

World Values Survey Project

Kathryn Haglich

10/17/2019

Report

Introduction

The World Values Survey (WVS) is an international to study how people's values change over time and the resulting impact on political and social aspects of a country. Originally starting in 1981, the common questionnaire has been distributed to nearly 100 countries to date from the poorest countries to the wealthiest ones. This provides a wealth of data that can be analyzed to understand how people's beliefs, motivations, and values evolve between countries and over time and providing key insights to a multitude of different questions. For this project, I performed simple Exploratory Data Analysis on WVS data sets from the United States, Russia, and Japan. After cleaning, I compared responses for individual questions and proportion of responses for pairs of variables and explored differences in the survey questions that were asked among the three countries.

Background

The WVS is a global survey project designed to understand people's societal values spanning over 100 countries over 7 Waves of surveys - the data for the most recent one set to be release in 2020. The standardized questions are designed to explore the participants' views on democracy, tolerance, gender and race equality, religion, the environment, politics, family, the community, etc. Once collected the data can be analyzed by experts from a wide variety of field to understand the socioeconomic, political, and cultural trends and impacts across the globe.

To develop the questionnaire, social scientists from the international community contribute their question suggests to create an English master version. The surveys were translated into various languages and then translated back to English to ensure accurate translations. After each wave, questions are updated with some being added and others being removed in order to allow the survey to evolve with the values and concerns of humanity. The topics of the Wave 7 survey is as follows:

- social values, attitudes & stereotypes (45 items)
- societal well-being (11 items)
- social capital, trust and organizational membership (49 items)
- economic values (6 items)
- corruption (9 items)
- migration (10 items)
- post-materialist index (6 items)
- science & technology (6 items)
- religious values (12 items)
- security (21 items)
- ethical values & norms (23 items)
- political interest and political participation (36 items)

- political culture and political regimes (25 items)
- demography (31 items)

To prevent any sampling bias and ensure accurate representation, 1200 interviews must be completed to meet the minimum sample size requirement. The WVS dictates that the sample must be representative of all of the country's residents 18 years or old, but allows the country to determine the specific method of sampling often deciding on a mix of stratified and random sampling. The surveys are completed either in person (primary option) or over telephone (if necessary to reach isolated areas).

Data Selection and Cleaning

The United States was chosen because that was the first data set that was provided at the start of the project. Additionally, since I am an American, I wanted to explore the values of my own country to see how closely they lined up with my perspective of them. Russia was chosen because of personal interest: I had traveled to the country and have grown fond of exploring more about the Russian culture. I also wanted to explore the differences, and similarities, between Russia and the United States. Japan was chosen last to add a starkly differing culture to the mix and to learn more about the values of the country that produces the animes my friends love. The data from all three countries were collected during Wave 6 (2010-2012) to provide a snapshot of the people's values at a certain point in time, in a way controlling for world events that may influence particular values. All data used for this project can be found at the WVS website (<http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp>).

