Predicting Movie Preferences from Personality Traits and Interests

Advanced Design of Experiments - MATH 567, December 1, 2019

Project Group/Authors:

Shikha Saxena (A20424636 - ssaxena13@hawk.iit.edu) - 50% Contribution
Saransh Kumar (A20424637 - skumar58@hawk.iit.edu) - 50% Contribution
Code: https://github.com/Saransh-git/movie-preference-analysis/blob/master/movie_pref.R

Abstract

In this project, we tried to create experimental design such that we find factors of personality traits, interests, phobia and music preferences that influence the choice of horror movie/serial genre among people.

The survey data used in the report contained a wide range of questions, hence we first performed factor screening and then carried out various data analysis such as residual analysis, interaction effect exploration. This helped us to achieve useful results of common personality traits and interests which overlap with liking of Horror genre.

All the steps performed in this project can be implemented for extracting similar information about liking of other movie genres such as Sci-fi, Action, War, Comedy, etc.

Introduction/ Recognition and Statement of the problem

Today, with increasing trend and popularity in online streaming services such as Netflix, Disney Plus, Amazon Prime , etc. everyone can choose and watch movies and serials of their own preferred genre.
One of the movie or serial genre is Horror, but we all know it is not everyone's favourite. When we are sitting in a group of friends there might be always someone who hate horror movies and other who might have finished watching all the released ones. This contrast liking of horror makes us sometimes wonder do people of certain interest hate horror movies. Hence, we thought to explore personality traits, phobias, music preferences and hobbies and Interests to find explainability towards liking of horror movies.

The objectives to our experimental design are therefore three folds:

- **Factor screening**: Identify the factors which most influence the affinity towards watching Horrors.
- **Optimization**: If we are hosting a movie party and we are planning on screening a Horror, then we need to be sure about who to invite and who not to invite so as to not make anyone uncomfortable in the party.
- **Generalization**: The same experimental design can be extended over to learn the affinity towards other genres such as Comedies, Sci-fis, Actions, Thrillers etc.

Design of Experiment

Selection of the response variable

Given we are targeting the affinity towards Horrors, the response variable was chosen on a 5 way likert scale ranging from Strongly Disagree (1), Disagree (2), Neutral (3), Agree(4), Strongly Agree (5). Respondents were asked whether they enjoy watching Horror movies or not. The choice of a 5-way likert scale was to ease the collection of responses from respondents, as it becomes subjective and difficult if someone has to score their affinity on a continuous scale.

Choice of factors, levels and range

Design factors

All the variables in the data contain responses as likert scale as shown in below screenshot of data -

```
> data_horror
# A tibble: 940 x 22
   like_music slow_fast_music watch_movies horror psychology politics physics sci_and_tech
        <int>           <int>        <int>  <int>      <int>    <int>   <int>        <int>
 1          5               3            5      4          5        1       3            4
 2          4               4            5      2          3        4       2            3
 3          5               5            5      3          2        1       2            2
 4          5               3            5      4          4        5       1            3
 5          5               3            5      4          2        3       2            3
 6          5               3            5      5          3        4       3            3
 7          5               5            4      2          3        1       1            4
 8          5               3            5      4          2        3       1            2
 9          5               3            5      1          2        1       1            1
10          5               3            5      2          2        3       1            3
# ... with 930 more rows, and 14 more variables: theatre <int>, history <int>,
#   fun_with_friends <int>, darkness <int>, snakes <int>, spiders <int>, rats <int>,
#   work_in_spare_time <int>, funny <int>, happy_life <int>, internet_usage <fct>,
#   finances <int>, entertainment_spend <int>, age <int>
```

As we are screening our factors, we plotted response distributions for each factors to segregate the 5-scale likert response to be considered as low level and the ones to be considered in high level. The plots are included in Factor level plots. We tried to balance the distribution of the responses across the lower and the higher level so as to not skew our analysis.

The following **20 design factors** are considered as experimental factors and corresponding responses as **low levels versus high levels** are included against them:

- **Likes Music** - This tells how much participant enjoys listening to music with 1 being enjoys least. 1,2,3,4 versus 5 ,
- **Slow songs or fast songs**- This tells whether the person prefer slow paced music(1) or fast paced music(5) - 1,2,3 versus 4,5
- **Horror** - This is our response variable with 1 means least preference to watch horror movies - 1,2 versus 3,4,5, (Response Variable)

- **Psychology** - This tells us the interest of participant in Psychology with 1 being least interested - 1,2,3 versus 4,5 ,
- **Politics** - This tells us the interest of participant in Politics with 1 being least interested 1,2 versus 3,4,5,
- **Physics** - This tells us the interest of participant in Physics with 1 being least interested 1 versus 2,3,4,5,
- **Science and Technology** - This tells us the interest of participant in Science and Technology with 1 being least interested 1,2,3 versus 4,5,
- **Theatre** - This tells us the interest of participant in Theatre with 1 being least interested. 1,2,3 versus 4,5,
- **History** - This tells us the interest of participant in History with 1 being least interested . 1,2,3 versus 4,5,
- **Fun with friends** - This tells us how much participant enjoys socializing with 1 being enjoys least. 1,2,3, 4 versus 5,
- **Darkness** - Participant has darkness phobia with 5 as maximum phobia.1,2 versus 3,4,5,
- **Snakes** - Participant has phobia of snakes with 5 as maximum phobia. 1,2 versus 3,4,5,
- **Spiders** - Participant has phobia of spiders with 5 as maximum phobia.1,2 versus 3,4,5,
- **Rats** - Participant has phobia of rats with 5 as maximum phobia.1,2 versus 3,4,5,
- **Workaholism** - Tells how much participants like to work in spare time with 5 being love to do it. 1,2 versus 3, 4,5,
- **Funniness** - Tells how much participants tries to be the funniest one, with 1 being least funny. 1,2,3 versus 4,5,
- **Happiness in life** - Tells how much happy participant is with life with 5 being 100% happy.1,2,3 versus 4,5,
- **Internet usage** - This tells how often participant uses internet and this variable has the following options - no time at all, less than an hour a day, a few hours a day and most of the day. few hrs versus others,
- **Finances** - tells how much participant saves money with 5 meaning saves as much as he/she can. 1,2,3 versus 4,5,
- **Entertainment spending** - This variable shows how much money is spent on partying and socializing with 5 being spends a lot. 1,2,3 versus 4,5,
- **Age** - This tells about the age of participants. Above 20 versus Below 20

Allowed to vary factors

Alongside the design factors considered for the experiment, data was also collected for several other traits demonstrating Music preferences, Hobbies, Interests, Health habits, views on life, opinions and spending habits. As we were focused on what could possibly impact the affinity towards watching Horrors, we considered the design factors as mentioned above only for the analysis and left the others to vary across the respondents (treatment units).

Nuisance factors

While we are interested in finding the most influencing factors, there can be variations in our analysis based on whether respondents like to watch movies overall or not. A respondent who is overall less likely to watch any sort of movie would also refrain from watching Horror.

As a result, respondents were asked if they enjoy **watching movies** on a scale of **1 (Strongly Disagree) to 5 (Strongly Agree)**. The same has been treated as **blocks** on our analysis.

Uncontrollable factors not accounted

Other traits as we know there are many besides what we have considered in our experimental design. Also, that bringing in factors based on regions, genders etc. brings in subjectivity to the problem and such kind of factors are not accounted and deemed uncontrollable.

Choice of experimental design

Given the first objective to our problem of screening influential factors, we restrict the levels/ treatments of factors to two levels each as described in Design factors. Also, given the presence of blocked variable in the form of watch movies or not, we have a blocking variable. Therefore, we initially deploy a fixed effects model to capture the main and interaction effects with a randomized block design. We are only considering the second order model and ignoring all the interaction effects involving more than two variables. This choice was made due to the presence of a lot of variables (~ 20 design factors), considering all the possible effects would have taken an enormous amount of time to execute. Please refer to Appendix to see the full model considered.

Data

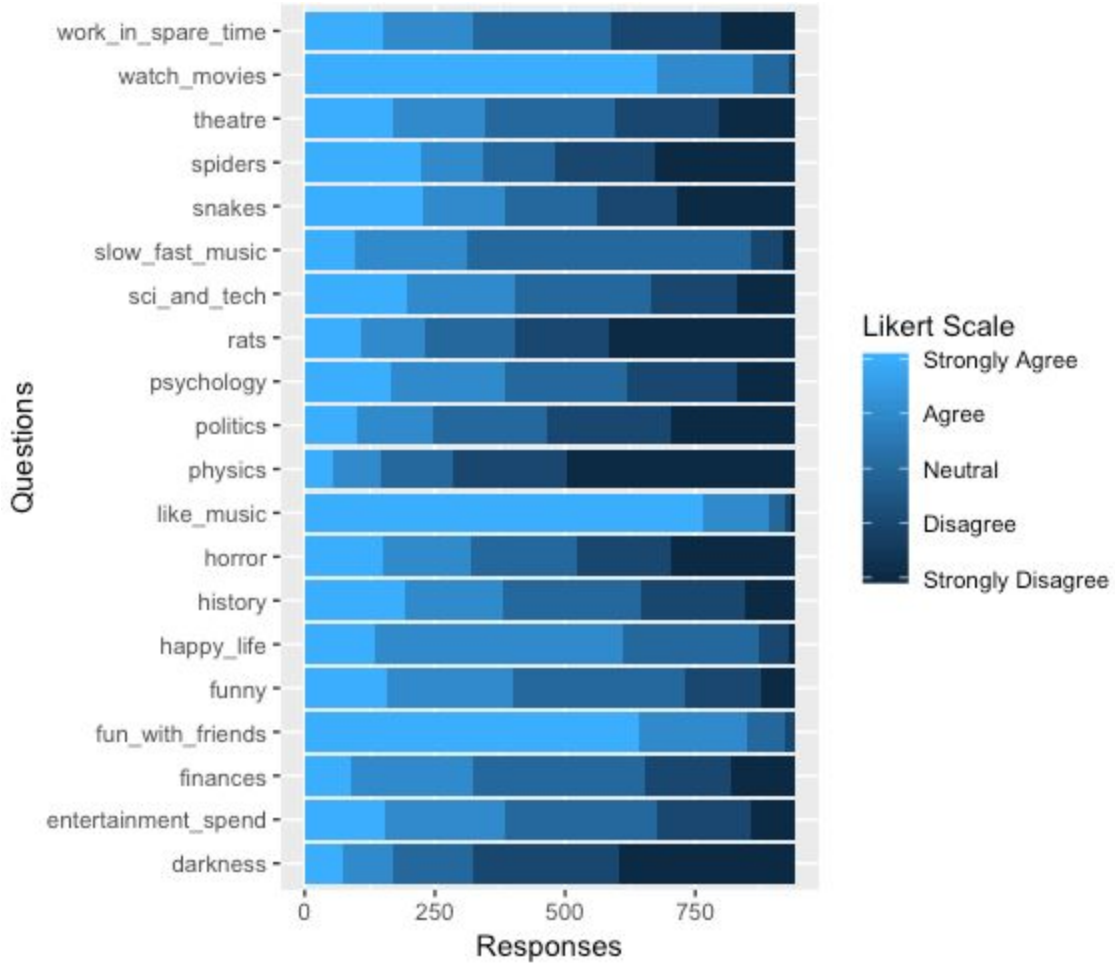
We had such a data available from a public survey where the 1,010 respondents in the age group of 15-30 were asked to have a survey on several items covering the following traits:

- Music Preferences (19 traits)
- Movie Preferences (12 traits)
- Hobbies & Interests (32 traits)
- Phobias (10 traits)
- Health Habits (3 traits)
- Personality traits, views on life, and opinions (57 traits)

- Spending habits (7 traits)
- Demographics (10 traits)

As mentioned in Choice of factors, levels and range, we selected the 20 design factors and 1 blocking variable as per the consideration in our current experimental design.

