Final Statistical Report

Kelseigh Powell-Hall

06/04/2019

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.4.4
```

```
## -- Attaching packages ------------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.0     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
## Warning: package 'tibble' was built under R version 3.4.4
```

```
## Warning: package 'tidyr' was built under R version 3.4.4
```

```
## Warning: package 'readr' was built under R version 3.4.4
```

```
## Warning: package 'purrr' was built under R version 3.4.4
```

```
## Warning: package 'dplyr' was built under R version 3.4.4
```

```
## Warning: package 'stringr' was built under R version 3.4.4
```

```
## Warning: package 'forcats' was built under R version 3.4.4
```

```
## -- Conflicts ---------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.4.4
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
indresp1 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
```

```
## See spec(...) for full column specifications.
```

```
wave1 <- indresp1%>%
  select(pidp, a_vote4,a_sex_dv, a_urban_dv)
rm(indresp1)

indresp2 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave2 <- indresp2%>%
  select(pidp, b_vote4,b_sex_dv, b_urban_dv)
rm(indresp2)

indresp3 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave3 <- indresp3%>%
  select(pidp, c_vote4,c_sex_dv, c_urban_dv)
rm(indresp3)

indresp4 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave4 <- indresp4%>%
  select(pidp, d_vote4,d_sex_dv, d_urban_dv)
rm(indresp4)

indresp5 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave5 <- indresp5%>%
  select(pidp, e_vote4,e_sex_dv, e_urban_dv)
rm(indresp5)

      indresp6 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave6 <- indresp6%>%
  select(pidp, f_vote4,f_sex_dv, f_urban_dv)
rm(indresp6)

indresp7 <- read_tsv("C:/Users/Kelseigh/Documents/THIRD YEAR/DATA ANALYSIS 3/Project 1/Data/UKDA-6614-ta
```

```
## Parsed with column specification:
## cols(
##   .default = col_double()
## )
## See spec(...) for full column specifications.
```

```
wave7 <- indresp7%>%
  select(pidp, g_vote4,g_sex_dv, g_urban_dv)
rm(indresp7)
```

```
allwaves <- wave1 %>%
  left_join(wave2, by="pidp")

allwaves <- allwaves %>%
  left_join(wave3, by="pidp")

allwaves <- allwaves %>%
  left_join(wave4, by="pidp")

allwaves <- allwaves %>%
  left_join(wave5, by="pidp")

allwaves <- allwaves %>%
  left_join(wave6, by="pidp")

allwaves <- allwaves %>%
  left_join(wave7, by="pidp")
```

```
allwaves <- allwaves %>%
  na.omit(allwaves)

allwaves$b_sex_dv <- NULL
allwaves$c_sex_dv <- NULL
allwaves$d_sex_dv <- NULL
allwaves$e_sex_dv <- NULL
allwaves$f_sex_dv <- NULL
allwaves$g_sex_dv <- NULL

allwaves <- allwaves %>%
      mutate(sex = ifelse(a_sex_dv == 1, "Male",
                          ifelse(a_sex_dv == 2, "Female", NA)))
```

```
## Warning: package 'bindrcpp' was built under R version 3.4.4
```

```
allwaves$a_sex_dv <- NULL

allwaves$b_urban_dv <- NULL
allwaves$c_urban_dv <- NULL
allwaves$d_urban_dv <- NULL
allwaves$e_urban_dv <- NULL
allwaves$f_urban_dv <- NULL
allwaves$g_urban_dv <- NULL

allwaves <- allwaves %>%
      mutate(urban = ifelse(a_urban_dv == 1, "Urban",
                           ifelse(a_urban_dv == 2, "Rural", NA)))
allwaves$a_urban_dv <- NULL

