Heresa me again

HERE IS ME

Parking\_JFK

Christopher Eshleman

5/25/2019

### Aiport Parking at JFK

## Aiport Parking at JFK

# Aiport Parking at JFK

library(leaps)   
library(zoo)

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library(olsrr)

## Warning: package 'olsrr' was built under R version 3.4.4

## Warning in as.POSIXlt.POSIXct(Sys.time()): unknown timezone 'zone/tz/2019a.  
## 1.0/zoneinfo/America/New\_York'

##   
## Attaching package: 'olsrr'

## The following object is masked from 'package:datasets':  
##   
## rivers

library(glmnet)

## Warning: package 'glmnet' was built under R version 3.4.2

## Loading required package: Matrix

## Loading required package: foreach

## Loaded glmnet 2.0-13

library(lubridate)

## Warning: package 'lubridate' was built under R version 3.4.4

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

library(ISLR)

## Warning: package 'ISLR' was built under R version 3.4.2

library(broom)   
  
#setwd("C:/Users/ceshleman/OneDrive - The Port Authority of New York & New Jersey/MAIN WORK ESHLEMANC/Forecasting/Port Exports 2019")   
setwd("~/OneDrive - The Port Authority of New York & New Jersey/MAIN WORK ESHLEMANC/Airport parking")   
a = read.csv("./Avg\_Domestic\_Ticket\_Prices.csv")   
 a$Year = as.numeric(as.character(a$Year))

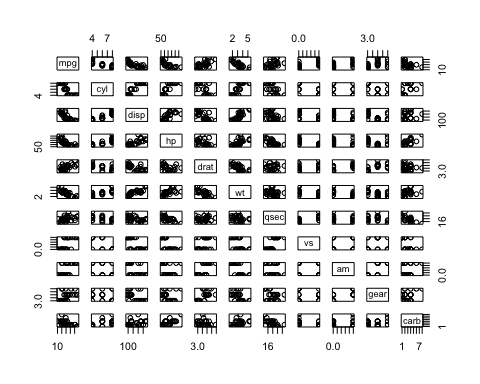
## Warning: NAs introduced by coercion

a$quarter = paste0("0", a$Quarter)  
a$quarter = as.Date(paste(a$Year,paste0("0", a$Quarter),"01",sep="-"))   
a = subset(a,a$quarter>="2009-01-01")   
b = read.csv("./EWR\_Data.csv")   
c = read.csv("./JFK\_Data.csv")   
d = read.csv("./Consumer\_Spending.csv")   
dd = read.csv("./Consumer\_Spending\_short.csv")   
c$month = as.Date(c$month, format="%m/%d/%y")   
d$month = as.Date(d$month, format="%m/%d/%y")   
cd = merge(c,d, by="month",all.x=TRUE)   
cd$Year = NULL   
cd$Month = NULL   
cd$m = month(cd$month)   
cd$quarter = substr(quarters(cd$month), 2, 2)  
cd$yr = year(cd$month)   
cd$quarter = paste(year(cd$month),cd$quarter,sep="-")   
cd$m = NULL   
cd$q = NULL   
cd$yr = NULL   
  
cd = aggregate(. ~ quarter, cd, FUN = mean, na.action=na.pass)   
cd$q = seq(as.Date("2006/1/1"), as.Date("2019/1/1"), by = "quarter")  
cd = subset(cd,cd$q<"2018-04-01")   
  
names(dd) = tolower(names(dd))   
dd$month = as.Date(dd$month, format="%m/%d/%y")   
cdd = merge(c,dd,by="month")   
  
cdd$m = month(cdd$month)   
cdd$quarter = substr(quarters(cdd$month), 2, 2)  
cdd$yr = year(cdd$month)   
cdd$quarter = paste(year(cdd$month),cdd$quarter,sep="-")   
cdd$m = NULL   
cdd$q = NULL   
cdd$yr = NULL   
cdd = aggregate(. ~ quarter, cdd, FUN = mean, na.action=na.pass)   
cdd$q = seq(as.Date("2006/1/1"), as.Date("2019/1/1"), by = "quarter")  
cdd = subset(cdd,cdd$q<"2018-04-01")

## Variable selection

When running remember to remove id column, response column, factors, anything with too many NAs.