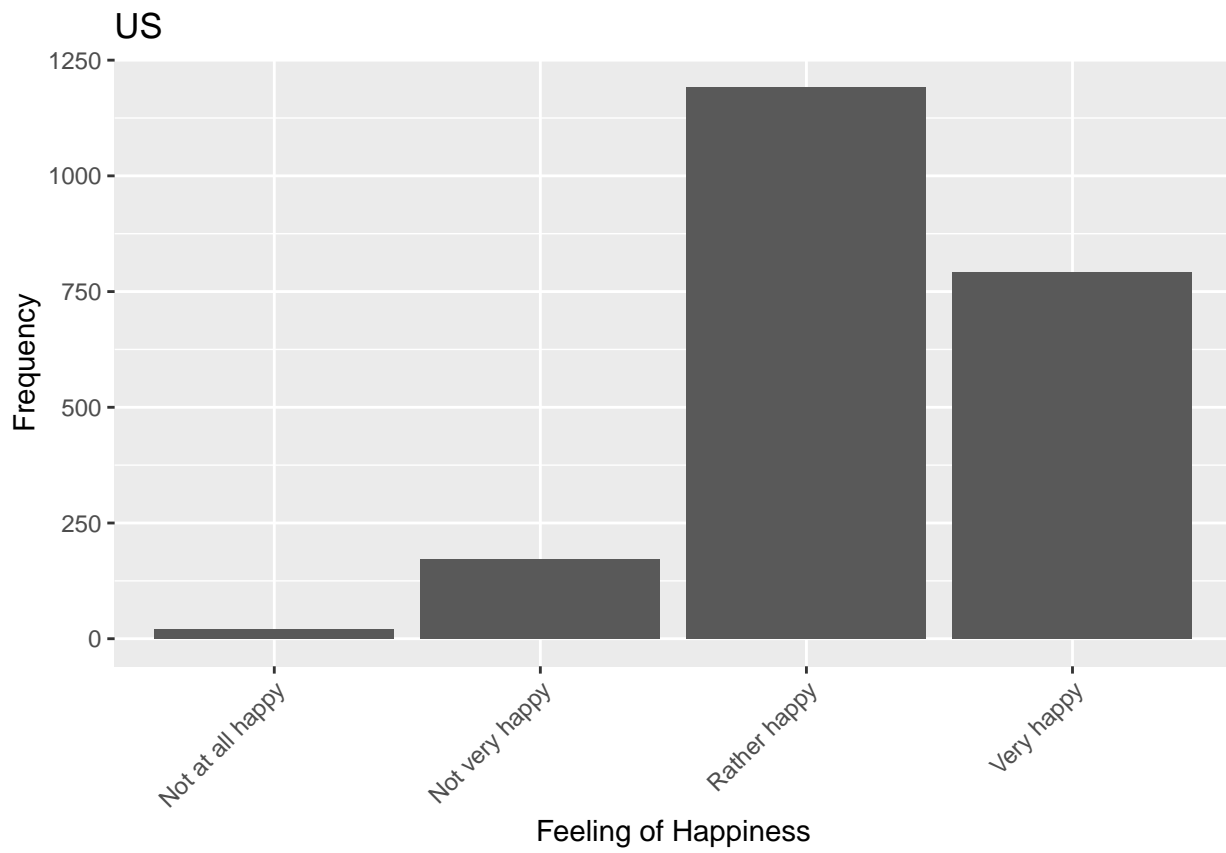
Cleaning

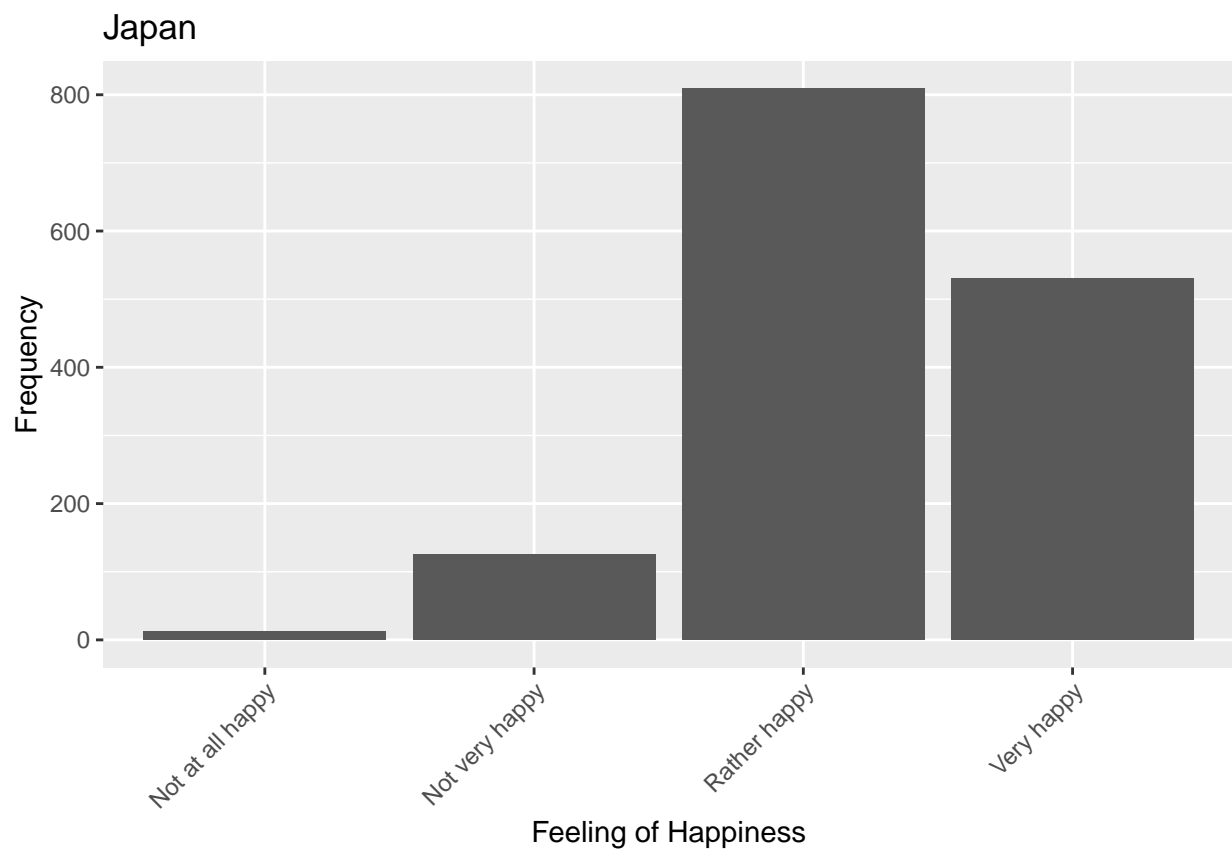
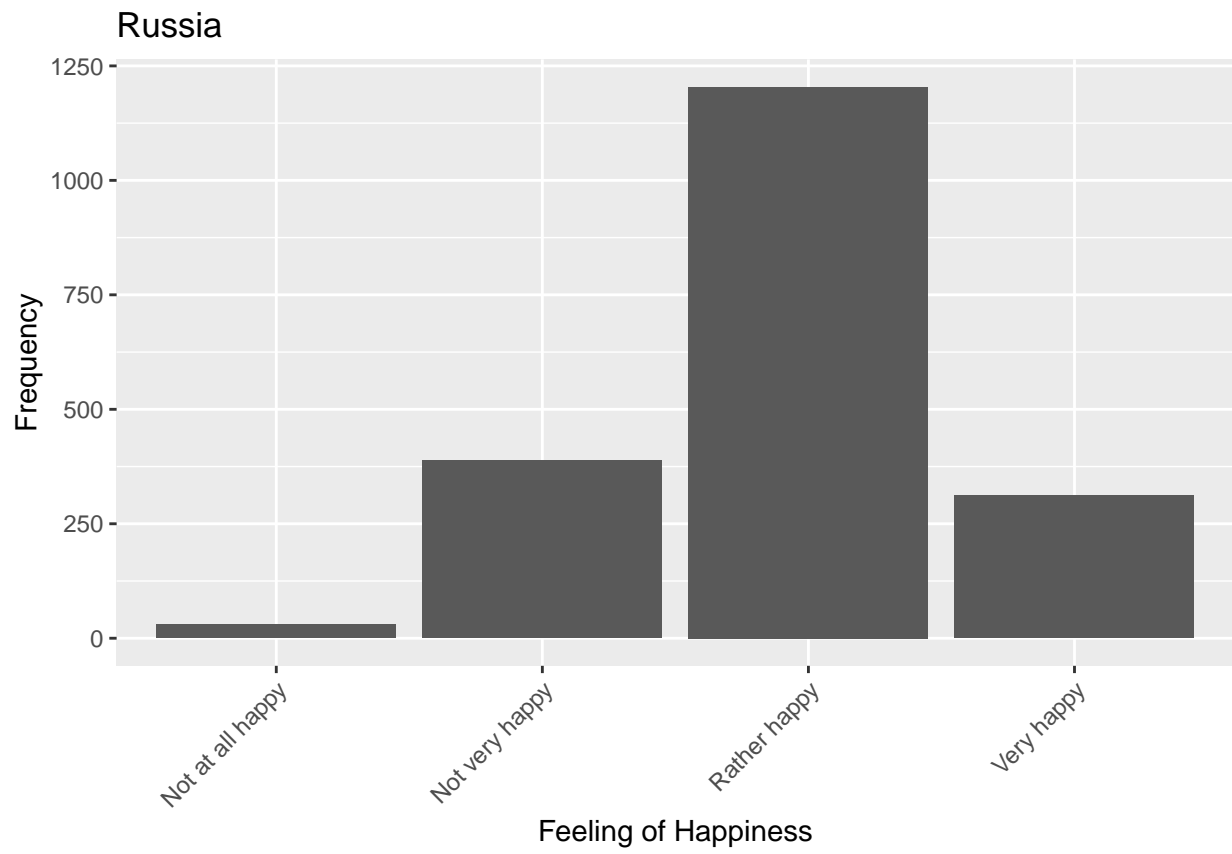
Once the data was inputted into R and converted to a data frame, the cleaning process began by removing columns with repetitive, no longer necessary information (Country Code, etc.). Next, questions that were not asked in each country were identified, saved to vectors by name, and deleted from the main data frame. Using the now smaller data set basic set theory was applied to find the union of all questions asked, intersections among the combinations of countries, and questions unique for each country. Since the WVS is dominated by categorical variables, each column of the data sets were set to be factors and reordered when necessary. The last step was intending to do extra cleaning in the Russian data set. Some of the responses included "DE,SE: Inapplicable;RU: Inappropriate response;SG:Missing;HT: Dropped out survey" and slight variations thereof. These values were supposed to be replaced with "No answer". Due to a lot of unforeseen technical difficulties, this step was not completed leaving the data much messier than intended. However, this project mainly focuses on initial EDA, and thus in future work, this cleaning step will be completed.

