U.S. national happiness post-covid*

Julie Nguyen, Missy Zhang, Linrong Li

15 March 2023

Abstract

add abstract

1 Introduction

Human beings strive for different things in life, but the ultimate goal is usually to achieve happiness. Happiness can be seen as a measure of life quality, and it is often associated with factors such as income, education, social interactions, and family. The relationship between happiness and age has been widely studied, and research suggests that there is not much change in happiness until around the age of 55, after which happiness starts to increase, peaking around 67, with a quite sharp decline around the age of 75 (Frijters and Beatton 2012). Studies examining the relationship between gender and happiness have found that individuals in the United States were more than twice as likely to experience well-being in states with high gender equality scores compared to states with low gender equality scores (Holter 2014). The relationship between race and happiness has also been studied, with research suggesting that the gap in happiness between whites and blacks is substantial, with whites more likely to report being very happy and less likely to report being not too happy (Icel and Ludwig-Dehm 2019). Several studies conducted in Europe suggest that working fewer hours is correlated with higher life satisfaction (Shao 2022). Regarding income, studies have shown that the relationship between perceived sufficient income and happiness is mediated by perceived health. Individuals with sufficient income are more likely to have better perceived health, which leads to increased happiness (Robert Weech-Maldonado and Miller 2017).

This paper delves into the complex relationship between age, gender, race, working hours, and income, and their impact on an individual's overall happiness. The findings reveal that while the level of happiness among various age groups remains stable, the 18-34 age group reported the highest level of extreme happiness. Additionally, the study suggests that male respondents tend to report higher levels of happiness compared to their female counterparts, whereas the level of happiness remains similar across different racial groups. [insert something here]

The remainder of the paper is split into five sections. Section 2 explains the data source and collection methodology, key features, our selected data's characteristics, strengths and weakness, as well as potential ethics issues. Section 3 discusses the methods used to produce the wanted results to investigate the relationship between happiness and the various factors. Section 4 presents the findings from our methodology.

2 Data

2.1 Data Source and Collection

To get better insights on what factors affect happiness, we utilized the 2021 US General Social Survey ("US General Social Survey" 2021) from NORC. The raw dataset is a Stata file and was imported in R using the package Haven (Wickham, Miller, and Smith 2022). It includes data from 568 survey questions for 4,032 survey participants. Using the R (R Core Team 2022) package tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2022), we were able to clean and perform exploratory data analysis on the dataset to get insights into the data. Further, we used R package ggplot2 (Wickham 2016), Knitr (Xie 2014) and KableExtra (Zhu 2021) to visualize the data for this paper. The dataset

*Code and data are available at:

used to explore the relationship between working conditions and happiness was directly extracted from the GSS Data Explorer website (https://gssdataexplorer.norc.org/home). All the interested variables were first selected from the website and then added to an extract, which was later exported in the Excel Workbook format. This dataset is saved in the job_happy folder inside the input folder. The R package readxl (Wickham and Bryan 2022) was used to read the Excel Workbook file in R.

The survey of interest in this report is the General Social Survey ("US General Social Survey" 2021), one that monitors pubic opinion and behavior in the United States. It has been conducted since 1972 by the NORC at the University of Chicago and funded by the National Science Foundation(NSF); aiming to minimize all changes via retaining similar sampling and questioning approaches. In each round, the GSS contains a set of repeating modules and a section of topical modules that is subject to change in every round and each participant will be given a subset of repeating modules and topical modules (it may not contain all modules).

2.2 Methodology

Since its inception, the GSS has traditionally used in-person data collection as its primary mode of data collection. However, to safeguard the health of staff and respondents during the COVID-19 pandemic, the 2021 GSS data collection used a mail-to-web methodology (supplemented with phone for respondents who needed the option) instead of its traditional in-person interviews.

