EDA GSS Lab

Gwen Thompson

GSS Lab Written Parts

**Description of Data**

For this exploratory data analysis lab, I have selected a subset of variables from the General Social Survey dataset that focus on happiness, health, work satisfaction, and socioeconomic status. My primary goal is to examine how demographic and financial factors and work and health conditions impact overall well-being and happiness.

Selected Variables:

The variables chosen for this analysis capture key aspects of mental well-being, personal health, job satisfaction, and socioeconomic status, including:

1. Personal & Demographic Information:
   1. Sex -gender of the respondent
   2. Year- year in which the survey was conducted
   3. Number of Children (childs) - Total number of children the respondent has
2. Socioeconomic Status:
   1. Total family income (income) - measure the financial standing of the respondent’s household
3. Mental & Emotional Well-Being
   1. General happiness (happy) - self-reported happiness level
   2. Happiness in marriage (hapmar) -happiness level in marriage, providing insight into personal relationships
4. Health & Job Satisfaction
   1. Condition of health (health) -self-reported overall health status
   2. Work satisfaction (satjob) -satisfaction with current job, reflecting work-life balance and career fulfillment
5. Identification & Metadata:
   1. Ballot used for interview (ballot) -Used to track survey methodology
   2. Respondent ID Number (id) -unique identifier for respondents

Why This Data?

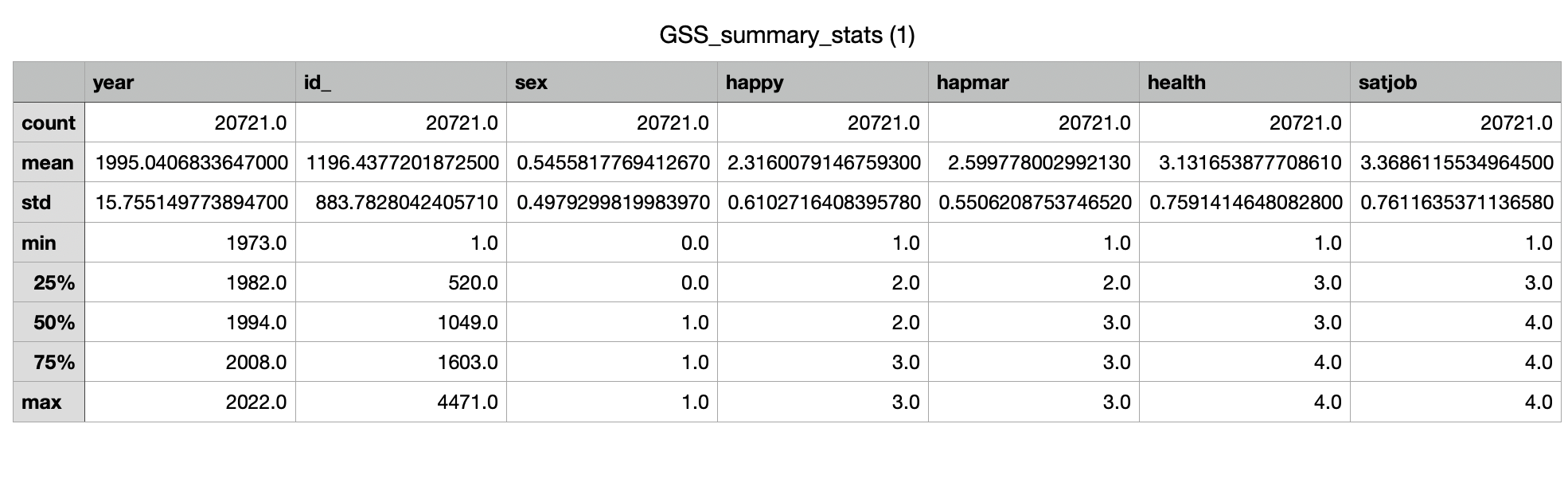
I selected these variables because they provide a holistic view of well-being, incorporating health, income, work satisfaction, and happiness levels. Many interconnected factors influence happiness, and through this dataset, I plan to explore relationships among these variables. By analyzing these variables, I aim to discover how different personal, professional, and financial factors contribute to happiness and well-being.

