Midterm Project

Jiun Lee

1. Introduction

This report aims to use multilevel modeling and evaluate it. The dataset in this report is from European Social Survey(ESS), which is the research infrastructure consortium in Europe. The behavior and attitudes survey of people over 14 is conducted every two years with 38 countries' participation. In ESS Data Portal, you can download variable datasets provided by ESS. The dataset in this report is the latest dataset among them. The variables of the dataset consist of respondents' different attitudes such as nationality, media, politics, understanding of democracy, social relationships, values, and so on. This project is conducted to see the effect of different attributes on happiness and health within countries and between countries.

1. EDA Process

The variables 'cntry', 'happy', 'health', 'ipmodst', 'impsafe', 'ipfrule', and 'stfmjob' are chosen for analysis. NA values (Refusal, No answer, Don't know) are all deleted and the variables are sorted in ascending order during the cleaning process.

This is the description of each variable.
- cntry: respondents' countries
    o BG(Bulgaria), CZ(Czechia), EE(Estonia), FI(Finland), FR(France), HR(Croatia), HU(Hungary), LT(Lithuania), SI(Slovenia), SK(Slovakia))
- happy: How happy are you
    o 0(extremely unhappy) – 3
- health: How is your health in general?
    o 1(Very bad) – 5(Very good)
- ipmodst: important to be humble and modest, not draw attention
    o 1(not like me at all) - 6(Very much like me)
- impsafe: Important to live in secure and safe surroundings
    o 1(not like me at all) - 6(Very much like me)
- ipfrule: Important to do what is told and follow rules
    o 1(not like me at all) - 6(Very much like me)
- stfmjob: How satisfied are you in your main job

To approximately scan the different variables' effects on health and happiness, I plotted the association of different variables with health and health within each country.
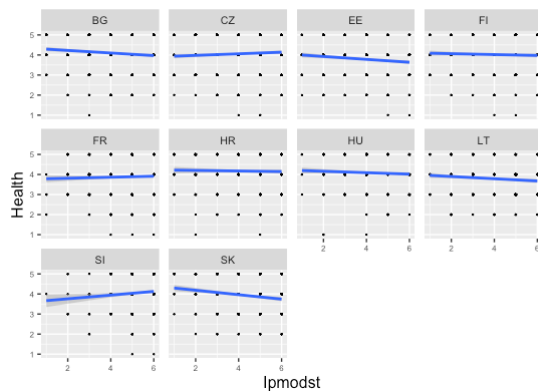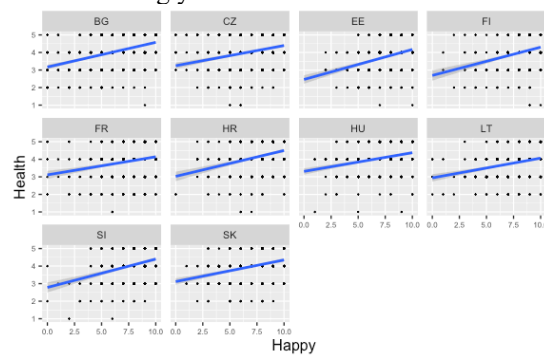

