

Puzzling Election Results?

Amit Nayak, Soomin Park, Junfei Zheng, Tianweibao Zheng, Ziqing Zhu

George Washington University

December 5, 2016

Data Table

State	Trump %	Hillary %	Income	GDP	Unemployment Rate	Education Level	White Ratio	M to F Index	LGBT Ratio
Alabama	62.9	34.6	38030	207248	5.4	24.2	65	94.33	2.8
Alaska	52.9	37.7	56147	50592	6.9	29.7	58	108.52	3.4
Arizona	49.5	45.4	39156	296555	5.5	27.7	51	98.74	3.9
Arkansas	60.4	33.8	38252	125759	4	21.8	74	96.45	3.5
California	33.1	61.6	53741	2513807	5.5	32.3	39	98.83	4
Colorado	44.4	47.3	50899	320176	3.6	39.2	69	100.48	3.2
Connecticut	41.2	54.5	68704	262240	5.4	38.3	70	94.83	3.4
Delaware	41.9	53.4	47633	69105	4.3	30.9	63	93.94	3.4
District of Columbia	4.1	92.8	73302	125838	6.1	56.7	37	89.52	10
Florida	49.1	47.8	44429	911025	4.7	28.4	54	95.6	3.5
Georgia	51.3	45.6	40306	508372	5.1	29.9	52	95.38	3.5
Hawaii	30	62.2	48288	81595	3.3	31.4	19	100.32	5.1
Idaho	59.2	27.6	38392	65915	3.8	26	84	100.39	2.7
Illinois	39.4	55.4	50295	784365	5.5	32.9	63	96.24	3.8
Indiana	57.2	37.9	41940	341092	4.5	24.9	80	96.83	3.7
Iowa	51.8	42.2	45902	173444	4.2	26.8	85	98.07	2.8
Kansas	57.2	36.2	47161	148473	4.4	31.7	75	98.45	3.7
Kentucky	62.5	32.7	38588	197649	5	23.3	85	96.85	3.9
Louisiana	58.1	38.4	42947	240028	6.4	23.2	58	95.9	3.2
Maine	45.2	47.9	42799	58125	4.1	30.1	91	95.84	4.8
Maryland	35.3	60.5	55972	371899	4.2	38.8	50	93.63	3.3
Massachusetts	33.5	60.8	62603	485474	3.6	41.5	73	93.66	4.4
Michigan	47.6	47.3	42812	478773	4.6	27.8	76	96.28	3.8
Minnesota	45.4	46.9	50871	336729	4	34.7	82	98.52	2.9
Mississippi	58.3	39.7	34771	108971	6	20.8	56	94.44	2.6
Missouri	57.1	38	42300	298158	5.2	27.8	80	96.01	3.3
Montana	56.5	36	41809	45786	4.3	30.6	90	100.8	2.6
Nebraska	60.3	34	48544	114327	3.2	30.2	79	98.51	2.7
Nevada	45.5	47.9	41889	143995	5.8	23.6	50	102	4.2
New Hampshire	47.3	47.6	55905	74916	2.9	35.7	91	97.35	3.7
New Jersey	41.8	55	59949	580377	5.3	37.6	56	94.84	3.7
New Mexico	40	48.3	37938	90978	6.7	26.5	38	97.66	2.9
New York	37.5	58.8	58670	1474493	5	35	58	93.76	3.8
North Carolina	50.5	46.7	40759	510248	4.7	29.4	62	95	3.3
North Dakota	64.1	27.8	55950	51503	3	29.1	85	102.15	1.7
Ohio	52.1	43.5	43566	618982	4.8	26.8	79	95.39	3.6
Oklahoma	65.3	28.9	45573	174776	5.2	24.6	75	98.03	3.4
Oregon	41.1	51.7	43783	221308	5.5	32.2	74	97.98	4.9
Pennsylvania	48.8	47.6	49745	698517	5.7	29.7	76	95.06	2.7
Rhode Island	39.8	55.4	50018	58295	5.6	32.7	72	93.43	4.5
South Carolina	54.9	40.8	38302	203631	4.9	26.8	64	94.73	2.9
South Dakota	61.5	31.7	47881	47216	2.9	27.5	83	100.14	4.4
Tennessee	61.1	34.9	42094	322944	4.6	25.7	73	95.11	2.6
Texas	52.5	43.4	46947	1568247	4.8	28.4	44	98.41	3.3
Utah	46.3	27.7	39308	159621	3.4	31.8	81	100.93	2.7
Vermont	32.6	61.1	48587	31171	3.3	36.9	94	97.06	4.9
Virginia	45	49.9	52052	492761	4	37	62	96.34	2.9
Washington	37.7	55.1	51898	455943	5.6	34.2	69	99.26	4
West Virginia	68.7	26.5	36758	71992	5.8	19.6	93	97.25	3.1
Wisconsin	47.9	46.9	45914	311191	4.1	28.4	78	98.53	2.8
Wyoming	70.1	22.5	56081	36219	5.3	26.2	85	104.07	2.9

