrow(df.temp))  #fill in blanks
df.temp <-cbind(df.temp ,level.rows) #lookup table
colnames(df.temp) <- c("issueaddress",paste("issueaddress2"))
df1 <- merge(x=df1, y=df.temp, by ="issueaddress", all.x=TRUE)
head(df1,3)
```

```
##         issueaddress carcolor carbodystyle carmake platestate routeid
## 1    ! % CULVER BLVD       BN           PA    JEEP         VT   00136
## 2      !& TOPSAIL ST       BK           PA    MNNI         CA   00136
## 3 !1700 NEBRASKA AVE       SL           PA    FORD         CA   00183
##        X ticketnumber           issuedate issueyear issuemonth issueday
## 1 245625   4332818545 2018-05-23T00:00:00      2018          5       23
## 2 591550   4344513143 2018-12-31T00:00:00      2018         12       31
## 3 390209   4337741634 2018-08-17T00:00:00      2018          8       17
##   issueweekday issuetime issuetimebin agencyid meterid issueaddresslat
## 1            4  12:24:00            4       51       0       -137.9131
## 2            2  10:31:00            3       51       0       -137.9131
## 3            6  12:19:00            4       51       0       -137.9131
##   issueaddresslon violationid        violationdesc violationfineamt
## 1        27.51751     80.69BS NO PARK/STREET CLEAN               73
## 2        27.51751     80.69BS NO PARK/STREET CLEAN               73
## 3        27.51751     80.69BS NO PARK/STREET CLEAN               73
##   plateexpiredate plateexpireyear plateexpiremonth plateexpireflag
## 1          201909            2019                9               1
## 2          201905            2019                5               1
## 3          201808            2018                8               0
##   carmakeimportflag routeid2 platestate2 carmake2 carbodystyle2 carcolor2
## 1                 0      117          58        6             1        10
## 2                 1      117           1       24             1         5
## 3                 0       97           1        3             1         9
##   issueaddress2
## 1        117357
## 2        112906
## 3        100872
```

```
#------CORRELATION----ELLIPSIS-----------------------------------
library(ellipse, warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'ellipse' was built under R version 3.5.3
```

```
library(RColorBrewer,warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'RColorBrewer' was built under R version 3.5.2
```

```
my_colors <- brewer.pal(11, "Spectral")  #build color panel
my_colors = colorRampPalette(my_colors)(100)
#x <- data.frame(colnames(df1))   #use colnames to get IDs
```

```
df.cor <-data.frame(df1[,c(11,12,13,15,16,18,19,26:33)])
cor.df1 <- cor(df.cor)   #build corrleation of data
ord <- order(df.cor[1,])  # Order the correlation matrix
data_ord = cor.df1[ord, ord]
plotcorr(data_ord , col = my_colors[data_ord * 50 + 30] , mar = c(1, 1, 1, 1))
#write.csv(round(cor.df1,2), "project.csv"): found coordinate issues had to fix
#-----MELT HISTROGRAM-----------------------------
library(sqldf, warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'sqldf' was built under R version 3.5.2
```

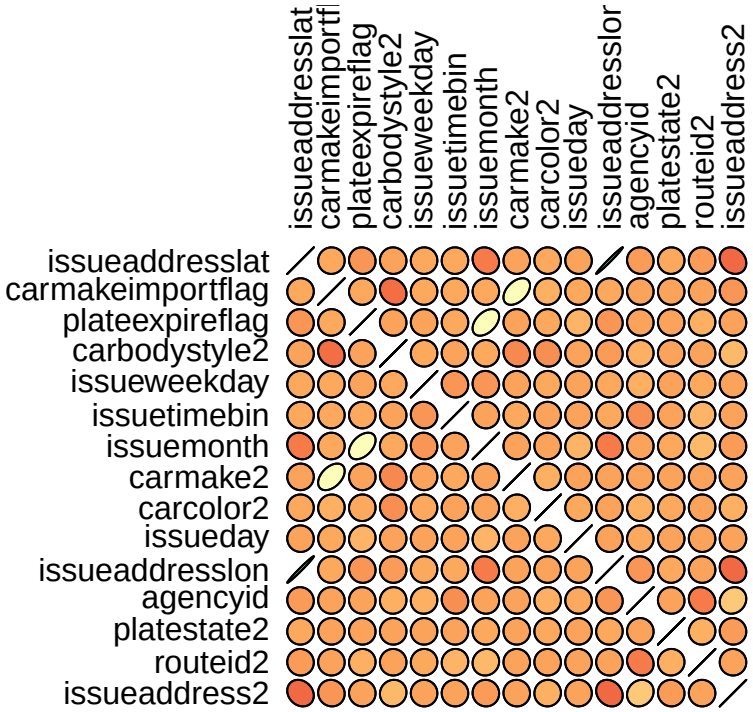
```
## Warning: package 'gsubfn' was built under R version 3.5.2
```

