Proposal

Songyi Ahn

2025-01-27

Research Questions

1. How does socioeconomic status (SES) influence individuals' perceptions of their own weight?
2. How does SES influence engagement in various weight management behaviors, and how is this behavior associated with their weight perception?
3. How does SES influence individuals' motivations and frequency of checking calorie and nutrition labels in stores, and how is this behavior associated with their weight perception?
4. How does SES influence individuals' choices of dining locations, and how is this decision associated with their weight perception?

Dataset Name and Keys for Joining the Data

National Health and Nutrition Examination Survey (NHANES) 2017-March 2020 from the National Center for Health Statistics of Centers for Disease Control and Prevention. Datasets are divided into several files based on NHANES's classification across demographic data, examination Data (body measures), and Questionnaire Data (consumer behavior, diet behavior & nutrition, health insurance, hospital utilization & access to care, income, weight history). However, all data have respondent sequence numbers, which will help join the data, such as using the inner_join function in R.

```r
library(rio)
library(here)
demo <- import(here("data", "P_DEMO.xpt"))
head(demo)
```

```
##      SEQN SDDSRVYR RIDSTATR RIAGENDR RIDAGEYR RIDAGEMN RIDRETH1 RIDRETH3
## 1 109263       66        2        1        2       NA        5        6
## 2 109264       66        2        2       13       NA        1        1
## 3 109265       66        2        1        2       NA        3        3
## 4 109266       66        2        2       29       NA        5        6
## 5 109267       66        1        2       21       NA        2        2
## 6 109268       66        1        2       18       NA        3        3
##   RIDEXMON DMDBORN4 DMDYRUSZ DMDEDUC2 DMDMARTZ RIDEXPRG SIALANG SIAPROXY
## 1        2        1       NA       NA       NA       NA       1        1
## 2        2        1       NA       NA       NA       NA       1        1
## 3        2        1       NA       NA       NA       NA       1        1
## 4        2        2        2        5        3        2       1        2
## 5       NA        2        3        4        3        3       1        2
## 6       NA        1       NA       NA       NA       NA       1        2
##   SIAINTRP FIALANG FIAPROXY FIAINTRP MIALANG MIAPROXY MIAINTRP AIALANGA
## 1        2       1        2        2      NA       NA       NA       NA
## 2        2       1        2        2       1        2        2        1
## 3        2       1        2        2      NA       NA       NA       NA
```

```
## 4         2         1         2         2         1         2         2         1
## 5         2         1         2         2        NA        NA        NA        NA
## 6         2         1         2         2        NA        NA        NA        NA
##     WTINTPRP   WTMECPRP SDMVPSU SDMVSTRA INDFMPIR
## 1  7891.762   8951.816       3      156     4.66
## 2 11689.747  12271.157       1      155     0.83
## 3 16273.826  16658.764       1      157     3.06
## 4  7825.646   8154.968       2      168     5.00
## 5 26379.992      0.000       1      156     5.00
## 6 19639.221      0.000       1      155     1.66
```

```r
whistory <- import(here("data", "P_WHQ.xpt"))
head(whistory)
```

```
##      SEQN WHD010 WHD020 WHQ030 WHQ040 WHD050 WHQ060 WHQ070 WHD080A WHD080B
## 1 109266     64    210      1      2    200     NA      1      NA      11
## 2 109267     63    130      3      1    125     NA      2      NA      NA
## 3 109268     65    120      3      3    120     NA      2      NA      NA
## 4 109271     72    222      3      3    180     NA      2      NA      NA
## 5 109273     72    165      2      1    155     NA      2      NA      NA
## 6 109274     75    219      1      2    230      1     NA      NA      NA
##   WHD080C WHD080D WHD080E WHD080F WHD080G WHD080H WHD080I WHD080J WHD080K
## 1      NA      13      NA      NA      NA      NA      NA      NA      NA
## 2      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 3      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 4      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 5      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 6      NA      NA      14      NA      NA      NA      NA      NA      NA
##   WHD080M WHD080N WHD080O WHD080P WHD080Q WHD080R WHD080S WHD080T WHD080U
## 1      34      30      NA      NA      NA      44      NA      46      NA
## 2      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 3      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 4      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 5      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 6      NA      NA      NA      NA      NA      NA      NA      NA      NA
##   WHD080L WHQ225 WHD110 WHD120 WHD130 WHD140 WHQ150
## 1      NA      2     NA    210     NA    218     28
## 2      NA      5     NA     NA     NA    148     19
## 3      NA      5     NA     NA     NA    135     11
## 4      NA      4    280    350     NA    350     24
## 5      NA      2    150    155     NA    300     16
## 6      NA      1    240    190     75    270     54
```

