

Case Study 2

Emily Gentles (Writer)

Weiye Liu (Coordinator and Checker)
Qinzhe Wang (Coder)

Jack McCarthy (Presenter)

10/28/2021

Introduction

Voter turnout by a number of different demographics such as age, race and party is a huge topic of interest for both pollsters and candidates, especially around presidential elections and with particular focus on swing states, such as North Carolina. In this case study we will use North Carolina registration and turnout data, by county, for the 2020 election year. We wish to investigate what demographic factors affect turnout and whether turnout differs by county.

Data Cleaning

Note that we renamed the history stats 2020-11-03 data as votes and the voterstats 2020-11-03 data as registers. Within the votes and registers data we aggregated the total voters variable by taking the sum, ignoring NA values, and then left joined votes to registers. We next replaced any NA values in total votes with 0. While these values are unlikely to be exactly 0, we chose to do this in order to preserve the information that was present as the alternative would be to throw the entire row away. After this we dropped all of the rows containing NA values in any of the other columns. Since total voters is one of the main variables we care about it made sense to preserve as many rows as possible but if we replaced every NA value in the data with 0 that may potentially introduce a lot of error into the data, as the values are unlikely to be exactly 0, so we instead dropped all of the rows containing NA values, leaving us with 51,883 rows. We next randomly sampled 30 counties to utilize in our analysis, leaving our final data with 16,439 rows and, after deaggregation, 1,908,907 individuals.

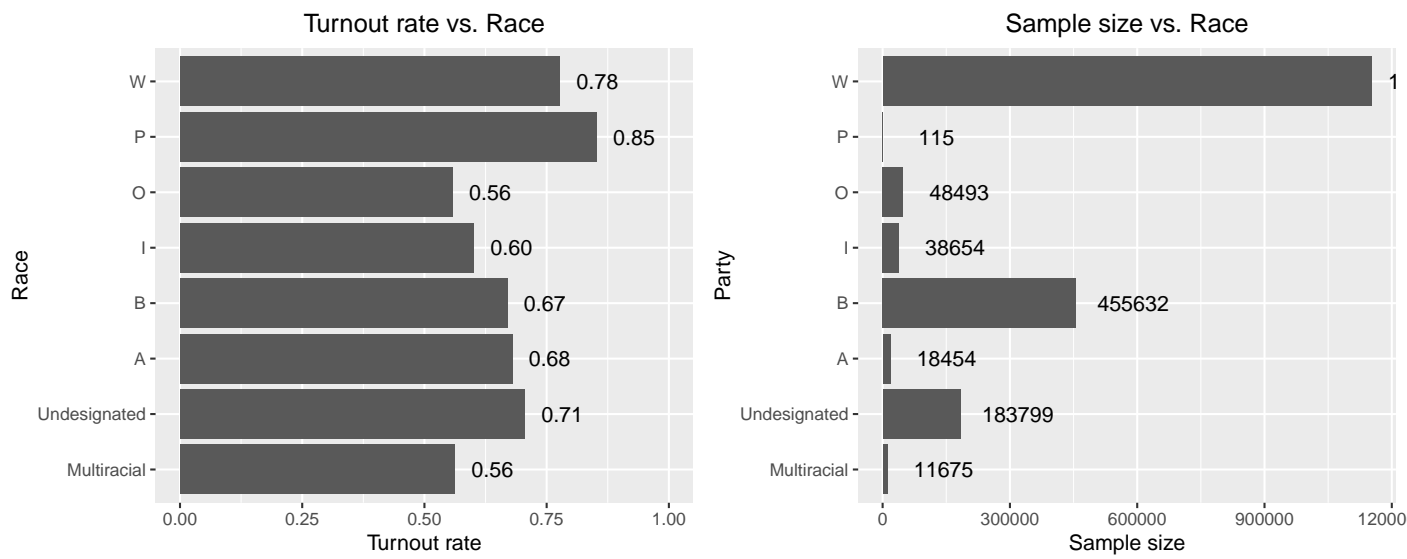
EDA

```
##
##           0           1
## 509502 1399405

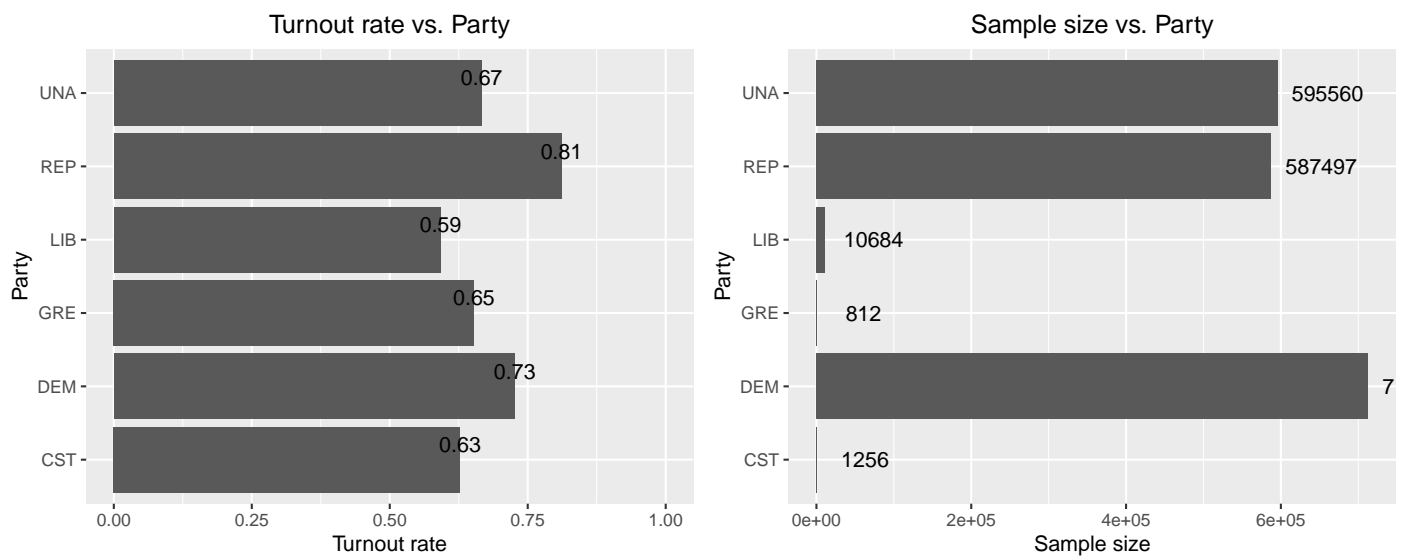
##   group turnout_rate
## 1 Total      0.7330923
```

