Factor Analysis of Different Kinds of Phobias on Slovak Citizens between 15 to 30 years old

Probability and Statistical Inference

[redacted]

Author Note

CAPartII – Second part of the continuous assessment requirements for the Probability and Statistical Inference module of the Master on Computing programme provided by the Dublin Institute of Technology. Lecturer: Deirdre Lawless

Abstract

The research provides an Exploratory Factor Analysis on the answers to specific fears and phobias questions by 1,010 young Slovak citizens. Two factors were extracted, and they were identified as: Specific Phobia – Animals Type and Specific Phobia – Natural Environment and Situational Type, following the classification of DSM-IV (1996). A Multiple Linear Regression Model was created to predict the Specific Phobia – Animals Type construct scores using Gender, and whether the person was raised in a city or in a village or town, as predictors. The results show that females increase the phobia scores almost in 50% compared to males, and in contrast to what was expected, people raised in villages have 10% more chance to develop animal-related phobias than people raised in cities.

Keywords: Factor Analysis, FA, Linear Regression, LM, Phobias.

Factor Analysis of Different Kinds of Phobias on Slovak Citizens between 15 to 30 years old

Probability and Statistical Inference

Through many years, psychologists and psychiatrists have devoted their time and effort studying human fears and phobias, and that has led to many different statistical studies pertaining these issues in an attempt to understand the elements that contribute in a positive or negative way to these conditions, and how they are experienced by different segments of the population: gender, age, ethnicity, wealth, etcetera. Fredrikson, Annas, Fischer, & Wik (1996) found that women compared to men show higher fear ratings to all objects and situations, and that inanimate object fears were more common in older than younger individuals, whereas animal fears were more intense in younger than in older individuals. Chapman, Kertz, Zurlage, Woodruff-Borden (2008) found that different factor loading patterns of fear-related stimuli may exist among African American and Caucasian American young adults.

Those studies, and the majority of modern studies, follow the guidelines of the **Diagnostic and Statistical Manual of Mental Disorders**, in both: its **fourth edition** (APA, 1996), and, the most recent, **fifth edition** (APA, 2013); which defines *Phobia* as a type of anxiety disorder, defined by a persistent fear of an object or situation, and classifies them into: Agoraphobia, Social phobia and Specific phobias; where the latter is subdivided in the following types: Animal, Natural Environment, Blood-injection-injury, Situational, and others.

The current analysis focuses in the results of a Personality Assessment Survey conducted by students of the Faculty of Social and Economic Sciences of the Comenius University in Bratislava (https://fses.uniba.sk/en/) on Slovak citizens, ranging from 15 to 30 years-old, where a set of ten specific questions regarding their fears were asked. The analysis objective is to conduct an Exploratory Factor Analysis on the elements used to measure the fears of the persons, and

validate if the constructs extracted are similar to the ones proposed by the DSM-IV (1996). In this regard, four constructs are expected to arise from the dataset: Specific Phobia – Animal Type, Specific Phobia – Natural Environment Type, Specific Phobia – Situational Type, and Social Phobia.

Finally, and depending on the results obtained from the Factor Analysis, the extracted variables will be used in a Regression model to predict different aspects of people's behaviors using other elements of the dataset.

Methodology

The Personality Assessment Survey used for this analysis was conducted in 2013 by students of the **Faculty of Social and Economic Sciences** of the **Comenius University in Bratislava**, using Snowball sampling technique, so the participants consist mainly on students and their network of friends and acquaintances: 1,010 responses were collected from Slovak citizens ranging from 15 to 30 years-old.

The sampling technique used suggests the sample is biased and thus it should not be used to draw conclusions on a wider population than that explained above. However, it is still very useful to analyze and validate trends, behaviors, correlations and other elements and patterns present in the information.

The survey was presented in both electronic and written form, and it contained 150 items grouped in 7 categories: Music preferences; Movie preferences; Hobbies & Interests; Phobias; Health habits Personality traits, views of life and opinions; Spending habits; and Demographics. From the 150 variables of the survey: 135 are ordinal, 11 are categorical, and 4 numeric. All ordinal variables are based on 5-point Likert Scales.

The data-set was published in Kaggle (http://www.kaggle.com) by Miroslav Sabo, a former student of the Comenius University in Bratislava (Sabo, M. 2016), under the Creative Commons CC0 1.0 License (https://creativecommons.org/publicdomain/zero/1.0/), which allows its universal use.

The original questionnaire was written in Slovak language and then translated to English for distribution among the international community through Kaggle.

The variables of interest for the analysis are the set defined for Phobias and the Demographic details. Table 1 shows the details of the variables used. In general, there are 10 variables on the "Phobias" group, and 4 on the "Demographics" group. The 10 Phobias-related variables are ordinal 5-point Likert scales, and the Demographics-related are: 1 interval variable and 3 categorical variables. Dimension reduction using factor analysis will be performed on the 10 Phobia-related variables. The 4 demographic-type variables will be used in a Regression model as predictors.

