**Research and Contributions Report**

****

By: Matthew Knight & Nathaniel Gosdin

Group 5 CSCI / ISAT B104 SPRING 2022

Contact Information

**Nathaniel Gosdin-**

**Phone #: 843-422-0935**

**Email:** [**nwgosdin@hotmail.com**](mailto:nwgosdin@hotmail.com)

**Matthew Knight-**

**Phone #: 843-338-8332**

**Email:** [**matthewkknight1@gmail.com**](mailto:matthewkknight1@gmail.com)

**Availability-**

**Matthew: Out of class Monday-Tuesday from 11 am – 7 pm, Wednesday-Sunday anytime during the day.**

**Nathaniel: Monday-Friday anytime during the day, Weekends depend on work.**

**The Project**

**The Question: What is the correlation between asthma and physical education attendance and how does it affect boy’s vs girls in high school?**

**Query:**

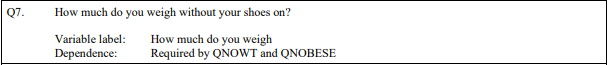
SELECT XXHq.raceeth, XXHq.q6orig, XXHq.q7orig, XXHq.record, XXHq.orig\_rec, XXHq.q1, XXHq.q2, XXHq.q3, XXHq.q4, XXHq.q5, XXHq.q6, XXHq.q7, XXHq.81, XXHq.87

FROM XXHq;

**Descriptive Questions:**

**Table

Description automatically generated**

****

**YRBS Questions:**

**Graphical user interface, text, application

Description automatically generated**

**Text, letter

Description automatically generated**

**What the query should look like: Graphical user interface, text, application

Description automatically generated**

**Graphical user interface, application, table, Excel

Description automatically generated**

**Our type of statistical test: The type of statistical test our group will be using is T-Test because we have one continuous variable, and one categorical variable**

|  |  |  |  |
| --- | --- | --- | --- |
| **Task set** | **Group Member** | **Contribution %** | **Components** |
| **ICE 13.1** | **Nate** | **50%** | **Question Definitions, Answer Mapping** |
|  | **Matthew** | **50%** | **Question Definitions, Answer Mapping** |
| **ICE 13.2** | **Nate** | **50%** | **Data Winnowing, SQL Query Development** |
|  | **Matthew** | **50%** | **Data Winnowing, SQL Query Development** |
| **ICE 13.3** | **Nate** | **50%** | **Identifying the type of statistical tests** |
|  | **Matthew** | **50%** | **Identifying the type of statistical tests** |
| **ICE 14.1a**  **ICE 14.1b** | **Nate** | **50%** | **Statistical Analysis,**  **Made official excel file** |
|  | **Matthew** | **50%** | **Statistical Analysis,**  **Made official excel file** |
| **ICE 14.2** | **Nate** | **50%** | **Working on completing Milestone 1** |
|  | **Matthew** | **50%** | **Working on completing Milestone 1** |
| **ICE 14.3 (Friday)** | **Nate** | **50%** | **Continue working on completing Milestone 1** |
|  | **Matthew** | **50%** | **Continue working on completing Milestone 1** |
| **ICE 15.1** | **Nate** | **50%** | **Working on Milestone II** |
|  | **Matthew** | **50%** | **Working on Milestone II** |
| **ICE 15.2** | **Nate** | **50%** | **Working on Milestone II** |
|  | **Matthew** | **50%** | **Working on Milestone II** |
| **ICE 15.3** | **Nate** | **50%** | **Working on Milestone II/Presentation Timeline** |
|  | **Matthew** | **50%** | **Working on Milestone II/Presentation Timeline** |
| **Apr 23 - 24** | **Nate** | **50%** | **Finishing up the Final Project Rough Draft** |
|  | **Matthew** | **50%** | **Finishing up the Final Project Rough Draft** |

**A Change:**

On the 25th of April we did not see any relationship between Asthma and PE attendance, no matter what way we viewed the data. Rather than wrapping the project up on an inconclusive note, we decided to ask a different question. After some debate, a new question was chosen.

