Tony Lim
BIOSTAT 203A
Professor Hilary Aralis
15 Nov 2019

Mid-Quarter Project

# Methods

## Initial Steps

First, I read in the provided datasets (*W2007*, *W2011*, and *FAMIDSubset)*. Then I subsetted *FAMIDSUBSET* to only have observations with a SubsetNumber equal to my SubsetNumber (34). Furthermore, I subsetted *W2007* and *W2011* to create *subset2007* and *subset2011*, and each resulting dataset only included observations containing FAMIDS that were within *FAMIDSubset*.

## Demographic and Employment Characteristics of the Sample in 2007

For **Highest Level of Education**, I created a new character variable that was filled with the education categories depending on the values. I also combined certain categories:

- High school or less (included (1) Elementary or junior high and (2) High school or GED)
- Some college (included (4) Associate degree and (5) Some college)
- Graduate degree (included (7) Master's degree and (8) Ph.D. or professional)

For **How secure is your primary job?** and **How satisfied with your job as a whole?**, I subsetted *subset2007* to create *employed2007* that included only individuals who indicated being currently employed (either part-time or full-time).

For **How satisfied with your job as a whole?**, I created a new character variable in *employed2007* that was filled with the satisfaction categories depending on the column values. I also combined the outer categories:

- Extremely or very dissatisfied (included (1) Extremely dissatisfied and (2) Very dissatisfied)
- Extremely or very satisfied (included (5) Very satisfied and (6) Extremely satisfied)

For **Age in years**, I created a new variable in *subset2007* containing the individual's birthdate by pasting the values from their birth month and birth year with the assumption that each individual was born on the first day of their birth month. Then I created another variable which calculated age in years by taking the difference between the birthdate variable and October 1, 2007 and dividing it by 365.25.

## Mental and Physical Health by Employment Characteristics for the 2007 Sample

For **Body Mass Index (BMI)**, I created a new variable in *subset2007* containing the BMI for each individual, calculated by dividing current weight in pounds by the total height in inches (calculated by combining the current height in feet and current height in inches) squared and multiplying the result by 703.

For **Mental Health Total Score**, I reversed the select items in both *subset2007* and *employed2007*. In both datasets, I created a new variable calculating the mental health total score by summing up the 15 items.

When tabulating **Body Mass Index (BMI)** and **Mental Health Total Score** by **Highest Level of Education** and **Currently employed (either part-time or full-time)**, I used *subset2007*. When tabulating the two variables by **How secure is your primary job?** and **How satisfied with your job as a whole?**, I used *employed2007* to only include individuals who were currently employed (either part time or full-time).

## Comparing Employment Characteristics between 2007 and 2011

First, I used the merge function on the 2007 and 2011 subsetted datasets (*subset2007* and *subset2011*) by FAMID and created a new dataset called *combine*, which retained individuals who were present in both datasets. 129 individuals completed a survey in 2007 but not 2011. 70 individuals completed a survey in 2011 but not 2007.

For **How much stress have you felt in meeting financial obligations?**, all individuals had a non-missing response in both years. For **How difficult is it for you to pay your bills on time?**, I filtered out *combine* to create a new subset called *paybill* that only included individuals with responses in both years and used that to tabulate.

For **Income as a Ratio of Poverty Threshold**, I first recreated the poverty threshold table in Microsoft Excel and created a file called poverty.csv and read it in. I filtered *combine* to create a subset called *irpt* to make sure individuals had responses in the annual household income and household size variables for both years. In irpt, I created two new variables containing the poverty threshold from 2006 and 2010. Then I created two new variables (ratio2007 and ratio2011) containing the income as a ratio of poverty threshold for 2007 and 2011 (calculated using the given formula).

## Employment-Related Concerns about the Future in 2007 and 2011

To find the frequencies for variables in 2007 and 2011, I used the table function and applied it to all the given variables.

## Bar Chart Depicting Average Number of Hours Worked by Day of Week in 2007 and 2011

First, for all the days, I replaced the hour values of -8 and -9 with NA. Then I converted the hour values from character to numeric type. For each day of the week, I created a vector filled with the mean hours worked for 2007 and 2011. Then I combined all the seven vectors into one column in a new data frame. I then created columns for day of the week, year, and number of records.

I would suggest replacing hour values greater than 24 with NA since it is not possible to work more than 24 hours in a day, although I did not perform this step when analyzing the data since this fact was brought to my attention later.

