**Data Appendix**

Dägg and Janné, "Money, Happiness and the Midlife Crisis"

The final data set (*Analysis.csv)* contains observations representing 10,355 individuals from 7 countries. The distribution of observations across countries is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Country | Individuals | Relative Frequency (Percent) | Cumulative Frequency(Percent) |
| China | 2837 | 27.39739256 | 27.39739256 |
| India | 1887 | 18.22308064 | 45.6204732 |
| Indonesia | 910 | 8.788025109 | 54.40849831 |
| Jordan | 931 | 8.990825688 | 63.399324 |
| Pakistan | 1629 | 15.73153066 | 79.13085466 |
| Russia | 916 | 8.845968131 | 87.97682279 |
| US | 1245 | 12.02317721 | 100 |
| <NA> | 0 | NA | NA |
| Total | 10355 | 100 | 100 |

The final analysis data set (*Analysis.csv* ) contains both individual level and country-level variables.

The final analysis data set contains two individual-level variables: *satis* and *age*.

Definitions, summary statistics and illustrations of the distributions of those two variables are shown below.

Variable name: *satis*

Source: Pew data set

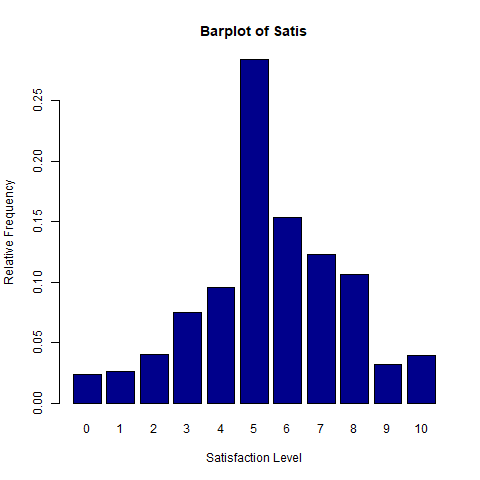
Values: Integers from 0 through 10

Coding: Respondent's assessment of her/his overall satisfaction with life on a scale of 0 (least satisfied) to 10 (most satisfied)

Frequency table:

|  |  |  |  |
| --- | --- | --- | --- |
| Satisfaction Level | Freq | Relative Frequency (Percent) | Cumulative Frequency(Percent) |
| 0 | 248 | 2.394978271 | 2.394978271 |
| 1 | 269 | 2.597778851 | 4.992757122 |
| 2 | 417 | 4.027040077 | 9.019797199 |
| 3 | 776 | 7.493964268 | 16.51376147 |
| 4 | 991 | 9.570255915 | 26.08401738 |
| 5 | 2947 | 28.45968131 | 54.5436987 |
| 6 | 1587 | 15.3259295 | 69.8696282 |
| 7 | 1272 | 12.28392081 | 82.15354901 |
| 8 | 1105 | 10.67117335 | 92.82472236 |
| 9 | 334 | 3.22549493 | 96.05021729 |
| 10 | 409 | 3.949782714 | 100 |
| <NA> | 0 | NA | NA |
| Total | 10355 | 100 | 100 |

Relative frequency chart:



Variable name: *age*

Source: Pew data set

Values: Positive integers

Coding: Respondent's age in years at last birthday

Summary statistics:

mean: 39.63

standard deviation: 12.17

minimum: 21

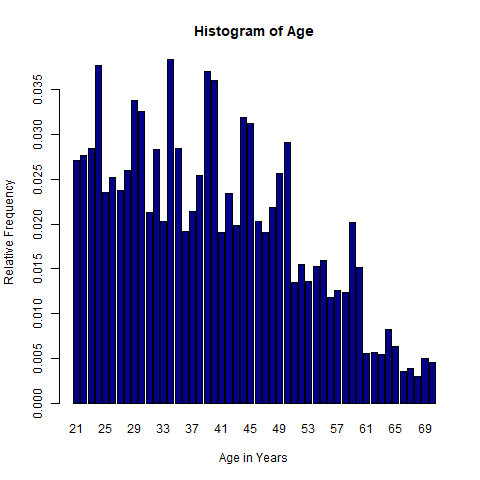
maximum: 70

25th percentile: 29

median: 39

75th percentile: 49

Histogram:



The final analysis data set (*Analysis.csv)* contains three country-level variables: *cm\_satis, gov\_cons*, and *gd\_pc*

Definitions, summary statistics and illustrations of the distributions of those three variables are shown below.

