|  |  |
| --- | --- |
|  | **CSCI/ISAT B104**  **Computer Programming Techniques, Practices, and Tools**  **Fall 2024**  **Project Documentation**  **Team XX** |

Contents

[Project Contributions by Team Member 1](#_Toc1538789207)

[Collaboration Tools 4](#_Toc1986994104)

[YRBS Questions Explored 4](#_Toc2144268284)

[YRBS Demographic Information Explored 5](#_Toc831551195)

[YRBS Data Winnowing & Retrieval 6](#_Toc1915595218)

[Research Question and/or Hypothesis 8](#_Toc671424772)

[Bibliography 8](#_Toc1468233840)

[Acknowledgements 10](#_Toc1832932150)

# Project Contributions by Team Member

**Members and their Contact Information**

|  |  |  |
| --- | --- | --- |
| **Member** | **Email** | **Text** |
| Jalyn Perry | jalynmp@email.uscb.edu | 8435993055 |
| Tabias Farrison | tabiasf@email.uscb.edu | 8039284887 |

**Overall**

Relative Contribution of each member over the course of the entire project

|  |  |  |
| --- | --- | --- |
| **Member** | **Contribution** | **Total Hours** |
|  |  |  |
|  |  |  |

**YRBS Question Identification & Documentation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 50 % | 4 | * Using critical thinking to determine which questions may produce an interesting correlation. * Using critical thinking to determine which demographic questions may relate to previously explored questions. |
| Tabias Farrison | 50 % | 4 | * Using critical thinking to determine which questions may produce an interesting correlation.   Using critical thinking to determine which demographic questions may relate to previously explored questions |

**YRBS Data Winnowing & Retrieval**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 50% | 4 | * Experimenting with pulling data directly from a SQL database into python. * Finalizing data by dropping unnecessary questions. * Determining which questions were necessary (helpful) to later analysis. |
| Tabias Farrison | 50% | 4 |  |

**Data Analysis (via Python)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 70 % | 8 | * Researching libraries is necessary for data analysis/visualization. * Creating hypothesis based on what we observed initially when exploring the data * Exploring data and correlations and ensuring the underlying logic behind each implementation was correct. |
| Tabias Farrison | 30 % | 5 | * Experimenting with pulling data directly from a SQL database into python. * Finalizing data by dropping unnecessary questions. * Determining which questions were necessary (helpful) to later analysis. |

**Data Visualization (via Python)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 70% | 5 | * Created visual graphs to represent data * Researched various graphs and decided which would be the best |
| Tabias Farrison | 30 % | 3 | * Created visual graphs to represent data * Researched various graphs and decided which would be the best |

**Python Script Creation**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 70% | 1.5 | * Created a menu script so that users can select the graph they want to see |
| Tabias Farrison | 30 % | 1 | * Tested and assisted in creating a menu script so that users can select the graph they want to see |

**Poster Development**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 70% | 4 | * Checked over poster and added findings * Reformatted poster * Added missing visuals/ findings * All completed on 12/4/24 |
| Tabias Farrison | 30% | 2 | * Added graphs, visuals, introduction, hypothesis, and data winnowing. * All completed on 12/4/24 |

**Presentation Preparation & Practice**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 50% | 1 | * Created notecards and practiced presentation * Practice scheduled for 12/5/24 at 5:45 * Added details to project timeline |
| Tabias Farrison | 50% | 1 | * Created notecards and practiced presentation * Practice scheduled for 12/5/24 at 5:45 |

**Project Documentation & Administration**

|  |  |  |  |
| --- | --- | --- | --- |
| **Member** | **Contribution** | **Hours** | **Components** |
| Jalyn Perry | 70% | 2 | * Added to project documentation |
| Tabias Farrison | 30% | 1 | * Added to project documentation |

# Collaboration Tools

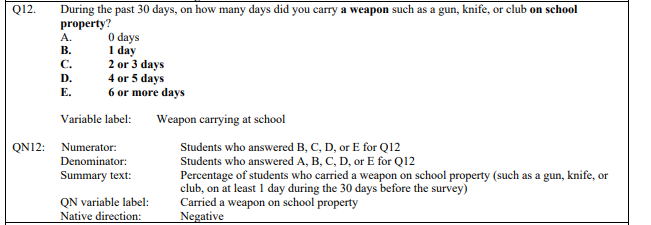
Tools to assist with task management, version control, communication, etc.