allwaves <- allwaves %>%
  mutate(a_vote4 =  ifelse(a_vote4 == -8, NA,
                    ifelse(a_vote4 == -7, NA,
                    ifelse(a_vote4 == -2, NA,
                    ifelse(a_vote4 == -1, NA,
                    ifelse(a_vote4 == 1, "C",
                    ifelse(a_vote4 == 2, "L",
                    ifelse(a_vote4 == 3, "LD",                          ifelse(a_vote4 == 4, "Other",
                    ifelse(a_vote4 == 5, "Other",
                    ifelse(a_vote4 == 6, "Other",
                    ifelse(a_vote4 == 7, "Other",
                    ifelse(a_vote4 == 8, "Other",
                    ifelse(a_vote4 == 9, "Other",
                    ifelse(a_vote4 == 10, "Other",
                    ifelse(a_vote4 == 11, "Other", NA
                  ))))))))))))))

allwaves <- allwaves %>%
  mutate(b_vote4 =  ifelse(b_vote4 == -8, NA,
                    ifelse(b_vote4 == -7, NA,
                    ifelse(b_vote4 == -2, NA,
                    ifelse(b_vote4 == -1, NA,
                    ifelse(b_vote4 == 1, "C",
```

```
                    ifelse(b_vote4 == 2, "L",
                    ifelse(b_vote4 == 3, "LD",                          ifelse(b_vote4 == 4, "Other",
                    ifelse(b_vote4 == 5, "Other",
                    ifelse(b_vote4 == 6, "Other",
                    ifelse(b_vote4 == 7, "Other",
                    ifelse(b_vote4 == 8, "Other",
                    ifelse(b_vote4 == 9, "Other",
                    ifelse(b_vote4 == 10, "Other",
                    ifelse(b_vote4 == 11, "Other", NA
                  ))))))))))))))

allwaves <- allwaves %>%
  mutate(c_vote4 =  ifelse(c_vote4 == -8, NA,
                    ifelse(c_vote4 == -7, NA,
                    ifelse(c_vote4 == -2, NA,
                    ifelse(c_vote4 == -1, NA,
                    ifelse(c_vote4 == 1, "C",
                    ifelse(c_vote4 == 2, "L",
                    ifelse(c_vote4 == 3, "LD",                          ifelse(c_vote4 == 4, "Other",
                    ifelse(c_vote4 == 5, "Other",
                    ifelse(c_vote4 == 6, "Other",
                    ifelse(c_vote4 == 7, "Other",
                    ifelse(c_vote4 == 8, "Other",
                    ifelse(c_vote4 == 9, "Other",
                    ifelse(c_vote4 == 10, "Other",
                    ifelse(c_vote4 == 11, "Other", NA
                  ))))))))))))))
allwaves <- allwaves %>%
  mutate(d_vote4 =  ifelse(d_vote4 == -8, NA,
                    ifelse(d_vote4 == -7, NA,
                    ifelse(d_vote4 == -2, NA,
                    ifelse(d_vote4 == -1, NA,
                    ifelse(d_vote4 == 1, "C",
                    ifelse(d_vote4 == 2, "L",
                    ifelse(d_vote4 == 3, "LD",                          ifelse(d_vote4 == 4, "Other",
                    ifelse(d_vote4 == 5, "Other",
                    ifelse(d_vote4 == 6, "Other",
                    ifelse(d_vote4 == 7, "Other",
                    ifelse(d_vote4 == 8, "Other",
                    ifelse(d_vote4 == 9, "Other",
                    ifelse(d_vote4 == 10, "Other",
                    ifelse(d_vote4 == 11, "Other", NA
                  ))))))))))))))


allwaves <- allwaves %>%
  mutate(e_vote4 =  ifelse(e_vote4 == -8, NA,
                    ifelse(e_vote4 == -7, NA,
                    ifelse(e_vote4 == -2, NA,
                    ifelse(e_vote4 == -1, NA,
                    ifelse(e_vote4 == 1, "C",
                    ifelse(e_vote4 == 2, "L",
                    ifelse(e_vote4 == 3, "LD",                          ifelse(e_vote4 == 4, "Other",
```