Results

An initial exploration of the correlations between the Phobia-related variables was conducted to assess if dimension reduction was plausible and valid for the existing data. Table 2 presents a matrix with the correlations between the 10 variables. The matrix was constructed using a "Complete Cases" approach, since it was identified that the cases excluded were about 2% of the total sample of 1,010 observations – in concrete, there were 992 complete cases for the study. The matrix shows that there are strong correlations between fears that relate to animals and insects, and between "Fear of Darkness" and "Fear of Storms" variables. "Fear of Flying" and "Fear of Heights" also show some correlation but these is rather weak. It is interesting, however, to note that "Fear of Ageing" and "Fear of Public Speaking" do not present any

significant correlation between them and with the rest of the variables, so they stand independently of the rest. Some initial constructs appeared to be arising at this point: those related to Specific Phobia – Animal Type (Fear of Spiders, Snakes, Rats, and Dangerous Dogs), Specific Phobia – Natural Environment Type (Fear of Storms and Darkness), and Specific Phobia – Situational Type (Fear of Flying and Heights). "Fear of Public Speaking" is related to the Social Phobia construct, and that supports the evidence found of weak correlation ratios with the rest. The Social Phobia construct is underrepresented in the sample. The same is happening with "Fear of Ageing", which is loading to a different concept.

A second correlation study was performed to analyze the shared variance between these variables but excluding the unrelated variables "Fear of Ageing" and "Fear of Public Speaking". Table 3 shows the correlation matrix for this analysis. This matrix shows stronger correlations than those in Table 2. All manifest variables are correlated to at least one of the others. The strongest correlation ratios are shown between: "Fear of Darkness" and "Fear of Storms", and "Fear of Rats" and "Fear of Snakes". "Fear of Heights" and "Fear of Flying" are the ones with weakest correlation ratios with the rests among all.

A Bartlett's Test of Sphericity was conducted to assess the usefulness of dimension reduction on the data and the results show that correlation between the manifest variables are sufficiently large to support it ($\chi^2 = 1{,}677.492$, $DF = 28$, $p\text{-value} < .05$).

The determinant of the second matrix was calculated to verify existence of multicollinearity and the data came through it successfully ($Det = .1829163$).

The usefulness of the dimension reduction was also assessed with Kaiser-Meyer-Olkin MSA Test concluding the dimension reduction was favorable ($KMO = .7950784$).

A Scree plot was analyzed to identify the number of factors to extract. It was identified that, even though ambiguous, the plot suggested the extraction of 2 factors since for more than those will render the *eigenvalues* weak (< 1). Figure 1 shows the Scree plot for the data.

Based on the results of the tests above, it was decided to extract only 2 factors instead of the 4 used by Sleziak, P., Sabo, M. (2014) on their research. Thus, the Factor Analysis executed, extracted 2 factors applying orthogonal "Varimax" rotation. The two factors together explain 38% of the variance. Table 4 shows the results obtained by the Factor Analysis showing high loading coefficients in the range of .37 to .75, and the values of the variance accounted for each one the variables by the factors (communality), which in average is .38. The factors obtained, mapped to DSM-IV (1996) definitions, represent: Specific Phobia – Animals Type (PA1), and Specific Phobia – Natural Environment and Situational Type (PA2). Figure 2 provides the mappings of the manifest variables to the obtained constructs.

Cronbach's Alpha Reliability Test returned acceptable results for both factors ($\alpha_{PA1} = .73$, $\alpha_{PA2} = .64$). The results also show that all variables within the factors will drop down the reliability of the factor when removed, so all variables contribute positively to the factors.

Based on the results obtained so far, it was decided to keep the 2 factors extracted, and a Linear Regression model was defined to predict the score of the most reliable factor extracted, which is PA1, based on the gathered demographic information. The goal is to identify if there is a differential effect of those values over the Specific Phobia – Animals Type construct. Of the demographic information in the dataset, Age and Gender have been identified by Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996) as having a differential effect on specific fears and phobias, thus it made sense to include them as part of the model. However, independent tests were carried out to assess the existence of a relation of the variance of Age values and Gender

groups, with the Specific Phobia – Animals Type scores. Thus, a Spearman's Correlation Test between Age and PA1 showed no evidence supporting a correlation of the two ($S = 163090000$, $p\text{-value} = 0.4534$). This could be explained by the age range of the people in the sample, which is capturing, in particular, young people (from 15 to 30 years old), with great predominance of people between 18 to 21 years old. Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996) found the effect of the age on greater age-spams. Spearman's Test was justified on the basis that both Age and Specific Phobia – Animals Type variables were not normal. Figure 3, Figure 4, and Figure 5 provide details on the distributions of those variables.

Additionally, a Wilcoxon Test was performed on Gender and PA1 to assess the existence of a significant difference in the means of Specific Phobia – Animals Type scores between the groups. The test showed that there is significant statistical evidence to support the existence of a difference between the mean values for males ($\mu_{males} = -.283$) and females ($\mu_{females} = .197$) ($W = 155380$, $p\text{-value} < 0.01$).

For the other demographic variables of interest similar tests were conducted to gain evidence supporting its inclusion in the model. A Wilcoxon Test performed on the Village/Town variable to identify a differential effect of the groups on the value the mean of PA1, showed statistical evidence supporting the hypothesis that the mean score of the Specific Phobia – Animals Type variable is different for people that spent most of their childhood in a city ($\mu_{cities} = .033$) than those that spent it in a village ($\mu_{males} = .081$). In this case, the significance level is lower than the one observed for Gender, but it is still satisfactory ($W = 93750$, $p\text{-value} < 0.05$). The findings for Gender and Village/Town variables confirms the results obtained by Sleziak, P., Sabo, M. (2014).