- SK(Slovakia) seems to have the biggest negative relationship of 'ipmodst' and 'health'.
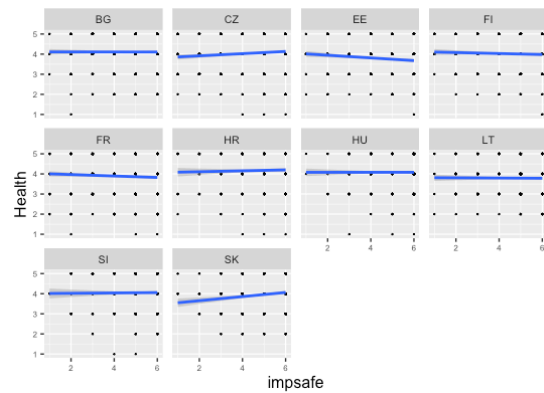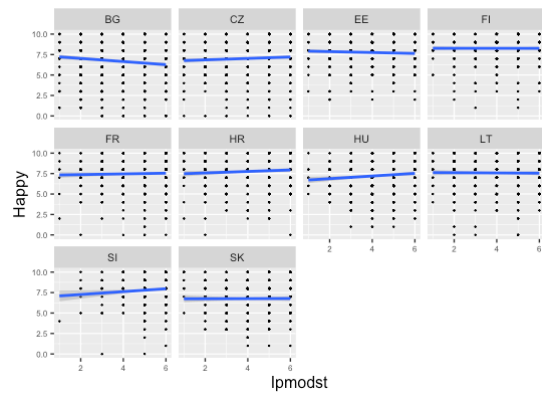
- SK(Slovakia) seems to have biggest positive relationship of 'impsafe' and 'health'.
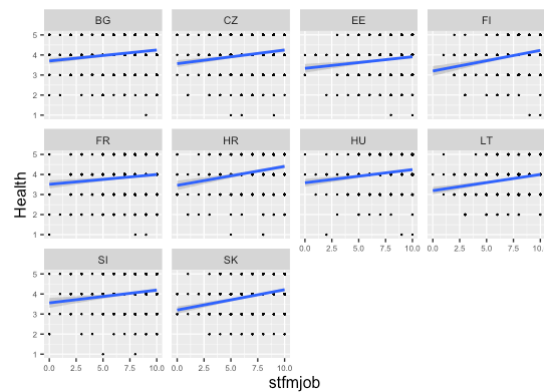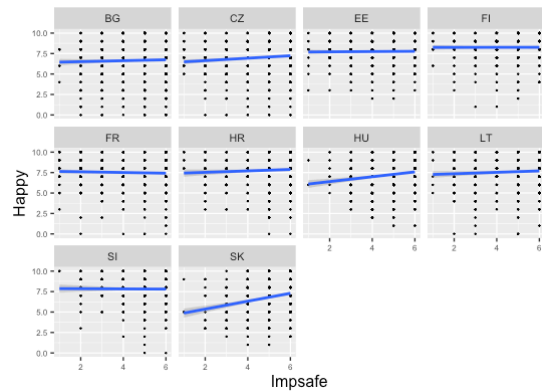

- All countries generally have a positive effect of 'stfmjob' on 'health'. The satisfaction of a job seems strongly associated with health.

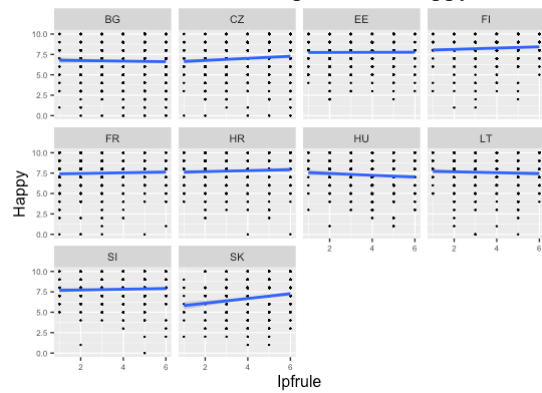- All countries generally have a significant positive effect of 'happy' on 'health'. Happiness seems strongly associated with health.
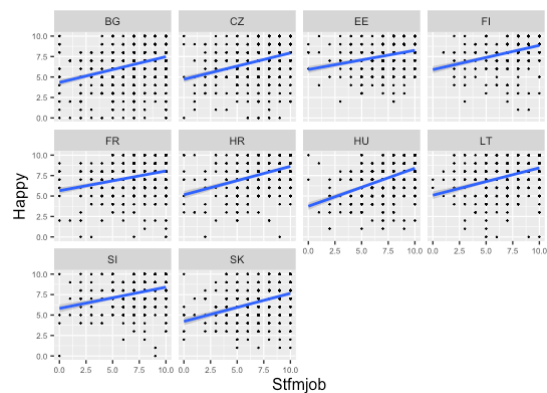
- The effects of 'ipmodst' on 'happy' in BG(Bulgaria) and SI(Slovenia) are different from other countries.
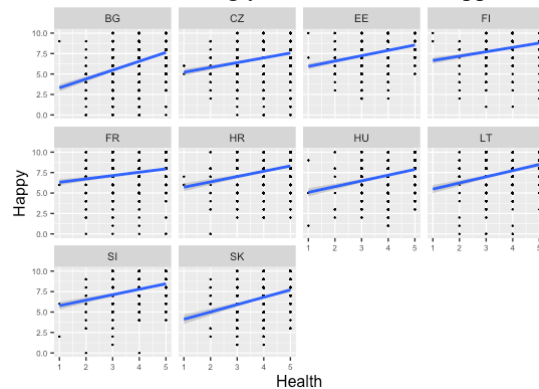