Outreach was conducted via mail and phone using commercially available phone number matches for addresses in the sample or inbound phone contact. GSS staff redesigned the mail-based outreach to respondents to introduce the GSS to fresh address-based sample and encourage them to participate, either by web or over the phone. Throughout the data collection period, selected households were sent postcards, invitation packets, and reminder letters using a combination of USPS and FedEx and urging them to complete the 2021 GSS survey. The sample was released in three batches, with evaluation of respondent recruitment protocol for each batch allowing the fine-tuning of protocol for subsequent batches.

With a response rate of 17.4%, a total of 4,032 surveys were completed; from December 1, 2020, to May 3, 2021. 88.3% of those respondents completed the survey via the web, and 11.7% completed it via the phone.

2.3 Key Features

The GSS collects data on contemporary American society to monitor and explain trends in opinions, attitudes, and behaviors. The GSS contains a standard core of demographic, behavioral, and attitudinal questions, plus topics of special interest. Among the topics covered are civil liberties, crime and violence, intergroup tolerance, morality, national spending priorities, psychological well-being, social mobility, and stress and traumatic events. The survey data has 4,032 respondents and counts 735 variables.

2.4 Variable Selection

06 variables were selected to conduct analysis for the report including respondents' age, sex, race, working hours, and income. We first viewed the relationship of happiness towards different age, sex, racial groups to investigate the national trend of happiness in the year 2021. Other variables are used to explore the relationship and their impact on happiness. For our study, alongside the comprehension of happiness nationwide in the year 2021, we wanted to focus on how work-related factors impact the population's happiness.

Happiness was self-measured by the respondents using the survey asking how happy they felt. It is scaled as "not too happy", "pretty happy", "very happy". Any type of responses that fell into "Inapplicable", "No Answer", "Do not Know/Cannot Choose", "Skipped on Web" were coded as "Non Applicable". Age of respondents was recorded each from 18 to 88 years old, and 89 and above. The participant's sex was measured as "male" and "female" with a small percentage (2.2%) of "Not Applicable" and "No answer". For variable regarding race, the survey separated the options for response as "Black", "White", and "Other". Only about 1.33% of the respondents left their answers as "Non Applicable".

Working hours refer to the hours the respondent had worked the week before they responded to the survey. Income of the respondent is also taken into account as we believe that it is an important factor in examining the happiness trend in overall and also act as a potential motive for happiness. Working hours. Income was divided into 12 levels, respectively lower then $1000, $1000 to 2999, $3000 to 3999, $4000 to 4999, $5000 to 5999, $6000 to 6999, $7000 to

$7999, $8000 to 9999, $10000 to 14999, $15000 to 19999, $20000 to 24999 and more than $25000. The respondents were asked to respond to which group their previous year's earnings fell into.

2.5 Strengths and Weaknesses

- Too many people having high incomes ($25k above) -> should narrow down into smaller clusters
- Subjective options for happiness level

2.6 Ethical Concerns & Limitations

The 2021 US General Social Survey has a few issues that should be considered. Since this is a voluntary survey, the appearance of non-sampling errors is inevitable. The type of non-sampling errors in the survey mostly falls into non-response error, where respondents left the answer blank or provide incomplete answers.

There are also ethical factors that should be taken into consideration. Regarding the race of respondents, there are only options for White, Black, and Others. Using "Others" option likely groups other racial identities, therefore could lead to interpretations not accurately capturing the current demographic.

The low response rate (17.4%) also heavily affects the final results, that it cannot generalize and reflect the whole population.

3 Discussion

In order to understand the context and trends, we observed the overall trend of happiness of the U.S. population in 2021. We also looked into the distribution of happiness among different age groups, sex groups, and racial groups.

3.1 National Happiness

Overall, the level of happiness in the United States in 2021 is relatively high. Figure 1 shows that among 4041 respondents, the majority of them (3118 or 77.2%) reported that they feel happy, while 923 people were not happy. While the number of people who reported feeling very happy is relatively low compared to those who reported feeling pretty happy, it still suggests that there is a sizeable portion of the population that is experiencing high levels of happiness. Conversely, the fact that a significant number of respondents reported not feeling happy highlights the importance of addressing mental health and well-being in the United States.