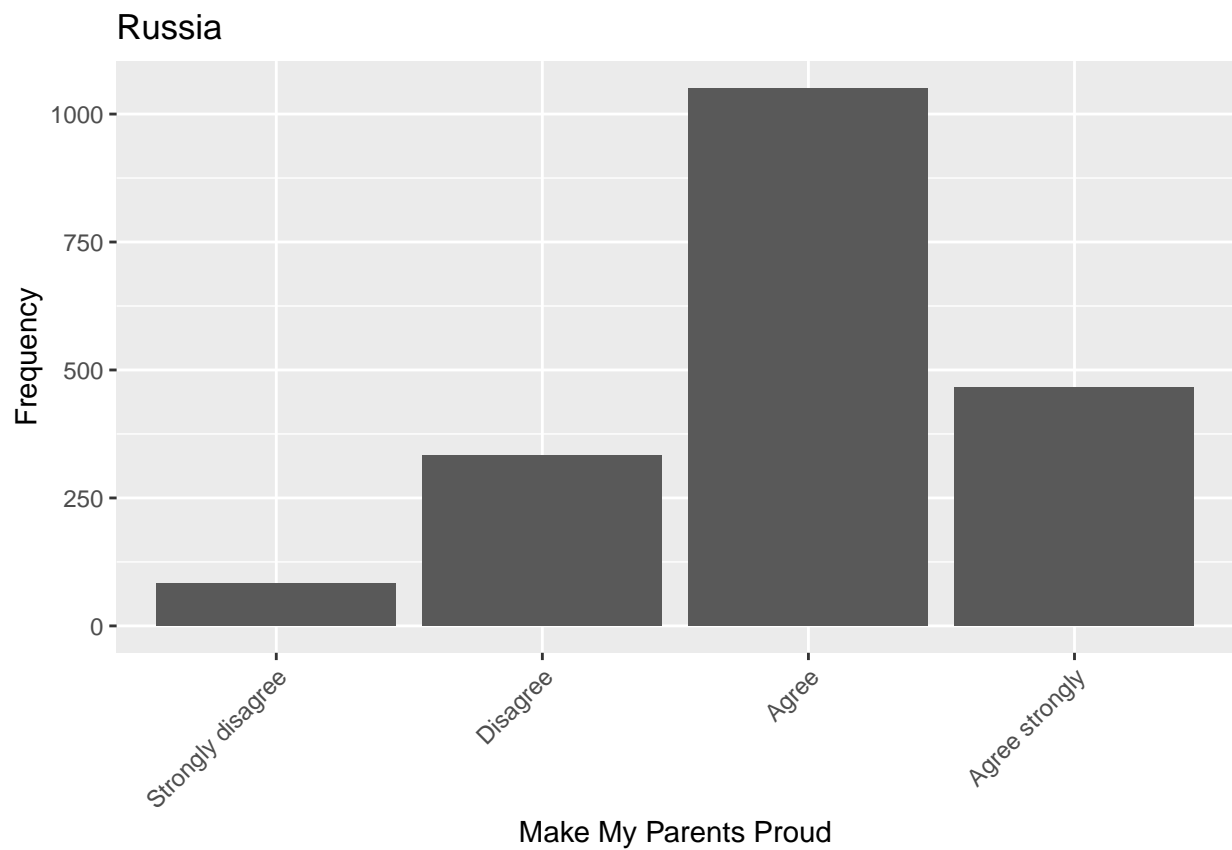
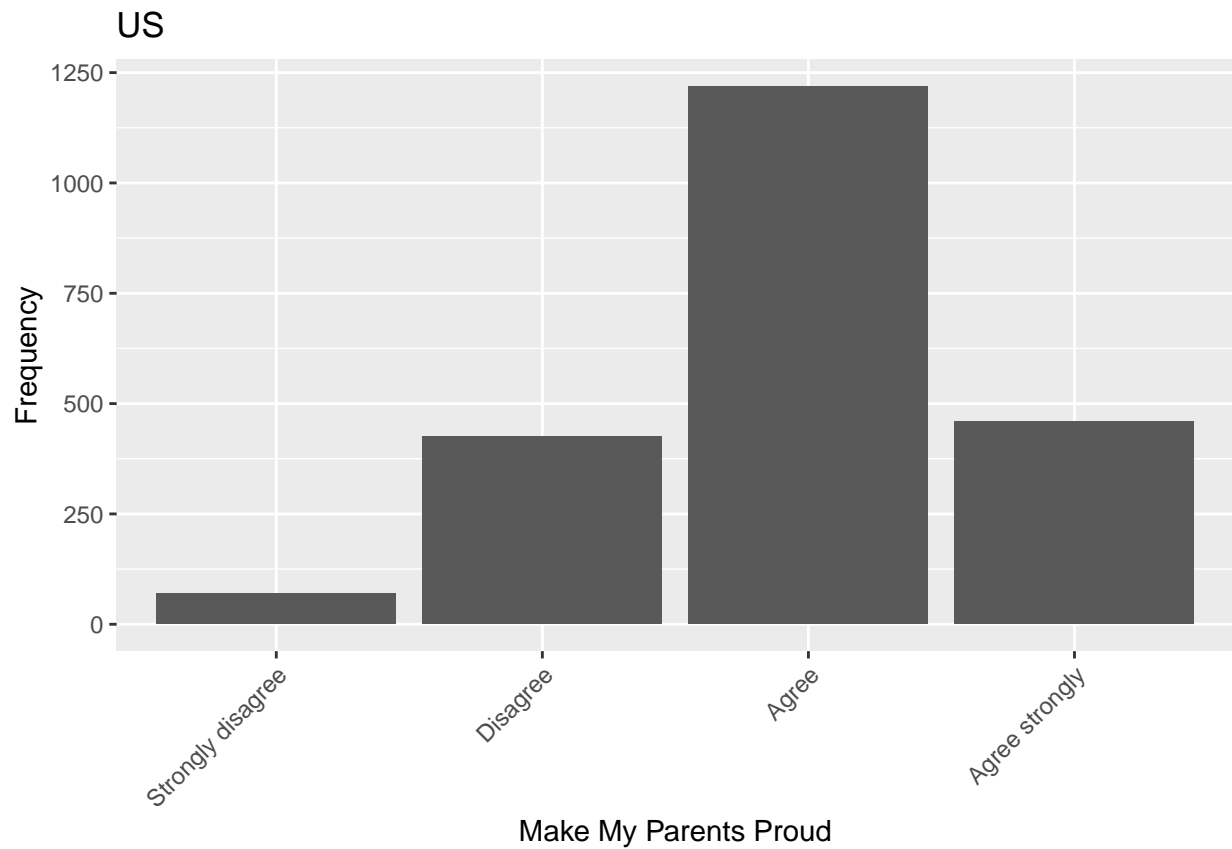
Exploration