- The effects of 'impsafe' on 'happy' are significant in HU(Hungary) and SK(Slovakia).

- The effect of 'ipfrule' on happy is significant in SK(Slovakia).

- All countries generally have a significant positive effect of 'stfmjob' on 'happy'. The satisfaction of a job seems strongly associated with happiness.

- All countries generally have a significant positive effect of 'health' on 'happy'. The health seems strongly associated with happiness.

2. Multilevel analysis

From the EDA process, I verified that there were differences in the effects between countries. To see the random intercept and random slope, we can set the model like this.

Firstly, the response to health.

m1 <- stan_lmer(health ~ stfmjob + happy + impsafe + ipmodst + (1|cntry) +(0+stfmjob|cntry)+(0+happy|cntry)+(0+impsafe|cntry)+(0+ipmodst|cntry),data=data2)

```
> coef(m1)
$cntry
   (Intercept)    stfmjob     happy      impsafe      ipmodst
BG   3.167973 0.02645480 0.1285551  0.002903975 -0.034064466
CZ   2.952433 0.03350599 0.1058616  0.018257439  0.011704099
EE   2.937629 0.03156357 0.1342348 -0.046758835 -0.053307305
FI   2.887504 0.04064534 0.1212647 -0.018022409 -0.027247545
FR   2.974122 0.03012753 0.1015564 -0.026996011  0.010737282
HR   2.996171 0.04054102 0.1226569  0.004034424 -0.029540654
HU   3.217603 0.03365285 0.1103501 -0.006804696 -0.035895907
LT   2.977008 0.03814633 0.1040416 -0.012584009 -0.049282637
SI   2.872669 0.03172376 0.1251332 -0.008299792  0.001592458
SK   3.067234 0.04583732 0.1095207  0.024125994 -0.081533040
```

```
> fixef(m1)
 (Intercept)       stfmjob         happy        impsafe        ipmodst
 3.002588628   0.035054785   0.116445126  -0.007444271  -0.028290544
> ranef(m1)
$cntry
     (Intercept)       stfmjob         happy         impsafe        ipmodst
BG   0.165384291  -0.008599987   0.012109991   0.0103482456  -0.005773923
CZ  -0.050155210  -0.001548798  -0.010583502   0.0257017094   0.039994643
EE  -0.064959790  -0.003491213   0.017789717  -0.0393145649  -0.025016761
FI  -0.115084510   0.005590559   0.004819590  -0.0105781387   0.001042999
FR  -0.028466636  -0.004927250  -0.014888736  -0.0195517405   0.039027826
HR  -0.006417169   0.005486238   0.006211783   0.0114786944  -0.001250111
HU   0.215014074  -0.001401938  -0.006095018   0.0006395749  -0.007605363
LT  -0.025580713   0.003091549  -0.012403482  -0.0051397381  -0.020992093
SI  -0.129919750  -0.003331025   0.008688090  -0.0008555216   0.029883001
SK   0.064645200   0.010782535  -0.006924470   0.0315702645  -0.053242497

with conditional variances for "cntry"
> VarCorr(m1)
 Groups    Name        Std.Dev.
 cntry     (Intercept) 0.207496
 cntry     stfmjob     0.014376
 cntry     happy       0.020498
 cntry     impsafe     0.034110
 cntry     ipmodst     0.044090
 Residual              0.727437
```

