

# Configuration Manual

MSc Research Project  
Programme Name

Utsoritselaju Koloh  
X20243821@student.ncirl.ie

School of Computing  
National College of Ireland

Supervisor: Brian Byrne

**National College of Ireland**  
**MSc Project Submission Sheet**  
**School of Computing**



**Student Name:** Utsoritselaju Koloh.....  
**Student ID:** x20243821.....  
**Programme:** MSc Fintech ..... **Year:** 2022.....  
**Module:** MSc Research Project.....  
**Supervisor:** Brian Byrne.....  
**Submission Due Date:** 15/08/2022.....  
**Project Title:** A Study on the impact of Remote Work on Internal Control Effectiveness in Ireland's Financial Sector.....  
**Word Count:** 1305..... **Page Count** 12.....

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:** Utsoritselaju Koloh .....  
**Date:** .....

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

Attach a completed copy of this sheet to each project (including multiple copies)	
<b>Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).</b>	
<b>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.</b>	

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

<b>Office Use Only</b>
Signature:
Date:
Penalty Applied (if applicable):

# CONFIGURATION MANUAL

Utsoritselaju Koloh  
X20243821

## 1 Introduction

This document is submitted as a part of the Research Project Report for the MSc in Fintech Module. The document describes each phase of the research carried out in the study including specifications of the machine, software and hardware of technologies used to achieve the study outcomes. In the case reproducibility, this configuration manual seeks to guide other researchers to accomplish objectives.

## 2 System Specification

### 2.1 System Hardware

The research project was carried out using an Apple MacBook Air. More specifications of the machine are provided below:

Version: macOS 12.0.1  
Processor: Dual-Core Intel Core i5 @  
Processor Speed: 1.6 GHz  
Memory: 8GB  
Storage: 256 GB

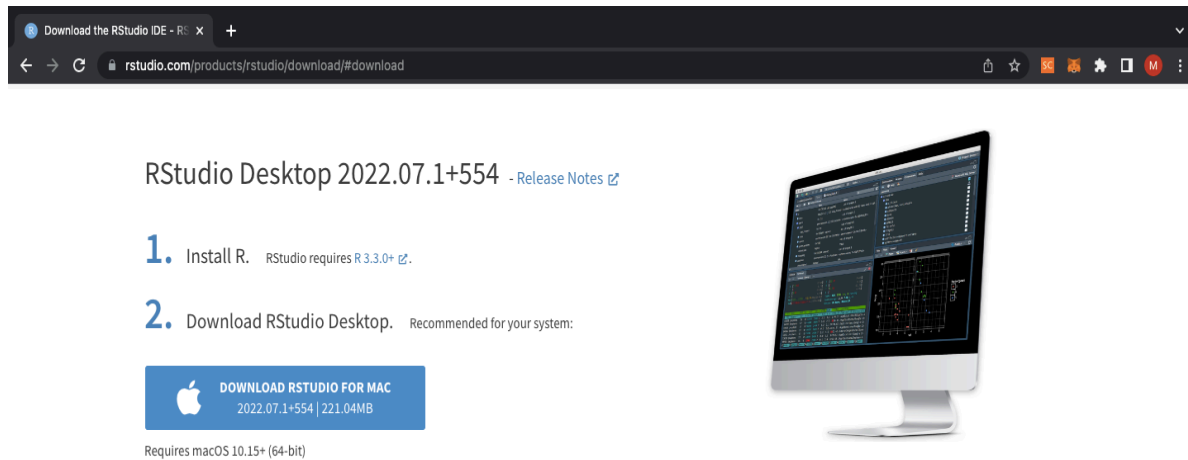
### 2.2 Tools and Software used for Analysis

- **Google Forms:** The questionnaire for the study was designed online using this tool. The tool enabled questions to be sent conveniently to the gather responses from the target population. The information gathered from respondents were automatically collated in an online spreadsheet. The collated information was extracted as a secure csv file into Microsoft Excel for further analysis.
- **Microsoft Excel for Mac:** Version 16.63.1 of Excel was used to analyze the extracted CSV file. No further analysis was done using Microsoft excel.
- **R programming Software:** The techniques used to analyze the data gathered for this research was carried out using the programming language called R. Version 4.1.2 for R studio was used to execute all statistical techniques that were reported in the study and they include Descriptive Statistics; Boxplots and Ggplots, Exploratory Staistics; Correlation Heat maps and Multiplot Ggplots , Inferential Statistics; Shapiro Wilk-Test, Regression Analysis, Spearman Rank Correlation-Coefficient Test, One-Way ANOVA Test, Pairwsie T-Test.