```
                    ifelse(e_vote4 == 5, "Other",
                    ifelse(e_vote4 == 6, "Other",
                    ifelse(e_vote4 == 7, "Other",
                    ifelse(e_vote4 == 8, "Other",
                    ifelse(e_vote4 == 9, "Other",
                    ifelse(e_vote4 == 10, "Other",
                    ifelse(e_vote4 == 11, "Other", NA
                  ))))))))))))))

allwaves <- allwaves %>%
  mutate(f_vote4 =  ifelse(f_vote4 == -8, NA,
                    ifelse(f_vote4 == -7, NA,
                    ifelse(f_vote4 == -2, NA,
                    ifelse(f_vote4 == -1, NA,
                    ifelse(f_vote4 == 1, "C",
                    ifelse(f_vote4 == 2, "L",
                    ifelse(f_vote4 == 3, "LD",                          ifelse(f_vote4 == 4, "Other",
                    ifelse(f_vote4 == 5, "Other",
                    ifelse(f_vote4 == 6, "Other",
                    ifelse(f_vote4 == 7, "Other",
                    ifelse(f_vote4 == 8, "Other",
                    ifelse(f_vote4 == 9, "Other",
                    ifelse(f_vote4 == 10, "Other",
                    ifelse(f_vote4 == 11, "Other", NA
                  ))))))))))))))

allwaves <- allwaves %>%
  mutate(g_vote4 =  ifelse(g_vote4 == -8, NA,
                    ifelse(g_vote4 == -7, NA,
                    ifelse(g_vote4 == -2, NA,
                    ifelse(g_vote4 == -1, NA,
                    ifelse(g_vote4 == 1, "C",
                    ifelse(g_vote4 == 2, "L",
                    ifelse(g_vote4 == 3, "LD",                          ifelse(g_vote4 == 4, "Other",
                    ifelse(g_vote4 == 5, "Other",
                    ifelse(g_vote4 == 6, "Other",
                    ifelse(g_vote4 == 7, "Other",
                    ifelse(g_vote4 == 8, "Other",
                    ifelse(g_vote4 == 9, "Other",
                    ifelse(g_vote4 == 10, "Other",
                    ifelse(g_vote4 == 11, "Other", NA
                  ))))))))))))))

allwaves <- allwaves %>%
  drop_na()

voterec <- allwaves %>%
        unite(VR, a_vote4, b_vote4, c_vote4, d_vote4, e_vote4, f_vote4, g_vote4) %>%
        count(VR) %>%
        arrange(-n)

voterecM <- filter(allwaves, sex =="Male")
voterecM<- voterecM %>%
```

```
        unite(VR, a_vote4, b_vote4, c_vote4, d_vote4, e_vote4, f_vote4, g_vote4) %>%
        count(VR) %>%
        arrange(-n)


voterecF <- filter(allwaves, sex =="Female")
voterecF<- voterecF %>%
        unite(VR, a_vote4, b_vote4, c_vote4, d_vote4, e_vote4, f_vote4, g_vote4) %>%
        count(VR) %>%
        arrange(-n)


voterecU <- filter(allwaves, urban =="Urban")
voterecU<- voterecU %>%
        unite(VR, a_vote4, b_vote4, c_vote4, d_vote4, e_vote4, f_vote4, g_vote4) %>%
        count(VR) %>%
        arrange(-n)


voterecR <- filter(allwaves, urban =="Rural")
voterecR<- voterecR %>%
        unite(VR, a_vote4, b_vote4, c_vote4, d_vote4, e_vote4, f_vote4, g_vote4) %>%
        count(VR) %>%
        arrange(-n)

voterectJoined <- voterec %>%
  left_join(voterecM, by = "VR")

voterectJoined <- voterectJoined %>%
  left_join(voterecF, by="VR")


voterectJoined <- voterectJoined %>%
  rename(Male = n.y) %>%
  rename(Female = n) %>%
  rename(All = n.x)

voterectJoined <- voterectJoined %>%
  left_join(voterecU, by = "VR")

voterectJoined <- voterectJoined %>%
  left_join(voterecR, by="VR")


voterectJoined <- voterectJoined %>%
  rename(Urban = n.x) %>%
  rename(Rural = n.y)