# Results

## Demographic and Employment Characteristics of the Sample in 2007

The 2007 sample was slightly skewed towards female (56.0%) (Table 1A). Individuals with technical or vocational degrees and graduate degrees tended to be underrepresented compared to the other categories (12.4% and 10.1%, respectively) (Table 1B). The majority of individuals were currently employed (68.4%) (Table 1C), were currently married or cohabiting in an intimate relationship (68.4%) (Table 1D), and had children (71.9%) (Table 1E). The majority of individuals in the sample felt that their job was either secure or very secure (totaling 70.8%) (Table 1F) and were somewhat satisfied or extremely or very satisfied with their jobs (totaling 77.4%) (Table 1G). The average age of the sample was 34.7 years old (Table 1H). The average annual household income was $71,620 (Table 1I).

## Mental and Physical Health by Employment Characteristics for the 2007 Sample

Individuals with a higher education level correlated with both a lower mean and median BMI level (Table 2A). Similarly, individuals with a higher education level correlated with both a lower mean and median mental health total score (Table 2B). However, individuals with a technical or vocational education have a mean mental health total score lower than those with some college education and have the same median mental health total score as individuals with a bachelor's degree (Table 2B). It is possibly there is some confounding protecting variable for those with a technical or vocational degree. Individuals with a higher primary job security correlated with a higher mean BMI score and lower mean/median mental health total score (Table 2E, F). However, there did not appear to be a correlation between primary job security and median BMI score (Table 2E). There was not an apparent trend between job satisfaction and mean/median BMI (Table 2G). For job satisfaction and mean/median mental health total scores, the extremes (extremely or very dissatisfied and extremely or very satisfied) had similar scores that were lower than the moderate job satisfaction categories (Table 2H).

## Comparing Employment Characteristics between 2007 and 2011

From 2007 to 2011, the average stress felt in meeting financial obligations rose slightly from 4.29 to 4.33 (a value of 4 meant moderately stressful) (Table 3A). From 2007 to 2011, the average difficulty in paying bills on time also rose slightly from 3.13 to 3.27 (a value of 4 meant moderately difficult) (Table 3B). The percentage of those currently employed (either part-time or full-time) decreased from 85.9% to 82.2% (Table 3C). From 2007 to 2011, income as a ratio of poverty threshold decreased from 4.37 to 4.04, indicating a decrease in annual household income compared to the poverty threshold (Table 3D).

## Employment-Related Concerns about the Future in 2007 and 2011

From 2007 to 2011, more individuals expressed more employment-related concerns about the future for some variables. For example, concern about lack of ability to get training or degree increased from 12.7% to 16.2%. Concern about being considered "overqualified" increased from 3.8% to 9.1%. Concern about lack of openings in the field increased from 13.1% to 30.2%. Concerns about illness, accident, or disability increased from 14.4% to 21.1% (Table 4A, B)

## Bar Chart Depicting Average Number of Hours Worked by Day of Week in 2007 and 2011

From 2007 to 2011, the average number of hours worked by day of week dropped for all days of the week. The decrease in average number of hours worked was especially severe for the weekdays. For Saturday, it dropped from 4.59 hours to 1.47 hours. For Sunday, it dropped from 3.84 hours to 1.20 hours.

## Table 1. Demographic and Employment Characteristics of the Sample in 2007

| Gender | | |
|---|---|---|
| | Frequency | Percentage |
| Female | 367 | 56.0% |
| Male | 288 | 44.0% |
| **Total** | **655** | **100.0%** |

Table 1A.

| Highest Level of Education | | |
|---|---|---|
| | Frequency | Percentage |
| High school or less | 153 | 23.4% |
| Technical or vocational | 81 | 12.4% |
| Some college | 202 | 30.8% |
| Bachelor's degree | 134 | 20.5% |
| Graduate degree | 66 | 10.1% |
| Missing values | 19 | 2.9% |
| **Total** | **655** | **100.0%** |

Table 1B.

| Currently employed (either part-time or full-time) | | |
|---|---|---|
| | Frequency | Percentage |
| Yes | 554 | 84.6% |
| No | 93 | 14.2% |
| Missing values | 8 | 1.2% |
| **Total** | **655** | **100.0%** |

Table 1C.

| Currently married or cohabiting in an intimate relationship | | |
|---|---|---|
| | Frequency | Percentage |
| Yes | 448 | 68.4% |
| No | 195 | 29.8% |
| Missing values | 12 | 1.8% |
| **Total** | **655** | **100.0%** |

Table 1D.