Some Documentation that I Have Played with the Data

NHANES data has responses that represent socioeconomic status, such as income, occupation, education level, ratio of family income to poverty, and health insurance status. Weight-related variables are diverse; diet behaviors, weight management activities, eating location, calorie or nutrition checking, weight change intention, and weight perception. Even though they are scattered in different dataset files, they can be joined and pivoted into tidy data.

```r
library(dplyr)
inner_joined <- inner_join(demo, whistory, by = "SEQN")
head(inner_joined)
```

```
##      SEQN SDDSRVYR RIDSTATR RIAGENDR RIDAGEYR RIDAGEMN RIDRETH1 RIDRETH3
## 1 109266       66        2        2       29       NA        5        6
## 2 109267       66        1        2       21       NA        2        2
## 3 109268       66        1        2       18       NA        3        3
## 4 109271       66        2        1       49       NA        3        3
## 5 109273       66        2        1       36       NA        3        3
## 6 109274       66        2        1       68       NA        5        7
##   RIDEXMON DMDBORN4 DMDYRUSZ DMDEDUC2 DMDMARTZ RIDEXPRG SIALANG SIAPROXY
## 1        2        2        2        5        3        2       1        2
## 2       NA        2        3        4        3        3       1        2
## 3       NA        1       NA       NA       NA       NA       1        2
## 4        2        1       NA        2        3       NA       1        2
## 5        2        1       NA        4        3       NA       1        2
## 6        1        1       NA        4        3       NA       1        2
##   SIAINTRP FIALANG FIAPROXY FIAINTRP MIALANG MIAPROXY MIAINTRP AIALANGA
## 1        2       1        2        2       1        2        2        1
## 2        2       1        2        2      NA       NA       NA       NA
## 3        2       1        2        2      NA       NA       NA       NA
## 4        2       1        2        2       1        2        2        1
## 5        2       1        2        2       1        2        2        1
## 6        2       1        2        2       1        2        2        1
##    WTINTPRP  WTMECPRP SDMVPSU SDMVSTRA INDFMPIR WHD010 WHD020 WHQ030 WHQ040
## 1  7825.646  8154.968       2      168     5.00     64    210      1      2
## 2 26379.992     0.000       1      156     5.00     63    130      3      1
## 3 19639.221     0.000       1      155     1.66     65    120      3      3
## 4  8481.590  8658.733       1      167       NA     72    222      3      3
## 5 20171.848 22163.597       1      155     0.83     72    165      2      1
## 6  7227.993  7801.600       2      167     1.20     75    219      1      2
##   WHD050 WHQ060 WHQ070 WHD080A WHD080B WHD080C WHD080D WHD080E WHD080F WHD080G
## 1    200     NA      1      NA      11      NA      13      NA      NA      NA
## 2    125     NA      2      NA      NA      NA      NA      NA      NA      NA
## 3    120     NA      2      NA      NA      NA      NA      NA      NA      NA
## 4    180     NA      2      NA      NA      NA      NA      NA      NA      NA
## 5    155     NA      2      NA      NA      NA      NA      NA      NA      NA
## 6    230      1     NA      NA      NA      NA      NA      14      NA      NA
##   WHD080H WHD080I WHD080J WHD080K WHD080M WHD080N WHD080O WHD080P WHD080Q
## 1      NA      NA      NA      NA      34      30      NA      NA      NA
## 2      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 3      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 4      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 5      NA      NA      NA      NA      NA      NA      NA      NA      NA
## 6      NA      NA      NA      NA      NA      NA      NA      NA      NA
##   WHD080R WHD080S WHD080T WHD080U WHD080L WHQ225 WHD110 WHD120 WHD130 WHD140
## 1      44      NA      46      NA      NA      2     NA    210     NA    218
## 2      NA      NA      NA      NA      NA      5     NA     NA     NA    148
## 3      NA      NA      NA      NA      NA      5     NA     NA     NA    135
## 4      NA      NA      NA      NA      NA      4    280    350     NA    350
## 5      NA      NA      NA      NA      NA      2    150    155     NA    300
## 6      NA      NA      NA      NA      NA      1    240    190     75    270
##   WHQ150
## 1     28
## 2     19
## 3     11
## 4     24
```
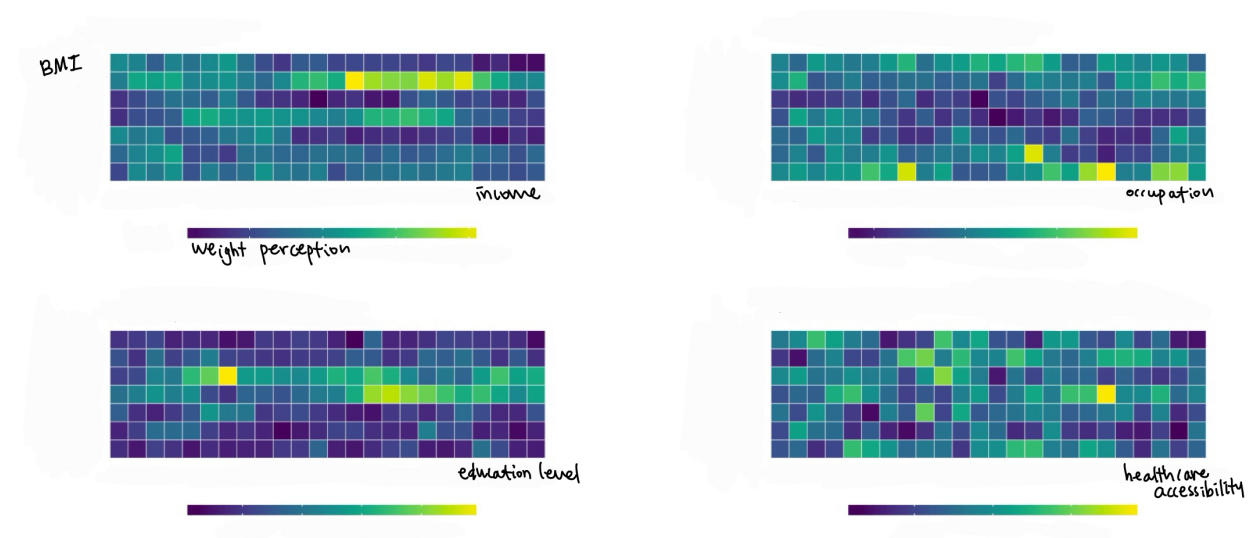
```
## 5      16
## 6      54
```

