

Configuration Manual

MSc Research Project
MSc In Fintech

Habil Shihoriwala
Student ID: X21197962

School of Computing
National College of Ireland

Supervisor: Theo Mendonca

National College of Ireland
MSc Project Submission Sheet



School of Computing

Student Name: Habil Shihoriwala

Student ID: X21197962

Programme: MSc In Fintech **Year:** 2022-2023

Module: MSc In Research

Lecturer: Theo Mendonca

Submission Due Date: 14th August '23

Project Title: Empowering the Crowd: Equity Crowdfunding's contribution to Startup Funding and Retail Engagement

Word Count: 1447 **Page Count:** 6

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: Habil Firozbbhai Shihoriwala

Date: 14th August '23

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
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

Habil Firozbbhai Shihoriwala
Student ID: x21197962

1 Introduction

The document highlights the specification of hardware, software used in data collection. The software consisting of R programming and data set collection done from the website. The manual covers the packages used, in order to create data visualization plots and machine learning analysis. It also covers the steps required to reproduce the results of the research “Empowering the crowd: Equity Crowdfunding’s contribution to startup funding and retail engagement.”

2 System Requirements

2.1 Hardware

Laptop : HONOR MagicBook X14
Processor : Intel(R) Core (TM) i5-10210U CPU @ 1.60GHz 2.11 GHz
RAM : 8GB
WIDOWS: WINDOWS 11

2.2 Software

- R studio 2022.07.01 for data visualization and machine learning analysis
- Microsoft word for report writing and configuration manual
- Microsoft excel for data sets used
- Google forms for survey creation

3 Data

The data set that is collected is from a US based Equity crowdfunding website known as Kingscrowd.com where we download datasets for 5 crowdfunding platforms based on their valuation and one more dataset consisting of campaign data of those platforms. The total campaigns are 4144. The data of the 1st data set of platforms was unstructured and we converted it tabular format while the other one consisting of campaigns had to be cleaned

processed and transformed. The last data we obtain is using a survey questionnaire created on google forms with a sample size of 50 participants.

1

4 Function

In R studio we install the relevant packages below:

Purpose	Function used
To find correlation between variables	Cor
To find the regression	Lm
To load training and testing data	Set.seed
Plot scatter and box plots	Ggplot , ggplot2
Logistic regression	Glm
Random forest	randomForest
Support vector machine	svm
Heatmap	Heatmap.2
Principal component analysis	Prcomp
Variance inflation factor	Car::vif

5 Steps

- We first copy the data sets and from Kings crowd platform and load it in a csv file in excel
- There are 2 data sets: Platform having 325 observations for monthly data, we first load it into R and perform data visualization by using we filter total valuation using filter and first find correlation. We find the average monthly data of total valuation vs number of raises
- Use date function to convert the dates, then use as.numeric and gsub function to treat non numeric values as well as commas
- We then perform a hypothesis test of each platform to find the p value
- After testing it at 5 percent confidence we use ggplot function to plot total valuation chart and number of raises for all platforms
- We need to convert the date as the dates are in 20180101 format using as.date function and then perform the Augmented dickey fuller test.
- We then take the second data set of the campaigns
- We first clean the file using gsub and numeric function to convert relevant columns to numeric
- We then perform data visualization by plotting charts and convert the dates in the format d-m-y then use scatter plots to plot different variables with others to make inference of our research question
- We use gsub and as.numeric function to clean two independent variables that is annual revenue and valuation/capitalization and plot its correlation.

¹ www.Kingscrowd.com-dataset

- We plot a number of charts and only selected ones that have impact on retail participation and amount raised.
- We create a multiple regression as amount raised as independent variable using lm function
- We then plot a variance inflation factor that checks for multicollinearity of independent variables using car::vif function
- The funding status is one of the dependent variables that has status as funded, not funded, active we need to convert it to binary and perform a logistic regression.
- We use set. Seed function and set the training and testing data as 70 percent and 30 percent and create a logistic regression with the confusion matrix, F1, precision, accuracy and recall value. However, we do not use it as there are many missing values.
- We play around with the data by plotting charts to check relationship with other variables and then perform logistic regression of status and investors, status and annual revenue.
- We further conduct analysis to do logistic regression once again but treat for missing values using mean function for each variable and set seed function with training 80% and 20% testing data
- We load library e1071 and randomForest, we perform random forest, support vector machine and logit model. We also use log function wherever there are valuation/capitalization, annual revenue, investors, and average daily raise.
- We plot confusion matrix, accuracy, precision, recall, F1 score, accuracy
- We then perform the random forest and support vector machine accuracy.
- Next, we create a subset with only numeric column so we use the has.na function as we want to create a correlation matrix and heatmap.
- In the end we perform Principal component analysis that is a dimensional reduction technique using prcomp function to check the relevant features that represent our point and print it.
- Please run the codes one after another as they are in sync with the above ones.
- As for the secondary validation-Survey, we load 19 questions in google forms and float it on LinkedIn for a sample size of 50

5.1 Survey questions:

Your name?
What is your age group?
How would you describe your investing experience?
What asset classes do you invest in?
Do you know anything about Equity crowdfunding ?
If given a chance to invest in startups, would you invest your money on a platform in exchange for equity shares just like seen on shark tank with the help of Equity crowdfunding?

How would you decide to invest in startups? Based on what parameters? What parameters you think are most relevant?
If you decide to invest in a startup through any of the platforms, on what basis would you choose the platform?
If you had a chance to invest in a company for as low as \$100 for its equity would you go for it?
Do you think equity crowd funding has allowed general public to participate in the growth phase of startups which was initially restricted to angel investors?
Do you think equity crowdfunding would help participation of general public in return for equity of a company around the world?
If you have participated in equity crowdfunding, what motivated you to invest through this platform?
In terms of Equity Crowdfunding Impact In your opinion, how has equity crowdfunding impacted the overall funding landscape for startups?
Would you recommend equity crowdfunding as a viable option for retail investors to participate in startup funding?
Are you a startup entrepreneur seeking funding?
If yes, have you considered using equity crowdfunding as a funding option?
You have an idea and a solid business plan; how would you want to raise funds?
As an entrepreneur what are the reasons for not considering equity crowdfunding?
Do you think that raising funds by the way of equity crowdfunding will see an upward trend in future?

References

KingsCrowd. (2023, May 6). *Home - KingsCrowd*. <https://kingscrowd.com/>