- Interpretation of fixed effects (fixef(m1))
    o Intercept: The average value of 'health' for every country is 3
    o Stfmjob: The average effect of stfmjob on health for every country is 0.04
    o Happy: The average effect of happy on health for every country is 0.12
    o Impsafe: The average effect of impsafe on health for every country is -0.01
    o Ipmodst: The average effect of ipmodst on health for every country is -0.03.
- Interpretation of random effects (ranef(m1),VarCorr(m1))
    o Intercept: The difference between the average health of all countries and the average health of each country
    o Stfmjob: The difference between all countries' effect of stfmjob on health and each country's effect of stfmjob
    o Happy: The difference between all countries' effect of Happy on health and each country's effect of Happy
    o Impsafe: The difference between all countries' effect of Impsafe on health and each country's effect of Impsafe
    o Ipmodst: The difference between all countries' effect of Ipmodst on health and each country's effect of Ipmodst
    o Std.Dev.(Intercept) : There are solid between-country differences in average value of health.(0.20)
    o Std.Dev(stfmjob,happy,impsafe,ipmodst): There are solid between-country differences in association between 'stfmjob,happy,impsafe,ipmodst' and health.
- Interpretation of coefficients (coef(m1))
    o Stfmjob: slopes are generally similar between countries.
    o Happy: slopes are generally similar between countries.
    o Impsafe: EE has the biggest negative impact and SK has the biggest positive impact on health among countries.
    o Ipmodst: SK has the biggest negative impact on health among countries.

To check the model's plausibility, we can use ICC(Intraclass correlation coefficient). If the ICC of each variable is over 0.1, we can consider the use of the multilevel model.

```
> RandomEffects <- as.data.frame(VarCorr(m1))
> RandomEffects
        grp           var1 var2          vcov        sdcor
1     cntry (Intercept) <NA> 0.0430544200 0.20749559
2     cntry      stfmjob <NA> 0.0002066733 0.01437614
3     cntry        happy <NA> 0.0004201696 0.02049804
4     cntry      impsafe <NA> 0.0011635018 0.03411014
5     cntry      ipmodst <NA> 0.0019439024 0.04408971
6  Residual         <NA> <NA> 0.5291644363 0.72743690
> ICC_int <- RandomEffects[1,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomE
ffects[3,4]+RandomEffects[4,4]+RandomEffects[5,4]+RandomEffects[6,4]
> ICC_int #1.531522
[1] 1.532899
> ICC_stfmjob<- RandomEffects[2,4]/RandomEffects[1,4]+RandomEffects[2,4]+Rand
omEffects[3,4]+RandomEffects[4,4]+RandomEffects[5,4]+RandomEffects[6,4]
> ICC_stfmjob #0.5361391
[1] 0.537699
> ICC_happy   <- RandomEffects[3,4]/RandomEffects[1,4]+RandomEffects[2,4]+Rand
omEffects[3,4]+RandomEffects[4,4]+RandomEffects[5,4]+RandomEffects[6,4]
> ICC_happy #0.5412589
[1] 0.5426577
> ICC_impsafe  <- RandomEffects[4,4]/RandomEffects[1,4]+RandomEffects[2,4]+Ra
ndomEffects[3,4]+RandomEffects[4,4]+RandomEffects[5,4]+RandomEffects[6,4]
> ICC_impsafe #0.5625203
[1] 0.5599227
> ICC_ipmodst  <- RandomEffects[5,4]/RandomEffects[1,4]+RandomEffects[2,4]+Ra
ndomEffects[3,4]+RandomEffects[4,4]+RandomEffects[5,4]+RandomEffects[6,4]
> ICC_ipmodst  #0.5818328
[1] 0.5780486
```
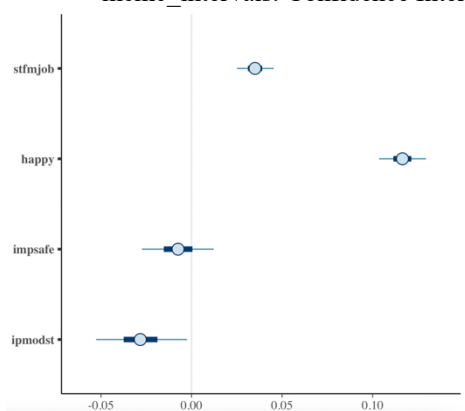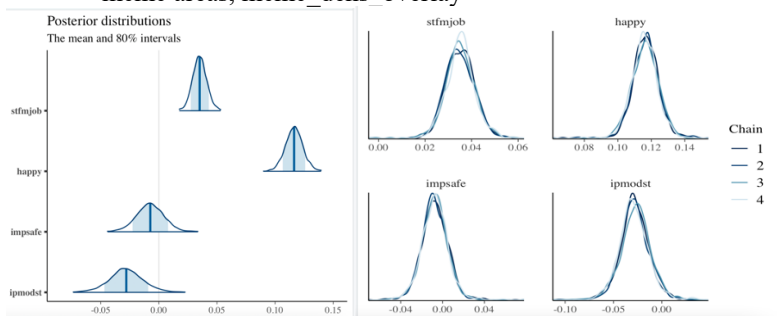
The ICC of each variable are all over 0.1. The multilevel model 'm1' is plausible.

