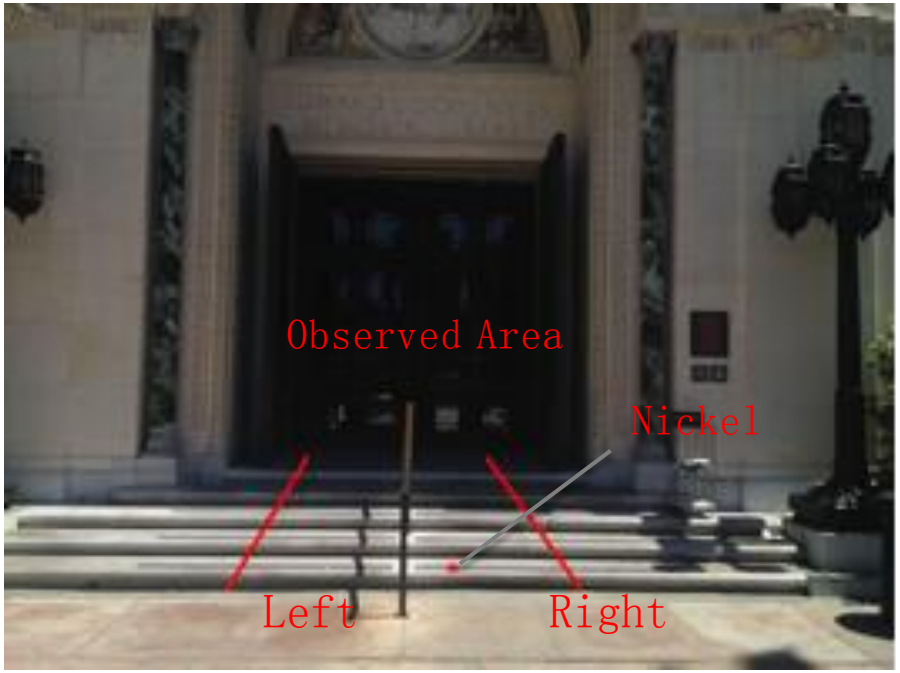
EE 517 —— PROJECT PICK UP THE NICKEL!

Wenhai Zhu
Chenxi Zhang

Data Collection

Collect time:
Apr. 8th - Apr. 21st 2pm – 5pm
Apr. 24th – Apr. 27th 10am – 5pm



Collection Rule

We have 6 rules when collecting data.

Variables

- Gender. Contain male and female.
- Age. Contain young, middle, and old.
- Race. Contain white, black, and yellow.
- Suit. Contain in suit and not in suit.
- Backpack. Contain no backpack, one-strip backpack, and two-strips backpack.
- Empty hands. Contain no hands are empty, one hand is empty, and two hands are empty.
- Doing something. Contain doing something and not doing something.
- Running. Contain running and not running.
- Friends. Contain alone, and with friends.
- Stranger. Contain no strangers and with strangers.
- Direction. Contain in and out.
- Side. Contain left side and right side.

What we found

Logit(odds) =

-2.846 + 3.006*doingsth – 1.483*friends – 1.303*direction

Observed	Predicted		
	Not Pick	Pick	Percentage Correct
Not Pick	933	30	96.9%
Pick	96	27	22.0%
Overall percentage			88.4%

We built a model to predict whether an individual will pick up the nickel.

Model Test

	B	S.E.	Wald	Sig.
doingsth	3.006	0.452	44.281	0.000
friends	-1.483	0.238	38.864	0.000
direction	-1.303	0.210	38.376	0.000

Test	Statistic	Sig.
Hosmer and Lemeshow Test	0.952	0.917
-2 log likelihood	628.476	--
Cox & Snell R^2	0.120	--
Nagelkerke R^2	0.237	--

Our model is a significant model

Multicollinearity

	Constant	doingSth	friends	direction
Constant	1.000	-0.875	-0.070	-0.166
dotingSth	-0.875	1.000	-0.304	-0.028
friends	-0.070	-0.304	1.000	-0.068
direction	-0.166	-0.028	-0.068	1.000

	Eigenvalue	Condition Index	Tolerance	VIF
doingSth	0.766	1.790	0.805	1.242
friends	0.432	2.382	0.793	1.261
direction	0.348	2.654	0.982	1.018

Eigenvalue !=0, CI < 10, VIF < 5

Our model does NOT have multicollinearity

Cross Validation(50/50)

Iteration	Constant	doingSth	friends	direction
1	0.000	0.000	0.000	0.001
2	0.000	0.000	0.000	0.003
3	0.000	0.000	0.000	0.003
4	0.000	0.000	0.017	0.000
5	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000
9	0.000	0.000	0.001	0.000
10	0.000	0.000	0.000	0.000

- Model is significant.

