Problem 1

We'll start off with questions over just the elections data. Which riding in Quebec has the most registered voters? Which has the fewest?

I summarized the data in the registered votes column and found the maximum value. Using this value, I filtered through the election data to find the riding that matched this population value. The riding that has the most registered voters is Brome-Missisquoi with 66769 people. I found the riding with the fewest registered voters the same way but with the minimum value instead. The riding that has the fewest registered voters is Îles-de-la-Madeleine with 11159 people.

Problem 2

What riding had the highest voter turnout among registered voters?

I summarize a column with percentage of the total votes out of the registered voters and found the maximum value in the new column. Then, I filtered through the election data to find a riding that matched this value. The riding that had the highest voter turnout was Louis-Hébert with 81.09% of their registered voters turned up.

Problem 3

Now let's see if we can get an overall picture of the results of the recent election. How many seats did each party win? What was the breakdown of the popular vote?

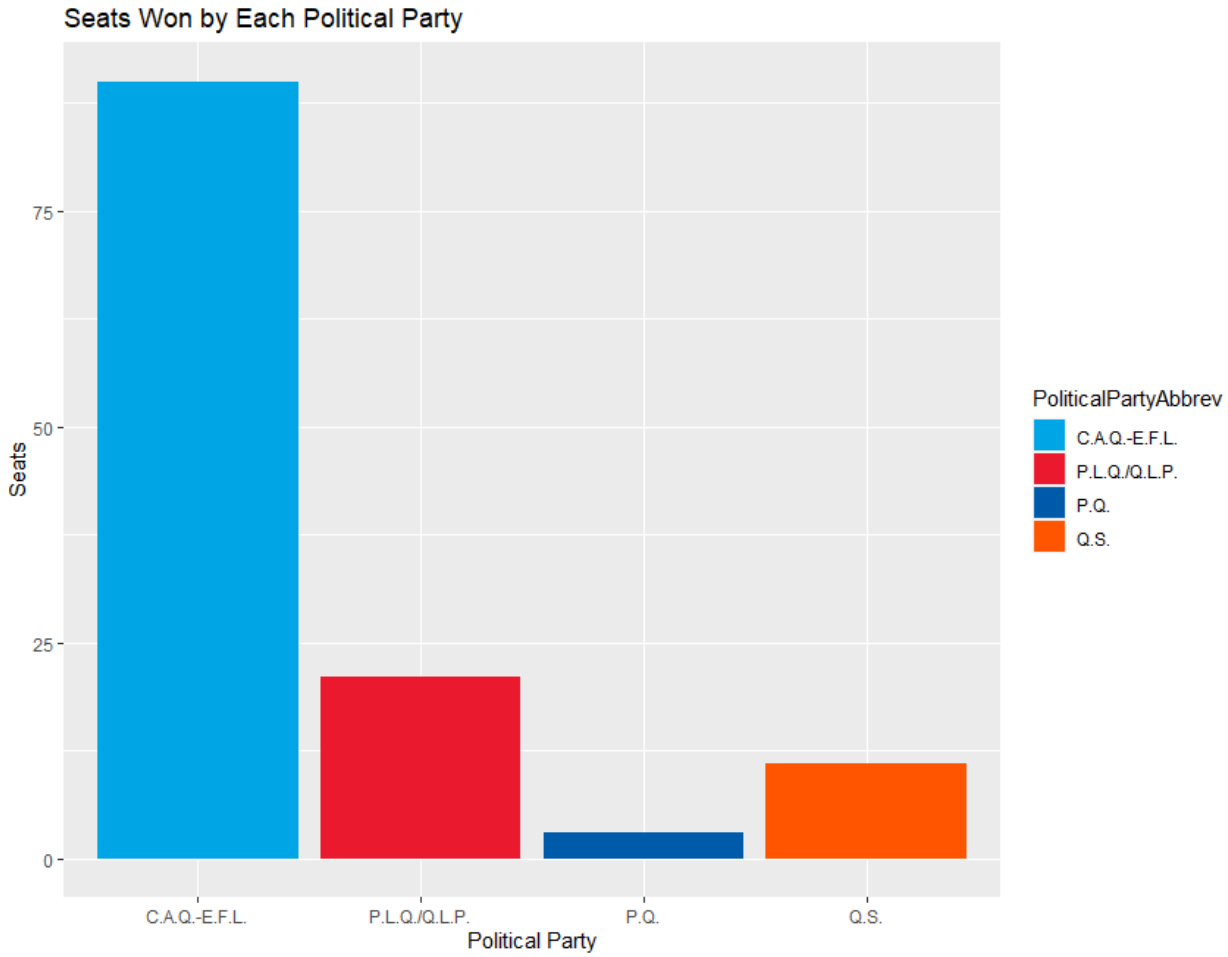
In order to find how many seats each party won, I filtered the election results so that only the winning parties remained. I grouped by political party and counted how many times each political party won.

Political Party	Seats Won
C.A.Q.-E.F.L.	90
P.L.Q./Q.L.P.	21
P.Q.	3
Q.S.	11

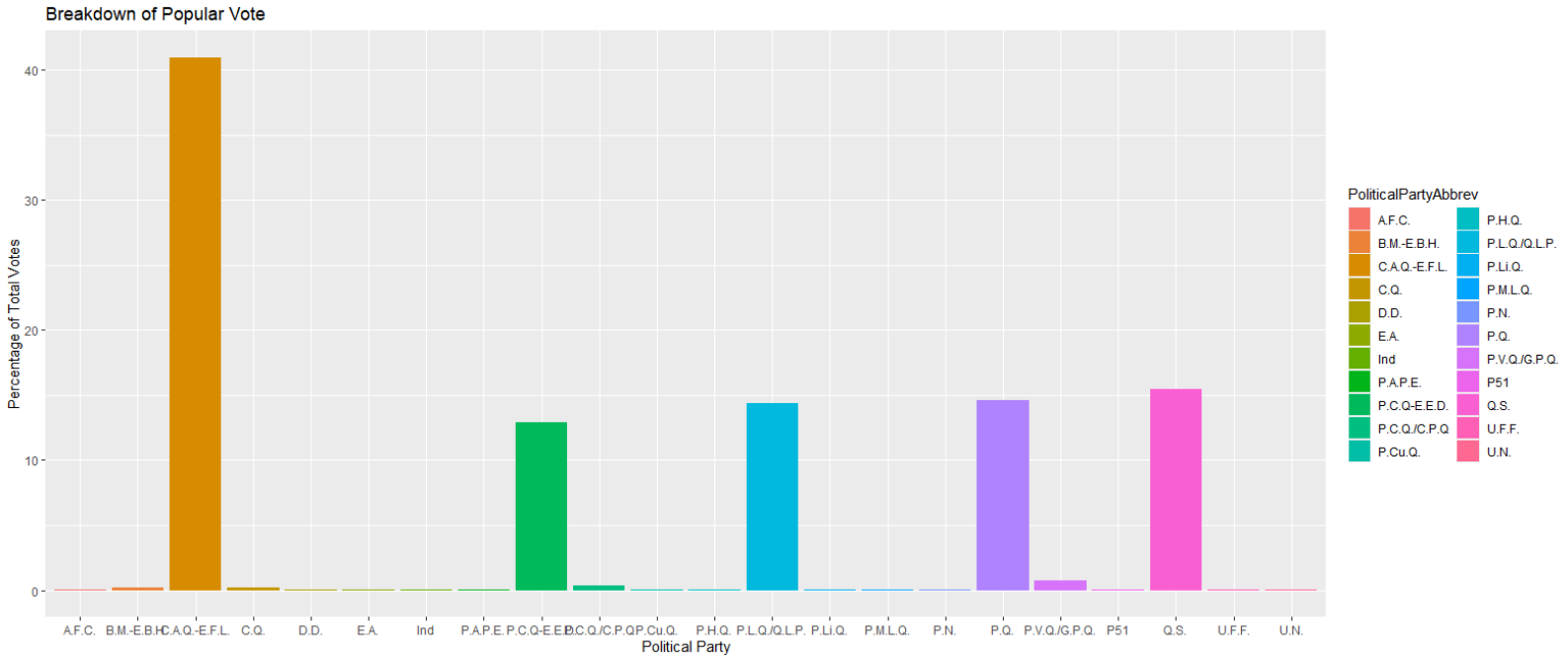
Then, I calculated the breakdown of the popular vote by finding the percentage of votes for each political party out of the total votes in the election. This table is called votesPerParty. The C.A.Q.-E.F.L. party had the highest percentage of votes.

