

Mini-Project #3  
Stat 212: Interim 2022

### **Introduction**

In this research we have focused on two questions “Is the proportion of residents aged 34 years or younger who think Northfield values culture the same as the proportion of residents over 34 years old sharing the same opinion?” As well as “Is the proportion of residents who agree that there is access to housing for people of various races significantly larger than 50% (a significant majority)?” We focused on the NREEC survey for our data. This survey was given to the residents of Northfield to fill out as a way to gather data about their perceptions and experiences. The survey was accessible through 20 organizations and it was also broadcasted through social media and posters with QR codes to get as many residents' responses as possible.

### **Methods**

Description: The Northfield Racial and Ethnic Equity Collaborative (NREEC) and the city of Northfield came together to create a survey. They sent it out to community centers, Hospitals, Colleges, and Churches. This survey sample also listed age, gender, race, etc. This allows us to look at specific demographics. We looked at age and housing in Northfield. We hope to reject our null hypotheses listed above. These demographics allowed us to do that.

Explanation: We created categories for age. We made two categories: residents under 34 year old and over 34 year old. We did this in order to look at the differences between the older and younger aged participants. The reason we changed this data is because we wanted to look at the relationship between culture and age. We thought that 34 was a good cut off. This is a little low, but in Northfield the demographic is mostly college students and we wanted to capture that part of the population. We also looked at housing and race. For this we looked at residents that were white and non-white and compared their thoughts on housing. In order to do this we change the data for race into those categories. Additionally we had to combine some age categories in order to conduct a Chi-Squared test of independence, there would have been a better outcome if we were able to include all age groups separately to have a fuller picture, however to do limitations of Chi-Squared test we divided age categories into three groups: Under 34, 35-54, over 55.

Tests: We used box plots to illustrate these differences. We are also going to do a chi-squared test for independence in order to compare our observed and expected results. We wanted to test our hypothesis that the proportion of residents who agree that there is access to housing for people of various races is not larger than 50% and The proportion of 34 yr olds who think Northfield values culture is the same as the proportion of over 34 years olds. We used these bar plots to evaluate these hypotheses. We ran “difference in proportions” for the proportion of 34 yr olds who think Northfield values culture is the same as the proportion of over 34 years olds. Then we ran a “proportion” for the people who believe that the access to housing varies by race.

These tests allowed us to reject and accept those hypotheses (more in results). The limitation of the Chi-Squared test of independence is that we cannot have a count lower than 5. That is the reason we combined all of the age groups into three. In the current version of the table all of the counts have a value of 5 or higher. We ran the Chi-Squared test to determine if there is difference in responses between those three age groups.

