Research topic & topic Parts:

Our topic will investigate the happiness and wellbeing for Thailand meaning that we want to tell whether a person is happy, quite happy or unhappy as well as assessing the relation between happiness and the other 5 explanatory variables to have a total of 6 variables as shown in the next pages.

The target population for the 7th wave of World Values Survey was eligible voters over the age of 18, where The sampling technique used in this survey was multi-stage systematic random sampling.

Firstly, I categorized different values together as mentioned in the proposal to prepare the data for our 'Categorical data analysis' which leaves us with the values written below.

Secondly, Regarding the missing values or unspecified answers like (Don't Know, Didn't answer,…) I removed them since their frequency was around 3-15 times which wouldn't affect the analysis significantly which leave us with n = 1454.

Moreover, The project will consist of 2 parts, 1st part is considered with the descriptive part of the variables(creating 2-way& 3-way tables , Checking independence& associations…),while the second will include the modeling part.

The Analysis was done with R and Python.

Variables & Distributions:

1. Q46: Feeling Happy. (Ordinal variable, response variable)

   With the following categories:

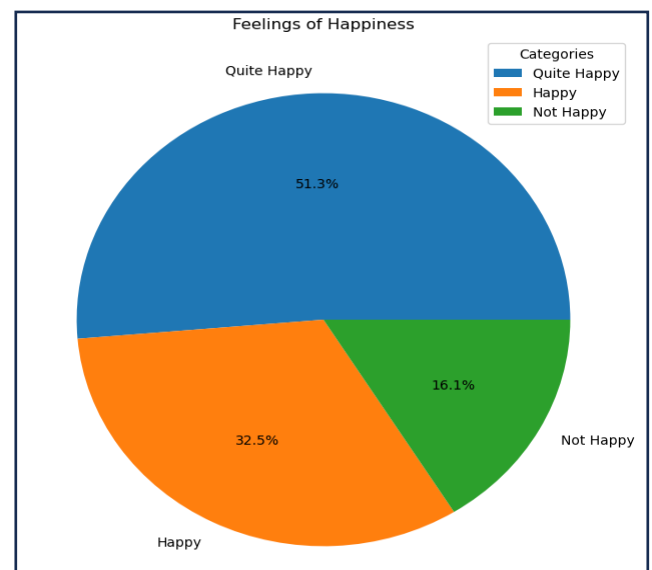   1- Happy.
   2- Quite happy.
   3- Not happy.

Figure 1: Feeling of Happiness Distribution

2. Q48: How much freedom of choice and control (Binary Variable, Explanatory Variable)

    With the following categories:
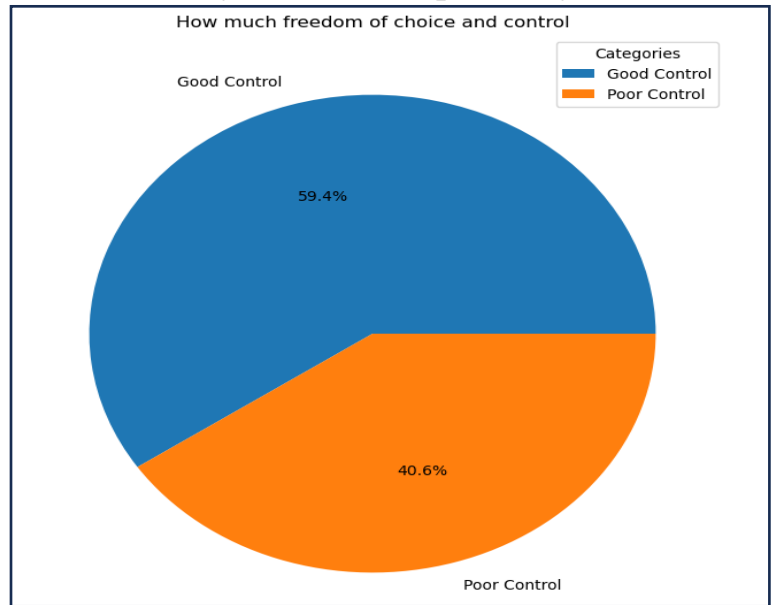
 1-Good control