Problem 4

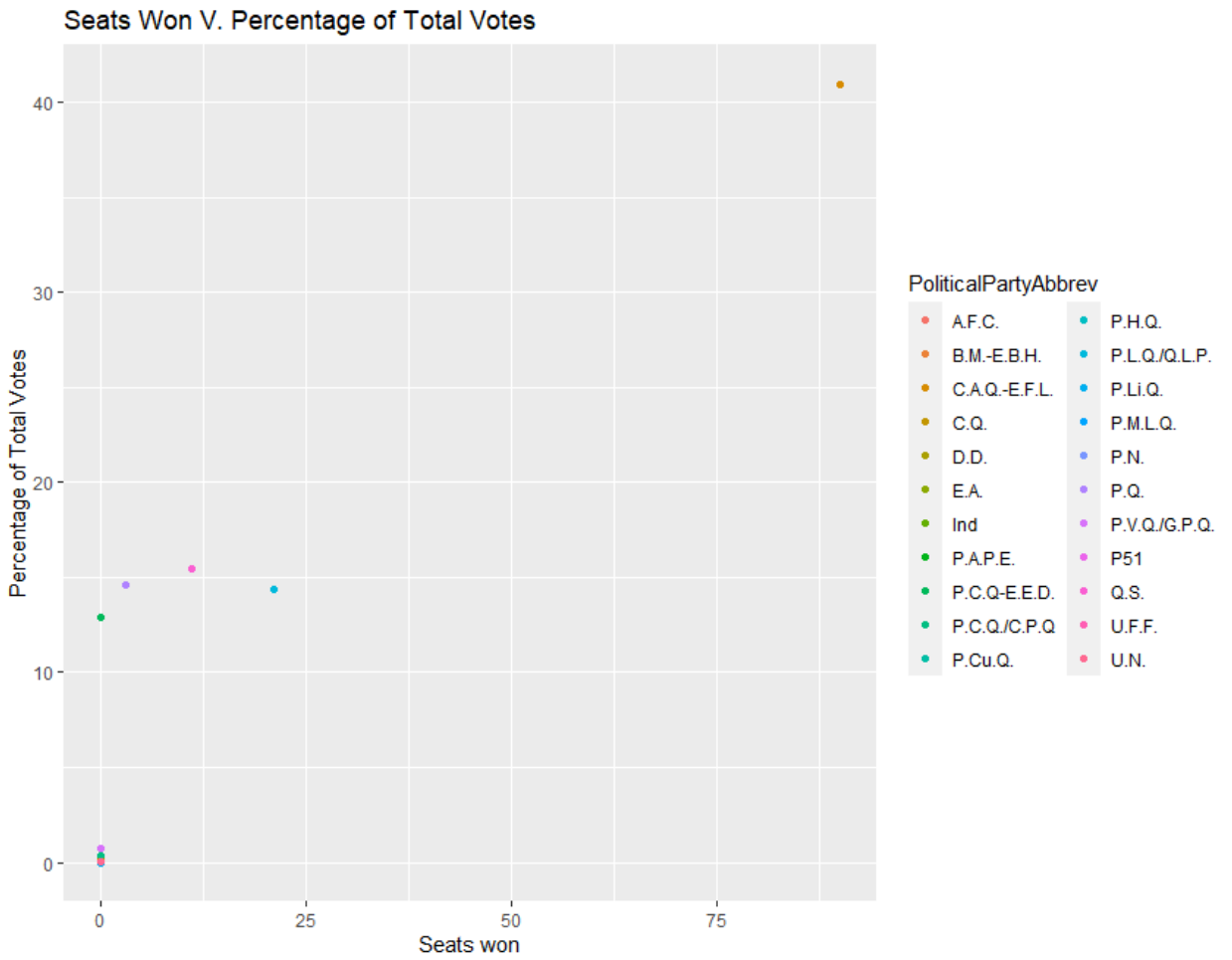
Create a plot/plots for both the popular vote and seat totals. Style points if you can convey both things in the same plot in a way that is both visually pleasing and informative?



I created this graph using the data frame from the previous question for seats won by each party.



I created this graph by using the data frame previously created in question 3 for the breakdown of the popular vote.



I combined both of the table with the x-axis being seats won by each party and the y-axis being the percentage of the total votes by each party. You can see that the P.C.Q.-E.E.D. party got about 13% of the popular vote but did not win any seats. Another party, P.L.Q./Q.L.P., got less of the popular vote than both Q.S. and P.Q. but won more seats.

Problem 5

Quebec Solidaire and the Parti Québecois are somewhat similar ideologically – so much so that there was a proposal for the parties to form an alliance in the 2018 provincial election. Although this proposal failed, it can be fun to speculate what might have been. Imagine that Quebec Solidaire and the Parti Quebecois had run as a single party in the most recent election. Imagine that their single candidate in each riding received as many votes as the sum of the two parties' votes in the actual election that took place. How many seats would this united party have won?

First, I recalculate the votes if both parties were combined. I named the name party the United party. Then I figured out which party would win in each riding with the new United party. I then created a similar data frame to the data frame I made for problem 3.

Political Party	Seats Won
C.A.Q.-E.F.L.	85
P.L.Q./Q.L.P.	21
United	19

The United party would win 19 seats.

Problem 6

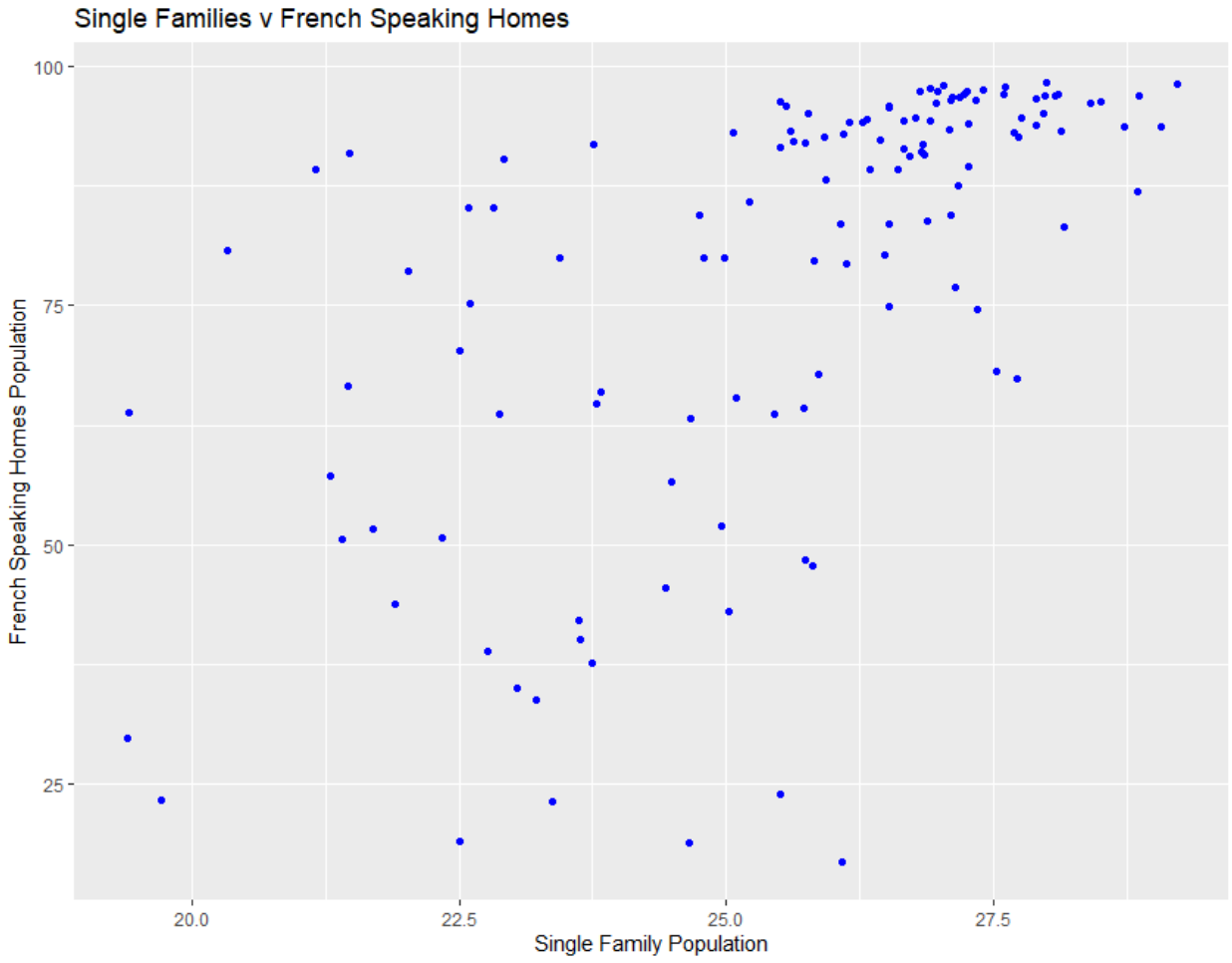
Now let's move on to census data. What is the youngest riding, measured by mean age? What is the oldest?