turnout rate

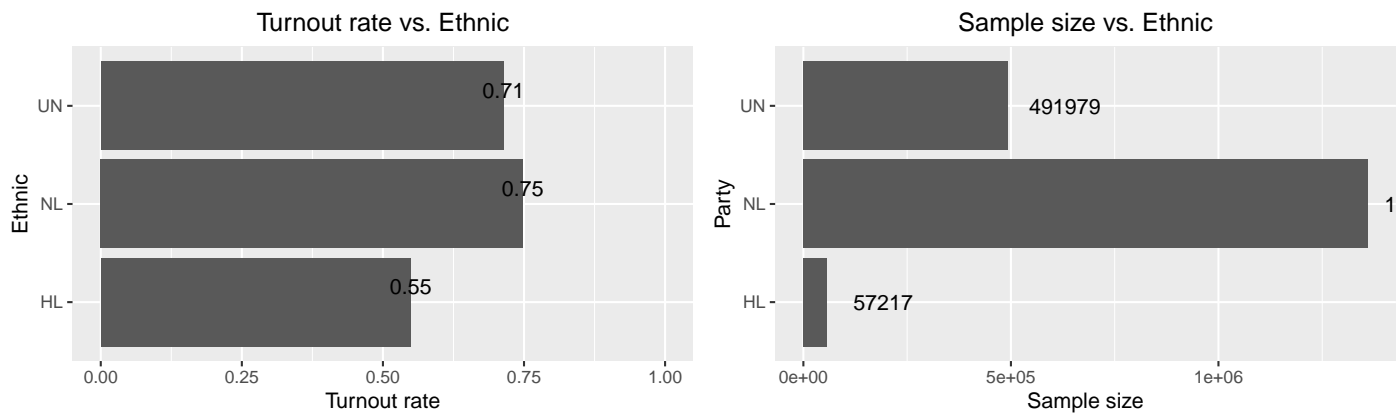
race



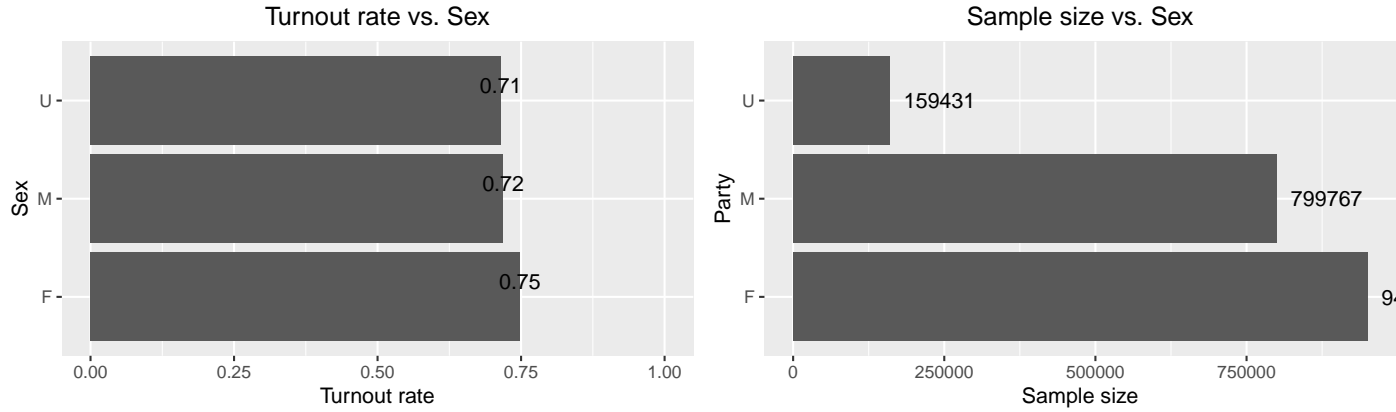
party