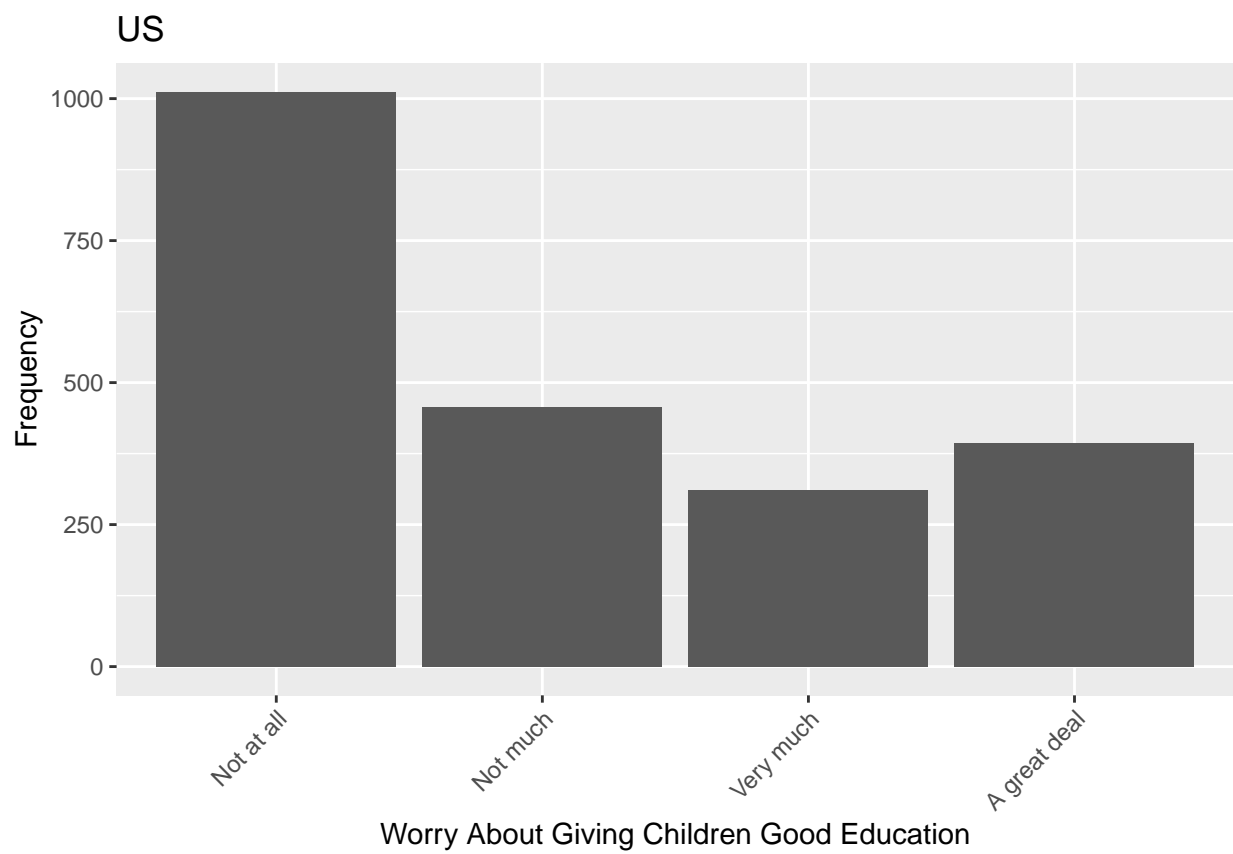
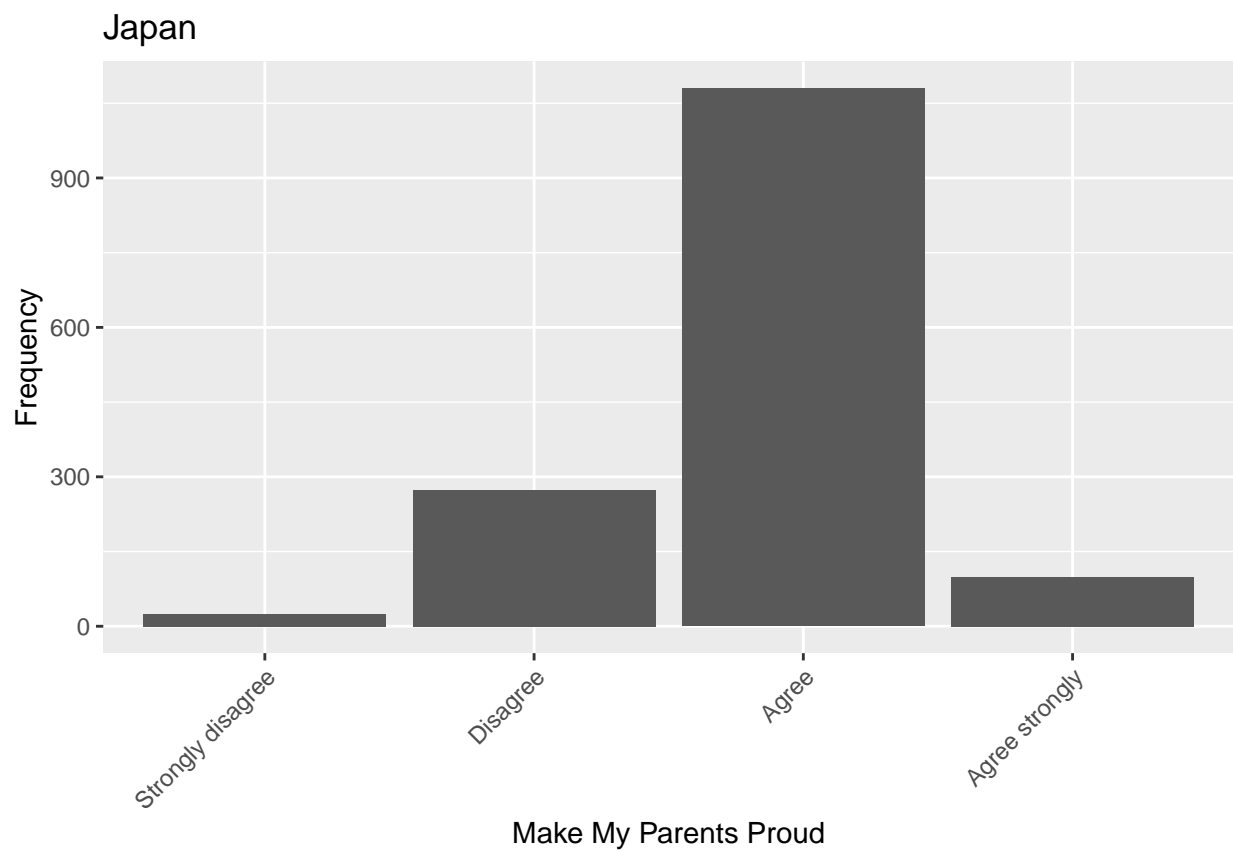
Since this data set is so large, the exploration done in this report will focus on the variables of "Feeling of Happiness", "Worries: Not being able to give one's children a good education", and "One of my main goals is to make my parents proud" and subsequent comparisons among the three countries. For relationships between different variables, please utilize the corresponding Shiny app for this project.

