

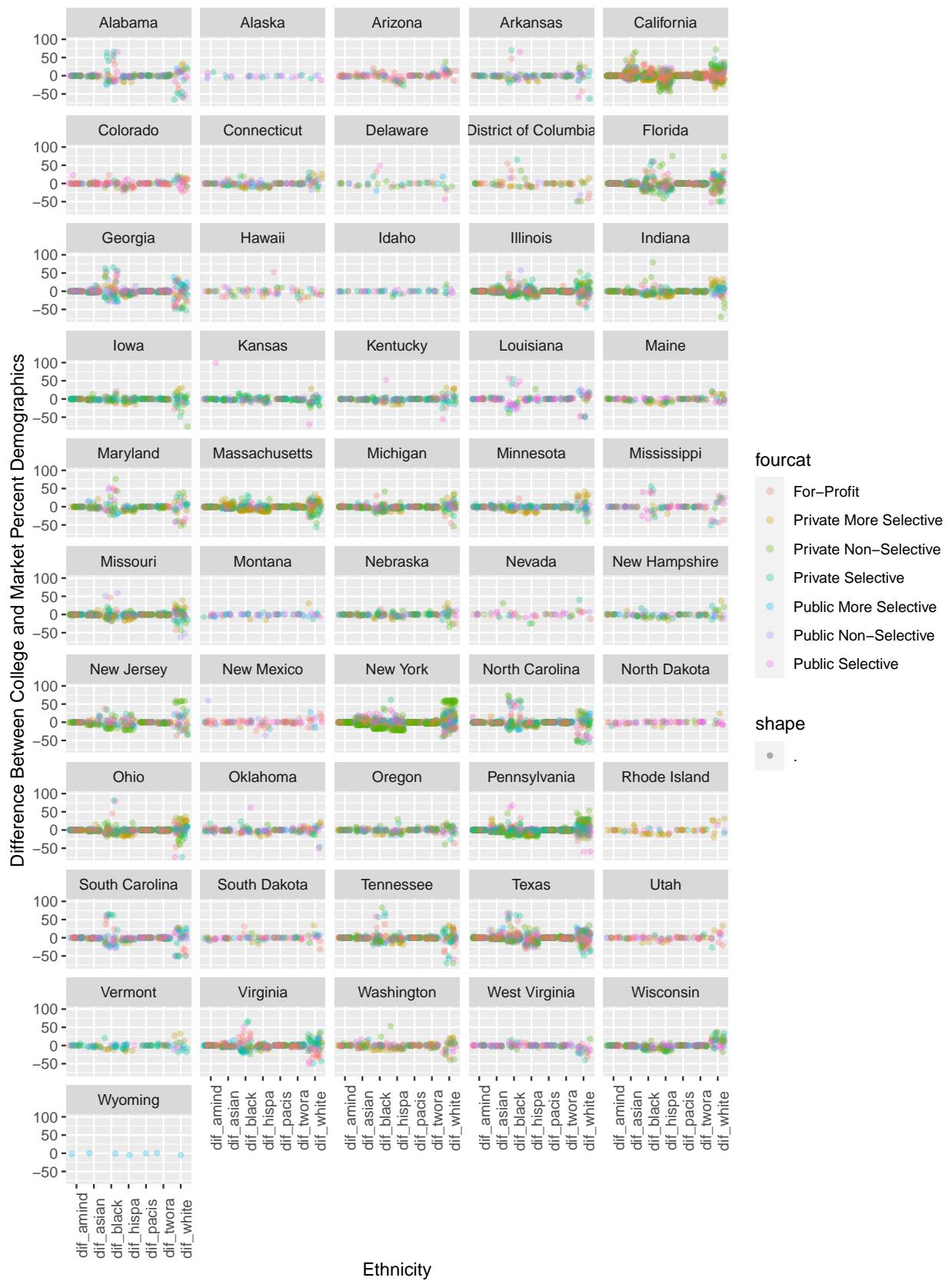
Midterm Project for Four Year College Data

Tobias Boggess, Carl Perry

3/16/2022

The task of data wrangling for the project was primarily defined by the verbs of dplyr. The first step in the project was using filter to subset the data to only include the academic year 2017, as the data set contained several years. After the initial filter subset, select was used to subset the data to include a select range of columns. For our project, we decided to include the columns fips_ipeds, inst_name, fourcat and all columns associated with difference, market, and college for their numeric values for computation. Mutate was used to add a categorical value to the long pivoted version of data to allow for grouping and faceting in graphs. Summarize was used to generate tables based on the subsets of data.

