

# Data Stories on “How Americans Vote?”

Zi Fang

1/13/2021

## 1. Introduction

Year 2020 has been full of surprises. Some people say 2020 is a magic year since so many things that you would never think about happened. The 2020 Presidential Election and the COVID-19 pandemic are two of the most important things that closely related to people in the United States. In this project, I will discuss data stories about “How Americans Vote” with a focus on 2020 Presidential Election. I am conducting the project using *2020 Exploratory Testing Survey* and *Time Series Cumulative Data File* from *The American National Election Studies* (ANES).

The 2020 Exploratory Testing Survey is a pilot study collected through the Internet and have combined 3080 sample data from three separate opt-in online panels. Time Series Cumulative Data File is a huge project that all cross-section cases and variables for select questions from the ANES Time Series studies conducted since 1948 have merged into. It's a quite huge project with meaningful variables. More information can be found in ANES official website.

## 2. How accurate did the votes be counted in 2020 Election?

2020 is a special year in comparing to the entire election history. On account of COVID-19, the election votes are collected generally through two methods: mail and in-person vote. The hybrid collecting mode causes some doubts about the accuracy in collecting the votes and counting the votes. Not only us, but also President Trump have the same concern, for example, Georgia's nearly 5 million votes in the presidential race will be counted for a third time, as President Trump's campaign has formally asked for a recount. Let's dive deep into the data and see what normal people think.

In order to get a general idea, we clean the data, and collect `votecount`, `votecount_mail` (also named `mail`) and `id` as three variables, the corresponding questions and the code for responses can be found here:

`votecount` –

In the November 2020 general election, how accurately do you think the votes will be counted?

☐ Not at all accurately [1]

☐ Not very accurately [2]

☐ Moderately accurately [3]

\_\_ Very accurately [4]

\_\_ Completely accurately [5]

votecount\_mail / mail –

How much do you trust that votes are counted accurately when people mail in their ballots?

1. A great deal
2. A lot
3. A moderate amount
4. A little
5. Not at all

voterid1 –

In your state, are people usually asked to show an ID when voting, or not asked to show an ID?

\_\_ Asked to show an ID [1]

\_\_ Not asked to show an ID [2]

\_\_ Unsure [3]

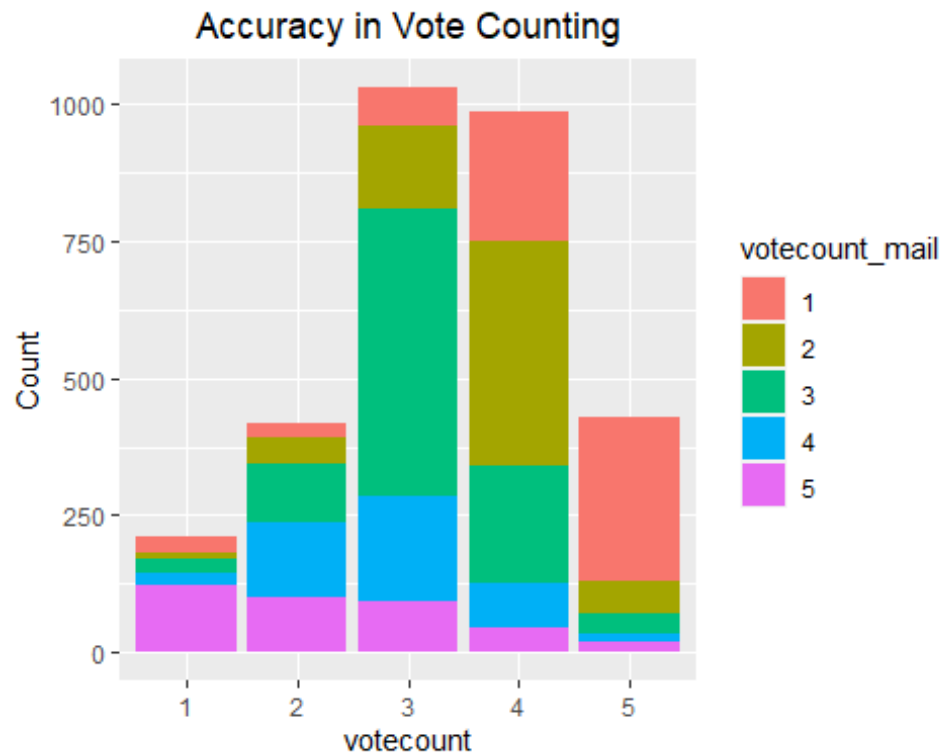
From the first plot named “Accuracy in Vote Counting”, if only focus on the y-values, we can find most of the people think the votes are moderately accurately or very accurately in the November 2020 general election.

