

Predicting Employee Happiness

Christopher Sanchez

Department of Economics

University of California, Davis

March 8th, 2025

Abstract

As mental health, employee satisfaction, and happiness gain greater attention in the modern workplace and across broader society, understanding, identifying, and predicting worker happiness has become increasingly important. The core usage of the prediction of worker happiness is the “happy-productive worker thesis,” which links employee happiness with increasing productivity. Using a dataset explicitly compiled for this analysis, this paper applies a logistic model combined with polynomial terms, step functions, and regularization, to analyze quantitative predictors like real income, and job tenure, along with qualitative predictors like industry type and job satisfaction in 2022 to predict worker happiness. The final model selected was a LASSO model which achieved a 68% predictive accuracy of identifying a happy worker, with key predictors being job satisfaction, years in education, full-time employment status, years of experience in the job, and weeks worked last year, along with a pruned, interpretable classification tree. These results align closely with other works predicting happiness—and the findings will help fine-tune predictive models for happiness to be used in finding causal links to increased productivity—to further motivate advances in workplace culture, practices, and norms.

1. Introduction

As society progresses, previously overlooked aspects of individual life experience are being brought center stage as the world revisits these once ignored aspects of life and analyzes their potential impact on people. Mental health, cultural expression, and sexuality have emerged at the forefront in this progressive analysis—but, arguably, underlying many of these parts is a universal chase of happiness. Historically neglected in the workplace and the field of economics, employers and researchers alike have recently been turning to the effects of happiness on workers to reveal ways to potentially invigorate the workforce, and at the core of this study is the happy-productive worker thesis. The thesis is basic in nature—happy workers tend to be more productive, because unhappy people tend to deal with life complications that, in turn, affect their performance at work. Despite its simplicity, this thesis has been widely studied, with works by Thompson and Bruk-Lee (2021) linking decreased organizational skills, lower task and conceptual performance, and increasing turnover rates to unhappy workers. At the center of validating this thesis is the challenge of accurately predicting happiness and providing a data-backed answer to the question, “How can we predict happiness?” If happiness can be accurately predicted, researchers can move forward into directly testing its effects on the workforce—and not just observing.

This paper contributes to the extensive field on happiness prediction by refining the methods used to predict employee happiness, specifically emphasizing a simple, interpretable, yet highly accurate model with fewer core predictors via regularization and shrinkage methods. The objective is to expand the general knowledge of extremely influential predictors that may be useful in tuning predictive models of happiness, and aid further research in developing methods of predicting happiness. This paper further builds on works such as Haley M. Ryan’s (2024)

approach to using quantitative and qualitative predictors to predict life satisfaction, which would, in turn, predict whether or not one is happy. Similar work has been done internationally, with Park et al. (2025) using leisure factors to predict happiness of self-employed workers in South Korea, with variables such as work-life balance and high economic status being the most significant predictors of happiness in a worker. Wright and Cropanzano (2000) furthered the importance of a person's psychological state and found that job satisfaction and physiological well-being are fundamental predictors of job performance and happiness. A more scientific approach has been taken to predicting happiness in people as a whole, with Caraballo-Arias et al. (2024) utilizing international resources and surveys on happiness to link job satisfaction, psychological well-being, life satisfaction, and around 34 other factors influencing happiness at work. Various methods of research have been applied to come to a plethora of different predictors—and while happiness is multifaceted and complex in nature, this paper seeks to hone in on the common and uncommon predictors of happiness in the workplace via regularization methods to establish a strong predictive model of happiness on easily surveyable aspects of the workforce.

The dataset for this study comprised survey information from the General Social Survey (GSS), data composed mainly of qualitative and quantitative metrics. Many variables in the dataset are transformations of observed survey variables, employing indicator functions, dummy transformations, and higher-order polynomial terms to obtain a robust dataset with a smaller number of core predictors. Principal component analysis was conducted to glean insight into the predictors which explained more variation in the happiness variable and better suited for regression. Various modeling methods were then applied to the data. Iterations of the models used included basic generalized linear models like logistic regressions, with modifications like

polynomials, step functions, and shrinkage methods being used to arrive at the model with the lowest estimated out-of-sample misclassification rate, which was a logistic LASSO model. Classification trees, bagged trees, and random forests were also utilized to provide a more interpretable and streamlined version of the regression analysis.

The resulting model had a 68% predictive accuracy when tested on held-out test sample data. The predictors selected by the LASSO included years of education, marital status, job satisfaction, part-time or full-time status, job tenure, and weeks worked last year. Step functions and polynomials captured nonlinear relationships in job tenure, age, and weeks worked. These results align clearly with previous findings, yet the use of nonlinear terms and shrinkage methods allowed for a much smaller and more interpretable model that has strong predictive power, giving light to critical regressors that are fundamental to the practice of predicting happiness. The pruned classification tree traded off predictive power for interpretability. While the tree only had 66% predictive accuracy when tested on out-of-sample data, a slight decrease from the LASSO, there were only four terminal nodes, with just three main predictors: job satisfaction, weeks worked last year, and part or full time status, providing an accurate and highly interpretable model to predict employee happiness.