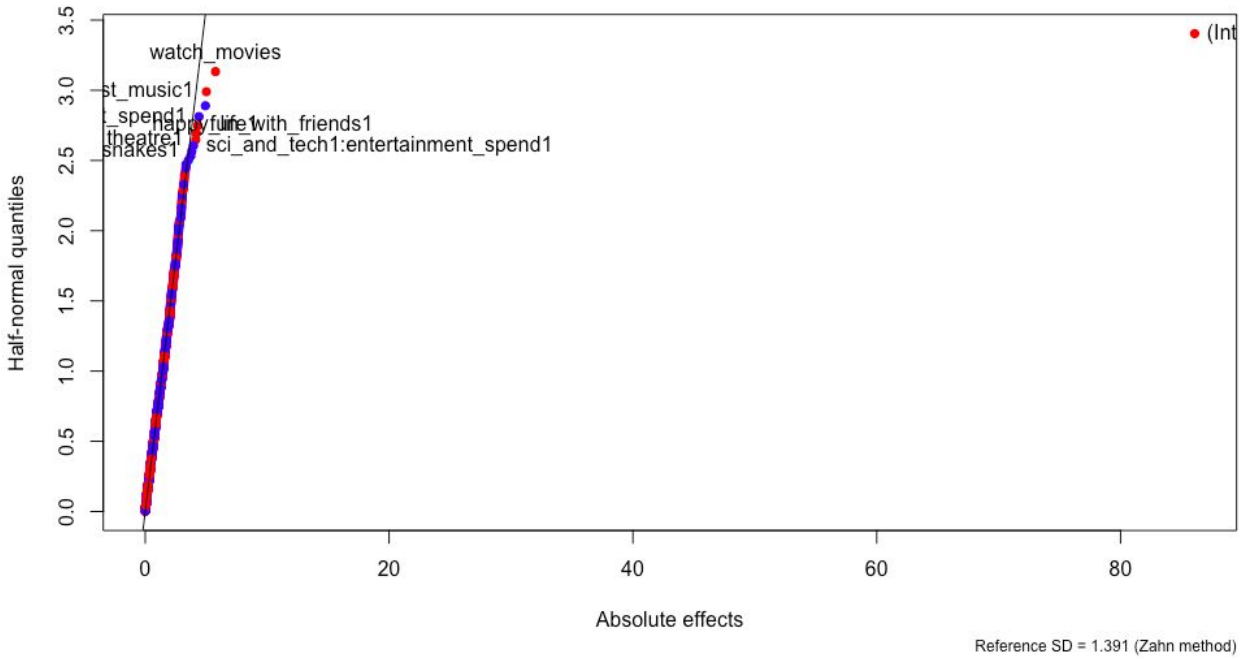
Following are the results as obtained from survey against the 20 design factors on a scale of 1 (Strongly Disagree) to 5 (Strongly Agree).



Analysis and Results

Models

We ran the full model as described in Full Model. Below is the plot as obtained with Half-normal probability plot on effects obtained from the Full Model.



Alongside, the half-normal probability plot, we also performed ANOVA analysis on the effects obtained from the full model. The ANOVA output for full model has been omitted from the report due to its length. Following factors were found significant at 95% significance level.

- Slow_fast_music
- Theatre
- Fun_with_friends
- Snakes
- Happy_life
- Entertainment_spend
- Age
- Watch_movies (Block effect)
- Like_music:history

- Slow_fast_music:psychology
- Slow_fast_music:entertainment_spend
- Politics:age
- Sci_and_tech:finances
- Sci_and_tech:entertainment_spend
- History:rats
- Darkness:snakes
- darkness:spiders

We fit the reduced model as obtained from the above analysis (described at Reduced Model 1). We further performed ANOVA analysis to find the significant factors at 95% significance and reduced our model further to Reduced Model 2.

```
> anova(fit_sig)
Analysis of Variance Table

Response: horror
                                    Df  Sum Sq Mean Sq F value    Pr(>F)
slow_fast_music                      1   24.92 24.9231 13.9528 0.0001991 ***
theatre                              1   15.50 15.4954  8.6748 0.0033085 **
fun_with_friends                     1   18.65 18.6450 10.4381 0.0012787 **
snakes                               1   17.32 17.3182  9.6953 0.0019052 **
happy_life                           1   24.16 24.1559 13.5233 0.0002494 ***
entertainment_spend                  1   22.75 22.7518 12.7372 0.0003772 ***
age                                  1    8.51  8.5121  4.7654 0.0292928 *
watch_movies                         4   35.47  8.8677  4.9644 0.0005813 ***
like_music                           1    3.19  3.1858  1.7835 0.1820494
history                              1    0.17  0.1707  0.0956 0.7572901
politics                             1    0.55  0.5489  0.3073 0.5794956
sci_and_tech                         1    0.43  0.4269  0.2390 0.6250463
finances                             1    0.25  0.2483  0.1390 0.7093615
rats                                 1    1.39  1.3901  0.7782 0.3779112
darkness                             1    1.61  1.6057  0.8989 0.3433274
spiders                              1    0.05  0.0539  0.0301 0.8621931
like_music:history                   1    4.80  4.7998  2.6871 0.1015101
slow_fast_music:psychology           2   10.55  5.2770  2.9542 0.0526192 .
slow_fast_music:entertainment_spend  1    5.57  5.5676  3.1169 0.0778181 .
age:politics                         1    6.21  6.2135  3.4785 0.0624928 .
sci_and_tech:finances                1    5.69  5.6894  3.1851 0.0746443 .
entertainment_spend:sci_and_tech     1   10.07 10.0710  5.6381 0.0177809 *
history:rats                         1    4.97  4.9671  2.7808 0.0957469 .
snakes:darkness                      1    6.00  6.0037  3.3611 0.0670807 .
darkness:spiders                     1    9.89  9.8930  5.5384 0.0188154 *
Residuals                          910 1625.49  1.7862
---
```

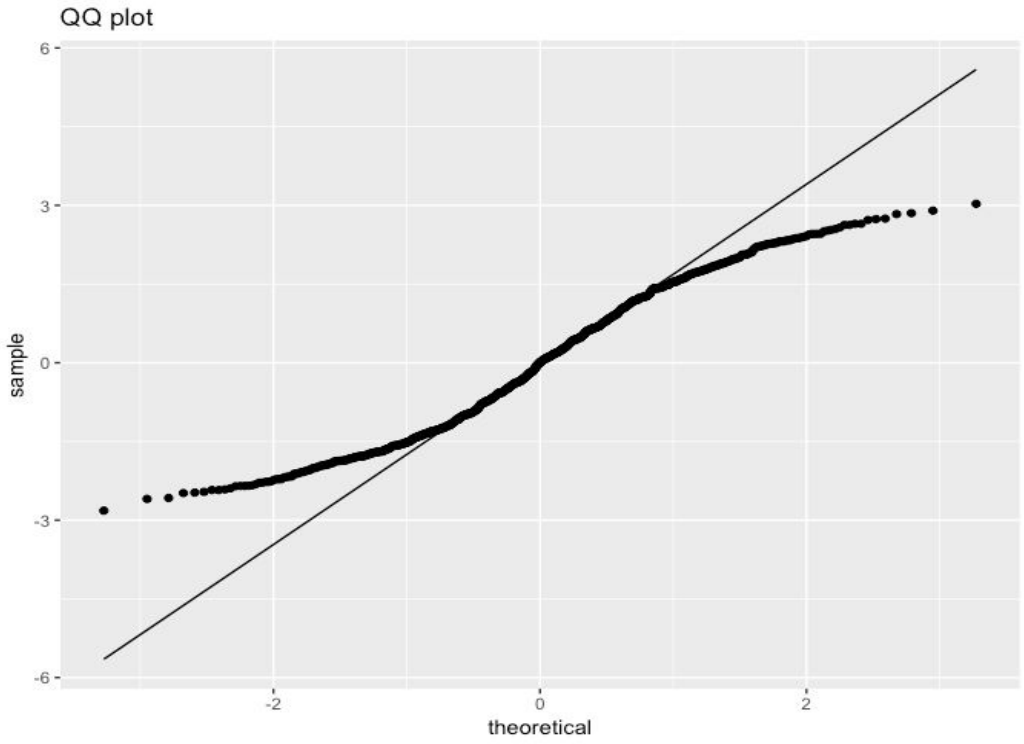
ANOVA analysis for Reduced Model 2:

```
> anova(fit_reduce)
Analysis of Variance Table

Response: horror
                                Df  Sum Sq Mean Sq F value    Pr(>F)
slow_fast_music                  1   24.92 24.9231 13.7207 0.0002247 ***
theatre                          1   15.50 15.4954  8.5305 0.0035774 **
fun_with_friends                 1   18.65 18.6450 10.2645 0.0014027 **
snakes                           1   17.32 17.3182  9.5340 0.0020773 **
happy_life                       1   24.16 24.1559 13.2983 0.0002805 ***
entertainment_spend              1   22.75 22.7518 12.5253 0.0004215 ***
age                              1    8.51  8.5121  4.6861 0.0306633 *
watch_movies                     4   35.47  8.8677  4.8818 0.0006724 ***
sci_and_tech                     1    0.53  0.5301  0.2918 0.5891765
darkness                         1    1.35  1.3544  0.7456 0.3880848
spiders                          1    0.13  0.1292  0.0711 0.7897583
entertainment_spend:sci_and_tech 1    6.79  6.7869  3.7363 0.0535469 .
darkness:spiders                 1   11.48 11.4786  6.3192 0.0121131 *
Residuals                      923 1676.60  1.8165
```