From the third plot named “Accuracy of Votes Counting By Mail”, we can find most of the people choose a moderate amount of trust, and then a lot trust, so most people have trust among the mail votes.

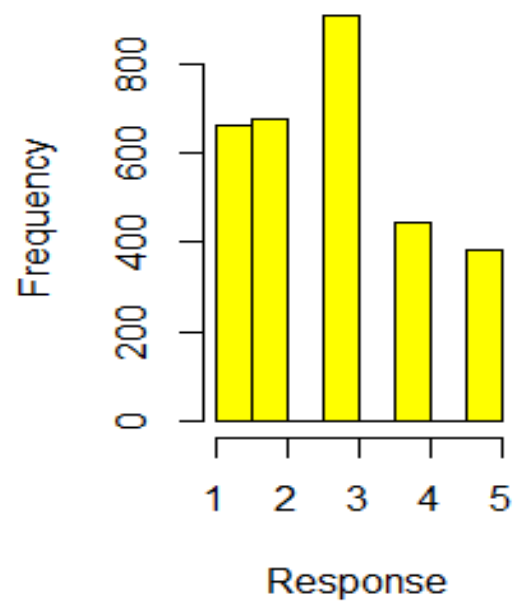
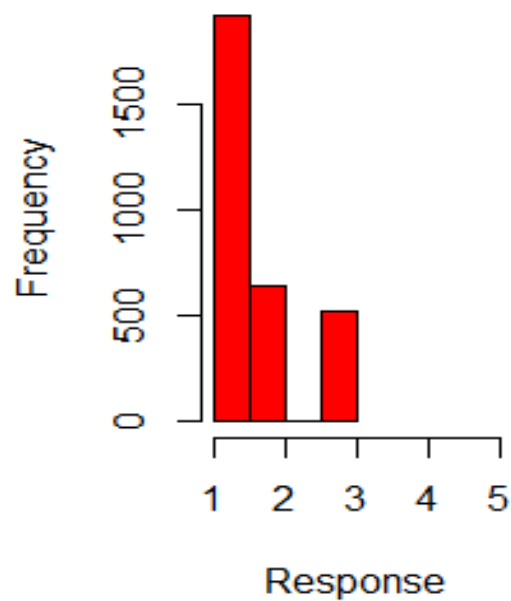
From the second plot named “Accuracy of Votes Counting”, more than a half of the people say that they have to show an ID when voting, which helps the accuracy of in-person voting.

In the first plot, when people choose “5. Completely accurately” inotecount they have a greater probability of choosing “1. A great deal” inotecount\_mail. When people choose “4. Very accurately” inotecount they have a greater probability of choosing “2. A lot” inotecount\_mail. The interesting phenomenon indicates that people think the accuracy of mail votes and all the votes are normally consistent.

Therefore, although there are some portion of people doubt the accuracy of votes, most people tend to trust the accuracy of counting votes.



**Accuracy of Votes Counting**



### 3. How did Trump's Approval Rating Change between 2016 and 2020?

Right now, we all know that Joe Biden has been appointed as the 46th and current president of the United States, but according to the general pattern, Trump should be able to maintain being the president for another four years. What happened during 2016 to 2020, the four year period of time? How did Trump's approval rating be changed?