The insights provided by this analysis offer significant value for businesses and employers aiming to enhance employee well-being and productivity. Given the growing research interest in the effects of workplace happiness and its implications, the key predictors identified by this study can be used to cultivate practices and foster workplace environments to promote happiness. By promoting job satisfaction, check-ins and evaluating workload management, employers can influence their employees to be happier, benefitting employee welfare and businesses while enabling a more robust analysis of the happy-productive worker thesis.

2. Data

The data used in this study was gathered from the GSS for the purpose of this analysis, consisting of quantitative and qualitative variables. The dataset name is “FinalDataGSS.” Core predictor selection was guided by previous research such as Haley M. Ryan’s (2024) study on life satisfaction. The dataset includes 15 primary predictors and 844 complete observations. The primary response variable is happiness (*happy*, 0 if *unhappy*, 1 if *happy*). Fifth degree and below polynomials were created for variables like years in education, job tenure, weeks worked last year, and age. Step functions were constructed for the same variables by grouping observations into the following categories: minimum to first quartile, first quartile to median, median to third quartile, and third quartile to maximum. Principal Component Analysis (PCA) was conducted to evaluate the variables, and a scree-plot indicated that 11 principal components explained around 71% of variance, with the remaining principal components explaining the remaining 29% of variation. (see Figure 1).

Summary statistics on the main variables in the analysis highlighted a notable skewness in the primary response variable, *happy*, indicating a large amount of “happy” (*response* = 1) to “unhappy” (*response* = 0) observations. In the original dataset, the *happy* variable was heavily skewed toward happy observations, resulting in positively biased models and inflated misclassification rates. To remedy this, a subset of 300 observations with a balanced response variable for a non-biased analysis was randomly extracted from the primary dataset (see Figure 2), but the sample size reduction limits the potential robustness of the study. Regardless, relationships between variables still demonstrate distinct differences in happiness per variable (see Figure 3). 10-fold cross-validation is conducted on training sets to tune parameters and train models, while the test set is used to evaluate model performance on untouched data.

3. Model

The first model applied to the dataset was a basic logistic regression using selected core predictors with no modifications to provide a baseline relationship between the response variable, happy, and the predictors. Models are presented in the log-odds form for interpretability.

$$\begin{aligned} \log\left(\frac{p_i}{1-p_i}\right) = & \beta_0 + \beta_1 (\text{gender}_i) + \beta_2 (\text{weekswrk}_i) + \beta_3 (\text{wwwhrs}_i) + \beta_4 (\text{partorfull}_i) \\ & + \beta_5 (\text{yearsjob}_i) + \beta_6 (\text{jobsatisfaction}_i) + \beta_7 (\text{realinc}_i) + \beta_8 (\text{financialsatisfaction}_i) \\ & + \beta_9 (\text{age}_i) + \beta_{10} (\text{educ_years}_i) + \beta_{11} (\text{married}_i) + \beta_{12} (\text{race}_i). \end{aligned}$$

To explore nonlinear relationships between predictors and the response, the second model was a polynomial regression, adding second-degree polynomial terms to the previous logistic regression to capture diminishing or increasing returns from weeks worked last year, hours teleworked, job tenure, and age. Using cross-validation to determine the optimal number of degrees for polynomials yielded mixed results, as limited computing power to cross-validate thousands of combinations of polynomials proved to be a significant hurdle in extensively testing the prime number of polynomials. As a blanket fix, fifth-degree polynomials are included in the dataset. Still, higher-order terms were reserved for the LASSO model to prevent overfitting and unnecessary complexity to models incapable of performing feature selection.

$$\log\left(\frac{p_i}{1-p_i}\right) = \dots + \beta_{13} (\text{weekswrk}_i^2) + \beta_{14} (\text{wwwhrs}_i^2) + \beta_{15} (\text{yearsjob}_i^2) + \beta_{16} (\text{age}_i^2).$$

The third model was a step function regression, focusing on adding more flexibility behind measuring nonlinear relationships by substituting the polynomial terms in the previous logistic regression with step function groupings with observations split along the minimum to first quartile, first quartile to median, median to third quartile, and third quartile to maximum, to mimic different “stages” in the data, like age groupings or job tenure groupings. Due to the

grouping, different relationships in different observation levels can be modeled, offering more insight into the relationship between the response and the predictors.

$$\log\left(\frac{p_i}{1-p_i}\right) = \dots + \beta_{13}(\text{week_group}_i) + \beta_{14}(\text{wwwhr_group}_i) + \beta_{15}(\text{yearsjob_group}_i) + \beta_{16}(\text{age_group}_i).$$

The fourth model was a LASSO logistic regression. The model was allowed to use all predictors in the dataset, along with all groupings and polynomial degrees of previous variables up to the fifth degree. The penalty term, lambda, was chosen via k-fold cross-validation with 10 folds.

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_{33} X_{33} + \lambda \sum_{j=1}^{33} |\beta_j|$$

As a more interpretable alternative, classification trees were utilized to obtain more straightforward predictions based on key variables. Trees were pruned via k-fold cross-validation with 10 folds to select the optimal amount of nodes with the lowest misclassification error and highest interpretability. Random forests and bagged trees are also used in comparison against classification trees, to improve upon the predictive performance of classification trees, which tend to overfit on training data and result in lower out-of-sample predictive performance. The random forest will be cross-validated using 10 folds to determine the optimal amount of predictors to randomize within the trees, ranging from using one predictor to using all predictors in the dataset.