Figure 1: National Happiness

3.2 Happiness among age groups

Comparing happiness among different age groups, it is apparent that the level of happiness remained stable throughout the groups (Figure 2). While in 18 to 64 and non-identifiable age groups, there are more people feeling unhappy than very happy, from 65 and above age group witnesses the same proportion of people feeling not too happy and very happy. In the 18-34 age group, there are 189 respondents (25.9%) reported to feel very happy, while that of 35-49, 50-64, and 65 and above age group take up 22.9%, 23.5%, and 21% respectively.

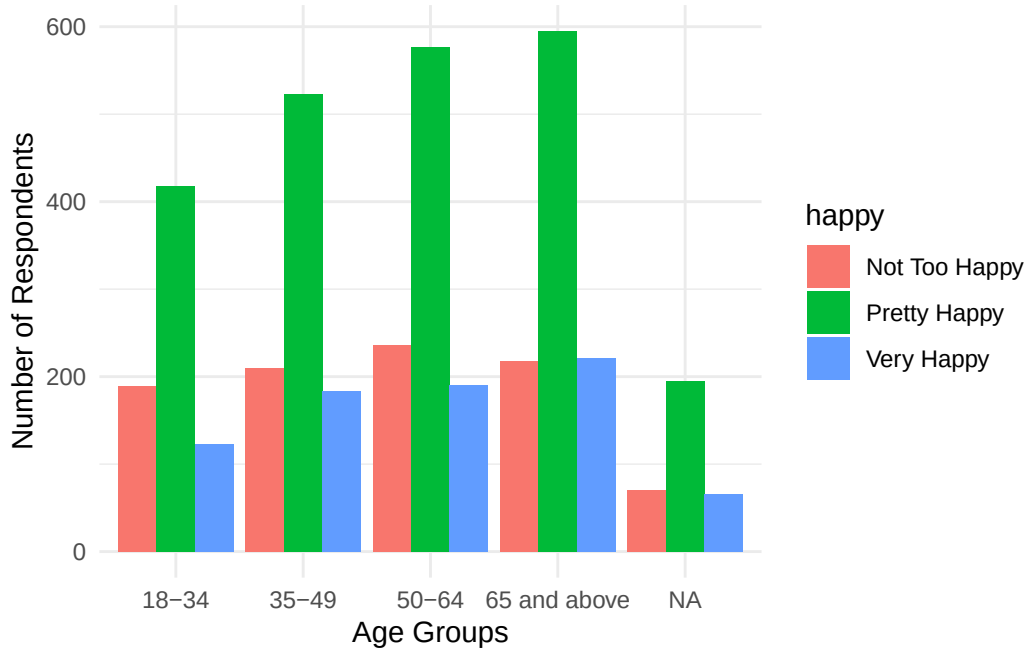

Figure 2: Happiness by age

3.3 Happiness among male, female, and others

Figure 3 shows that male tend to be more happy than female [more elaboration is needed here about why male are more happy than female]. In the year 2021, the U.S. population reported to relatively happy among different sex groups. There is a great disparity between female respondents feeling unhappy and very happy - more female experienced unhappiness than very happy, while that of male respondents is quite similar. In addition, there are 92 people who responded their level of happiness without identifying their sex. In this group, there are 62 respondents feel happy, 18 respondents feel very happy, and 12 respondents feel unhappy.