Percent Difference b/w College and Market Vs Ethnicity

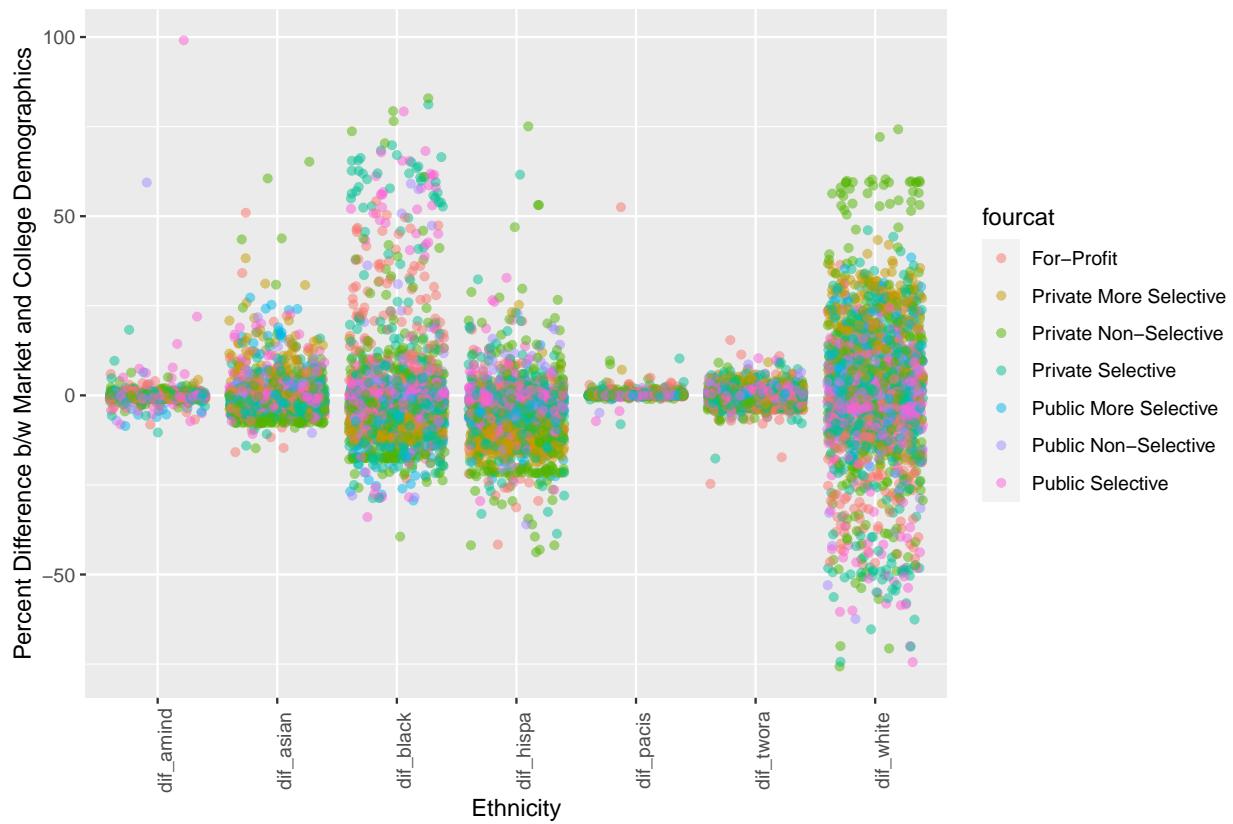


fips_ipeds	mean(dif_white)	mean(dif_hispa)	mean(dif_black)	mean(dif_asian)	mean(dif_amind)	mean(dif_pacis)	mean(dif_twora)
Alabama	-8.005	-3.579	9.360	-0.557	-0.091	0.180	-0.026
Alaska	-4.470	-2.525	-1.600	-2.218	2.053	-0.703	-1.101
Arizona	10.210	-11.936	2.398	0.096	-1.219	0.319	0.014
Arkansas	-1.350	-3.967	3.591	-0.622	-0.023	-0.002	-0.015
California	2.736	-8.624	0.181	4.459	-0.256	0.453	0.106
Colorado	3.369	-7.862	0.282	-0.059	0.573	0.083	0.377
Connecticut	6.744	-6.128	-1.774	-0.700	-0.199	0.005	-0.416
Delaware	-10.720	-6.385	14.199	-3.771	0.118	0.125	0.891
District of Columbia	-20.075	-4.229	16.584	-1.578	-0.258	0.013	-0.651
Florida	-0.034	-4.848	-0.194	-0.637	-0.044	0.144	-1.032
Georgia	-6.525	-4.154	5.730	-0.263	-0.084	0.134	-0.414
Hawaii	-4.683	-3.003	0.065	-4.412	0.100	6.191	-6.400
Idaho	1.424	-4.920	0.322	-0.147	-0.408	0.255	0.520
Illinois	1.017	-1.619	-1.714	0.065	-0.048	0.158	-0.329
Indiana	3.383	-3.863	-0.729	-1.603	-0.088	0.006	-0.483
Iowa	-4.630	-2.758	0.106	-1.229	-0.117	0.133	-0.542
Kansas	-6.938	-1.939	1.438	-1.921	3.320	0.031	-1.476
Kentucky	0.589	-4.049	-1.105	-1.297	-0.085	-0.002	-0.185
Louisiana	-1.898	0.269	-3.957	0.556	0.114	0.022	0.617
Maine	2.913	-5.071	-3.236	-1.535	0.251	0.016	-0.098
Maryland	-4.728	-6.259	7.367	-0.064	-0.181	0.054	-0.189
Massachusetts	0.151	-5.245	-1.839	0.497	-0.303	-0.035	-0.371
Michigan	0.342	-2.745	-4.058	-0.819	-0.045	0.045	-0.901
Minnesota	4.491	-5.358	-3.053	-0.469	-0.294	-0.022	-0.467
Mississippi	-5.320	-2.465	6.169	-0.254	-0.194	0.022	-0.236
Missouri	-1.698	-2.885	-0.241	-0.549	-0.161	-0.011	-0.306
Montana	1.284	-2.461	-0.982	-0.450	-3.019	0.303	0.021
Nebraska	-0.120	-3.155	-0.518	-1.066	-0.322	0.159	0.576
Nevada	9.162	-10.123	-0.980	5.096	0.257	0.574	-1.492
New Hampshire	4.339	-5.778	-1.776	-1.962	0.211	0.056	-1.053
New Jersey	8.390	-4.004	-2.900	-1.063	-0.128	0.101	-1.442
New Mexico	4.658	-8.178	0.455	-0.011	3.867	0.489	-0.411
New York	11.683	-6.448	-5.693	-0.478	-0.190	0.034	-1.115
North Carolina	-7.213	-4.602	9.159	-0.629	-0.122	0.068	-0.311
North Dakota	0.367	-0.817	-0.599	-1.644	-3.200	0.302	0.776
Ohio	-0.074	-3.193	-1.160	-0.935	-0.104	0.009	-0.557
Oklahoma	-1.976	-4.267	1.433	-1.167	0.604	0.058	0.507
Oregon	-0.646	-3.722	-1.419	0.057	-0.138	0.512	0.663
Pennsylvania	6.916	6.605	-2.770	-0.731	-0.148	0.015	-0.305
Rhode Island	9.404	-8.441	-6.388	0.417	-0.466	-0.061	-0.543
South Carolina	-4.074	-4.095	5.147	-1.034	0.047	0.123	-0.261
South Dakota	1.876	-2.560	3.279	-1.095	-3.359	0.108	-0.480
Tennessee	-2.875	-3.773	3.844	-0.220	0.020	0.012	-0.357
Texas	-0.594	-6.146	5.228	0.987	-0.107	0.161	-0.379
Utah	3.048	-5.490	-1.116	-1.009	0.087	0.295	0.532
Vermont	0.470	-5.195	-2.241	-2.635	0.429	-0.017	0.536
Virginia	-2.458	-2.683	0.977	-0.597	0.160	0.229	-0.319
Washington	-3.572	-3.022	-1.644	0.435	0.094	0.257	-0.280
West Virginia	-11.753	0.153	1.880	-0.948	0.364	0.181	-0.130
Wisconsin	11.994	-6.349	-4.880	-0.864	-0.100	0.048	-0.313
Wyoming	-5.087	-4.829	-0.878	0.045	-2.046	-0.364	0.571