Variable name: *cm\_satis*

Source: Pew data set

Values: Real numbers from 0 through 10

Definition: Mean value of *SATIS* for all respondents from the country

Summary statistics:

mean: 5.51

standard deviation: 0.74

minimum: 4.74

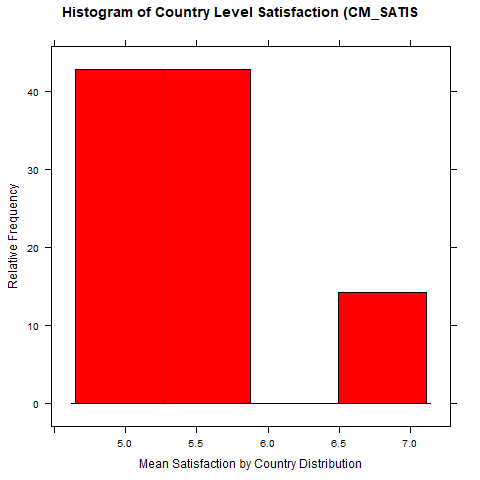
maximum: 7.02

25th percentile: 5.09

median: 5.34

75th percentile: 5.76

Histogram:



Variable name: *gov\_cons*

Source: WDI

Values: Positive real numbers

Definition: General government consumption expenditure for 2002, as a percentage of GDP

Summary statistics:

mean: 14.16

standard deviation: 5.36

minimum: 7.26

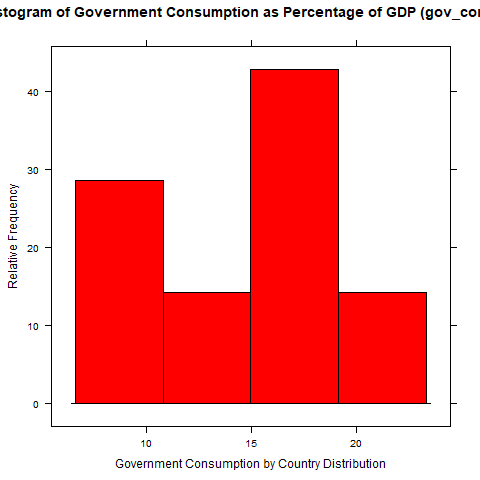
maximum: 22.70

25th percentile: 8.72

median: 15.04

75th percentile: 17.95

Histogram:



Variable name: *gdp\_pc*

Source: WDI

Values: Positive real numbers

Definition: Income per capita for 2002, current US dollars

Summary statistics:

mean: 6495

standard deviation: 13978

minimum: 483

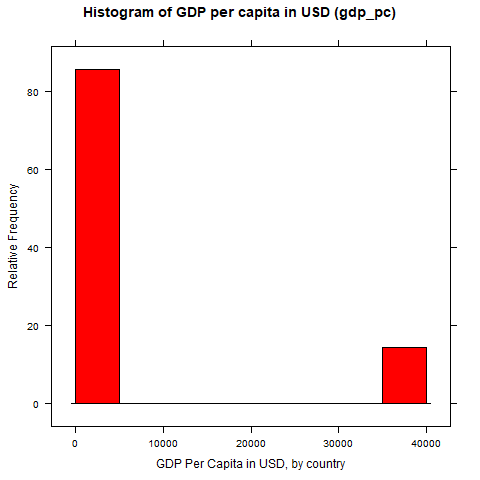
maximum: 38175

25th percentile: 487

median: 1135

75th percentile: 2375

Histogram:



Created by Richard Ball