

Configuration Manual

MSc Research Project
Programme Name

Diana Lorena Gutierrez
x21104263@student.ncirl.ie

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Diana Lorena Gutierrez Ramirez.....

Student ID: x21104263@student.ncirl.ie.....

Programme:MSc in Fintech **Year:**2022/2023.....

Module:Research Project

Lecturer: Noel Cosgrave.....

Submission Due Date:17 September 2023.....

Project Title: ...Investigating the Influence of Fintech Solutions on Pension Fund Savings Growth in Mexico

Word Count:538..... **Page Count:**5.....

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: 

Date:13/08/2023.....

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	X
Attach a Moodle submission receipt of the online project submission , to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project , both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Diana Lorena Gutierrez
x21104263@student.ncirl.ie

1 Configuration and system requirements.

This configuration manual describes all the requirements and procedures for the reproducibility of the analysis performed in the study to ensure the validity of this research.

1.1 Hardware

Hardware/Software	Version	Description
Procesador	11th Gen Intel(R) Core(TM) i3-1125G4 @ 2.00GHz, 1190 Mhz, 4 procesadores principales, 8 procesadores lógicos	Workstation.
Memoria física installed (RAM):	8.00 GB.	
System type:	Microsoft Windows 11 Home Single Language	
Google Form		A platform on the internet for creating online questionnaires
Microsoft Excel	Microsoft 365	Software for spreadsheet that displays and manages the survey's data
R version 4.2.1 (2022-06-23 ucrt)		software for statistical computation and graphics was used to conduct the data analysis process.

2 Data

The study used primary data, which were gathered through an online survey made available through Google Forms.

3 Analysis

As part of the data analysis, the Table below shows the packages used in the code.

R packages	Description
<code>install.packages("Amelia")</code>	Used for imputation of missing data in data sets.
<code>install.packages("ppcor")</code>	to calculate and visualize different types of partial correlations between variables.
<code>install.packages("nFactors")</code>	Used to perform factor analysis.
<code>install.packages("GPARotation")</code>	Provides tools to perform rotations in factor analysis.
<code>install.packages("polycor")</code>	Used to calculate different types of polychetic correlations.
<code>install.packages("ggcorrplot")</code>	Useful for visualizing correlation matrices using the ggplot2 library.
<code>install.packages("corrplot")</code>	Corrplot is also used to display correlation matrices,

R Library	Version
library(readxl)	1.4.1
library(dplyr)	1.0.10
library(forcats)	0.5.2
library(Amelia)	1.8.1
library(ggplot2)	3.4.2
library(psych)	2.3.6
library(ppcor)	1.1
library(nFactors)	2.4.1.1
library(lattice)	0.20.45
library(MASS)	7.3.57
library(boot)	1.3.28
library(GPArotation)	2023.3.1
library(polycor)	0.8.1
library(ggcorrplot)	0.1.4
library(corrplot)	0.92

3.1.1 Preprocessing data.

1. The data collected from the survey was displayed using Microsoft Excel and an initial preprocessing was performed by changing the variable names as well as the question names to determine the correlation table required to run the Chi-square test.

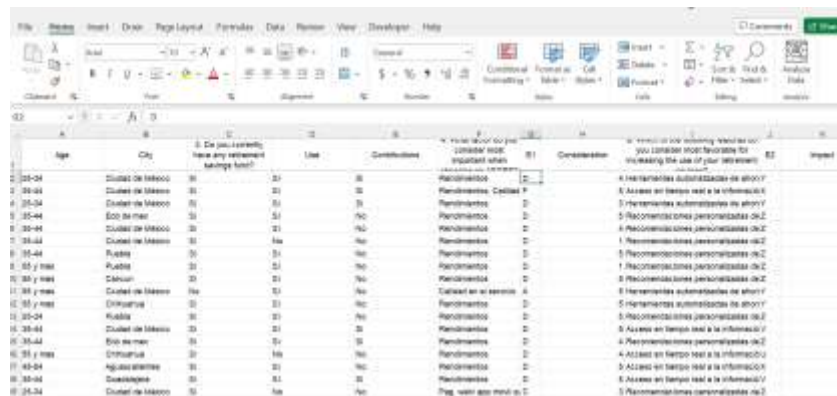


Figure 1. Transformation using replace function in Excel.