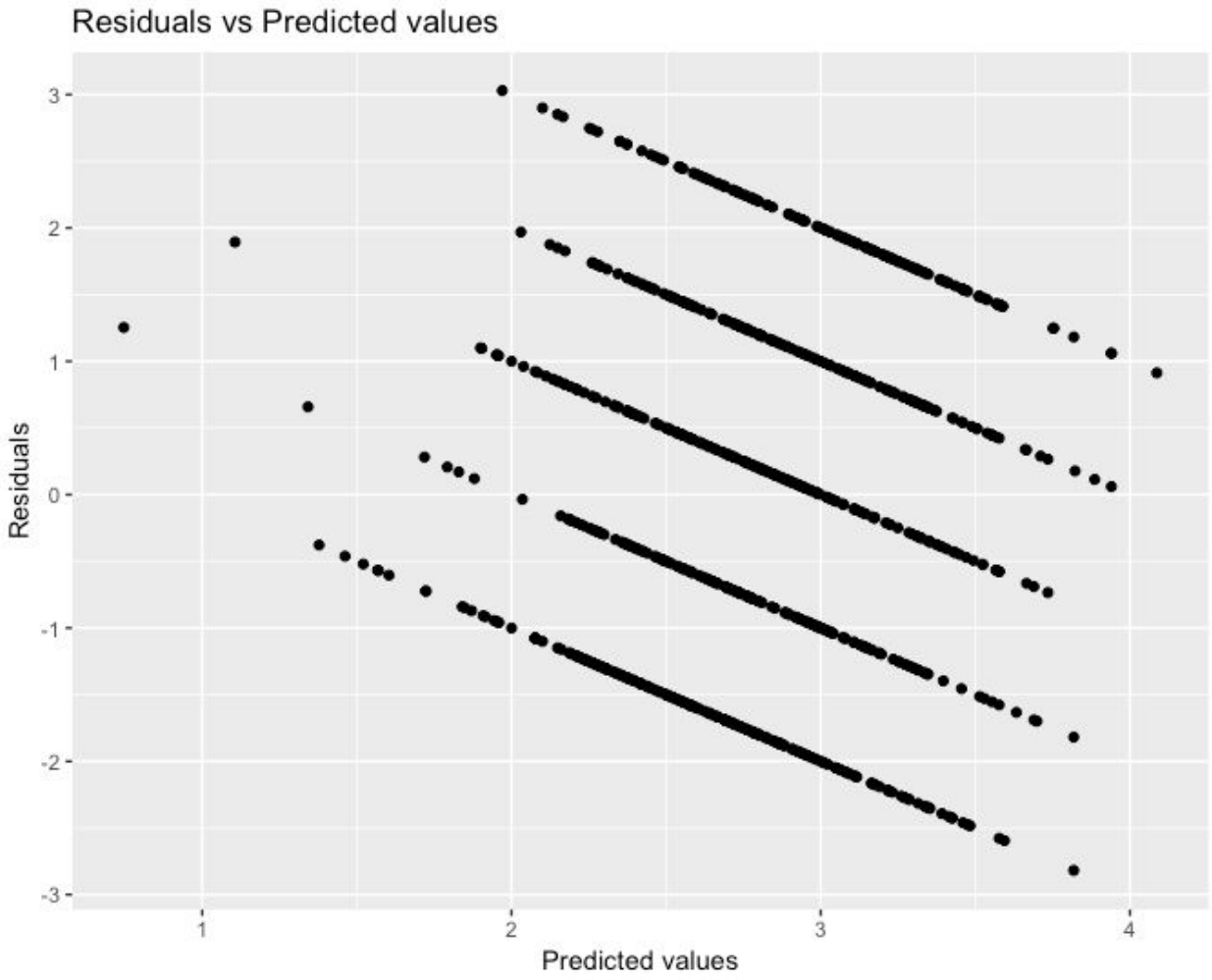
Residual Analysis

Using the Reduced Model 2, we performed a residual **analysis** to validate our model assumptions.



In the above Quantile-Quantile plot, we notice that most of the residuals within the interquartile range lie on the straight line with deviations along the tails which tells that the normality assumption is correct for the new fit with only significant factors.

We further analyzed the residuals against the predicted values to check our assumptions about the constant variance, curvatures, missed interaction effects. Below we plot the Residuals vs Predicted values:

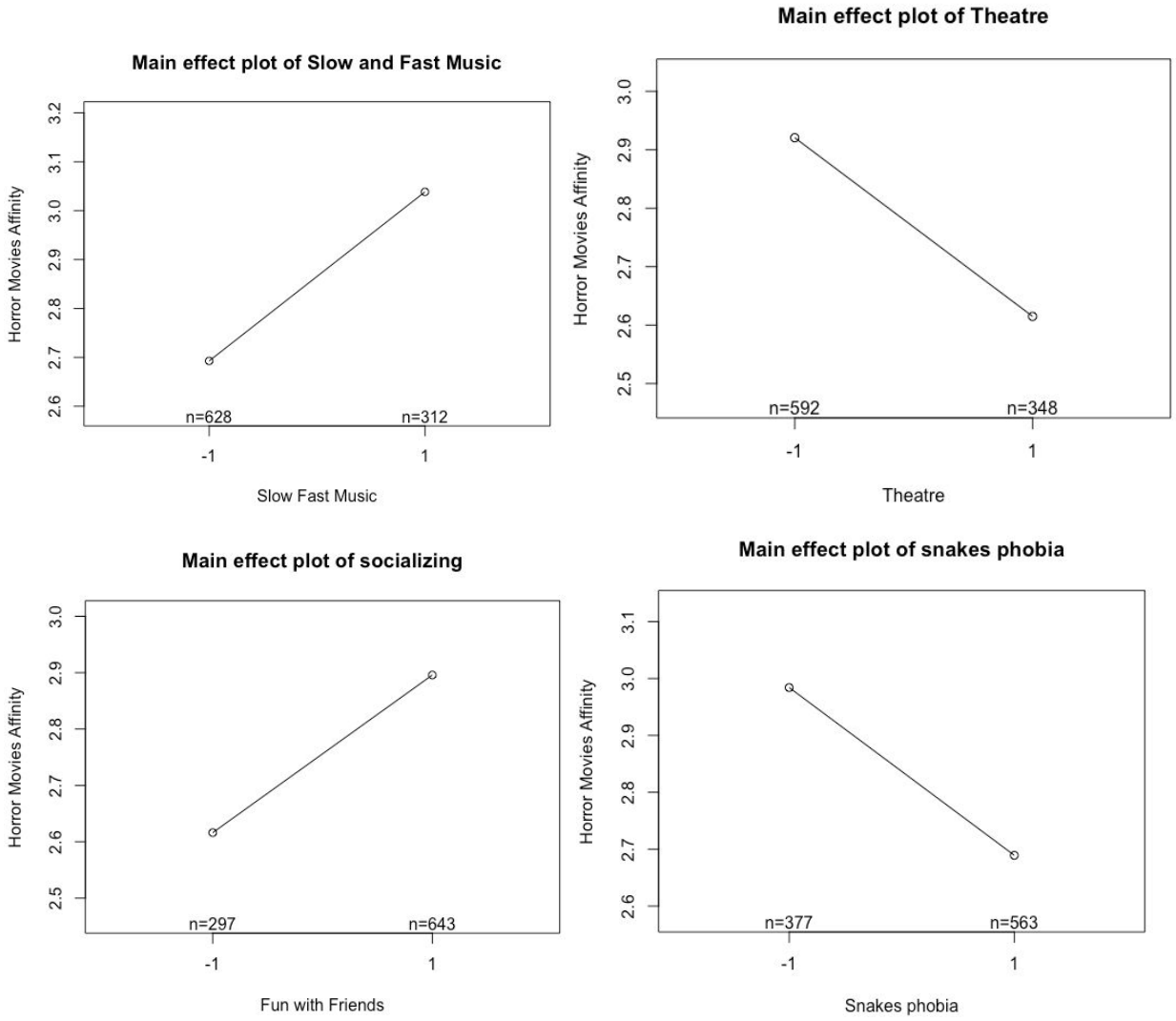


From the above plot, we observe parallel trends amongst residuals to predicted values. We were initially suspicious of these results.

We further plotted residuals against factors (included here: Residual to factor plots) to further inspect our model. From the plots of residuals to factor, we observed that residuals were all randomly scattered across the levels and that there was no curvature found with respect to any of the factors. Hence, this suggested to us that no transformation could help improve the model.

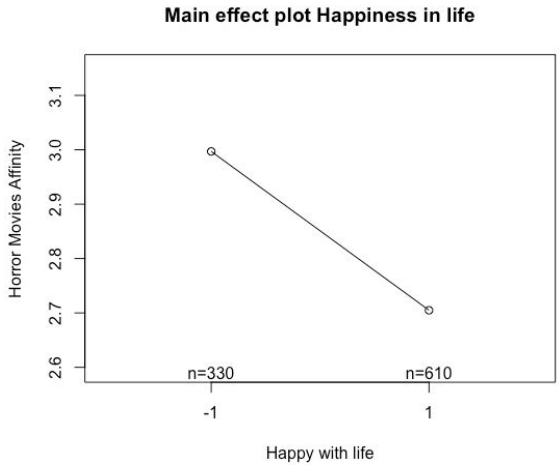
On further analysis, we learnt that the trend is evident as our response variable is ordinal and due to that the predicted values contain few possible values. This calls for the use of Ordinal regression (Generalized linear models) which has been included in Future Work as well.

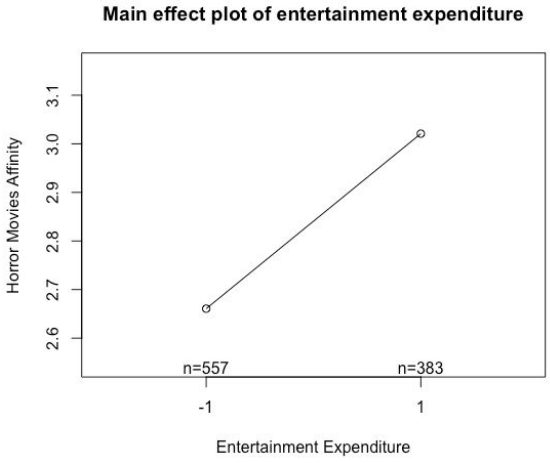
Effect Plots

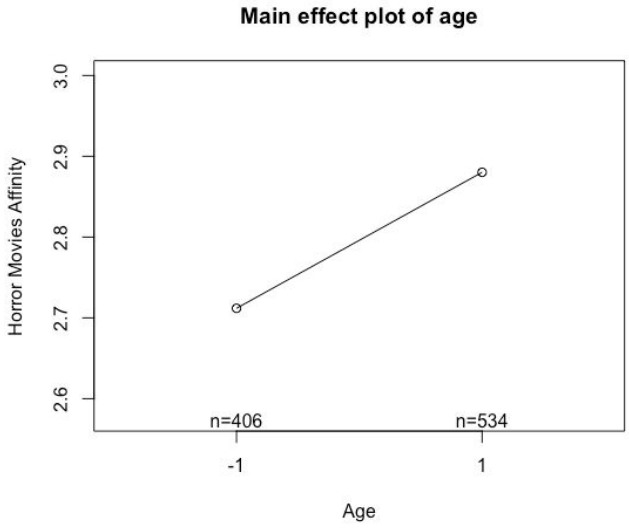


- First, the plot of the main effect of Slow or Fast music versus response variable (Horror) tells us that the people who listen to fast music also watch horror movies, maybe because both bring thrill to their life.
- In the second plot (right side) of Theatre versus Response variable tells us that the people who are interested in Theatre do not watch horror movies. Maybe because horror movies have low IMDB ratings indicating less inclusion of cinematic art.

- Plot of Fun with Friends verus Horror tells us that the people who like to socialize also prefer to watch horror movies.
- Then right side plot of Snake versus Horror tells us that as the phobia of snakes is more among people they tend to avoid horror movies.

