Project 2 Rubric: Data Transformation (90 points)

Use three datasets chosen in Week 5 Discussion (3 points)

Data storage (5 points by dataset)

Create a .csv file containing the information (4 points)
Using “wide” data structure (1 point)
Or, optionally, an SQL DB (extra credit: 1 point)
If DB, normalized (long) data structure (extra point: 1 point)
Implement Database in the Cloud (e.g. on AWS RDS) (extra credit: 1 point)

Importing and Preparing Data (14 points by dataset)

Read the information from your .CSV file into R (4 points)
Use tidyr as needed to tidy your data. (5 points)
Use dplyr as needed to transform your data. (5 points)

Reproducibility (2 points by dataset)

Data is available on the web (1)
Using R Markdown text and headers (1 points)

Data Analysis (3 points by dataset)

Include publication quality graphics or tables (2 points)
Find patterns or trends (1 point)

Documentation (3 points by dataset)

Include an overview of your approach, that describes the assigned problem and clearly explains the workflow (1 point)

Include an appropriate citation to the sourced data (using APA or MLA citation standards) (1 point)

Provide a conclusion (including any findings and recommendations) (1 point)

Submission (2 points by dataset)

Publish to rpubs and provide a link in your assignment submission (1 point)

Publish to GitHub and provide a link in your assignment submission (1 point)