```
## Warning: package 'proto' was built under R version 3.5.2
```

```
## Warning: package 'RSQLite' was built under R version 3.5.2
```

```
library(ggplot2, warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.3
```



```
library(reshape2, warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'reshape2' was built under R version 3.5.3
```

```
library(ggmap, warn.conflicts = FALSE, quietly = TRUE)
```

```
## Warning: package 'ggmap' was built under R version 3.5.2
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.
```

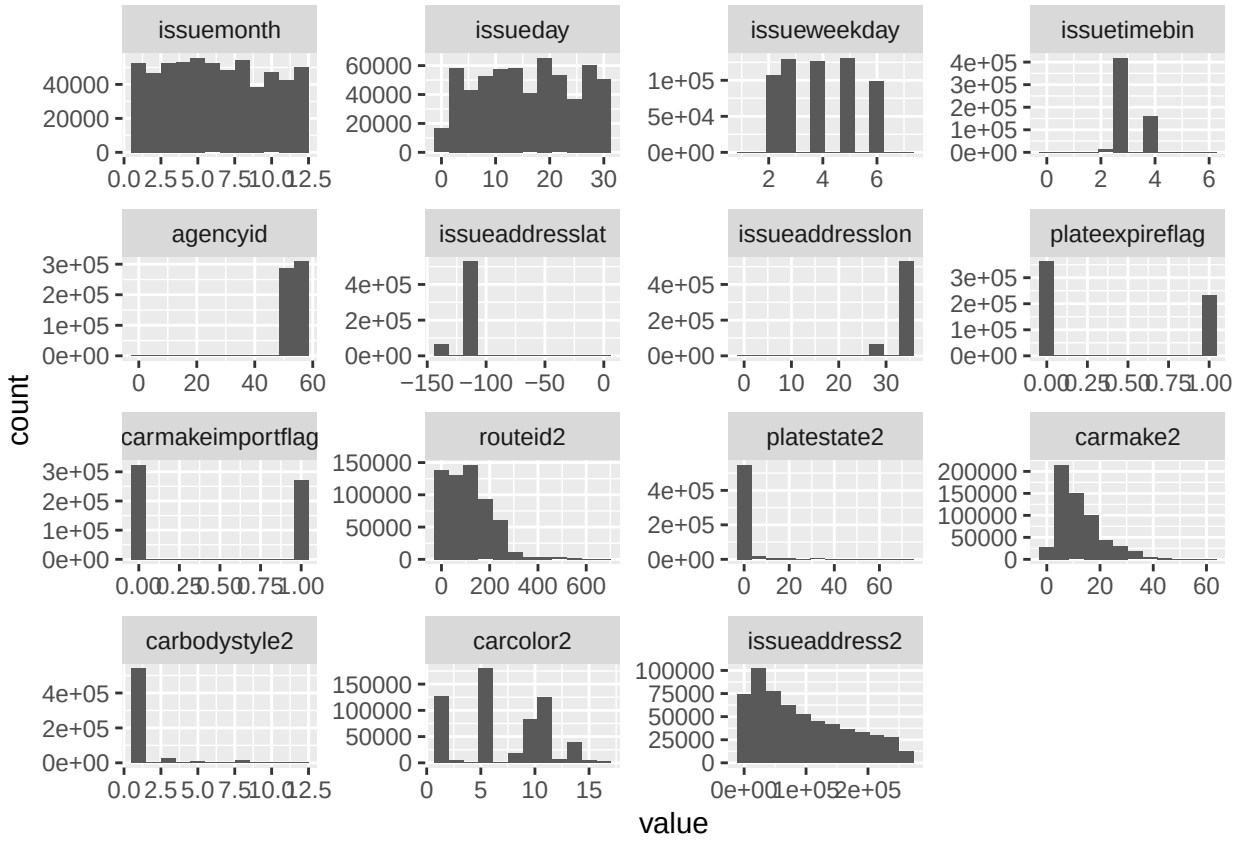
```
## Please cite ggmap if you use it! See citation("ggmap") for details.
```