Preliminary Ideas for Data Visualizations

All visualizations are simple sketches created by borrowing similar visual charts/graph images and manually labeling them with the variables relevant to current project.
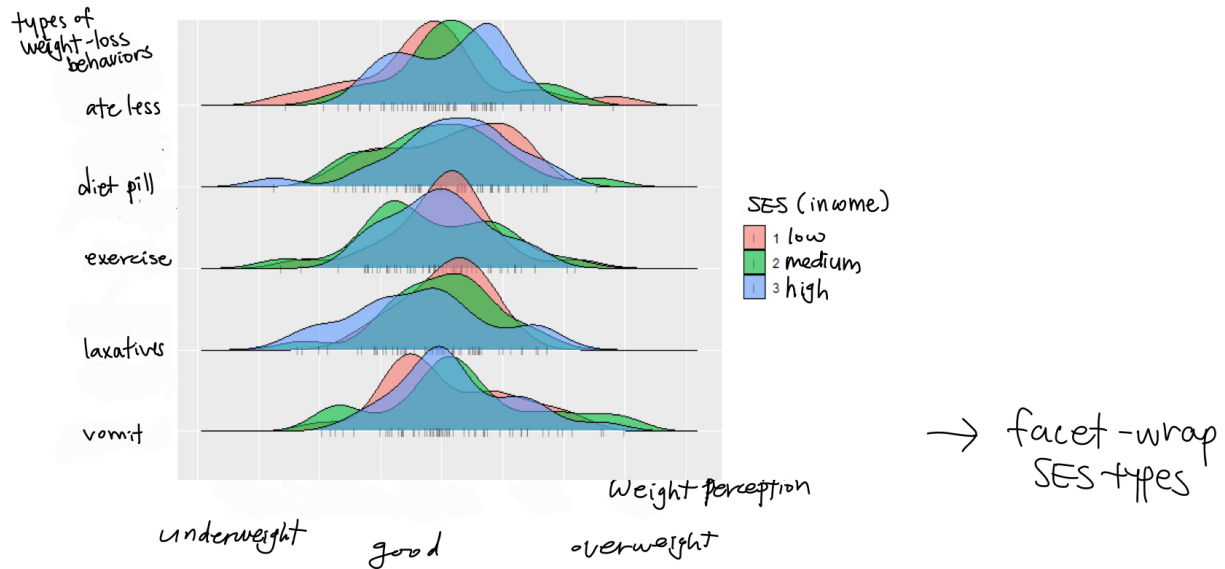
1. SES and Weight Perception