Conclusion

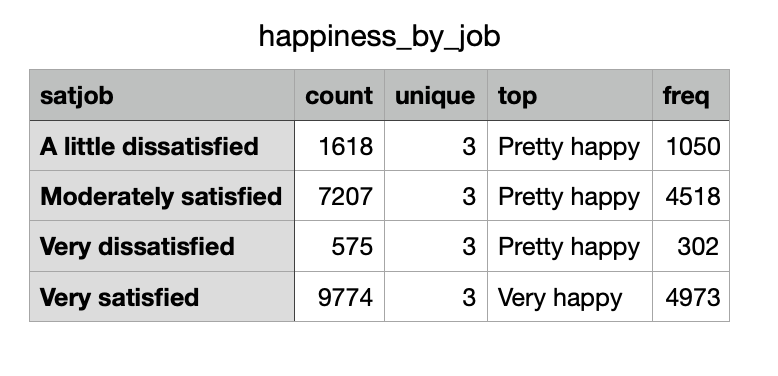
This dataset offers valuable opportunities to explore the intersection of socioeconomic conditions, health, and happiness. Through cleaning, summarization, and visualization, I aim to identify key trends and understand how different aspects of life impact overall well-being. These insights could contribute to mental health awareness, workplace policies, and social well-being distinctions.

**Summaries and Visualisation of Data**

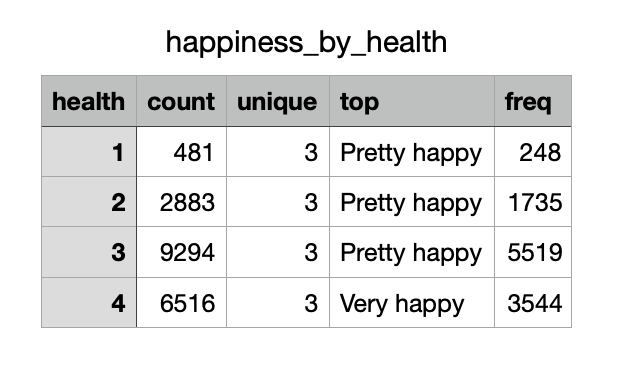
### **Summary of GSS Dataset Statistics**

****This table presents descriptive statistics for key variables in the GSS dataset, including happiness, marital happiness, health, and job satisfaction. The mean happiness score is 2.32 (between "Pretty Happy" and "Very Happy"), while self-reported health averages 3.13 (close to "Good" health). The data also indicates that job satisfaction is relatively high, with a mean score of 3.37 (approaching "Very Satisfied"). These statistics provide an overview of general well-being and work-life satisfaction among respondents.

### **Happiness by Job Satisfaction**

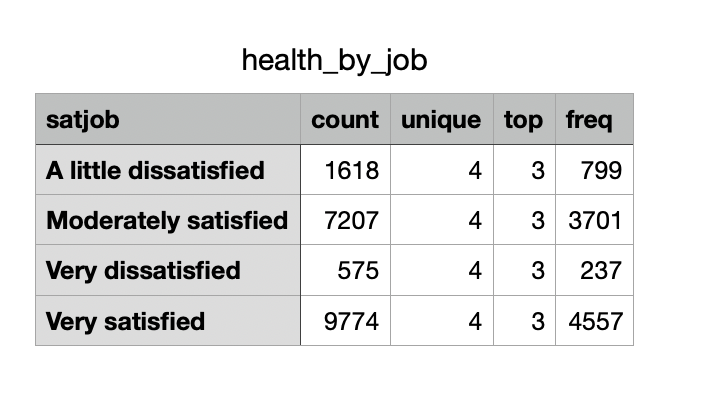
****This table summarizes happiness levels across job satisfaction categories. The most common response for each group is: "Very Happy" among those who are "Very Satisfied" with their jobs and "Pretty Happy" among those who are moderately or slightly dissatisfied. Interestingly, even among those who are "Very Dissatisfied," "Pretty Happy" is the most frequent response, though the number of "Very Happy" respondents is much lower. This suggests that while job satisfaction positively correlates with happiness, some people may maintain a high level of happiness despite workplace dissatisfaction.