A Kruskal-Wallis test was executed for the multilevel variable Education to identify a significant effect of each group on the PA1 scores. However, the test showed no evidence of the existence of such an effect, hence the null hypothesis cannot be rejected ($\chi^2 = 5.7133$, *DF = 5, p-value = 0.3351*).

An initial Multiple Linear Regression Model was conducted on all candidate variables, including Age based on previous research findings. The results in Table 6 show that the model is capable of predict the values of the dependent variable (*F = 28.51 on 3 and 972 DF, p-value = < 0.01*), although it is only explaining about 8% of its variance ($R^2 = .08087$, *Adjusted* $R^2 = 0.07803$). The model also shows that Gender and Village/town variables have a significant differential effect on the values of PA1 scores (*Gender p-value < .001, Village/town p-value < .05*). However, the model indicates that Age variable does not work well as a predictor of the dependent variable for this dataset (*Age p-value = .596602*).

A second Multiple Linear Regression Model was executed excluding Age variable as a predictor. Table 7 shows the results of this iteration. The results show the model is still relevant for predicting the PA1 score (*F = 43.16 on 2 and 979 DF, p-value = < 0.01*), and point to a slight improvement of the predictive power ($R^2 = .08103$, *Adjusted* $R^2 = 0.07915$). However, this improvement is marginal. It is worth noticing that the exclusion of the Age variable did not bring any major positive or negative effect to the model. The model is represented by the following equation:

$$\text{PA1} = -0.19840 + 0.48114 * \text{female} + (-0.12018) * \text{city}$$

The model proposes an increase of .48 units of the PA1 phobias score when a person is a female and then a decrease of .12 units when the person has been raised in a city. It is important

to notice that in the case of a female raised in a village the phobias score gets incremented by .6 units, which is 70% of 1 standard deviation of the score distribution.

Discussion

The results presented in this report show an Exploratory Factor Analysis on the fears and phobias data collected by the mentioned survey on young Slovak citizens. The Factor Analysis extracted 2 factors that were identified as: Specific Phobia: Animals -Type and Specific Phobia: Natural Environment and Situational Type; following the definitions and classification provided in DSM-IV (1996). It was found that only 8 out of 10 phobia-related variables were viable for factor analysis. The other two, “Fear of Ageing” and “Fear of Public Speaking”, seem to be related to different constructs than the ones found, and that were underrepresented in the dataset. The finding goes in line with the Exploratory Factor Analysis performed by Sleziak, P., Sabo, M. (2014) although they extracted 4 factors instead of 2.

The Specific Phobia – Animals Type construct was selected to be used in a Multiple Linear Regression Model for two reasons: first, it was the strongest factor of the two, and second, to confirm some of the findings Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996). The Multiple Linear Regression Model goal was to identify differential effects of some of the demographic variables of the dataset – such as: Gender, Age, Education, etc.; on the Specific Phobia – Animals Type latent variable scores.

In contrast to the findings of Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996), there was no correlation found between the person’s age and the its phobia-towards-animals score. This could be explained by the fact that the dataset focuses in people between 15 to 30 years old, and 50% of the observations lie between 19 and 22 years old. Then the dataset does not provide enough variety of the samples to find an effect by this variable.

A Multiple Linear Regression Model was created using the person's gender, and whether the person was raised in a city or in a village, as predictors. The model is not powerful, since it can only explain around 8% of the variance of the score, but it was found that it provides a better fit compared to the mean. The main insights derived by the model are:

1) Being a female increases the chances of having a specific animal-related phobia. Females add almost a 50% contribution to the phobia score compared to males.
2) People raised in cities have a lower chance of develop an animal-related phobia. People raised in villages have 10% more chances of developing an animal-related phobia.

The latter is quite surprising since goes against what we think is the common belief, i.e. that people raised in cities would have greater odds in developing phobias towards animals than people raised in villages and town. This result could be due to the higher representation of people raised in cities in the sample (70%), but certainly opens the path for further investigation.

References

Sabo, M. (2016). Young People Survey. Retrieved from https://www.kaggle.com/miroslavsabo/young-people-survey/data on October 24th, 2017.

Sleziak, P., Sabo, M. (2014). Gender differences in the prevalence of specific phobias. *Forum Statisticum Slovacum. 2014, Vol. 10, No. 6*.

Sabo, M. (2014). Multivariate Statistical Methods with Applications. *Diss. Slovak University of Technology in Bratislava*.

American Psychiatric Association (1994). Diagnostic and Statistical Manual of Mental Disorders (4th ed.) (DSM-IV). Washington, DC: Author.

American Psychiatric Association (2013), Diagnostic and Statistical Manual of Mental Disorders (5th ed.), Arlington: American Psychiatric Publishing, pp. 190, 197-202, ISBN 0890425558.

Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour research and therapy. 34*. 33-9. 10.1016/0005-7967(95)00048-3. Available from: https://www.researchgate.net/publication/14638289_Gender_and_age_differences_in_the_prevalence_of_specific_fears_and_phobias

Chapman, L.K., Kertz, S.J., Zurlage, M.M., & Woodruff-Borden, J. (2008). A confirmatory factor analysis of specific phobia domains in African American and Caucasian American young adults. Journal of Anxiety Disorders, 22. 763-771. Available from: https://www.psychologytoday.com/files/attachments/74833/confirmatory-factor-analysis-specific-phobia-domains.pdf

Hadley, W., Romain, F., Lionel, H., & Kirill, M. (2017). dplyr: A Grammar of Data Manipulation. R package version 0.7.4. Available from: https://CRAN.R-project.org/package=dplyr

Schafer, J., Opgen-Rhein, R., Zuber, V., Ahdesmaki, M., Duarte Silva, A. P., & Strimmer, K.. (2017). corpcor: Efficient Estimation of Covariance and (Partial) Correlation. R package version 1.6.9. Available from: https://CRAN.R-project.org/package=corpcor

Bernaards, C., & Jennrich, R. (2014). GPArotation: GPA Factor Rotation. R package version 2014.11-1. Available from: https://CRAN.R-project.org/package=GPArotation

Revelle, W. (2017). psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 1.7.8. Available from: https://CRAN.R-project.org/package=psych

Raiche, G., & Magis, D. (2011). nFactors: Parallel Analysis and Non Graphical Solutions to the Cattell Scree Test. R package version 2.3.3. Available from: https://CRAN.R-project.org/package=nFactors

Wickham, H., & Chang, W. (2016). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. R package version 2.2.1. https://CRAN.R-project.org/package=ggplot2

Tables

Table 1

Variables used on the Analysis

Group	Variable	Description
Demographics	Age	Age of the participant. Interval.
	Gender	Gender of the participant. Categorical. Binary. Values: "male", "female".
	Education	Highest education achieved. Categorical. Values: "Currently a Primary school pupil", "Primary school", "Secondary school", "College/Bachelor degree", "Masters degree", "Doctorate degree".
	Village/Town	Question "I spent most of my childhood in a …". Categorical. Values: "city", "town".
Phobias	Flying	Assess how strong is the participant's fear towards flying. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Storms: Thunder, lightning	Assess how strong is the participant's fear towards thunder and lightning. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Darkness	Assess how strong is the participant's fear towards darkness. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Heights	Assess how strong is the participant's fear towards heights. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)

	Spiders	Assess how strong is the participant's fear towards spiders. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Snakes	Assess how strong is the participant's fear towards snakes. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Rats, mice	Assess how strong is the participant's fear towards rats and mice. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Ageing	Assess how strong is the participant's fear towards ageing. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Dangerous dogs	Assess how strong is the participant's fear towards dogs. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)
	Public speaking	Assess how strong is the participant's fear towards speaking in public. Ordinal. 5-point Likert scale ranging from 1 (Not afraid at all) to 5 (Very afraid of)

Note: The table above shows the definition of each one of the variables used during the analysis of this document. Each variable represents a question of the survey used to collect the information.

Table 2

Phobias Correlation Matrix Using Complete Observations

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Flying (1)	*									
Storm (2)	**.331**	*								
Darkness (3)	.187	**.503**	*							
Heights (4)	.246	.273	**.305**	*						
Spiders (5)	.123	.258	**.312**	.17	*					
Snakes (6)	.215	**.321**	.264	.21	**.434**	*				
Rats (7)	.184	.295	**.3**	.195	**.375**	**.573**	*			
Ageing (8)	.138	.177	.219	.133	.149	.16	.231	*		
Dang. Dogs (9)	.174	.286	.207	.215	.218	**.376**	**.416**	.262	*	
Pub.Speak. (10)	.137	.086	.158	.143	.15	.148	.137	.107	.191	*

Note: The table shows that there is a strong correlation between the fears that relate to animals and insects and there is another strong correlation between “Fear of Darkness” and “Fear of Storms”. These suggest the rise of two possible factors that might be aligned to the DSM-IV (1996) Phobias classification: “Specific phobias: Animals” and “Specific phobias: Natural Environment” respectively. The table also shows that “Fear of Flying” and “Fear of Heights” manifest variables experience lower correlation factors than the rest, while “Fear of Ageing” and “Fear of Public Speaking” do not present any significant correlation.

Table 3

Phobias Correlation Matrix Using Complete Observations Without Not-Correlated Items

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Flying (1)	*							
Storm (2)	**.328**	*						
Darkness (3)	.186	**.506**	*					
Heights (4)	.243	.275	**.305**	*				
Spiders (5)	.121	.26	**.313**	.172	*			
Snakes (6)	.214	**.324**	.268	.21	**.434**	*		
Rats (7)	.183	**.3**	**.304**	.195	**.375**	**.575**	*	
Dang. Dogs (8)	.173	.289	.212	.215	.219	**.379**	**.419**	*

Note: The correlation table shows stronger correlations than those in Table 2. All manifest variables are correlated to at least one of them and the ones that show the highest correlation ratios are: “Fear of Darkness” vs. “Fear of Storms”, “Fear of Snakes” vs. “Fear of Spiders”, “Fear of Rats” vs. “Fear of Snakes”, and “Fear of Dangerous Dogs” vs. “Fear of Rats”. “Fear of Heights” and “Fear of Flying” are still the ones with weakest correlation ratios among all.