4. Results

The first baseline logistic regression had a misclassification rate of approximately 49%, indicating relatively low predictive performance in the basic models. The most impactful variables were financial satisfaction, multiplying the odds of being happy by 4.768348, job satisfaction, multiplying the odds of being happy by 3.061791, and race. People who were not black or white multiplied the odds of being happy by 3.077138.

$$\begin{aligned} \log\left(\frac{p_i}{1-p_i}\right) = & \dots + 6.453e-01 (\text{week_group}(48, 52]_i) + 1.960e-01 (\text{wwwhr_group}(4, 10]_i) + 3.114e-01 (\text{wwwhr_group}(10, 24]_i) \\ & + 5.056e-01 (\text{wwwhr_group}(24, 100]_i) - 5.810e-01 (\text{yearsjob_group}(3, 5]_i) - 5.197e-01 (\text{yearsjob_group}(5, 13]_i) \\ & + 1.824e-01 (\text{yearsjob_group}(13, 42]_i) - 4.538e-01 (\text{age_group}(33, 43]_i) - 3.097e-01 (\text{age_group}(43, 56]_i) - 4.564e-01 (\text{age_group}(56, 83]_i) \end{aligned}$$

The second regression with polynomials yielded similar results—financial satisfaction, job satisfaction, and people who were not black or white had the most significant impacts on the odds of being happy, with similar magnitudes as before. However, this regression had an improved misclassification rate of 44%, showing that nonlinearity helped improve out-of-sample predictive power (see Figure 4).

$$\log\left(\frac{p_i}{1-p_i}\right) = \dots + 1.331e-03 (\text{weekswrk}_i^2) - 3.557e-04 (\text{wwwhrs}_i^2) - 3.240e-03 (\text{yearsjob}_i^2) - 9.482e-04 (\text{age}_i^2).$$

The third regression with step functions further improved the misclassification rate, and resulted in errors only 42% of the time. The most significant impacts remained the same, though magnitudes shifted, with financial satisfaction multiplying the odds of being happy by 5.337. Not identifying as black or white multiplied the odds of being happy by 3.068599, now more significant than the effect of job satisfaction, which only multiplied odds of being happy by 2.776981. Different groups in variables resulted in drastically different effects (see Figure 5).

$$\log\left(\frac{p_i}{1-p_i}\right) = \dots + 6.453e-01 (\text{week_group}(48, 52]_i) + 1.960e-01 (\text{wwwhr_group}(4, 10]_i) + 3.114e-01 (\text{wwwhr_group}(10, 24]_i)$$

$$\begin{aligned}
& + 5.056e-01 (\text{wwwhr_group}(24, 100|_i) + 5.810e-01 (\text{yearsjob_group}(3, 5|_i) + 5.197e-01 (\text{yearsjob_group}(5, 13|_i) \\
& + 1.824e-01 (\text{yearsjob_group}(13, 42|_i) + 4.538e-01 (\text{age_group}(33, 43|_i) + 3.097e-01 (\text{age_group}(43, 56|_i) + 4.564e-01 (\text{age_group}(56, 83|_i)
\end{aligned}$$

The LASSO logistic regression significantly improved the out-of-sample misclassification rate. While other models averaged misclassification rates around 45%, the LASSO model yielded a misclassification rate of only 32%. The lambda chosen via cross-validation was 0.03109968. Again, the most impactful variables were financial satisfaction, job satisfaction, and not being black or white. Magnitudes shifted again, however, with financial satisfaction and job satisfaction multiplying the log odds of being happy by 3.140871 and 2.07929, respectively.

$$\begin{aligned}
\log\left(\frac{p_i}{1-p_i}\right) = & -9.587635e-01 + -8.657809e-02 (\text{Consumer Services}_i) + 2.826101e-01 (\text{married}_i) \\
& + -2.067398e-02 (\text{educ_years}_i) + 1.144500e+00 (\text{financial_satisfaction}_i) \\
& + 5.567609e-01 (\text{raceOther}_i) + 4.227567e-03 (\text{weekswrk}_i) \\
& + -3.232750e-01 (\text{partorfull}_i) + 7.320263e-01 (\text{jobsatisfaction}_i) \\
& + 1.344047e-26 (\text{realinc}_i^5) + -1.562548e-08 (\text{yearsjob}_i^5) \\
& + -1.417169e-10 (\text{age}_i^5) + -1.175888e-02 (\text{yearsjob_group} [5-13]_i) \\
& + 0.03109968 \sum_{j=1}^{33} |\beta_j|
\end{aligned}$$

The pruned classification tree had a misclassification rate of 34%, significantly better than logistic regression models but worse than the LASSO regression model. Out-of-sample performance is traded for an incredibly easy-to-interpret classification tree that can arguably be used as a more efficient prediction method than the LASSO regression model, as the pruned classification tree had an optimal number of four terminal nodes (see Figure 6). Random forests yielded a misclassification rate of approximately 33%, marginally worse than the LASSO regression but marginally better than the classification trees. The optimal amount of randomized predictors was 12, with other levels averaging misclassification rates of around 35-38%. Bagged

trees were an outlier, achieving an estimated out of sample performance rate of 40%. Although random forests achieve a marginally better estimated out of sample performance (around 1%) than classification trees, the visualization and interpretability of classification trees provide more value to the results and analysis, especially when the LASSO logistic regression outperforms both methods.

