

FORMULA

for Happiness

GOAL

What makes a person happy? While everyone has different personal criteria for happiness, plenty of research exists on attempting to find consensus on the factors that influence the happiness of a population.^{1/2} A common methodology, which we implement, is the analysis of survey data to build a linear model that begins to explain the relationships between self-reported happiness and other factors in life, such as marital status, income, or religiosity. We use clustering and decision tree regression to

support our findings. We highlight factors like health status and financial satisfaction, factors that have considerable “governmental influence”. **The conclusions drawn** could mean that some factors that are adjustable on a large scale are linked to individual happiness, **which could be excellent news for policymakers, business owners, and other societal decision-makers**, helping them to have a different impact across populations.

AT A GLANCE: OUR DATASET

The data analysed are from the GSS - General Social Survey.³ They collect hundreds of opinion/status questions in an annual survey of Americans since 1972. We will limit our data to 15 relevant variables, and only on one year: 2018. This will be to exclude any affect from the COVID pandemic on data collection, and allow us to treat the data as cross-sectional.

Data Description

- 15 variables
- 2348 entries
- Focus: America

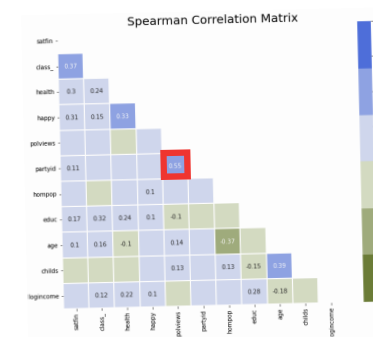
Variables

- **10** categorical VS **5** numerical variables
- TARGET = 'Happy'
- Social, economical, and political variables

GETTING THE MAIN FACTORS



First, we select the most relevant features for determining happiness. Using a correlation matrix and the Lasso method, we decide to drop *Party ID*, which is correlated to another feature. Based on the Chi-squared test, we are dropping 'male' from categorical variables because gender doesn't seem to be impacting the reported level of happiness.



NUMERICAL DATA

Predictor	Spearman	Lasso	Decision Tree	What to do?
Health	Keep	Keep	Keep	Keep
Polviews	Discard	Keep	Keep	Keep
PartyID	Discard	Keep	Keep	Discard
HomePop	Keep	Keep	Keep	Keep
Education	Keep	Discard	Keep	Keep
Age	Keep	Keep	Keep	Keep
Childs	Keep	Keep	Keep	Keep
Logincome	Keep	Keep	Keep	Keep
SatFin	Keep	Keep	Discard	Keep
Class	Keep	Discard	Discard	Keep

CATEGORICAL DATA

Predictor	Chi-Square
Race	Keep
Married	Keep
Male	Discard
Working	Keep
Religious	Keep

MEASURING THE INFLUENCE

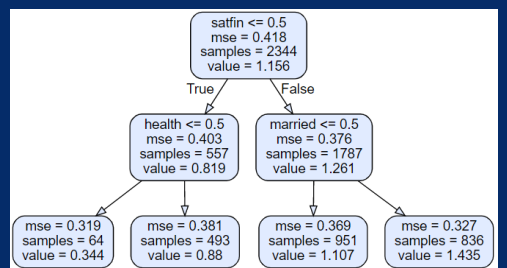
LINEAR REGRESSION

Now we can run a number of linear regressions, taking into account the survey population weights, to understand the influence of our selected variables. Using backwards elimination and weighted least squares,⁴ we find the following **statistically significant** (5%) and **practically significant** factors.

$$\text{Happy} = .176\text{satfin} + .196\text{health} + .299\text{married} + .487$$

DECISION TREE

Our regression findings are supported by a decision tree regression, which uses factors "satfin", "health", and "married" when constrained to a depth of two.

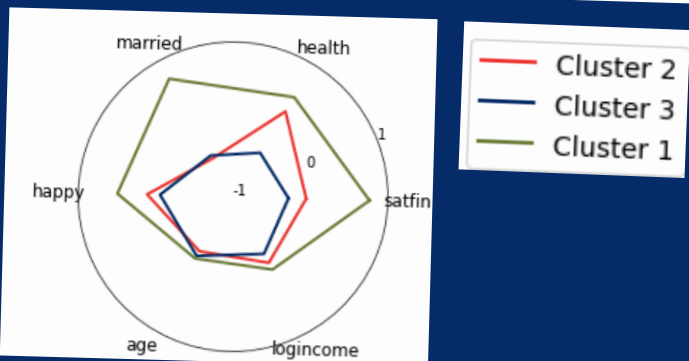


Variables:

- 'satfin': self-reported level of satisfaction with financial situation (0-2)
- 'health': self-reported health level (0-3)
- 'married': dummy variable, marital status = "married"

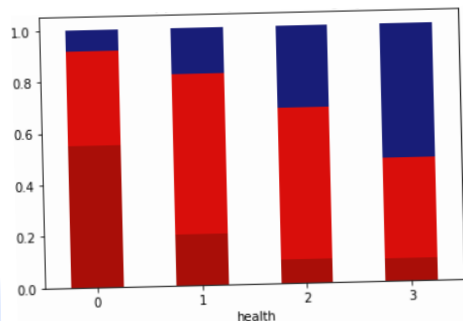
NOW LET'S CLUSTER

From our cluster analysis, we can deepen our understanding. **In Cluster 1**, the happiest people tend to combine very good health, good satisfaction of their financial situation, and are married, reflecting the positive coefficients of our regression variables. **In Clusters 2 and 3**, individuals are less likely to be married, and have lower reported health and financial statuses. **Overall**, we can also see that neither age nor income earned impacts reported happiness.



We can also highlight the nuance between income earned and financial satisfaction. Seeing that the first one has no impact on happiness and the second one does, we can conclude that the perception Americans have on their financial situation is more important than the actual income they earn.

IN CONCLUSION



We see that a number of factors have influence on happiness, ceteris paribus. One factor worth further study is health.

As we can see in the plot above, reported **happiness increases as health level increases**, and we have proven using clustering and regression that the difference is **statistically significant**. If the U.S. government wants to positively affect happiness in the adult population, it is worth improving access to healthcare. Additionally, our findings support the Boyce paper's conclusion that an individual's **perception of their financial situation has more effect on happiness than their income level**.

REFERENCES

1. Boyce, C. J., Brown, G. D., & Moore, S. C. (2010). Money and Happiness. *Psychological Science*, 21(4), 471–475. <https://doi.org/10.1177/0956797610362671>
2. Napier, J. L., & Jost, J. T. (2008). Why Are Conservatives Happier Than Liberals? *Psychological Science*, 19(6), 565–572. <https://doi.org/10.1111/j.1467-9280.2008.02124.x>
3. University of Chicago. (2021). About the GSS | NORC. <https://Gss.Norc.Org/About-The-GSS>. Retrieved January 30, 2022, from <https://gss.norc.org/About-The-GSS>
4. Shalizi, C. (2015). Lecture 24–25: Weighted and Generalized Least Squares. Retrieved January 30, 2021 from: <http://www.stat.cmu.edu/~cshalizi/mreg/15/lectures/24/lecture-24-25.pdf>