A Multivariate Analysis on U.S Presidential Election Results From 1992

Jared Thacker

Objective

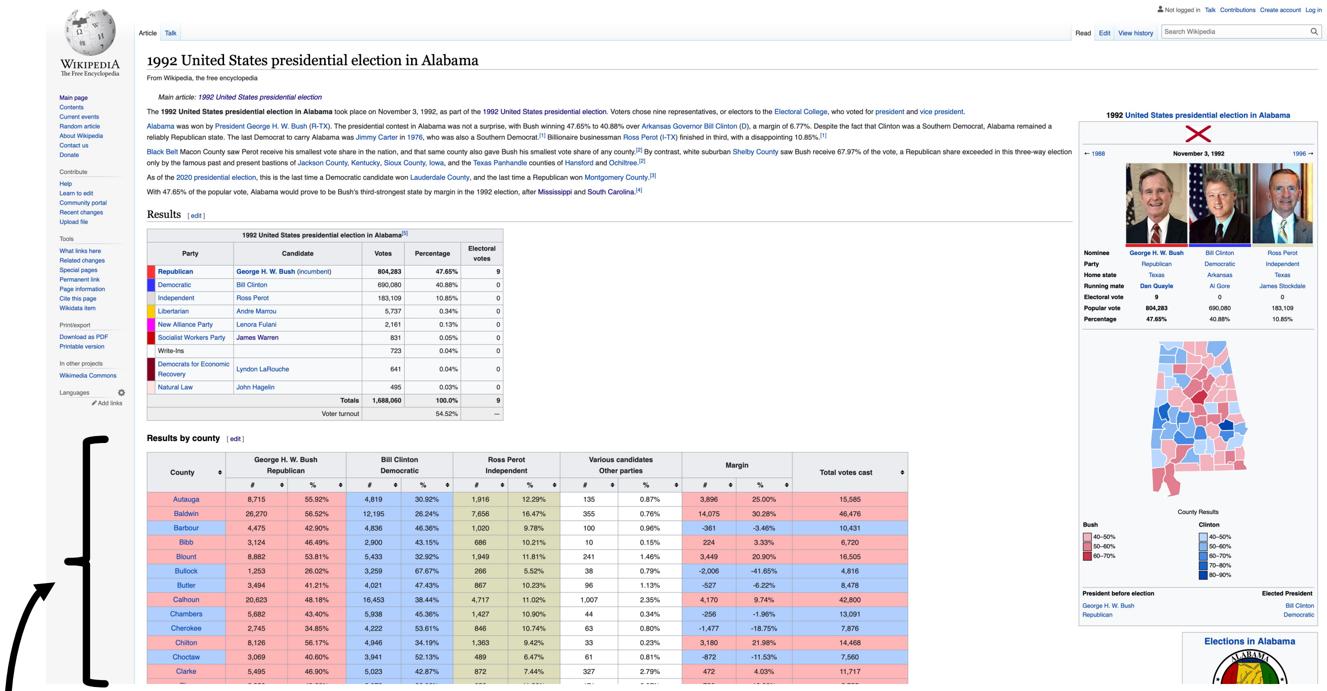
1. Is there a difference in census information between counties that Bill Clinton lost and won?
2. Can we predict the results of an election using *only* census information ignoring the temporal structure? Baseline model?
3. Which variables are the most important?

Democratic Candidate: Bill Clinton

Republican Candidate: George H. W. Bush

Data-Sourcing

- Two Sources
  - Ufl.edu – census variables by county
  - Wikipedia – election results by county



Not logged in Talk Contributions Create account Log in

Article Talk

Read Edit View history Search Wikipedia

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export
Download as PDF
Printable version

In other projects
Wikimedia Commons

Languages
Add links

1992 United States presidential election in Alabama

From Wikipedia, the free encyclopedia

Main article: 1992 United States presidential election

The **1992 United States presidential election in Alabama** took place on November 3, 1992, as part of the 1992 United States presidential election. Voters chose nine representatives, or electors to the Electoral College, who voted for president and vice president.

Alabama was won by President George H. W. Bush (R-TX). The presidential contest in Alabama was not a surprise, with Bush winning 47.65% to 40.88% over Arkansas Governor Bill Clinton (D), a margin of 6.77%. Despite the fact that Clinton was a Southern Democrat, Alabama remained a reliably Republican state. The last Democrat to carry Alabama was Jimmy Carter in 1976, who was also a Southern Democrat.[1] Billionaire businessman Ross Perot (I-TX) finished in third, with a disappointing 10.85%.[1]

Black Belt Macon County saw Perot receive his smallest vote share in the nation, and that same county also gave Bush his smallest vote share of any county.[2] By contrast, white suburban Shelby County saw Bush receive 67.97% of the vote, a Republican share exceeded in this three-way election only by the famous past and present bastions of Jackson County, Kentucky, Sioux County, Iowa, and the Texas Panhandle counties of Hansford and Ochiltree.[2]

As of the 2020 presidential election, this is the last time a Democratic candidate won Lauderdale County, and the last time a Republican won Montgomery County.[3]

With 47.65% of the popular vote, Alabama would prove to be Bush's third-strongest state by margin in the 1992 election, after Mississippi and South Carolina.[4]

Results [edit]

1992 United States presidential election in Alabama[5]

Party	Candidate	Votes	Percentage	Electoral votes
Republican	**George H. W. Bush** (incumbent)	**804,283**	**47.65%**	**9**
Democratic	Bill Clinton	690,080	40.88%	0
Independent	Ross Perot	183,109	10.85%	0
Libertarian	Andre Marrou	5,737	0.34%	0
New Alliance Party	Lenora Fulani	2,161	0.13%	0
Socialist Workers Party	James Warren	831	0.05%	0
Write-Ins		723	0.04%	0
Democrats for Economic Recovery	Lyndon LaRouche	641	0.04%	0
Natural Law	John Hagelin	495	0.03%	0
	Totals	**1,688,060**	**100.0%**	**9**
	Voter turnout		54.52%	—

Results by county [edit]