2. The pre-processed data was then imported into RStudio for further preparation which involved checking for:
 - a. Missing values using the Amelia package.
 - b. 'dplyr' package to inspect the demographic profile of the respondents (Age and City).
 - c. Frequency was obtained by using 'ages_counts' and 'city_counts'.
 - d. ggplot2 package was used to graphically display the demographic profile.

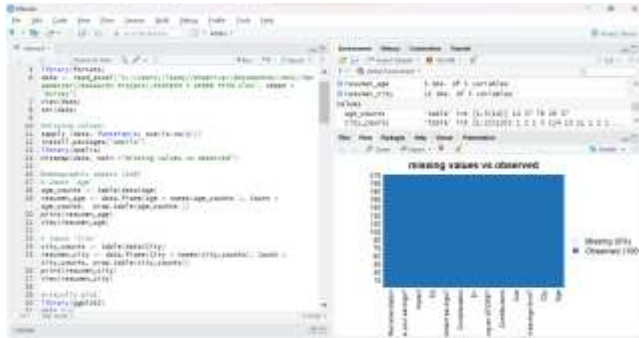


Figure 2. Exploratory Analysis

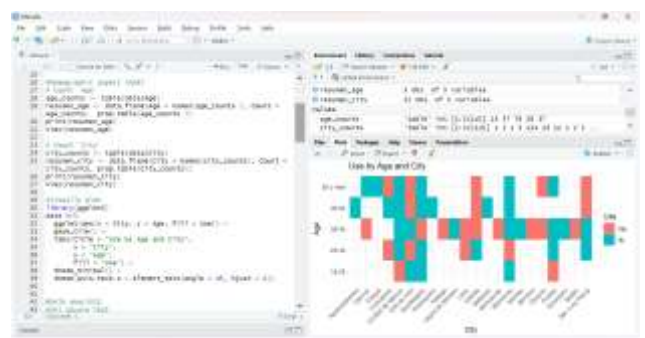


Figure 3. Frequency of demographic aspects

3.2 Chi-square Test

Subsequently, a new data frame was created: 'data' to evaluate the established hypotheses. The `chisq.test()` function in R was used for this process.

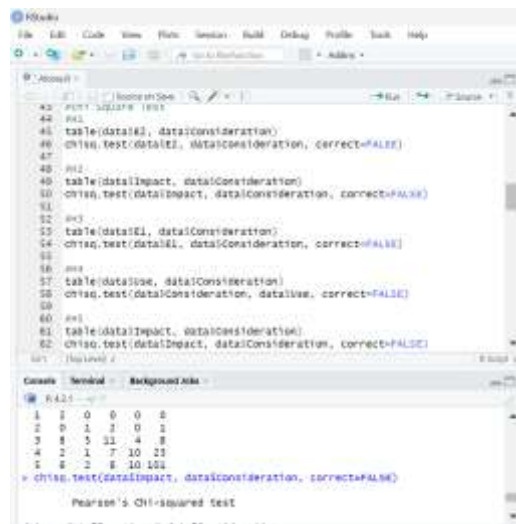


Figure 4. Chi-square test of the hypothesis

3.3 Factor Analysis

- After obtaining the results of the Chi-squares, the database was reloaded, in which the questions were also renamed in Excel, and all the missing answers were parameterized to have a complete exploratory factor analysis.
- A data frame called 'dataFA' was created again and once the variables that were below what was established by the MSA (less than 50) were identified, the correlation matrix was run with the `View()` code.

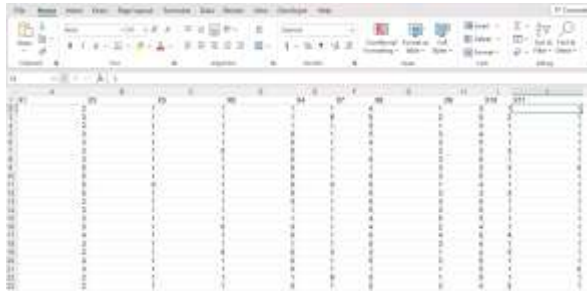


Figure 5. Renamed variables in Excel.

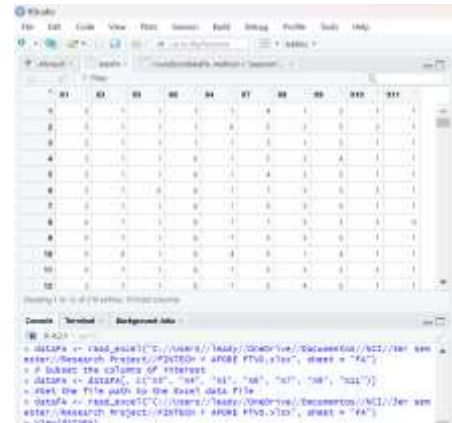


Figure 6. Validation of the updated data

- As well as the generation of the correlation plot, Bartlett's Test for Sphericity and KMO were also run on the same data frame using `cortest.bartlett()` and `kmo()` respectively and run the plot for determining the number of factors `plotnScreen()`.