```
us <-map_data("state")
```

```
## Warning: package 'maps' was built under R version 3.5.2
```

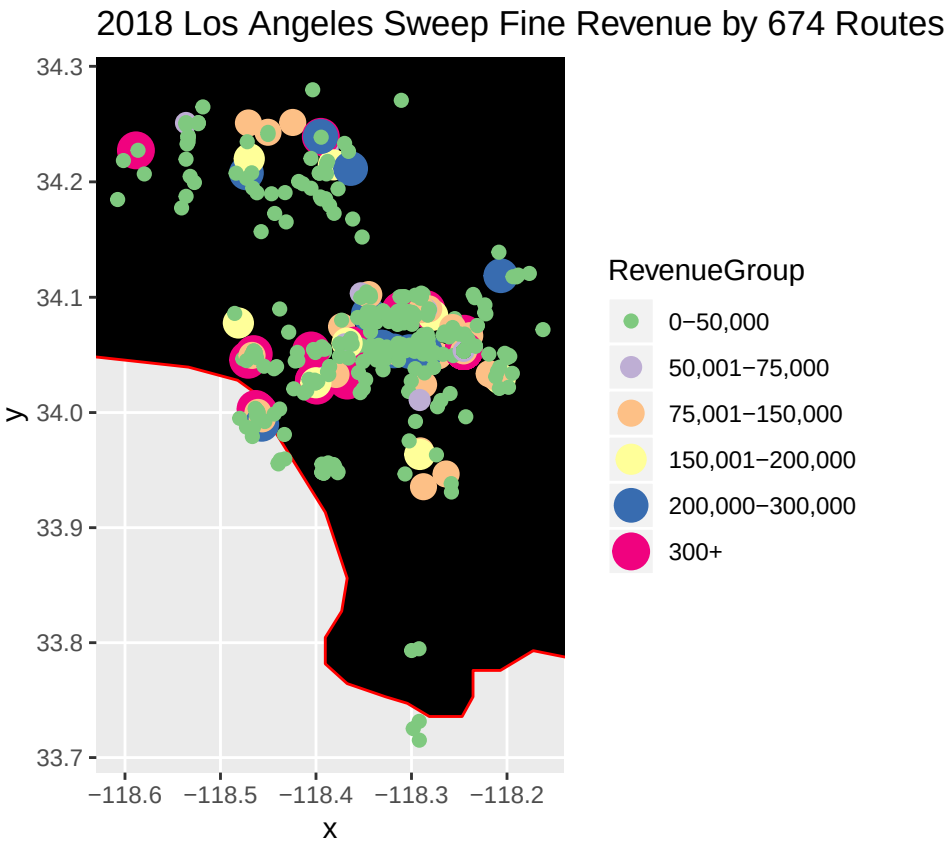
```
suppressWarnings(require(RColorBrewer)) #install.packages("RColorBrewer")
ggplot(data = melt(df1[,c(11,12,13,15,16,18,19,26:33)]), mapping = aes(x = value)) +
  geom_histogram(bins=12)+ facet_wrap(~variable, scales = "free")
```