Also, we can use MCMC(Markov Chain Monte Carlo) for testing single parameters.
- mcmc_intervals: Confidence Interval of each parameter

- mcmc areas, mcmc_dens_overlay

Now, try the response to happiness.

```
> coef(m2)
$cntry
    (Intercept)    stfmjob     health     impsafe      ipmodst
BG    1.431195  0.2646585  0.8772629  0.057842816 -0.112365386
CZ    2.537934  0.2807243  0.4763525  0.063351821  0.019013124
EE    3.589537  0.2194887  0.5962911  0.047363713 -0.007051088
FI    3.940069  0.2641114  0.4621626  0.037930676 -0.011601962
FR    4.057467  0.2356893  0.3875400  0.005422104  0.008677632
HR    3.012651  0.2907083  0.5048596  0.048180040  0.048096819
HU    1.865173  0.3760426  0.5402039  0.099870739  0.054049240
LT    3.086490  0.2823545  0.5664392  0.051028389  0.004412759
SI    3.292259  0.2368370  0.5795603  0.022326920  0.026148597
SK    1.409376  0.2558089  0.5821326  0.129708008  0.011250061
        ipfrule
BG -0.037506971
CZ  0.043747143
EE  0.006019139
FI  0.066849069
FR  0.028563678
HR  0.008464631
HU -0.100762307
LT -0.001698429
SI  0.029676490
SK  0.103038779

> fixef(m2)
 (Intercept)       stfmjob        health       impsafe       ipmodst       ipfrule
2.820555438   0.270514942   0.557875254   0.054402397   0.004274779   0.014588640


> ranef(m2)
$cntry
    (Intercept)      stfmjob       health       impsafe      ipmodst
BG  -1.3893600 -0.005856478  0.319387679  0.003440419 -0.1166401646
CZ  -0.2826211  0.010209368 -0.081522747  0.008949425  0.0147383457
EE   0.7689813 -0.051026215  0.038415884 -0.007038684 -0.0113258663
FI   1.1195134 -0.006403521 -0.095712700 -0.016471720 -0.0158767407
FR   1.2369113 -0.034825619 -0.170335209 -0.048980293  0.0044028537
HR   0.1920953  0.020193369 -0.053015635 -0.006222357  0.0438220408
HU  -0.9553820  0.105527627 -0.017671395  0.045468342  0.0497744614
LT   0.2659346  0.011839582  0.008563975 -0.003374008  0.0001379804
SI   0.4717031 -0.033677905  0.021685080 -0.032075477  0.0218738183
SK  -1.4111798 -0.014706008  0.024257323  0.075305611  0.0069752824
        ipfrule
BG -0.052095611
CZ  0.029158503
EE -0.008569501
FI  0.052260428
FR  0.013975038
HR -0.006124010
HU -0.115350947
LT -0.016287070
SI  0.015087850
SK  0.088450139

> VarCorr(m2)
 Groups    Name        Std.Dev.
 cntry     (Intercept) 1.204992
 cntry     stfmjob     0.062345
 cntry     health      0.179067
 cntry     impsafe     0.070384
 cntry     ipmodst     0.075519
 cntry     ipfrule     0.084030
 Residual              1.590703
```