| Do you have any children? | | |
|---|---|---|
| | Frequency | Percentage |
| Yes | 471 | 71.9% |
| No | 179 | 27.3% |
| Missing values | 5 | 0.8% |
| **Total** | **655** | **100.0%** |

Table 1E.

| How secure is your primary job? | | |
|---|---|---|
| | Frequency | Percentage |
| Not at all secure | 25 | 4.4% |
| Somewhat secure | 126 | 22.4% |
| Secure | 200 | 35.6% |
| Very secure | 198 | 35.2% |
| Missing values | 13 | 2.3% |
| **Total** | **562** | **100.0%** |

Table 1F.

| How satisfied with your job as a whole? | | |
|---|---|---|
| | Frequency | Percentage |
| Extremely or very dissatisfied | 55 | 9.8% |
| Somewhat dissatisfied | 60 | 10.7% |
| Somewhat satisfied | 175 | 31.1% |
| Extremely or very satisfied | 260 | 46.3% |
| Missing values | 12 | 2.1% |
| **Total** | **562** | **100.0%** |

Table 1G.

| Age in Years | | |
|---|---|---|
| Number of records | Mean | Standard Deviation |
| 650 | 34.7 | 0.581 |

Table 1H.

| Annual Household Income | | |
|---|---|---|
| Number of records | Mean | Standard Deviation |
| 582 | $71,620.47 | $75,124.66 |

Table 1I.

## Table 2. Mental and Physical Health by Employment Characteristics for the 2007 Sample

| BMI and Highest Level of Education | | | |
|---|---|---|---|
| n = 626 | Mean | Median | Standard Deviation |
| High school or less | 29.6 | 28.3 | 8.09 |
| Technical or vocational | 28.0 | 27.5 | 5.92 |
| Some college | 28.4 | 26.8 | 6.36 |
| Bachelor's degree | 27.0 | 25.8 | 5.78 |
| Graduate degree | 25.8 | 25.1 | 4.30 |

Table 2A.

| Mental Health Total Score and Highest Level of Education | | | |
|---|---|---|---|
| n = 636 | Mean | Median | Standard Deviation |
| High school or less | 41.5 | 41.0 | 10.9 |
| Technical or vocational | 38.3 | 36.0 | 10.1 |
| Some college | 39.5 | 38.0 | 9.20 |
| Bachelor's degree | 36.6 | 36.0 | 8.90 |
| Graduate degree | 35.8 | 35.0 | 7.93 |

Table 2B.

| BMI and Employment Status | | | |
|---|---|---|---|
| n = 626 | Mean | Median | Standard Deviation |
| Currently employed (either part-time or full-time) | 28.2 | 26.8 | 6.65 |
| Not currently employed | 27.7 | 26.8 | 6.35 |
| Missing value | 28.5 | 28.9 | 5.48 |

Table 2C.

| Mental Health Total Score and Employment Status | | | |
|---|---|---|---|
| n = 647 | Mean | Median | Standard Deviation |
| Currently employed (either part-time or full-time) | 38.1 | 37.0 | 9.28 |
| Not currently employed | 43.2 | 40.0 | 11.2 |
| Missing value | 40.4 | 35.5 | 13.0 |

Table 2D.

| BMI and Primary Job Security | | | |
|---|---|---|---|
| n = 529 | Mean | Median | Standard Deviation |
| Not at all secure | 27.2 | 27.2 | 3.16 |
| Somewhat secure | 27.2 | 26.4 | 5.10 |
| Secure | 28.7 | 27.4 | 7.92 |
| Very secure | 28.5 | 27.3 | 6.41 |
| Missing values | 23.9 | 22.7 | 3.50 |

Table 2E.

| Mental Health Total Score and Primary Job Security | | | |
|---|---|---|---|
| n = 549 | Mean | Median | Standard Deviation |
| Not at all secure | 44.9 | 45.0 | 11.1 |
| Somewhat secure | 41.0 | 41.0 | 9.10 |
| Secure | 38.0 | 37.0 | 8.32 |
| Very secure | 35.6 | 34.0 | 9.19 |
| Missing values | 34.6 | 34.0 | 10.0 |

Table 2F.

| BMI and Job Satisfaction | | | |
|---|---|---|---|
| n = 529 | Mean | Median | Standard Deviation |
| Extremely or very dissatisfied | 28.0 | 26.9 | 5.36 |
| Somewhat dissatisfied | 29.0 | 27.1 | 10.3 |
| Somewhat satisfied | 28.2 | 27.0 | 6.18 |
| Extremely or very satisfied | 28.1 | 26.8 | 6.23 |

Table 2G.

| Mental Health Total Score and Job Satisfaction | | | |
|---|---|---|---|
| n = 550 | Mean | Median | Standard Deviation |
| Extremely or very dissatisfied | 35.6 | 34.0 | 9.89 |
| Somewhat dissatisfied | 41.1 | 41.0 | 9.47 |
| Somewhat satisfied | 41.2 | 40.0 | 9.43 |
| Extremely or very satisfied | 35.9 | 35.0 | 8.17 |

Table 2H.

