Programming Assignment 3

Andrew Shaw

09.11.25

### Create a table that shows the means of ecig use by each category and shows the p-value of a t-test from a two-sample difference test for the following categories. Use R-Markdown to complete the assignment.

1. Male versus Female
2. 12th versus 9th grade
3. Hispanic versus non-Hispanic
4. Black versus non-Black
5. White versus non-White
6. Smokers versus non-Smokers

### Tables

table\_sex

Table 1: E-cigarette Use by Sex

| Group | % E-cig users |
| --- | --- |
| Female | 18.5 |
| Male | 19.8 |
| p-value: 2.14e-02 | |

table\_grade

Table 1: E-cigarette Use by Grade

| Group | % E-cig users |
| --- | --- |
| 9th | 32.7 |
| 12th | 22.0 |
| p-value: 2.78e-17 | |

table\_hispanic

Table 1: E-cigarette Use by Hispanic & Non-Hispanic

| Group | % E-cig users |
| --- | --- |
| Hispanic | 18.3 |
| Non-Hispanic | 19.7 |
| p-value: 1.87e-02 | |

table\_black

Table 1: E-cigarette Use by Black & Non-Black

| Group | % E-cig users |
| --- | --- |
| Black | 15.5 |
| Non-Black | 20.6 |
| p-value: 8.31e-14 | |

table\_white

Table 1: E-cigarette Use by White & Non-White

| Group | % E-cig users |
| --- | --- |
| Non-White | 15.3 |
| White | 21.5 |
| p-value: 1.15e-24 | |

table\_smoker

Table 1: E-cigarette Use by Smokers & Non-Smokers

| Group | % E-cig users |
| --- | --- |
| Smokers | 16.6 |
| Non-Smokers | 82.6 |
| p-value: 2.9e-229 | |

### Summary

Throughout all the various populations observed, all six were found to be statistically significant to varying degrees. The groups with the largest differences between the observed populations were; smokers and non-smokers, 9th and 12th graders, Black and non-Black, and White & non-White. The tests for male and female, and Hispanic and non-Hispanic, were statistically significant but less so than the other groups with p-values of 0.0214 and 0.0187, respectively. This represents a 2% probability that the difference in males’ higher usage rates from women is due to random variation. Given the large sample size of the data across all subgroups, there is a significant amount of data to detect differences among these populations.