financial Well Being Analysis*

Xinxiang Gao

November 26, 2024

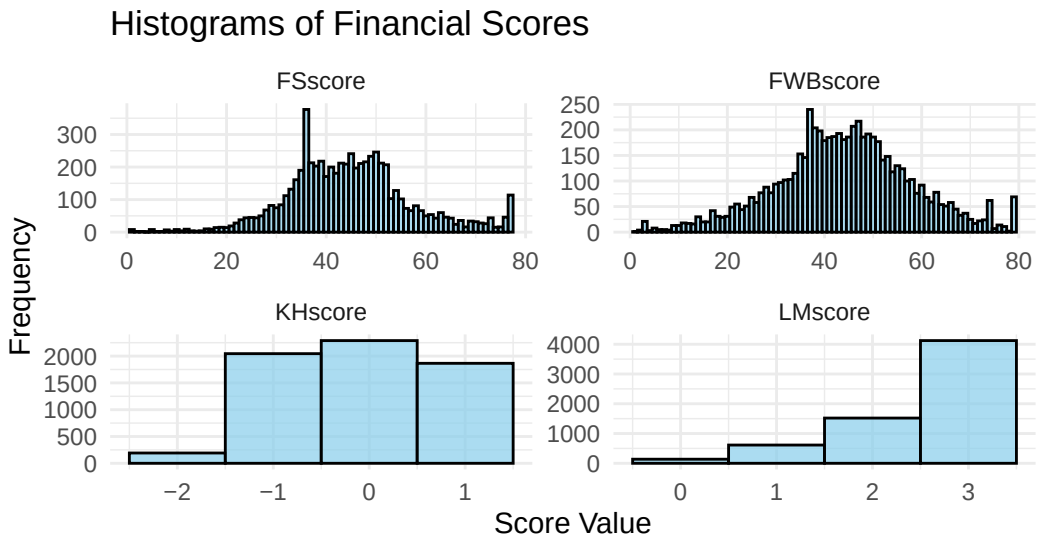
This should be an abstract. The report is still being finalized.

1 Introduction

2 Data

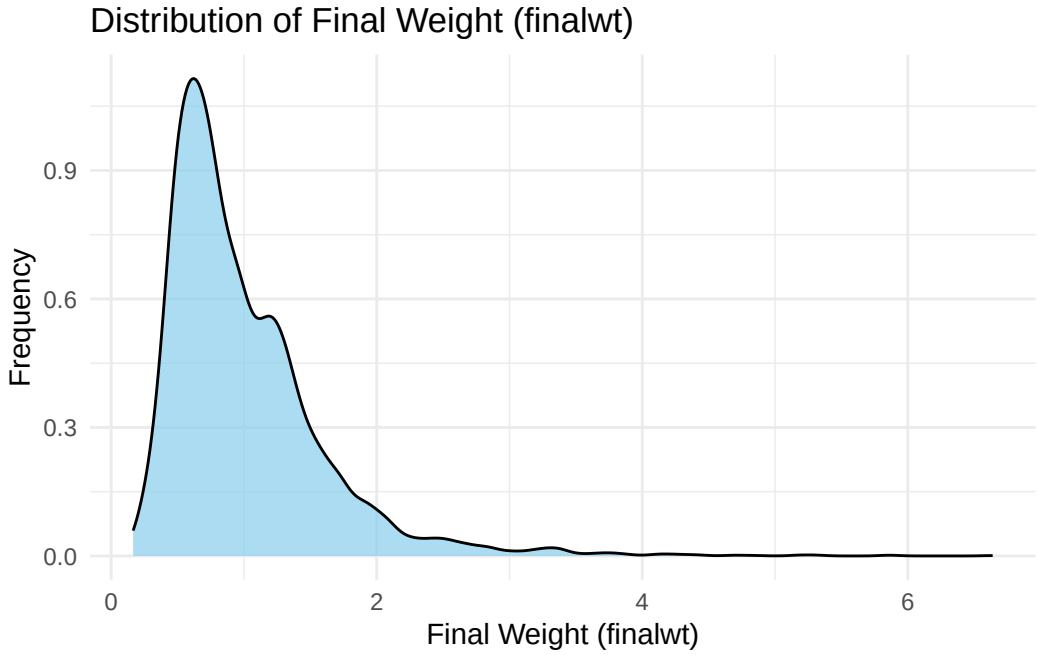
```
-- Data Summary ------------------------
                           Values
Name                       Piped data
Number of rows             6394
Number of columns          5
-----------------------
Column type frequency:
  numeric                  5
-----------------------
Group variables            None

-- Variable type: numeric -------------------------------------------------------
  skim_variable   mean     sd     p0    p25    p50    p75  p100
1 FWBscore      44.0   14.0    1     36     44     53     79
2 FSscore       44.7   12.5    1     36     44     51     77
3 LMscore        2.51   0.755  0      2      3      3      3
4 KHscore       -0.057  0.815 -2.05  -0.57  -0.188  0.712  1.27
5 finalwt        1      0.585  0.166  0.601  0.845  1.25   6.64
```