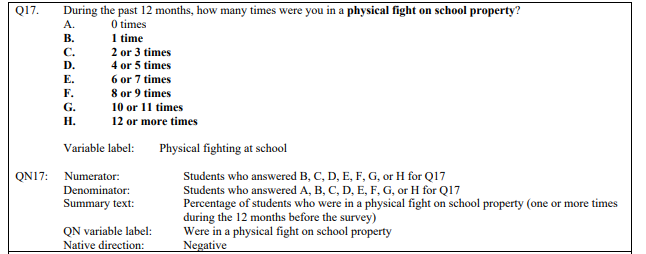
* Trello ([B104 Project | Trello](https://trello.com/b/YIUJzZwH/b104-project))
* GitHub (https://github.com/Jalyn-Perry/B104)
* Teams
* Word
* Spyder

# YRBS Questions Explored

Questions & responses explored over the course of the project.

*Place screenshots of the two questions below. Be sure to include their “response keys”.*



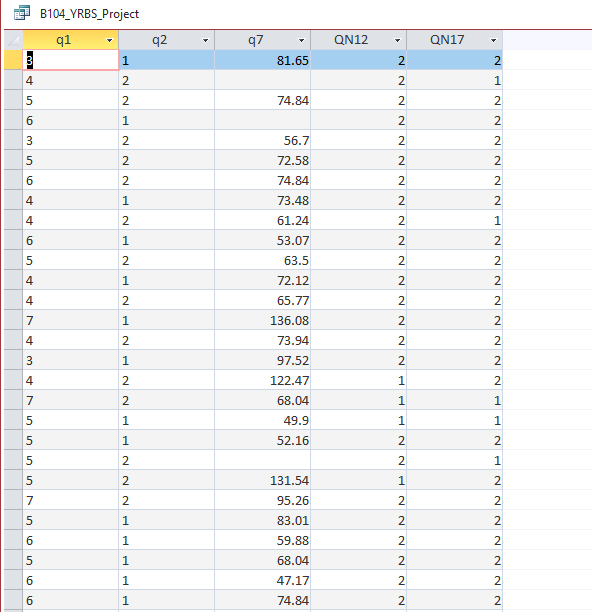


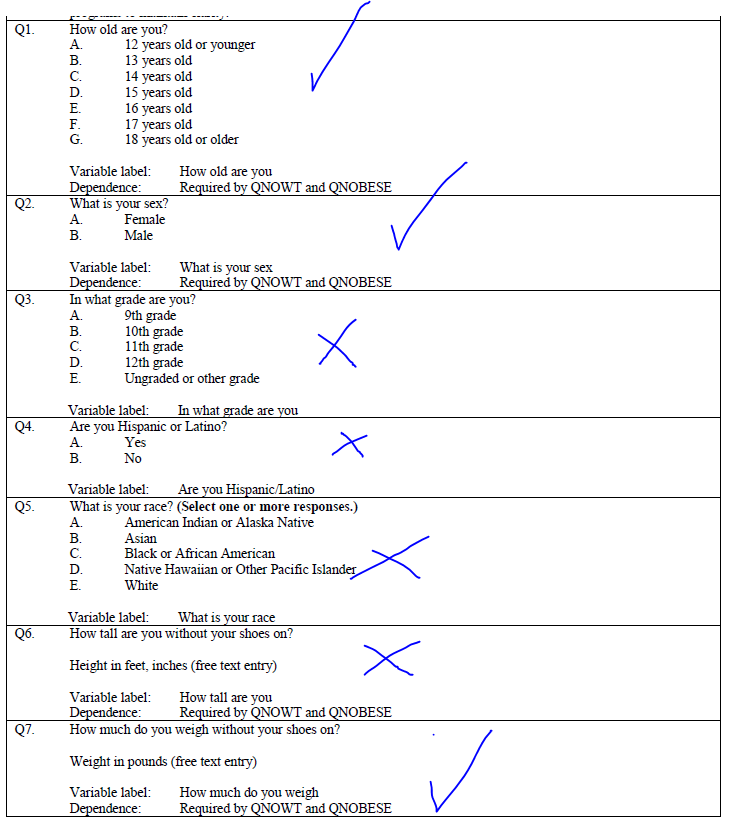
# YRBS Demographic Information Explored

Demographic questions & responses explored over the course of the project

*Place screenshots of the questions below. Be sure to include their “response keys”.*

*Demographic information includes gender, age, height, weight, ethnicity, etc.*



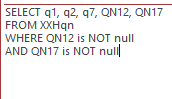
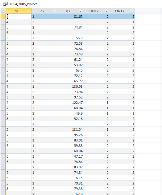


# YRBS Data Winnowing & Retrieval

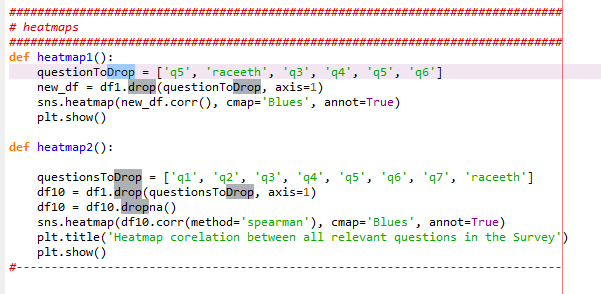
How we winnowed & retrieved the data explored over the course of the project.