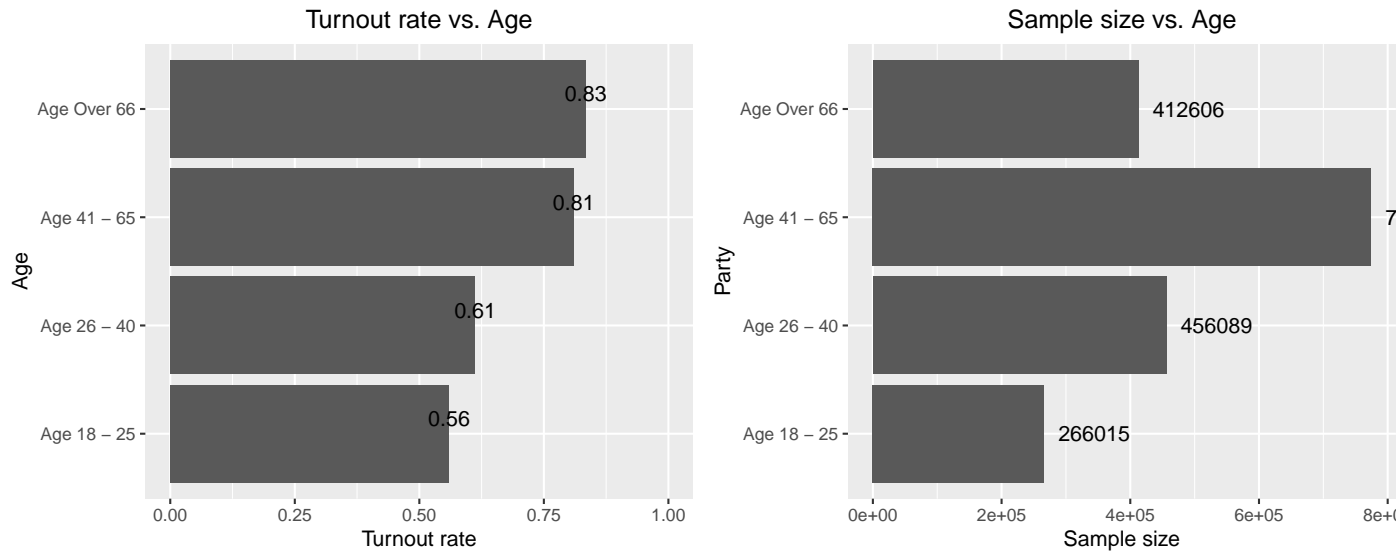
ethnic groups



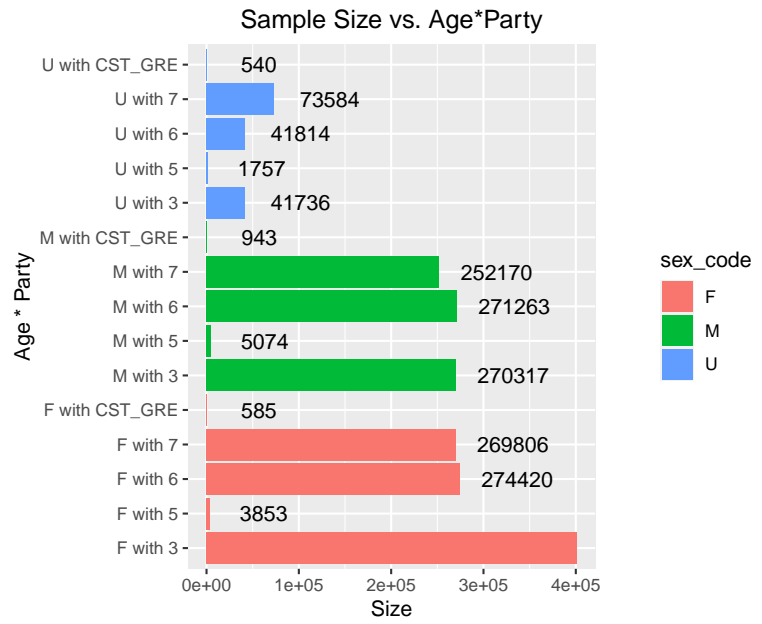
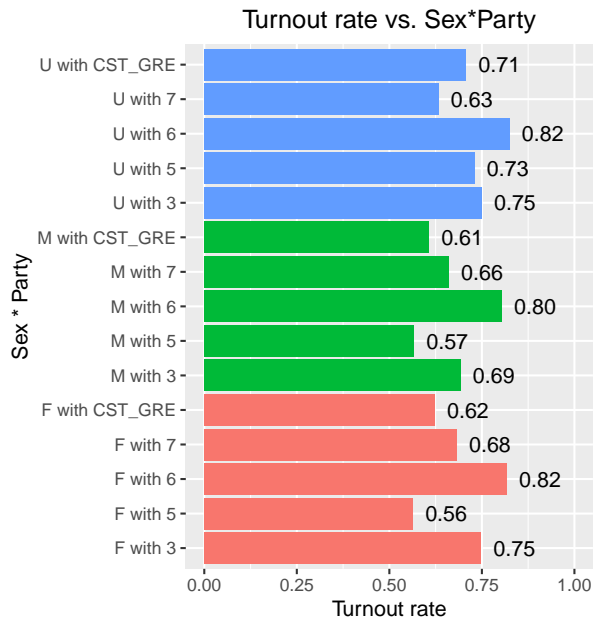
Sex



