To Do’s Cristina & Laura

Check Cristina

Check Laura

Step 1 – Data upload

|  |  |  |
| --- | --- | --- |
| **Task** | **Details** | **Check** |
| Find data |  |  |
| Load data |  |  |
| Clean data | Remove missing valuables |  |
| Rename names (if nessessary) |  |  |
| Identify key variables |  |  |
| Remove not needed variables |  |  |
| Check if riskitem needs to be reverted | VWS File |  |
| Create a list for hardship (Countryfacts) | Years:   * WVS5: 2005 – 2009 * WV6: 2010 – 2014 * GPS: 2012 - 2013 |  |
| Add countryfacts to data set |  |  |

|  |  |  |
| --- | --- | --- |
| **Open Questions** | **Explanation** | **Question from** |
| Countryfacts: do we need 3 different countryfacts-files per wave/dataset? | Countryfacts such as GDP, infant mortality, etc. can change over time. In the paper there is no explanation why only one dataset for countryfacts was created.  We think it might be useful if there were two times of countryfacts for each datafile to understand the hardship better  How should we combine the courtyfacts per year and calculate the mean of all of the years (e.g. cf 2005, 2006, 2007, 2008, 2009/2005-2009)? | Laura & Cristina |
| Countryfacts | How can you add the countryfacts to the countries in the original dataset? | Cristina |

Step 2 – Descriptive Description

|  |  |  |
| --- | --- | --- |
| **Task** | **Details** | **Check** |
| Describe the data | * Countries * Participants * Age (range, distribution, mean, median, etc.) * Gender (per country, per age) * Distribution of age per country * Risk-item (difference in gender, age, country) * Hardship per country |  |
| Compare the datafiles | * Which countries are in all three datafiles |  |

|  |  |  |
| --- | --- | --- |
| **Open Questions** | **Explanation** | **Question from** |
| Visualisazion | Should we visualize in total:   * Gender * Age * Countries * Gender / Risk * Age / Risk * Countries / Risk   Note/Info Cristina: analysis only with the combined waves  What else? | Laura & Cristina |

Step 3 – Analysis Country, Gender, Age

|  |  |  |
| --- | --- | --- |
| **Task** | **Details** | **Check** |
| Create table | * Country * Isocode * n (count of participants) * female percentage (%) * mean age * age range * risktaking * intercept * r\_squared * slope (age, gender) |  |
| Plot regression for each country |  |  |
| Analyze the intercept |  |  |

|  |  |  |
| --- | --- | --- |
| **Open Questions** | **Explanation** | **Question from** |
| intercept | Is the table ok like that or did I miss something? | Laura |
| Intercept interpretation | How do I find out about the countries I want to analyse in detail?  🡪 e.g. is the intercept the relevant number to evaluate which coutries fall out of the norm? |  |

Step 4 - Comparison

|  |  |  |
| --- | --- | --- |
| **Task** | **Details** | **Check** |
| Compare the countries from the different datafiles | * n countrys * risktaking per country |  |

|  |  |  |
| --- | --- | --- |
| **Open Questions** | **Explanation** | **Question from** |
|  |  |  |

Notes Cristina

**Countries**

Wave 5: 51 countries

Wave 6: 60 countries

Wave 5 + 6: 77 countries

**Participants per Wave**

Wave 5: 70 308

Wave 6: 86 220

Wave 5 + 6: 156 528

**Participants per country (Waves 5 + 6)**

See country\_facts

**Participants per age (Waves 5 + 6)**

See age\_counts

**Age (Wave 5 + Wave 6)**

Median: 39 y

Mean: 41.62

Range: 15 - 102

**Age distribution in countries**

See agepercountries\_counts

**Gender per wave**

**Wave 5:** 36 520 (females); 33 788 (males)

**Wave 6:** 44 864 (females); 41 356 (males)

**Combined:** 81 384 (females); 75 144 (males)

**Genderdistribution per country**

See sexpercountries\_counts

**Age distribution per country (mean, median, minimum, maximum age**

See age\_distribution\_per\_country

**Risk by gender**

Females: 3.02

Males: 3.41

**Risk by countries**

See risk\_by\_country\_lab

**Risk by age**

See risk\_by\_ageacat

**Hardship index distribution**

Regression lines per country (age vs ristaking)