Table 4

Factor Analysis Results

Variable "Fear of …"	**Item**	**Factor 1**	**Factor 2**	**Communalities**	**Uniqueness**
Snakes	6	0.75		0.60	0.40
Rats	7	0.73		0.57	0.43
Spiders	5	0.46		0.27	0.73
Dangerous dogs	8	0.45		0.26	0.74
Storm	2		0.72	0.56	0.44
Darkness	3		0.61	0.42	0.58
Heights	4		0.41	0.20	0.80
Flying	1		0.37	0.16	0.84
SS loadings		1.66	1.39		
Proportion Variance		0.21	0.17		
Cumulative Variance		0.21	0.38		
Proportion Explained		0.55	0.45		
Cumulative Proportion		0.55	1.00		

Note: The table shows the results of the Factor Analysis extraction for 2 factors on the dataset. "Varimax" orthogonal rotation was applied to maximize the allocation of manifest variables to a single factor.

Table 5

Cronbach's Alpha Test Results.

Factor 1		**Factor 2**	
Variable	**Alpha**	**Variable**	**Alpha**
Spiders	0.72	Flying	0.62
Snakes	0.60	Storm	0.49
Rats	0.61	Darkness	0.54
Dangerous dogs	0.72	Heights	0.61
Overall	**0.73**	**Overall**	**0.64**

This table shows the overall values of Alpha for each factor and then the specific value Alpha would take if the given variable was removed from the factor. It is worth noticing that every variable will cause a decrease of Alpha if removed from the factor.

Table 6

Multiple Linear Regression Model 1 Results

Residuals:					
	Min	**1Q**	**Median**	**3Q**	**Max**
	-1.98555	-0.69494	-0.04037	0.62483	2.07337
Coefficients:					
Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Sig.**
(Intercept)	-0.20834	0.0575	-3.623	0.000306	***
fsevukPhobiasCCExtended$female	0.48531	0.05371	9.035	< 2e-16	***
fsevukPhobiasCCExtended$zAge	0.01406	0.02655	0.529	0.596602	
fsevukPhobiasCCExtended$city	-0.11391	0.05744	-1.983	0.047642	*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Residual standard error: 0.8176 on 972 degrees of freedom					
Multiple R-squared:	0.08087				
Adjusted R-squared:	0.07803				
F-statistic: 28.51 on 3 and 972 DF, p-value: < 2.2e-16					

The table shows the results obtained in the first iteration of the regression model including all identified predictor variables. First, the results show that the model is a good fit for predicting the scores (*p-value < .001*). Second, it shows that Gender and Village/town variables have a significant differential effect on the values of PA1 scores. In particular, there is a .48 increase in the Phobia scores for females compared to males, and, what is quite surprising, people that have been raised most of their childhood in a city have .11 less chance to have animal-related phobias. Finally, the result shows that Age does not have a particular influence on the phobias score for this particular dataset.

Table 7

Multilevel Regression Model 2 Results

Residuals:					
	Min	**1Q**	**Median**	**3Q**	**Max**
	-1.94364	-0.68651	-0.0144	0.63024	2.1
Coefficients:					
Variable	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**	**Sig.**
(Intercept)	-0.1984	0.05717	-3.47	0.000543	***
fsevukPhobiasCCExtended$female	0.48114	0.05306	9.068	< 2e-16	***
fsevukPhobiasCCExtended$city	-0.12018	0.05715	-2.103	0.035737	*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Residual standard error: 0.817 on 979 degrees of freedom					
Multiple R-squared:	0.08103				
Adjusted R-squared:	0.07915				
F-statistic: 43.16 on 2 and 979 DF, p-value: < 2.2e-16					

The table presents the results of the second iteration of the multilevel regression model excluding Age variable as predictor. The results show that the model is still relevant for predicting the Specific Phobia – Animals Type score, and point to a slight fit improvement from the first iteration (Table 6) in terms of the variance explained by the model. It is worth noticing that the exclusion of the Age variable did not bring any major positive or negative effect to the model.