## Table 3. Comparing Employment Characteristics between 2007 and 2011

| How much stress have you felt in meeting financial obligations? | | | |
|---|---|---|---|
| Year | Number of records | Mean | Standard Deviation |
| 2007 | 526 | 4.29 | 1.86 |
| 2011 | 526 | 4.33 | 1.85 |

Table 3A.

| How difficult is it for you to pay your bills on time? | | | |
|---|---|---|---|
| Year | Number of records | Mean | Standard Deviation |
| 2007 | 524 | 3.13 | 1.95 |
| 2011 | 524 | 3.27 | 2.02 |

Table 3B.

| Employment Status | 2007 | | 2011 | |
|---|---|---|---|---|
| | Frequency | Percentage | Frequency | Percentage |
| Currently employed (either part-time or full-time) | 445 | 85.9% | 426 | 82.2% |
| Not currently employed | 73 | 14.1% | 92 | 17.8% |
| **Total** | **518** | **100.0%** | **518** | **100.0%** |

Table 3C.

| Income as a Ratio of Poverty Threshold | | | |
|---|---|---|---|
| Year | Number of records | Mean | Standard Deviation |
| 2007 | 448 | 4.37 | 5.66 |
| 2011 | 448 | 4.04 | 3.30 |

Table 3D.

## Table 4. Employment-Related Concerns about the Future in 2007 and 2011

| Concerned about | 2007 | | | | | |
|---|---|---|---|---|---|---|
| | Circled | | Not circled | | Total | |
| | Frequency | Percentage | Frequency | Percentage | Frequency | Percentage |
| Lack of ability to get training or degree | 67 | 12.7% | 459 | 87.3% | 526 | 100% |
| Lack of money to complete education or get started in my chosen career field | 135 | 25.7% | 391 | 74.3% | 526 | 100% |
| I am considered "overqualified" | 20 | 3.8% | 506 | 96.2% | 526 | 100% |
| Lack of openings in my field | 69 | 13.1% | 457 | 86.9% | 526 | 100% |
| Relocation is difficult or impossible | 96 | 18.3% | 430 | 81.7% | 526 | 100% |
| Illness, accident, or disability | 76 | 14.4% | 450 | 85.6% | 526 | 100% |
| Caring for a sick parent or relative | 83 | 15.8% | 443 | 84.2% | 526 | 100% |
| Transportation problems - difficulty in getting to or from work | 37 | 7.0% | 489 | 93.0% | 526 | 100% |

Table 4A.

| Concerned about | 2011 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Statement applies to me | | Statement does not apply to me | | Missing | | Total | |
| | Frequency | Percentage | Frequency | Percentage | Frequency | Percentage | Frequency | Percentage |
| Lack of ability to get training or degree | 85 | 16.2% | 435 | 82.7% | 6 | 1.1% | 526 | 100% |
| Lack of money to complete education or get started in my chosen career field | 134 | 25.5% | 375 | 71.3% | 17 | 3.2% | 526 | 100% |
| I am considered "overqualified" | 48 | 9.1% | 465 | 88.4% | 13 | 2.5% | 526 | 100% |
| Lack of openings in my field | 159 | 30.2% | 361 | 68.6% | 6 | 1.1% | 526 | 100% |
| Relocation is difficult or impossible | 102 | 19.4% | 412 | 78.3% | 12 | 2.3% | 526 | 100% |
| Illness, accident, or disability | 111 | 21.1% | 405 | 77.0% | 10 | 1.9% | 526 | 100% |
| Caring for a sick parent or relative | 108 | 20.5% | 408 | 77.6% | 10 | 1.9% | 526 | 100% |
| Transportation problems - difficulty in getting to or from work | 46 | 8.7% | 471 | 89.5% | 9 | 1.7% | 526 | 100% |

Table 4B.

## Figure 1. Bar Chart Depicting Average Number of Hours Worked by Day of Week in 2007 and 2011