County	George H. W. Bush Republican		Bill Clinton Democratic		Ross Perot Independent		Various candidates Other parties		Margin		Total votes cast
	#	%	#	%	#	%	#	%	#	%	
Autauga	8,715	55.92%	4,819	30.92%	1,916	12.29%	135	0.87%	3,896	25.00%	15,585
Baldwin	26,270	56.52%	12,195	26.24%	7,656	16.47%	355	0.76%	14,075	30.28%	46,476
Barbour	4,475	42.90%	4,836	46.36%	1,020	9.78%	100	0.96%	-361	-3.46%	10,431
Bibb	3,124	46.49%	2,900	43.15%	686	10.21%	10	0.15%	224	3.33%	6,720
Blount	8,882	53.81%	5,433	32.92%	1,949	11.81%	241	1.46%	3,449	20.90%	16,505
Bullock	1,253	26.02%	3,259	67.67%	266	5.52%	38	0.79%	-2,006	-41.65%	4,816
Butler	3,494	41.21%	4,021	47.43%	867	10.23%	96	1.13%	-527	-6.22%	8,478
Calhoun	20,623	48.18%	16,453	38.44%	4,717	11.02%	1,007	2.35%	4,170	9.74%	42,800
Chambers	5,682	43.40%	5,938	45.36%	1,427	10.90%	44	0.34%	-256	-1.96%	13,091
Cherokee	2,745	34.85%	4,222	53.61%	846	10.74%	63	0.80%	-1,477	-18.75%	7,876
Chilton	8,126	56.17%	4,946	34.19%	1,363	9.42%	33	0.23%	3,180	21.98%	14,468
Choctaw	3,069	40.60%	3,941	52.13%	489	6.47%	61	0.81%	-872	-11.53%	7,560
Clarke	5,495	46.90%	5,023	42.87%	872	7.44%	327	2.79%	472	4.03%	11,717

1992 United States presidential election in Alabama

← 1988 November 3, 1992 1996 →

Nominee	**George H. W. Bush**	Bill Clinton	Ross Perot
Party	Republican	Democratic	Independent
Home state	Texas	Arkansas	Texas
Running mate	**Dan Quayle**	Al Gore	James Stockdale
Electoral vote	**9**	0	0
Popular vote	**804,283**	690,080	183,109
Percentage	**47.65%**	40.88%	10.85%

County Results

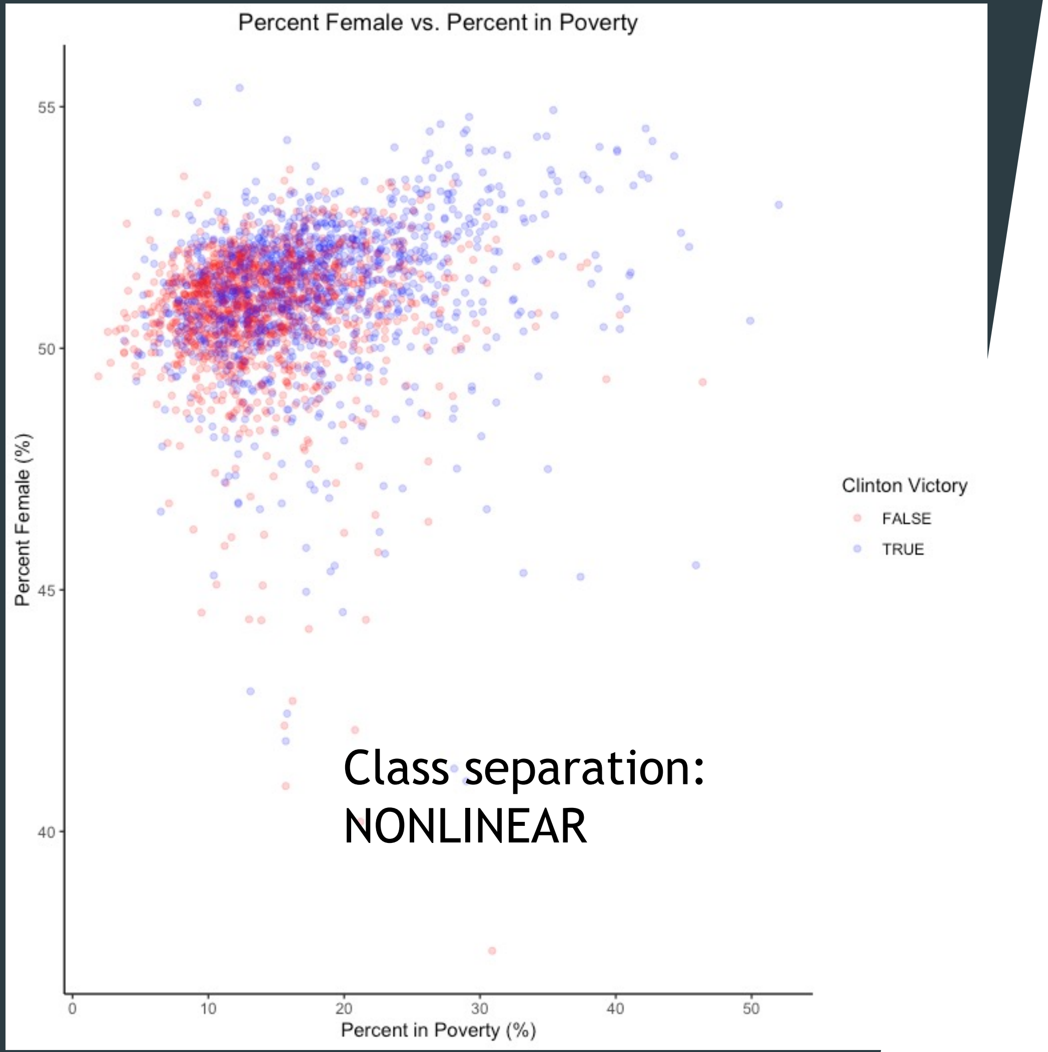
Bush	Clinton
40–50%	40–50%
50–60%	50–60%
60–70%	60–70%
	70–80%
	80–90%

President before election	Elected President
George H. W. Bush Republican	Bill Clinton Democratic