First, check for strong correlations.



# First remove things with NAs

# Then remove things that are characters (not numeric)

## term estimate std.error  
## 1 (Intercept) 5.174126e+05 1.831682e+05  
## 2 Total.Revenue 6.937866e-02 2.484049e-03  
## 3 Avg..Parking.Paid -2.336965e+04 1.769950e+03  
## 4 Cell.Service -7.495096e-02 4.086535e-01  
## 5 DOMESTIC.Enplaments 5.183288e-02 4.565728e-02  
## 6 INTERNATIONAL.Enplanements -7.563308e-02 3.829138e-02  
## 7 JFK\_Airtrain 5.961604e-02 6.902750e-02  
## 8 personal.consumption.expenditures -1.683461e+00 8.757067e-01  
## 9 all.goods 2.308434e+00 1.569246e+00  
## 10 durable.goods -6.209570e-01 8.262559e-01  
## 11 nondurable.goods -6.595183e-01 8.601136e-01  
## 12 all.services 1.759395e+00 8.893179e-01  
## 13 household.services -9.367718e-02 1.469912e-01  
## statistic p.value  
## 1 2.8247957 7.667156e-03  
## 2 27.9296675 5.400464e-26  
## 3 -13.2035680 2.292308e-15  
## 4 -0.1834096 8.555058e-01  
## 5 1.1352598 2.637686e-01  
## 6 -1.9751987 5.595170e-02  
## 7 0.8636564 3.934946e-01  
## 8 -1.9224029 6.249580e-02  
## 9 1.4710468 1.499662e-01  
## 10 -0.7515312 4.572183e-01  
## 11 -0.7667804 4.482116e-01  
## 12 1.9783643 5.557893e-02  
## 13 -0.6372979 5.279633e-01

## Backward Elimination Method   
## ---------------------------  
##   
## Candidate Terms:   
##   
## 1 . Total.Revenue   
## 2 . Avg..Parking.Paid   
## 3 . Cell.Service   
## 4 . DOMESTIC.Enplaments   
## 5 . INTERNATIONAL.Enplanements   
## 6 . TOTAL.Enplanments   
## 7 . JFK\_Airtrain   
## 8 . personal.consumption.expenditures   
## 9 . all.goods   
## 10 . durable.goods   
## 11 . nondurable.goods   
## 12 . all.services   
## 13 . household.services   
##   
## We are eliminating variables based on p value...  
##   
## Variables Removed:   
##   
## - Cell.Service   
## - household.services   
## - durable.goods   
## - JFK\_Airtrain   
##   
## No more variables satisfy the condition of p value = 0.3  
##   
##   
## Final Model Output   
## ------------------

## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length

## Model Summary   
## --------------------------------------------------------------------  
## R 0.994 RMSE 7469.596   
## R-Squared 0.989 Coef. Var 1.916   
## Adj. R-Squared 0.987 MSE 55794866.794   
## Pred R-Squared 0.977 MAE 5121.448   
## --------------------------------------------------------------------  
## RMSE: Root Mean Square Error   
## MSE: Mean Square Error   
## MAE: Mean Absolute Error   
##   
## ANOVA   
## ---------------------------------------------------------------------------------  
## Sum of   
## Squares DF Mean Square F Sig.   
## ---------------------------------------------------------------------------------  
## Regression 201148427587.241 8 25143553448.405 450.643 0.0000   
## Residual 2231794671.766 40 55794866.794   
## Total 203380222259.007 48   
## ---------------------------------------------------------------------------------  
##   
## Parameter Estimates   
## ---------------------------------------------------------------------------------------------------------------------------  
## model Beta Std. Error Std. Beta t Sig lower upper   
## ---------------------------------------------------------------------------------------------------------------------------  
## (Intercept) 497606.535 144057.986 3.454 0.001 206454.486 788758.585   
## Total.Revenue 0.069 0.002 1.296 31.084 0.000 0.065 0.073   
## Avg..Parking.Paid -23700.443 1440.406 -0.695 -16.454 0.000 -26611.613 -20789.274   
## DOMESTIC.Enplaments 0.040 0.030 0.053 1.302 0.200 -0.022 0.101   
## INTERNATIONAL.Enplanements -0.052 0.023 -0.172 -2.278 0.028 -0.097 -0.006   
## TOTAL.Enplanments NA 0.486 -4.170 -1.900 0.065 NA NA   
## personal.consumption.expenditures -0.924 0.508 10.832 1.870 0.069 -1.907 0.059   
## all.goods 0.950 0.115 -0.646 -1.099 0.278 -0.077 1.978   
## nondurable.goods -0.127 0.483 1.927 1.936 0.060 -0.360 0.106   
## all.services 0.934 NA 0.449 NA NA -0.041 1.909   
## ---------------------------------------------------------------------------------------------------------------------------

##   
##   
## Elimination Summary   
## ---------------------------------------------------------------------------------------  
## Variable Adj.   
## Step Removed R-Square R-Square C(p) AIC RMSE   
## ---------------------------------------------------------------------------------------  
## 1 Cell.Service 0.9897 0.9866 11.0336 1028.2494 7541.2214   
## 2 household.services 0.9895 0.9868 9.4702 1026.8395 7486.2710   
## 3 durable.goods 0.9893 0.9868 8.2314 1025.8517 7466.3961   
## 4 JFK\_Airtrain 0.989 0.9868 7.2188 1025.1343 7469.5962   
## ---------------------------------------------------------------------------------------

## Length Class Mode   
## mallows\_cp 4 -none- numeric   
## removed 4 -none- character  
## rsquare 4 -none- numeric   
## indvar 13 -none- character  
## steps 1 -none- numeric   
## sbic 4 -none- numeric   
## adjr 4 -none- numeric   
## rmse 4 -none- numeric   
## aic 4 -none- numeric   
## sbc 4 -none- numeric   
## model 12 lm list

## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length

## Backward Elimination Method   
## ---------------------------  
##   
## Candidate Terms:   
##   
## 1 . Total.Revenue   
## 2 . Avg..Parking.Paid   
## 3 . Cell.Service   
## 4 . DOMESTIC.Enplaments   
## 5 . INTERNATIONAL.Enplanements   
## 6 . TOTAL.Enplanments   
## 7 . JFK\_Airtrain   
## 8 . personal.consumption.expenditures   
## 9 . all.goods   
## 10 . durable.goods   
## 11 . nondurable.goods   
## 12 . all.services   
## 13 . household.services

## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length  
  
## Warning in b \* sx: longer object length is not a multiple of shorter object  
## length

## Warning: `as\_dictionary()` is soft-deprecated as of rlang 0.3.0.  
## Please use `as\_data\_pronoun()` instead  
## This warning is displayed once per session.

## Warning: `new\_overscope()` is soft-deprecated as of rlang 0.2.0.  
## Please use `new\_data\_mask()` instead  
## This warning is displayed once per session.

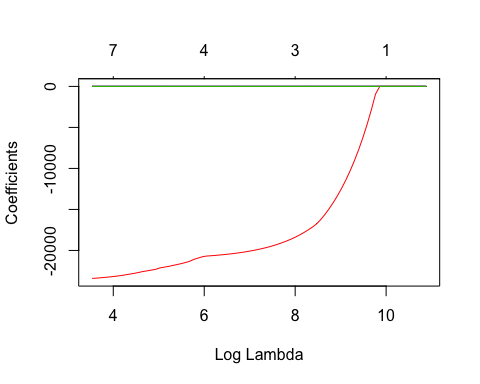
## Warning: The `parent` argument of `new\_data\_mask()` is deprecated.  
## The parent of the data mask is determined from either:  
##   
## \* The `env` argument of `eval\_tidy()`  
## \* Quosure environments when applicable  
## This warning is displayed once per session.