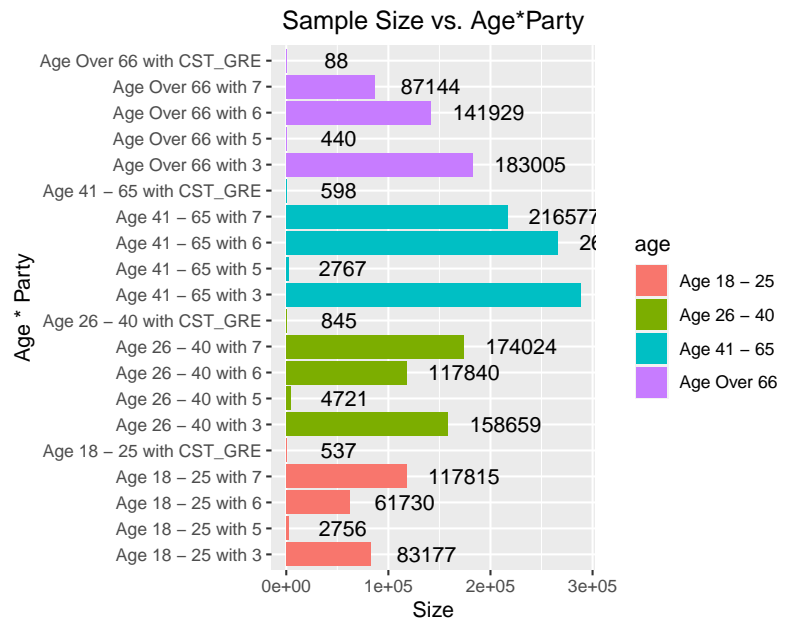
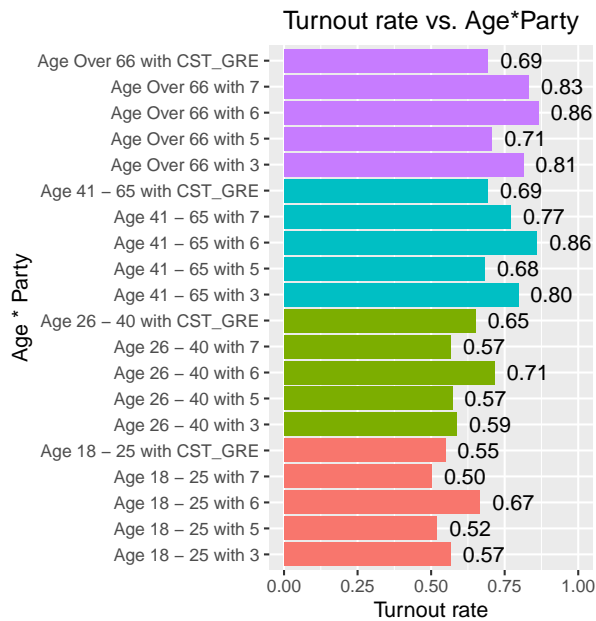
age