To start off, I pivoted the age data frame to long format, filtered so the constituency is "Mean Age" and used the summary function to find the lowest of all of the mean ages. The riding with the youngest mean age is Ungava with a mean age of 32.6 I found the oldest riding the same way but looking for the greatest of all the mean ages. Two ridings have the same highest mean age, Gaspé and Îles-de-la-Madeleine, with a mean age of 49.1.

Problem 7

Make a plot comparing the proportion of households that are single-family homes within a riding to the proportion of the population of the riding speaking French at home. Do single-family homes seem to correlate with a riding being more or less Francophone?

I took the house hold data, pivoted longer, and filtered so it was only the data for house holds with a single family and no additional people. Then I took the language data, pivoted longer, and filtered so the data was for French speaking households. I joined the two table together, and added two columns. One is for the percentage of the riding that lives in a house hold with a single family and no additional people. The second column is for the percentage of the riding which speaks French at home.

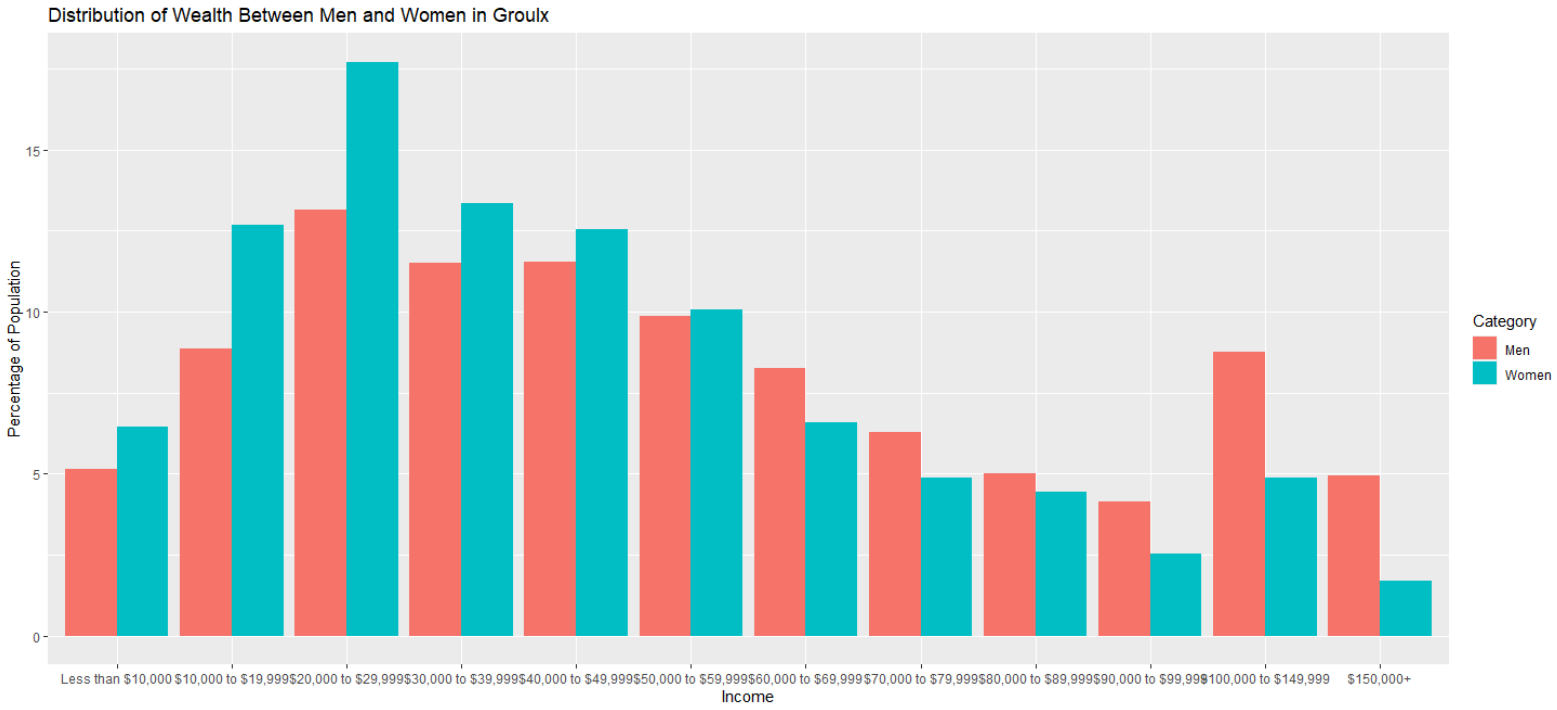


I made this graph using the joined table with the x-axis of single families and the y-axis of french speaking homes. This graph shows that there is a correlation between if a household is a single family and if the household speaks French. If a household is a single family, it is more likely to speak French at home.

Problem 8

Make a plot showing the distribution of income among men and women, side by side in any district. Mention the district name in the plot.