Figures

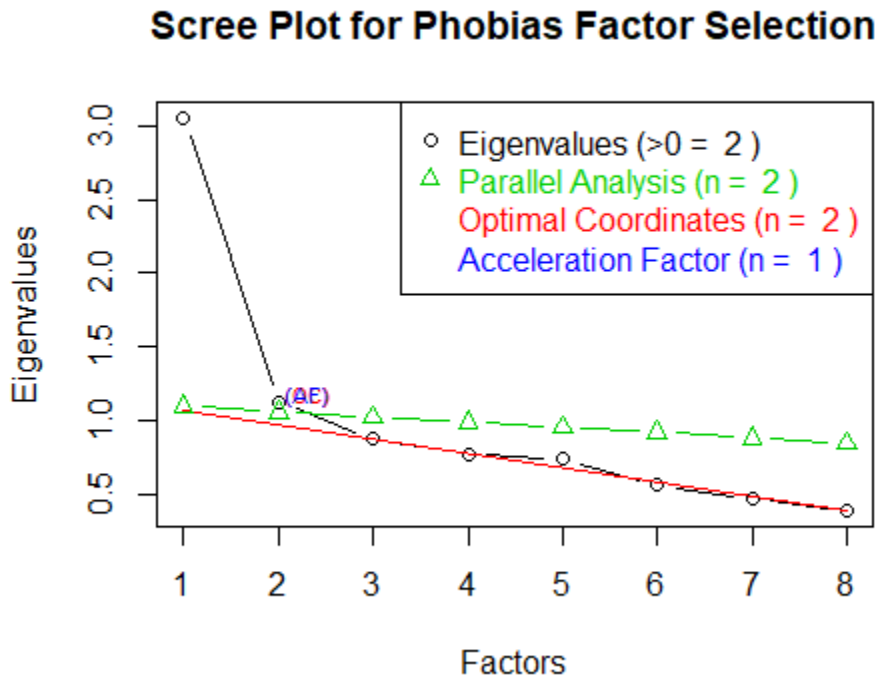

Figure 1. Scree plot for selection of factors from 8 phobias manifest variables.

The figure shows that the data is better fit for two factors as the first inflection point is found at this value and the corresponding eigenvalues are above 1. Additional inflection points can be found for 4 and 6 number of factors.

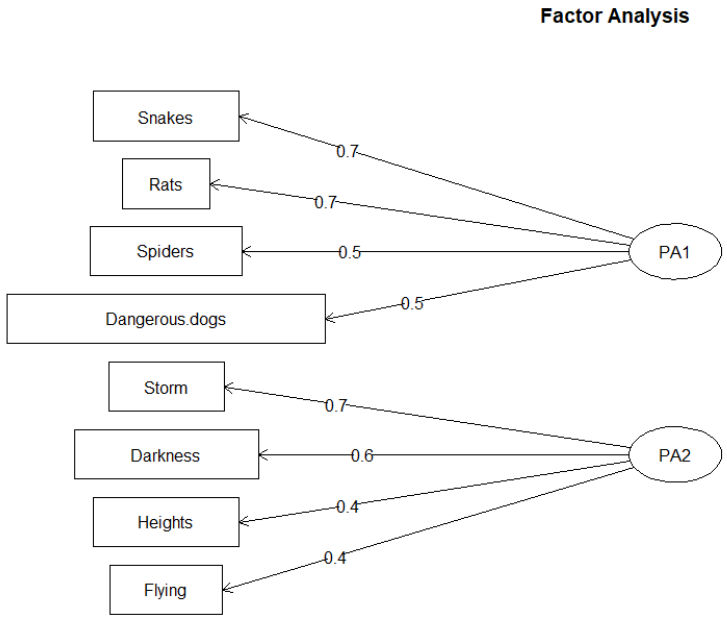

Figure 2. Factor analysis extraction for 2 factors using "Varimax" rotation.

The figure shows the 2 factors extracted and how each manifest variable load to each one of them. Based on this result, and according to the DSM-IV (1996) definition, the factors obtained represent: Specific Phobia – Animals Type (PA1), and Specific Phobia – Natural Environment and Situational Type (PA2).

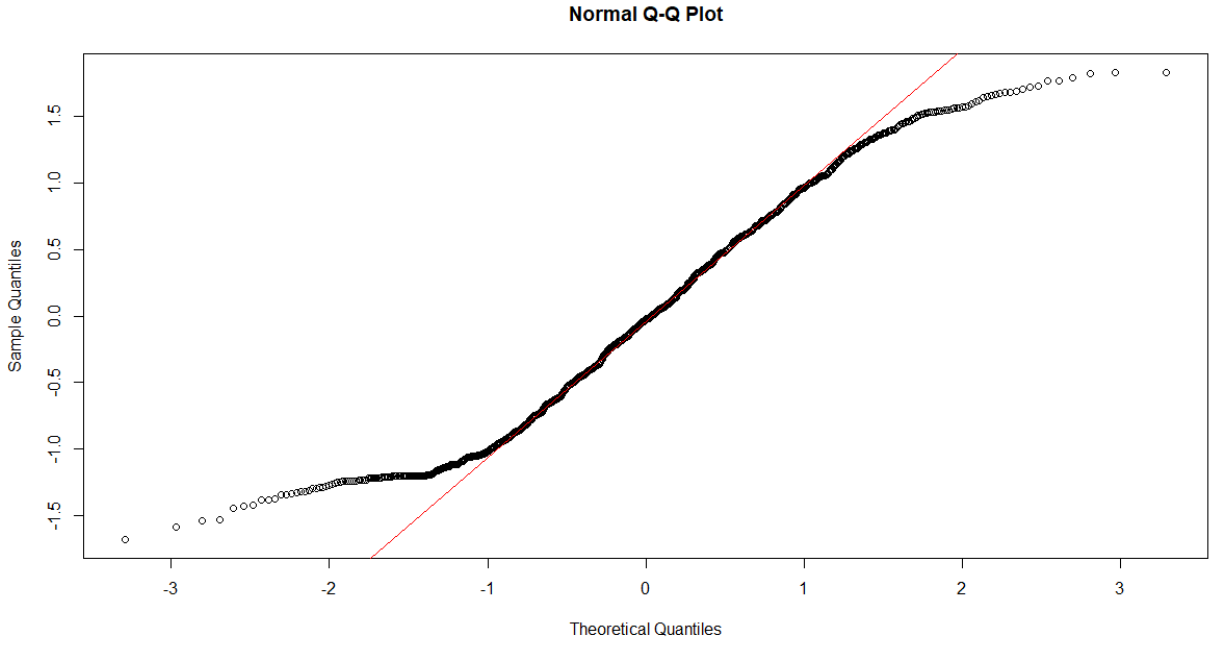

Figure 3. Quantile-Quantile plot for Fear of Animals distribution.