```
## No id variables; using all as measure variables
```



```
#------------LA REVENUE MAP-----------------------------
state <-map_data("state")
route.df1 <- sqldf('select routeid2, issueaddresslat as long,issueaddresslon as lat,
                   SUM(violationfineamt) as fine from df1 group by routeid2')
RevenueGroup <- c(1:nrow(route.df1))  #fill in blanks to add a group size variable
state <- c(1:nrow(route.df1))     #write.csv(route.df1, "Route_Fine.csv")
longnew <-c(1:nrow(route.df1))
latnew <-c(1:nrow(route.df1))
route.df1 <-cbind(route.df1,RevenueGroup)   #Gadd state full name back for mapping please
route.df1 <-cbind(route.df1,state)
route.df1 <-cbind(route.df1,longnew)
route.df1 <-cbind(route.df1,latnew)
remove(RevenueGroup,state,latnew, longnew)  #----ADD FIXING OF LAT/LON CONVERSION HERE
#str(route.df1)
# there are both bad and blank coordinates; I am just grouping for revenue diagram
#old values = -137.9131, 27.51751
#new values = -118.24532, 34.05349
n <-1
while (n <= nrow(route.df1))
  {  route.df1[n,7]= round(route.df1[n,2],4)
     route.df1[n,8]= round(route.df1[n,3],5)
     #route.df1[n,2]= 0
     #route.df1[n,3]= 0
     n <- n+1                 #fixing some bad lat long positions - moving over from ocean
  }                               #REVENUE DOT PLOT FIXING LAT/LONG
n <- 1
while (n <= nrow(route.df1))
  { #if (route.df1[n,7]== -137.9131) { baddates <- baddates+1 }
    if (route.df1[n,7]== -137.9131)  {route.df1[n,2]= -118.24532  }
    if (route.df1[n,8]== 27.51751)   {route.df1[n,3]= 34.05349  }
    if (route.df1[n,7]== 0)          {route.df1[n,2]= -118.24532  }
    if (route.df1[n,8]== 0)         {route.df1[n,3]= 34.05349  }
    n <- n+1  }
    #colnames(route.df1) <- c("routeid2","long","lat","fine","grpsize","state","fix")
#head(route.df1)
route.df1$state <- "california"   #expand to get whole view on R windows
route.df1$RevenueGroup <- cut(route.df1$fine,           #make buckets for mapping
                    breaks = c(-Inf, 50000, 75000, 150000, 200000, 300000, Inf),
  labels = c("0-50,000","50,001-75,000","75,001-150,000","150,001-200,000","200,000-300,000","300+"),
                          right = FALSE)
mycolors <- brewer.pal(6, "Accent")   #head(df4,1)
names(mycolors) <-levels(route.df1$grpsize)  #getting the color names
ggplot(route.df1, aes(map_id=state)) +
  expand_limits(x=route.df1$long, y=route.df1$lat) + coord_map()+
  geom_map(map = us, fill="black", color="red" ) + geom_point(data=route.df1,
                    aes(x=long,y=lat, color=RevenueGroup, size=RevenueGroup ))+
  #stat_density2d(data=route.df1,aes(x=long,y=lat), geom="density_2d")+
  scale_colour_manual( values=mycolors) +  #name="Color",
  ggtitle("2018 Los Angeles Sweep Fine Revenue by 674 Routes")
```

```
## Warning: Using size for a discrete variable is not advised.
```




```
      #nice revenue map - had some coordinate issues fixed
#=============================================================================
#-------------->>>>> HEATMAP WORK  <<<<=======================
#GET A GRID OF THE CARS AS NEED TO MAKE NUMERIC THEN NORMALIZE 0-1
#
#       head(df1$carmake)
#       hist(df1$carmake2)
#       nrow(table(df1$violation.fine.amt))
#       hist(df1$violation.fine.amt)
#       df.temp <- data.frame(tapply(df1$carmake2, df1$carmake,min))
#       unique(df1$carmake)
#       tapply(df1$carmake2, df1$carmake,min)

car.df1 <- data.frame(sqldf('select carmake,SUM(carmake2) as CarQty
                            from df1 group by carmake') )
#explore & performing grouping/normalization in Excel
#would have done in R but running out of time on delivery to team! - doing rest in R
      #write.csv(df.carid.master, "df.carid.master.csv") #spped up normalization
car.df2 <- read.xls("C:/Users/Brian P Hogan Jr/Desktop/IST687+Final+Project+Car+Group+Normalize.xlsx",
                perl="C:/strawberry/perl/bin/perl.exe", verbose=TRUE)
```

```
## Using perl at C:\STRAWB~1\perl\bin\perl.exe
## Using perl at C:\STRAWB~1\perl\bin\perl.exe
##
## Converting xls file
##     "C:/Users/Brian P Hogan Jr/Desktop/IST687+Final+Project+Car+Group+Normalize.xlsx"
```

```
## to csv  file
##     "C:\Users\BRIANP~1\AppData\Local\Temp\RtmpKiQ7RZ\filec3466a2c10.csv"
## ...
##
## Executing ' "C:\STRAWB~1\perl\bin\perl.exe" "C:/Users/Brian P Hogan Jr/Documents/R/win-library/3.5/g
##
## 0
##
## Done.
##
## Reading csv file  "C:\Users\BRIANP~1\AppData\Local\Temp\RtmpKiQ7RZ\filec3466a2c10.csv" ...
## Done.
```