 2-Poor control

Figure 2: Freedom of Choice & Control Distribution

3. Q50: Satisfaction with financial situation of household (Binary variable, explanatory variable)

    With the following values:
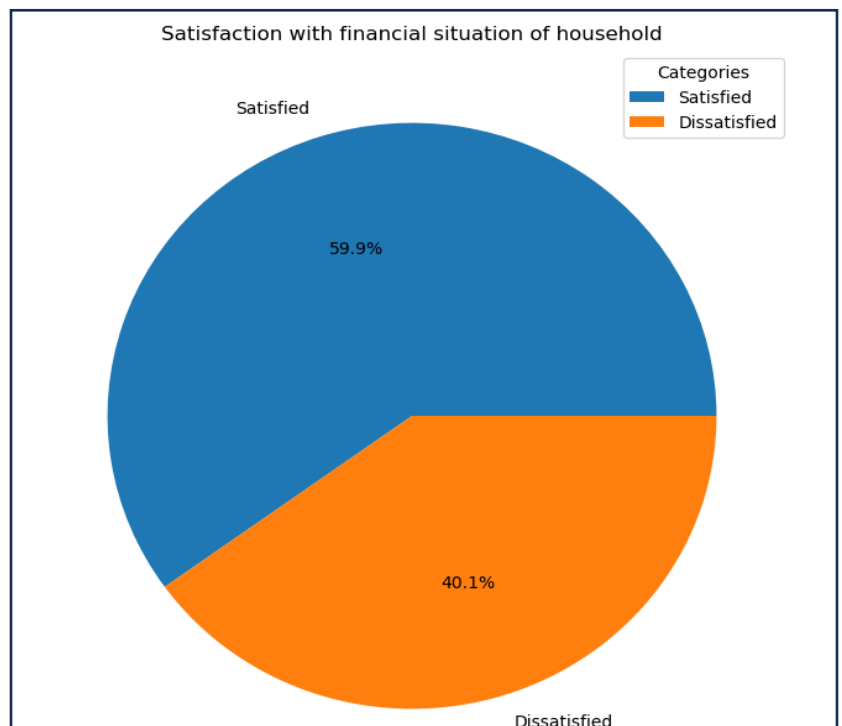
    1-Satisfied

    2-Dissatisfied

Figure 3: Satisfaction with financial situation of household Distribution

4. Q54: Frequency you/family (last 12 month): Gone without a cash income (Nominal variable, explanatory variable)

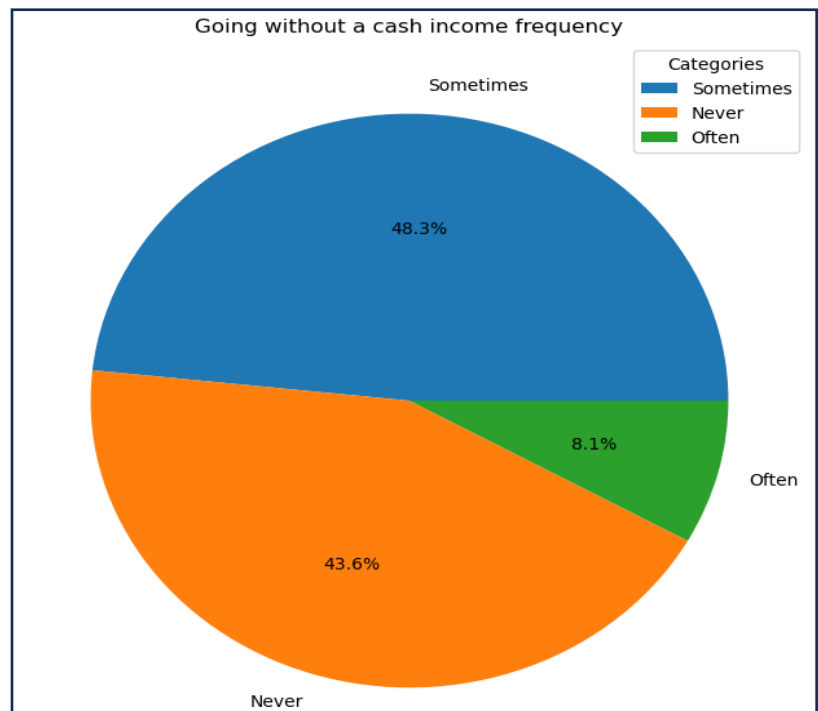With the following categories:
   1-Often
   2-Somtimes
   3-Never

Figure 4: Going without a cash income Frequency.