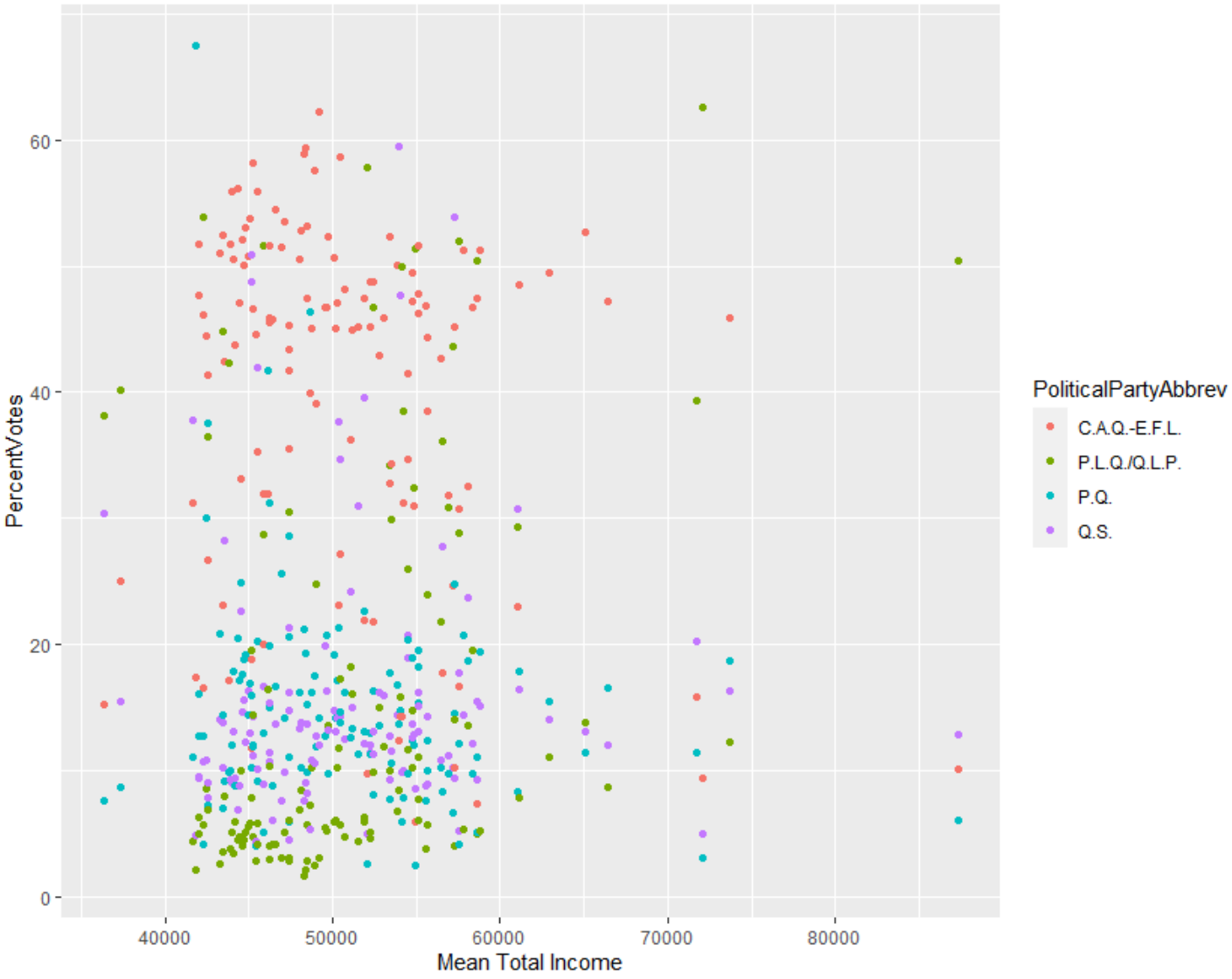
The riding I picked to demonstrate the distribution of income among men and women was Groulx. I separate the income data into a women's and men's table. In each table I added a percentage column which gives a percentage of the women who earn a specific income out of the total women who earn an income. I then joined both of the men and women tables together and displayed it in a bar chart, with two bars for each level of income, one showing the percentage of women who earn that income, the other showing the percentage of men who earn that income.