Write down a step-by-step guide explaining how you winnowed & retrieved the data from the YRBS site.

*Place screenshots of the SQL code and other relevant details here*

* *Download msdb file*
* *Open in msdb file in Microsoft access*
* *Create new query* 
  + 
* Run query
* Resulting data
  + 
* *Save query*
* *Export as excel sheet*

*This approach models the questions we used, however, unnecessary questions were mostly dropped via python (see below)*



*-*

# Research Question and/or Hypothesis

The specific research questions we sought to answer.

What correlations can exploratory data analysis of YRBSS data reveal about age, gender, and violent behavior, including the relationship between bringing a weapon to school and engaging in frequent fights?

H1: Someone who has been involved in a school fight has a higher chance of carrying a weapon to school.

H2: Someone who has been involved in a school fight has a lower chance of carrying a weapon to school.

H3: There is a correlation between gender and violent tendencies.

Null-H1: There is no correlation between school fights and carrying a weapon to school

Null-H2: There is no correlation between gender and violent tendencies.

Findings

* Our findings suggest that the correlation between questions 12 and 17 is very loose in the positive direction. This correlation is not necessarily proof of anything significant.
* Out of the male students who brought weapons 4 or more days (149), 13 of them got into frequent fights. 8.7%
* Out of the 324 female students who brought weapons 4 or more days, 42 got into frequent fights.12.3%
* There is not a strong correlation between a person's gender, and their propensity for violent behavior.

# Bibliography

Resources which we used over the course of the project.

Citations for:

Python Library Documentation

Python Tutorials

Best Practices

Poster Templates

Stack Overflow questions/answers

Etc

References

Abba, I. (2023, January 12). *Matplotlib Figure Size – How to Change Plot Size in Python with plt.figsize()*. freeCodeCamp.org. <https://www.freecodecamp.org/news/matplotlib-figure-size-change-plot-size-in-python/>

Belkin, J. (2023, October 30). Working with survey data in Python: Easy pandas project. *Medium*. <https://medium.com/@juliabelkin/working-with-survey-data-in-python-easy-pandas-project-d70089cce1ae>

GeeksforGeeks. (2024, January 9). *Plotting Histogram in Python using Matplotlib*. GeeksforGeeks. <https://www.geeksforgeeks.org/plotting-histogram-in-python-using-matplotlib/>

*How to plot pie chart using data frame group by different range?* (n.d.). Stack Overflow. <https://stackoverflow.com/questions/46542572/how-to-plot-pie-chart-using-data-frame-group-by-different-range>

Savcı, U. (2022, April 3). Exploratory Data Analysis using Python | Medium. *Medium*. <https://medium.com/@ugursavci/complete-exploratory-data-analysis-using-python-9f685d67d1e4>

*Seaborn heatmap | How to make a heatmap in Python Seaborn and adjust the heatmap style - YouTube*. (n.d.). Youtube. <https://www.youtube.com/watch?v=0U9cs2V-Mqc&t=217s>

*Seaborn Heatmap - How to Visualise Correlations and Data With Heatmaps in Python*. (n.d.). Youtube. <https://www.youtube.com/watch?v=J7cd1-g1O7A&t=1s>

*seaborn: statistical data visualization — seaborn 0.13.2 documentation*. (n.d.). <https://seaborn.pydata.org/>

*seaborn.kdeplot — seaborn 0.13.2 documentation*. (n.d.). <https://seaborn.pydata.org/generated/seaborn.kdeplot.html>

*W3Schools.com*. (n.d.). Retrieved November 29, 2024, from <https://www.w3schools.com/python/matplotlib_pie_charts.asp>

# Acknowledgements

Specific peers who helped you debug

Upper class students, parents, siblings, friends, etc. who provided feedback and guidance

Etc.

* Amber Ravenell – Project Management and Documentation Assistant
* Daniel Scheer - Understanding how to process data