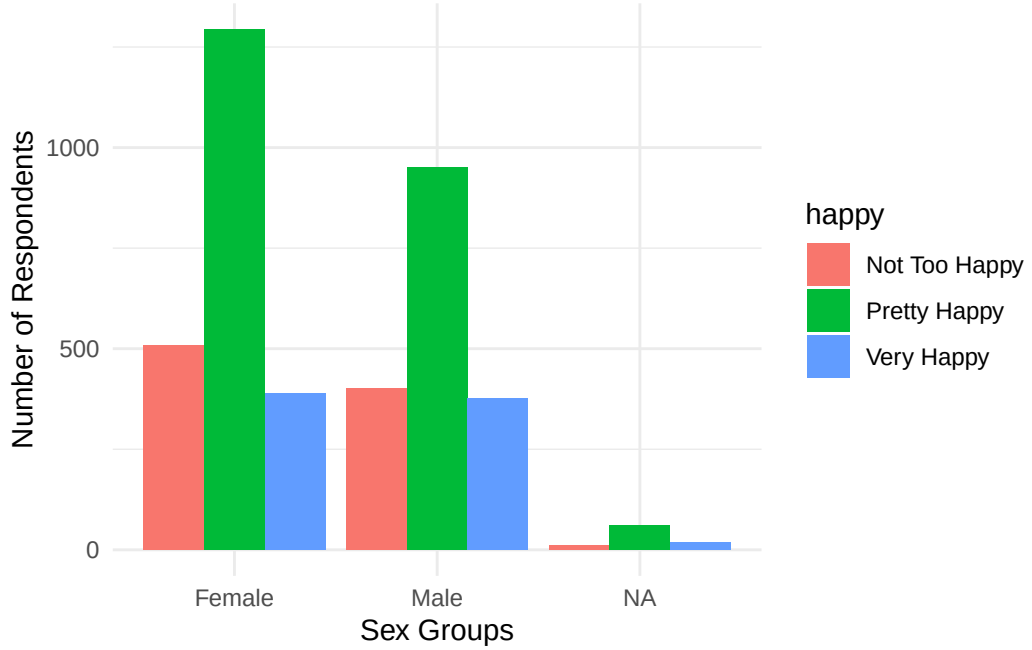

Figure 3: Happiness by sex in 2021

3.4 Happiness among different races

As a multicultural country, happiness in different races is also worth observing as it shows the level of happiness in different racial communities. From Figure 4 We notice that there is a similar pattern between the black and white racial groups, and between other races and non-identifiable group. [which is?]