```
car.df2
```

```
##    carmake2 carmake carqty cargroup cargroupnormalize
## 1         9    CHEV 361926        1            1.0000
## 2        28    CHRY 195104        1            0.5378
## 3         3    FORD 154884        1            0.4264
## 4        50     GEO   8250        1            0.0202
## 5        41    SMRT  18819        1            0.0494
## 6        20    STRN  48960        1            0.1329
## 7        61    TESL    976        1            0.0000
## 8         4    VOLV  22876        1            0.0607
## 9        27    BUIC  74925        2            0.5516
## 10       31    CADI 128743        2            1.0000
## 11        6    JEEP  92004        2            0.6939
## 12       25    LINC  57375        2            0.4054
## 13       37    OLDS  21312        2            0.1049
## 14       12    OTHR  80208        2            0.5956
## 15       48    PLYM  13632        2            0.0409
## 16       30    PONT  74430        2            0.5475
## 17       57    TRIU   8721        2            0.0000
## 18       46    TSMR  52348        2            0.3635
## 19        1    AUDI  10731        3            0.0102
## 20       33    FIAT  82203        3            0.0838
## 21       13    HOND 972023        3            1.0000
## 22       22    HYUN 460790        3            0.4736
## 23       32    INFI 242208        3            0.2486
## 24       21     KIA 336651        3            0.3458
## 25       58    LAMO    812        3            0.0000
## 26       26    LEXS 427804        3            0.4396
## 27       16    MAZD 203664        3            0.2089
## 28       34    MERC  53414        3            0.0542
## 29        7    MERZ 140364        3            0.1437
## 30       11    MITS  54428        3            0.0552
## 31       19    NISS 870979        3            0.8960
## 32       49    SCIO   4410        3            0.0037
## 33       15    SUBA 140475        3            0.1438
## 34       44    SUZI  40172        3            0.0405
## 35        8    TOYT 917096        3            0.9434
## 36       14    VOLK 307762        3            0.3160
## 37       18    ACUR 125658        4            0.5345
## 38       55    ALFA  11495        4            0.0431
```

```
## 39       10     BMW 233790        4            1.0000
## 40       23    DATS   1909        4            0.0018
## 41       35    HUMM   6510        4            0.0216
## 42       36    JAGU  55944        4            0.2344
## 43       42    KAWK  11928        4            0.0449
## 44       17    LNDR  63869        4            0.2685
## 45       24    MNNI 114360        4            0.4859
## 46       38    PORS  71174        4            0.3000
## 47       62    RROV   1488        4            0.0000
## 48       43    SAAB  33368        4            0.1372
## 49       52    YAMA  15964        4            0.0623
## 50       54    ASTO   2160        5            0.1223
## 51       51    BENT   5100        5            0.5218
## 52       39    BENZ   8619        5            1.0000
## 53       60    BUGA   1260        5            0.0000
## 54       56    FERR   2072        5            0.1103
## 55       59     ROL   2478        5            0.1655
## 56        2    DODG  33588        6            0.9218
## 57       45    FRHT   9990        6            0.2563
## 58        5     GMC  36360        6            1.0000
## 59       47    INTL   6768        6            0.1655
## 60       40     ISU  29200        6            0.7981
## 61       53    MACK    901        6            0.0000
## 62       29    WINN   2987        6            0.0588
```

```
df.temp <- data.frame(car.df2)
#head(df.temp)
              # carqty <- c(1:nrow(df1))
              # cargroupnormalize <- c(1:nrow(df1))
              # df2.heat <-cbind(df1 ,carqty,cargroup,cargroupnormalize)
              # df2.heat$carqty <- -199
              # df2.heat$cargroup <- -199
              # df2.heat$cargroupnormalize <- -199
#============================================================
#FINDING===> WAS USING 2 DIFFERENT TABLE IDS PULLED AT DIFFERETN TIME PIONTS!
#===> R resorts data on subsequent data merges
#============================================================
#fixed the merge - was using 2 different tables w differing assignment values..
df2.heat <-merge(x=df1, y=df.temp, by=c("carmake2","carmake"), all.x = TRUE)
#head(df2.heat,10)
remove(car.df1,car.df2)
cargroupname <- c("Dom.EconY","Dom.EconN","Intl.EconY","Intl.EconN","Fancy","Trucks")
cargroup<- c(1,2,3,4,5,6)
df.car.group <- data.frame(cargroup, cargroupname)
df.car.group
```