Utilizing the LASSO logistic regression model for its low misclassification rate, a simulated out-of-sample observation can be predicted based on the variables chosen within the model. If this employee worked in consumer services (*belongs to the “Consumer Services” category*), is aged 50 (*age = 50*), with 20 years of education (*educ_years = 20*) and 10 years of experience at their job (*yearsjob = 10*), is married (*married = 1*), was financially satisfied (*financialsatisfaction = 1*), was not black or white (*belongs to the “Other” category*), worked 40 weeks last year (*weekswrk = 40*) as a part time worker (*partorfull = 0*), was satisfied with their job (*jobsatisfaction = 1*), and had an income of \$70,000 (*realinc = 70000*), the LASSO model would have a predicted probability of $\text{happy} = 1$ given the predictor values of 0.7906597, meaning we predict this unseen employee to be happy.

If the same simulated out-of-sample observation was to be predicted using the classification tree, again, someone that was financially satisfied (*financialsatisfaction = 1*), worked 40 weeks last year (*weekswrk = 40*), and was a part time worker (*partorfull = 0*), that person is predicted to be happy (see Figure 6), the same result as the LASSO model.

Regardless of the method, both results are plausible. A person who has attained financial and job satisfaction, being able to work much fewer weeks per year than a typical full-time worker, along with being a part-time worker during those weeks, and earning a stable, high income of \$70,000 per year, is doing exceptionally well. Save for any personal struggles, that

person is definitely expected to be happy. It is interesting to consider why people who are not black or white have had such a considerable impact on the models of happiness. It is equally provoking to consider why real income—something every person in the world seeks to maximize, is not as impactful on the odds of someone being happy. Regardless, the results are undeniable.

5. Conclusion

Thus, the final models are the LASSO model, with an estimated out-of-sample performance of 68%, and a classification tree, with an estimated out-of-sample performance of 66% but extremely straightforward interpretability. Both models highlighted variables essential in predicting happiness: job satisfaction, financial satisfaction, part or full time status, and weeks worked last year. That is not to say research on this subject of predicting happiness is over—many key questions arise. For example, why are people who are not black or white predicted to be considerably happier? Is there a way to predict job or financial satisfaction? Does higher income truly not matter after a certain amount to people? These are questions that must be further evaluated before work on happiness and the factors that go into it can be considered “finished.” While qualitative variables were used, personable traits like mental health status, number of children, number of friends, or number of hobbies, things that can vary wildly from person to person, were not included, due to poor sampling and availability. Future studies should strive to include more volatile and personal traits, along with including a larger sample size, to unravel happiness on a deeper level fully—but from a business and economic standpoint, job satisfaction, financial satisfaction, race, industry choice, and part or full time status are essential components that should be considered in many models of happiness in the future. Employers looking to maximize happiness in their workers should guide workplace policies to check-in with employees about their workloads, their satisfaction with their positions within the company, and their current financial situations. By checking in and attempting to remedy these situations with employees, employees evidently would have a higher chance of being happy, and if the happy-productive worker thesis holds, would be more productive, engaged, and likely to contribute more to the company.

Appendix

Figure 1: A Scree Plot constructed for principal component analysis shows that 11 principal components explain around 70% of the variance within predictors within the dataset.