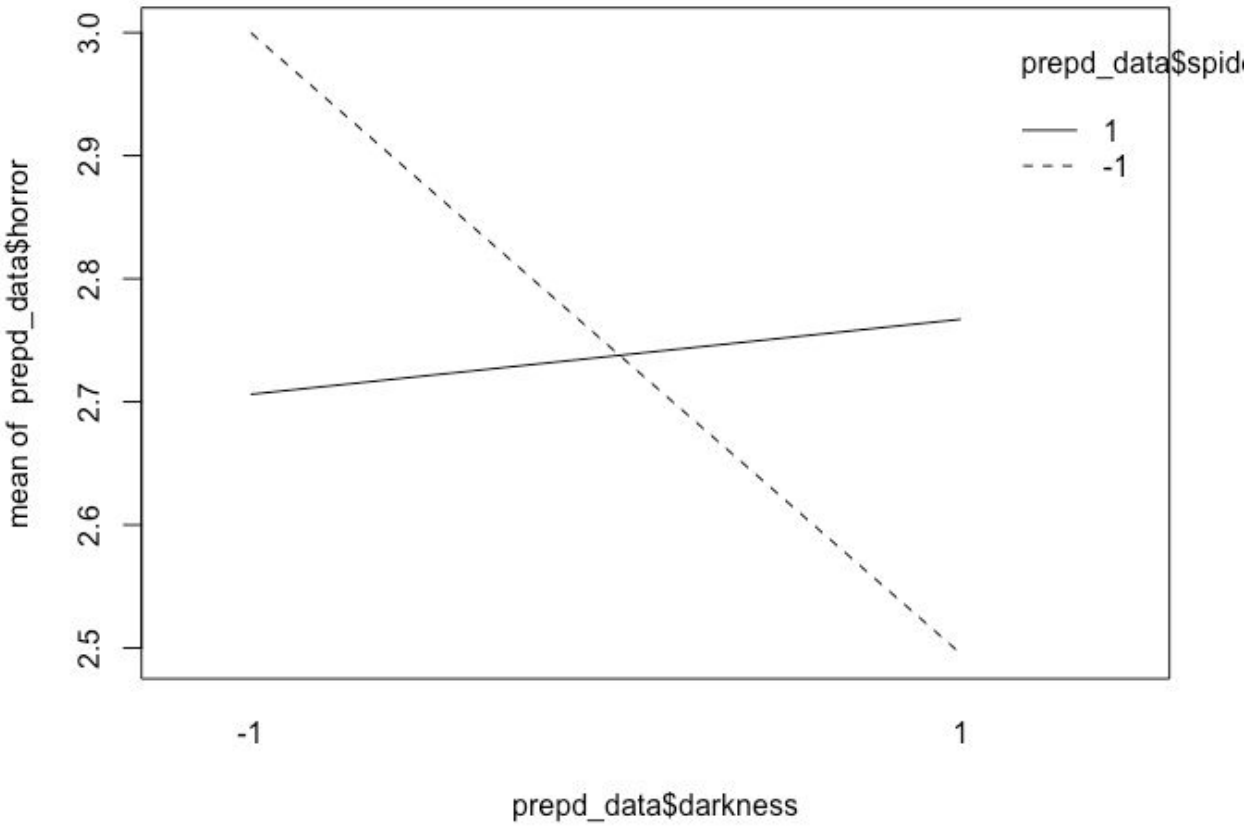






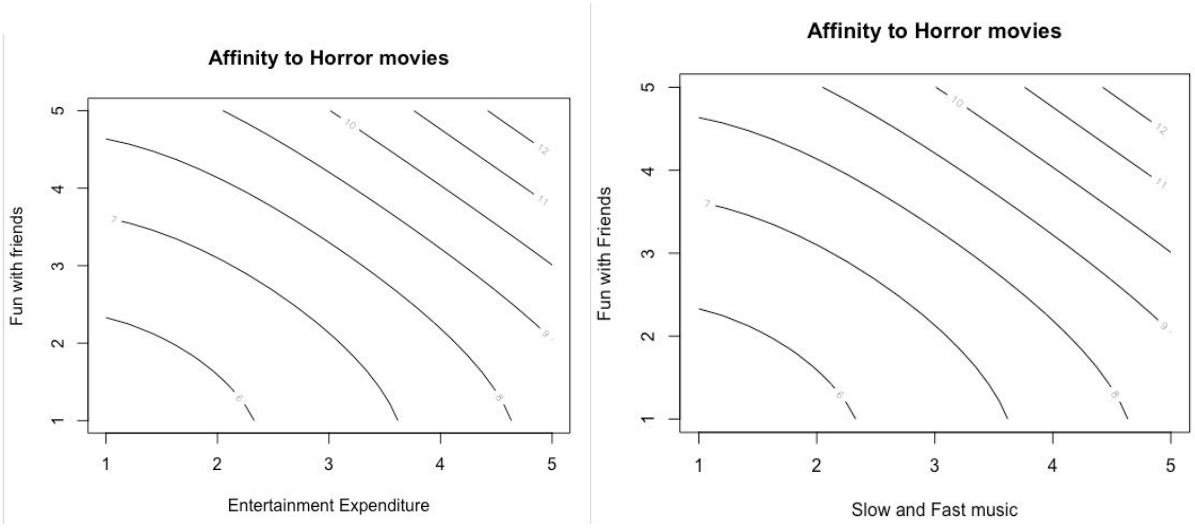
- When we plot Happy with life versus Horror, we do observe result that people who are not 100% happy with life watch horror movies more than people who are happy. Maybe the people who are happy they are spending their time with their source of happiness.
- Next, plotting Entertainment Spend versus Horror, we get pretty obvious result that people who spend more on entertainment also watch horror movies.

- Moreover, lastly among main effects plot, we can notice from plot of Age versus Horror that as the age of person increases they are more likely to watch horror movies. This could be because in childhood we might be scared to watch them.



In the interaction plot above of Spider, Darkness versus Horror, we can observe that the dotted line representing people with no spider phobia, when have fear from darkness do not watch horror movies but the ones with no fear of darkness strongly prefer to watch horror movies. Now, in dark line representing phobia of spiders no matter people have darkness phobia or not they do not watch much horror movies as there is no large change in slope of the dark line.

Counter Plots



In the two counter plot above show that all three factors of likability to socialize more with friends, listening to fast pace music as well as more expenditure for entertainment contribute to more preference of horror movie genre.

We didn't generate further contours as 3-dimensional and 2-dimensional contours couldn't accommodate the impact on response as interplayed together by 8 significant factors obtained in Reduced Model 2.

Conclusions

Final Conclusions

We concluded interesting results that the people who are of more age in the interval of 15-30 years, who spend more on entertainment, like socializing with friends and listen to fast pace music, while less happy with their lives, less interest in theatre and have no phobia of darkness and snakes prefer to watch more horror movies.

Future Work

- Firstly, in this project we worked only with horror movie preferences, using the same data all the steps could be repeated to find out the interests, traits and phobia of leading to interest in other genres of movies such as sci-fi, comedy, war, action, etc.
- For moving ahead with same problem statement to incorporate even better model we learnt from our residual plot analysis that ordinal regression can be performed in the future. As, it might be more suitable for the ordinal survey data we are dealing with in this project.

References

- Parallel straight Lines on Residuals vs. Predicted Values Plot ...." 28 Jun. 2017, https://stats.stackexchange.com/questions/287669/parallel-straight-lines-on-residuals-vs-predicted-values-plot. Accessed 1 Dec. 2019.
- "Douglas C. Montgomery-Design and Analysis of Experiments ...." http://faculty.business.utsa.edu/manderso/STA4723/readings/Douglas-C.-Montgomery-Design-and-Analysis-of-Experiments-Wiley-2012.pdf. Accessed 1 Dec. 2019.
- MATH 567 - Design and Analysis of Experiments, Illinois Institute of Technology
- "ANOVA." 17 Sep. 2018, https://stat.ethz.ch/~meier/teaching/anova/. Accessed 1 Dec. 2019.

Appendix

Full Model

horror ~
like_music + slow_fast_music + psychology + politics + physics + sci_and_tech + theatre + history + fun_with_friends + darkness + snakes + spiders + rats + work_in_spare_time + funny + happy_life + internet_usage + finances + entertainment_spend + age + watch_movies

like_music*slow_fast_music + like_music*psychology + like_music*politics
+ like_music*physics + like_music*sci_and_tech + like_music*theatre + like_music*history
+ like_music*fun_with_friends + like_music*darkness + like_music*snakes + like_music*spiders
+ like_music*rats + like_music*work_in_spare_time + like_music*funny + like_music*happy_life
+ like_music*internet_usage + like_music*finances + like_music*entertainment_spend
+ like_music*age +

slow_fast_music*psychology* + slow_fast_music*politics+slow_fast_music*physics +
slow_fast_music*sci_and_tech+slow_fast_music*theatre+slow_fast_music*history +
slow_fast_music*fun_with_friends+slow_fast_music*darkness+slow_fast_music*snakes+slow_fast_music*spiders+slow_fast_music*rats+slow_fast_music*work_in_spare_time +
slow_fast_music*funny+slow_fast_music*happy_life+slow_fast_music*internet_usage
+slow_fast_music*finances + slow_fast_music*entertainment_spend+slow_fast_music*age

+ psychology*politics + psychology*physics + psychology*sci_and_tech + psychology*theatre
+ psychology*history + psychology*fun_with_friends + psychology*darkness +
psychology*snakes + psychology*spiders + psychology*rats + psychology*work_in_spare_time
+ psychology*funny + psychology*happy_life + psychology*internet_usage +
psychology*finances + psychology*entertainment_spend + psychology*age +

politics*physics +

politics*sci_and_tech+politics*theatre+politics*history +
politics*fun_with_friends+politics*darkness + politics*snakes+politics*spiders +
politics*rats+politics*work_in_spare_time+politics*funny+politics*happy_life +
politics*internet_usage+politics*finances+politics*entertainment_spend+politics*age +

physics*sci_and_tech+physics*theatre+physics*history + physics*fun_with_friends +
physics*darkness + physics*snakes+physics*spiders +
physics*rats+physics*work_in_spare_time +
physics*funny+physics*happy_life+physics*internet_usage+physics*finances +
physics*entertainment_spend + physics*age +

sci_and_tech*theatre + sci_and_tech*history +
sci_and_tech*fun_with_friends+sci_and_tech*darkness + sci_and_tech*snakes +
sci_and_tech*spiders + sci_and_tech*rats+sci_and_tech*work_in_spare_time +
sci_and_tech*funny+sci_and_tech*happy_life+sci_and_tech*internet_usage+
sci_and_tech*finances+sci_and_tech*entertainment_spend+sci_and_tech*age +

theatre*history + theatre*fun_with_friends+theatre*darkness +
theatre*snakes+theatre*spiders+theatre*rats+theatre*work_in_spare_time + theatre*funny +
theatre*happy_life+theatre*internet_usage+ theatre*finances + theatre*entertainment_spend +
theatre*age +

history*fun_with_friends+history*darkness+history*snakes+history*spiders+history*rats +
history*work_in_spare_time+history*funny+history*happy_life+history*internet_usage +
history*finances + history*entertainment_spend + history*age +

fun_with_friends*darkness + fun_with_friends*snakes + fun_with_friends*spiders +
fun_with_friends*rats + fun_with_friends*work_in_spare_time + fun_with_friends*funny +
fun_with_friends*happy_life + fun_with_friends*internet_usage + fun_with_friends*finances +
fun_with_friends*entertainment_spend + fun_with_friends*age +

darkness*snakes + darkness*spiders + darkness*rats + darkness*work_in_spare_time +
darkness*funny + darkness*happy_life+darkness*internet_usage + darkness*finances +
darkness*entertainment_spend+darkness*age +