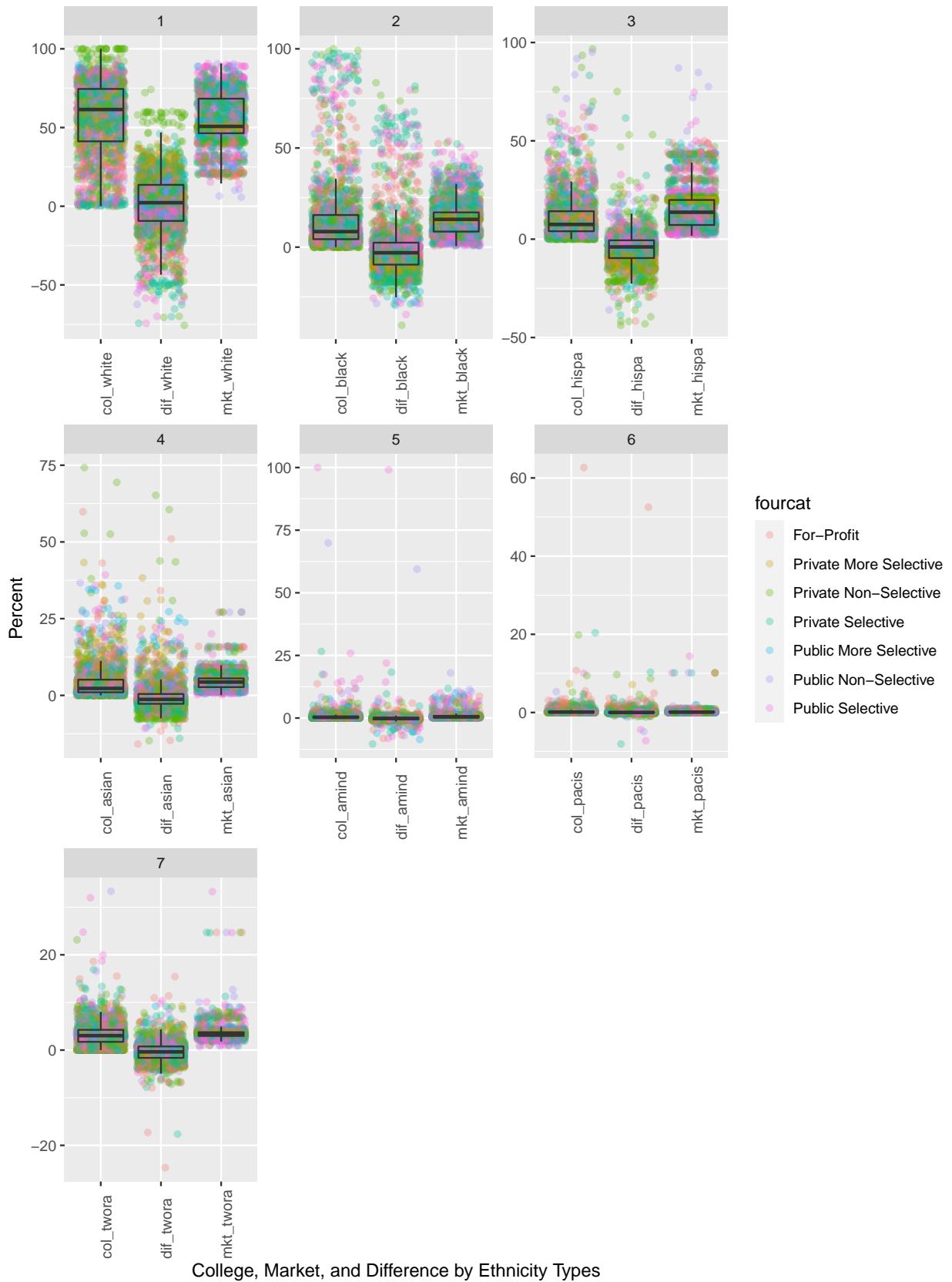
fips_ipeds	sd(dif_white)	sd(dif_hispa)	sd(dif_black)	sd(dif_asian)	sd(dif_amind)	sd(dif_pacis)	sd(dif_twora)
Alabama	30.151	5.125	30.552	1.735	0.370	0.559	1.717
Alaska	6.182	1.758	1.488	3.274	7.966	0.600	2.853
Arizona	10.463	8.439	6.921	2.746	1.639	0.395	3.168
Arkansas	25.259	5.430	25.040	1.630	0.995	0.167	2.617
California	16.601	14.786	8.259	13.717	0.645	0.964	3.388
Colorado	12.283	7.568	6.430	2.129	4.492	0.259	2.588
Connecticut	12.976	5.468	8.735	4.617	0.426	0.161	1.370
Delaware	20.620	2.951	24.734	2.618	0.480	0.186	2.341
District of Columbia	25.477	7.156	27.403	3.870	0.530	0.246	1.589
Florida	22.798	17.935	20.174	2.232	0.422	0.370	1.612
Georgia	25.350	5.219	26.203	4.485	0.376	0.307	1.740
Hawaii	7.370	5.601	5.813	10.512	0.250	18.792	11.212
Idaho	9.373	6.104	0.725	0.623	0.691	0.429	1.479
Illinois	18.048	9.554	13.383	4.671	0.445	0.379	2.352
Indiana	19.496	6.727	15.910	2.422	0.416	0.137	1.363
Iowa	20.792	5.394	6.994	3.138	0.498	0.264	1.178
Kansas	16.653	5.390	6.518	1.842	19.167	0.233	1.941
Kentucky	17.677	4.907	11.194	1.232	0.317	0.105	1.904
Louisiana	26.049	6.634	29.389	2.206	1.196	0.074	1.403
Maine	11.117	6.018	4.499	2.378	1.064	0.140	2.158
Maryland	25.733	4.672	26.785	6.501	0.341	0.182	1.870
Massachusetts	16.924	6.199	8.743	6.050	0.443	0.130	1.735
Michigan	16.389	4.295	10.992	3.073	0.876	0.180	1.622
Minnesota	19.003	6.694	7.537	3.232	1.527	0.111	2.399
Mississippi	30.323	4.832	31.375	0.641	0.324	0.108	1.415
Missouri	19.732	5.410	15.277	2.631	0.811	0.232	1.915
Montana	11.900	5.430	4.671	1.760	2.667	0.302	1.799
Nebraska	15.231	5.710	5.451	1.702	0.424	0.414	2.222
Nevada	15.544	10.028	4.136	12.450	1.207	0.577	4.191