5. Q158: Science and technology are making our lives healthier, easier, and more comfortable (Binary variable, explanatory variable)

With the following categories:
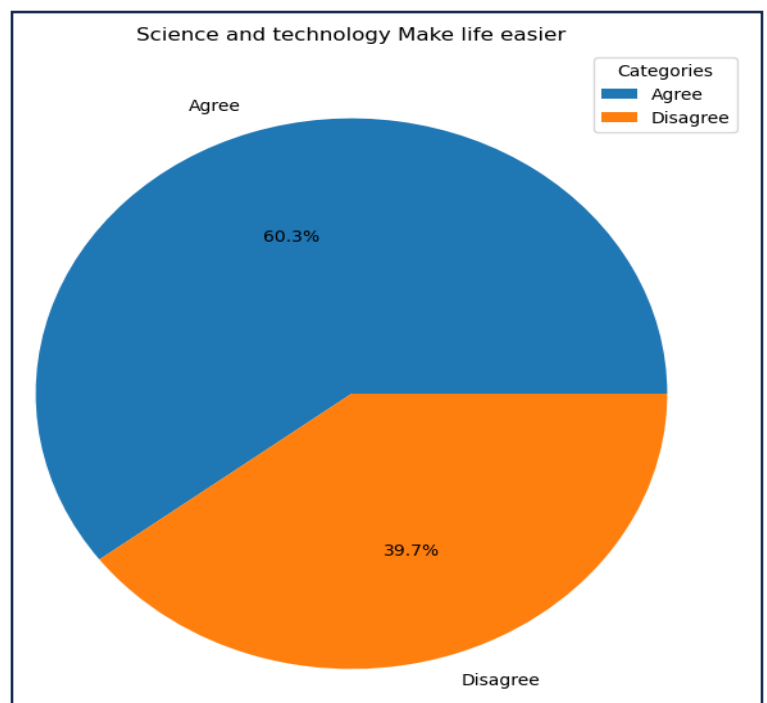   1- Agree.
   2- Disagree.

Figure 5: Science & technology make life easier Distribution.

6. Q260: Sex (Nominal variable, explanatory variable)

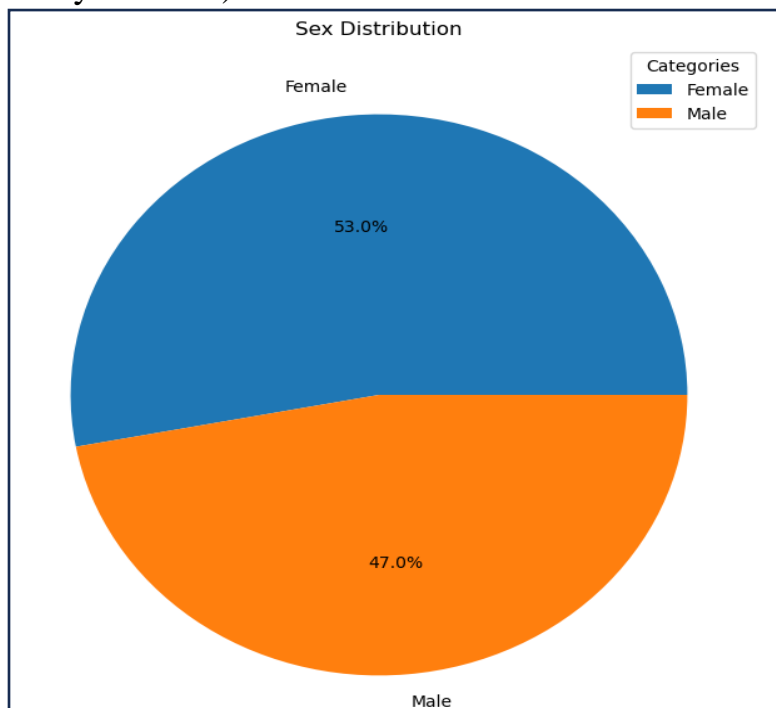With values:

  1- Male

  2- Female

Figure 6: Sex Distribution

1ˢᵗ Part: Contingency tables

2-Way Tables:

Testing independence & association measures:

Independence hypothesis: H_0: $\pi_{ij} = \pi_{i+} \pi_{+j}$ *For all i & j* (α = 0.05)

1 - Feeling of happiness& how much freedom of choice and control:

how much freedom of choice and control	Feeling of Happiness		
	Happy	Quite Happy	Unhappy
Good Control	280	477	108
Poor Control	196	265	128

Table 1: Feeling of happiness & how much freedom of choice and control contingency tables.

Pearson Chi-square test	Gamma measure of ordinal association ($\widehat{Y}$)
P-value = 0.000002730	0.088480037

Table 1-1: Gamma ($\widehat{Y}$) & Pearson chi-square test statistic.

From the P-value(less than 0.05) we **reject** H_0 with 0.05 significance level, therefore, with 95% conf.level, we have enough evidence to say that there's an association between Feeling of happiness & freedom of choice, **hence** we can describe that relationship with gamma $\widehat{Y}$ which indicates that there's a weak positive relationship between the 2 variables such that as freedom of choice & control decreases(poorer control), people tend to be more unhappy.

2-Feeling of Happiness & Satisfaction with financial situation of household

Satisfaction with financial situation of household	Feeling of Happiness		
	Happy	Quite Happy	Unhappy
Satisfied	306	461	106
Dissatisfied	170	281	130

Table 2: Feeling of happiness & Satisfaction with financial situation of household

Pearson Chi-square test	Gamma measure of ordinal association($\widehat{Y}$)
P-value = 0.000001106	0.1908475

Table 2-1: Gamma ($\widehat{Y}$) & Pearson chi-square test statistic.

From the P-value(less than 0.05) we **reject** H_0 with 0.05 significance level, therefore, with 95% conf.level, we have enough evidence to say that there's an association between Feeling of happiness & Satisfaction with financial situation of household, **hence** we can describe that relationship with gamma $\widehat{Y}$ which indicates that there's a weak positive association between the 2 variables such that as satisfaction of financial situation decreases people tend to be more unhappy with their lives.

3-Feeling of Happiness & Frequency you/family (last 12 month): Gone without a cash income

Frequency you/family (last 12 month): Gone without a cash income	Feeling of Happiness		
	Happy	Quite Happy	Unhappy
Often	35	43	37
Sometimes	188	366	152
Never	253	333	47

Table 3: Feeling of happiness & Frequency you/family (last 12 month): Gone without a cash income.

Pearson Chi-square test	Uncertainty Coefficient($\widehat{U}$)
P-value = 0.000000	0.03093574

Table 3-1: Uncertainty coefficient & Pearson chi-square test statistic.

From the P-value(less than 0.05) we **reject** H_0 with 0.05 significance level, therefore, with 95% conf.level, we have enough evidence to say that there's an association between Feeling of happiness & frequency of Going with cash income, **hence** we can describe that relationship with $\widehat{U}$ which indicates that there's a weak relationship between the 2 variables where knowing the frequency of going without cash income reduces the entropy of feeling of happiness by almost 3%.

Note to mention that I used uncertainty coefficient since Y is ordinal **but** X has more than 2 categories (isn't binary) which means we can't use the gamma association measure.

4- Feeling of Happiness & Science and technology are making our lives healthier, easier, and more comfortable

Science and technology are making our lives healthier, easier, and more comfortable	Feeling of Happiness		
	Happy	Quite Happy	Unhappy
Agree	258	514	107
Disagree	218	228	129

Table 4: Feeling of happiness & Science make our lives healthier, easier and more comfortable.