snakes*spiders + snakes*rats + snakes*work_in_spare_time + snakes*funny +
snakes*happy_life + snakes*internet_usage + snakes*finances + snakes*entertainment_spend
+ snakes*age + spiders*rats + spiders*work_in_spare_time + spiders*funny +
spiders*happy_life + spiders*internet_usage + spiders*finances + spiders*entertainment_spend
+ spiders*age +

rats*work_in_spare_time + rats*funny + rats*happy_life + rats*internet_usage + rats*finances +

rats*entertainment_spend + rats*age +

work_in_spare_time*funny + work_in_spare_time*happy_life +
work_in_spare_time*internet_usage + work_in_spare_time*finances +
work_in_spare_time*entertainment_spend + work_in_spare_time*age +

funny*happy_life + funny*internet_usage + funny*finances + funny*entertainment_spend +
funny*age +

happy_life*internet_usage + happy_life*finances + happy_life*entertainment_spend +
happy_life*age +

internet_usage*finances + internet_usage*entertainment_spend + internet_usage*age +

finances*entertainment_spend + finances*age +

entertainment_spend*age

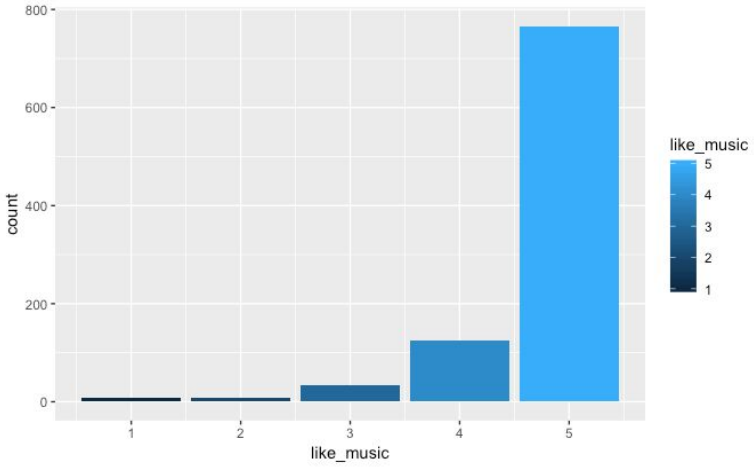
Reduced Model 1

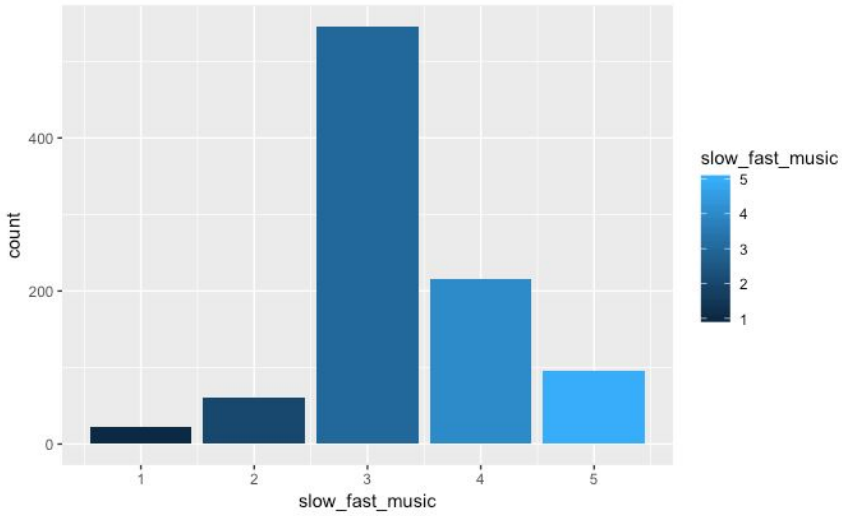
horror~ slow_fast_music+ theatre + fun_with_friends + snakes + happy_life +
entertainment_spend + age + watch_movies + like_music*history +
slow_fast_music:psychology + slow_fast_music*entertainment_spend + politics*age +
sci_and_tech*finances + sci_and_tech*entertainment_spend + history*rats + darkness*snakes
+ darkness*spiders

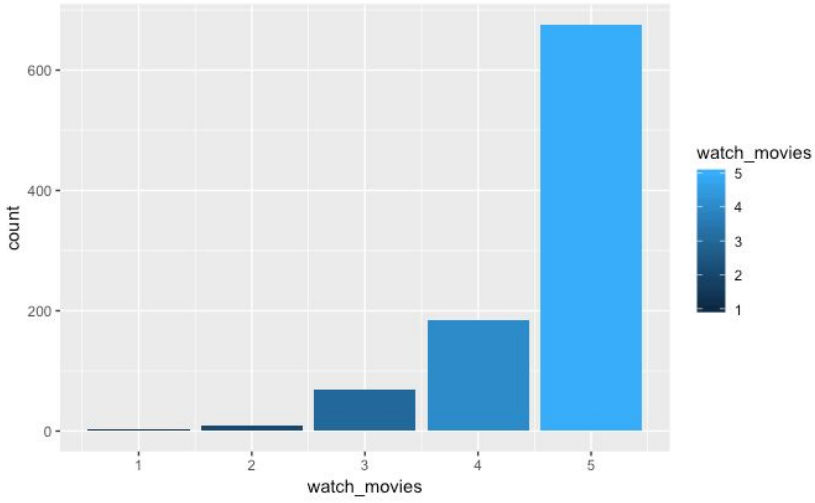
Reduced Model 2

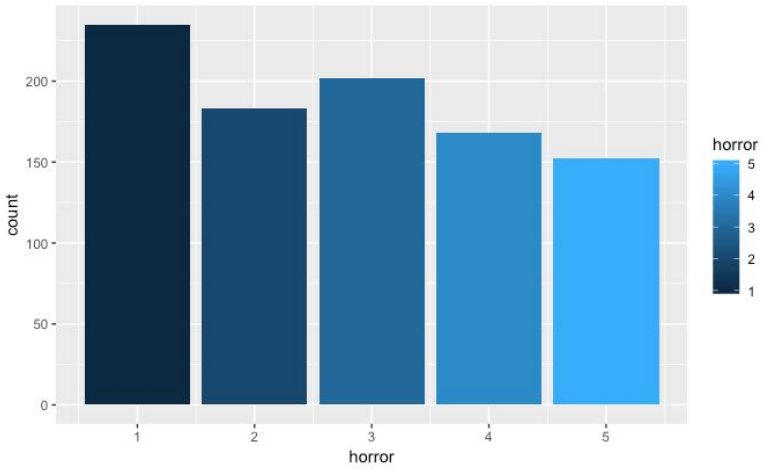
horror ~ slow_fast_music + theatre + fun_with_friends + snakes + happy_life +
entertainment_spend + age + watch_movies + entertainment_spend*sci_and_tech +
darkness*spiders

