**Demographic Testing**

**Introduction**

This analysis will focus on a cleaned subset of the 2022 Hennepin County Public Health survey. This survey is taken every four years and gathers data about residents’ mental health, physical well being, and the social determinants of health. This data has been cleaned in order to focus on individuals who had one or more poor mental health days within 31 days of the start of the survey. This is a numerical measure that ranges from 0.5 to 31. A planned subsequent analysis will identify the correlation, if any, between a resident’s number of bad mental health days and two factors. One such factor is their insecurity score. This is determined through a series of questions asking about food, housing, and transportation insecurity as well as community safety. This is a numerical factor rated on a scale of zero to thirty six. The other factor is rather the resident felt they had been discriminated against when seeking medical, mental, or dental health care. This is a categorical response of ‘yes’ or ‘no’. The purpose of this analysis is to determine if there are confounding factors present that may impede the accuracy of subsequent analysis. **The confounding factors examined in this analysis include gender identity, if they identify as transgender, and their race/ethnicity.**

**Testing for associations between demographics and the Insecurity predictor**

This analysis will first focus on confounders of insecurity score. The three main confounders being examined are a resident’s gender identity, if they identify as transgender, and their race/ethnicity. The ‘gender’ response category lists respondents as male, female or agender/genderqueer/nonbinary. These are numbered 1 through three respectively in the data. The transgender status variable assigns cisgender respondents a grouping of 0 and transgender respondents 1. The race/ethnicity category lists respondents as Native/American Indian (including Hispanic), Hispanic/Latino/a/e, Asian (non-Hispanic), Black (non-Hispanic), or White (non-Hispanic). Within the data, these groups are given a number 1-5 respectively. Participant insecurity scores were in a range of zero to twenty six, with thirty five being the highest possible score.

In table 1, summary statistics for all groups were found using the command describeby() within the ‘psych’ library in R. Further analysis was then performed based on the results.

Beginning with the gender category, figure one shows a similar mean insecurity score for both males and females; the insecurity score of agender/genderqueer/nonbinary respondents seems to be higher. Within the gender category there were 1336 males, 2967 females, and 87 agender/genderqueer/nonbinary residents (table 1). The highest mean insecurity score was agender/genderqueer/nonbinary residents’ 5.9 and the lowest was female residents’ score of 4.44 (table 1). Standard deviations were relatively similar, ranging from 3.9 to 4.54 (table 1). A two-sample t-test was determined to be the most appropriate method of comparison as the explanatory variable is categorical and the response variable is numerical. Additionally, the necessary assumptions (independent sampling, normal distribution, and equal variance) were met.The test was performed using the summary(aov()) function in R. The ANOVA test returned an insignificant p-value, insinuating there is not a significant difference of any gender groups’ insecurity score (table 1).

Grouping respondents by transgender status, figure 2 shows an elevated insecurity score for those identifying as transgender in comparison to those identifying as cisgender. Given the size of the response groups, the CLT will hold. There were 4338 cisgender respondents and 62 transgender respondents with mean insecurity scores of 4.49 and 6.47 respectively (table 1). The standard deviation of responses were very similar at 4.39 and 4.44 (table 1). An ANOVA test was determined to be most appropriate as the explanatory variable is categorical and the response variable is numeric. Necessary assumptions (independent sampling, normal distribution, and equal variance) were met. This test was performed using the summary(aov()) function in R. The test returned a p-value of , meaning there is a significant difference between the average insecurity score of cisgender and transgender residents (table 1).