```
##   cargroup cargroupname
## 1        1    Dom.EconY
## 2        2    Dom.EconN
## 3        3   Intl.EconY
## 4        4   Intl.EconN
## 5        5        Fancy
## 6        6       Trucks
```

```
df2.heat <-merge(x=df2.heat, y=df.car.group, by="cargroup", all.x = TRUE)
head(df2.heat,1)
```

```
##   cargroup carmake2 carmake      issueaddress carcolor carbodystyle
## 1        1       20    STRN 531 KENMORE AVE S       GY           PA
##   platestate routeid      X ticketnumber           issuedate issueyear
## 1         CA   00400 120557   4327894885 2018-03-14T00:00:00      2018
##   issuemonth issueday issueweekday issuetime issuetimebin agencyid meterid
## 1          3       14            4  08:58:00            3       54       0
##   issueaddresslat issueaddresslon violationid        violationdesc
## 1       -118.2966        34.06485     80.69BS NO PARK/STREET CLEAN
##   violationfineamt plateexpiredate plateexpireyear plateexpiremonth
## 1               73          201807            2018                7
##   plateexpireflag carmakeimportflag routeid2 platestate2 carbodystyle2
## 1               0                 0       48           1             1
##   carcolor2 issueaddress2 carqty cargroupnormalize cargroupname
## 1        11         40558  48960            0.1329    Dom.EconY
```

```
#=================>>>>>remove(df.car.group)
#Heatmap
#car type grouping - 0-1 scaling was done in Excel to spped up as had issue mergeing
#issue wasn't mergeing but was 2 different data tables pulled at different time points
colnames(df2.heat)
```

```
##  [1] "cargroup"          "carmake2"          "carmake"
##  [4] "issueaddress"      "carcolor"          "carbodystyle"
##  [7] "platestate"        "routeid"           "X"
## [10] "ticketnumber"      "issuedate"         "issueyear"
## [13] "issuemonth"        "issueday"          "issueweekday"
## [16] "issuetime"         "issuetimebin"      "agencyid"
## [19] "meterid"           "issueaddresslat"   "issueaddresslon"
## [22] "violationid"       "violationdesc"     "violationfineamt"
## [25] "plateexpiredate"   "plateexpireyear"   "plateexpiremonth"
## [28] "plateexpireflag"   "carmakeimportflag" "routeid2"
## [31] "platestate2"       "carbodystyle2"     "carcolor2"
## [34] "issueaddress2"     "carqty"            "cargroupnormalize"
## [37] "cargroupname"
```

```
# table(df2.heat$issuetimebin)
# hist(df2.heat$cargroupnormalize)

#Step 4 Heatmap---using normalization approach between 0-1 as didnt' want negative scale
df3.heat <- data.frame(df2.heat[,c(9,12,13,15,17,1,37,36)])
head(df3.heat,1)
```

```
##        X issueyear issuemonth issueweekday issuetimebin cargroup
## 1 120557      2018          3            4            3        1
##   cargroupname cargroupnormalize
## 1    Dom.EconY            0.1329
```

