Parking_JFK

Christopher Eshleman 5/25/2019

Aiport Parking at JFK

Load data on airport parking activity, ticket prices and consumer spending. Remove things with too many blanks (NAs) and factor variables.

Variable selection

When running remember to remove id column, response column, factors, anything with too many NAs. First, check for strong correlations.

1. Kitchen sink (less useful)

```
##
                                             estimate
                                                         std.error
                                                                     statistic
## 1
                            (Intercept)
                                         1.613672e+06 8.952553e+05
                                                                     1.8024717
## 2
                           Cell.Service -1.569221e+00 1.954613e+00 -0.8028298
## 3
                    DOMESTIC.Enplaments -3.475196e-01 2.106457e-01 -1.6497827
## 4
             INTERNATIONAL.Enplanements 5.045144e-01 1.573618e-01 3.2060795
## 5
                           JFK_Airtrain -1.252246e-01 3.446887e-01 -0.3632977
     personal.consumption.expenditures -1.337968e+00 4.140082e+00 -0.3231744
##
## 7
                              all.goods 2.133976e+00 7.052804e+00 0.3025713
## 8
                          durable.goods -4.199836e-01 3.690041e+00 -0.1138154
## 9
                       nondurable.goods -2.393256e+00 3.703669e+00 -0.6461853
## 10
                           all.services 7.682370e-01 4.271735e+00 0.1798419
## 11
                     household.services 9.425703e-01 7.115910e-01 1.3245956
##
         p.value
     0.07940720
## 1
## 2
     0.42706613
## 3
     0.10723074
## 4
     0.00272796
## 5
     0.71839579
## 6
     0.74833495
## 7
     0.76386670
## 8
     0.90998332
## 9
     0.52204302
## 10 0.85823228
## 11 0.19321665
```

2. Stepwise

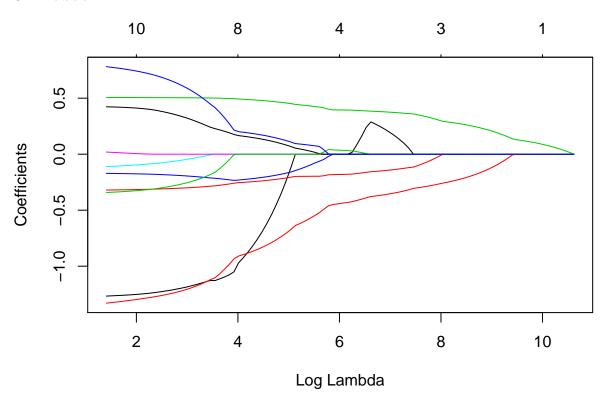
```
## Backward Elimination Method
## ------
```

```
##
## Candidate Terms:
##
## 1 . Cell.Service
## 2 . DOMESTIC.Enplaments
## 3 . INTERNATIONAL.Enplanements
## 4 . TOTAL.Enplanments
## 5 . JFK_Airtrain
## 6 . personal.consumption.expenditures
## 7 . all.goods
## 8 . durable.goods
## 9 . nondurable.goods
## 10 . all.services
## 11 . household.services
## We are eliminating variables based on p value...
## Variables Removed:
##
## - durable.goods
## - all.services
## - JFK Airtrain
## - Cell.Service
## No more variables satisfy the condition of p value = 0.3
##
## Final Model Output
                      Model Summary
## R
                     0.839
                              RMSE
                                                  37916.752
                  0.703 Coef. Var
0.661 MSE
0.569 MAE
## R-Squared
                                                   9.725
## Adj. R-Squared
                                               1437680110.009
## Pred R-Squared
                                                    25417.052
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
                                ANOVA
##
                      Sum of
                     Squares DF Mean Square F Sig.
## ------
## Regression 142997657638.634 6 23832942939.772 16.577 0.0000
                                42 1437680110.009
## Residual 60382564620.373
            203380222259.007
                                 48
##
                                             Parameter Estimates
                                      Beta Std. Error Std. Beta t
##
                         model
                                                                             Sig
```

:									
!		(Intercept)	1940670.318		21.635	-0.348	3.107	0.003	
:	DOMESTIC.Enplament		-0.260		0.156		-1.672	0.102	
<u> </u>	INTERNATIONAL.En		0.429		0.083	1.428	5.171	0.000	
:	${\tt TOTAL.Enplanment} \\ {\tt personal.consumption.expenditure} \\$		NA		0.639	11.952	-1.525	0.135	
-			-0.722				1.642	0.108	
	_	all.goods	1.049		0.788	-9.122 2.024	-2.268		
	nondurable.goods		-1.788				1.653		
	househol	ld.services 	0.981	l 	NA 	-1.564 	NA 	NA	
		Elim	ination Summa	ary					
	Variable								
Step	Variable Removed	R-Square	Adj. R-Square	C(p)	AIC	RM	SE		
1	durable.goods	0.7102	0.6433	9.0130	1187.5424	 38873	 .6613		
2	all.services	0.7102	0.6521						
3		0.7086	0.6521		1183.8087				
4	Cell.Service	0.7031	0.6607		1182.7311				
3 . IN 4 . TO 5 . JF 6 . pe 7 . al 8 . du 9 . no 10 . a	MESTIC.Enplaments TERNATIONAL.Enpla TAL.Enplanments K_Airtrain rsonal.consumptic l.goods rable.goods ndurable.goods ll.services ousehold.services	anements on.expenditur	res						
Variab	les Removed:								
- TOTA	L.Enplanments								
No mor	e variables to be	e removed.							
			Backward El	liminatio	n Summarv				
Variab							 R-Sa		
		AIC	; 			 	R-Sq	нај 	
Full M	odel	1189.	526 589152	216359.44	3 1.	44465e+11	0.7103	2 0	.6

## INTERNATIONAL.Enplanements	1187.526	74851698145.534	128528524113.473	0.63196	0.5470
## TOTAL.Enplanments	1187.526	NA	NA	NA	N
##					

3. Lasso



12 x 1 sparse Matrix of class "dgCMatrix" ## ## (Intercept) 1.450614e+06 ## Cell.Service -1.266763e+00 ## DOMESTIC.Enplaments -3.201383e-01 ## INTERNATIONAL.Enplanements 5.066759e-01 ## TOTAL.Enplanments ## JFK_Airtrain -1.711005e-01 ## personal.consumption.expenditures -1.104534e-01 ## all.goods 1.916618e-02 ## durable.goods 4.232946e-01 ## nondurable.goods -1.329763e+00 ## all.services -3.418267e-01 ## household.services 7.812882e-01

Stepwise guidance suggests sticking with international emplanements and goods, and personal consumption is close.