voterectJoined <- voterectJoined[-c(3),]
```

Social scientists have long been interested in the way people vote in elections, and which characteristics are linked to voting for a particular party. Knowing information on the characteristics of those who vote for

particular parties is useful as it allows social scientists to make predictions about peoples voting patterns as well as explain differences in people political preferences. Some of the characteristics often analysed by researchers in relation to voting preferences include gender, income and region. Social scientists assert that people's political preferences rarely change overtime, so for example an individual who votes conservative will likely continue to vote conservative. In order to test this assertion this paper aims to analyse data from the Understand Society questionnaire to determine whether people's political preferences change overtime. Another aim of this paper is to find out if characteristics of individuals have an effect on whether their political preferences change. The characteristics this study will look at are sex and whether individuals live in a rural or urban area. Sex is a good characteristic to analyse as it is a clear binary variable that stays the same over the course of the voting records I will be looking at. It is also interesting to explore whether the assertion of social scientists that people rarely change their political preferences is true for both men and women. The second variable used in this analysis (whether individuals live in rural or urban areas), is important as different cultural aspects can influence the different voting patterns of individuals, and it is interesting to see if these cultural influences also affect the stability of an individual's political preferences.

The dataset used for this analysis is the Understanding Society Survey. This survey is completed by households across the UK, at the first wave of the dataset approximately 40,000 households responded. This is a longitudinal study and therefore the households are visited each year to be updated on any changed of circumstances. Interviews are carried out alongside the questionnaire by trained interviewers, this allows for the collection of more in-depth data than using a survey alone would allow. The survey aims to produce reliable longitudinal data on a range of subjects such as health, work, education, income and family/social life (Understanding Society – Main Survey). Given the large sample size (40,000 households) this is a good dataset to use, it is representative of the UK as a whole, this is positive as it means the conclusions drawn from this dataset are likely to be more reliable. This analysis will be using three variables from the dataset. The first variable is called 'vote4', this variable asked the participant which political party they feel closest to, the answers included political parties in the UK. To simplify this analysis the vote4 answers include 'Conservative', 'Labour', 'Liberal Democrat' and 'Other'. The second variable that will be included in this analysis is called 'sex_dv', this variable contains the sex of the participant. The third and final variable included in this analysis is called 'urban_dv,' this variable contains either rural or urban depending on where the individual lives. There are eight waves of the understanding society data, this analysis will only include the first seven waves of the survey as the eighth wave does not contain the 'vote4' variable. For this analysis all observations with missing values have been removed to ensure the results are not skewed. Also, the voting record 'Other_Other_Other_Other_Other_Other_Other' has been removed as it is impossible to say whether this includes consistent votes for the same party or different, therefore this analysis only focuses on three political parties ('Conservative', 'Labour' and 'Liberal Democrat').

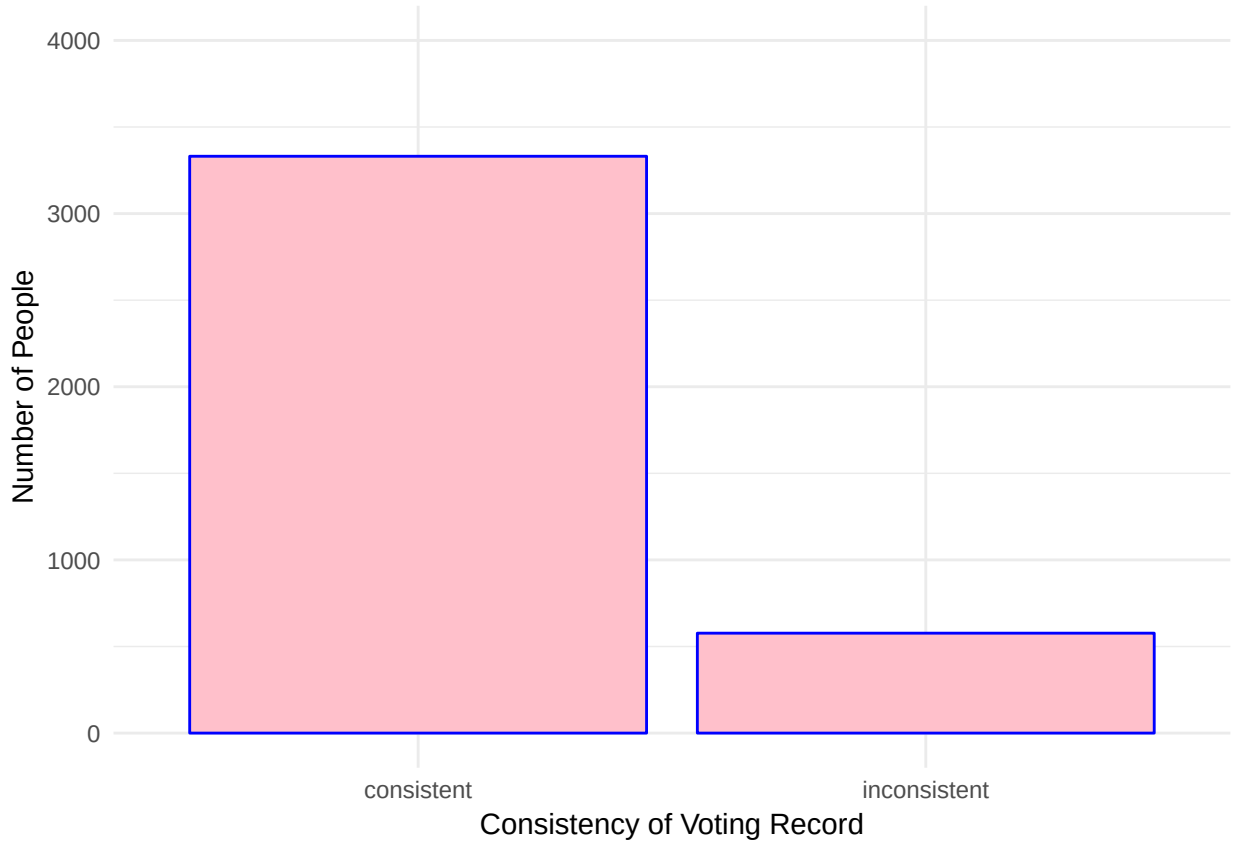
This analysis aims to find out whether the assertion by social scientists, that people rarely change their political preferences, is true. In order to determine it is necessary to compare the number of people in the sample that consistently voted for one party with those who changed their political preferences over the course of the survey. Secondly, in order to determine whether gender and location has an influence on an individual's political preference it is necessary to compare both men's political stability with women's. As well as comparing the political preference stability of those in rural areas as well as those in urban area.

For the first part of this analysis the voting records have been categorised into two groups. The first is called 'consistent' this contains the people who voted for the same party in all seven waves. The second group is called 'inconsistent', containing those who voted for more than one party over the sevens waves of the survey. The graph below shows these two categories and then the number of individuals that fall into each category. If the assertion of social scientists is true, we would expect to see a majority of people to vote for the same party over all seven waves and thus be in the 'consistent' category'.

