Weekly Status Submission Template

Week Number: 5

Date: September 30 2025

Student Name: Abdullah Siddiqui

Original Plan for This Week: (5 points)

My plan for this week was to extend the Week 4 correlation analysis by running both linear and logistic regression models using the cleaned BRFSS dataset. I planned to check regression assumptions, save diagnostic plots, calculate odds ratios, and improve my GitHub documentation with outputs and tables.

Tasks Accomplished This Week: (10 points)

Ran a multiple linear regression predicting number of poor mental health days (MENTHLTH) from income, education, exercise, smoking, alcohol, and age.
Checked multicollinearity using VIF and exported results.
Generated diagnostic plots (residuals vs. fitted, Q–Q plot).
Built a logistic regression model predicting frequent distress (binary >14 poor mental health days).
Extracted and saved coefficients, odds ratios, and confidence intervals.
Plotted the predicted probability distribution from the logistic model.

Organized evidence into /plots, /tables, and /outputs folders in my GitHub repo.

Comparison of Planned vs Actual: (10 points)

Planned (Week 5)	Actual (Week 5)
Run linear regression and check assumptions	Completed linear regression, saved summary, VIF, and diagnostics
Run logistic regression	Completed logistic regression with odds ratios and coefficients
Save outputs (tables/plots)	Saved all outputs to /tables, /plots, /outputs
Document results in GitHub	Updated repository structure with code and evidence

I accomplished all of the tasks I had planned.

Self-Rating of Progress: (5 points)

- [Done] Met the Planned Tasks
- [ ] Did Not Meet the Planned Tasks

(If you did not meet the planned tasks, please explain why and how you plan to get back on schedule)

I am on schedule with my timeline and did not encounter major issues.

Plan for Next Week: (5 points)

Refine regression models by considering interaction effects (e.g., education × income).
Explore stratified analysis by demographic groups (age, sex).
Add more polished visualizations (e.g., grouped bar plots, regression coefficient plots).
Continue improving GitHub documentation with clear explanations and markdown notes.

Evidence of Tasks Completed: (15 points)

All code, outputs, and plots are available in my GitHub repository:
https://github.com/abdullahs1357/SeniorThesis.git

- Code: BRFSS2023_regression_week5.R
- Tables: regression coefficients, VIF results, odds ratios
- Plots: residuals vs fitted, Q–Q plot, predicted probability histogram
- Outputs: linear and logistic regression summaries