In order to figure out the the approval rate changes, I cleaned the data and select three variables: vote16, vote20jb and vote20\_party to be the main focus. Detailed description about the variables and the meaning of the numbers can be find here:

vote16 –

In the 2016 presidential election, who did you vote for? Donald Trump, Hillary Clinton, or someone else?

☐ Donald Trump [1]

☐ Hillary Clinton [2]

☐ someone else [3]

vote20jb –

If the 2020 presidential election were between Donald Trump for the Republicans and Joe Biden for the Democrats, would you vote for Donald Trump, Joe Biden, someone else, or probably not vote?

☐ Donald Trump [1]

☐ Joe Biden [2]

☐ Someone else [3]

☐ Probably not vote [4]

vote20\_party –

If the election for the U.S. House of Representatives were being held today, and you had to make a choice, would you be voting for the Republican candidate or the Democrat candidate in your district?

☐ Democrat [1]

☐ Republican [2]

☐ Other [3]

☐ Won't vote [4]

\_\_ Don't know [5]

The following three consecutive plot gives us an overview of vote16, vote20jb and vote20\_party.

In the 2016 election, votes for Trump and Clinton are almost the same, but Trump has a little bit more votes, which proves Trump's victory in 2016 Presidential Election.

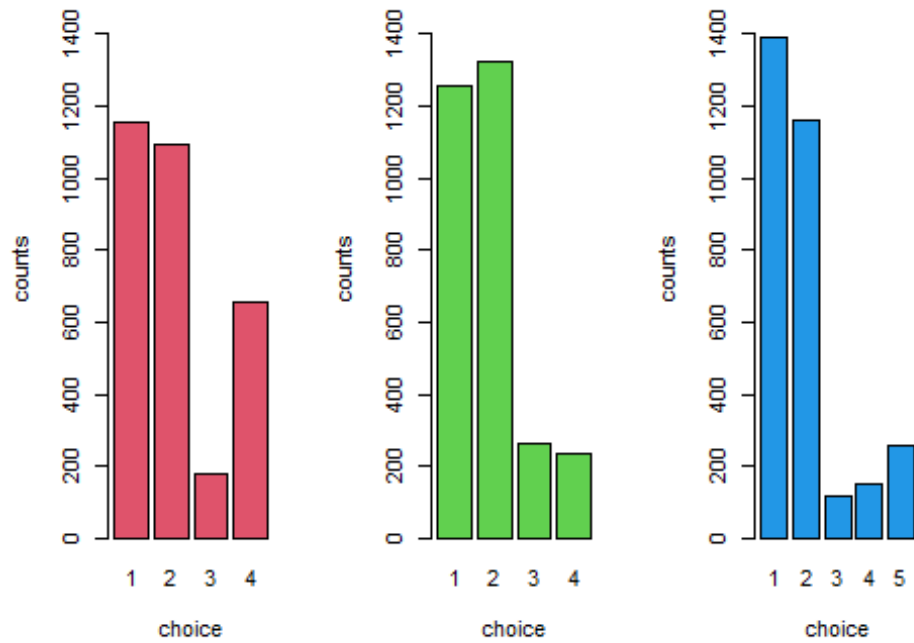
From the second graph, in 2020 election, votes between Trump and Biden are still close, but the gap is greater than the gap that we got in 2016 vote Condition. Biden has more votes, and it is proved by the 2020 election result.

Then, the graph leads me to the next question: is that because people prefer Biden than Trump or because people favor Democrat party so they vote for Biden? Biden himself or the Democrat party weights more in the election?