Grouping residents by race/ethnicity in figure 3, it can be seen that residents identifying as white have a significantly lower average insecurity score than other groups. Those identifying as native American/Indian American (including hispanic) have the highest average insecurity score with Black (non-hispanic) respondents not too far behind (fig. 3). 193 respondents identify as Native (2022 inclusive), 218 as Hispanic, 190 as Asian (non-hispanic), 404 as Black (non-hispanic), and 3158 as White (table 1). The mean insecurity score ranged from Native respondents’ 10.05 to White respondents’ 3.34; Black residents had the second highest mean insecurity score of 9.28. Asian respondents had a mean insecurity score of 5.32 and Hispanic respondents of 6.39 (table 1). Standard deviation also varied widely between White respondents’ 3.25 and Native respondents’ 5.7 (table 1). Given the explanatory variables are categorical and the response variable is numerical, an ANOVA test would generally be used. However, the equal variance assumption required for an ANOVA analysis was not met, so a Kruskal-Wallace test was performed using the kruskal.test() command in R. This returned a significant p-value of , meaning there is a difference in average insecurity score between recorded races/ethnicities (table 1). To identify which groups had significant differences, a pairwise test was performed using the pairwise.wilcox.test() command with the p.adjust.method = "BH" option. The only groups to not have a significant difference were White and Black residents. The p-value of Hispanic and Asian residents was . All other groups had a p-value below 0.0001 (table 1).

**Testing for Associations Between Demographics and the Discrimination in Health Care Predictor**

The next part of this analysis will focus on the association between the aforementioned demographic categories and the discrimination these individuals face in a healthcare setting.

This will be done by analyzing the relative risk of the categories above, meaning all categories will be reduced to two groups. For this section, gender will be listed as either agender/genderqueer/nonbinary or male/female. Trans status will remain between either cisgender or transgender. Groupings of ethnicity/race will be split between BIPOC/AI and White. Relative risk is acceptable to be used in this situation as this is not a case/-control study and all assumptions are met: random sampling, accurate data, independent observations, and all expected values are greater than 5, making them ‘large enough’.

In table 2, relative risks, confidence intervals, and p-values were found using the command (summary(epi2x2)) from the epibasix library in R.

Beginning once again with gender, figure four shows the majority of both groups did not face discrimination during medical care. The differences in proportionality in discrimination between the groups is difficult to ascertain due to the large difference in sample sizes, so a qualitative analysis will tell us much more. Using the p-value of 0.513 given from the command given above, there is not enough evidence to say that there is a significant difference between medical discrimination for those identifying as agender/genderqueer/nonbinary and male/female.

Much like figure four, figure five shows that the majority of respondents did not experience discrimination in a medical setting. The difference in medical discrimination between those identifying as transgender and cisgender are still hard to see given the discrepancy in number of respondents (fig. 5). The p-value from a relative risk calculation for these two groups is 0.091, which is smaller than that above, but still not statistically significant.

Looking to race/ethnicity, figure six shows that the majority of both groups say they have not faced discrimination. However, BIPOC/AI respondents have reported more instances of medical discrimination then White respondents, which is notable given the difference in number of respondents per group. From a visual inspection, there seems to be a significant difference in medical discrimination between the two groups. This is confirmed by a relative risk assessment, where the p-value returns as zero. With 95% certainty, the relative risk of facing medical discrimination for those who are BIPOC/AI are between 3.72 and 5.611 times the risk than those who are White.

**Testing for Associations Between Demographics and the Number of Not Good Mental Health Days Outcome**

This final section of the analysis looks to determine the association between demographic (as defined in the section ‘Testing for associations between demographics and the Insecurity predictor’) and the number of days a participant’s health was not good. As in the section mentioned, summary statistics for all groups in table 3 were found using the command describeby() within the ‘psych’ library in R and further analysis was then performed based on those results. An ANOVA analysis was determined to be the most appropriate method of comparison for all categories as the explanatory variables are categorical and the response variable is numerical. Additionally, the necessary assumptions (independent sampling, normal distribution, and equal variance) were met by all demographic groupings (table 3).