The figure shows that Fear of Animals distribution can not be considered normal since the values start to spread out far from the normal distribution line after passing 1 standard deviation from the mean.

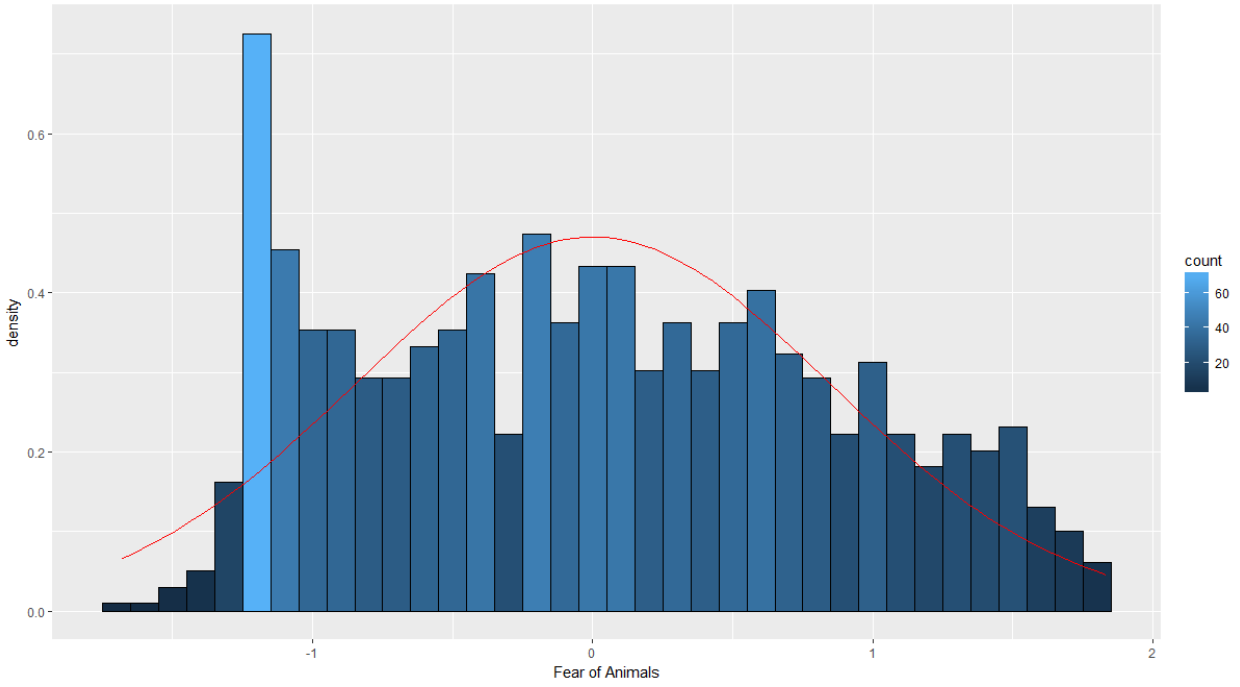

Figure 4. Fear of Animals distribution histogram.

The histogram above clearly shows that the distribution is far from close to a bell-shaped symmetrical distribution. Mean (*0*) and median (*-0.0242671*) values are different from each other, and the mode (*-1.2*) is far from both. Standardized values for skewness (*6*) and kurtosis (*-33.34*) are far from the acceptable range (-2 to 2) to be considered normal.