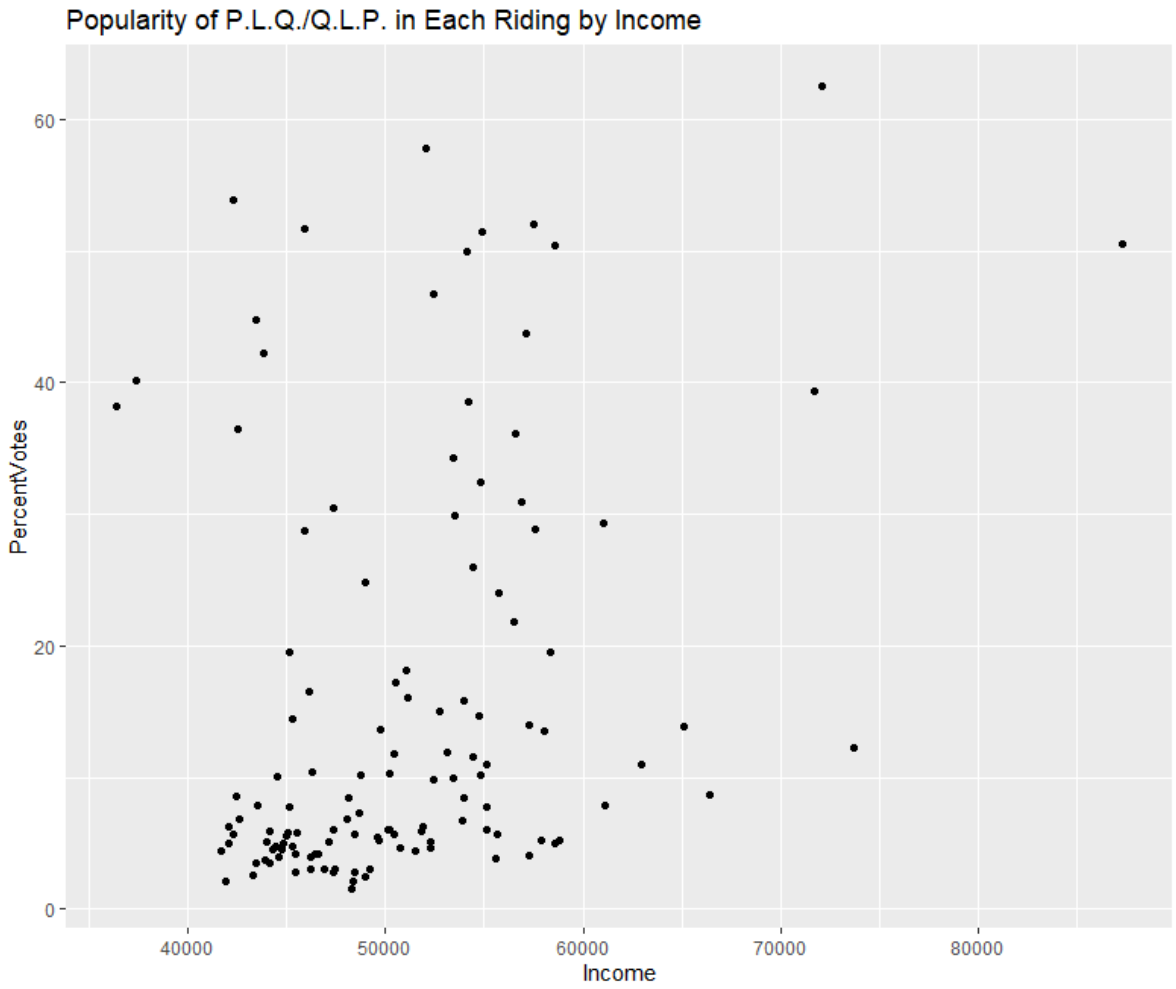
Problem 9

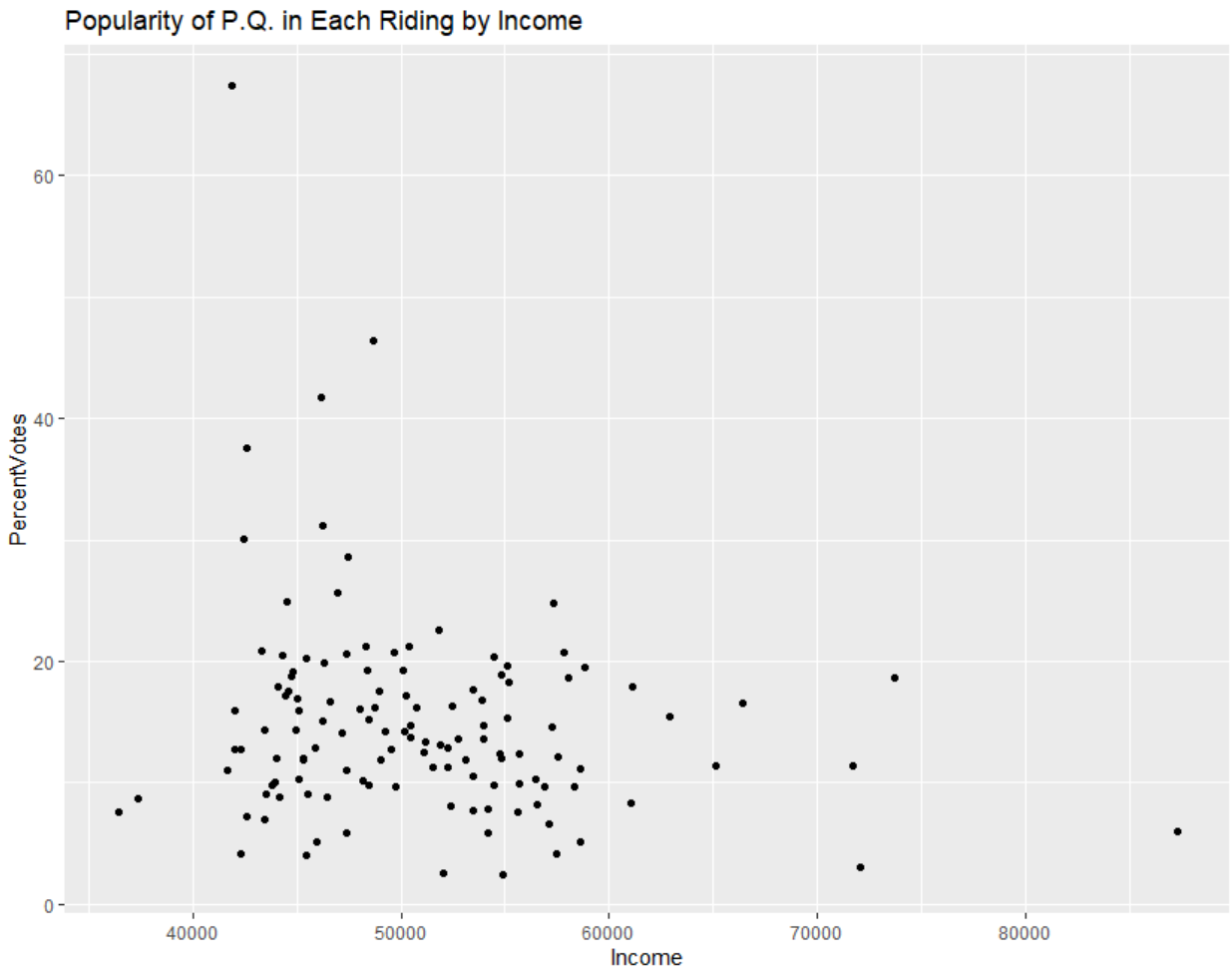
Now for questions that involve both politics and demographics! Often times political parties intersect largely with class. Taking a look at the income data supplied. Is there a party that seems to particularly appeal to richer voters? Is there a party that seems to particularly appeal to poorer voters?

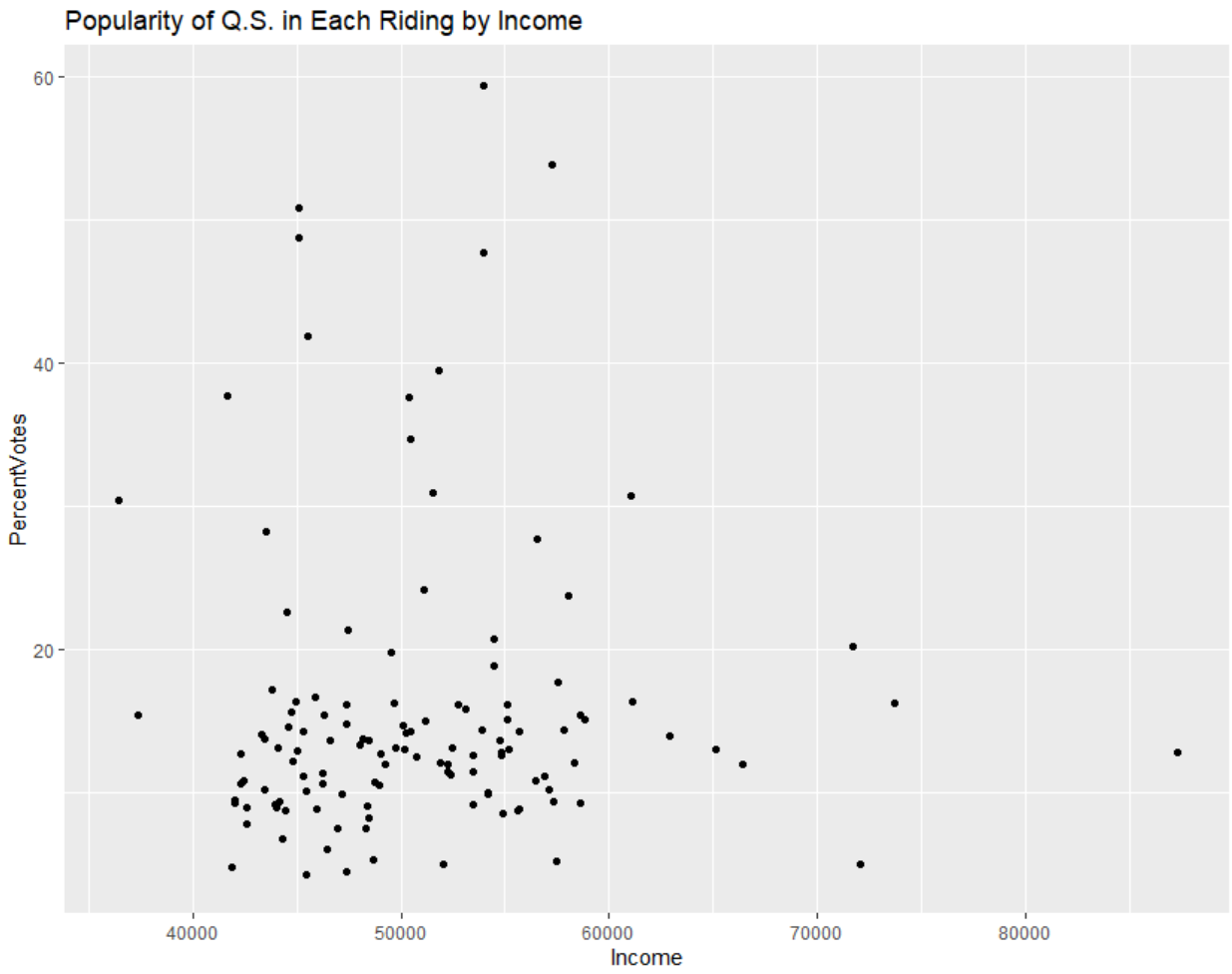
First, I took just the mean total income of each riding from the full income data table. Then I created a table that took the percentage of votes for each party in each riding to show each riding's political party preferences. I joined the table with the mean incomes and the table for each riding's political party preferences. I then used this new joined table in a point graph with a x-axis of total mean income and y-axis of percentage of votes for each party.