Looking first to the gender demographic, the number of bad mental health days for those identifying as agender/genderqueer/nonbinary looks to be significantly higher than those identifying as male or female, which seem to be relatively even with one another (fig. 7). There were 1381 respondents identifying as male, 3057 as female, and 90 as agender/genderqueer/nonbinary (table 3). The mean number of bad mental health days for each group, respectively, is 8.37. 8364, and 15.27 (table3). An ANOVA analysis returned a p-value of , meaning there were significant differences between the prescribed gender categories (table 3). A TukeyHSD post-hoc analysis was performed to find what groups had significant differences. It was found that males and agender/genderqueer/nonbinary respondents as well as female and agender/genderqueer/nonbinary respondents had significant differences in number of bad mental health days; both of which were near zero (table 3).

Moving to the transgender status category, figure 8 shows a large discrepancy in the number of bad mental health days between those who identify as cisgender and transgender, with those identifying as transgender having a higher amount. There were 4471 respondents who identify as cisgender and 68 who identify as transgender (table 3). The mean number of bad mental health days for cisgender respondents is 8.58; the mean for transgender respondents is nearly double at 16.31 days. An ANOVA test was performed and returned a p-value of , meaning there is a statistically significant difference between the mean number of mental health days for cisgender and transgender residents.

Lastly, differences in mean number of mental health days between races/ethnicities were examined. Looking ro figure 9, there is a fair bit of variance between the groups, but no groups stand out as having exceptionally more or less bad mental health days. Native Americans/American indians (including Hispanic) look to have the highest number of bad mental health days while Whites (non-Hispanic) have the least. 200 respondents identify as Native (2022 inclusive), 229 as Hispanic, 200 as Asian (non-hispanic), 427 as Black (non-hispanic), and 3239 as White (table 1). The mean number of bad mental health days ranged from Native respondents’ 11.57 to White respondents’ 8.06; Black residents had the second highest mean number of mental health days with 10.9 days. An ANOVA test returned a p-value of , which indicates that there are significant differences between demographic groups. A post-hoc TukeyHSD test determined that There were a number of differences. These differences included Native & Asian participants with a p-value of . Native & White participants with a p-value of . Hispanic & White participants with a p-value of . Asian & Black participants with a p-value of: . Asian & White participants with a p-value of , as the p-value is too small for the software to calculate.

**Conclusion**

A number of the variables explored today showed an association with a predictor (insecurity or health care discrimination) or the outcome of interest (bad mental health days). For the predictor insecurity score, both transgender status and race/ethnicity were confounding factors. The predictor of medical discrimination had the confounding factor of BOPIC/AI status. All demographics (gender, transgender status, and race/ethnicity) were confounders for the number of bad mental health days a person could have. Transgender status and race/ethnicity had an association with both a predictor and the outcome of interest. These will need to be taken into account as confounding variables in subsequent analyses.

**Appendix**

Figure 1.