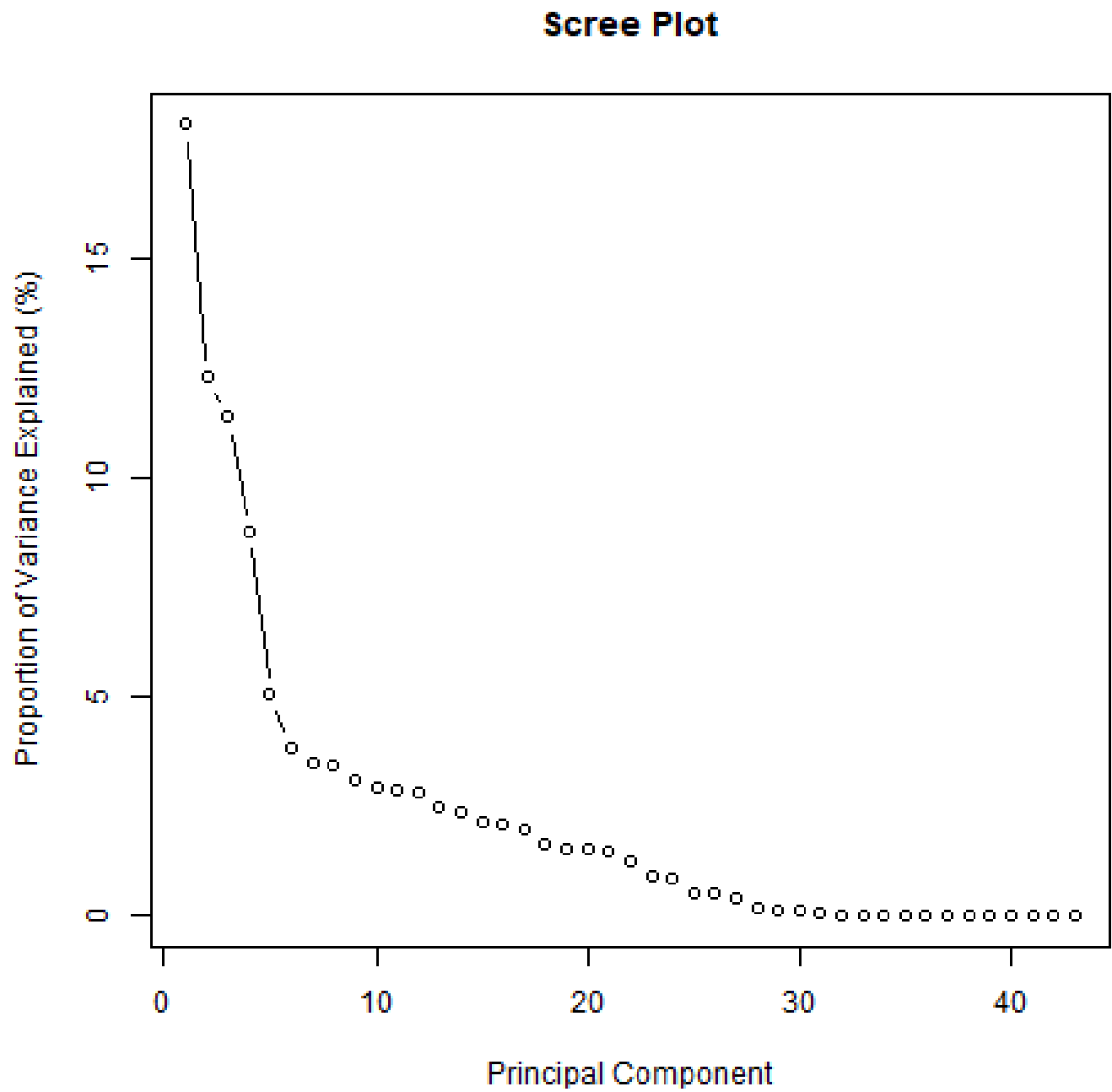


Figure 2: Summary statistics of the main predictors in the auxiliary dataset with balanced dummy variables.

Variable	n	Min.	Mean	Max.	Std. Dev.	Levels/Desc
industry	300	/	/	/	/	1. Manual Labor 2. Retail Labor 3. Consumer Serv. 4. Healthcare & Education 5. Finance, Admin
married	300	/	/	/	/	0. Not married 1. Married
educ_years	300	7.00	14.82	20.00	2.54	Years in education
financialsatisfaction	300	/	/	/	/	0. Not satisfied 1. Satisfied
race	300	/	/	/	/	1. White 2. Black 3. Other
age	300	18.00	44.48	83.00	13.98	/
gender	300	/	/	/	/	0. Female 1. Male
weekswrk	300	5.00	47.30	52.00	9.21	Weeks worked last year
partorfull	300	/	/	/	/	0. Part-Time 1. Full-Time
yearsjob	300	1.00	8.76	42.00	8.25	Job tenure
jobsatisfaction	300	/	/	/	/	0. Not satisfied 1. Satisfied
realinc	300	204.50	42319.15	135195.96	37992.90	Adjusted income
wwwhrs	300	0	17.71	100	18.88	Hours teleworked
happy	300	/	/	/	/	0. Unhappy 1. Happy

Figure 3: Barplots displaying relationships between categorical variables and their differences in happiness levels. There are some key differences—happy people tend to be more satisfied with their job and financial status, as denoted by the low amount of unhappy people in those categories. Gender does not seem to have a stark difference in happiness vs unhappiness. More part-time workers are happy, and there are significantly fewer people who are unhappy while married than those who are happy and married (though the amount of unhappy and happy non-married individuals does not have a stark difference).