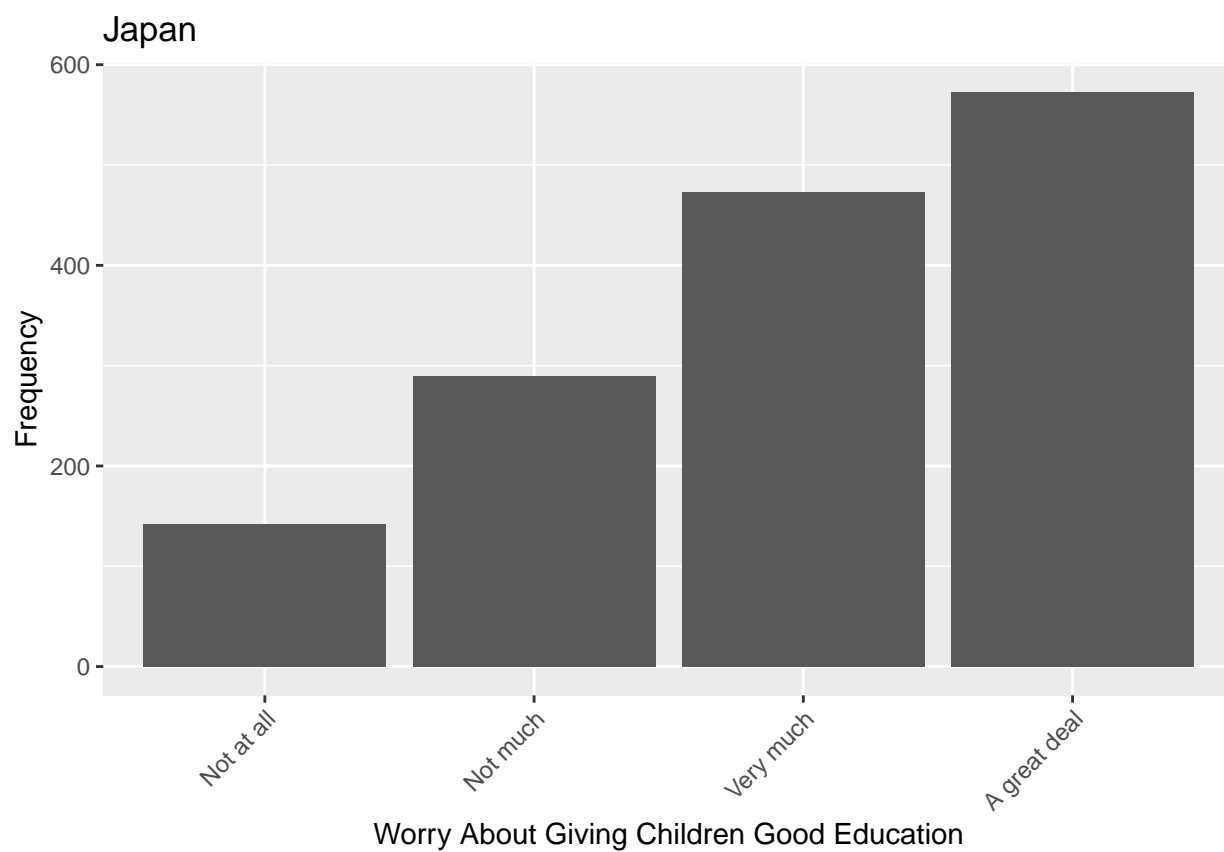
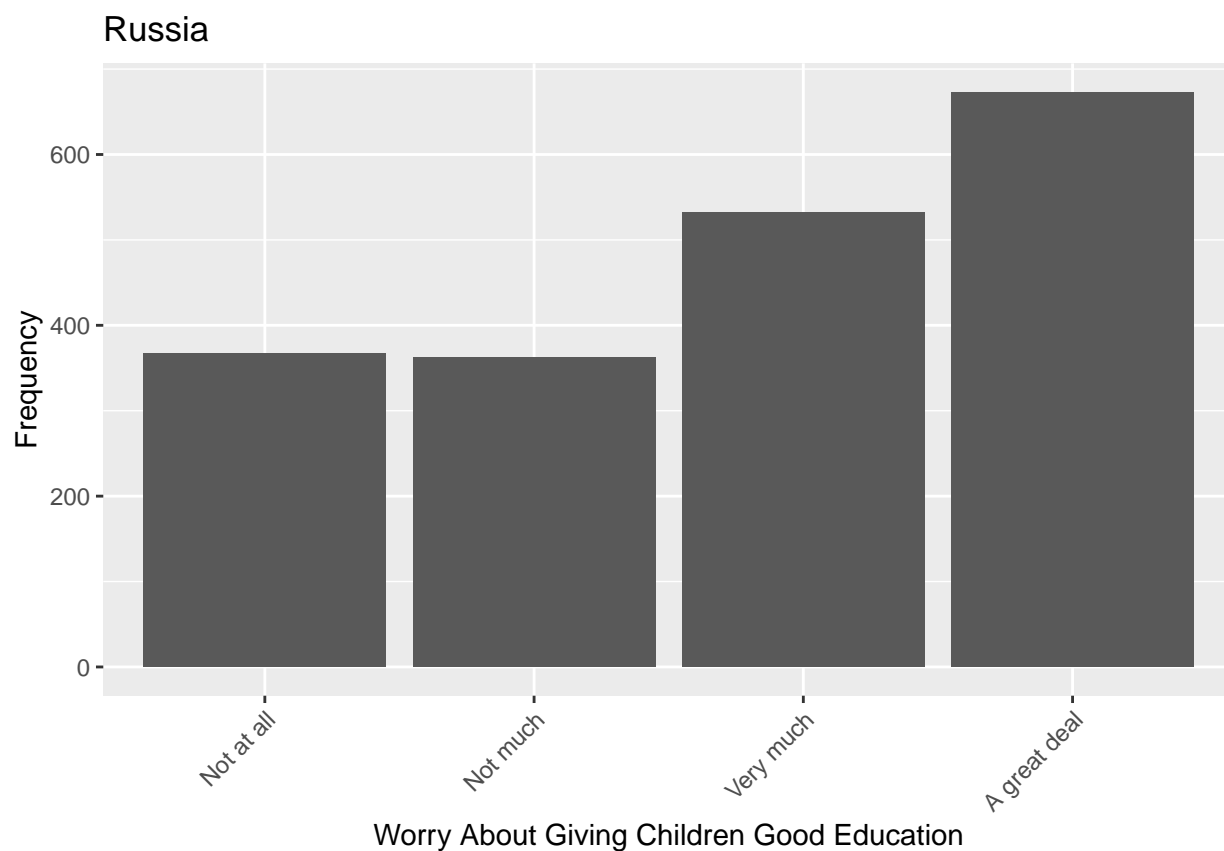
Single Variables