**The New Question:**

**What is the relationship between** **Alcohol use and physical altercations in high school students?**

**Query:**

SELECT XXHq.raceeth, XXHq.q6orig, XXHq.q7orig, XXHq.record, XXHq.orig\_rec, XXHq.q1, XXHq.q2, XXHq.q3, XXHq.q4, XXHq.q5, XXHq.q6, XXHq.q7, XXHq.17, XXHq.41

FROM XXHq;

**Abstract:**

**In this project, we seek to examine the relationship between Alcohol use and physical altercations in high school students using the data from the YRBS 2019 survey and python to visualize the data via graphs. We found a noticeable connection between Alcohol use and physical altercations and conclude that there is a relationship between the two using the data given.**

Text

Description automatically generated

**Data Retrieval and Winnowing**

The first step we took in the data retrieval process was downloading the database file from the YRBS website and running it in Microsoft Access. The next step we took was to use an SQL Query to select the data we wanted to use. This included the demographic questions (Q 1-7) and the questions we chose for this project (Q17, Q41).

**-Contributions-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Task Set** | **Group Member** | **Contribution %** | **Components** |
| **Finding a new question and updating query (04/25)** | **Nate** | **50%** | **Question and Query** |
|  | **Matthew** | **50%** | **Question and Query** |
| **Updated coding and Heatmap** | **Nate** | **50%** | **Reviewed and tested code** |
|  | **Matthew** | **50%** | **Created and fixed the new code** |
| **Coding new graphs for data (04/26-28)** | **Nate** | **50%** | **Reviewed and tested the graphs and pie chart** |
| **--------------------** | **Matthew** | **50%** | **Created and fixed new code** |
| **Beginning work on the poster/R&C report and updating code (04/29-05/02)** | **Nate** | **50%** | **Worked on poster and R&C report** |
|  | **Matthew** | **50%** | **Coding and Reviewing poster and work. Updated and created bar graph and pie chart. tweaked heat map.** |
| **Final day: Checkup and quality insurance (05/03)** | **Nate** | **50%** | **Review content** |
|  | **Matthew** | **50%** | **Review content** |

**Graphs:**

**Diagram

Description automatically generated with low confidence**

**Chart, bar chart

Description automatically generated**

**Chart, pie chart

Description automatically generated** **Chart, pie chart

Description automatically generated**

**Results:**

**There is considerable correlation in the relationship between alcohol consumption and physical altercations in high school students.**

**The number of students who consume alcohol and/or have had a physical altercation is likely far higher due to the sensitivity of the subject.**

**Many students (21.2%) did not answer the question as to if they have had a physical altercation, as is shown in the Pie chart.**

**There exists a small but noticeable correlation between physical altercations and sex, but not for Alcohol consumption and sex.**

**There exists a small (0.1) but noticeable correlation between weight and both Alcohol consumption and physical altercations.**

**References:**

**“Youth Risk Behavior Surveillance — United States, 2019.” CDC, CDC,** [**https://www.cdc.gov/mmwr/volumes/69/su/pdfs/su6901-H.pdf**](https://www.cdc.gov/mmwr/volumes/69/su/pdfs/su6901-H.pdf)**.**

**“Matplotlib Pie Charts.” Matplotlib Pie Charts, w3schools,** [**https://www.w3schools.com/python/matplotlib\_pie\_charts.asp**](https://www.w3schools.com/python/matplotlib_pie_charts.asp)**.**

**“Python - Read CSV Columns into List.” GeeksforGeeks, GeeksforGeeks, 17 Sept. 2021,** [**https://www.geeksforgeeks.org/python-read-csv-columns-into-list/**](https://www.geeksforgeeks.org/python-read-csv-columns-into-list/)**.**

**“How to Plot Bar Graph in Python Using CSV File?” GeeksforGeeks, 25 Feb. 2021,** [**https://www.geeksforgeeks.org/how-to-plot-bar-graph-in-python-using-csv-file/**](https://www.geeksforgeeks.org/how-to-plot-bar-graph-in-python-using-csv-file/)**.**

**Indian AI Production. “Seaborn Heatmap Using Sns.heatmap(): Python Seaborn Tutorial.” Indian AI Production, 8 Sept. 2019,** [**https://indianaiproduction.com/seaborn-heatmap/**](https://indianaiproduction.com/seaborn-heatmap/)**.**