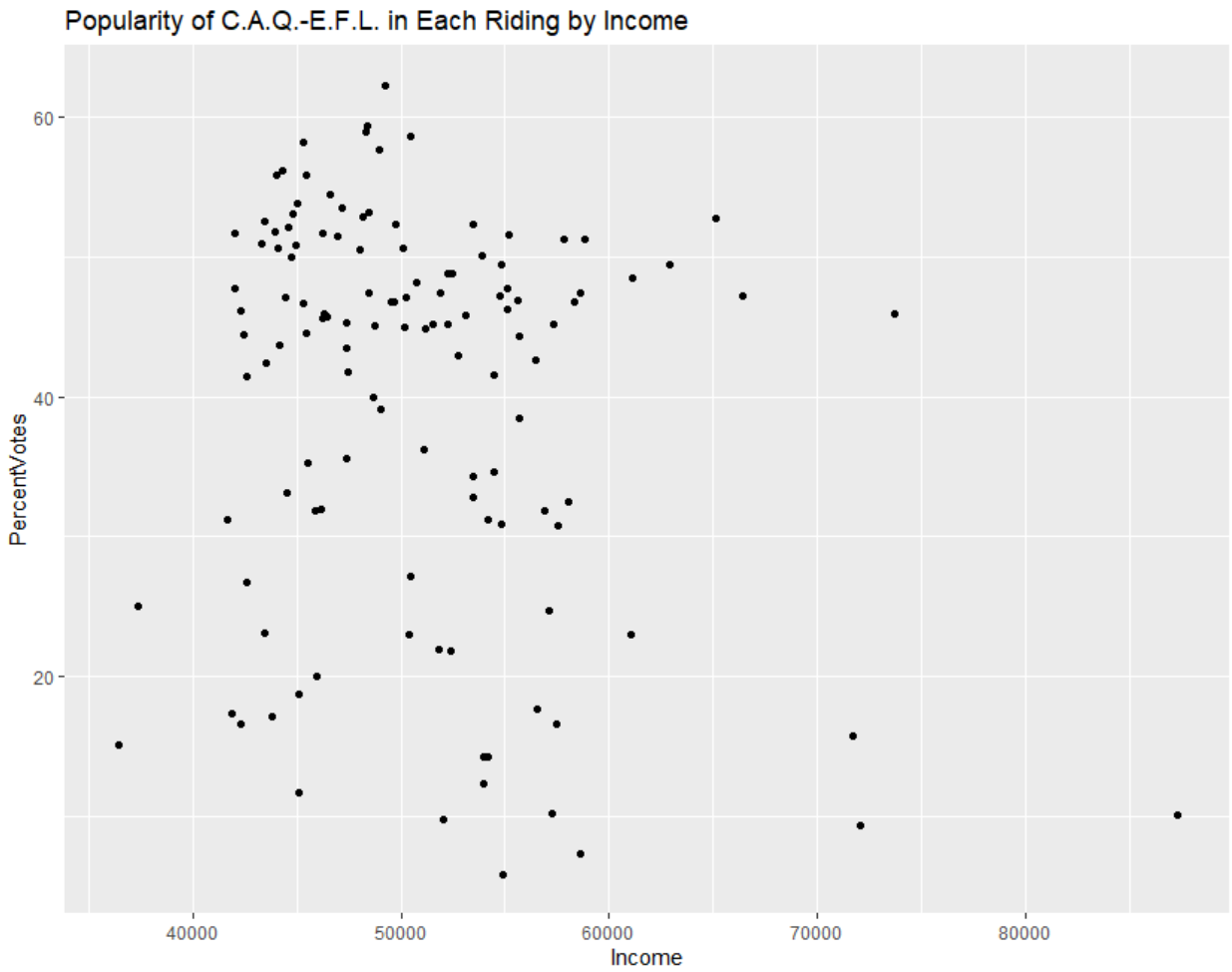


This is the graph for all 4 major parties. The next 4 graphs are for each of the major political parties.


Popularity of P.L.Q./Q.L.P. in Each Riding by Income
PercentVotes
0
20
40
60
40000
50000
60000
70000
80000
Income


Popularity of P.Q. in Each Riding by Income
PercentVotes
0
20
40
60
40000
50000
60000
70000
80000
Income


Popularity of Q.S. in Each Riding by Income
PercentVotes
60
40
20
40000
50000
60000
70000
80000
Income

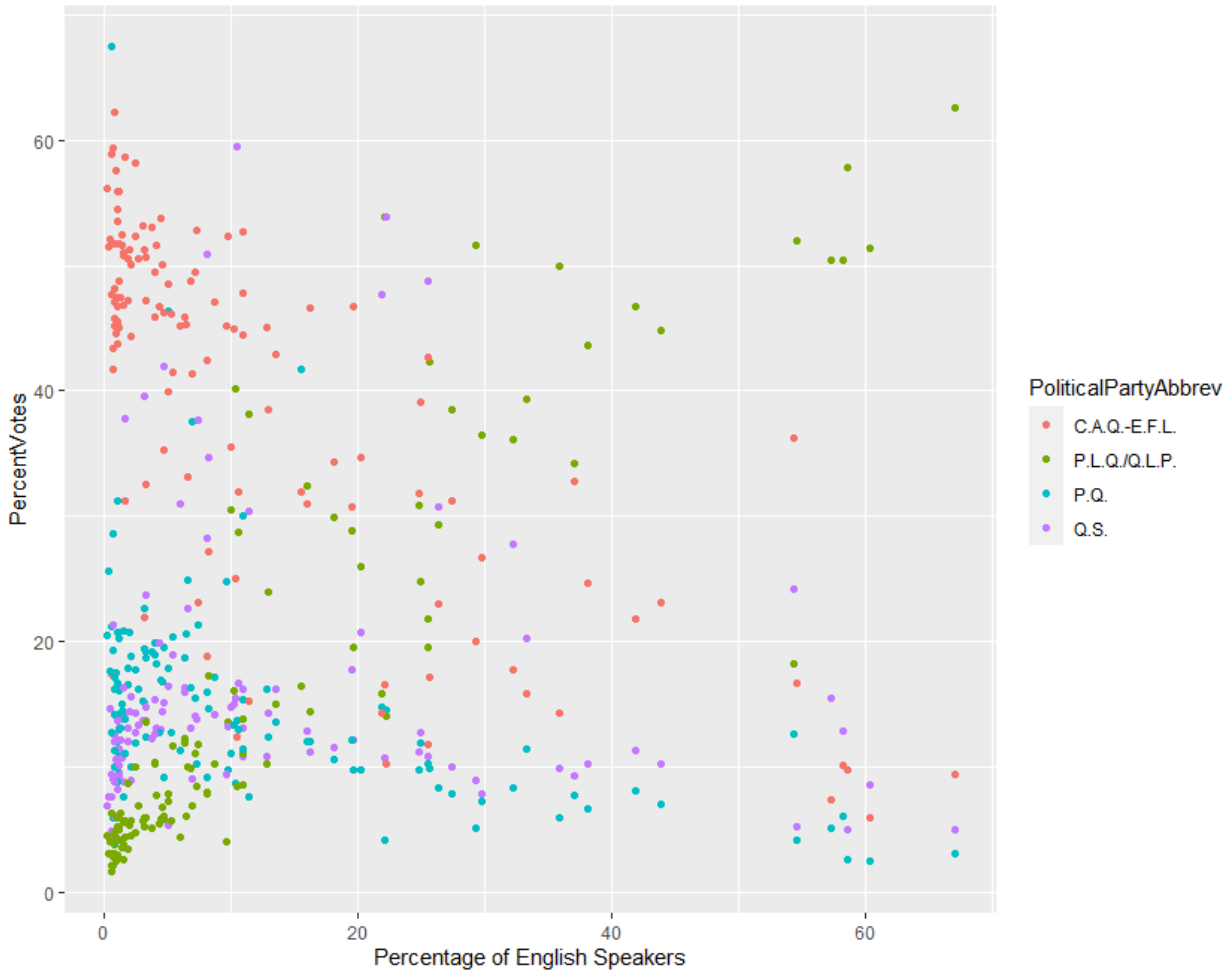


After analyzing all of these graphs, there is no clear winner in either of the poorer or richer distinct, but it seems as if the P.L.Q./Q.L.P. party appeals to the two poorest and the two richest districts.

Problem 10

There is one party that is often said to mostly represent the English-speaking minority in Quebec. Which party do you think this is? Why? You can come up with any way to test this – visual, numerical, etc. But make sure to explain and defend your idea!

I took the language data and filtered it so only the data where the primary language is English remains. I joined the English data and the population data for each riding, joined them together, and added a column for the percentage of people whose primary language is English. Then I joined this table and the table for the percentage of votes for each party in each riding. The second of these joined tables was used to make a graph showing the correlation.