Residual Test

Test	Maximum Value
Cook’ s Distance	0.2166
Leverage Value	0.01204
DFBETA for Constant	0.17103
DFBETA for doingSth	0.02707
DFBETA for friends	0.04902
DFBETA for direction	0.02909
Standard Residual in [-1.96,1.96]	96.8%
Standard Residual in [-2.58,2.58]	99.8%
Standard Residual in [-3,3]	100%

No case exert an undue influence on the model.

Conclusion and Weakness

Conclusion:

We can build a model to predict whether an individual will pick up the nickel with variables, doingSth, friends, direction

An individual with friends, doing nothing, when s/he exits the library are more possible to pick up the nickel.

Weakness:

- Some variables are subjective, such as age, race. Sometimes it is difficult to decide the values of these variables.
- Maybe the front door of Doheny is not a good place to put the nickel, because only less than 10% people could notice the nickel.
- As a few people could notice the nickel, it is not easy to collect data. We need more time and more data.

Appendix

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	138.841	3	.000
	Block	138.841	3	.000
	Model	138.841	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	628.476[a]	.120	.237

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.952	4	.917

We have a significant model.

Appendix

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	doingsth(1)	3.006	.452	44.281	1	.000	20.208
	friends(1)	-1.483	.238	38.864	1	.000	.227
	direction(1)	-1.303	.210	38.376	1	.000	.272
	Constant	-2.846	.423	45.197	1	.000	.058

Classification Table[a]

			Predicted		
			pick		Percentage Correct
	Observed		0	1	
Step 1	pick	0	933	30	96.9
		1	96	27	22.0
	Overall Percentage				88.4

Appendix

Correlation Matrix

		Constant	doingsth(1)	friends(1)	direction(1)
Step 1	Constant	1.000	-.875	-.070	-.166
	doingsth(1)	-.875	1.000	-.304	-.028
	friends(1)	-.070	-.304	1.000	-.068
	direction(1)	-.166	-.028	-.068	1.000

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	direction	doing sth.	friends
1	1	1.577	1.000	.21	.21		
	2	.423	1.930	.79	.79		
2	1	1.953	1.000	.11	.11	.11	
	2	.691	1.681	.00	.42	.58	
	3	.356	2.344	.89	.47	.31	
3	1	2.453	1.000	.06	.05	.06	.06
	2	.766	1.790	.04	.45	.20	.12
	3	.432	2.382	.25	.08	.28	.73
	4	.348	2.654	.65	.42	.46	.09

Appendix

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	Collinearity Statistics
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.063	.011		5.525	.000		
	direction	.150	.020	.223	7.523	.000	1.000	1.000
2	(Constant)	.100	.013		7.911	.000		
	direction	.151	.020	.224	7.705	.000	1.000	1.000
	doing sth.	-.131	.021	-.186	-6.400	.000	1.000	1.000
3	(Constant)	.085	.013		6.727	.000		
	direction	.134	.019	.200	6.927	.000	.982	1.018
	doing sth.	-.195	.022	-.276	-8.688	.000	.805	1.242
	friends	.150	.023	.206	6.429	.000	.793	1.261

Appendix

	Analog of Cook's influence DFBETA for statistics direction(1)	Leverage Standard value residual	Normalized residual	DFBETA for constant	DFBETA for doingsth(1)	DFBETA for friends(1)
Grand Total						
Minimum	.00000	.00082	-1.08306	-.00993	-.19811	-.03312
	-.01803	-1.25346				
Maximum	.21660	.01204	8.71320	.17103	.02607	.04902
	.02909	2.95132				
In -1.96 to 1.96	100.0%	100.0%	96.8%	100.0%	100.0%	100.0%
	100.0%	96.8%				

Appendix

	Analog of Cook's influence DFBETA for statistics direction(1)	Leverage Standard value residual	Normalized residual	DFBETA for constant	DFBETA for doingsth(1)	DFBETA for friends(1)
Grand Total						
Minimum	.00000	.00082	-1.08306	-.00993	-.19811	-.03312
	-.01803	-1.25346				
Maximum	.21660	.01204	8.71320	.17103	.02607	.04902
	.02909	2.95132				
In -2.58 to 2.58	100.0%	100.0%	96.8%	100.0%	100.0%	100.0%
	100.0%	99.8%				