New Hampshire	14.621	4.635	7.051	3.733	0.512	0.182	1.625
New Jersey	28.891	12.626	13.004	6.673	0.297	0.208	1.459
New Mexico	13.598	12.289	4.669	1.450	16.378	0.866	3.997
New York	26.095	11.908	9.749	6.506	0.430	0.194	1.912
North Carolina	25.506	4.025	27.747	2.864	2.010	0.373	1.860
North Dakota	9.770	4.838	3.906	1.494	3.791	0.350	2.004
Ohio	21.293	6.188	16.845	2.644	0.380	0.138	1.544
Oklahoma	16.615	6.013	14.368	1.778	5.147	0.262	4.818
Oregon	14.273	8.984	5.074	3.451	0.769	1.805	3.191
Pennsylvania	17.928	5.908	12.611	5.571	0.475	0.136	1.603
Rhode Island	18.638	5.642	6.476	6.685	0.293	0.092	1.863
South Carolina	27.639	3.583	30.374	1.224	0.535	0.464	1.895
South Dakota	14.958	5.218	10.274	1.243	2.761	0.311	2.219
Tennessee	25.520	5.912	23.857	2.502	0.489	0.195	1.864
Texas	17.378	11.688	18.803	6.023	0.714	0.371	1.882
Utah	16.474	4.004	5.609	1.730	0.989	1.166	3.403
Vermont	15.129	6.789	7.951	2.004	1.352	0.122	2.114
Virginia	21.117	4.155	20.520	3.926	0.636	0.541	2.252
Washington	17.866	11.881	7.458	6.901	1.628	1.149	3.582
West Virginia	12.859	2.184	5.870	1.133	0.472	0.417	2.199
Wisconsin	13.372	6.390	5.376	1.893	0.573	0.295	1.220
Wyoming	NA						

To answer the question of the degree of difference between college racial and ethnic compositions and the compositions of the corresponding geographic markets we decided that the information should be broken down by several variables. The largest factor in relation to geographical markets was the racial breakdown by state (fips_ipeds), as there was quite a lot of variation in population demographics. We then used the variables of ethnicity and college type (fourcat) to further define the data.