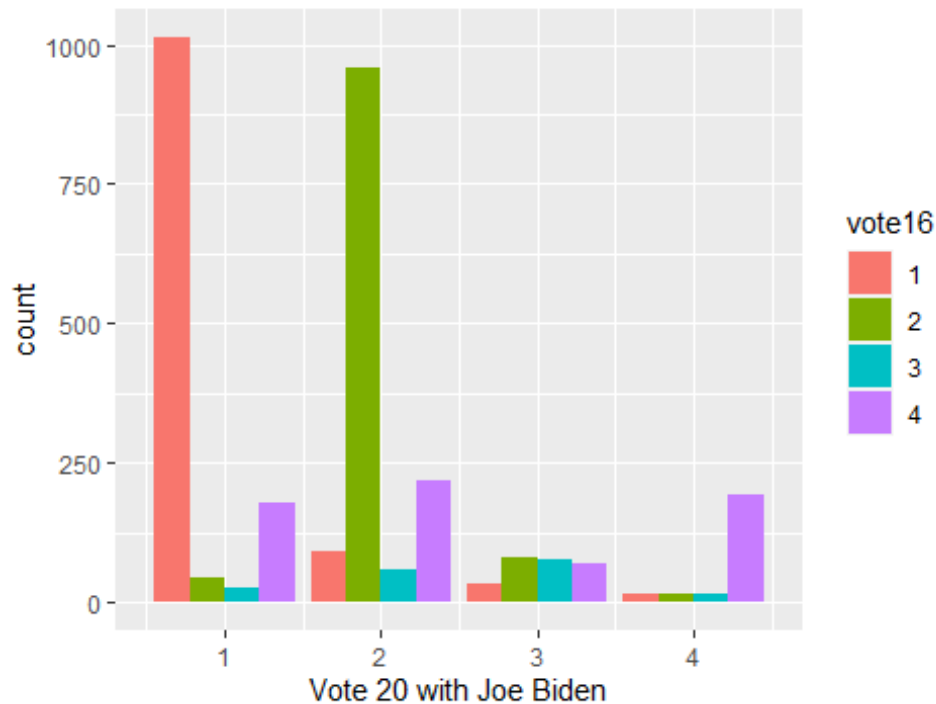
Based on the third graph, most of the people votes for democrat party than republican, and in comparing with the second graph, the gap between the two parties are larger then the gap between Biden and Trump.

From the last graph in the section named "Trump 2016 Votes & 2020 Votes(Joe Biden) Comparison", we are comparing vote16 and vote20jb. We can find that most of the people who voted for Trump in 2016 will still vote for him, and most of the people who voted for Clinton will vote for Biden in 2020. There are some people switched from Trump to Biden and some others switch from Clinton to Trump, but the extend of switching from Trump to Biden than switching from Clinton to Trump is higher, which is also the main reason in winning the competition. This part of people, especially people living in a swing state, is the people who have the real deciding power in hand, since gathering their opinion together decides corresponding state's decision.

2016 Vote Condition 2020 Vote Condition-Per 2020 Vote Condition-Pai



Trump 2016 Votes & 2020 Votes(Joe Biden) Comparison



#### **4. How well did President Trump handle his job in the past four years?**

It seems that Trump's performance was not satisfying enough, so that he lost some swing states in the 2020 election. Some people wrote articles saying that President Trump lost the election mainly because of COVID-19. COVID-19 is influential, and may have some influence on Trump's support rate, but is that the main reason? Despite COVID-19, did President Trump do all other job well?

In order to answer the question, I select job performance, relations, immigration and economy as the main object with variable names job, relations, imm and eco. Detailed description about the questions and the code meaning can be found below.

job –

Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling his job as president?

relations –

Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling relations with foreign countries?

imm –

Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling immigration?

eco –

Do you approve, disapprove, or neither approve nor disapprove of the way Donald Trump is handling the economy?