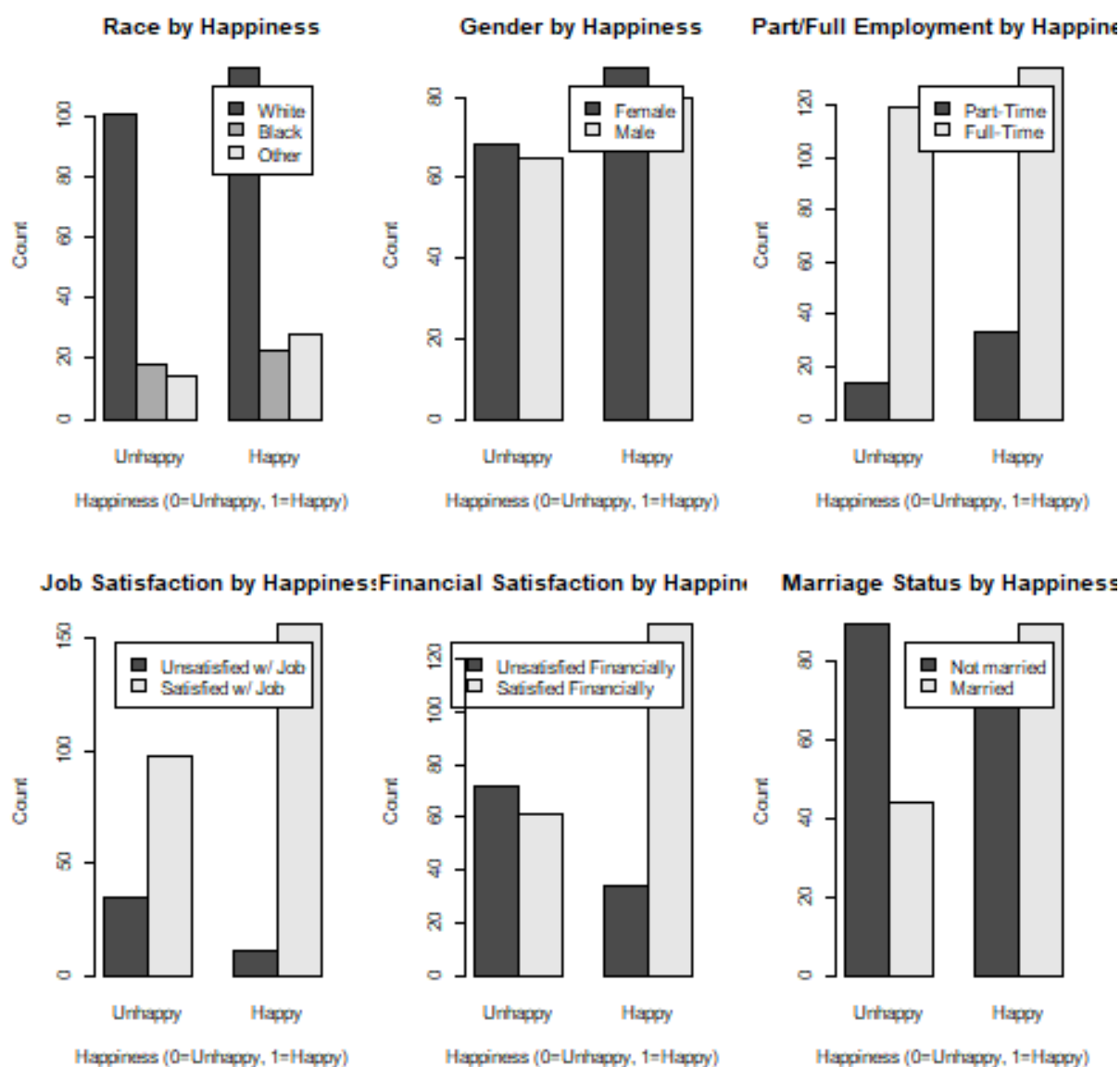


Figure 4: The full polynomial regression equation.

$$\begin{aligned}
\log\left(\frac{p_i}{1-p_i}\right) = & -1.616e+00 - 2.743e-01 (\text{gender}_i) - 5.606e-02 (\text{weekswrk}_i) \\
& + 3.180e-02 (\text{wwwhrs}_i) - 1.524e+00 (\text{partorfull}_i) + 7.742e-02 (\text{yearsjob}_i) \\
& + 1.256e+00 (\text{jobsatisfaction}_i) + 8.986e-06 (\text{realinc}_i) + 1.534e+00 (\text{financialsatisfaction}_i) \\
& + 6.749e-02 (\text{age}_i) - 1.099e-01 (\text{educ_years}_i) + 5.979e-01 (\text{married}_i) \\
& + 5.283e-01 (\text{raceBlack}_i) + 1.302e+00 (\text{raceOther}_i) + 1.331e-03 (\text{weekswrk}_i^2) \\
& - 3.557e-04 (\text{wwwhrs}_i^2) - 3.240e-03 (\text{yearsjob}_i^2) - 9.482e-04 (\text{age}_i^2).
\end{aligned}$$

Figure 5: The full step-function regression equation.

$$\begin{aligned}
\log\left(\frac{p_i}{1-p_i}\right) = & -2.516e-01 - 4.880e-02 (\text{gender}_i) - 8.481e-01 (\text{partorfull}_i) \\
& + 1.021e+00 (\text{jobsatisfaction}_i) + 9.794e-06 (\text{realinc}_i) \\
& + 1.675e+00 (\text{financialsatisfaction}_i) - 1.230e-01 (\text{educ_years}_i) + 6.861e-01 (\text{married}_i) \\
& + 4.438e-01 (\text{raceBlack}_i) + 1.121e+00 (\text{raceOther}_i) \\
& + 6.453e-01 (\text{week_group}(48, 52]_i) + 1.960e-01 (\text{wwwhr_group}(4, 10]_i) + 3.114e-01 (\text{wwwhr_group}(10, 24]_i) \\
& + 5.056e-01 (\text{wwwhr_group}(24, 100]_i) - 5.810e-01 (\text{yearsjob_group}(3, 5]_i) - 5.197e-01 (\text{yearsjob_group}(5, 13]_i) \\
& + 1.824e-01 (\text{yearsjob_group}(13, 42]_i) - 4.538e-01 (\text{age_group}(33, 43]_i) - 3.097e-01 (\text{age_group}(43, 56]_i) - 4.564e-01 (\text{age_group}(56, 83]_i).
\end{aligned}$$

Figure 6: The final classification tree, pruned to the optimal number of 4 terminal nodes via k-fold cross validation with 10 folds. The main predictors are job satisfaction ($=0$ if not satisfied, $=1$ if satisfied), weeks worked last year, and part or full time status for an employee ($=0$ if part time, $=1$ if full time). The response variable is whether an employee is happy or not ($=0$ if unhappy, $=1$ if happy).