```
VR <- c( "consistent", "inconsistent")
All <- c(3331, 577)
Alltab <- data.frame(VR, All)
View(Alltab)

Alltab %>%
```

```
  ggplot(aes(x=VR, y=All))+
  geom_bar(stat = "identity", colour = "blue", fill="Pink") +
  theme_minimal() +
  ylim(0,4000) +
  xlab("Consistency of Voting Record")+
  ylab("Number of People")
```



From the table we can see 3331 people voted consistently over the seven waves of the survey, this is 85.2% of people. This is more than half of the sample population. The sum of the remaining voting records accounts is 577 this is 14.8% of the sample population therefore this is the minority. This graph clearly shows that a majority of people have consistent voting records, meaning they did not change which party they voted for over all seven waves. The assertion of social scientists that people rarely change their political preferences is therefore true.

Now that it is clear that a majority of people vote for the same party over all seven waves of the survey, the focus of the analysis shifts to the effect gender has on the consistency of an individuals voting records. The graph below shows voting records of the sample population according to gender. The x-axis shows the two categories of voting consistency and the y-axis shows the percentage of each gender that fall into each category. This graph uses the percentages as there is a different amount of men and women in the sample.

If gender does have an effect on the consistency of an individual's voting records, then we would expect the number of people in each category to be drastically different for men and women. If gender does not have an effect, we would expect the bars for men and women to be roughly the same height.