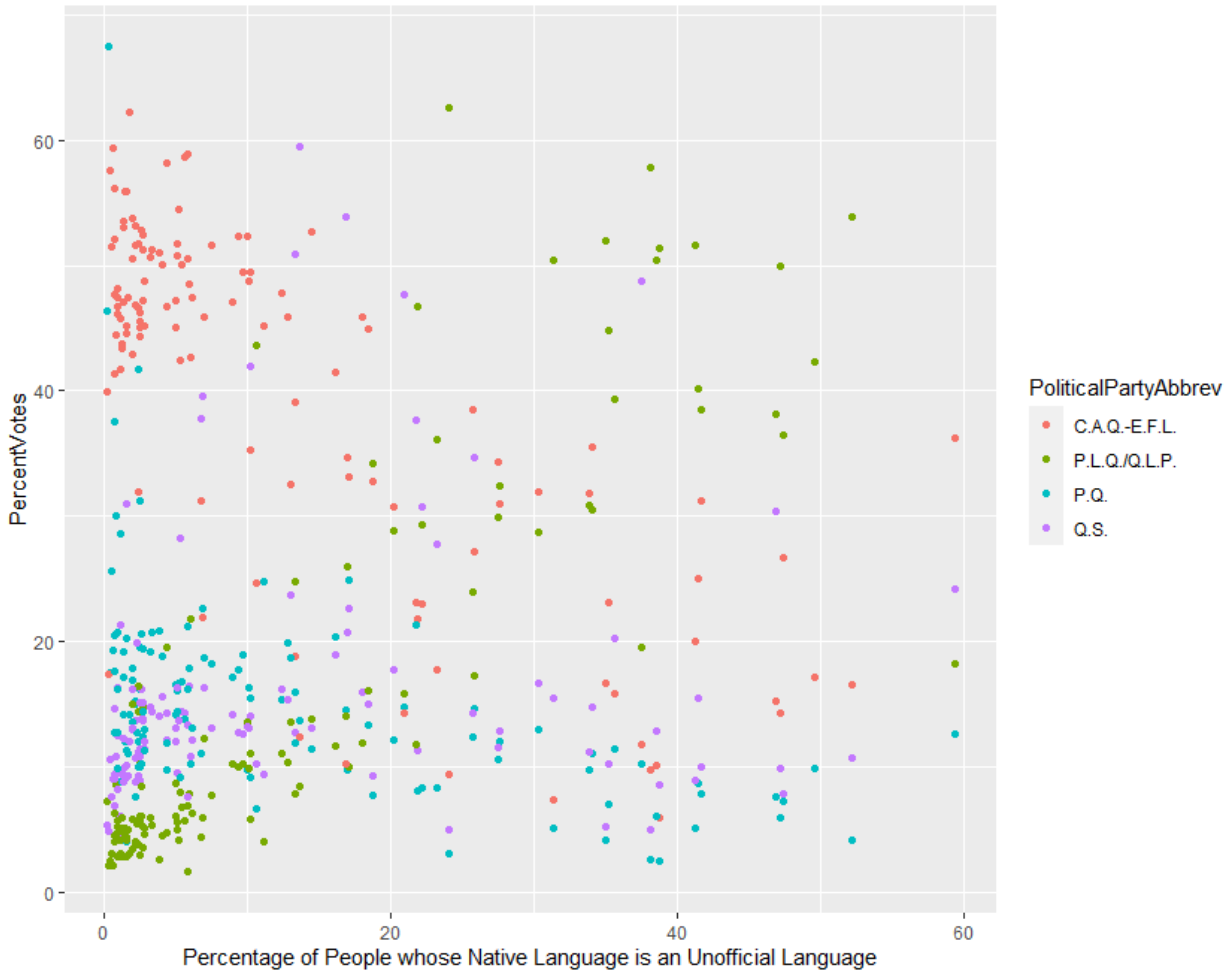


This graph clearly shows that there is strong correlation between the English speaking population and their preference for the P.L.Q./Q.L.P. party. People whose primary language is English are most likely to vote for the P.L.Q./Q.L.P.

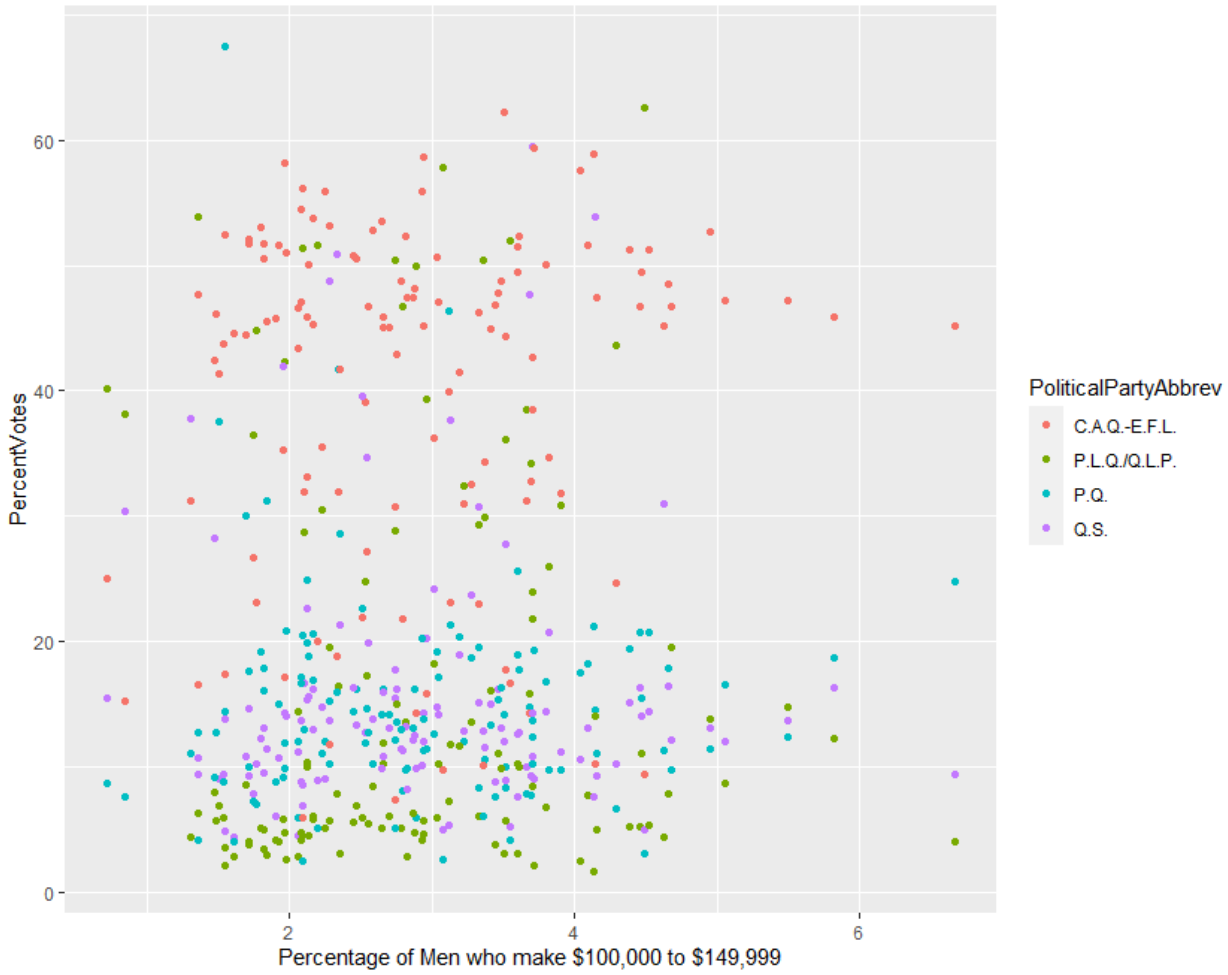
Problem 11

A man approaches you and gives you a challenge: His native language is neither English nor French. He makes $100,000 a year. He lives in an apartment building with more than 5 floors. He is 29 years old. He asks you to guess what party he voted for based on these facts. What would be your guess and why?