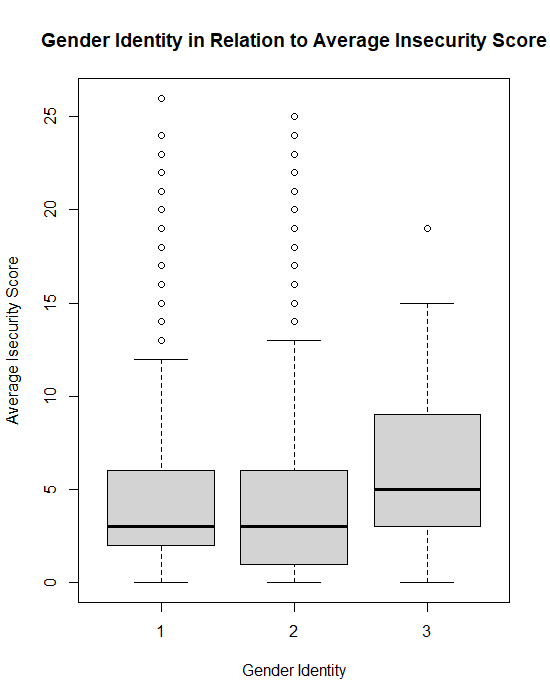


Figure 2

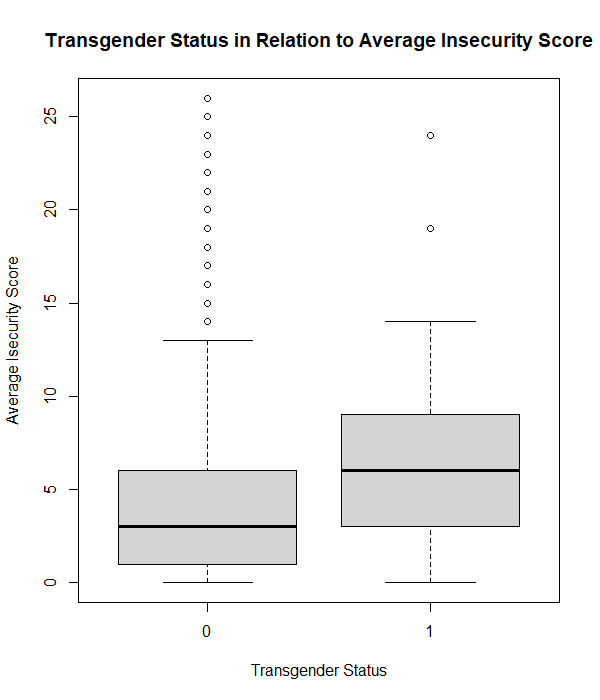
****

Figure 3

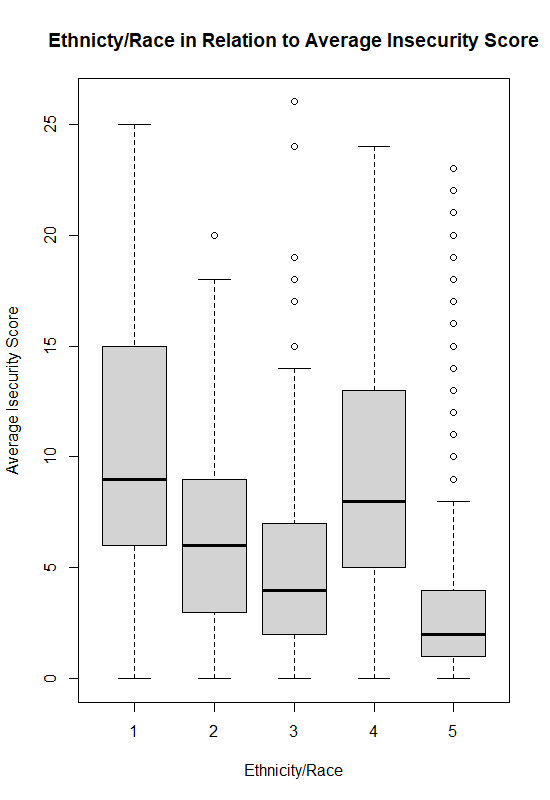


Figure 4

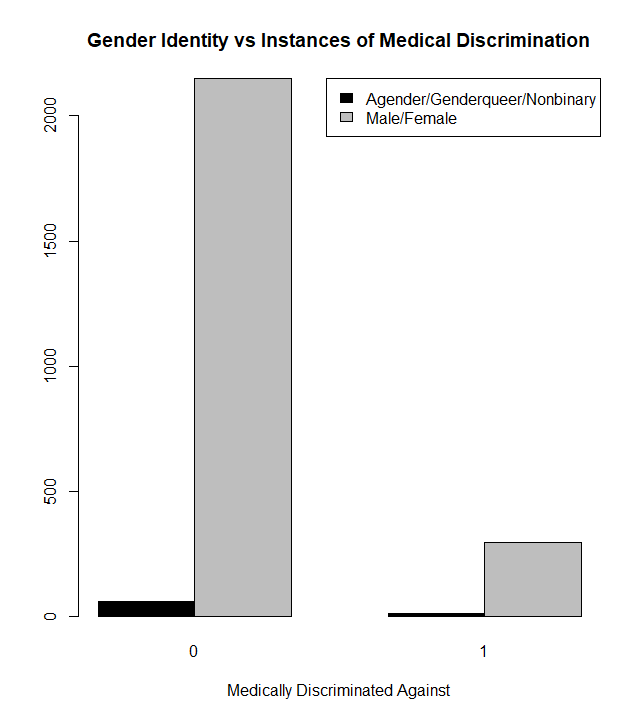
****

Figure 5

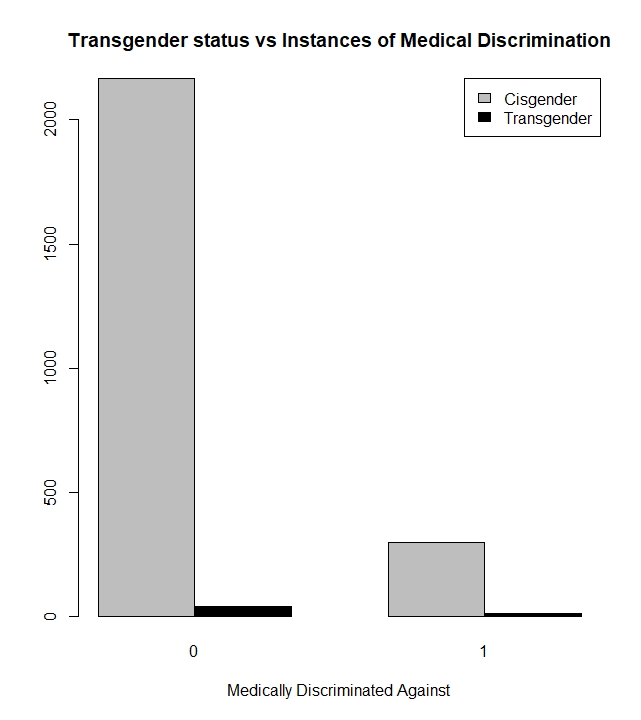


Figure 6

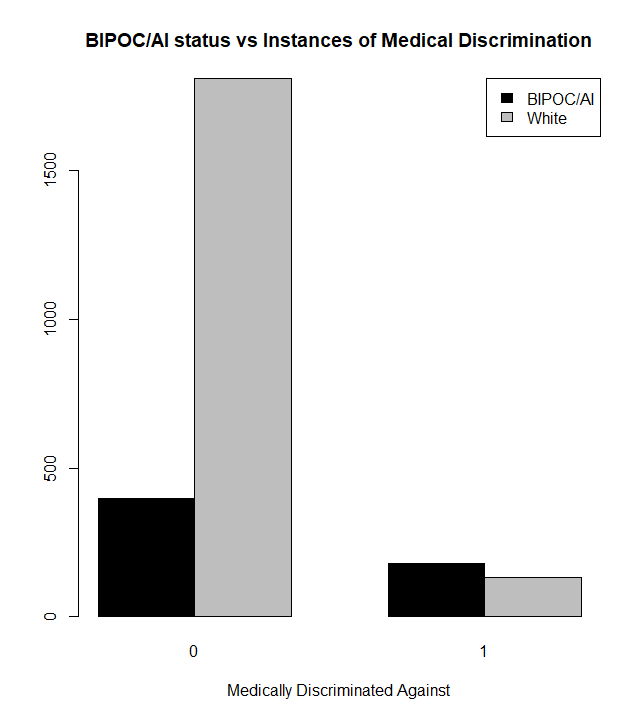


Figure 7

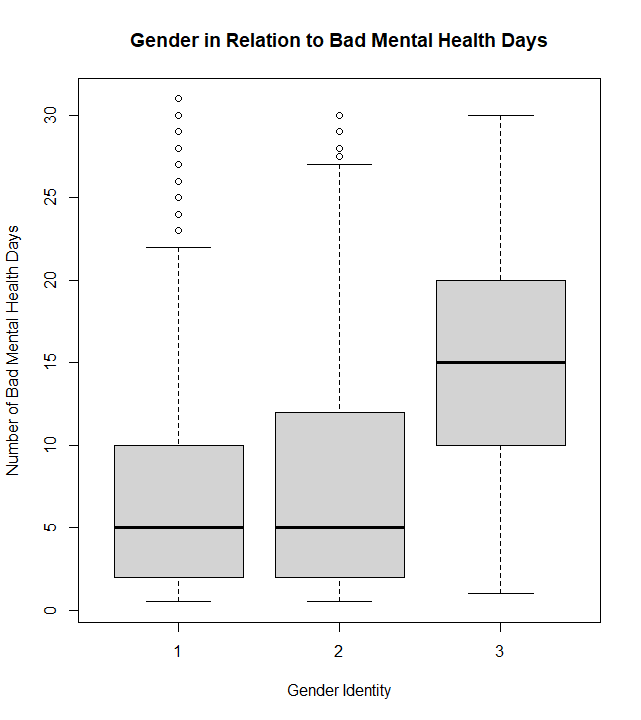


Figure 8

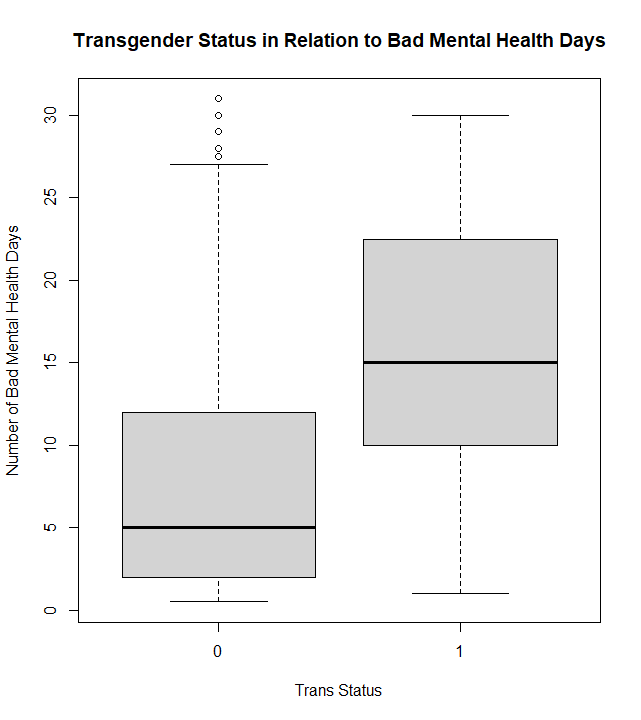


Figure 9

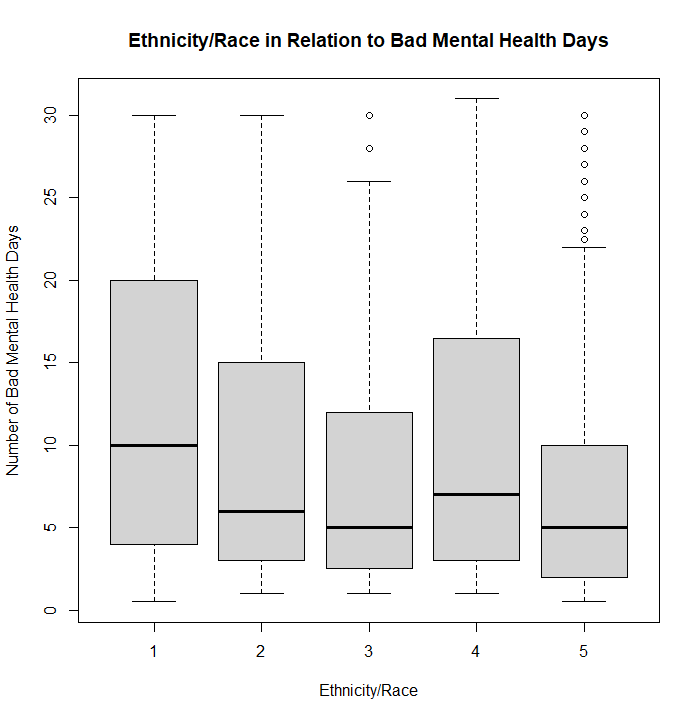


Table 1: Demographic Categories versus Insecurity Score

| **Variable/Categories** | **n** | **Mean** | **SD** | **p-value** | **Differences** |
| --- | --- | --- | --- | --- | --- |
| **Gender** |  |  |  | 0.786 |  |