*Code and data are available at: https://github.com/xgao28/financial_well_being_analysis.



FWBscore: Financial Well−being Scale
FSscore: Financial Skill Scale
LMscore: Lusardi and Mitchell Financial Knowledge Scale
KHscore: Knoll and Houts Financial Knowledge Scale

The composition of finalwt remains on investigation.



3 Model

4 Linear Regression Model

The linear regression model can be expressed mathematically as:

$$y_i = \beta_0 + \beta_1 \cdot \text{LMscore}_i + \beta_2 \cdot \text{KHscore}_i + \beta_3 \cdot \text{FWBscore}_i + \beta_4 \cdot \text{FSscore}_i + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

4.0.1 Model Components

- y_i: The dependent variable, finalwt_i, for observation i.
- β_0: The intercept term (constant in the model).
- $\beta_1, \beta_2, \beta_3, \beta_4$: The slopes (coefficients) of the respective predictor variables:
  - LMscore_i: Lusardi and Mitchell financial knowledge scale score.
  - KHscore_i: Knoll and Houts financial knowledge scale score.
  - FWBscore_i: Financial well-being scale score.
  - FSscore_i: Financial skill scale score.

Table 1: Summary of Linear Regression Model

	Term	Estimate	Std_Error	t_value	p_value
(Intercept)	(Intercept)	1.335	0.042	32.138	0.000
LMscore	LMscore	-0.024	0.012	-2.061	0.039
KHscore	KHscore	-0.135	0.011	-12.094	0.000
FWBscore	FWBscore	-0.006	0.001	-10.547	0.000
FSscore	FSscore	0.000	0.001	-0.151	0.880

5 Results

6 Discussion

7 Appendix

7.1 Data cleaning

7.2 Surveys, sampling, and observational data

7.3 Acknowledgements

We would like to express our gratitude to the developers and contributors of R (R Core Team 2023) as well as several R packages that were essential for the analysis and visualization of the data in this report. The following R packages provided indispensable tools and functionalities:

- `tidyverse` (Wickham et al. 2019): A collection of R packages designed for data science, including dplyr, ggplot2, readr, purrr, and others, which greatly facilitated data manipulation, analysis, and visualization.
- `ggplot2` (Wickham 2016): An implementation of the Grammar of Graphics, which allowed us to create complex and aesthetically pleasing visualizations with ease.
- `knitr` (Xie 2023): This package enabled us to perform data demonstration with tables.
- `styler` (Müller and Walthert 2023): This package is helpful for styling the code.
- `arrow` (Richardson et al. 2024): This package provides a convenient and efficient way to work with parquet format.

We are grateful to the Consumer Financial Protection Bureau (CFPB) for conducting the Financial Well-Being Survey (Consumer Financial Protection Bureau 2017) and making the data publicly available. Their dedication to collecting and sharing such vital information significantly contributes to research on financial well-being.

Finally, we would like to thank all those who contributed to the development and maintenance of the R programming language and its ecosystem, as well as the broader open-source community, whose efforts make such research possible.

References

Consumer Financial Protection Bureau. 2017. "Financial Well-Being Survey Data." https://www.consumerfinance.gov/data-research/financial-well-being-survey-data/.

Müller, Kirill, and Lorenz Walthert. 2023. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.