College–Market Difference of Percent Versus Ethnicity



Percentages by Ethnicity (Separated)



$\min(\text{dif_white})$	$\max(\text{dif_white})$	$\min(\text{dif_hispa})$	$\max(\text{dif_hispa})$	$\min(\text{dif_black})$	$\max(\text{dif_black})$	$\min(\text{dif_asian})$	$\max(\text{dif_asian})$
-75.668	74.25	-43.805	75.065	-39.422	82.922	-15.846	65.207

$\min(\text{dif_amind})$	$\max(\text{dif_amind})$	$\min(\text{dif_pacis})$	$\max(\text{dif_pacis})$	$\min(\text{dif_twora})$	$\max(\text{dif_twora})$
-10.365	99.096	-8.066	52.522	-24.692	15.425

In order to determine which racial or ethnic groups had the largest discrepancies between representation in colleges and the markets we organized the data into two separate graphical representations. The first being a simple comparison of differences between racial or ethnic groups, the second being separate graphs per each racial or ethnic group comparing the college representation, market representation, and the difference. This allowed for a general overview as well as some more specific data per demographic.

Percentage by Institution Type

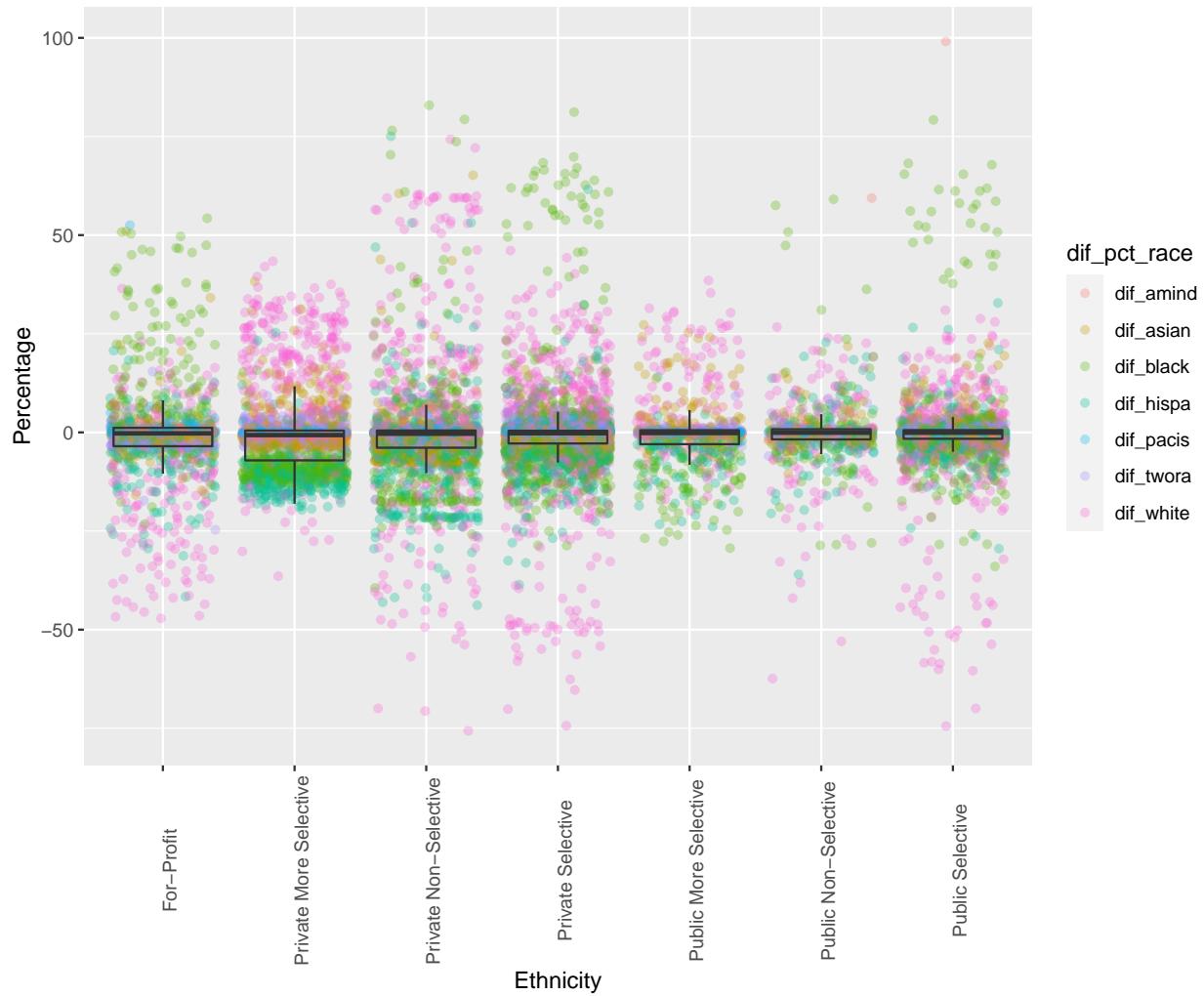


Table 1: White difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-47.145	36.454	-11.496	-12.062	17.497
Private More Selective	-36.443	43.368	14.721	12.248	15.807
Private Non-Selective	-75.668	74.250	4.605	5.156	27.464
Private Selective	-74.348	46.089	1.520	-1.581	20.065
Public More Selective	-19.914	38.509	8.327	8.954	13.366
Public Non-Selective	-62.414	26.712	0.065	-1.658	15.332
Public Selective	-74.434	30.985	-0.320	-3.511	17.676

Table 2: Hispanic difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-41.642	20.701	-4.343	-5.414	8.780
Private More Selective	-19.660	25.274	-11.304	-10.425	5.678
Private Non-Selective	-43.805	75.065	-5.230	-5.967	13.478
Private Selective	-38.616	61.586	-3.041	-3.569	7.657
Public More Selective	-23.959	14.062	-4.155	-5.651	6.187
Public Non-Selective	-36.011	22.814	-1.707	-1.279	8.360
Public Selective	-29.487	32.800	-1.354	-1.413	5.769

Table 3: Black difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-23.233	54.269	7.230	12.346	15.589
Private More Selective	-13.991	15.785	-9.266	-8.690	3.683
Private Non-Selective	-39.422	82.922	-2.963	-1.204	16.117
Private Selective	-28.686	81.186	-1.911	2.390	18.617
Public More Selective	-29.353	2.844	-7.937	-9.224	7.395
Public Non-Selective	-28.635	59.072	-0.478	1.212	13.628
Public Selective	-33.977	79.224	-0.675	2.439	17.788

Table 4: Asian difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-15.846	50.950	-1.607	-0.954	6.472
Private More Selective	-5.013	38.275	-0.084	2.159	6.627
Private Non-Selective	-14.764	65.207	-2.181	-1.355	7.519