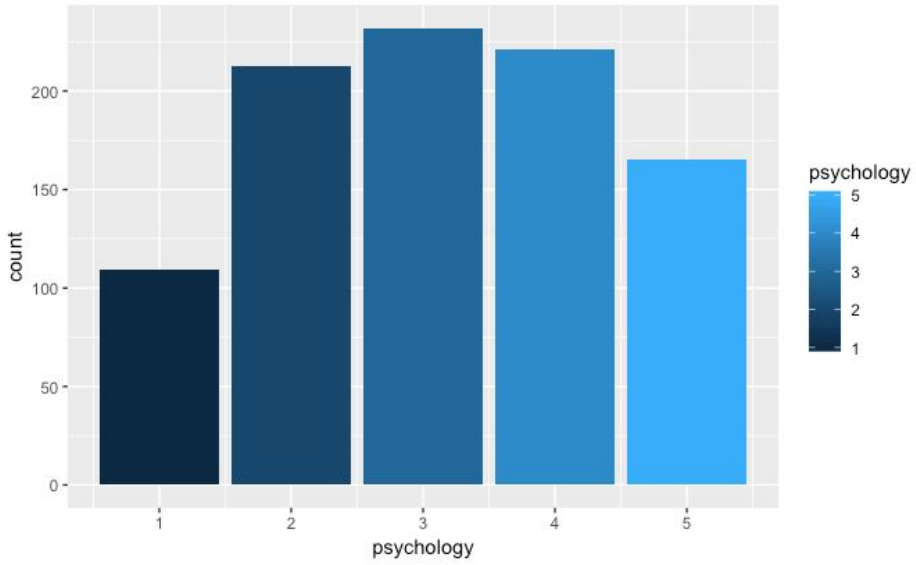
Factor level plots

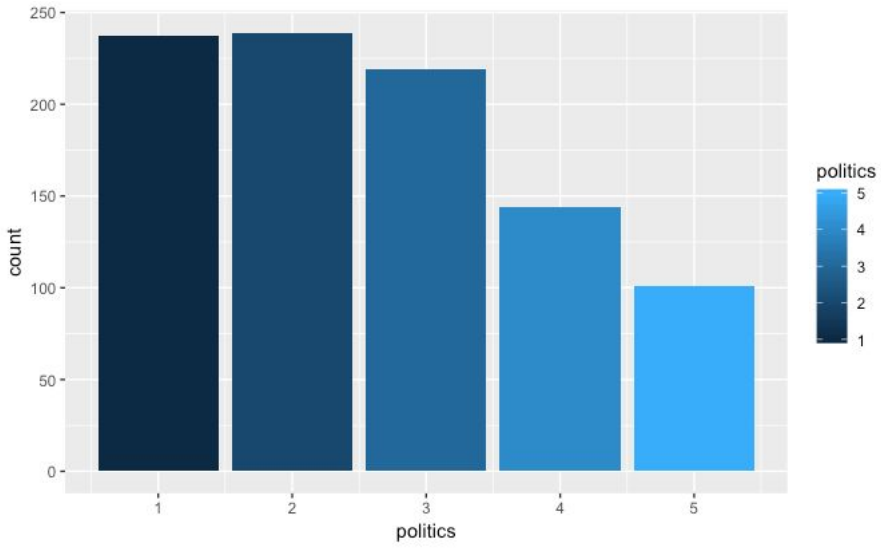

800
600
400
200
0
count
1
2
3
4
5
like_music
like_music
5
4
3
2
1

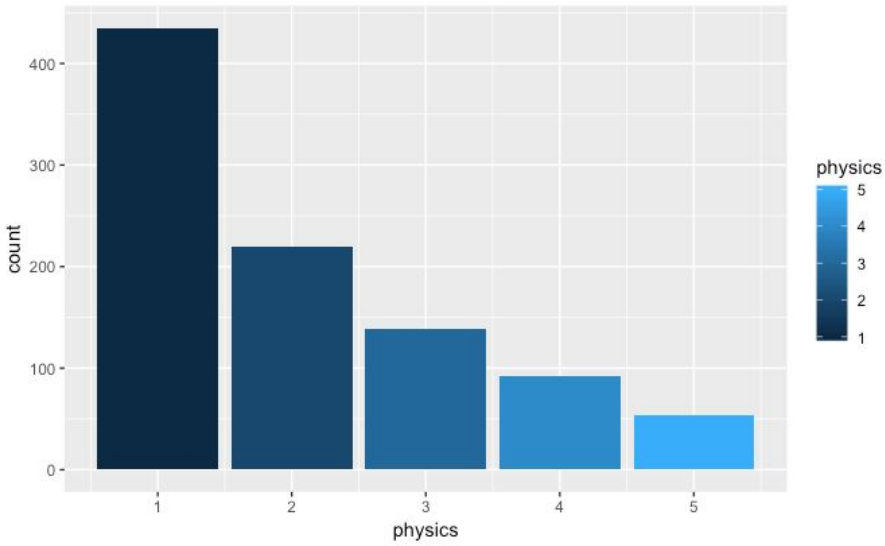

400
200
0
count
1
2
3
4
5
slow_fast_music
slow_fast_music
5
4
3
2
1

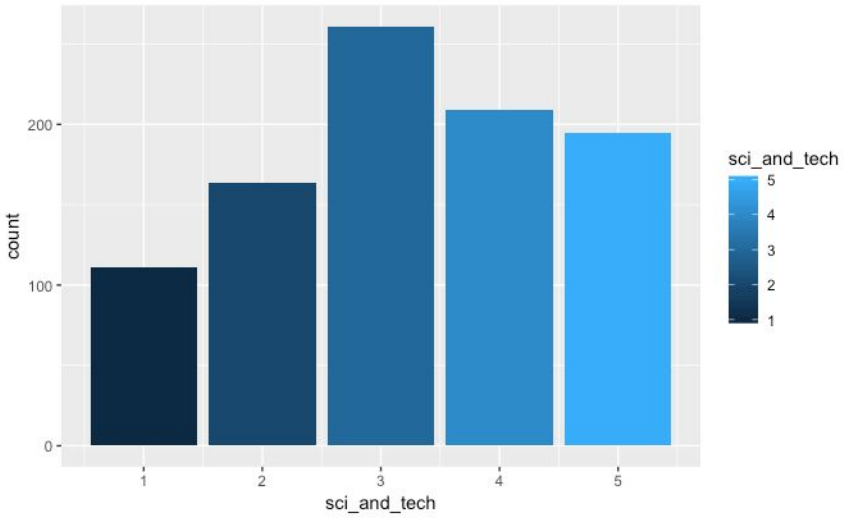

600
400
200
0
count
1
2
3
4
5
watch_movies
watch_movies
5
4
3
2
1

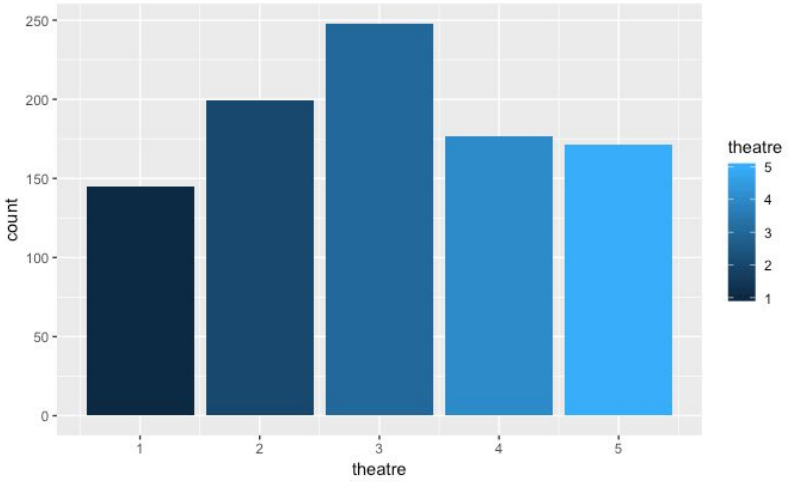

count
200
150
100
50
0
1
2
3
4
5
horror
horror
5
4
3
2
1

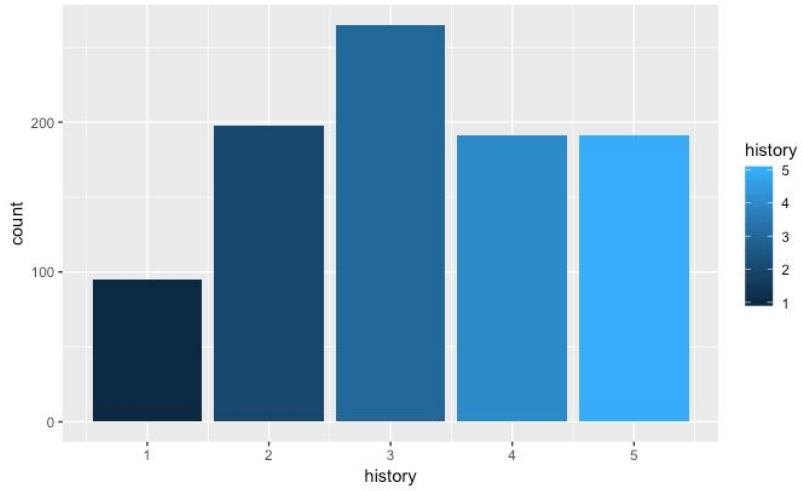

count
200
150
100
50
0
1
2
3
4
5
psychology
psychology
5
4
3
2
1

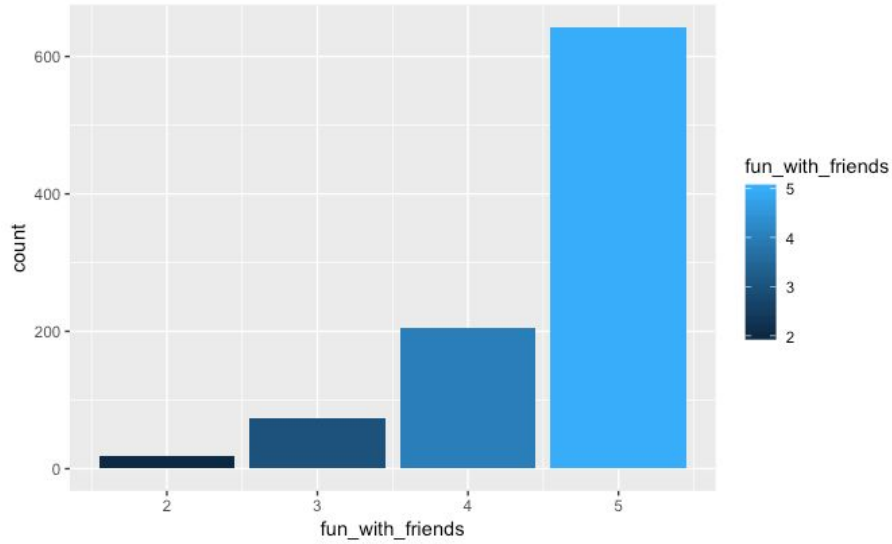

count
250
200
150
100
50
0
1
2
3
4
5
politics
politics
5
4
3
2
1

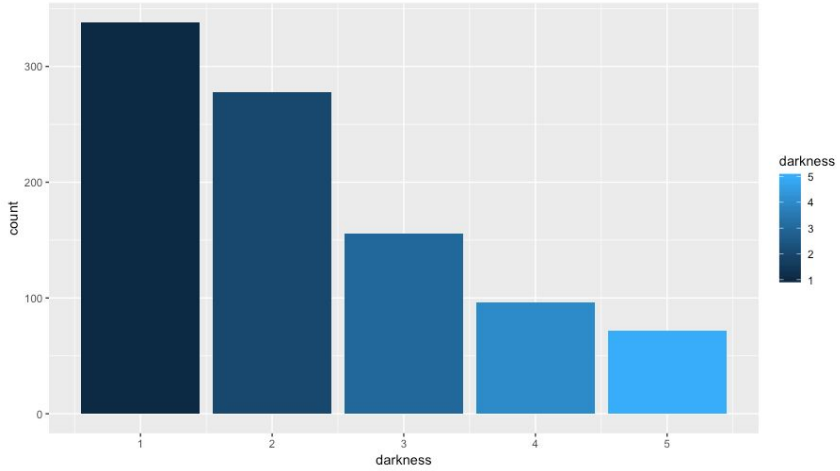

400
300
200
100
0
count
1
2
3
4
5
physics
physics
5
4
3
2
1

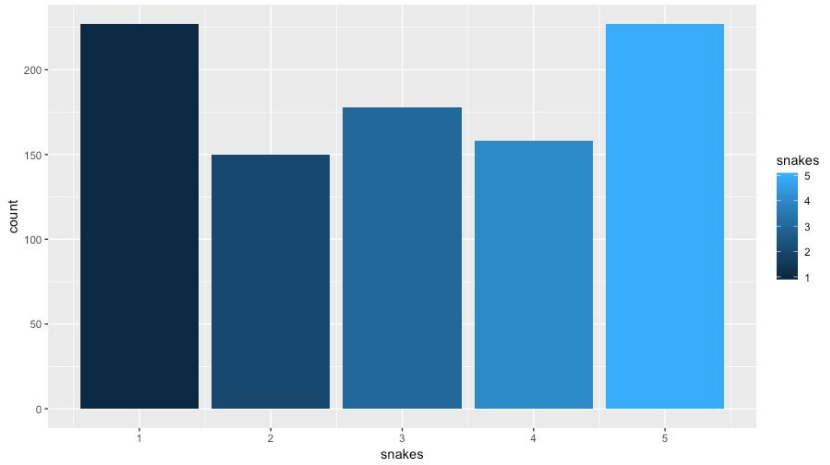

200
100
0
count
1
2
3
4
5
sci_and_tech
sci_and_tech
5
4
3
2
1

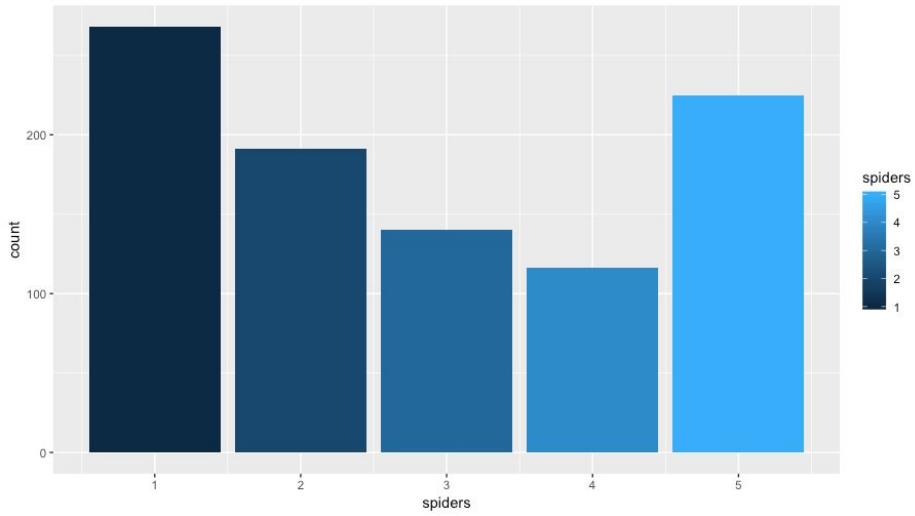

250
200
150
100
50
0
count
1
2
3
4
5
theatre
theatre
5
4
3
2
1

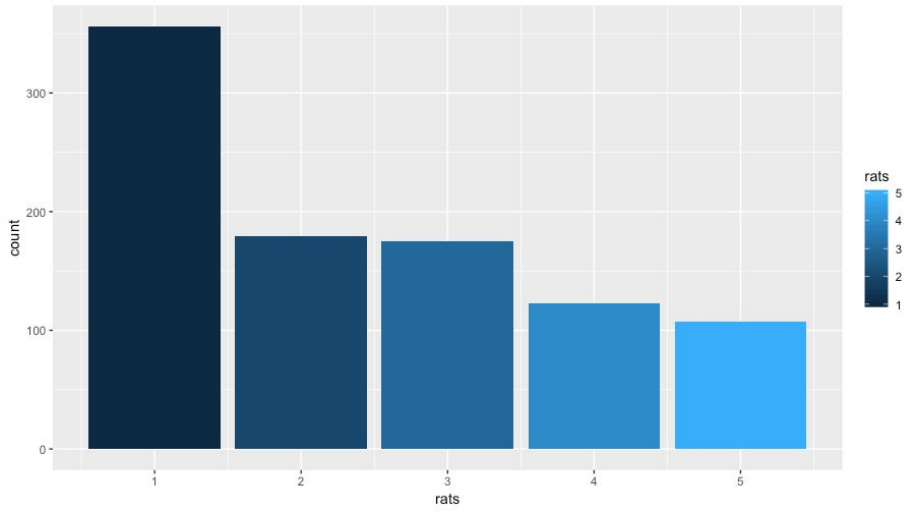

count