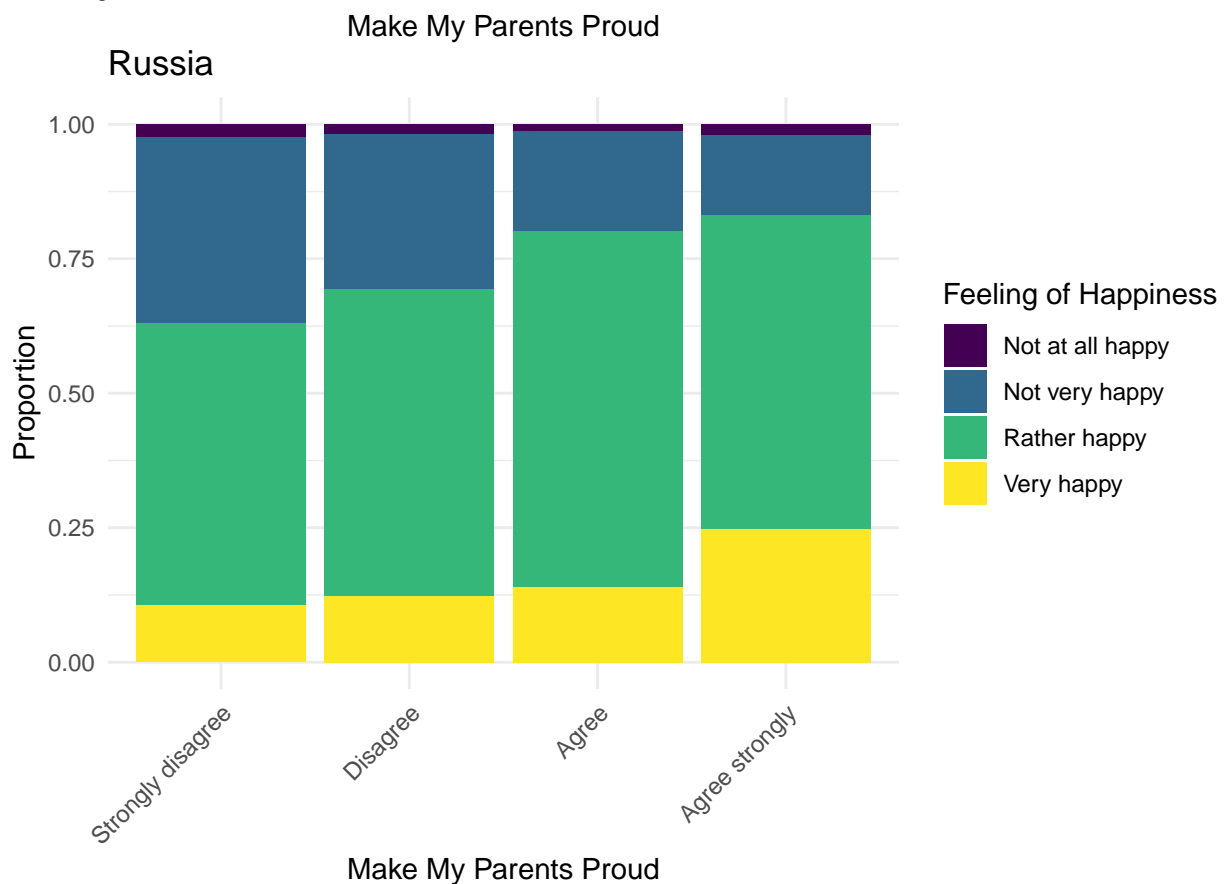
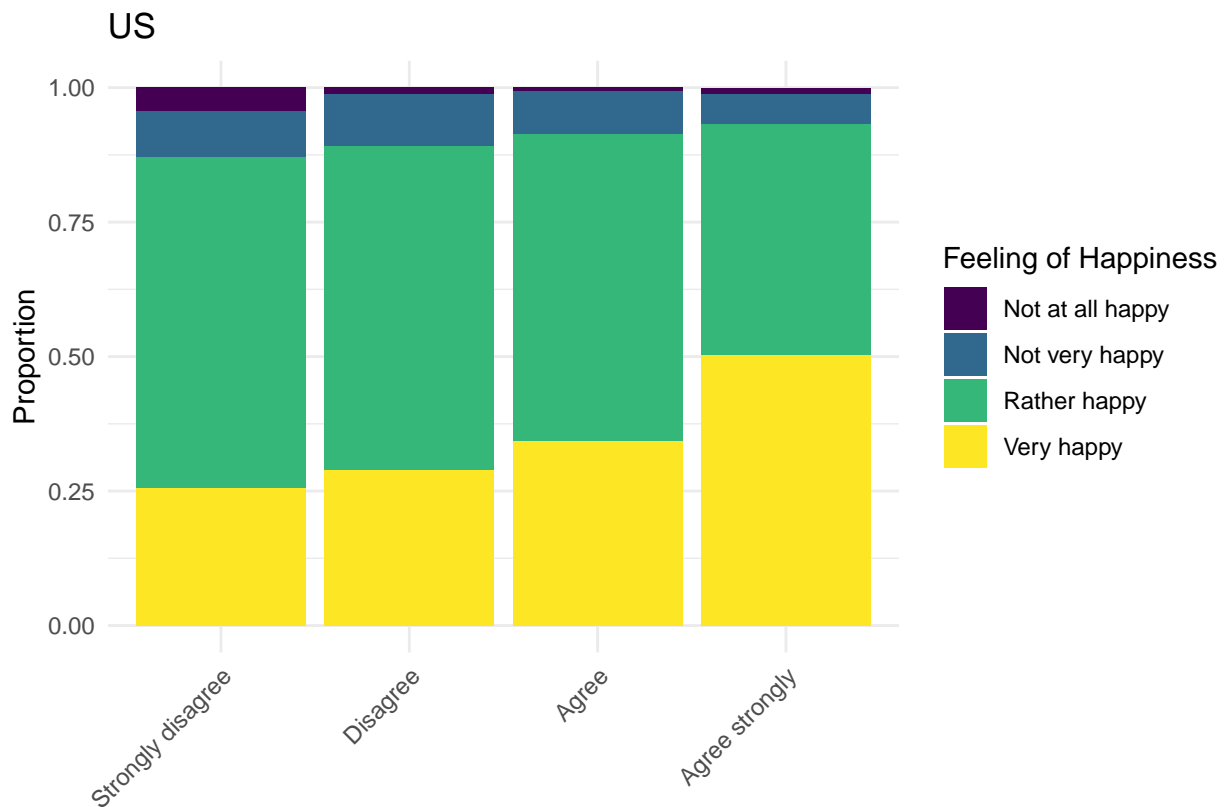