Private Selective	-14.040	15.873	-1.919	-1.735	2.729
Public More Selective	-6.198	27.273	1.352	3.912	7.273
Public Non-Selective	-10.498	21.809	-0.764	0.164	4.314
Public Selective	-10.779	22.656	-0.523	0.258	4.209

Table 5: American Indian difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-8.152	5.953	0.119	0.155	1.352
Private More Selective	-0.895	4.342	-0.730	-0.582	0.468
Private Non-Selective	-4.476	6.353	-0.194	-0.072	0.873
Private Selective	-10.365	18.286	-0.035	0.000	1.238
Public More Selective	-8.552	6.443	-0.236	-0.591	1.507
Public Non-Selective	-7.055	59.403	-0.080	0.413	5.599
Public Selective	-8.123	99.096	-0.114	0.200	5.787

Table 6: Pacific Islander difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-0.792	52.522	0.277	0.810	4.008
Private More Selective	-0.208	7.147	-0.139	-0.052	0.473
Private Non-Selective	-1.055	9.700	-0.031	0.195	0.875
Private Selective	-8.066	10.293	0.056	0.152	0.706
Public More Selective	-0.656	0.549	0.008	-0.010	0.130
Public Non-Selective	-4.873	2.088	0.029	0.028	0.554
Public Selective	-7.214	2.205	0.024	-0.001	0.502

Table 7: Multiracial difference summary

fourcat	minDif	maxDif	medDif	meanDif	sdDif
For-Profit	-24.692	15.425	-2.412	-1.766	3.686
Private More Selective	-3.753	9.260	-0.076	0.216	2.168
Private Non-Selective	-7.195	9.120	-1.278	-1.214	2.303
Private Selective	-17.637	9.127	-0.489	-0.368	2.232
Public More Selective	-3.578	9.216	0.190	0.208	1.412
Public Non-Selective	-6.474	8.638	-0.117	-0.007	2.158
Public Selective	-6.080	10.372	0.250	0.228	1.575

To examine discrepancies in colleges by institution level, degree of selectivity, and/or public/private/for profit status, while still making the data regarding individual racial or ethnic groups visible, we decided to make each data point color coded by ethnicity. This allowed us to examine not only the amount of discrepancy but also trends involving racial and ethnic groups.

Conclusions

Q1. Many mean and median values lie near zero, but there is significant deviation present. For approximately 50% of the states the difference in the market vs college attendance is primarily within 25% of zero difference. There appears to be a correlation between population density and greater deviation.

Q2. There are three groups where the discrepancies can be considered large, each with slight differences in distribution. The first group is the White students. This category probably has the largest amount of discrepancy without a skew and no particular trend to the variation. The other two groups with large discrepancies are Black and Hispanic students. For the most part, they have similar negative skews and ranges but there are a significantly larger amount of positive outliers among Black students.