☐ Approve extremely strongly [1]

☐ Approve moderately strongly [2]

☐ Approve slightly [3]

☐ Neither approve nor disapprove [4]

☐ Disapprove slightly [5]

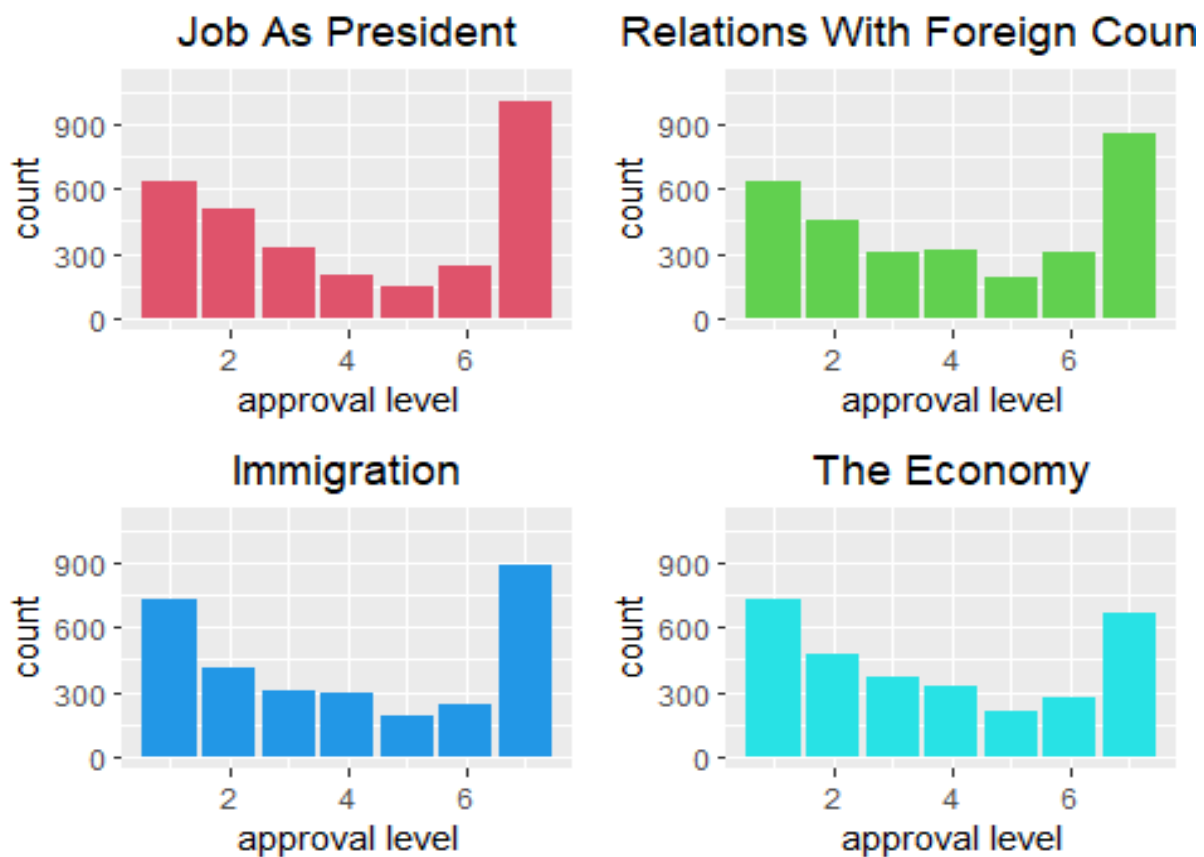
☐ Disapprove moderately strongly [6]

☐ Disapprove extremely strongly [7]

The pattern of the four bar plots are similar, and they all have U-shapes. The high proportions are mainly on the two ends, meaning most of the people either approve extremely strongly or disapprove extremely strongly. Let's take a look by part.

Most of the people disapprove extremely strongly that President Trump is performing well during the four year. Most of the people disapprove extremely strongly Trump's way of handling relations with foreign countries and handling immigration, but they approve extremely strongly Trump's way of handling economy. Actually, the gap between the first place and the second place in economy is much much lower than the gaps in general job performance, relations and immigration.