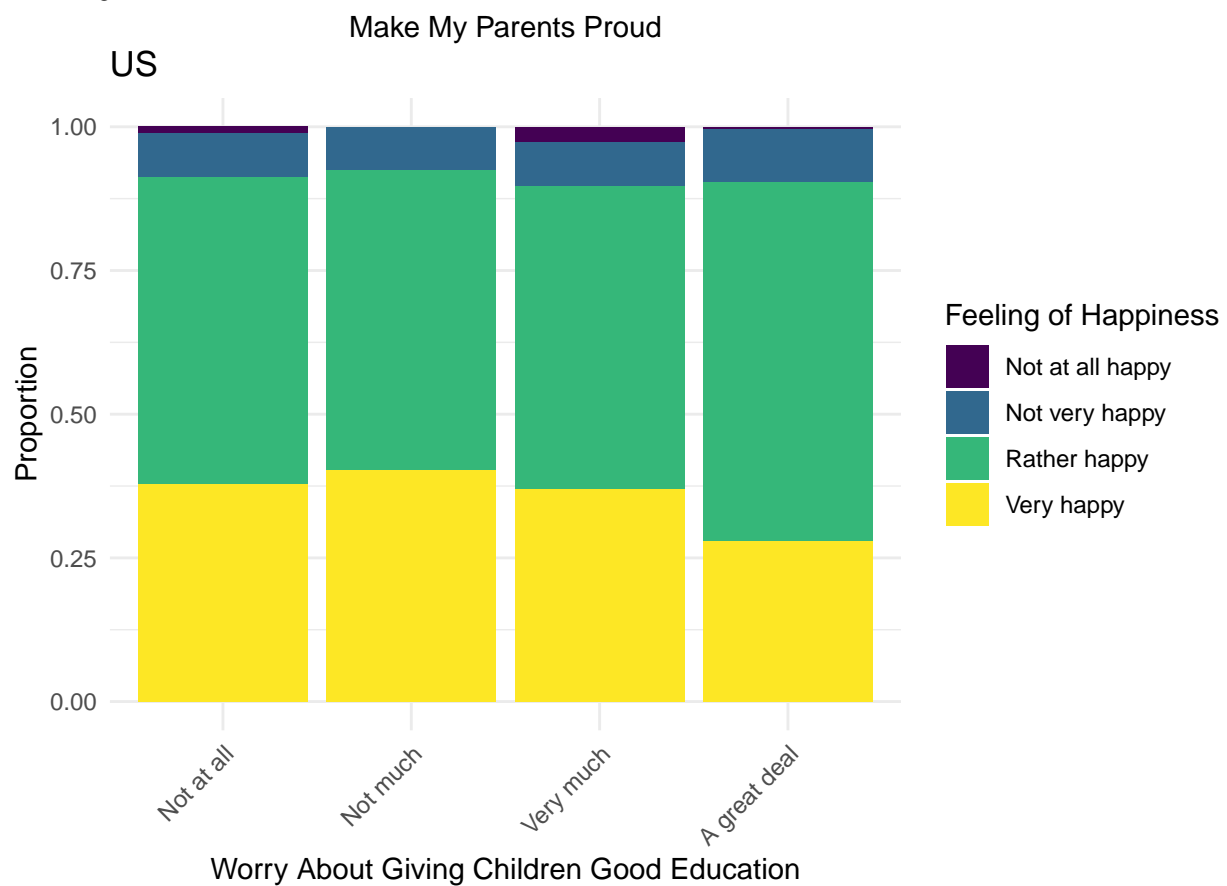
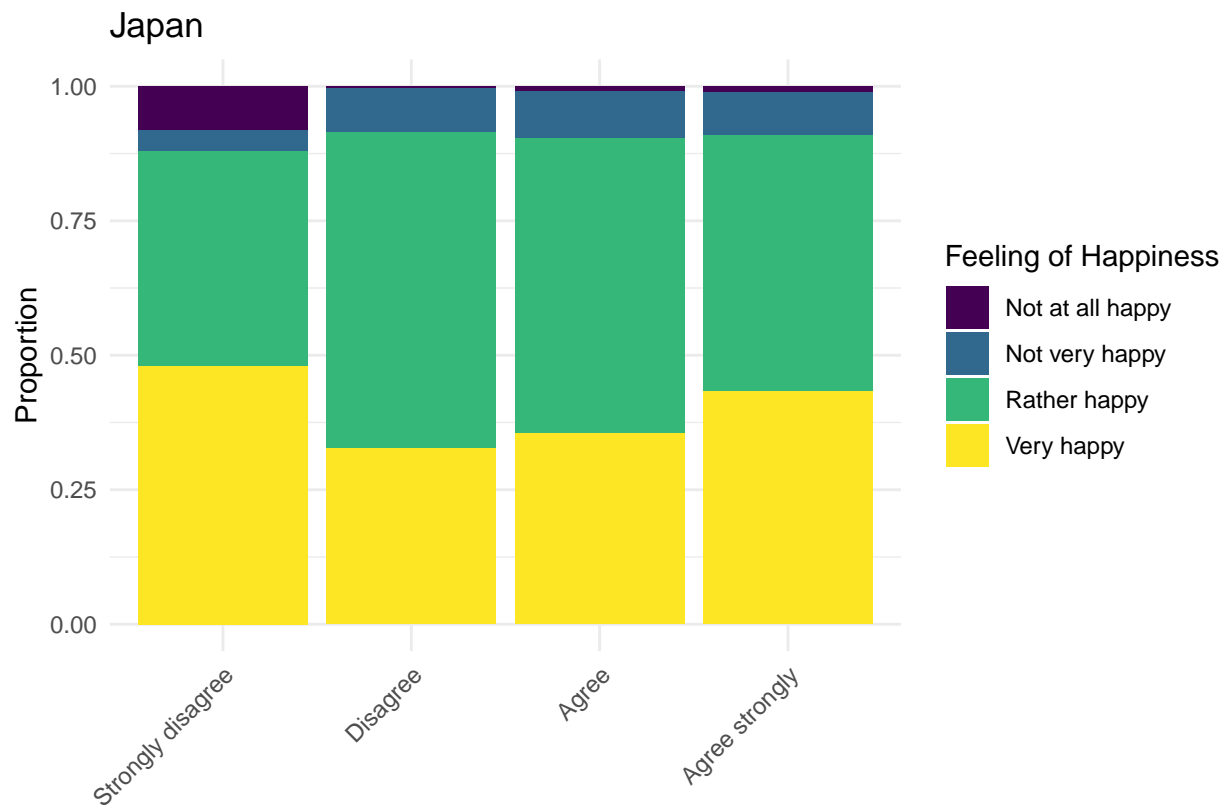
Overall, the plots regarding the happiness follow the same relative shape for all three countries with the concentration of answers centered around the “rather happy” response. And it can be noticed that Russia has much lower responses for “very happy” than the other two countries. (One possibly reason as for why this might be is the significant lack of sunshine in the winter months, but further research is needed for evidence.) These plots suggest that the general distribution of the happiness answers are universal, centered at an average happy level, but the slight differences imply that it is not completely independent of country.