```
voterecM <- voterecM %>%
  mutate(perc = (n/(sum(voterecM$n)))*100)
```

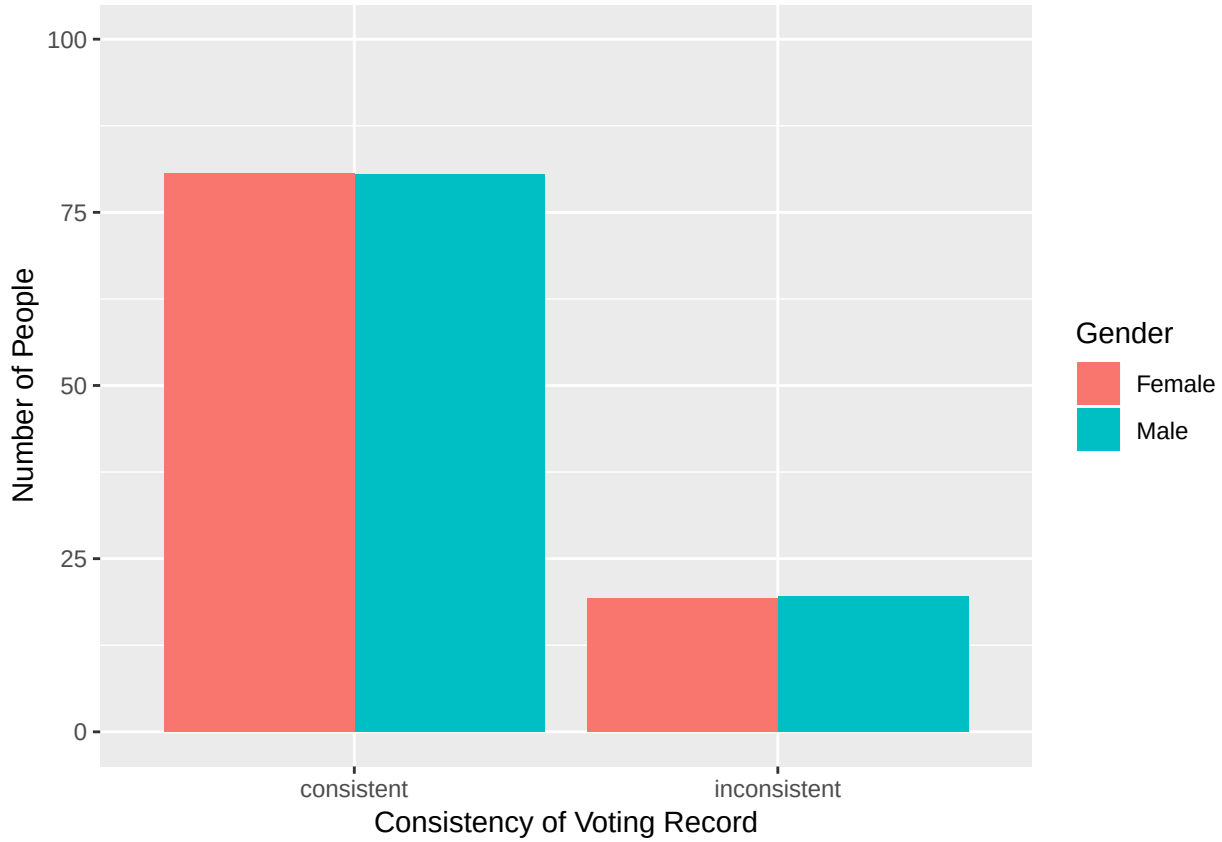
```
voterecF <- voterecF %>%
  mutate(perc = (n/(sum(voterecF$n)))*100)
voterecM <- voterecM[-c(3),]
voterecF <- voterecF[-c(3),]

VR <- c("consistent", "inconsistent")
Gender <- c("Male", "Male", "Female", "Female")
n <- c(80.5, 19.5, 80.7, 19.3)
Gendertab <- data.frame(VR, Gender, n)



