



# Does Financial Wellness = Happiness?

Tanner Kunz  
Oct. 04, 2021 Cohort



## Problem: Can Financial Wellness data be used to predict Happiness?

- Happiness is complex and not determined the same way for each person.
- Financial wellness impacts many aspects of our lives:
  - Housing
  - Diet
  - Health
  - Downtime/Entertainment
- Poor financial wellness may also lead to stress and decrease happiness.
- **Goal:** Use financial wellness data to accurately predict happiness.



## Data

- Financial Wellness survey conducted by Consumer Financial Protection Bureau ([CFPB](#)) in 2016.
  - Over 6000 respondents
  - Targeted sample demographics to accurately represent the national population
- Included median income per region data from the Federal Reserve Economic Data ([FRED](#)).



## Method

- The survey has 3 main questions about happiness:
  - I am satisfied with my life.
  - I am optimistic about my future.
  - If I work hard today, I will be more successful in my future.
- These questions are rated on a 1-7 scale with 1 being “Strongly Disagree” and 7 being “Strongly Agree.”
- Split the values to “above 6” and “below 5” to balance the target variable.



# Data Cleaning



## Data Wrangling

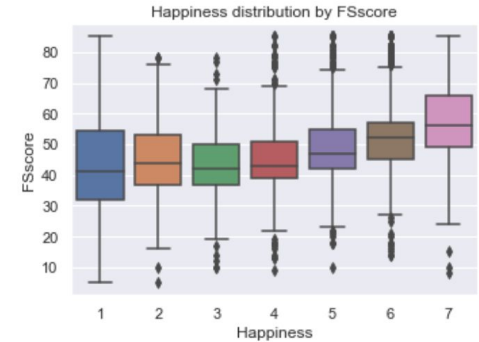
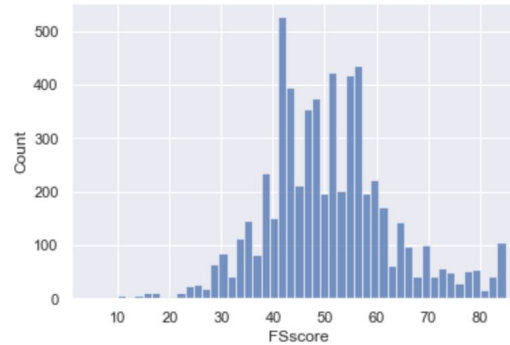
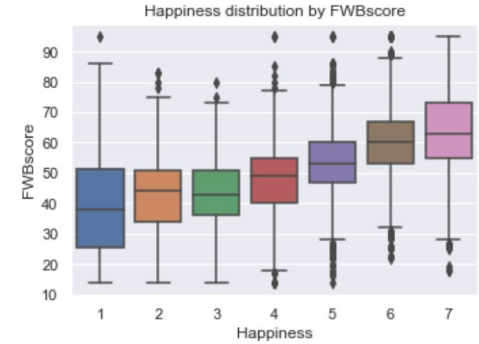
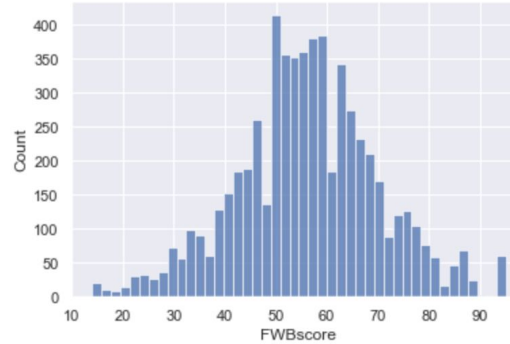
- Started with over 200 features which could lead to overfitting.
- Methods for feature reduction:
  - Combined financial wellness tests to an overall score.
  - Summed instances within the same category.
  - Averaged scores of features in the same category.
  - Removed features that were not relevant to happiness or contained redundant information.
- This resulted in 3 target variables and 61 features.
- Replaced categorical features with descriptions rather than numerical values.
- Renamed features with better descriptions.



# Exploratory Data Analysis (EDA)

# Financial Tests:

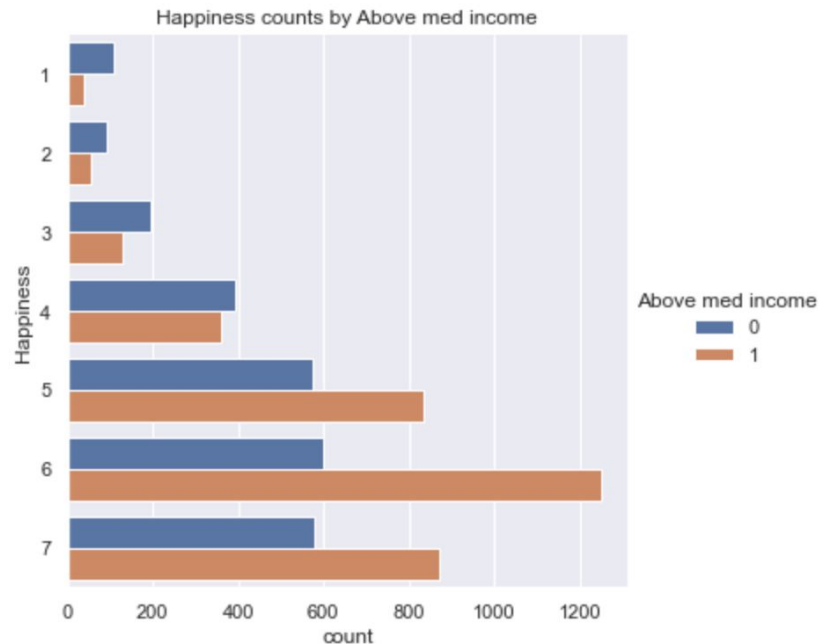
- Both the Financial Well-being and Financial Skill test scores have a normal distribution.
- The higher score looks correlated with a higher happiness.
- T-test showed that people with higher happiness also had a higher average test score for both tests.





## Above Median Income

- The distribution of the happiness seems to generally be higher for those that have an income above the median income level.
- Using a t-test, those that had a higher than median income also had higher happiness than those with income less than the median income.





# Modeling



## Model

- These are the Happiness variable prediction results.
- Logistic regression, random forest, and XGBoost had about the same performance.
- Logistic regression outperformed the other models in computation time.
- The training score and accuracy are approximately the same indicating that the model did not overfit.

Model	Accuracy	Best training score
Logistic Regression	0.748	0.756
Random Forest	0.735	0.745
XGBoost	0.745	0.747



## Results

- The advanced models could not outperform the logistic regression model.
- Most models could accurately predict Happiness for ~75% of respondents.
- Financial wellness is a good indicator for Happiness for most of the sample.
- But only so much Happiness can be explained by financial wellness.
- There are other variables that affect Happiness that are outside the scope of this model that cause the misclassification of the remaining ~25% of the sample.



## Next Steps

- Use clustering algorithms to see if there are underlying trends in the data.
- Analyze the most important features for the respondents that were misclassified.
  - Were they outliers in these categories?
  - Are they similar to each other?
  - Are they vastly different from each other?
- Create a shortened financial wellness survey that questions the most important features and then predicts Happiness.
  - If they are in the lower Happiness criteria then introduce financial wellness resources.