## 3 Installing the Software

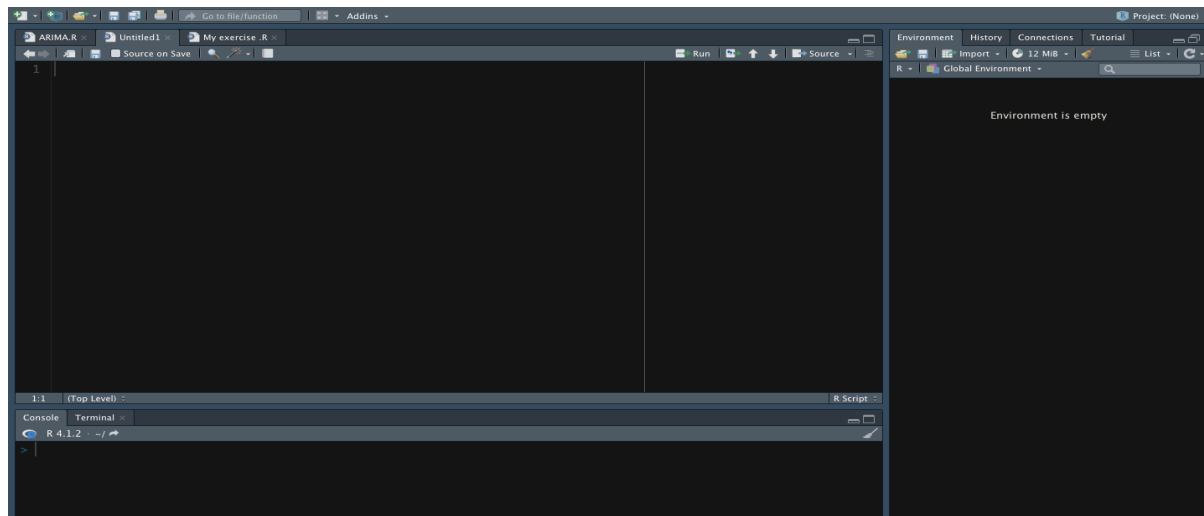
### 3.1 R Studio

R programming language is a major tool for data analytics and research. It has a vast library that contains several statistical models and visualization tools for data analysis. The software provides various statistical models, most of which were used in the study. Figure 1 shows the process of downloading R Studio:



**Figure 1. R Studio Download**

The raw file is first downloaded from the internet after which it is then converted into the app and successfully installed. The version for R studio is 4.1.2 and can be easily assessed from the Mac Launchpad. The R script/file containing all codes used for performing the statistical tests and the study report are both submitted on Moodle. In Figure 2 is a snapshot of R Studio



**Figure 2. R Studio**

## 4. Data Collection

The data used for the study was quantitative and was collected by sending online surveys to respondents.

## 5. Consent Form

The first section of the online survey contained a brief description of the research, ethics validation, declaration, and consent for respondents who decide to participate in the study. All respondents approved their consent to participate in and the data was extracted from the online spreadsheet to a csv file.

## 6. Reading the Excel csv in R Studio

The data downloaded from the online spreadsheet was stored as a csv file in Excel and imported to R studio using a code. Several statistical tests were conducted after the data was imported into R Studio.

## 7. Installing Packages

Packages were installed into R studio using the `install.packages()` function. The function `library()` was used to load in the installed package. The following packages were installed for the analysis:

- Plm - For estimating quantitative variables in the linear regression model in R
- Pheatmap – This function is used to cluster heat maps in R and has a wide parameter for graphics control like the ggplot. It was used to draw the correlation heatmap
- Psych – The package was used for descriptive statistics in R to provide summary statistics of the ordinal data using the `psych::describe()` function. The `describeBy()` function was used to group a variable and report summary statistics
- Rsample and Rstatix packages were used to provide additional parameters for drawing the correlation heatmap
- Dplyr – Streamlines data manipulation. The function `sep()` from the package was used to read in data
- The packages Ggpubr, Ryouready and Ez were installed to test assumptions for the statistical tests conducted

## 8. Missing Values

The function `sum(is.na())` was used to check for missing values and 19 columns were found to contain NA values. Respondents provided no answer to a question. NA values were assumed by the study to be neutral since respondents skipped the question.

## 9. Data Cleaning

Since The study used a Seven point Likert scale, columns that contained missing values were replaced with the Neutral value of chosen scale '4'. The imputation was computed using the function `x[is.na(x)] <-x`. Outliers were discovered using boxplots and were treated using the function `winsor()`. The questions were renamed in R to avoid text overlapping and allow for easier analysis

## 10. Techniques for Analysis

- **Descriptive statistics**

Summary statistics are obtained from nominal and ordinal variables in the dataset using codes in R. Nominal data includes demographic information of respondents and multiple choice questions relating to Internal Control Effectiveness while ordinal data includes Likert scale questions that measure Internal Control Effectiveness, both of which are discussed in the paper. The codes and visualizations are included in R script that were submitted on Moodle

### Descriptive statistics by Group

The psych package was installed for descriptive statistics and the `psych::describe()` function was used to provide summary statistics of the metric scales. The `describeby()` function was used to provide summary statistics by grouping the work arrangement variable to derive the mean and standard deviation for each