**Happiness by Self-Reported Health**

****

This table shows the relationship between happiness and health status. The data reveals a clear trend. Individuals who rated their health as 4 (Excellent) are most frequently "Very Happy." Those with lower health ratings (1 = Poor, 2 = Fair) were more likely to be "Pretty Happy" rather than "Very Happy." This indicates that better health is associated with greater happiness, reinforcing the idea that physical well-being plays a key role in overall life satisfaction.

**Health by Job Satisfaction**

****

This table summarizes the distribution of self-reported health across different job satisfaction levels. The most common health rating (mode) is 3 (Good) across all satisfaction levels, with those who are "Very Satisfied" at work having the highest count of responses. This suggests that job satisfaction may be associated with better self-reported health.

**Distribution of Happiness Levels (Bar Chart)**

****

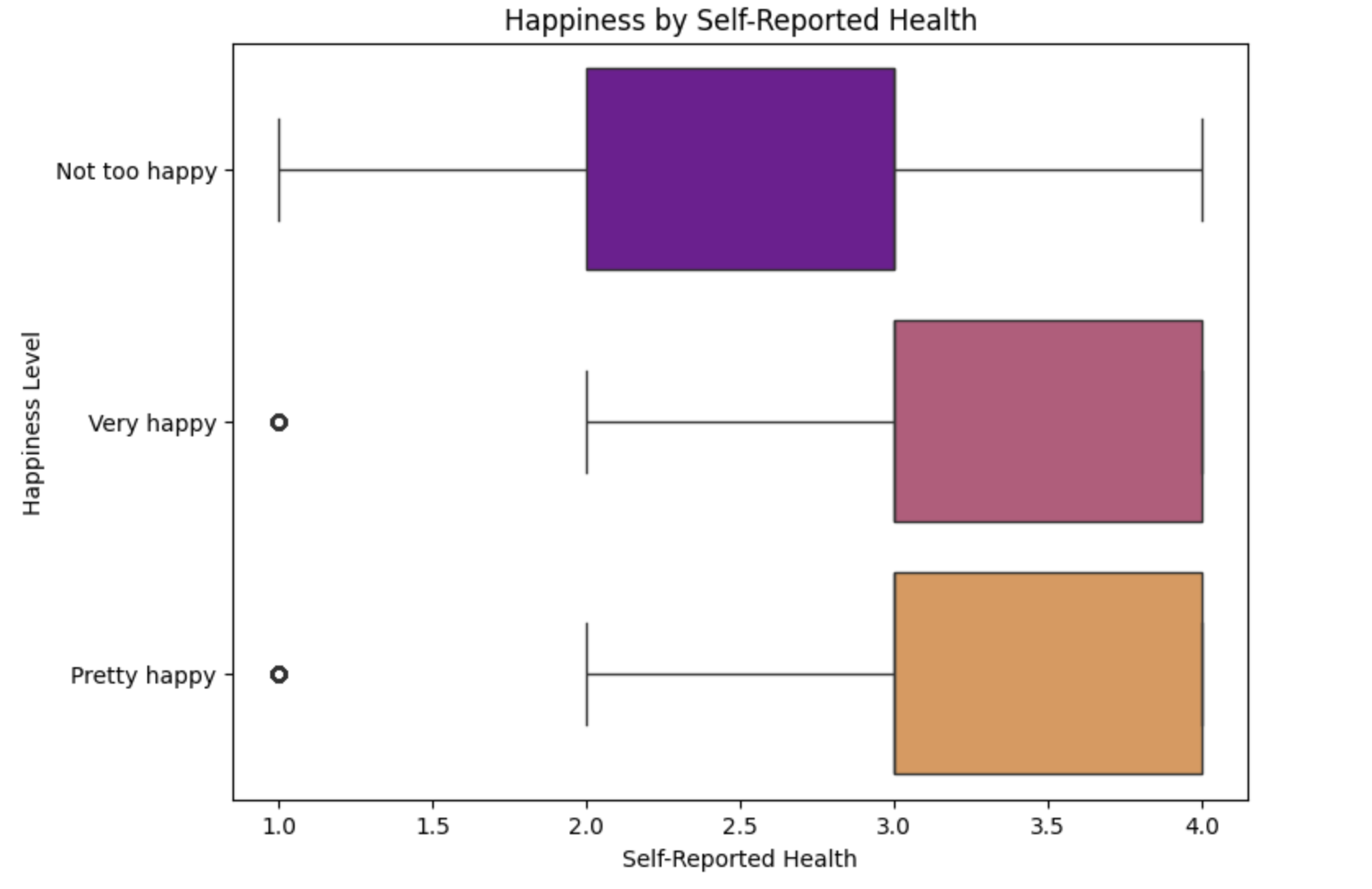
This bar chart visualizes the overall distribution of happiness levels in the dataset. Most respondents reported being "Pretty Happy" or "Very Happy", while a smaller proportion reported being "Not Too Happy". This suggests that, on average, most individuals in the survey report moderate to high happiness levels.