Figure : State-by-State Data

Simple Linear Regression

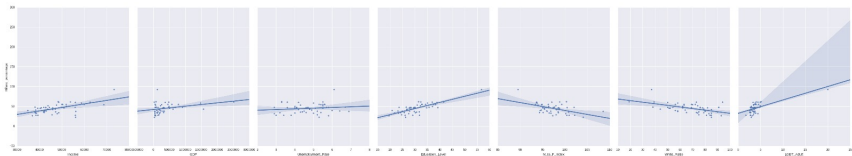


Figure : Single Linear Regression for Clinton

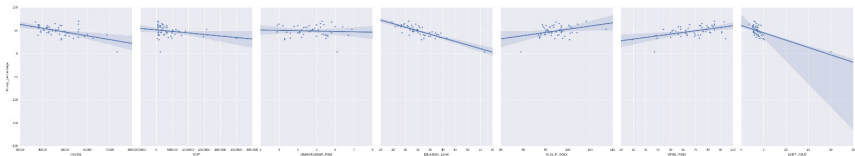


Figure : Single Linear Regression for Trump

Stepwise Regression

```
> stepAANOVA # display results
Stepwise Model Path
Analysis of Deviance Table

Initial Model:
Hillary.percentage ~ Income. + GDP. + Unemployment.Rate + Education.Level +
  White.Ratio + M.to.F.Index + LGBT.Ratio

Final Model:
Hillary.percentage ~ GDP. + Education.Level + White.Ratio + M.to.F.Index +
  LGBT.Ratio
```

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
1				43	1144.508	174.6561
2	- Income.	1	2.127739	44	1146.636	172.7509
3	- Unemployment.Rate	1	5.410736	45	1152.047	170.9910

Figure : Stepwise Regression for Clinton

```
> stepAANOVA # display results
Stepwise Model Path
Analysis of Deviance Table

Initial Model:
Trump.percentage ~ Income. + GDP. + Unemployment.Rate + Education.Level +
  White.Ratio + M.to.F.Index + LGBT.Ratio

Final Model:
Trump.percentage ~ Income. + GDP. + Education.Level + White.Ratio +
  LGBT.Ratio
```

	Step	Df	Deviance	Resid. Df	Resid. Dev	AIC
1				43	1174.175	175.9613
2	- M.to.F.Index	1	1.688724	44	1175.864	174.0346
3	- Unemployment.Rate	1	3.703405	45	1179.567	172.1949

Conclusion:

Figure : Stepwise Regression for Trump

Variable Dependency Calculations

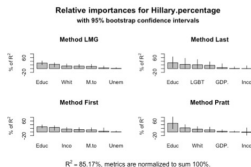


Figure : Variable Impact on Model for Clinton

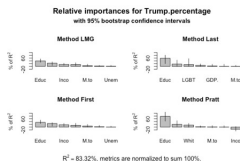


Figure : Variable Impact on Model for Trump

Correlation Matrix

```
> cor(election[,c(4:10)],method="pearson")
```

	Income.	GDP.	Unemployment.Rate	Education.Level	White.Ratio	M.to.F.Index	LGBT.Ratio
Income.	1.00000000	0.185697574	-0.04245328	0.8124473	-0.1620168	-0.1110420	0.434192871
GDP.	0.18569757	1.00000000	0.15708893	0.1192266	-0.3910250	-0.1532202	-0.001754075
Unemployment.Rate	-0.04245328	0.157088931	1.00000000	-0.1477452	-0.4043772	-0.1061634	0.146856231
Education.Level	0.81244726	0.119226646	-0.14774518	1.00000000	-0.2193178	-0.3029699	0.610995020
White.Ratio	-0.16201677	-0.391025004	-0.40437719	-0.2193178	1.00000000	0.1567227	-0.316154984
M.to.F.Index	-0.11104199	-0.153220228	-0.10616344	-0.3029699	0.1567227	1.00000000	-0.344215426
LGBT.Ratio	0.43419287	-0.001754075	0.14685623	0.6109950	-0.3161550	-0.3442154	1.00000000

Figure : Correlation matrix of all our independent variables

Clinton Regression Model

Call:

```
lm(formula = Hillary.percentage ~ Education.Level + White.Ratio +  
    M.to.F.Index + LGBT.Ratio, data = election)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.3504	-2.8903	0.1377	3.1665	11.9085

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	107.83264	25.85945	4.170	0.000133 ***
Education.Level	0.98092	0.14857	6.603	3.60e-08 ***
White.Ratio	-0.23527	0.04672	-5.036	7.79e-06 ***
M.to.F.Index	-0.88197	0.24980	-3.531	0.000954 ***
LGBT.Ratio	2.62890	0.83944	3.132	0.003019 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.165 on 46 degrees of freedom

Multiple R-squared: 0.841, Adjusted R-squared: 0.8272

F-statistic: 60.84 on 4 and 46 DF, p-value: < 2.2e-16

Trump Regression Model

Call:

```
lm(formula = Trump.percentage ~ Education.Level + White.Ratio +  
    LGBT.Ratio, data = election)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.9049	-4.7800	0.3942	3.7334	11.2421

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	78.09813	5.60551	13.932	< 2e-16 ***
Education.Level	-1.13518	0.15218	-7.460	1.65e-09 ***
White.Ratio	0.21058	0.04816	4.372	6.76e-05 ***
LGBT.Ratio	-2.55086	0.84989	-3.001	0.00429 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.33 on 47 degrees of freedom

Multiple R-squared: 0.8103, Adjusted R-squared: 0.7982

F-statistic: 66.93 on 3 and 47 DF, p-value: < 2.2e-16

Figure : Final Regression Model for Donald Trump

Verification Graphs

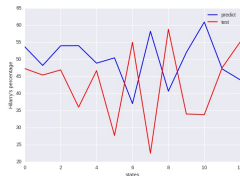


Figure : Comparison of Clinton Model with the Actual Outcomes

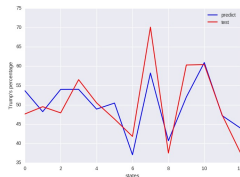
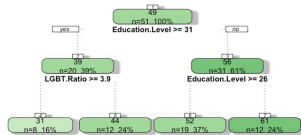
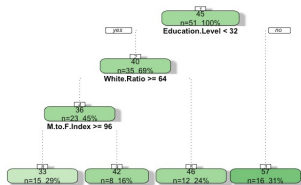


Figure : Comparison of Trump Model with the Actual Outcomes



Rattle 2016-Nov-28 11:06:37 JeffZheng

Figure : Stepwise Model for Hillary Clinton



Rattle 2016-Nov-28 11:12:57 JeffZheng

Figure : Stepwise Model for Donald Trump

Prediction vs. Actual Map

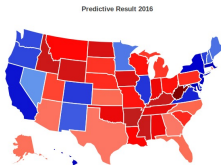


Figure : Map our Model Predicts

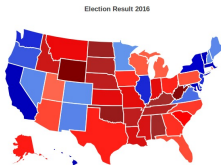


Figure : Actual Election Results Map

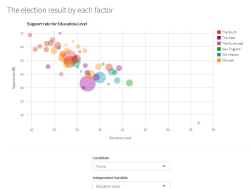


Figure : Graph of Education vs. Trump Support Rate

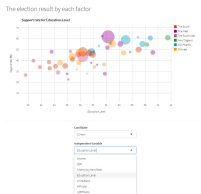


Figure : Graph of Education vs. Clinton Support Rate