Q3. Overall it seems that public colleges tend to have narrower ranges of discrepancy than private and for profit colleges. When looking at outliers there is a trend across all but one type of college where there tend to be more density of positive outliers for White and Asian students and more density of negative outliers for Black and Hispanic students. The exception in this trend is the for profit schools where the outliers for White and Black students are reversed. The Private, more selective schools have the largest range of discrepancies excluding outliers.

```

library(tidyr)
library(dplyr)
library(ggplot2)
library(kableExtra)
#####
#           General setup for choosing files and setup          #
#####

# Finding the path for the file
my.file <- file.choose()

# Reading file into a data gram
# File has a header
fryr_collg_demogr <-
  read.csv(my.file,
    header = TRUE,
    sep = ",",
    stringsAsFactors = F)

# Filters down the results by the year 2017
fryr_collg_demo_2017 <- fryr_collg_demogr %>% filter(year == 2017)

# Shortens the number of variables to columns starting with dif, col, mkt
fryr_2017_short <-
  fryr_collg_demo_2017 %>% select(fips_ipeds,
    inst_name,
    fourcat,
    starts_with("dif_"),
    starts_with("mkt_"),
    starts_with("col_"))
  )

# Lengthens the columns starting with dif to use in plots
fryr_2017_short_dif <-
  fryr_2017_short %>% pivot_longer(cols = dif_white:dif_twora,
    names_to = "dif_pct_race",
    values_to = "dif_in_pct")

#####
#           Plots and Tables for Research Question 1          #
#####

# Groups college demographics by state
fryr_collg_demo_2017 <- fryr_collg_demo_2017 %>% group_by(fips_ipeds)

# Filters out Private More Selective for more accurate min/max in col columns
#fryr_collg_demo_2017 <- fryr_collg_demo_2017 # %>%
#  filter(fourcat != "Private More Selective")

# Setting up data for all ggplots
fryr_plot <- ggplot(data = fryr_2017_short_dif)

# Plot for research question 1

```

```

# Displays a graph of ethnicity on the x axis and the percent difference
# between College demographic percentages and Market demographic percentages
# separated by state
fryr_plot + geom_point(
  aes(x = dif_pct_race, y = dif_in_pct, color = fourcat, shape="."),
  alpha = 0.3,
  position = 'jitter') +
  facet_wrap(~ fips_ipeds, ncol = 5) + xlab("Ethnicity") +
  ylab("Difference Between College and Market Percent Demographics") +
  ggtitle(label = "Percent Difference b/w College and Market Vs Ethnicity") +
  theme(axis.text.x = element_text(
    angle = 90,
    vjust = 1,
    hjust = -0.1
  ))

# Table for research question 1
# Summary of statistics for each state
# This will have the min, max, median, mean, and standard deviation of different
# ethnicities. Excludes Private More Selective as this will impact the data
sum_fryr_cllg_2017_mean <- fryr_collg_demo_2017 %>% summarise(
  mean(dif_white),
  mean(dif_hispa),
  mean(dif_black),
  mean(dif_asian),
  mean(dif_amind),
  mean(dif_pacis),
  mean(dif_twora),
)
sum_fryr_cllg_2017_sd <- fryr_collg_demo_2017 %>% summarise(
  sd(dif_white),
  sd(dif_hispa),
  sd(dif_black),
  sd(dif_asian),
  sd(dif_amind),
  sd(dif_pacis),
  sd(dif_twora)
)
knitr::kable(sum_fryr_cllg_2017_mean, format = "latex", digits = 3) %>%
  kable_styling(latex_options = "scale_down")
knitr::kable(sum_fryr_cllg_2017_sd, format = "latex", digits = 3) %>%
  kable_styling(latex_options = "scale_down")
# Writes the summary table above to a .txt file
# write.table(
#   sum_fryr_cllg_2017,
#   file = "E:/MTH 3270 Data Science/Midterm Project/resQuest1.txt",
#   sep = ",",
#   quote = F,
#   row.names = T
# )

#####

```

```

#           Plots and Tables for Research Question 2           #
#####
# Plot for research question 2
# x-axis: ethnicity
# y-axis: Difference in percents b/w College demographics and Market demographics
# graph isn't separated by states
# will show the different types of schools by color
fryr_plot + geom_point(mapping = aes(x = dif_pct_race,
                                      y = dif_in_pct,
                                      color = fourcat),
                        alpha = 0.5, position = 'jitter') +
  xlab("Ethnicity") +
  ylab("Percent Difference b/w Market and College Demographics") +
  ggtitle(label = "College-Market Difference of Percent Versus Ethnicity") +
  theme(axis.text.x = element_text(angle = 90), aspect.ratio = 0.8)

# Table to put differences between college, market, and difference of the two
coll4_long<-fryr_2017_short%>%
  pivot_longer(cols = dif_white:col_twora)%>%
  mutate(rID = case_when(name=="col_white" | name=="mkt_white" | name=="dif_white" ~ 1,
                         name=="col_black" | name=="mkt_black" | name=="dif_black" ~ 2,
                         name=="col_hispa" | name=="mkt_hispa" | name=="dif_hispa" ~ 3,
                         name=="col_asian" | name=="mkt_asian" | name=="dif_asian" ~ 4,
                         name=="col_amind" | name=="mkt_amind" | name=="dif_amind" ~ 5,
                         name=="col_pacis" | name=="mkt_pacis" | name=="dif_pacis" ~ 6,
                         name=="col_twora" | name=="mkt_twora" | name=="dif_twora" ~ 7)) %>%
  group_by(rID)

# Plot for both research question 1 and research question 2
# x-axis: ethnicity type
# y-axis: percentages of races
# separated by type of school
ggplot(data = coll4_long, mapping=aes(x=reorder(name,rID),y=value,))+ 
  geom_point(mapping=aes(color=fourcat),alpha=.3,position = "jitter")+
  geom_boxplot(outlier.shape = NA,fill=NA)+ 
  theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust = 0)) +
  facet_wrap(~rID,scales = "free") +
  xlab("College, Market, and Difference by Ethnicity Types") +
  ylab("Percent") + ggtitle(label = "Percentages by Ethnicity (Separated)")

# Summary table to show min/max of each race
sum_fryr_2017_by_race_first <- fryr_collg_demo_2017 %>% ungroup() %>%
  summarise(
    min(dif_white),
    max(dif_white),
    min(dif_hispa),
    max(dif_hispa),
    min(dif_black),
    max(dif_black),
    min(dif_asian),
    max(dif_asian)
  )