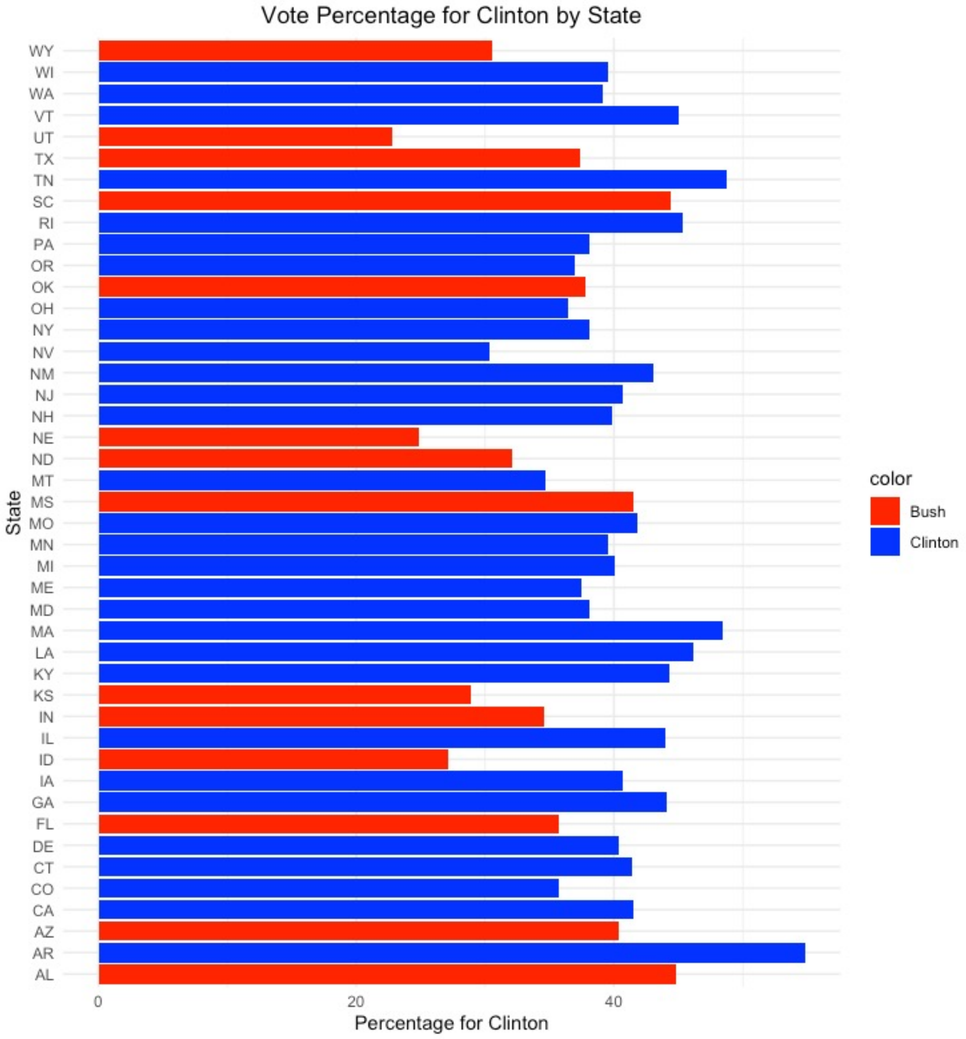
Elections in Alabama

An example of a Wikipedia table that was scraped

Exploratory Data Analysis (EDA)