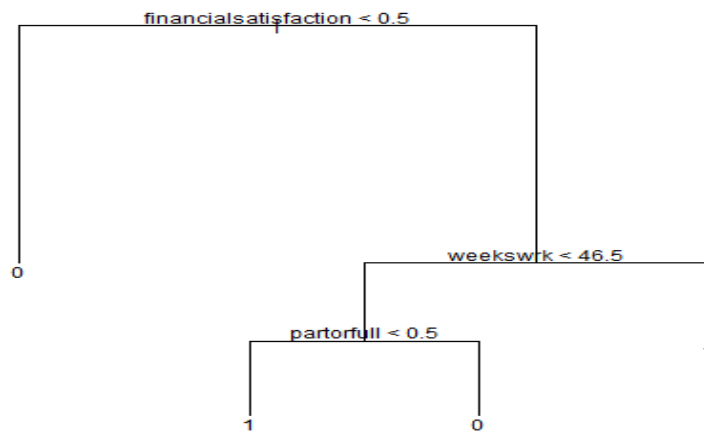


Figure 7: A barplot showcasing the relationship between the most commonly used variable, financial satisfaction in relation to the primary response variable, happiness. Here, when people are unhappy, the left bars of the plot, those people tend also to be financially unsatisfied. Meanwhile, happy people, the right bars of the plot, tend to be financially satisfied as well. Very few unsatisfied people financially are also happy—showcasing a clear correlation between financial satisfaction and happiness.



Figure 8: Boxplots showcase relationships between numerical variables and happiness. There aren't as drastic differences between happiness levels and numerical variables as there are in the categorical variables, especially in job tenure (*yearsjob*), years in education (*educ_years*), or age. There are slight differences in virtual work hours (*wwwhrs*), real income, and weeks worked last year (*weekswrk*), with the interquartile ranges differing for happy and unhappy people.