```

```

sum_fryr_2017_by_race_last <- fryr_collg_demo_2017 %>% ungroup() %>%
  summarise(
    min(dif_amind),
    max(dif_amind),
    min(dif_pacis),
    max(dif_pacis),
    min(dif_twora),
    max(dif_twora)
  )

knitr::kable(sum_fryr_2017_by_race_first, format = "latex", digits = 3) %>%
  kable_styling(latex_options = "scale_down")

knitr::kable(sum_fryr_2017_by_race_last, format = "latex", digits = 3) %>%
  kable_styling(latex_options = "scale_down")

# Writes the summary table above to a .txt file
# write.table(
#   sum_fryr_cllg_2017,
#   file = "E:/MTH 3270 Data Science/Midterm Project/resQuest2.txt",
#   sep = ",",
#   quote = F,
#   row.names = T

#####
#          Plots and Tables for Research Question 3
#####

# Setup data frame for graph
fryr_2017_short_dif <- fryr_2017_short %>% group_by(fourcat)

# Setup data frame for table
fryr_2017_type_cllg <- fryr_2017_short %>% group_by(fourcat)

# Plotting of percent difference versus type of school
fryr_plot + geom_point(mapping = aes(x = fourcat,
                                       y = dif_in_pct,
                                       color = dif_pct_race),
                        position = 'jitter',
                        alpha = 0.3) +
  ggtitle(label = "Percentage by Institution Type") +
  xlab("Ethnicity") + ylab("Percentage") +
  geom_boxplot(mapping=aes(x = fourcat,
                           y = dif_in_pct ),
               outlier.shape = NA,
               fill=NA, alpha = 0.8) +
  theme(axis.text.x = element_text(angle = 90), aspect.ratio = 0.8)

# Table to summarize data based on type of college including min, max, median
# mean, and standard deviation of each race
sum_type_cllg_white <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_white),
            maxDif=max(dif_white),

```

```

    medDif=median(dif_white),
    meanDif=mean(dif_white),
    sdDif=sd(dif_white))

sum_type_cllg_hispa <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_hispa),
            maxDif=max(dif_hispa),
            medDif=median(dif_hispa),
            meanDif=mean(dif_hispa),
            sdDif=sd(dif_hispa))

sum_type_cllg_black <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_black),
            maxDif=max(dif_black),
            medDif=median(dif_black),
            meanDif=mean(dif_black),
            sdDif=sd(dif_black))

sum_type_cllg_asian <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_asian),
            maxDif=max(dif_asian),
            medDif=median(dif_asian),
            meanDif=mean(dif_asian),
            sdDif=sd(dif_asian))

sum_type_cllg_amind <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_amind),
            maxDif=max(dif_amind),
            medDif=median(dif_amind),
            meanDif=mean(dif_amind),
            sdDif=sd(dif_amind))

sum_type_cllg_pacis <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_pacis),
            maxDif=max(dif_pacis),
            medDif=median(dif_pacis),
            meanDif=mean(dif_pacis),
            sdDif=sd(dif_pacis))

sum_type_cllg_twora <- fryr_2017_type_cllg %>%
  summarise(minDif=min(dif_twora),
            maxDif=max(dif_twora),
            medDif=median(dif_twora),
            meanDif=mean(dif_twora),
            sdDif=sd(dif_twora))

knitr::kable(sum_type_cllg_white,caption = "White diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_hispa,caption = "Hispanic diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_black,caption = "Black diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_asian,caption = "Asian diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_amind,caption = "American Indian diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_pacis,caption = "Pacific Islander diffrence summary",format = "latex",digits = 3)
knitr::kable(sum_type_cllg_twora,caption = "Multiracial diffrence summary",format = "latex",digits = 3)
# Writes the summary table above to a .txt file

```

```
# write.table(  
#   sum_fryr_cllg_2017,  
#   file = "E:/MTH 3270 Data Science/Midterm Project/resQuest3.txt",  
#   sep = ",",  
#   quote = F,  
#   row.names = T
```