```
max(df3.heat$issuetimebin)
```

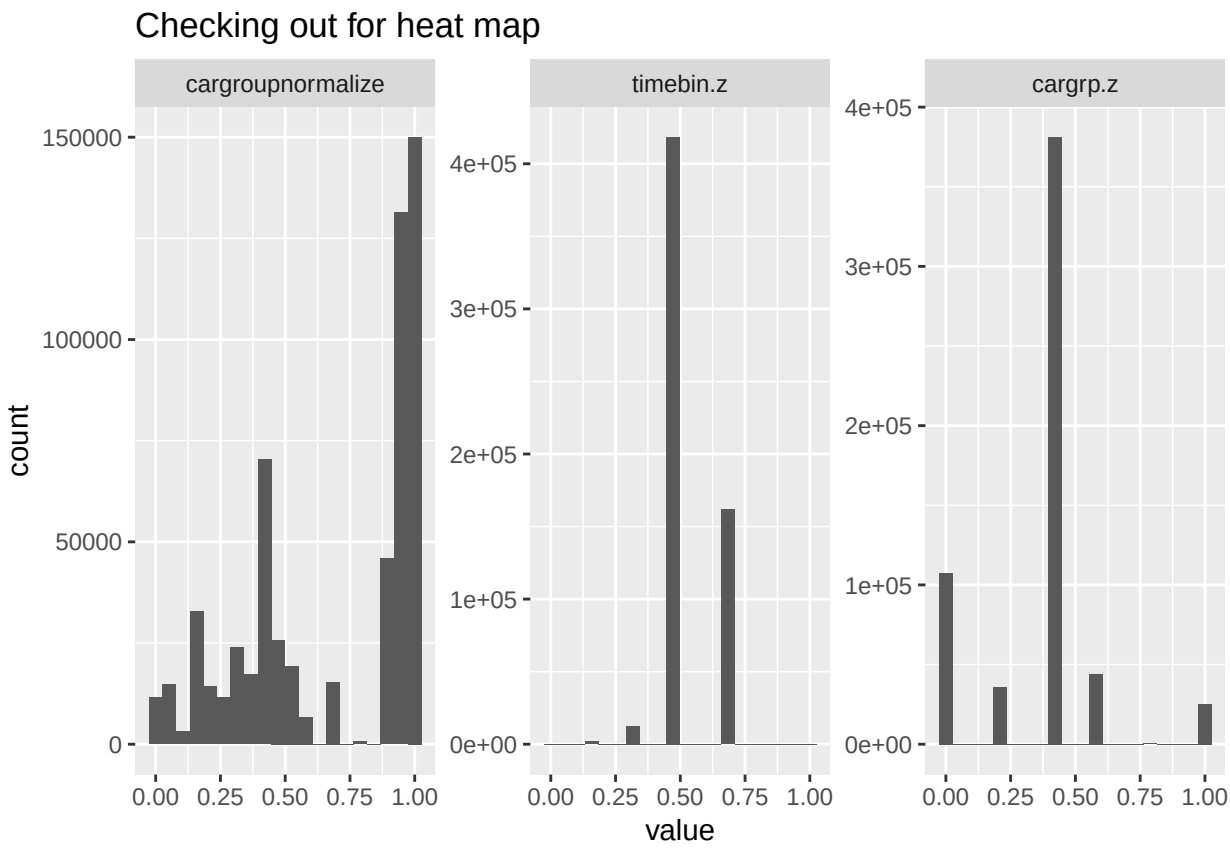
```
## [1] 6
```

```
range(df3.heat$issuetimebin)
```

```
## [1] 0 6
```

```
df3.heat$timebin.z <- (df3.heat[,5]-min(df3.heat$issuetimebin))/
  (max(df3.heat$issuetimebin)-min(df3.heat$issuetimebin))
df3.heat$cargrp.z <- (df3.heat[,6]-min(df3.heat$cargroup))/
  (max(df3.heat$cargroup)-min(df3.heat$cargroup))
# head(df3.heat,1)
# length(df3.heat)
dftemp <-data.frame(df3.heat) #    <----switch to dataframe w no dots in it
 #============================================================
#hwat I really want is by car groups........
cargrp <- sqldf('select issuemonth as month, issueweekday as day,
        cargroupname , cargroupnormalize as zscore from dftemp ')
#Comparison to learn if need to transform data to interpret better in heatmap
ggplot(data = melt(df3.heat[,c(8:10)]), mapping = aes(x = value)) + geom_histogram(bins=20)+
  facet_wrap(~variable, scales = "free") +
  ggtitle("Checking out for heat map")
```