**Happiness by Job Satifaction (Boxplot)**

****

This boxplot displays the distribution of happiness levels across different job satisfaction categories. The data suggests that higher job satisfaction corresponds to higher happiness levels. Those who are very dissatisfied with their jobs exhibit a wide range of happiness levels, including many who are not happy at all, while those who are very satisfied tend to report higher happiness levels.

**Happiness by Self-Reported Health (Boxplot)**

****

This boxplot illustrates the relationship between happiness levels and self-reported health. The general trend shows that individuals who report better health tend to have higher happiness levels. Those who are "Very Happy" tend to have better self-reported health, while those who are "Not Too Happy" exhibit a wider range of health conditions, including poorer health.

**Overall Findings**

In this analysis of the GSS, I explored the relationships between happiness, self-reported health, and job satisfication. Using a cleaned dataset of over 20,000 respondent, I examined numeric summaries and visulizations to understand how well-being indictars interact.

**Happiness and Self-Reported Health**

My findings show a strong positive correlation between self-reported health and happiness. Respondents who rate their health as “excellent” were far more likely to identify as “Very Happy,” while those reporting “poor” or “fair” health had lower happiness scores. This suggests that physical well-being plays a significant rolle in overall life satisfaction. The boxplot analysis confirm this trend, showing a gradual increase in happiness levels as health imporves.

**Happiness and Job Satisfaction**

A similar trend was observed between happiness and job satisfaction. Individuals who reported being "Very Satisfied" with their jobs were more likely to be "Very Happy," whereas dissatisfaction at work corresponded with lower happiness levels. Interestingly, even among those who were "Very Dissatisfied," the most frequent happiness level was still "Pretty Happy," suggesting that while job satisfaction impacts happiness, it is not the sole determining factor.

**Numeric Summaries and Correlations**

My statistical summaries reinforce these findings. The mean self-reported happiness score was 2.32, aligning between "Pretty Happy" and "Very Happy," while job satisfaction had a mean of 3.37, close to "Very Satisfied." The analysis revealed moderate positive correlations between happiness, job satisfaction, and self-reported health, confirming that both work and health significantly influence happiness.

**Conclusion**

This analysis underscores the importance of both job satisfaction and health in shaping overall happiness. While a fulfilling job contributes to well-being, good health appears to be a stronger determinant of happiness. Future research could explore additional factors, such as income or social relationships, to gain a more comprehensive understanding of what drives happiness.