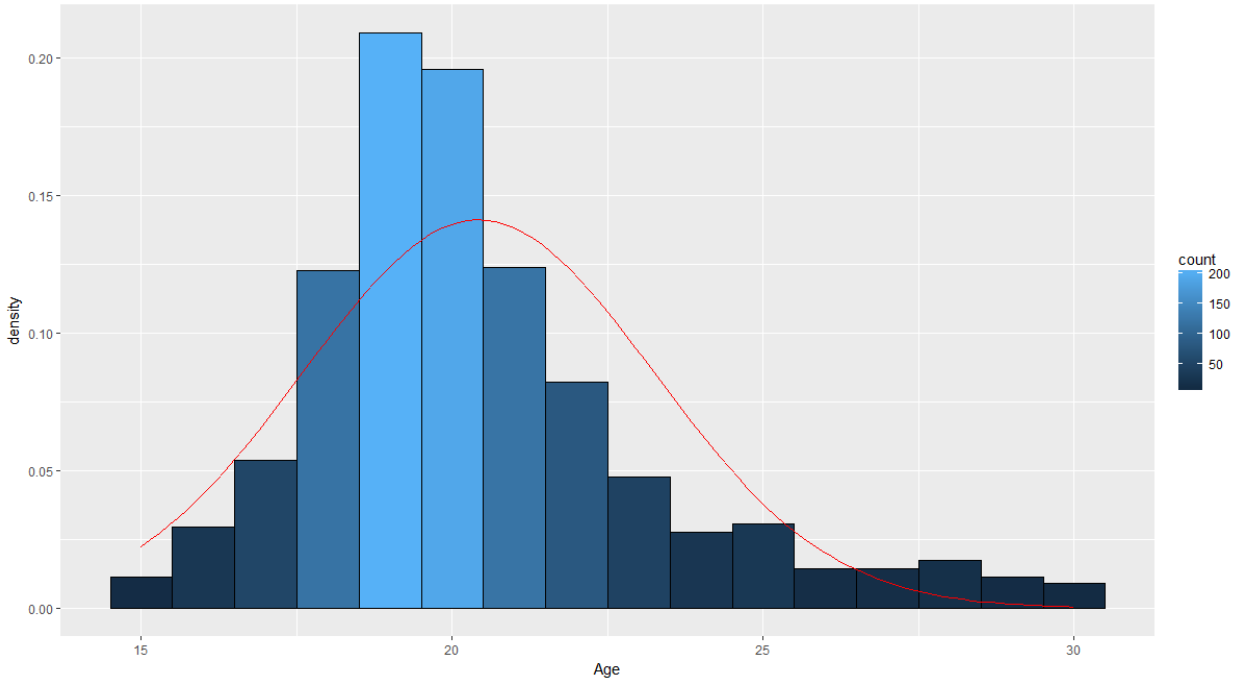

Figure 5. Age distribution histogram

The Histogram shows the distribution of the “Age” variable in the dataset. The distribution is slightly skewed to the right, no symmetrical, and leptokurtic. The mean value (*20.41827*) is different than the median (*20*) and the mode (*19*). The standardized values for skewness (*12.67*) and kurtosis (*17.12*) fall way off the acceptable range of values -2 to 2.

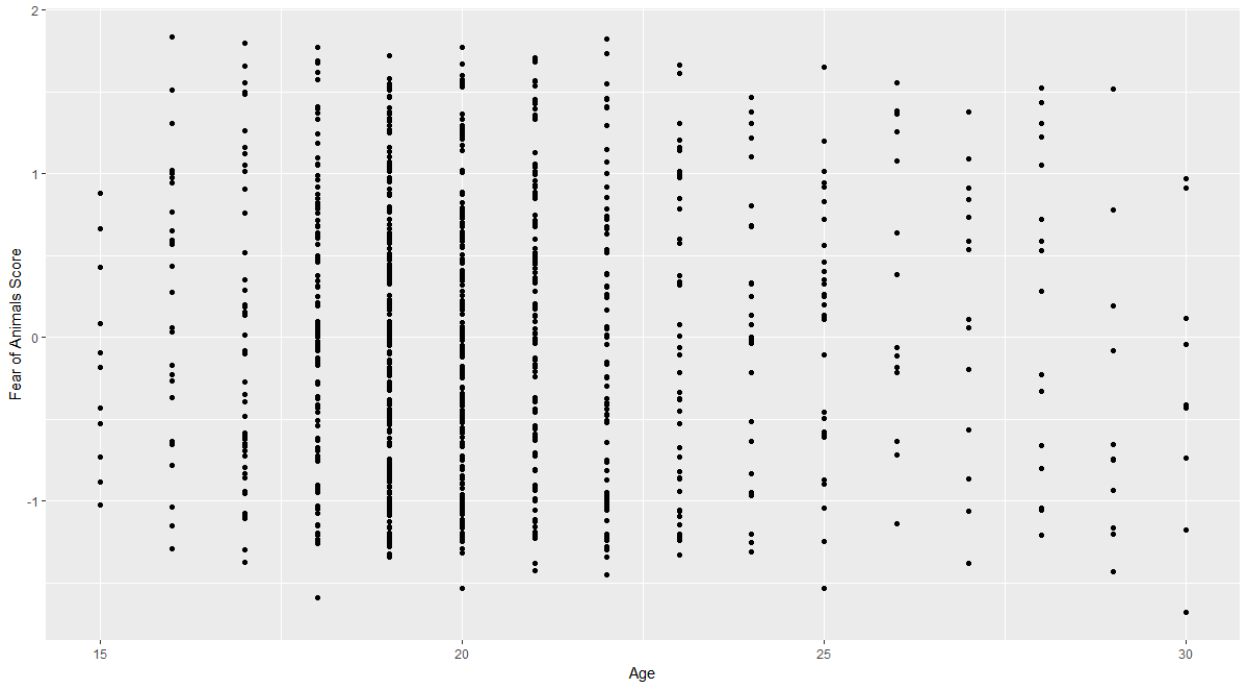

Figure 6. Scatterplot of the relationship between "Age" and "Fear of Animals".

The scatterplot shows no evidence of the existence of correlation between "Age" and "Fear of Animals" variables in the current dataset.