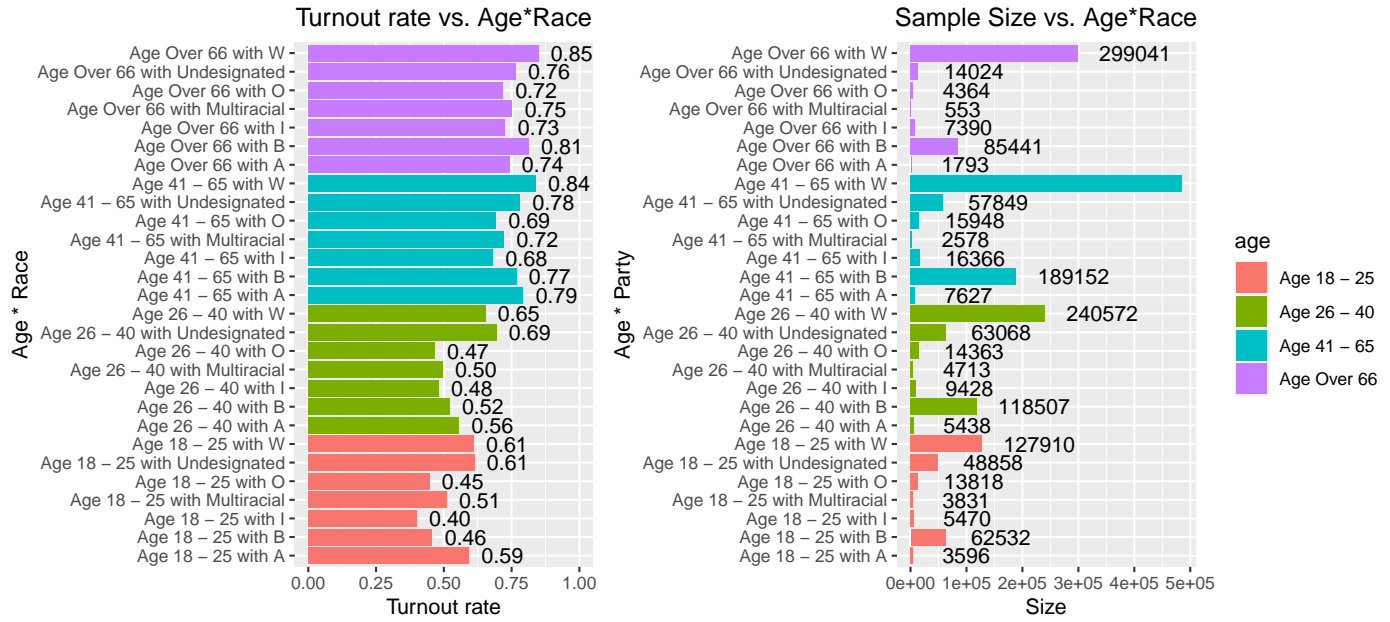
sex & party



age & party



age & race



Model

Table 1: Forward model selection

Model	LRT.p.value	AIC	BIC
Base model		96904.01	96988.73
Add race	0	88191.94	88322.86
Add race and ethnic	0	86526.87	86673.20
Without intercept	0	99839.91	99955.43
Add the interaction of sex and party_cd, and age and party	0	83825.28	84125.65
Add the interaction of sex and party_cd, age and party, and age and race	0	79688.45	80127.45

##	Estimate	Std. Error	z value
## (Intercept)	0.06229984	0.046545701	1.33846603
## party_cd5	-0.63633686	0.047331626	-13.44422148
## party_cd6	-0.02297380	0.013754022	-1.67033287
## party_cd7	-0.52686363	0.011060013	-47.63680007
## party_cdCST_GRE	-0.34773766	0.112648109	-3.08693739
## race_codeUndesignated	0.35843539	0.034981706	10.24636677
## race_codeA	0.15171144	0.047742968	3.17767103
## race_codeB	-0.34413455	0.034128646	-10.08345146
## race_codeI	-0.24086614	0.043653043	-5.51774010
## race_codeO	-0.10978982	0.037422701	-2.93377603
## race_codeW	0.23184339	0.033694463	6.88075650
## ethnic_codeNL	0.44375223	0.010691808	41.50394615
## ethnic_codeUN	0.36184997	0.010948056	33.05152740
## sex_codeM	-0.28136933	0.005776245	-48.71145670
## sex_codeU	0.10247015	0.014214942	7.20862225
## ageAge 26 - 40	-0.18986654	0.044805737	-4.23754980