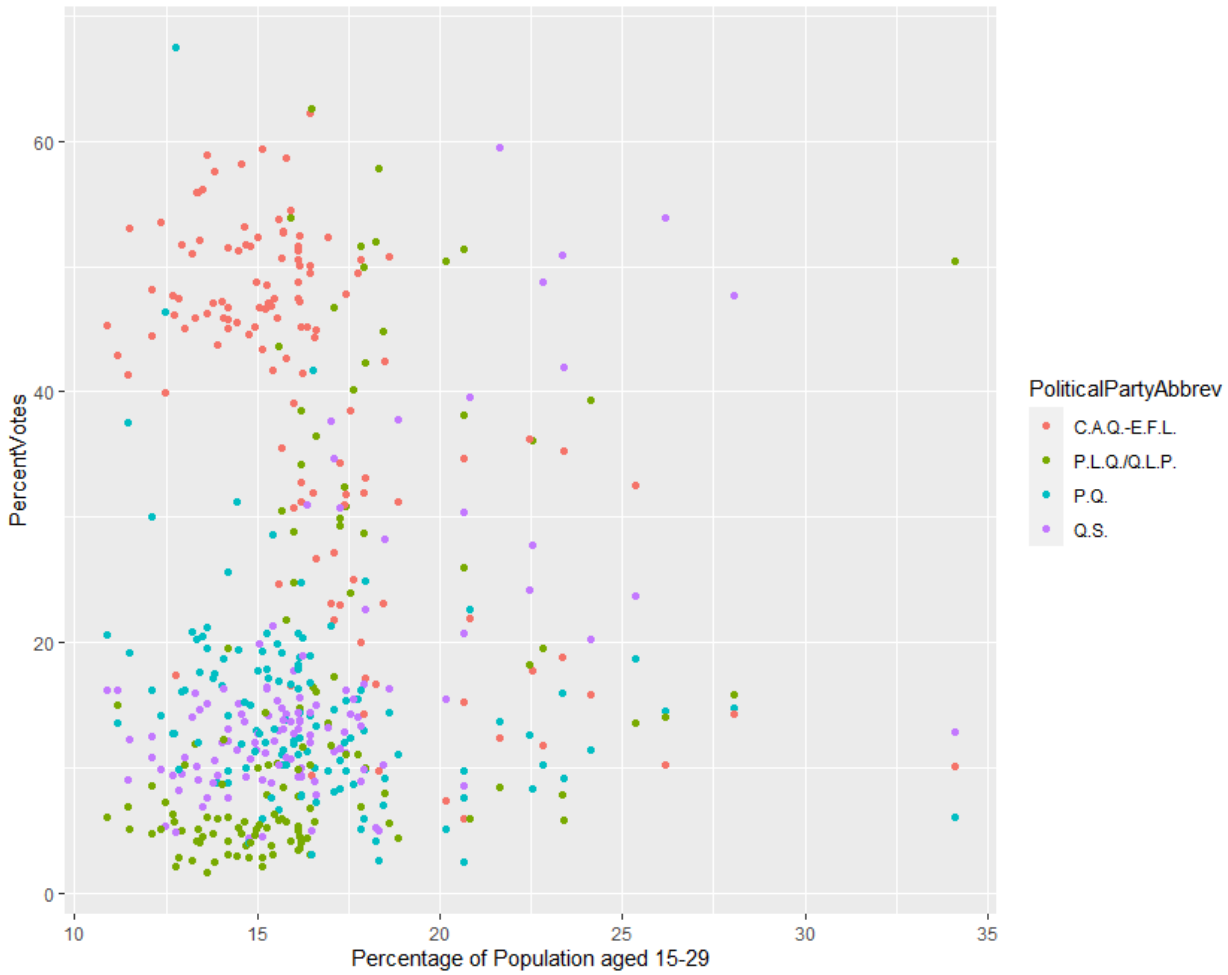
I approached this problem by making 4 graphs. Each graph had a x-axis with the percentage of each riding that had his specific housing, income, language, age, or gender. The y-axis of each graph was the percentage of total votes for each political party in each riding.



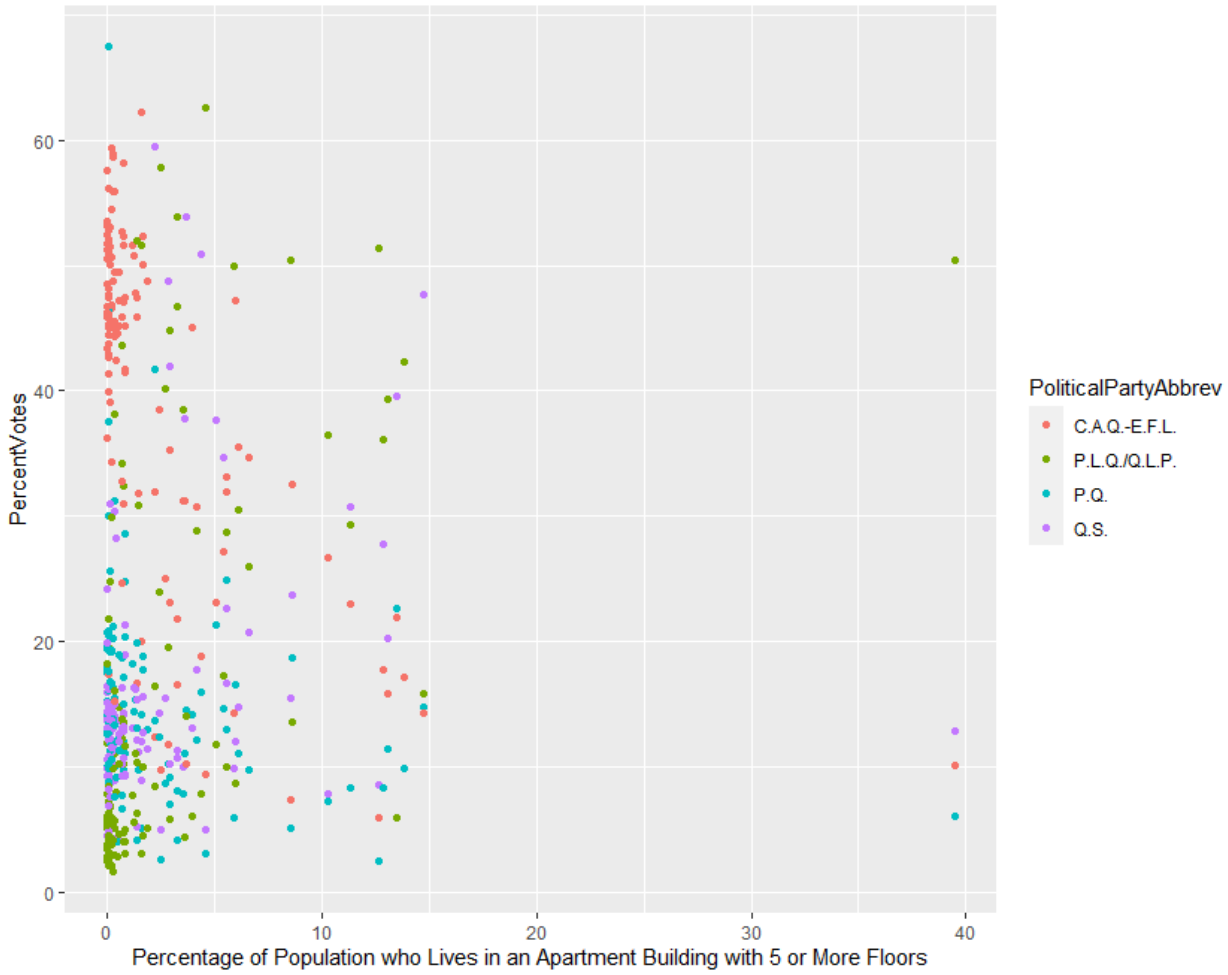
This graph shows a clear correlation between people whose native language is unofficial and voting for the P.L.Q./Q.L.P.



This graph does not have a clear correlation between men who earn $100,000 to $149,999. It seems that the C.A.Q.-E.F.L. seems the most popular over this entire group of people.



In this graph, its as though when the percentage of the riding that is aged 15-29 gets higher, the percentage of votes for the P.L.Q./Q.L.P. party also gets higher.



In this graph, it seems that the percentage of votes for the C.A.Q.-E.F.L. definitely goes down when the percentage of the population of people who live in an apartment building of 5 or more floors goes up. We can rule out this party from our considerations. After looking at all of the graphs, I would guess that this man voted for the P.L.Q./Q.L.P. party.