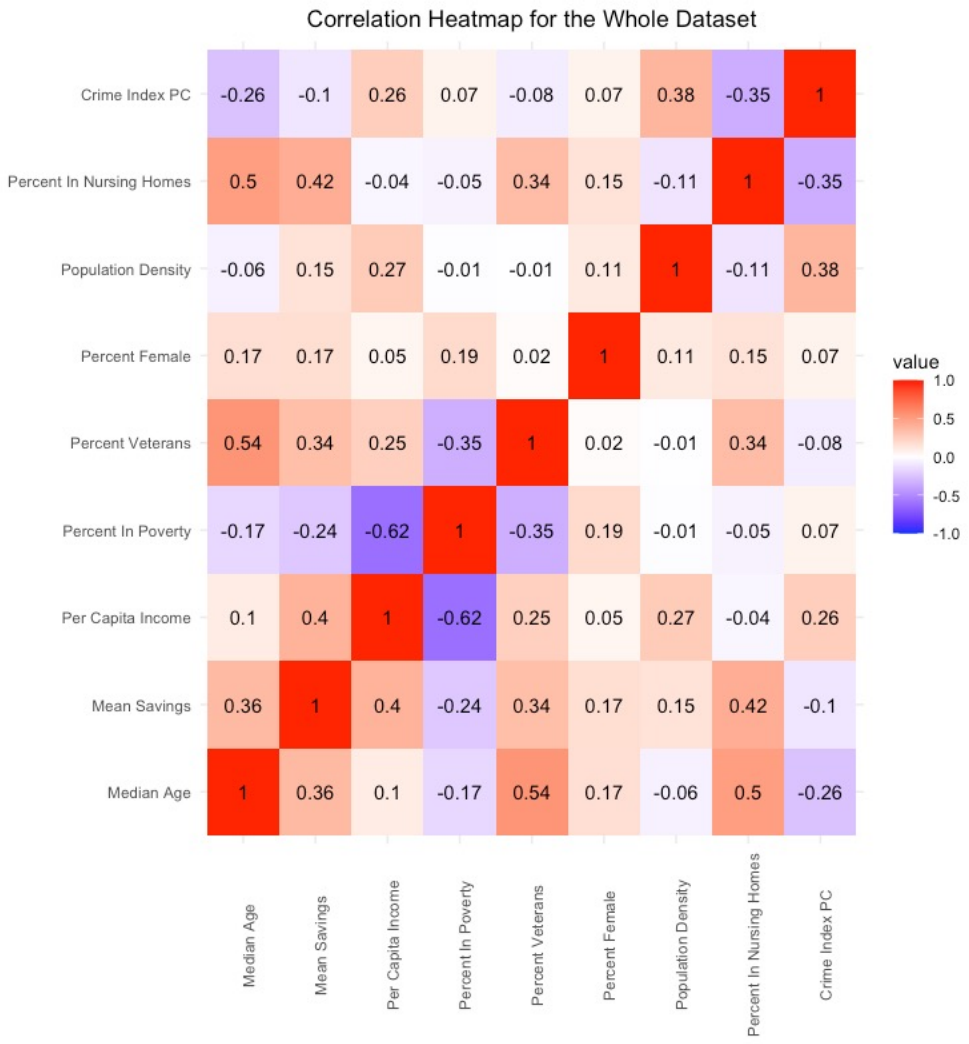
Class separation: NONLINEAR



More EDA

- Moderately strong correlation between variables
- PCA might be appropriate

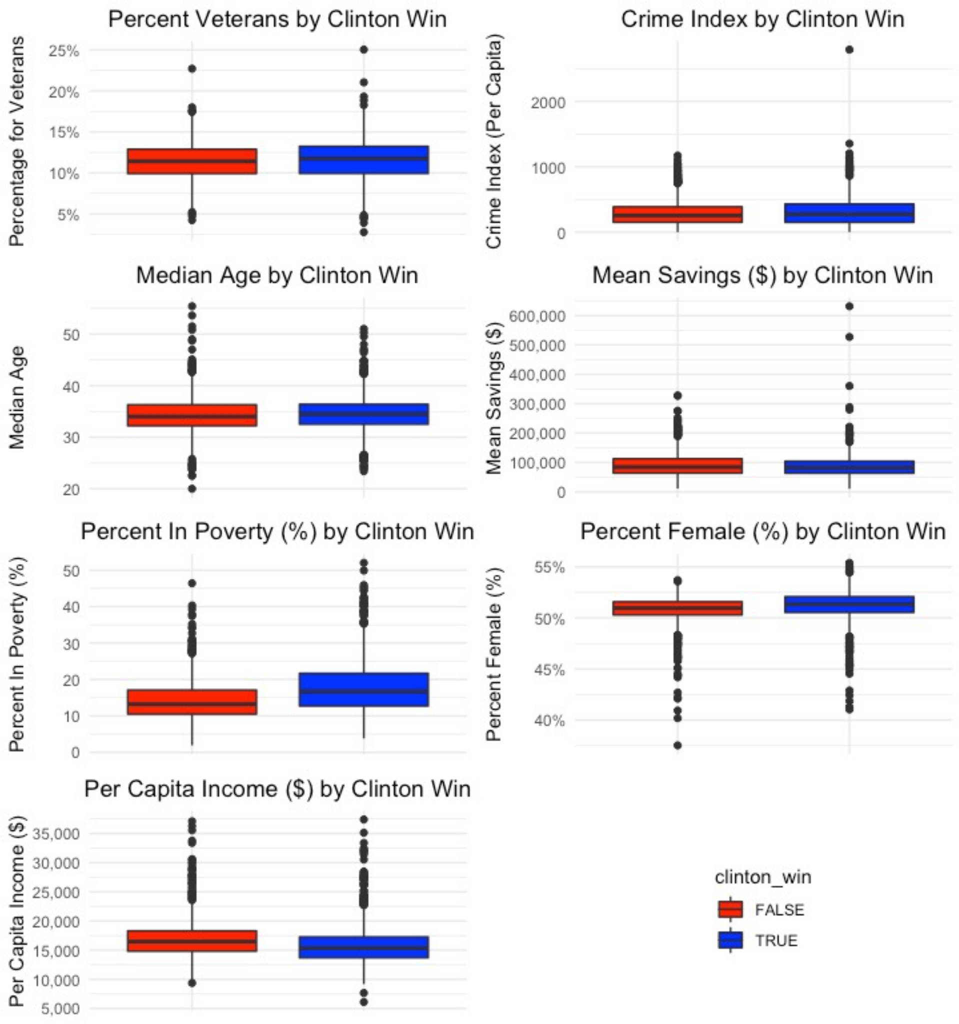
Correlation Heatmap for the Whole Dataset



	Median Age	Mean Savings	Per Capita Income	Percent In Poverty	Percent Veterans	Percent Female	Population Density	Percent In Nursing Homes	Crime Index PC
Crime Index PC	-0.26	-0.1	0.26	0.07	-0.08	0.07	0.38	-0.35	1
Percent In Nursing Homes	0.5	0.42	-0.04	-0.05	0.34	0.15	-0.11	1	-0.35
Population Density	-0.06	0.15	0.27	-0.01	-0.01	0.11	1	-0.11	0.38
Percent Female	0.17	0.17	0.05	0.19	0.02	1	0.11	0.15	0.07
Percent Veterans	0.54	0.34	0.25	-0.35	1	0.02	-0.01	0.34	-0.08
Percent In Poverty	-0.17	-0.24	-0.62	1	-0.35	0.19	-0.01	-0.05	0.07
Per Capita Income	0.1	0.4	1	-0.62	0.25	0.05	0.27	-0.04	0.26
Mean Savings	0.36	1	0.4	-0.24	0.34	0.17	0.15	0.42	-0.1
Median Age	1	0.36	0.1	-0.17	0.54	0.17	-0.06	0.5	-0.26