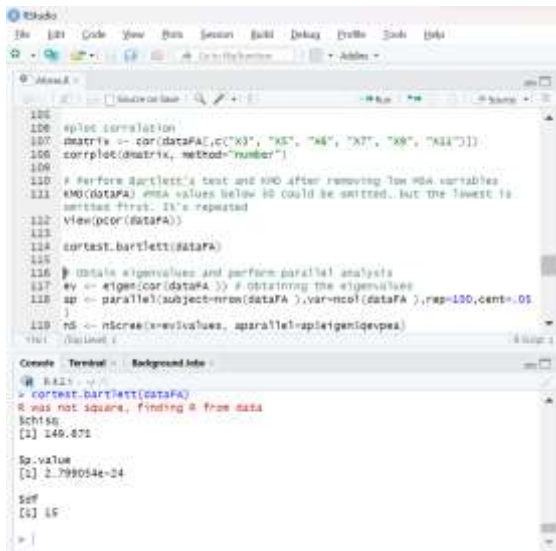


Figure 7. Bartlett's Test for Sphericity.

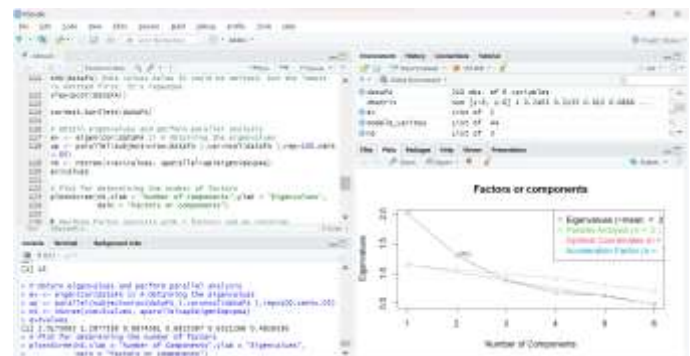


Figure 8. Parallel Plot

- Using the two factors based on the data obtained using `factanal()`.
- To provide a better understanding and graphical interpretation, the function different rotations were created biplots using `(fa(dataFA, nfactores = 2, fm="minres", rotate = tipo), main = paste("Biplot con rotación ", tipo), col=c(2,3,4), pch = c(21,18))`
`}`
`par(mfrow = c(2, 2))`
`sapply(rot, bi_mod)`



Figure 9. Varimax, Quartimax, Promax Plot Rotation

- Finally, using the function fa() using the varimax we can further facilitate its interpretation. With the latest data obtained, no further analysis is required.

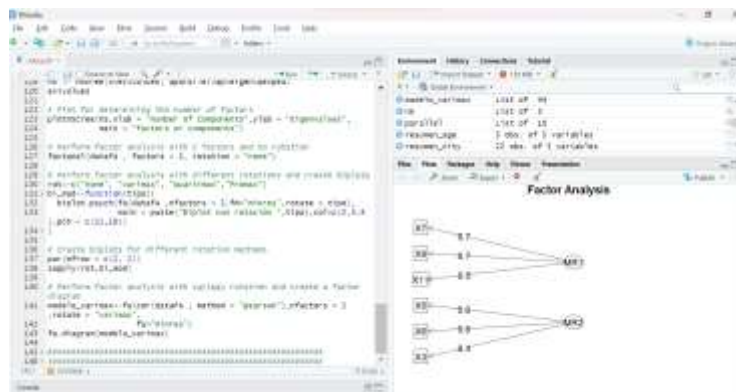


Figure 10. Plotting the Factor Analysis chart

References

Hair, J.F., Black, W.C. and Babin, B.J. (2019). *Multivariate Data Analysis*. Andover, Hampshire, United Kingdom: Cengage Learning Emea. Copyright.

Appendix

Link of the survey:

https://docs.google.com/forms/d/e/1FAIpQLSc0Wd1c9j0a7pYRzEwbEFmC8EVP03T1h6iTyHBogpfaD2XCw/viewform?usp=pp_url