- Interpretation of fixed effects (fixef(m2))
  o Intercept: The average value of happiness for every country is 2.82
  o Stfmjob: The average effect of stfmjob on happiness for every country is 0.27

- o   health: The average effect of health on happiness for every country is 0.56
- o   Impsafe: The average effect of impsafe on happiness for every country is 0.05
- o   Ipmodst: The average effect of ipmodst on happiness for every country is 0.004
- o   ipfrule: The average effect of ipfrule on happiness for every country is 0.01
  - Interpretation of random effects (ranef(m2),VarCorr(m2))
    - o   Intercept: The difference between the average happiness of all countries and the average happiness of each country
    - o   Stfmjob: The difference between all countries' effect of stfmjob on happiness and each country's effect of stfmjob
    - o   health: The difference between all countries' effect of health on happiness and each country's effect of health
    - o   Impsafe: The difference between all countries' effect of Impsafe on happiness and each country's effect of Impsafe
    - o   Ipmodst: The difference between all countries' effect of Ipmodst on happiness and each country's effect of Ipmodst
    - o   Ipfrule: The difference between all countries' effect of Ipmodst on happiness and each country's effect of ipfrule
    - o   Std.Dev.(Intercept) : There are solid between-country differences in average value of happiness.(1.20)
    - o   Std.Dev(stfmjob,happy,impsafe,ipmodst,ipfrule): There are solid between-country differences in the association between 'stfmjob, health, impsafe, ipmodst' and happiness.
  - Interpretation of coefficients (coef(m2))
    - o   Intercept: FR has the biggest average happiness among countries.
    - o   Stfmjob: slopes are generally similar between countries.
    - o   health: slopes are generally similar between countries.
    - o   Impsafe: SK has the biggest positive impact on happiness among countries.
    - o   Ipmodst: BG has the biggest negative impact on happiness among countries.
    - o   Ipfrule: HU has the biggest negative impact and SK has the biggest positive impact on happiness among countries.