200
100
0
1
2
3
4
5
history
history
5
4
3
2
1

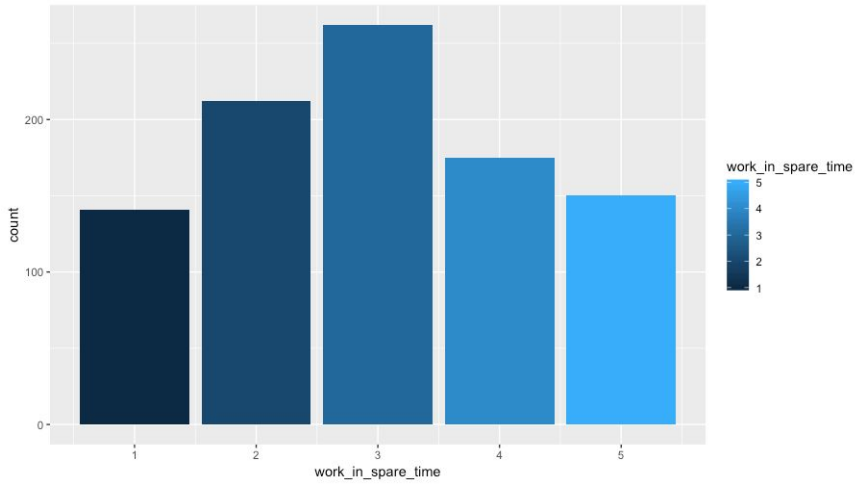

count
600
400
200
0
2
3
4
5
fun_with_friends
fun_with_friends
5
4
3
2

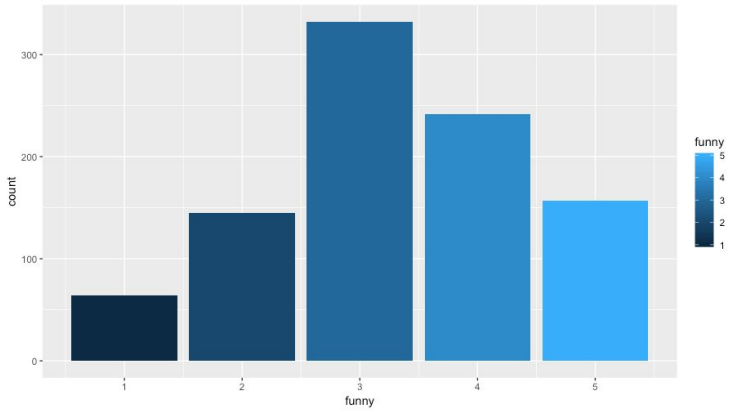

count
300
200
100
0
1
2
3
4
5
darkness
darkness
5
4
3
2
1

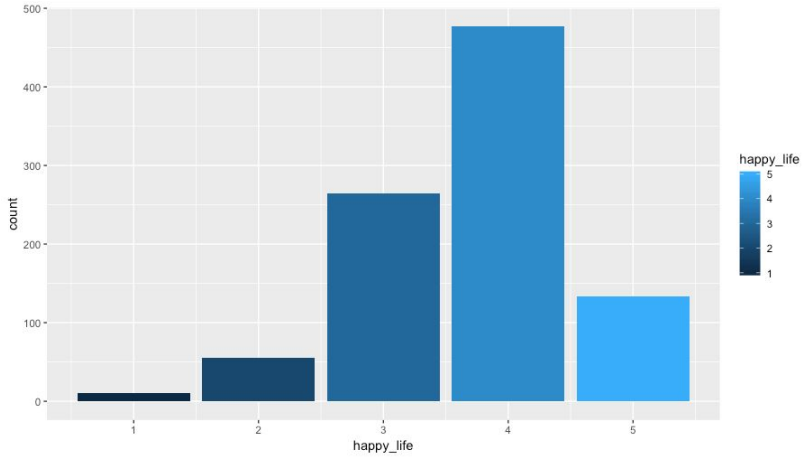

count
200
150
100
50
0
1
2
3
4
5
snakes
snakes
5
4
3
2
1

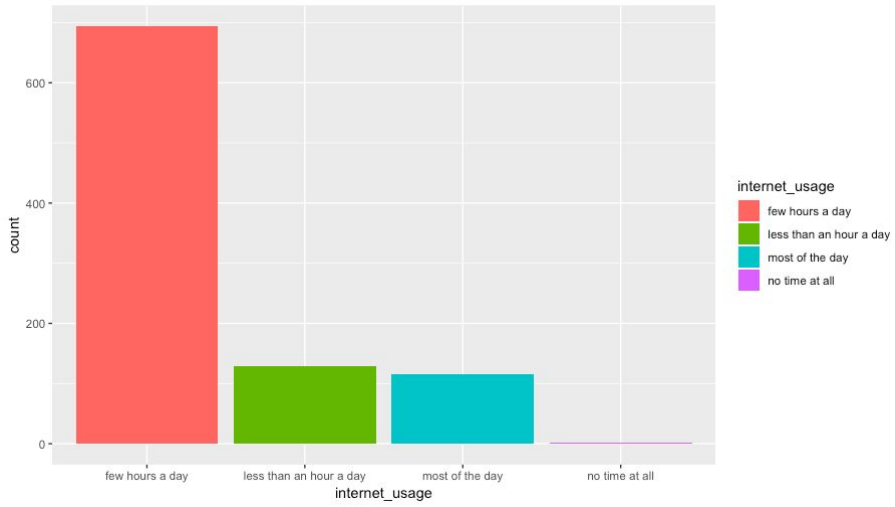

count
200
100
0
1
2
3
4
5
spiders
spiders
5
4
3
2
1

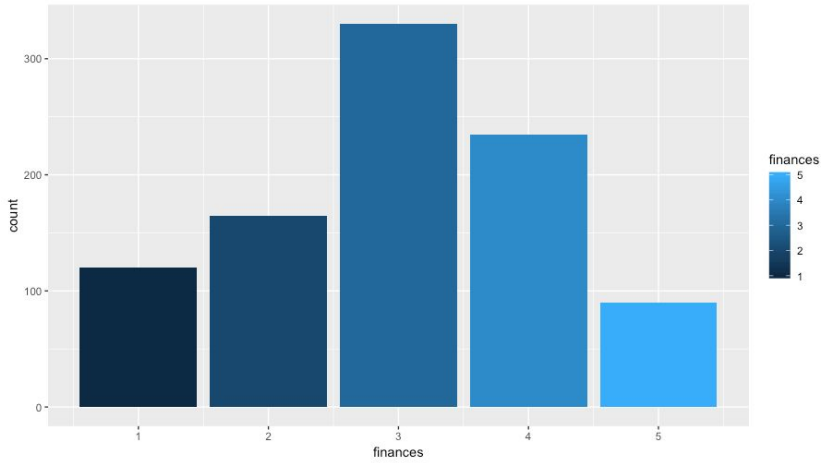

count
300
200
100
0
1
2
3
4
5
rats
rats
5
4
3
2
1

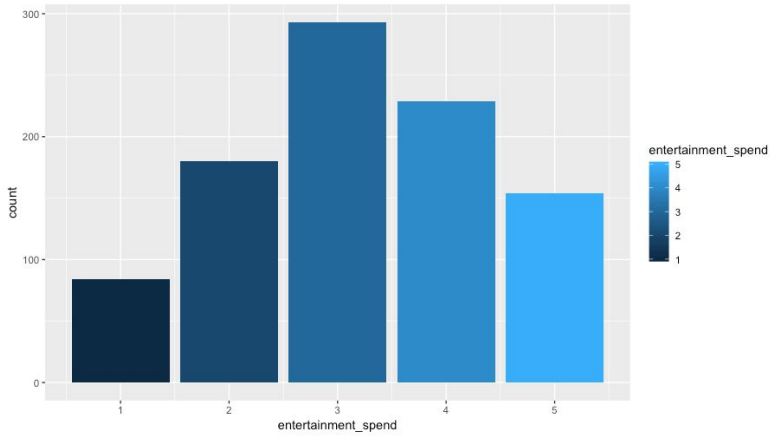

count
200
100
0
1
2
3
4
5
work_in_spare_time
work_in_spare_time
5
4
3
2
1

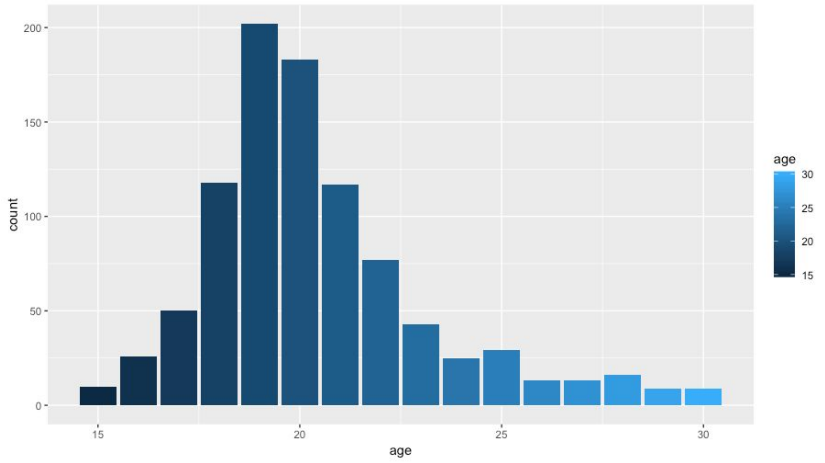

count
300
200
100
0
1
2
3
4
5
funny
funny
5
4
3
2
1


count
500
400
300
200
100
0
1
2
3
4
5
happy_life
happy_life
5
4
3
2
1

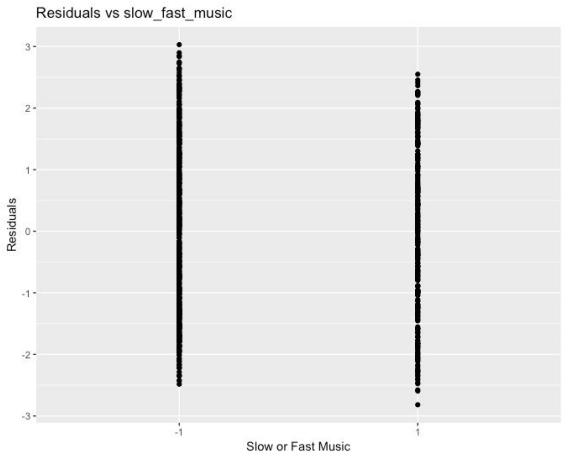

count
600
400
200
0
few hours a day
less than an hour a day
most of the day
no time at all
internet_usage
internet_usage
few hours a day
less than an hour a day
most of the day
no time at all