## Results

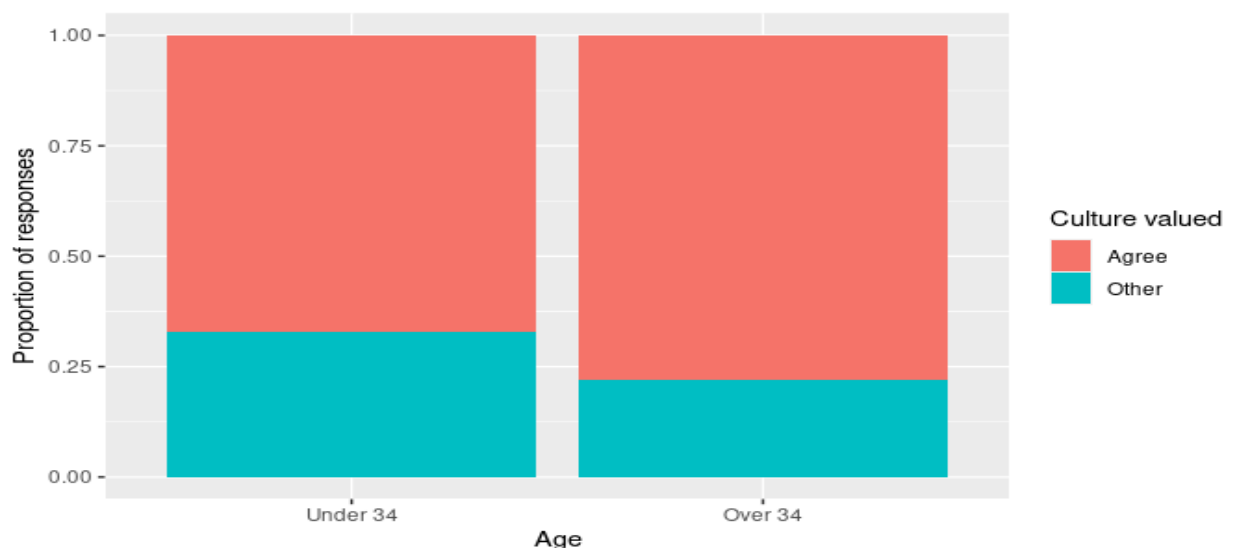
### EDA

#### Q1:

This data set shows the relationship of the age of Northfield residents and how much they feel culture is valued. From the bar plot we can see that the age group of 34 or older feel that they agree that Northfield values culture while those 34 or under disagree with it more. Overall there seems to also show that there are more 34 or older responses which may show a reason for the differences in the responses. But we can see that generally the age group of 34 or older and 34 or younger are not the sample proportions and more 34 or younger residents think culture isn't valued while more 34 or older residents do feel like culture is valued. There were a total of 252 responders under 34 years old and 619 responders aged over 34 years old. At the same time there were 652 respondents who agreed that Northfield values their culture and 219 respondents who have any opinion different from that (disagree and neutral).