Gendertab %>%
  ggplot(aes(x=VR, y=n, fill= Gender))+
  geom_col(stat = "identity", position = "dodge") +
  ylim(0,100) +
  xlab("Consistency of Voting Record")+
  ylab("Number of People")
```

```
## Warning: Ignoring unknown parameters: stat
```



80.5% of males and 80.7% of females voted for the same political party in every wave of the Understanding Society survey. 19.5% of males and 19.3% of females voted for more than one political party over the course of the survey. The male and female bars in this graph are almost identical for both consistent voting records and inconsistent voting records. This graph shows that both men and women rarely change their political preferences and therefore gender has no effect on voting record consistency.

The next characteristic that this analysis focuses on is location of residence, more specifically whether living in a rural or urban area has an effect on the consistency of an individual's voting record. The graph below shows the percentage of the sample population that fall into consistent and inconsistent voting categories and is split into those who live in rural and urban parts of the country. We would expect to see a large difference between the bars for rural and urban if location of residency has an effect on voting consistency.

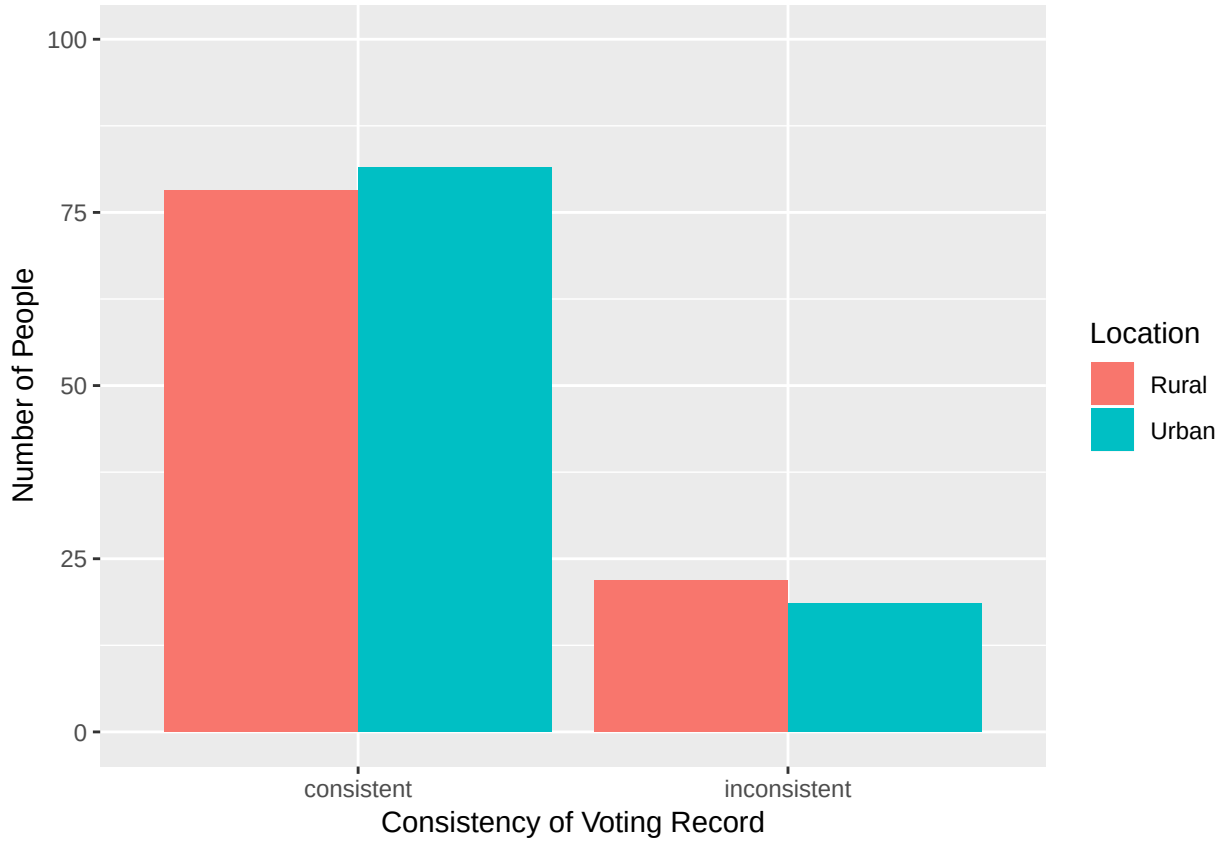
```
voterecU <- voterecU %>%
  mutate(perc = (n/(sum(voterecU$n)))*100)
voterecR <- voterecR %>%
  mutate(perc = (n/(sum(voterecR$n)))*100)
voterecR <- voterecR[-c(3),]
voterecU <- voterecU[-c(4),]