Heatmap

Intended Audience: Social scientists and researchers

Intended Message: Individuals with higher SES might be more likely to perceive themselves as overweight, regardless of their BMI, as thinness is often seen as a form of social privilege or cultural capital in modern society.

2. SES, Weight Management Behaviors and Weight Perception
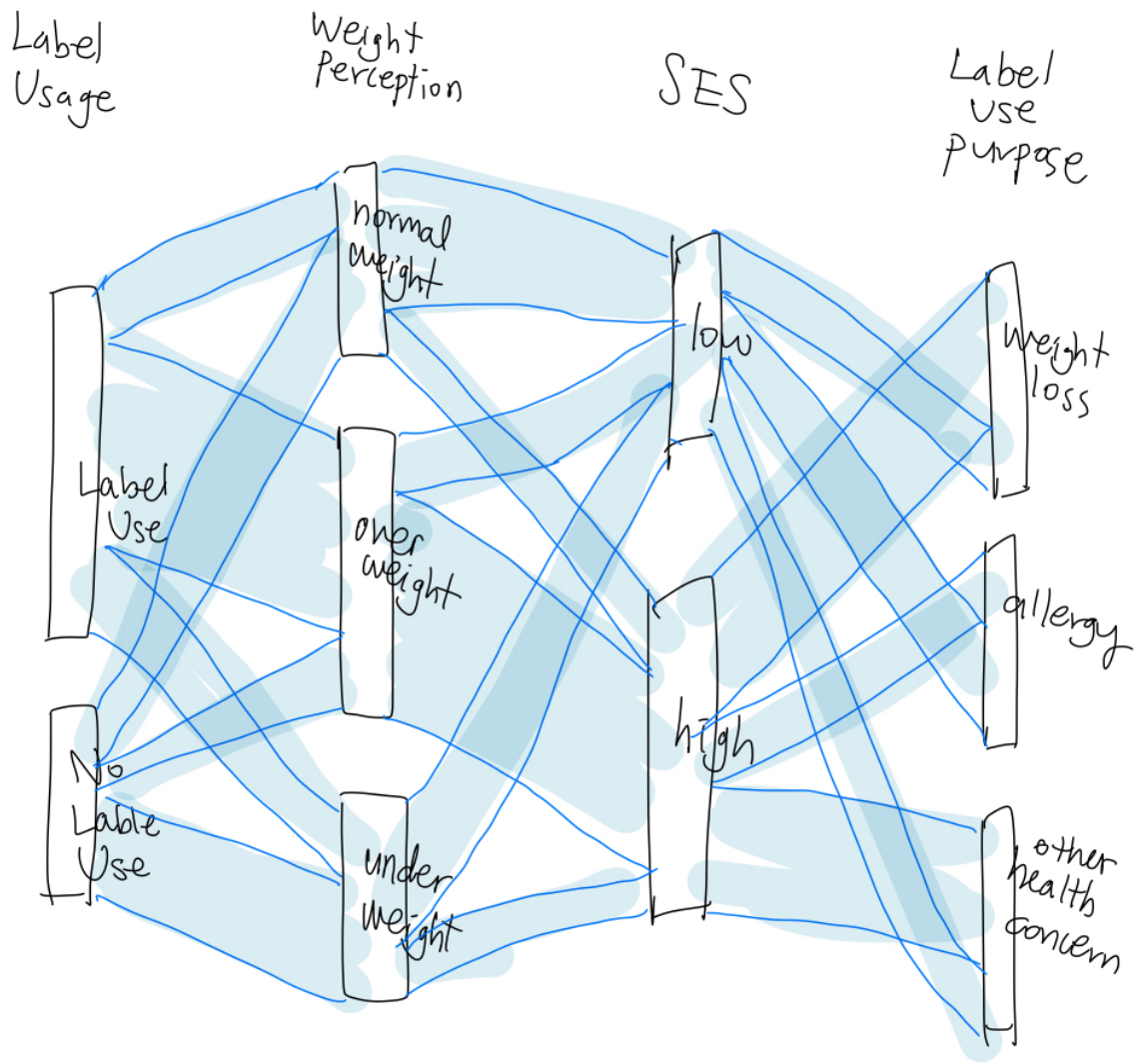
Ridgeline Plot

Intended Audience: Healthcare professionals

Intended Message: Individuals from different SES might engage in varying types of weight management behaviors, and appropriate interventions based on their social background are needed to address potentially harmful behaviors.