Table 1: Descriptive Statistics by Working Arrangement Group

Statement	Fully Remote		Hybrid		On-Site Office	
	Mean	Sd	Mean	Sd	Mean	Sd
Existing fraud controls have been effective in managing occupational fraud since the start of the COVID-19 pandemic till date?	5.18	1.14	5.19	1.18	5.50	1.14
Identification process effectively identifies internal risk your organization is exposed to.	5.41	0.96	5.51	1.22	5.88	1.03
The audit committee effectively performs their duty to monitor internal controls	5.91	0.75	5.95	0.72	5.96	0.71
Your department manager explicitly describes your departments obligation in managing its risk profile.	5.41	0.80	5.72	1.01	6.00	0.78
You understand your obligation to manage the fraud risks that your department is exposed to	5.23	0.87	5.79	0.91	6.00	0.59
Whistleblowing process ensures the confidentiality of the employee who reports suspected fraud	4.68	1.13	4.91	1.54	4.96	1.37
Fraud perpetrating employees are appropriately punished when caught in your organization.	5.03	1.06	4.70	1.25	4.88	1.23



DATA/ANALYTICS																
YES	10	11.23	4	4.494	9	10.11	5	5.617	4	4.494	4	4.494	8	8.988	5	5.617
NO	-	-	6	6.741	1	1.123	4	4.494	6	6.741	6	6.741	1	1.123	4	4.494
I DON'T KNOW	-	-	-	-	-	-	1	-	-	-	-	-	1	1.123	1	1.123
FUNDING/TRADING																
YES	4	4.49	1	1.123	2	2.247	2	2.247	3	3.370	1	1.123	4	4.494	4	4.494
NO	-	-	3	3.370	2	2.247	2	2.247	1	1.123	3	3.370	-	-	-	-
I DON'T KNOW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
INSURANCE/INSURTECH																
YES	6	6.741	1	1.123	4	4.494	5	5.617	3	3.370	5	5.617	3	3.370	6	6.741
NO	1	1.123	5	5.617	3	3.370	1	1.123	4	4.494	2	2.247	3	3.370	-	-
I DON'T KNOW	-	-	1	1.123	-	-	1	1.123	-	-	-	-	1	1.123	1	1.123
PAYMENTS																
YES	8	8.988	9	-	7	7.865	5	5.617	5	5.617	6	6.741	7	7.867	8	8.988
NO	1	1.123	-	10.11	2	1.123	4	4.494	4	4.494	2	2.247	2	2.247	1	1.123
I DON'T KNOW	-	-	-	-	-	-	-	-	-	-	1	1.123	-	-	-	-
REGTECH/COMPLIANCE																
YES	8	8.988	3	5.617	7	7.865	8	8.988	3	3.370	5	5.617	7	7.867	3	3.370
NO	0	-	5	3.370	1	1.123	-	-	5	5.617	3	3.370	1	1.123	5	5.617
I DON'T KNOW	-	-	-	-	-	-	-	-	-	-	-	-	4	4.494	-	-
SAVINGS/INVESTMENT																
YES	3	3.370	2	2.247	4	4.494	3	3.370	2	2.24	3	3.370	4	4.494	1	1.123
NO	1	1.123	2	2.247	-	-	1	1.123	2	2.24	1	1.123	-	-	3	3.370
I DON'T KNOW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.123

- **The Shapiro-Wilk Test**

This test was performed in the study to test normality of the data. The results from the test are reported and discussed in the result section of the paper. The Stats package was installed for the Shapiro-Wilk test and the function `shapiro.test()` was used to conducted the test in R studio. Microsoft Word was used to summarize results in tabular form.

- **Spearman Rank Correlation-Coefficient Test**

The Spearman rank Correlation was test did not require a package to be installed in R. It is one of the statistical models that is included in R base. The `cor.test()` function was conducted to derive the results for measuring the strength and direction of the relationship between two ranked variables. . The results from the test are reported and discussed in the result section of the paper.

- **Regression Analysis**

A linear regression was conducted in the study using the `lm()` function in R studio and the plm package was installed to conduct this test. This analysis was used to test for the significant impact of X and Y.

- **One-Way ANOVA or Analysis of Variance Test**

In this study, one-way ANOVA was conducted in R to investigate if a single dependent variable differs among groups. The stats package was installed and the `aoov()` function was used for the test.



- **Pairwise T-Test**

The Mosaic package was installed to conduct the pairwise t-test. The test was used to pairwise compare variables grouped by their working arrangement with Internal Control Effectiveness. The `pairwise_t_test()` function was used in R to conduct the test.

## **11. Analysis of Existing Internal Controls**

This chapter analyses categorical variables of ICE. The `ggplot()` and `table()` function are used to analyze the results of the variables. The summary statistics are presented in this configuration manual using tables from Microsoft word and statistical plots from R studio

The study gathered responses across 10 organization types in Irelands financial sector as shown in Table 3

## **12. Conclusions**

This document details technology used for the analysis of the study, steps of the research and additional analysis on internal control effectiveness. The study adopted an online survey to gather primary data for evaluation. This configuration manual shall aid researchers reproduce the result of this study in the future.