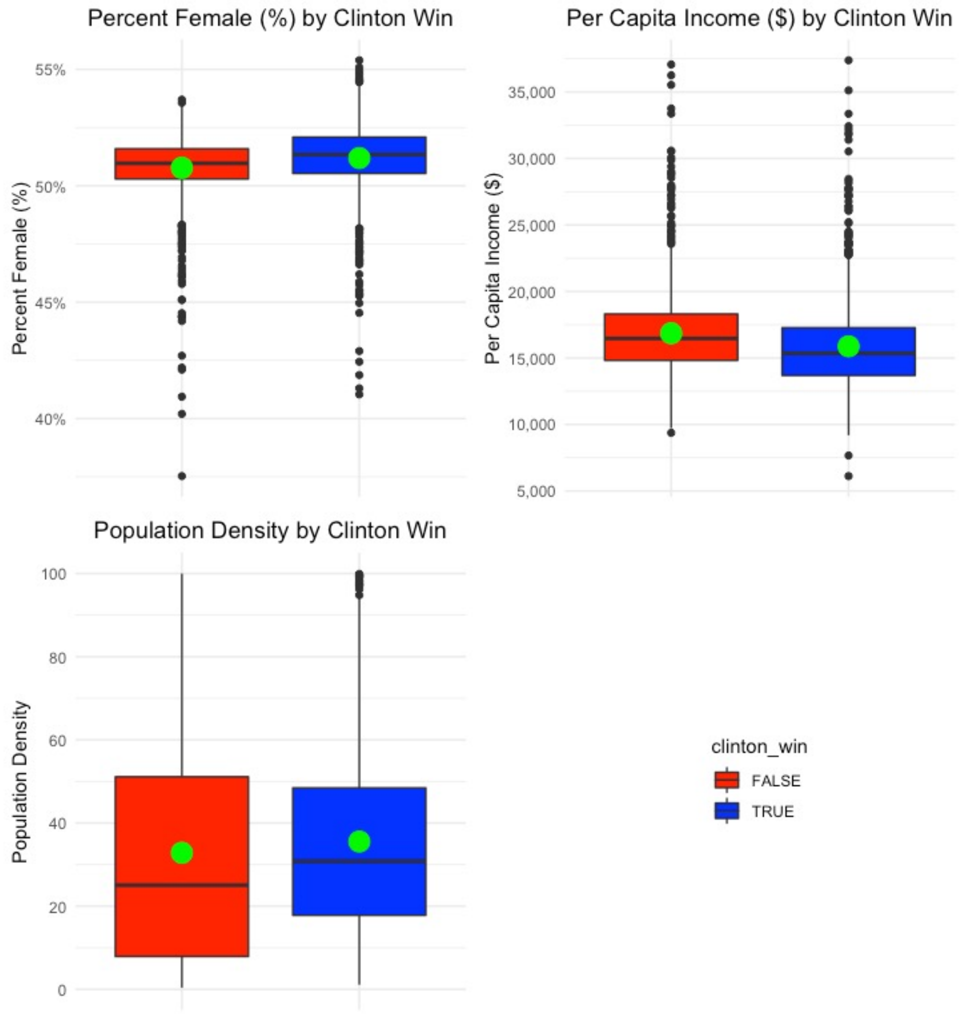
More EDA

- Appearance: small difference between different census measure
- Large sample size -> more power



Question 1: MANOVA

- Use Wilk's test
- F test statistic: 11.493
- *P*-value < 0.00001
- Important variables (ANOVA)
  - Mean Savings: *p* < 0.001
  - PC Income: *p* < 0.0001
  - % Female: *p* < 0.0001
  - Population Dens.: *p* < *0.005*
  - % in poverty: p < 0.0001



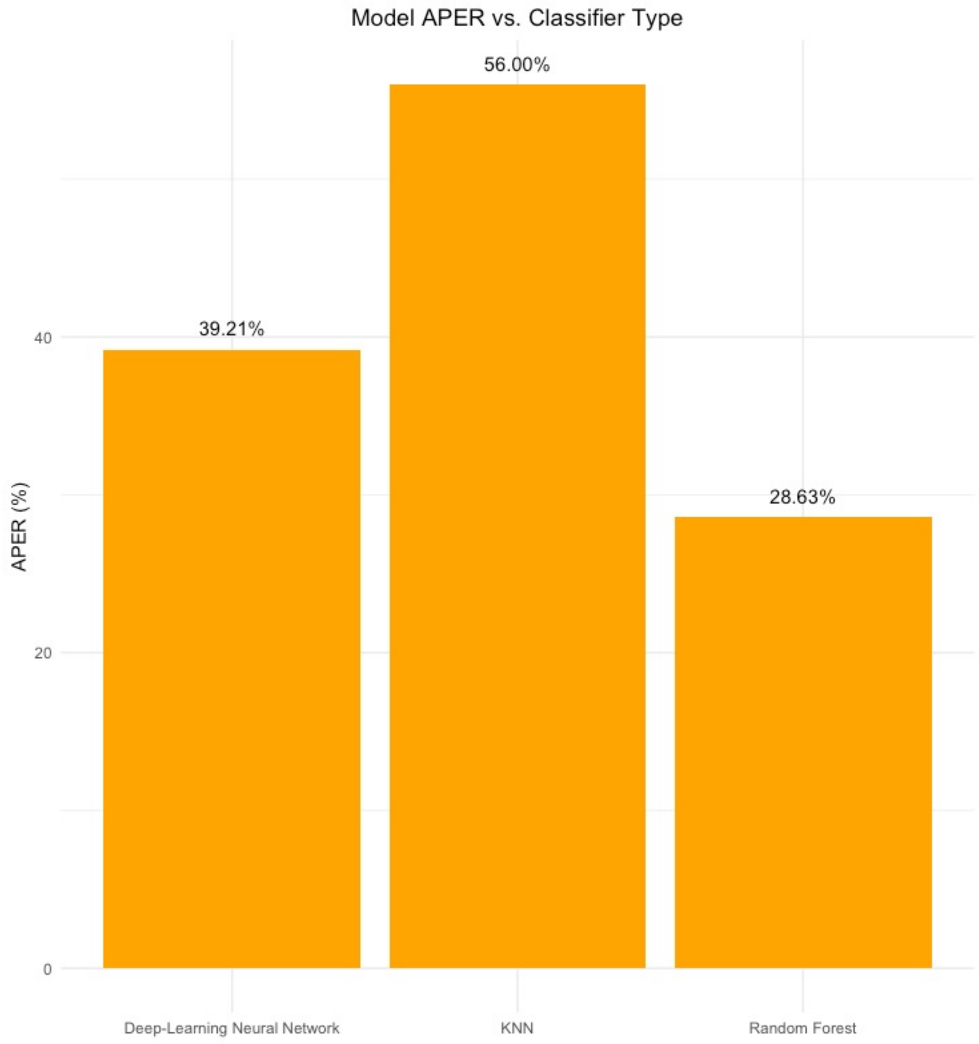
Question 2: Predictive Modeling

- Three Models
  - KNN, Random Forest, Dense Deep-Learning Neural Network (DLNN)
  - Train test split: 80%/20%
- KNN
  - K=5
- Random Forest
  - # of Trees: 150
  - variables at each split : 1

- DLNN (All were similarly bad)
  - # of hidden layers: 5
  - # of hidden units:
    - 32
    - 64
    - 128
    - 64
    - 32
  - Dropout rate: 10%
  - Activation function: Sigmoid
  - Loss function = “binary cross entropy”
  - Optimizer = “RMSprop”

