STA130 Course Project Proposal

Introduction: In modern societies, the distance between people seems to grow larger and larger, and thus more people are feeling lonely. What causes people to feel lonely? What types of people are more likely to feel lonely? These questions piqued our interest and therefore became our main focus when approaching the CSCS dataset. In order to find out the answer, we’ll be examining the relationship between age, social interactions / relationships, immigration and loneliness. Throughout this project, we hope to have a better understanding of the cause of loneliness, as it is a prerequisite to finding a cure for such problems.

Main research question: What is the most significant factor that leads to loneliness?

Analysis #1: Are old people lonelier?

Data used: DEMO\_age , LONELY\_ucla\_loneliness\_scale\_score and LONELY\_ucla\_loneliness\_scale\_score\_y\_n

Process: In order to find whether old people are more likely to feel lonelier, we first need to provide a definition for old people. In this case, we’ll use the UN definition, which is over 60 years old. The value for loneliness that we’ll be using comes from the UCLA Loneliness Scale, in which the larger the value, the lonelier one feels. Our hypothesis is that old people tend to feel lonelier compared to young people.

We’ll be using three plots in this analysis, in order to determine whether there is a correlation between age and loneliness:

1. Scatter plot: We’ll plot a scatter plot with age on the x-axis and loneliness score value on the y-axis. Then, a line of best fit will be plotted. In order to prove the line of best fit is reasonable, we’ll analyse r^2. If r^2 value is too small, then it means the line of best fit cannot represent the data well, therefore there is no correlation between age and loneliness. If the line of best fit is a good fit, there are two cases: If it is increasing, then it means older people tend to feel lonely more than young people. If it is decreasing, then it means the opposite.
2. Box plot: We’ll plot a box plot to compare the spread of loneliness scale values between young people (Group A) and old people (Group B). To find whether the elderly feel more lonely, we’ll compare the median of both data. If the median of Group B is greater than Group A, it means that old people tend to feel lonelier than young people.
3. 100% Stacked bar chart: With an 100% stacked bar chart one can easily compare the percentage of young and old people who fall into each result category of the survey, in this case y/n to loneliness. If the results for old people are more likely “yes”, then the percentage of yes answers among old people will be greater than the percentage of such answers among young people.

Analysis #2: Does social relationships help to mitigate loneliness?

Data used: RELATIONSHIP\_identity\_person\_(1~10), RELATIONSHIPS\_close\_person\_(1~10) LONELY\_ucla\_loneliness\_scale\_score

Process: Many people say that having meaningful social relationships will make one feel less lonely. In this analysis, we aim to figure out if such a statement stands for people living in Canada. In the relationship section of the survey, participants are required to list up to 10 people that they have a meaningful interaction with within 6 months and how close they are to each of these people. We’ll perform a hypothesis testing with the null hypothesis: Social relationships don’t help mitigate loneliness. The alternative hypothesis will then be: Social relationships will help mitigate loneliness. We’ll perform a two sample t-test, with two groups being people with high-quality relationships and people with low-quality relationships. In this case, people with high-quality relationships are defined as having more than median people who they feel are close with, while people with low-quality relationships have less than median people they feel are close with. We’ll then compare the difference of the median of loneliness scale score between two groups of people with the null hypothesis and find a p-value. If the p-value is less than 0.05, we’ll have enough evidence to reject the null hypothesis. Elsewise, we won’t have enough evidence against the null hypothesis.

Analysis #3: Do immigrants feel lonelier?

Data used: DEMO\_immigration, DEMO\_immigration\_age, LONELY\_ucla\_loneliness\_scale\_score, LONELY\_ucla\_loneliness\_scale\_score\_y\_n

Process: As immigrants are moving into a completely foreign country, it is sometimes normal for them to feel lonely. This analysis aims to find out if this statement holds true for immigrants to Canada. Our hypothesis will be that immigrants tend to feel lonelier compared to locals. We come up with two graphs that can be used to approach this problem.

1. 100% Stacked Bar Chart: The chart will be divided into two columns, one for immigrants and one for locals. Each column will display the corresponding percentage of people who fall in the yes and no category of the UCLA Loneliness Scale. If the percentage of immigrants who feel lonely is more than the percentage of locals who feel lonely, it can be inferred that immigrants tend to feel lonelier.
2. Scatter Plot: In order to have a deeper understanding of the issue, a scatter plot will be plotted to examine whether there is a correlation between immigration age and loneliness. The x-axis for the scatter plot will be age of immigration and the y-axis will be the loneliness scale score. We’ll then try to plot a line of best fit. If there is such a line that can represent the distribution of the data well, we can derive our results from the slope of the line. If the slope is positive, then immigrants who moved to Canada at an older age tend to feel lonelier. If it is negative, then immigrants who moved to Canada at a young age tend to feel lonelier.