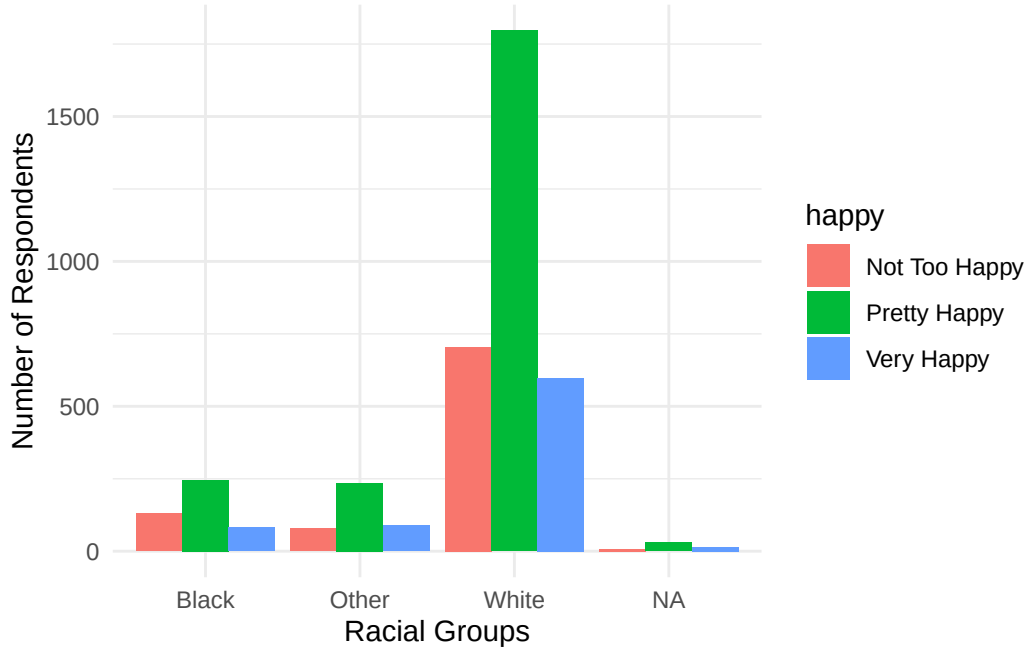

Figure 4: Happiness by race

3.5 Happiness and working hours

Table 1 shows the mean and standard deviation of working hours in different happiness level, as well as the total number of observation of each happiness level. Though the mean of working hours for the very happy group was slightly lower than the rest, there are no significant difference between different happy levels. In addition, it can be told from the standard deviation that the variability of working hours for the three groups are similar. However, it can be found that from people responded to all three questions (happiness, working hours and income level), the population of pretty happy group takes the majority of total number of observations.

Table 1: The means of hours worked last week in different levels of happiness

happy	Mean	Sd	Total_observed
Not too happy	40.68378	13.04215	370
Pretty happy	40.16744	11.95678	1075
Very happy	39.87363	13.63153	364

3.6 Happiness and income level

Figure 5 presents the levels of happiness across different income groups. **?@tbl-income** provides a breakdown of the total number of people in each income group who reported being in different levels of happiness. Figure 5 visualizes these results and shows a clear trend.

The data shows that, for most income groups, the highest number of respondents reported feeling "pretty happy" compared to the other two happiness levels. However, there were some exceptions. For individuals earning less than $1000, their reported happiness levels tended to be more polarized towards "not so happy" and "very happy". Moreover, for those earning between $6000 and $6999, the majority reported being "very happy".

Overall, the data suggests that, across income groups, respondents tended towards the "not so happy" level of happiness. However, it's worth noting that there were variations in happiness levels across different income groups.

Figure 5: Relationship of income level and happiness

4 Conclusion

- Only records that responded to all three variables of interest were kept when discovering the relationship between working conditions and happiness.
- Omit NA for happiness

5 Appendix

References

Frijters, Paul, and Tony Beatton. 2012. "The Mystery of the u-Shaped Relationship Between Happiness and Age." *Journal of Economic Behavior & Organization* 82 (2): 525–42. https://doi.org/https://doi.org/10.1016/j.jebo.2012.03.008.

Holter, Øystein Gullvåg. 2014. "'What's in It for Men?': Old Question, New Data." *Men and Masculinities* 17: 515--548. https://doi.org/https://doi-org.myaccess.library.utoronto.ca/10.1177/1097184X14558237.

Icel, John, and Sarah Ludwig-Dehm. 2019. "Black-White Differences in Happiness, 1972–2014." *Social Science Research* 77: 16--29. https://doi.org/https://doi-org.myaccess.library.utoronto.ca/10.1016/j.ssresearch.2018.10.004.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Robert Weech-Maldonado, Justin C. Lord, and Michael J. Miller. 2017. "The Relationships Among Socio-Demographics, Perceived Health, and Happiness." *Applied Research Quality Life* 12: 289–302. https://doi.org/https://doi-org.myaccess.library.utoronto.ca/10.1007/s11482-017-9517-8.

Shao, Qinglong. 2022. "Does Less Working Time Improve Life Satisfaction? Evidence from European Social Survey." *Health Economics Review Volume* 12. https://doi.org/https://doi.org/10.1186/s13561-022-00396-6.

"US General Social Survey." 2021. *NORC*. https://gss.norc.org/Get-The-Data.

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.

Wickham, Hadley, and Jennifer Bryan. 2022. *Readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Evan Miller, and Danny Smith. 2022. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. https://CRAN.R-project.org/package=haven.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.