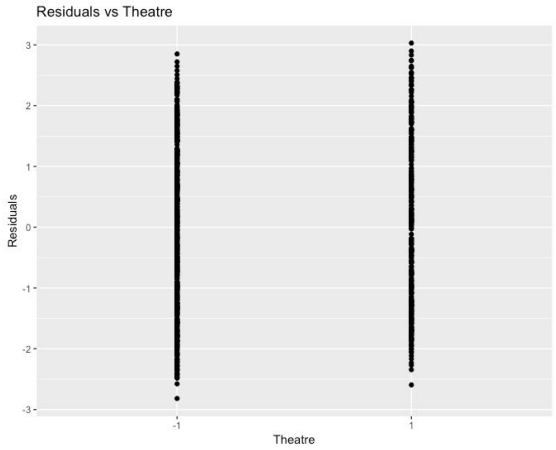

count
300
200
100
0
1
2
3
4
5
finances
finances
5
4
3
2
1

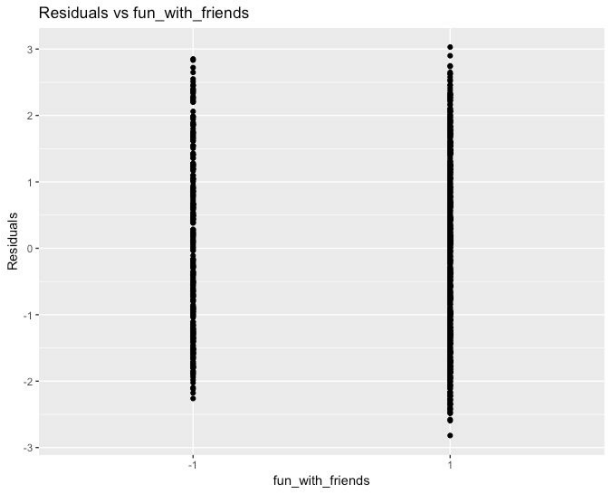

count
300
200
100
0
1
2
3
4
5
entertainment_spend
entertainment_spend
5
4
3
2
1

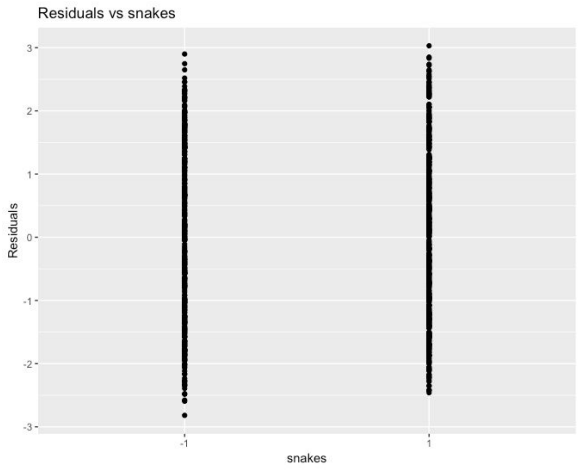


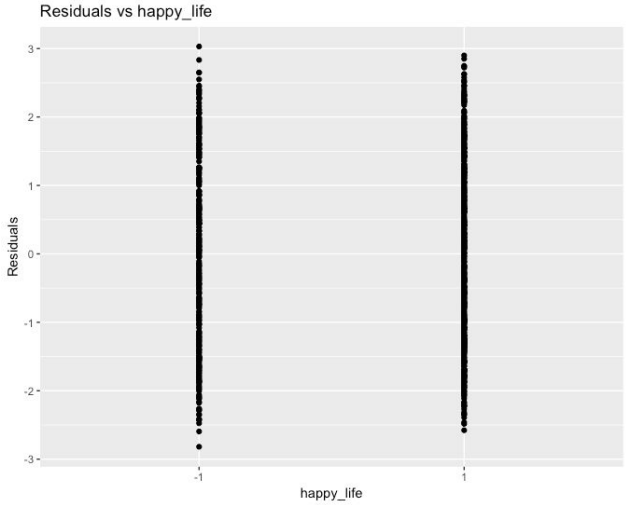
Residual to factor plots

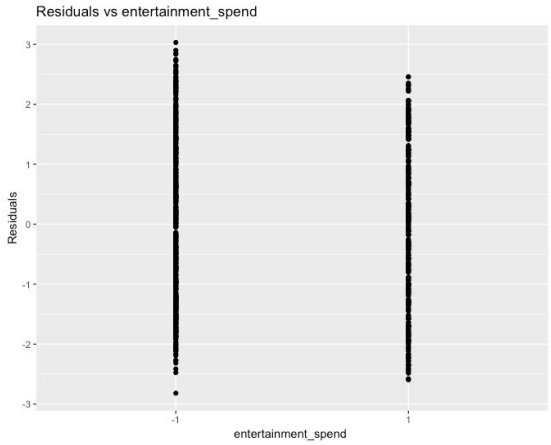


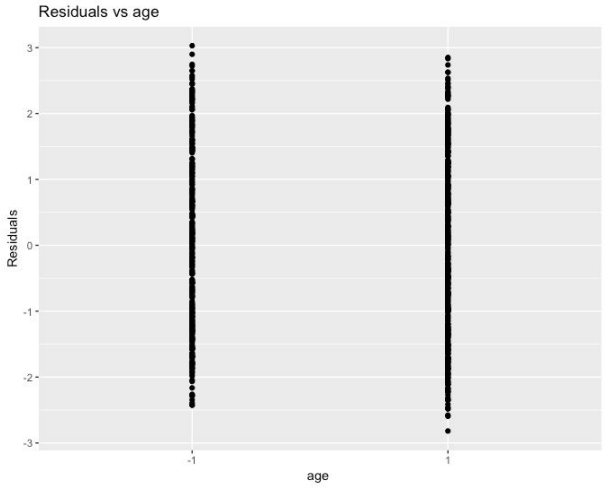


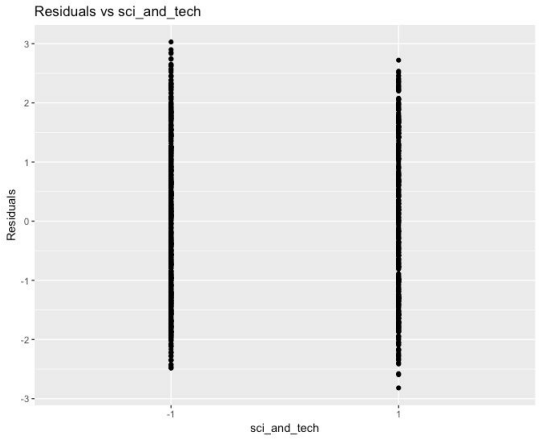

Residuals vs fun_with_friends
Residuals
fun_with_friends
3
2
1
0
-1
-2
-3
-1
1

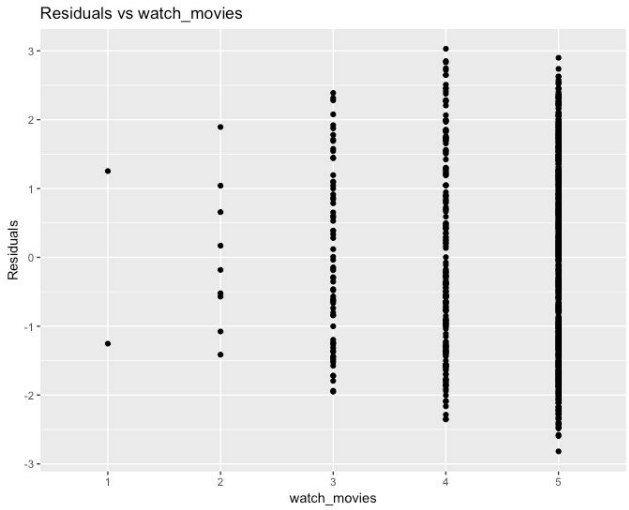

Residuals vs snakes
Residuals
snakes
3
2
1
0
-1
-2
-3
-1
1

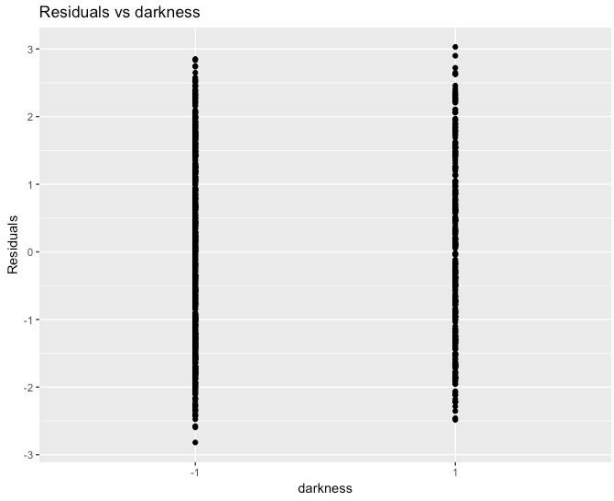

Residuals vs happy_life
Residuals
happy_life
3
2
1
0
-1
-2
-3
-1
1

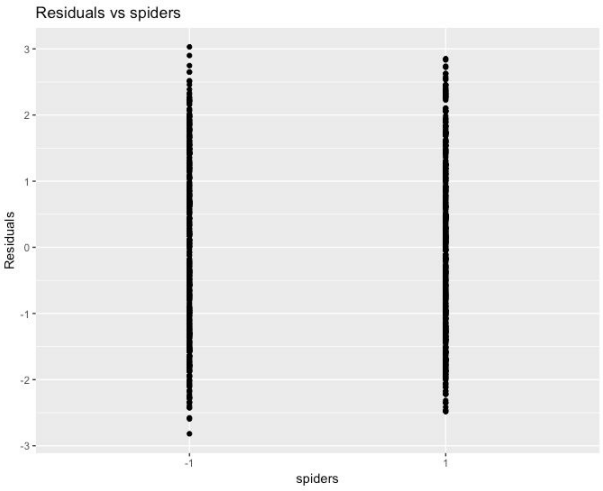

Residuals vs entertainment_spend
Residuals
entertainment_spend
3
2
1
0
-1
-2
-3
-1
1


Residuals vs age
Residuals
age
3
2
1
0
-1
-2
-3
-1
1


Residuals vs sci_and_tech
Residuals
sci_and_tech
3
2
1
0
-1
-2
-3
-1
1


Residuals vs watch_movies
Residuals
watch_movies
3
2
1
0
-1
-2
-3
1
2
3
4
5


Residuals vs darkness
Residuals
darkness
3
2
1
0
-1
-2
-3
-1
1


Residuals vs spiders
Residuals
spiders
3
2
1
0
-1
-2
-3
-1
1