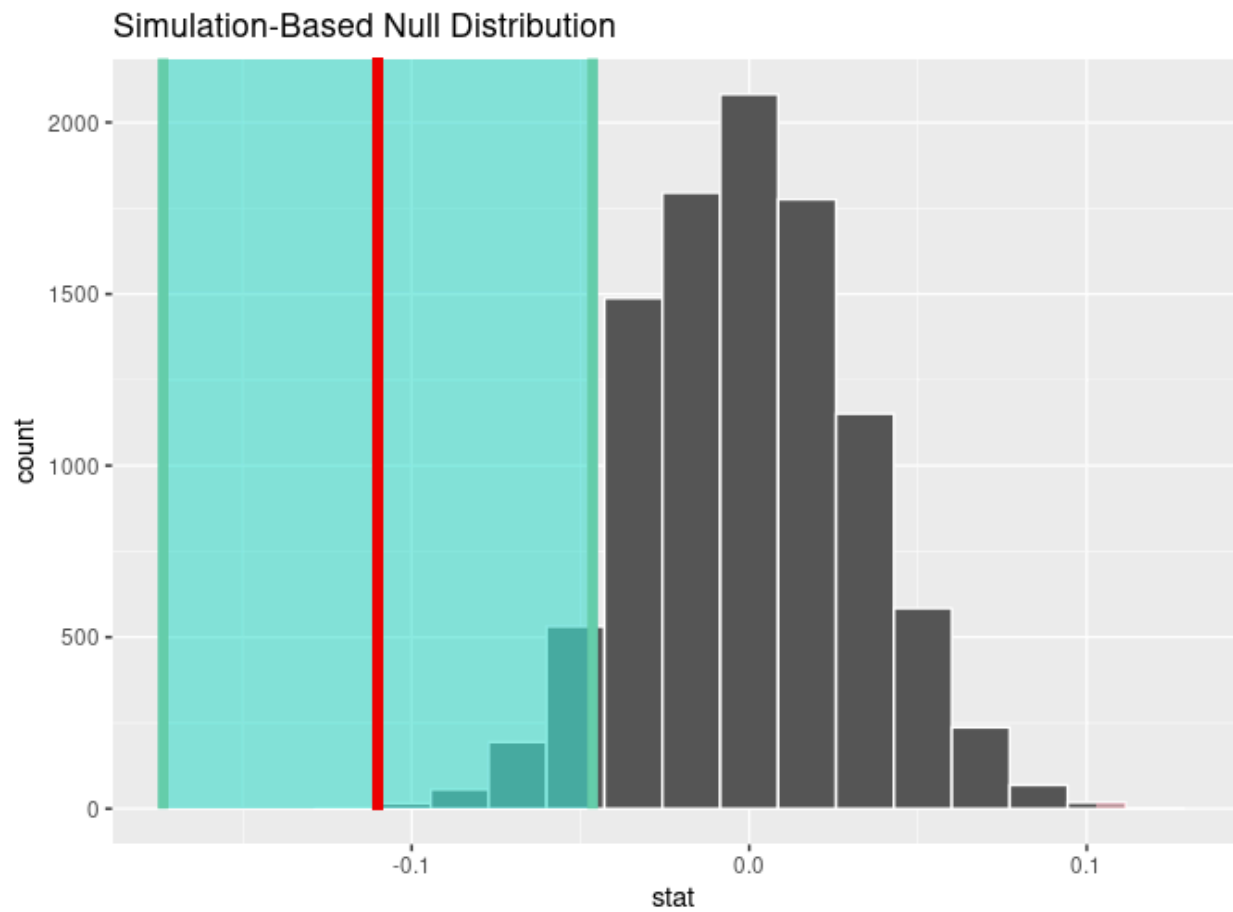
Under 34	Over 34
252	619

Agree	Other
652	219



P-value: 0.0006

Hypothesis testing with confidence interval of 95%:



Chi Squared test of independence:

Age	Area					
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	
under 34	75	94	57	21		5
35-54	133	115	49	22		7
over 55	132	103	38	13		7

X-squared = 20.473

p-value = 0.008687

According to the randomization test we are 95% confident that the actual difference between the proportions of Northfield residents aged under and over 34 is in the interval between -0.174 and -0.046 (17.4% and 4.6% where residents over 34 have higher proportion of residents who think their culture is valued).

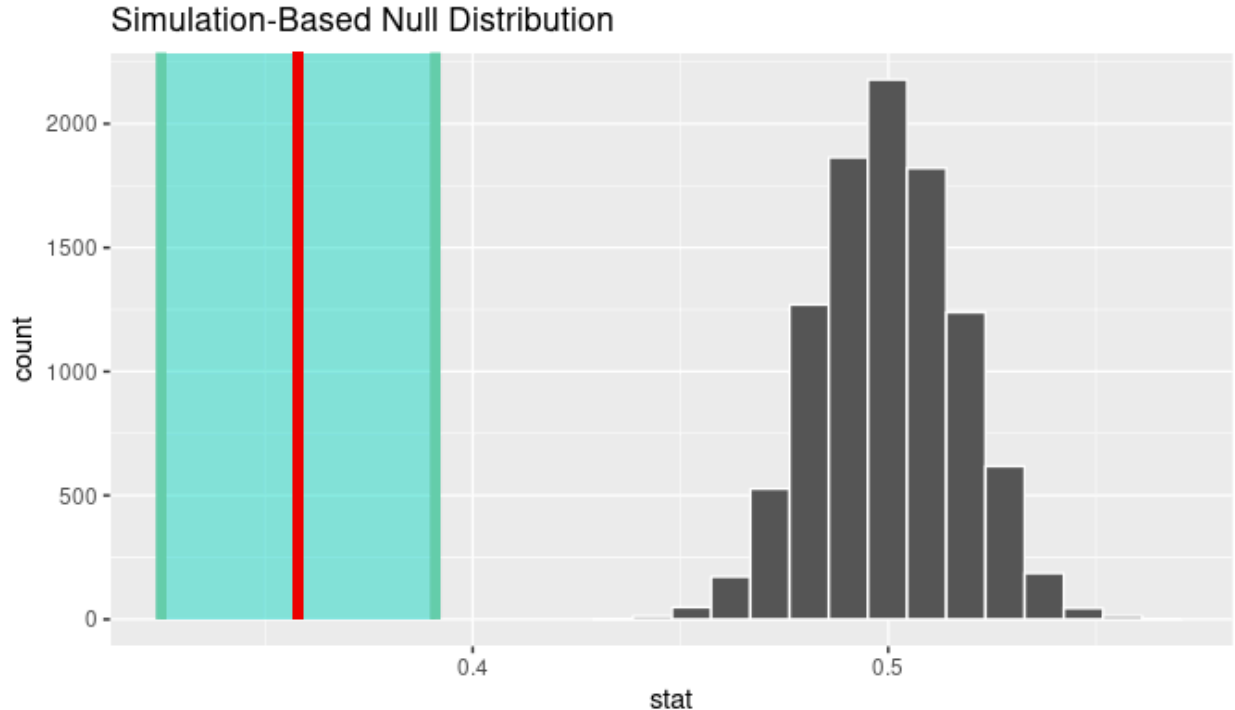
## Q2:

The bar plot shows us that more residents of Northfield feel that the housing offered in the area is not offered to various races. Though the survey originally has 5 categories "strongly disagree" "agree" "neutral" "disagree" and "strongly disagree" they're split into 2 categories, of just "agree" and "disagree", taking out the "neutral" answers. Through this bar plot and the data table we can conclude that there is a little over half of residents who don't feel that the housing offered is open to various races but it isn't a significant amount more who feel they disagree. There were 312 responders who agreed with the statement and 559 responders who had different opinions (disagree and neutral).