## Warning: `overscope\_clean()` is soft-deprecated as of rlang 0.2.0.  
## This warning is displayed once per session.

##   
## Variables Removed:   
##   
## - INTERNATIONAL.Enplanements   
## - Cell.Service   
## - household.services   
## - durable.goods   
## - JFK\_Airtrain   
## - nondurable.goods   
##   
## No more variables to be removed.

##   
##   
## Backward Elimination Summary   
## ----------------------------------------------------------------------------------------------------  
## Variable AIC RSS Sum Sq R-Sq Adj. R-Sq   
## ----------------------------------------------------------------------------------------------------  
## Full Model 1030.204 2102226397.549 2.01278e+11 0.98966 0.98622   
## INTERNATIONAL.Enplanements 1028.204 2102226397.549 2.01278e+11 0.98966 0.98622   
## Cell.Service 1026.249 2104190757.305 2.01276e+11 0.98965 0.98658   
## household.services 1024.839 2129681640.833 201250540618.174 0.98953 0.98677   
## durable.goods 1023.852 2174135737.018 201206086521.989 0.98931 0.98684   
## JFK\_Airtrain 1023.134 2231794671.766 201148427587.241 0.98903 0.98683   
## nondurable.goods 1022.592 2299179566.461 2.01081e+11 0.98870 0.98677   
## ----------------------------------------------------------------------------------------------------

## Paid.Parking.Activity Total.Revenue Avg..Parking.Paid Cell.Service  
## 1 382943.7 6555556 17.16667 68595.00  
## 2 366933.0 5362625 14.66667 70657.33  
## 3 437769.3 5858919 13.38333 72224.67  
## 4 321068.3 4982998 15.55667 74172.67  
## 5 117007.3 1534286 13.29667 75925.33  
## 6 287156.3 4714868 16.64000 76952.00  
## DOMESTIC.Enplaments INTERNATIONAL.Enplanements TOTAL.Enplanments  
## 1 856900.3 646472.7 1503373  
## 2 922923.7 845178.3 1768102  
## 3 1012010.0 937801.7 1949812  
## 4 1005413.0 797049.0 1802462  
## 5 957618.0 721472.0 1679090  
## 6 1092813.0 915841.3 2008654  
## JFK\_Airtrain personal.consumption.expenditures all.goods durable.goods  
## 1 281482.0 10288892 3478259 1037859  
## 2 329285.7 10341016 3483329 1038292  
## 3 352022.0 10403798 3513963 1053159  
## 4 357215.0 10504481 3563194 1068013  
## 5 320680.7 10563261 3588229 1082376  
## 6 380516.3 10582839 3596629 1093736  
## nondurable.goods all.services household.services  
## 1 2442281 6807418 6584077  
## 2 2447242 6855184 6623770  
## 3 2461263 6886717 6651318  
## 4 2495483 6937420 6703669  
## 5 2503947 6970971 6735709  
## 6 2498651 6982089 6750518



## 14 x 1 sparse Matrix of class "dgCMatrix"  
## 1  
## (Intercept) 3.682926e+05  
## Total.Revenue 6.874375e-02  
## Avg..Parking.Paid -2.340875e+04  
## Cell.Service 2.151638e-01  
## DOMESTIC.Enplaments 2.682980e-02  
## INTERNATIONAL.Enplanements -4.295897e-02  
## TOTAL.Enplanments .   
## JFK\_Airtrain .   
## personal.consumption.expenditures .   
## all.goods .   
## durable.goods -1.876006e-02  
## nondurable.goods -6.971073e-03  
## all.services .   
## household.services .