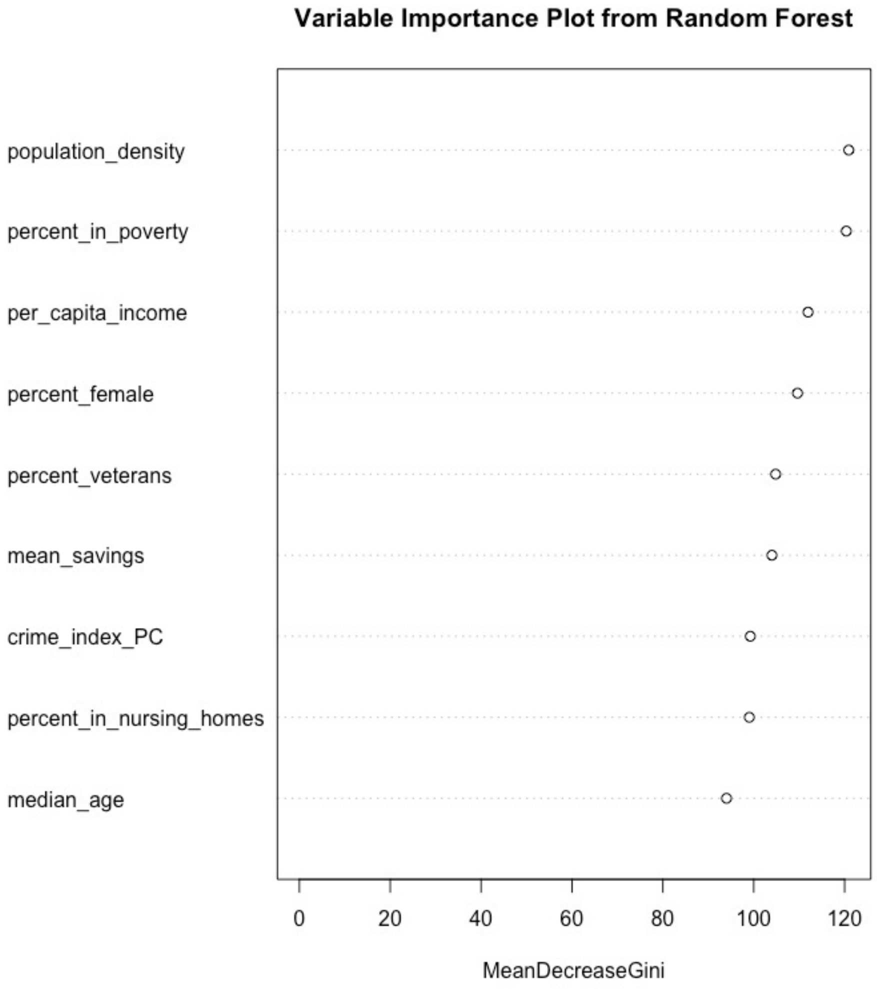
Question 2: Predictive Modeling

- Random forest – highest performer
- Neural Network
  - Not enough data (inconclusive)
- There are more ML models
  - I chose just three
- I should've considered statistical models
  - ML offers no advantage sometimes
- We're ignoring the temporal and cyclical nature of election cycles – this is the future of this study (Time series models, RNNs)



Question 3: Variable Importance

- Random forest – highest performer: Population Density
- Random forest importance is *the same as* ANOVA results



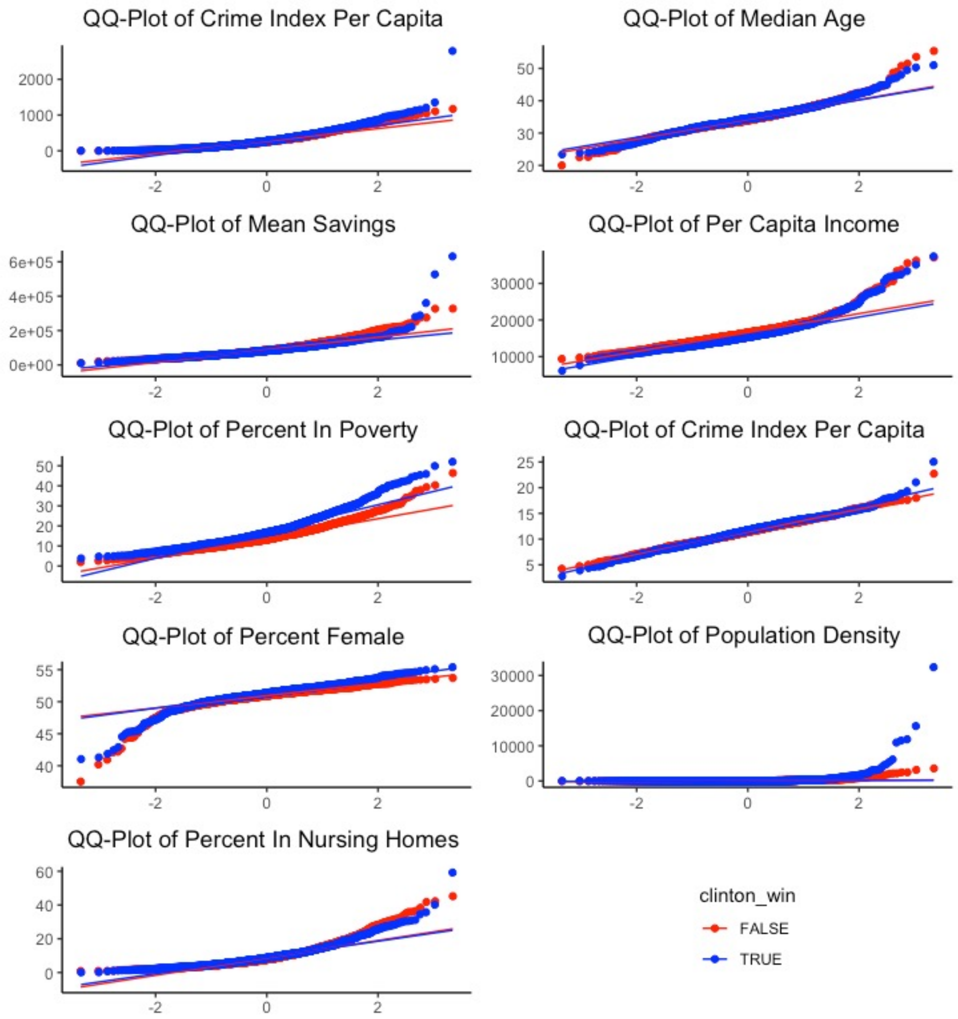
Conclusion

- There are differences in census measurement that voted for democratic vs. republican
- Modeling elections is difficult, but possible
  - Current baseline: random forest
  - Future: Time-series model, RNN
- Population density, % female, PC income, % in poverty – same as individual ANOVA results

- Future Work:
  - Add interaction effects between important variables
  - Time-series models
  - Optimize Parameters (grid search)
  - Use cross-validation