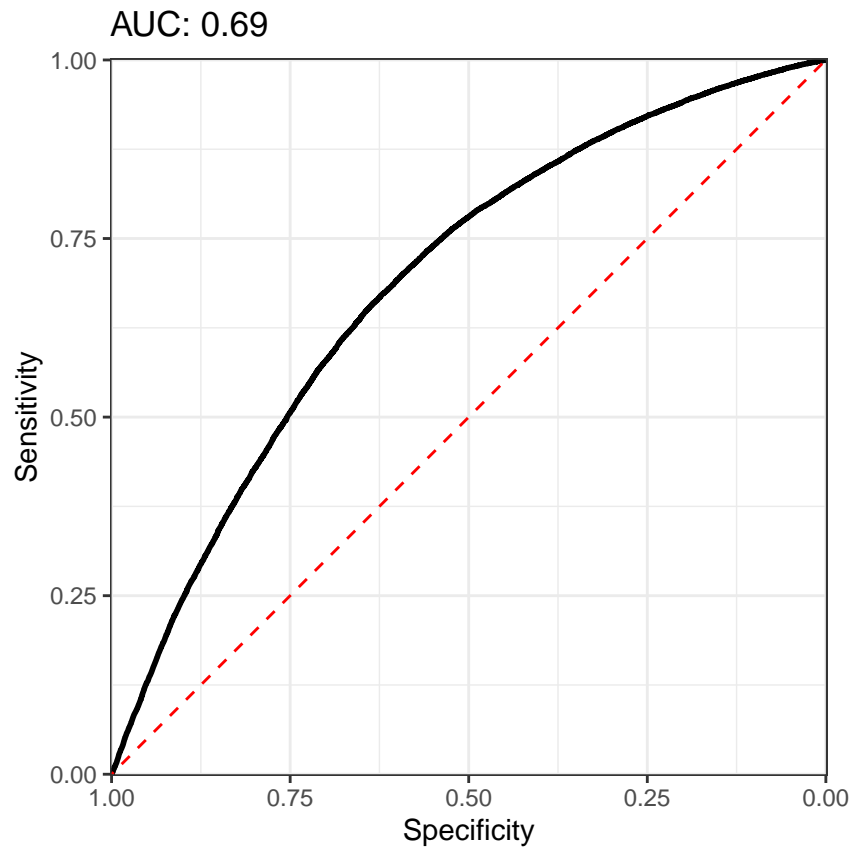
## ageAge 41 - 65	0.80335846	0.055851305	14.38387967
## ageAge Over 66	0.75255948	0.105028495	7.16528864
## party_cd5:sex_codeM	0.23876425	0.044657959	5.34651057
## party_cd6:sex_codeM	0.23127356	0.009157990	25.25374697
## party_cd7:sex_codeM	0.16840260	0.008492820	19.82881910
## party_cdCST_GRE:sex_codeM	0.20854818	0.111252289	1.87455178
## party_cd5:sex_codeU	0.55193499	0.064992570	8.49227840
## party_cd6:sex_codeU	0.23002653	0.018796275	12.23787816
## party_cd7:sex_codeU	-0.17332351	0.015311892	-11.31953601
## party_cdCST_GRE:sex_codeU	0.07655570	0.130391411	0.58712223
## party_cd5:ageAge 26 - 40	0.11234865	0.050613692	2.21972841
## party_cd6:ageAge 26 - 40	0.17561057	0.015890632	11.05120100
## party_cd7:ageAge 26 - 40	0.18349987	0.012808716	14.32617180
## party_cdCST_GRE:ageAge 26 - 40	0.23906764	0.115994189	2.06103115
## party_cd5:ageAge 41 - 65	-0.31268181	0.058077696	-5.38385354
## party_cd6:ageAge 41 - 65	0.08558492	0.015437770	5.54386547
## party_cd7:ageAge 41 - 65	0.14003954	0.012832337	10.91301944
## party_cdCST_GRE:ageAge 41 - 65	-0.41624836	0.126672120	-3.28602980
## party_cd5:ageAge Over 66	-0.08716228	0.113437642	-0.76837176
## party_cd6:ageAge Over 66	0.21850850	0.016957513	12.88564593
## party_cd7:ageAge Over 66	0.58023791	0.015610329	37.17012615
## party_cdCST_GRE:ageAge Over 66	-0.29168492	0.250904123	-1.16253539
## race_codeUndesignated:ageAge 26 - 40	0.33270792	0.046164601	7.20699210
## race_codeA:ageAge 26 - 40	-0.05830307	0.062497997	-0.93287905
## race_codeB:ageAge 26 - 40	0.36029253	0.045479664	7.92205790
## race_codeI:ageAge 26 - 40	0.37342195	0.056328180	6.62939842
## race_codeO:ageAge 26 - 40	0.13930692	0.050506162	2.75821629
## race_codeW:ageAge 26 - 40	0.21759883	0.045035200	4.83175017
## race_codeUndesignated:ageAge 41 - 65	-0.19184163	0.057169943	-3.35563790
## race_codeA:ageAge 41 - 65	0.11402344	0.071171349	1.60209743
## race_codeB:ageAge 41 - 65	0.47074092	0.056374598	8.35023114
## race_codeI:ageAge 41 - 65	0.22724283	0.064321149	3.53294105
## race_codeO:ageAge 41 - 65	0.06721625	0.060632844	1.10857829
## race_codeW:ageAge 41 - 65	0.24805032	0.055978775	4.43114953
## race_codeUndesignated:ageAge Over 66	-0.46652769	0.107126565	-4.35492069
## race_codeA:ageAge Over 66	-0.33419230	0.122941340	-2.71830695
## race_codeB:ageAge Over 66	0.73290817	0.105543805	6.94411352
## race_codeI:ageAge Over 66	0.38959349	0.111587103	3.49138455
## race_codeO:ageAge Over 66	-0.00542286	0.111534850	-0.04862032
## race_codeW:ageAge Over 66	0.17763266	0.105116278	1.68986828
##	Pr(> z)		
## (Intercept)	1.807446e-01		
## party_cd5	3.328908e-41		
## party_cd6	9.485352e-02		
## party_cd7	0.000000e+00		
## party_cdCST_GRE	2.022302e-03		
## race_codeUndesignated	1.228758e-24		
## race_codeA	1.484631e-03		
## race_codeB	6.538804e-24		
## race_codeI	3.433868e-08		
## race_codeO	3.348658e-03		
## race_codeW	5.953554e-12		

```