VR <- c("consistent", "inconsistent")
Location <- c("Urban", "Urban", "Rural", "Rural")
n <- c(81.5, 18.5, 78.2, 21.8)
UrbanTab <- data.frame(VR, Location, n)



UrbanTab %>%
  ggplot(aes(x=VR, y=n, fill= Location))+
  geom_col(stat = "identity", position = "dodge") +
  ylim(0,100) +
  xlab("Consistency of Voting Record")+
  ylab("Number of People")
```

```
## Warning: Ignoring unknown parameters: stat
```



The bars are very similar for both location groups for consistent and inconsistent voting records. This graph clearly shows very little difference in voting consistency for those living in rural and urban areas. Overall, this analysis shows clearly that people, in most cases, always vote for the same political party. One explanation for consistent voting records is the concept that people continue to vote for the party they first registered to. This loyalty to the first party registered to is shown in a study carried out by Kaplan and Mukand, who looked at the political preferences of Californians who turned 18 just before or after 9/11 (which caused many people in the USA to join the Republican party). They found more people born after 9/11 chose to join the republican party and a majority of them chose to vote Republican in 2006 and 2008. This study outlines how even though one event caused people to register for a political party, they then continued to vote for the same party log after the event. This loyalty to first registered party could be due to the fact that once registered to a particular party, the individual will receive plenty of propaganda supporting said political party. This constant exposure to propaganda from the political party may promote confirmation bias whereby people use information to conform their existing views.

Individuals are very resistant to changing their political preferences, Kingerlee states than this may be due to the fact that intrinsic factors influence an individual's voting choices. According to Kingerlee, biology, neurology, cultural conditioning, motivation and personality all play a role in an individual's voting decision. Due to the intrinsic nature of these factors it in unlikely they will change and thus unlikely an individual will change their political preferences.

Another reason for the voting consistency displayed by a majority of people could be that people chose political parties because they feel their own views and opinions align with those of the party (Krosnick, 1988). Peoples world views rarely change drastically and therefore if their world view aligns with one particular party then they will continue to vote for that party as it is in their best interest.

In conclusion, individuals usually vote for the same party in every election and are very resistant to changing their political preferences. This study showed that an overwhelming majority of those who participate din the Understanding Society survey voted for the same party in all seven waves. This study endeavoured to

determine the effects of location and gender the political consistency of the participants. There was little difference in the political stability of men and women, therefore it can be concluded that gender has no significant effect on the political stability of individuals. The same is true for urban and rural location, little difference was found. In the analysis of both characteristics it is clear that a majority of people vote consistently. The findings of this study mirror the well documented findings prevalent in social science. Many reasons outlined above attempt to explain this phenomenon, namely, the study carried out by Kaplan and Mukand described above. To build on this study in the future social scientists could look at other characteristics and determine their effect on political consistency. One particularly interesting characteristic to study would be age, as it is a common belief that people become more conservative with age. Another interesting variation of this study would be to determine whether specific political parties promote more political consistency.

Works Cited

Kaplan, E. and Mukand, S., 2011. The persistence of political partisanship: Evidence from 9/11. Mimeo, University of Maryland.

Krosnick, J.A., 1988. The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. Journal of personality and social psychology, 55(2), p.196.

The Psychology of Changing Political Viewpoints, 2013, Alan Martin, https://www.wired.co.uk/article/changing-political-beliefs Dataset Documentation, Understanding Society, https://www.understandingsociety.ac.uk/documentation/mainstage/dataset-documentation

https://github.com/kelseigh399/Final-Report