  - ICC

```
> #ICC
> VarCorr(m2)
 Groups    Name        Std.Dev.
 cntry     (Intercept) 1.204992
 cntry     stfmjob     0.062345
 cntry     health      0.179067
 cntry     impsafe     0.070384
 cntry     ipmodst     0.075519
 cntry     ipfrule     0.084030
 Residual              1.590703
> RandomEffects <- as.data.frame(VarCorr(m2))
> RandomEffects
      grp          var1 var2        vcov      sdcor
1   cntry (Intercept) <NA> 1.452004878 1.20499165
2   cntry      stfmjob <NA> 0.003886904 0.06234504
3   cntry       health <NA> 0.032064820 0.17906652
4   cntry      impsafe <NA> 0.004953882 0.07038382
5   cntry      ipmodst <NA> 0.005703108 0.07551893
6   cntry      ipfrule <NA> 0.007061108 0.08403040
7 Residual         <NA> <NA> 2.530334451 1.59070250
> ICC_int <- RandomEffects[1,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffects[4,
4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_int #3.564563
[1] 3.584004
> ICC_stfmjob<- RandomEffects[2,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffects
[4,4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_stfmjob #2.566859
[1] 2.586681
> ICC_health  <- RandomEffects[3,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffects
[4,4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_health #2.583978
[1] 2.600087
> ICC_impsafe  <- RandomEffects[4,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffect
s[4,4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_impsafe #2.56716
[1] 2.587416
> ICC_ipmodst  <- RandomEffects[5,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffect
s[4,4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_ipmodst  #2.567771
[1] 2.587932
> ICC_ipfrule  <- RandomEffects[6,4]/RandomEffects[1,4]+RandomEffects[2,4]+RandomEffects[3,4]+RandomEffect
s[4,4]+RandomEffects[5,4]+RandomEffects[6,4]+RandomEffects[7,4]
> ICC_ipfrule #2.568804
[1] 2.588867
```
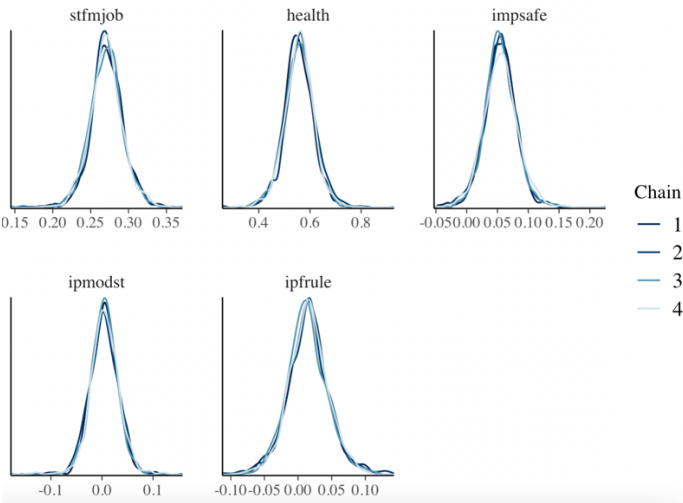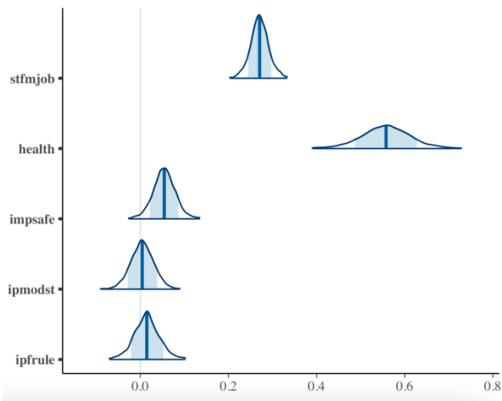
The ICC of each variables are all over 0.1. The multilevel model 'm2' is plausible.

Also, we can use MCMC(Markov Chain Monte Carlo) for testing single parameters.
- mcmc_intervals: Confidence Interval of each parameter
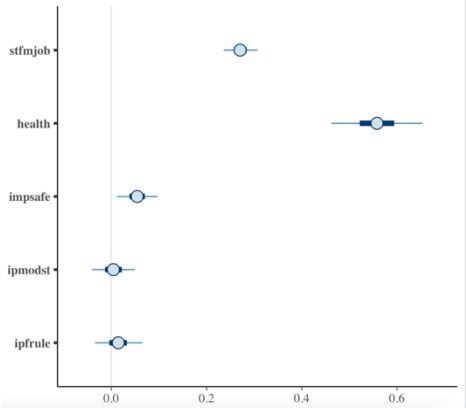
- mcmc areas, mcmc_dens_overlay

Posterior distributions
The mean and 80% intervals

3. Result

In model "m1", we could find out that satisfaction from a job and happiness are significantly related to health in every country with a similar amount of effect. The countries' average health is similar as we can see from each intercept. However, it seemed that people who think it's important to live in secure and safe surroundings are less healthy in Estonia but are healthier in Slovakia. Also, people who think it's important to be humble and modest tend to live less healthy in Slovakia

In model "m2", we could also find out that satisfaction from a job and health are significantly related to happiness in every country with a similar amount of effect. But the average happiness in France is the biggest among other countries. Also, people in Slovakia who think it's important to live in secure and safe surroundings are happier. In Bulgaria, people who find it important to be humble are less happy. People who think it's important to follow rules are less happy in Hungary but are happier in Slovakia.

In conclusion, this report has handled the process of multilevel regression. We confirm that there are differences among countries with multiple parameters.