3. SES, Nutritional Label Usage, and Weight Perception

Sankey diagram (I need more feedback on this visualization. When I first learned Sankey diagram, I thought it would be so cool, but now it looks so complicated and messy.)
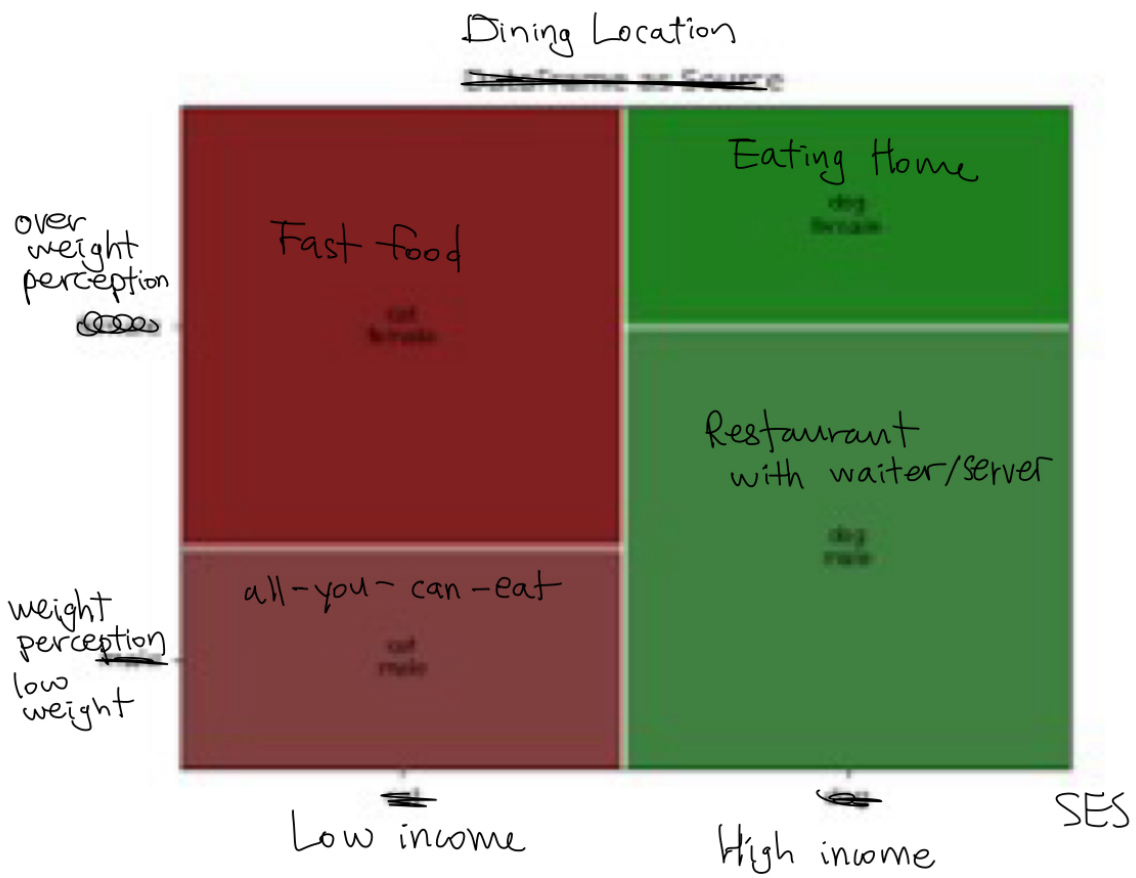
Intended Audience: Public health officials, nutritionists

Intended Message: Individuals from different SES use nutrition/calorie labels for varying reasons, with certain SES groups more likely to internalize societal thinness ideals by prioritizing weight control motivation.

4. SES, Dining Location Choice, and Weight Perception

Mosaic plot

Intended Audience: Public health officials, social scientist and researchers

Intended Message: Individuals from different SES make dining location choices based on their weight perception, societally pressured by thinness ideals.