Therefore, even without COVID-19, there are other parts such as relations with foreign countries and immigration policy that Trump was using opposite methods than what people desire him to handle. The general evaluation result is that most of the people disapprove strongly about the Trump's job, which makes the results that we find in the previous section reasonable.





## 5. Is the data collected reliable?

When talking about the *2020 Exploratory Testing Survey*, the website says, “As distinct from many ANES pilot surveys, the primary aim of this study was to allow for more targeted experimentation and testing of longer batteries of questions, with less concern for estimation of population characteristics.” So I start to think about if the data collected is reliable enough? Did the data really reduce enough influence of the population characteristics?

In this section, we use data from both *2020 Exploratory Testing Survey* and *Time Series Cumulative Data File*. Since in the following plots, all the variables are labeled appropriately, I will not use spaces to introduce what questions are asked when collecting the corresponding data. I choose race and sex as the main study focus, and want to make a comparison about the data pattern in *2020 Exploratory Testing Survey* and the trend in *Time Series Cumulative Data File*.

In the plot named “Different Race Groups Participated Over The Years”, we can find the portion of people voted is changing over the years. The proportion is consistently high for white people, and have been increased a lot for other groups of people. However, since about 76% of people living in the US are white, and 14% of people are black people, no wondering although the participation rates are increasing, the race allocation in the first pie chart named “Pie Chart of Race Distribution 2020” looks unevenly.

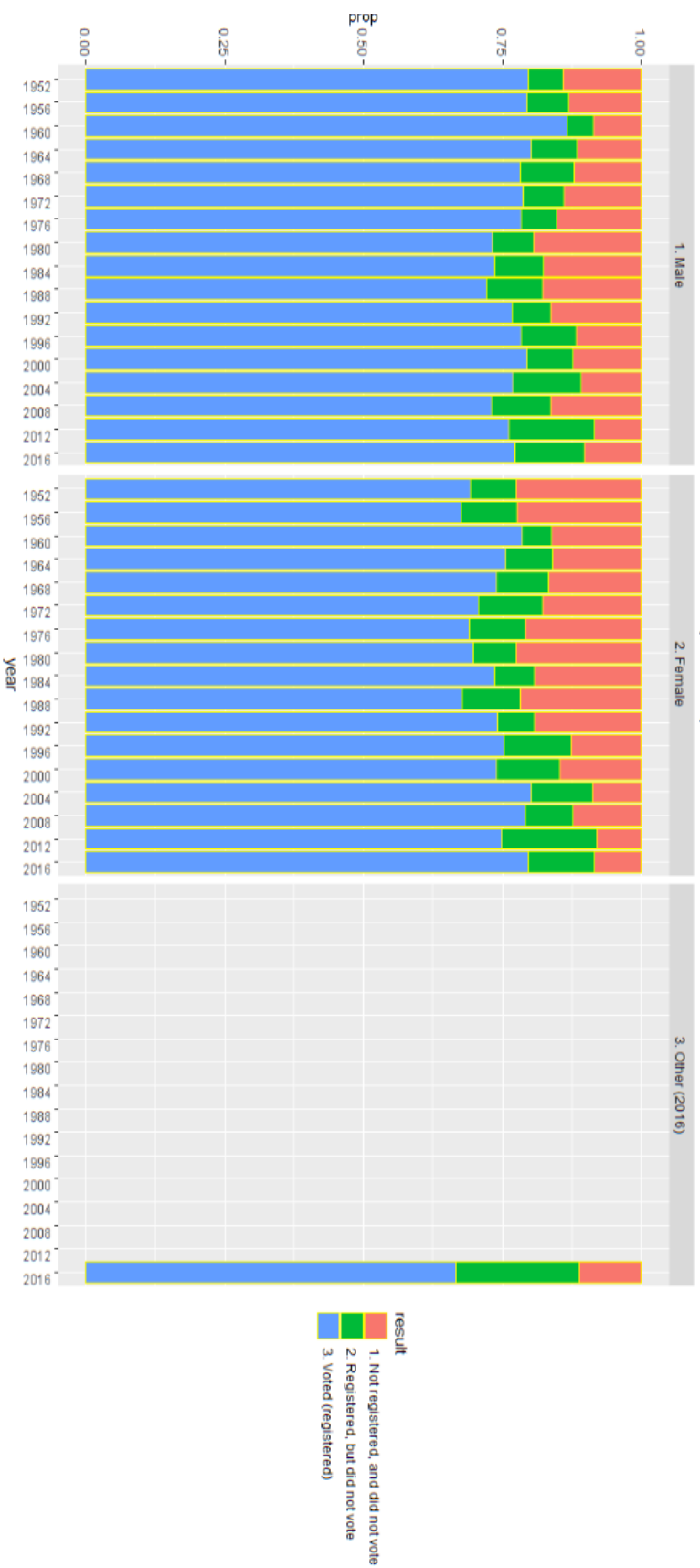
For the sex part, from the plot named “Different Gender Groups Participated Over The Years”, the proportion of male and female’s voting rate are almost the same. Plus US people’s genders are almost evenly distributed, so we expect to see a approximately evenly distributed pattern in the *2020 Exploratory Testing Survey*. And the graph called “Pie Chart of Sex 2020” proves that.