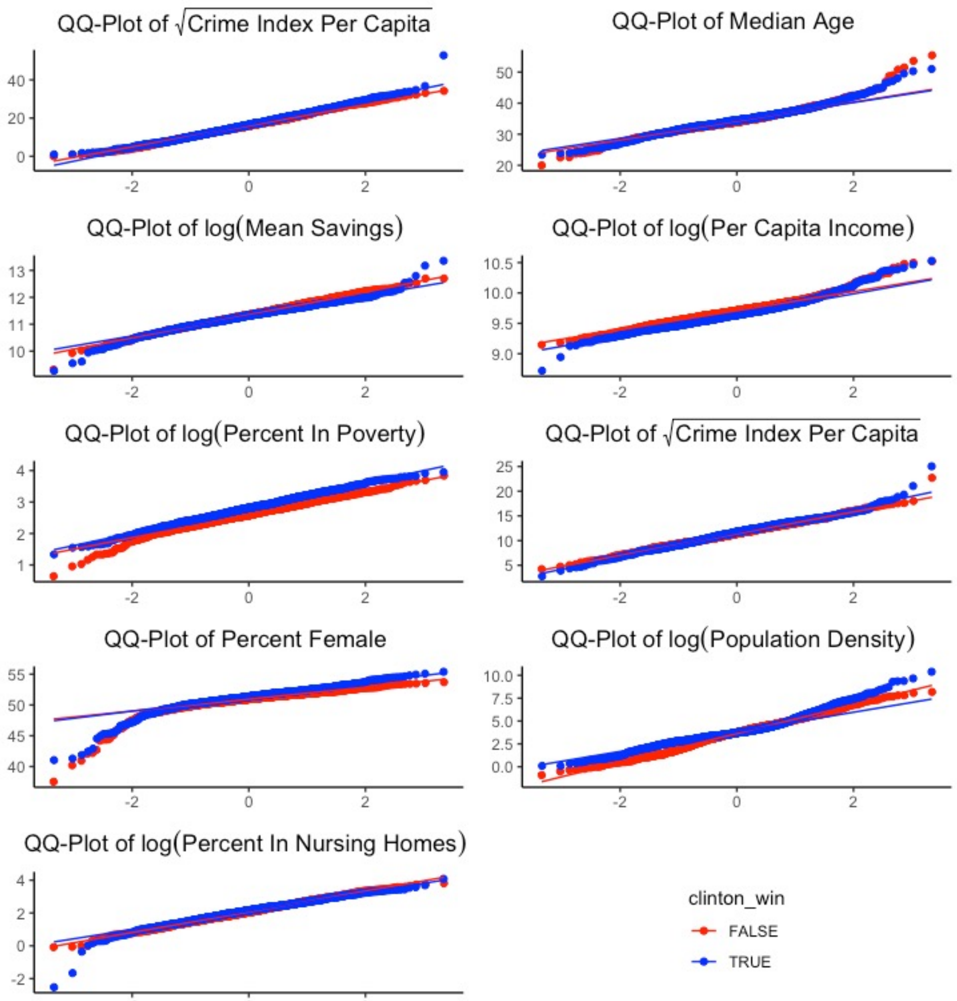
Additional Graphs

- Grouped raw QQ - plots



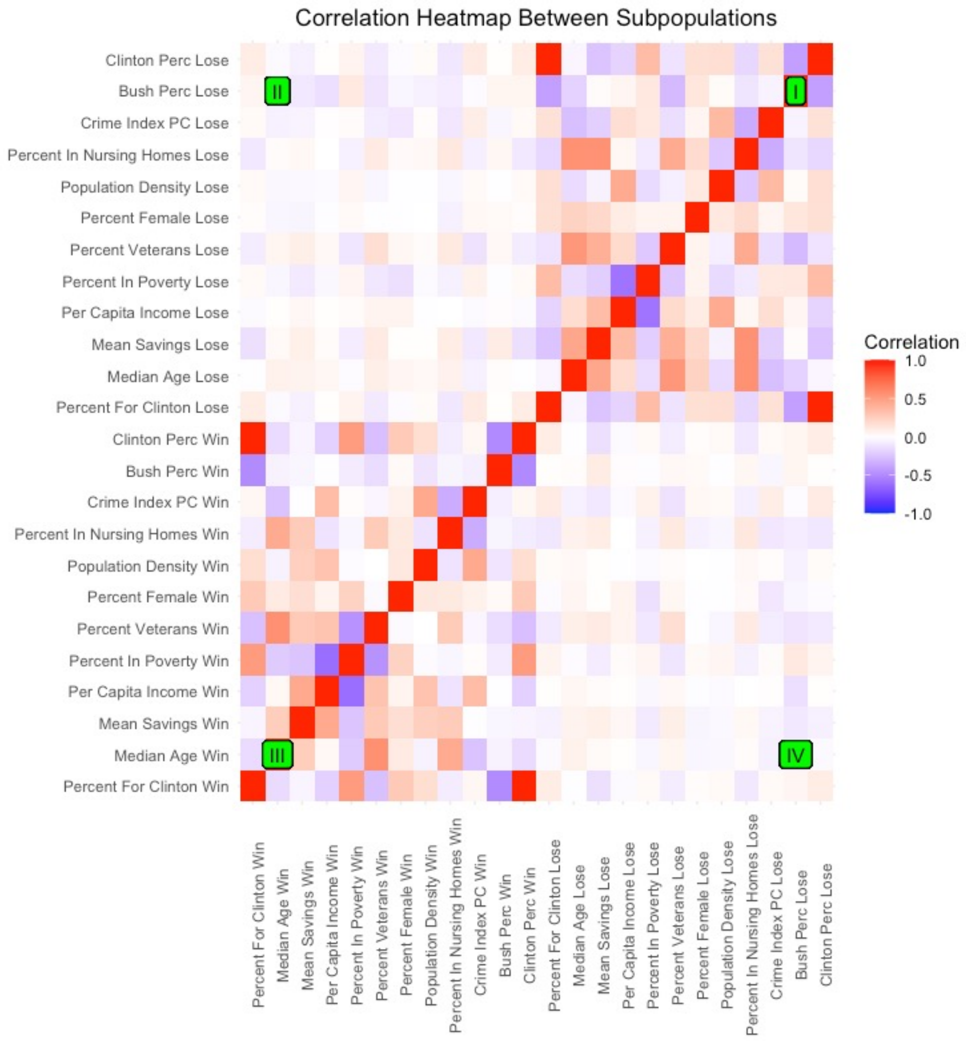
Additional Graphs

- Grouped transformed QQ - plots



Additional Graphs

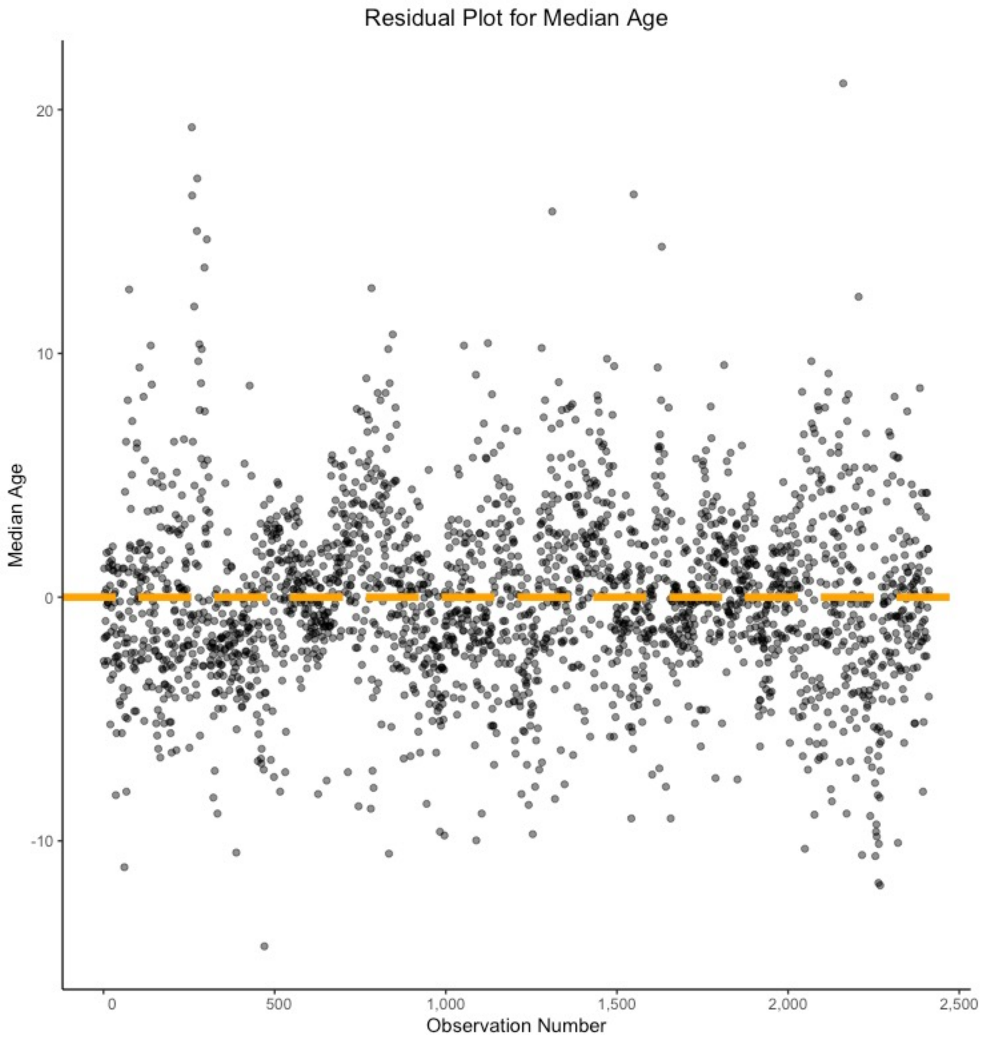