| Male | 1336 | 4.57 | 4.54 |  | None |
| Female | 2967 | 4.44 | 4.34 |  |
| Agender/Genderqueer/  Nonbinary | 87 | 5.9 | 3.9 |  |
| **Trans Status** |  |  |  |  |  |
| Cis | 4338 | 4.49 | 4.39 |  | Cisgender & Transgender |
| Trans | 62 | 6.47 | 4.44 |  |
| **Race/Ethnicity** |  |  |  |  |  |
| Native 2022 inclusive | 193 | 10.05 | 5.7 |  | Native & Hispanic:  Native & Asian:  Native & Black: None  Native & White:  Hispanic & Asian:  Hispanic & Black:  Hispanic & White:  Asian & Black:  Asian & White:  Black & White: |
| Hispanic | 218 | 6.39 | 4.37 |  |
| Asian-NH | 190 | 5.32 | 4.52 |  |
| Black-NH | 404 | 9.28 | 5.4 |  |
| White-NH | 3158 | 3.34 | 3.25 |  |

Table 2: Demographic Risk of Discrimination in Health Care

| **Variable/Categories** | **Relative Risk** | **95% Confidence Interval for RR** | **p-value** |
| --- | --- | --- | --- |
| **Gender2** |  |  | 0.513 |
| Agender/Genderqueer/Nonbinary | 1.272 | [0.731, 2.212] |  |
| Male/Female | 1 | N/A |  |
| **Trans Status** |  |  | 0.091 |
| Trans | 1.716 | [1.003, 2.935] |  |
| Cis | 1 | N/A |  |
| **BIPOC/AI** |  |  | 0 |
| BIPOC/American Indian | 4.569 | [3.72, 5.611] |  |
| White | 1 | N/A |  |

Table 3: Demographics versus number of not good mental health days

| **Variable/Categories** | **n** | **Mean** | **SD** | **p-value** | **Differences** |
| --- | --- | --- | --- | --- | --- |
| **Gender** |  |  |  |  |  |
| Male | 1381 | 8.37 | 8.36 |  | Male & Agender/Genderqueer/Nonbinary: 0  Female & Agender/Genderqueer/Nonbinary: 0 |
| Female | 3057 | 8.64 | 8.3 |  |
| Agender/Genderqueer/Nonbinary | 90 | 15.27 | 8.67 |  |
| **Trans Status** |  |  |  |  |  |
| Cis | 4471 | 8.58 | 8.33 |  | Cisgender & Transgender |
| Trans | 68 | 16.31 | 9.3 |  |
| **Race/Ethnicity** |  |  |  |  |  |
| Native 2022 inclusive | 200 | 11.57 | 9.27 |  | Native & Asian:  Native & White:  Hispanic & White:  Asian & Black:  Asian & White: |
| Hispanic | 229 | 9.77 | 8.67 |  |
| Asian-NH | 200 | 8.71 | 8.72 |  |
| Black-NH | 427 | 10.9 | 9.36 |  |
| White-NH | 3239 | 8.06 | 7.96 |  |