Agree	Other
312	559

p-value = 0

Hypothesis testing with confidence interval of 95%:



According to the randomization test we are 95% confident that the actual proportion of Northfield residents who agree that there is access to housing for people of various races is between 0.333 and 0.383 (33.3% and 38.3%)

## **Discussion**

The goal of this research was to understand how the residents of Northfield feel about culture and housing in the area and as we can see from the results of the survey for the first research question majority of residents both 34 years and older and 34 years and younger felt that Northfield did value culture. However, the p-value for the test was low enough to conclude that residents over 34 years old agreed to that statement more often than residents under 34 years old. In fact, we were able to determine that we have a 95% confidence that the actual difference between agreement responses of residents aged over 34 and under 34 is between 4.6% and 17.4% which is quite significant. Using the Chi-Squared test we were able to determine that in fact there was also a significant difference between responders aged under 34, from 34 to 55 and over 55 for that statement. The test statistic that we got is big enough to conclude that the strong difference exists between three groups. All of those conclude to the fact that Northfield residents have a different perspective of the city being welcoming based on their age. It is actually a logical result to get as the perception of cultural identity for the younger generation has increased over the last years. Older people have a tendency to be more acceptable towards cultural discrimination compared to younger generations. We can also see in the second set of data for the second research question that many of the residents felt that various races were not given equal opportunities to housing in the area. Using randomization we came to the conclusion that we are 95% sure that the actual value of Northfield residents who agreed that there are equal housing opportunities for people of all races is between 33.3% and 38.3%, which allows us to say that the majority of residents do not agree with that statement. Despite all of the results there are several limitations for this study. Though the survey itself is done as a retrospective observational study, which means we cannot make causal inferences in the data, nor is it representative of the Northfield population as it was not a random sample but rather voluntary responses we can still see from the population surveyed that many in sample seem to lean one way more than the other in both research questions. From this research and data it can be used to understand more demographics of the Northfield residents such as the disparities of race and ethnicity within the community as well as the difference in opinions of those younger or older such as in the first research question. We can also see that for those who don't feel a gap in housing for people of color would once again be due to the lack of diversity or understanding of the area which would be another factor of the results in the data.

The limitations within this research project starts with the sample being non-representative of the Northfield population. Though the survey is posted and sent out to various areas within Northfield and even open to the public through QR codes and social media because the responses are not randomly sampled but rather consists of volunteers or a group of residents from one area or institution it doesn't reflect the diversity within Northfield. The areas the survey was handed out to also shows an issue as most of the locations focus on communal areas or institutions that not a diverse population of Northfield residents would go to such as churches and educational institutions, and though the clinics may be more frequented not a large or diverse enough population would go there often. Another limitation is with the method itself as a survey, though this allows more voices to be heard anonymously there are still chances of no responses as well as discrepancies in the answers whether due to wording of the survey itself or dishonest answers of participants.

Because this survey was not given to a random sample in future research it would be more effective to give surveys out to a random sample spread throughout various areas of Northfield for equal representation. It would also be more effective if future surveys had more specific detail or a better scale to allow more room for opinion such as rating numerically or adding more space for comments in each section. Another improvement would be specifying questions or terms such as in question one which talks about culture being valued which can be vague so focusing more on what ways it is valued, what types of cultures are valued or where people see culture being valued would be good insight. This also leads to making it more detailed and significant if we can understand the places the participants come from specifically in the area or areas they visit or use often to understand what parts of the city they most frequent and how well they know the city or areas of Northfield to give a more concrete idea of their perspective. Overall the results of these research questions and the survey have been able to give a push in the right direction of trying to understand balance of race and equity in Northfield but future research would be better if there were more focus on details within the communities and a representative sample which can speak to Northfield as a whole.