- Subpopulation heatmap



Correlation Heatmap Between Subpopulations

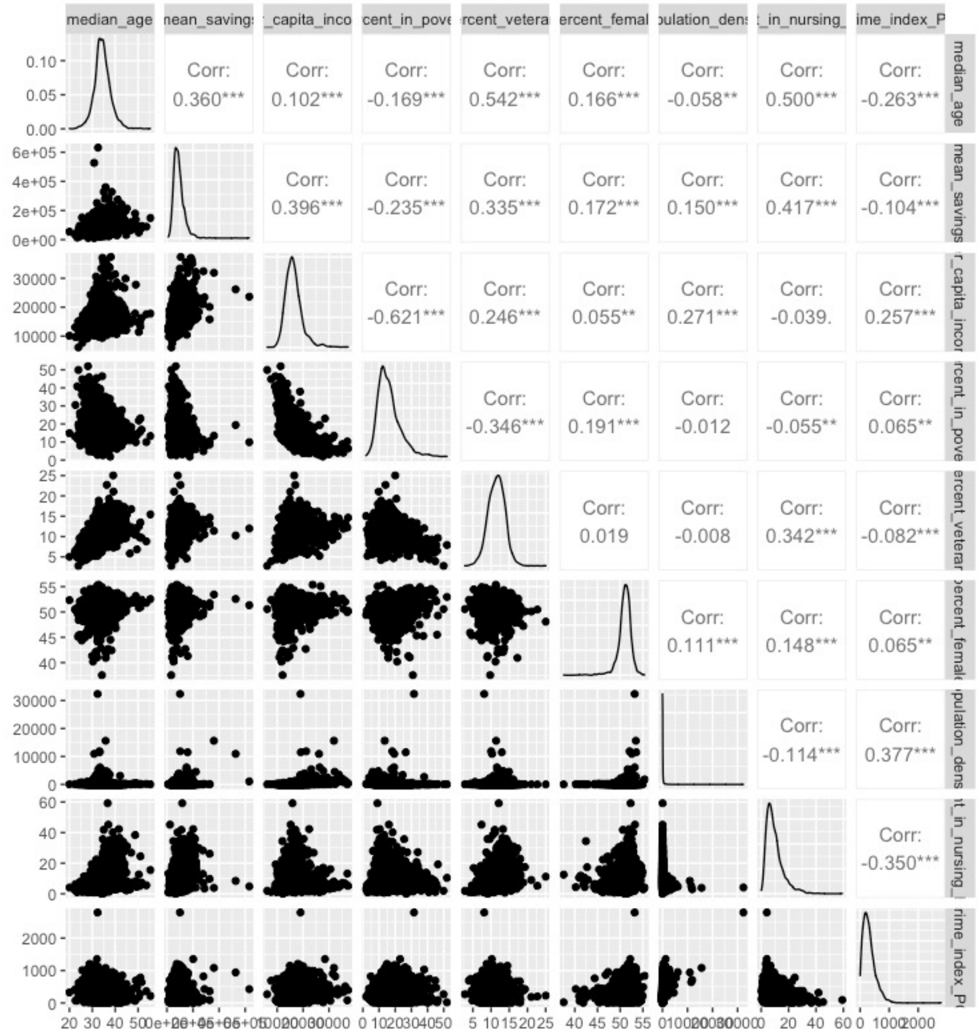
Additional Graphs

- Residual Plot - MANOVA



Additional Graphs

- Scatterplot matrix



My R Package VIP List

- Rvest
- Dplyr
- Stringr
- ggplot2
- Keras