```
## No id variables; using all as measure variables
```



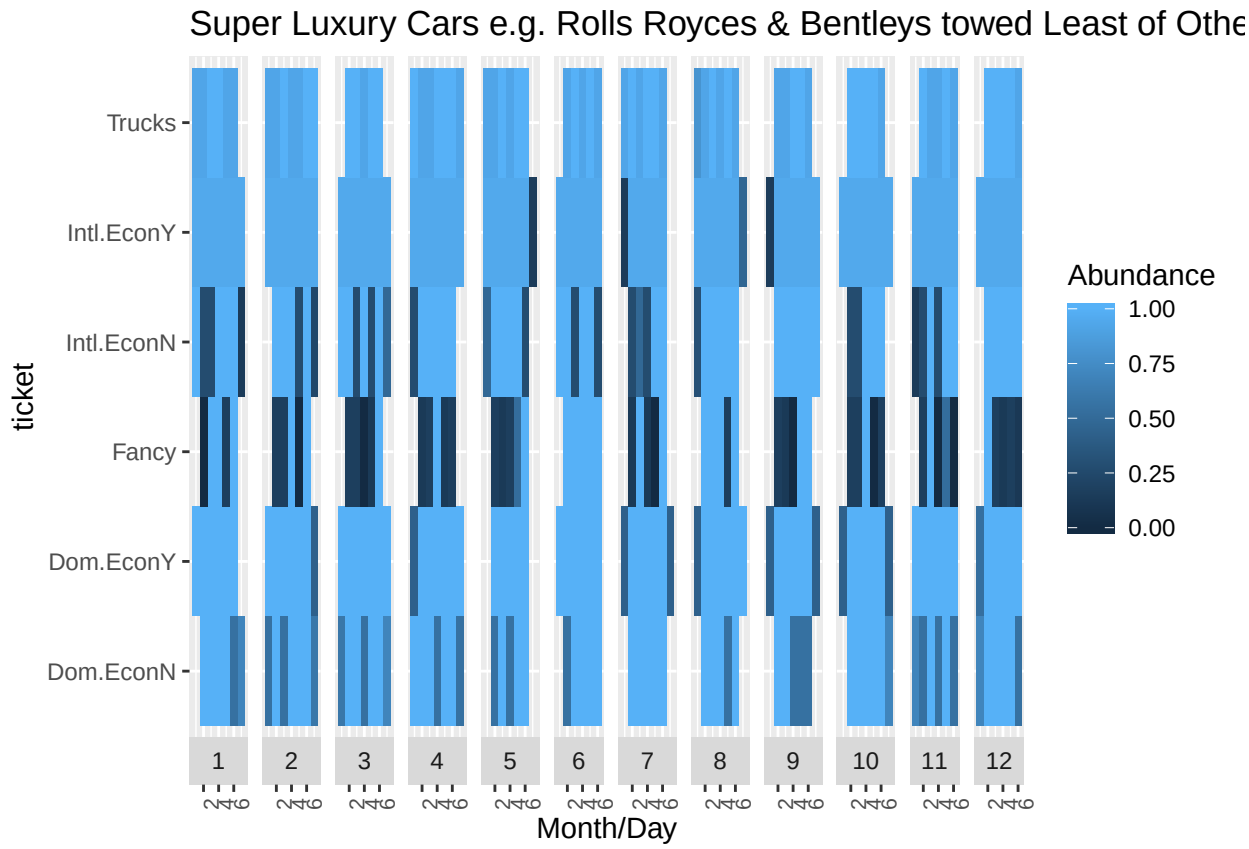
```
#install.packages("tidyr")      #CLASS SLACK HELP was INVALUABLE !
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.5.2
```

```
##
## Attaching package: 'tidyr'
```

```
## The following object is masked from 'package:reshape2':
##
##     smiths
```

```
#Abundance = grouping of all air quality factor values 0-1
#Felt Square Root Transformation did help with graph read but could vary by person
df4.heat <- gather(data = cargrp, key = Class, value= Abundance,-c(1:3))
df4.heat$Sqrt.Abundance <- sqrt(df4.heat$Abundance)
heat.reg <-ggplot(data=df4.heat, mapping= aes(x=day,y=cargroupname,fill=Abundance))+
  geom_tile() + xlab(label="Month/Day") + ylab(label="ticket") +
  ggtitle("Super Luxury Cars e.g. Rolls Royces & Bentleys towed Least of Others but Still 3% or 21,600 c
  theme(axis.text.x = element_text(size=8,angle = 90, hjust = 1)) +
  facet_grid(~ month, switch = "x", scales="free_x", space="free_x")
heat.reg #non transformed data
```



```
# heat.reg <-ggplot(data=df4.heat, mapping= aes(x=day,y=cargroupname,fill=Sqrt.Abundance))+
#   geom_tile() + xlab(label="Month/Day") + ylab(label="ticket") +
#   ggtitle("Super Luxury Cars e.g. Rolls Royces & Bentleys are towed Least of All Others
#           but 3% of Total Cars or 21.6k/7.6M Tows in 2018") +
#   theme(axis.text.x = element_text(size=8,angle = 90, hjust = 1)) +
#   facet_grid(~ month, switch = "x", scales="free_x", space="free_x")
# heat.reg #non transformed data

#HOUSE KEEPING-------------------------------
#remove(df0, df1,df.temp, cor.df1, route.income.df1, df.carid.master)
```

[redacted]+BRIAN+HOGAN+CORRELATION+HEATMAP+PREDIC