Pearson Chi-square test	Gamma measure of ordinal association ($\widehat{Y}$)
`P-value = 0.000000`	-0.004280515

Table 4-1: Gamma ($\widehat{Y}$) & Pearson chi-square test statistic

From the P-value(less than 0.05) we **reject** H_0 with 0.05 significance level, therefore, with 95% conf.level, we have enough evidence to say that there's an association between Feeling of happiness & Science and technology are making our lives healthier, easier, and more comfortable, **hence** we can describe that relationship with $\widehat{Y}$ which indicates that there's a weak negative relationship between the 2 variables such that as people tend to disagree that science& technology make life easier, they tend to be happier.

5- Feeling of Happiness & Sex

Sex	Feeling of Happiness		
	Happy	Quite Happy	Unhappy
Male	211	350	123
Female	265	392	113

Table 5: Feeling of happiness & Sex.

Pearson Chi-square test
P-value = 0.1456

Table 5-1: Pearson Chi-square test statistic.

From the P-value (greater than 0.05) we **fail to reject** H_0 with 0.05 significance level, therefore, with 95% conf.level, we have enough evidence to say that **there's no** association between Feeling of happiness & sex, therefore, I won't calculate any association measure between the 2 variables since the p-value is much larger than the 0.05(α).

Note the Chi square approximation conditions in all tables were satisfied such that:
- large sample (n=1454).
- No more than 20% of the expected frequencies are less than 5.

3-Way Table:

I have decided to make a 3-way contingency table between Q46 , Q48 & Q50 since Both Q48 and Q50 had the highest association measure value($\widehat{Y}$) among the other variables.

Satisfaction with financial situation of household	How much freedom of choice and control	Feeling of Happiness		
		Happy	Quite Happy	Unhappy
Satisfied	Good Control	211	352	73
	Poor Control	95	109	33
Dissatisfied	Good Control	69	125	35
	Poor Control	101	156	95

Table 6: 3-way table for Feeling of happiness , Freedom of choice & Satisfaction with financial situation

Testing Mutual, joint & conditional independence :

Mutual independence:

H_0: $\pi_{ijk} = \pi_{i++} \ \pi_{+j+} \ \pi_{++k}$ For all i, j& k   (α = 0.05)

joint independence (Y is jointly independent of X &Z) :

H_0: $\pi_{ijk} = \pi_{i+k} \ \pi_{+j+}$ For all i, j& k   (α = 0.05)

Conditional Independence (Y is conditionally independent of X given Z):

H_0: $\pi_{ij|k}^{YX|Z} = \pi_{i|k}^{X|Z} \pi_{j|k}^{Y|Z}$ For all i, j& k   (α = 0.05)

Where i = 1 , 2 (Good control , Poor control)

j = 1 , 2 , 3 (happy , quite happy ,unhappy)

k = 1 , 2 (satisfied , dissatisfied)

I will use the Pearson chi-square test statistic (X^2) to test these hypotheses.

	Mutual independence	Joint independence	Conditional independence
Calculated value X^2	225.109	48.202	17.668
Tabulated value $\chi^2_{df}(0.05)$	$\chi^2_7(0.05) = 14.067$	$\chi^2_6(0.05) = 12.595$	$\chi^2_4(0.05) = 9.488$
Decision	**Reject** H_0	**Reject** H_0	**Reject** H_0

Table 6-1 : Independence testing & Decisions

For mutual independence:

With 0.05 significance level we **reject** H_0 therefore, with 95% conf.level, we have enough evidence to say that **there's no** mutual independence between the 3 variables.

For joint independence:

With 0.05 significance level we **reject** H_0 therefore, with 95% conf.level, we have enough evidence to say that **Feeling of happiness isn't jointly independent** of How much freedom of choice and control **&** Satisfaction with financial situation of household.

For conditional independence:

With 0.05 significance level we **reject** H_0 therefore, with 95% conf.level, we have enough evidence to say that there's no significant conditional independence between Feeling of happiness **&** How much freedom of choice and control **controlling for** Satisfaction with financial situation of household.

Note that I used chi-square tests instead of the Cochran mantel Haenszel test for conditional independence, **because** we have a 2 by 3 partial table which we can't get the odds ratio for just to make sure the odds ratios have the same sign in order to use Cochran mantel Haenszel test.