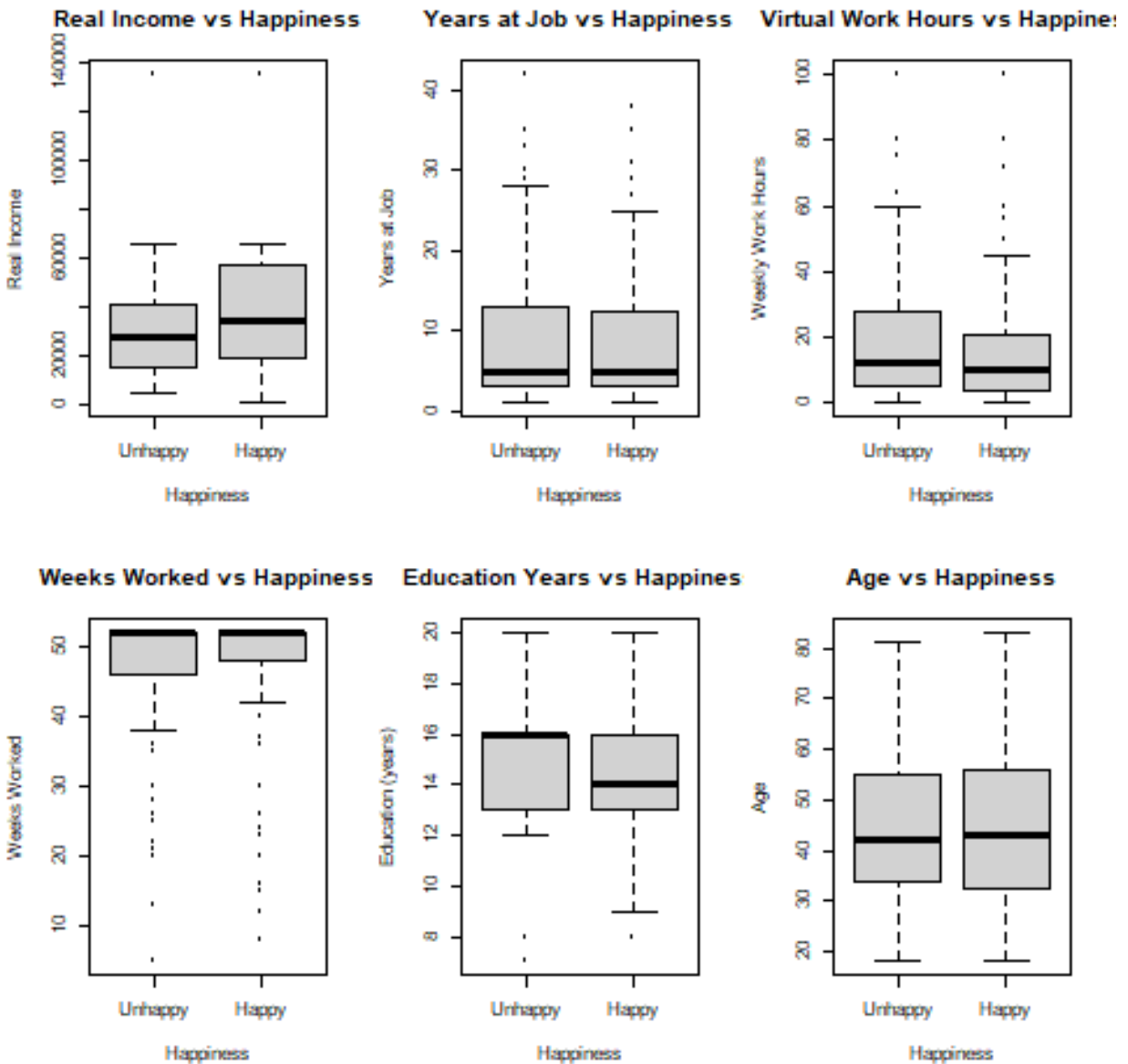
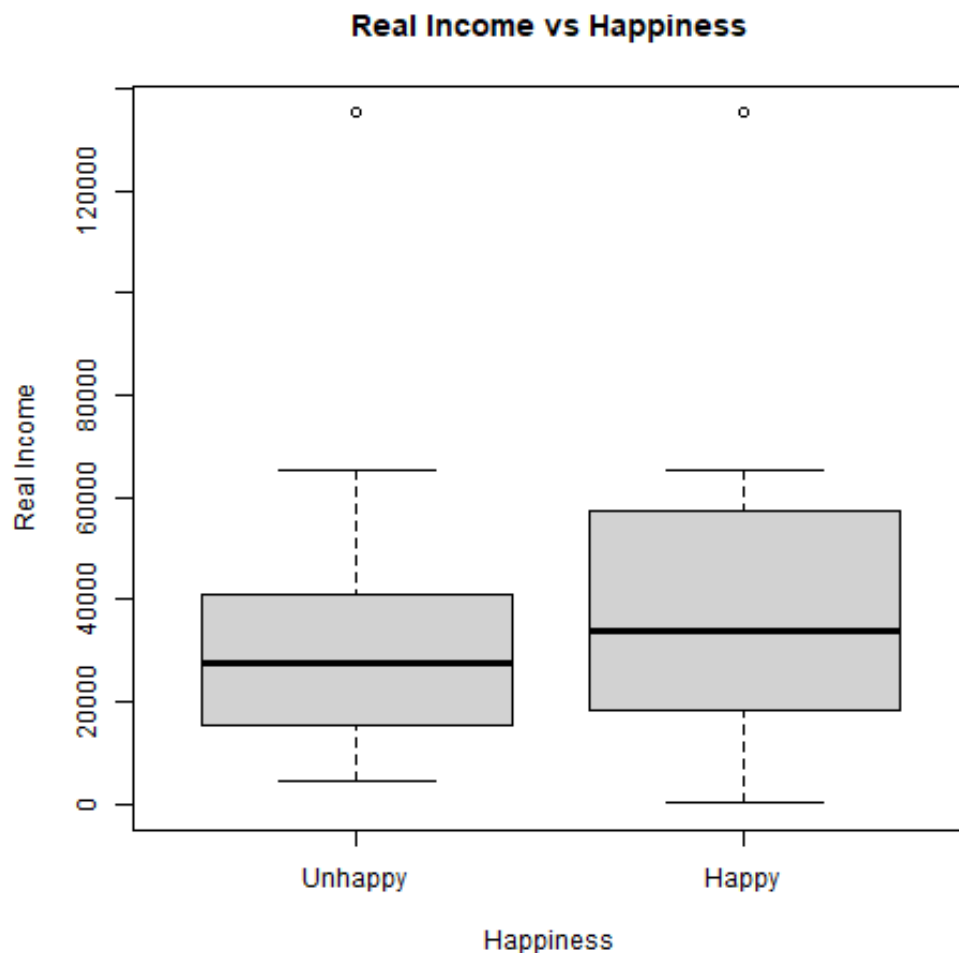


Figure 9: A boxplot between real income and happiness levels. People that are happy tend to have a higher interquartile range, and a higher median income than people who are unhappy.



Acknowledgements: To whoever reads this, professor or TA, thank you for a great quarter, and thank you for all your help throughout the creation of this project. My interest in this subject has never been greater, and that would not have been possible without this class. I wish you luck in all of your future endeavors, and hope I get to take another class with either of you in the future.

All my best,

Christopher Sanchez

References

- Haley, M. (2024). Identifying Worker and Workplace Predictors of Life Satisfaction in Age-Sex Sub-Populations: An “Emelometrics” Approach.
- Fogaça, N., & Junior, F. A. C. (2016). Is “happy worker” more productive. *Management Studies*, 4, 149–160.
- Thompson, A., & Bruk-Lee, V. (2021). Employee happiness: Why we should care. *Applied Research in Quality of Life*, 16, 1419–1437.
- Bayona, J. A. (2025). Happy-Productive worker thesis: The role of work characteristics, gender, and age. *PloS One*, 20.
- Kwang-Hi, H., & Kim, S. (2021). Leisure factors predicting the happiness of self-employed workers in South Korea. *International Journal of Environmental Research and Public Health*, 18.
- Wright, T. A., & Cropanzano, R. (2000). Psychological well-being and job satisfaction as predictors of job performance. *Journal of Occupational Health Psychology*, 5.
- Caraballo-Arias, Y., Feola, D., & Milani, S. (2024). The science behind happiness at work. *Current Opinion in Epidemiology and Public Health*, 3, 11–24.