Therefore, I take race and gender as two examples to test the reliability of the *2020 Exploratory Testing Survey* dataset, and it passes the test, which provides more confidence in analyzing and explaining the results stated in the previous sections.

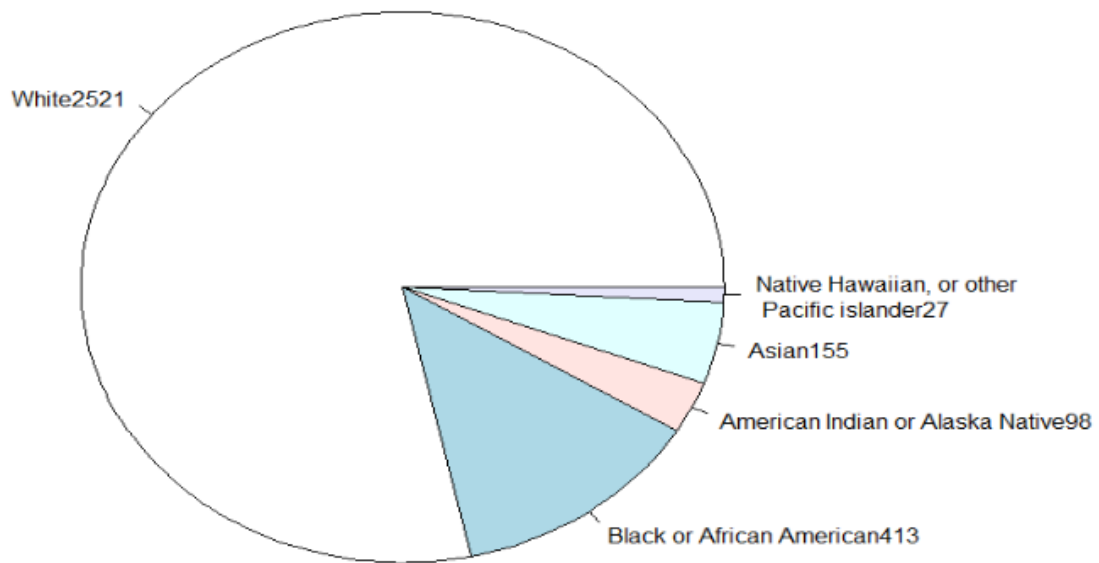
### Different Race Groups Participated Over The Years



Different Gender Groups Participated Over The Years



**Pie Chart of Race Distribution 2020**



**Pie Chart of Sex 2020**