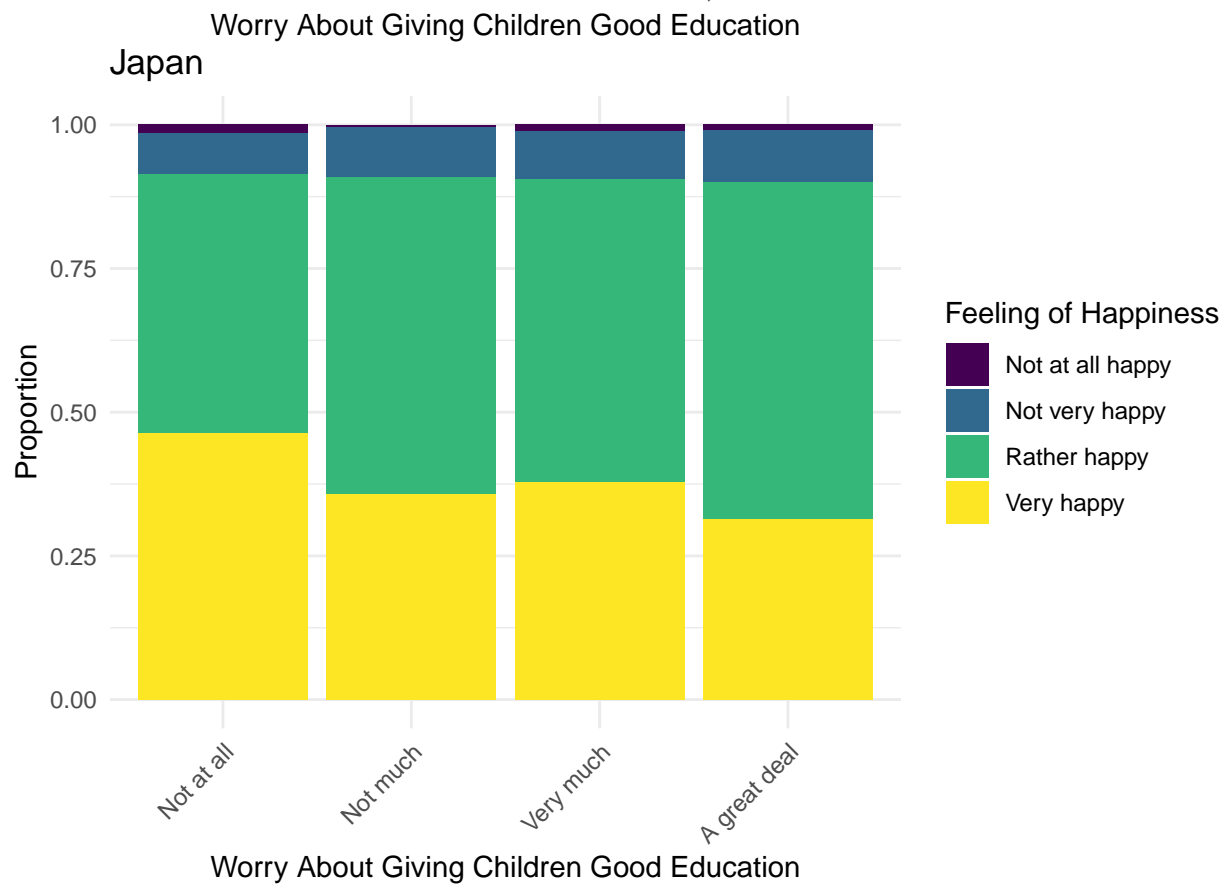
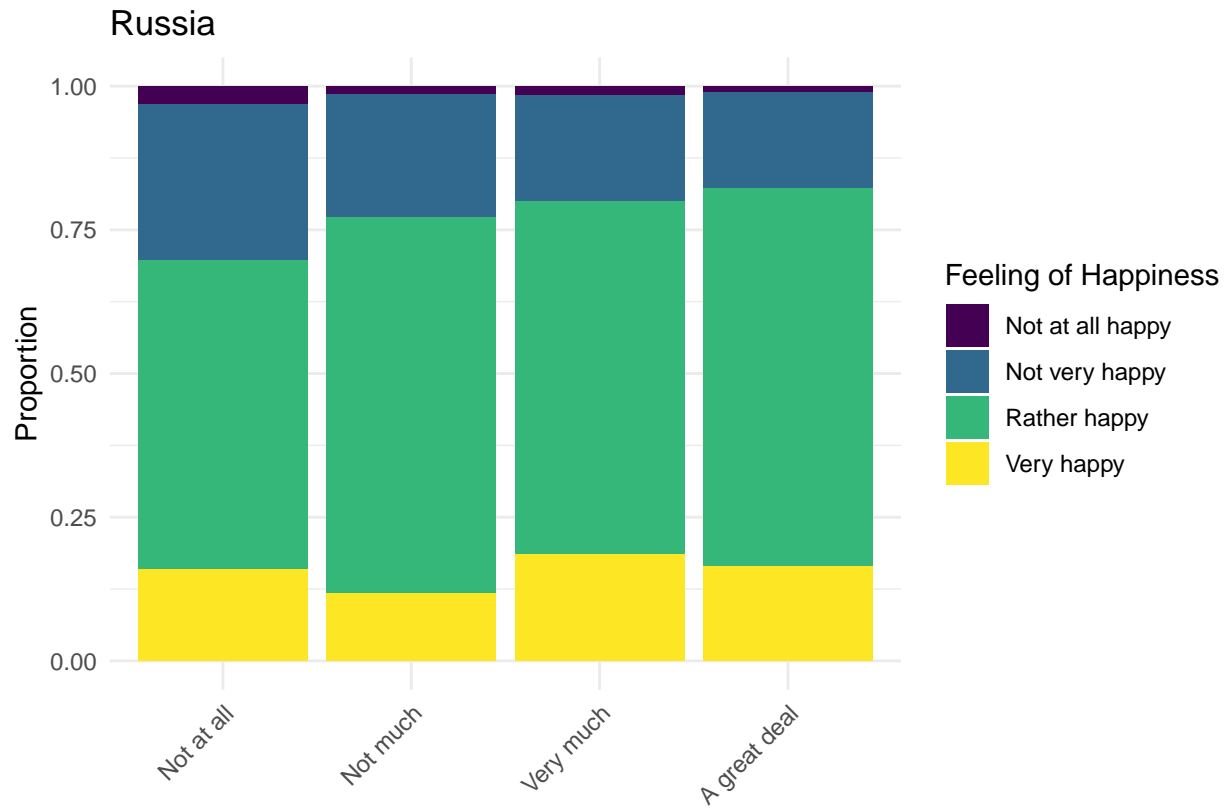
Similarly, the plots regarding the “parents proud variable” again follow the same relative shape for all three countries with the concentration of answers centered around the “agree” response. This time Japan is the one with a significant decrease in the number of “strongly agree”. Further research and data analysis are needed to explore exactly why this is, but it may be due to differences in cultural expectations of the parental figures.

Last, there are significant differences in the distribution of responses when pressed on worries about not being able to give one’s children a good education. First, participants from the United States are most confident that they will be able to give their children a good education. The Russian distribution is more evenly distributed with a gentle uptick at the end, allowing the “worry a great deal” to become the most frequent choice. Finally Japan has an almost positive linear relationship between the variable and the frequency. The reasons for these differences are numerous and further research is needed. Some ideas may be that American schools are better than those in Japan and Russia, government student loans are more available to American students than in other countries providing relief to low income families, or Americans are more laid-back when it comes to their children’s education.

Double Variables







For the plots comparing happiness and parent's pride, there is a very slight increase in the proportion of those who are "very happy" when concern about one's parents' pride goes up for US and Russia. In Japan, this distribution shows little relationship. When considering just general happiness, there is no significant difference when increasing concern about parental pride. The WVS survey consists of participants 18+ years old, and they may not be as concerned with pleasing their parents as they once did as children.

For the plots comparing happiness and worrying about children's education, there is a lack of an overall trend for the three countries. The only significant point is that Russia has an overall fewer proportion of "very happy" responders, as we addressed in the previous section.

Question Sets

The questions that were not asked in all of the countries include an interesting mix of topics. There are some questions that are logically excluded such as a person's confidence level on certain organizations that the three countries are not members since those do not pertain to those three countries. The more intriguing ones are those that seem to be more general in their scope such as additional personality traits, election integrity and government corruption inquiries. One reason for the personality questions to be excluded is that the current ones already account for the attribute being measured and thus unnecessary to ask a second time. It can also be argued that the election and corruption inquiries are more focused towards less developed countries where the governments and economies tend to be more unstable. Thus, it was not deemed necessary to ask these questions to the industrialized nations of the United States, Russia, and Japan. However, governments can be corrupt and elections can be unfair just as easily as struggling democracies; thus legitimizing a concern that this data set is missing potentially missing key information and trends by not asking these questions in the three studied countries.

The unique values for the United States are rather straightforward in that they are questions that are appropriate for the only North American country in the data set. In regards to Russia's three unique variables, two of them relate to the Commonwealth of Independent States (CIS) - an organization of ten former members Soviet Union. The last asks about the justifiability of prostitution. This is surprising to see as a unique question for Russia since prostitution happens everywhere in the world. Similar to the US, Japan has additional questions specifically over the Asian concerns. However, Japan has the additional personality question "It is important to this person to do something good for society". This one is most surprising since the desire to help others and better the community should be a goal everyone should have regardless of nation of origin. By excluding this question for other countries, the survey is failing to fully understand the nature of people's values on compassion and community support.