## ethnic_codeNL          0.000000e+00
## ethnic_codeUN          1.478951e-239
## sex_codeM              0.000000e+00
## sex_codeU              5.652081e-13
## ageAge 26 - 40         2.259723e-05
## ageAge 41 - 65         6.532740e-47
## ageAge Over 66         7.762283e-13
## party_cd5:sex_codeM    8.966605e-08
## party_cd6:sex_codeM    1.030163e-140
## party_cd7:sex_codeM    1.679363e-87
## party_cdCST_GRE:sex_codeM 6.085441e-02
## party_cd5:sex_codeU    2.026249e-17
## party_cd6:sex_codeU    1.950656e-34
## party_cd7:sex_codeU    1.050263e-29
## party_cdCST_GRE:sex_codeU 5.571216e-01
## party_cd5:ageAge 26 - 40 2.643721e-02
## party_cd6:ageAge 26 - 40 2.163016e-28
## party_cd7:ageAge 26 - 40 1.501735e-46
## party_cdCST_GRE:ageAge 26 - 40 3.930007e-02
## party_cd5:ageAge 41 - 65 7.290784e-08
## party_cd6:ageAge 41 - 65 2.958662e-08
## party_cd7:ageAge 41 - 65 9.988263e-28
## party_cdCST_GRE:ageAge 41 - 65 1.016103e-03
## party_cd5:ageAge Over 66 4.422664e-01
## party_cd6:ageAge Over 66 5.421964e-38
## party_cd7:ageAge Over 66 2.074126e-302
## party_cdCST_GRE:ageAge Over 66 2.450181e-01
## race_codeUndesignated:ageAge 26 - 40 5.720132e-13
## race_codeA:ageAge 26 - 40 3.508824e-01
## race_codeB:ageAge 26 - 40 2.336115e-15
## race_codeI:ageAge 26 - 40 3.370577e-11
## race_codeO:ageAge 26 - 40 5.811773e-03
## race_codeW:ageAge 26 - 40 1.353380e-06
## race_codeUndesignated:ageAge 41 - 65 7.918217e-04
## race_codeA:ageAge 41 - 65 1.091341e-01
## race_codeB:ageAge 41 - 65 6.812951e-17
## race_codeI:ageAge 41 - 65 4.109641e-04
## race_codeO:ageAge 41 - 65 2.676121e-01
## race_codeW:ageAge 41 - 65 9.373206e-06
## race_codeUndesignated:ageAge Over 66 1.331151e-05
## race_codeA:ageAge Over 66 6.561693e-03
## race_codeB:ageAge Over 66 3.808442e-12
## race_codeI:ageAge Over 66 4.805241e-04
## race_codeO:ageAge Over 66 9.612219e-01
## race_codeW:ageAge Over 66 9.105316e-02

```

	Estimate	95% CI
(Intercept)	1.173	[1.098, 1.248]
party_cd5	-0.888	[-0.952, -0.823]
party_cd6	0.314	[0.302, 0.327]
party_cd7	-0.384	[-0.395, -0.373]
party_cdCST_GRE	-0.629	[-0.798, -0.461]
party_cd3:sex_codeM	-0.277	[-0.288, -0.266]
party_cd5:sex_codeM	0.006	[-0.079, 0.091]
party_cd6:sex_codeM	-0.089	[-0.103, -0.076]
party_cd7:sex_codeM	-0.102	[-0.114, -0.09]
party_cdCST_GRE:sex_codeM	-0.048	[-0.262, 0.165]
party_cd3:sex_codeU	-0.021	[-0.045, 0.002]
party_cd5:sex_codeU	0.728	[0.604, 0.852]
party_cd6:sex_codeU	0.048	[0.02, 0.075]
party_cd7:sex_codeU	-0.219	[-0.236, -0.201]
party_cdCST_GRE:sex_codeU	0.338	[0.088, 0.589]



county_desc

