

Configuration Manual

MSc Research Project
Programme Name

Girish kandan Murugan
Student ID: X19199791

School of Computing
National College of Ireland

Supervisor: Jorge Basilio

National College of Ireland
MSc Project Submission Sheet
School of Computing

Student Name: Girish kandan Murugan

Student ID:x19199791.....

Programme:.....M.sc Data
Analytics..... **Year:**2021.....

Module:

Supervisor:JORGE BASILIO.....

Submission Due Date:16.08.2021.....

Project Title:Configuration
Manual.....

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:

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

Girish kandan Murugan
X19199791

1 Introduction

This Configuration Manual contains all of the criteria and specifics for our research " Creation of a recommendation system to recommend cryptocurrency portfolio using Association rule mining." This manual consists of the following sections. The hardware requirements are shown in the second section. The third section discusses all of the software needed to complete this project. The fourth section explains where to get the data. The project's library module needs are indicated in the fifth section. The final section is the Exploratory data analysis section. With this the manual will be concluded.

2 Hardware Configuration

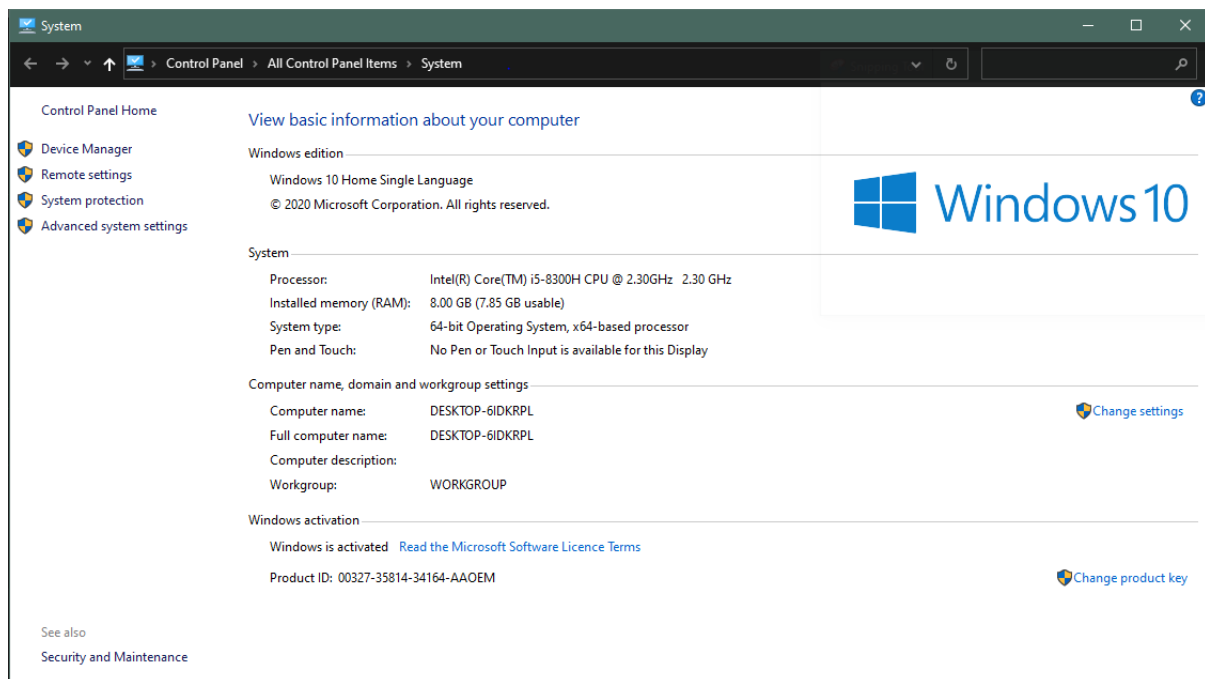


Figure 1 Hardware Configuration

OS	Windows 10 home single language
Installed Memory	8 GB RAM
Hard disk (HDD)	1 TB
Processor	Inter® Core™ i5-8300H CPU @ 2.30GHz
Software's & Languages	Python 8 or above, Jupyter notebook (anaconda3), Google collab, VS studio

3 Software Requirements

Python 3.8 was used to create the entire script. Because the entire program was written in Jupyter Notebook, which is available with Anaconda Navigator and for this Anaconda Navigator is necessary too.

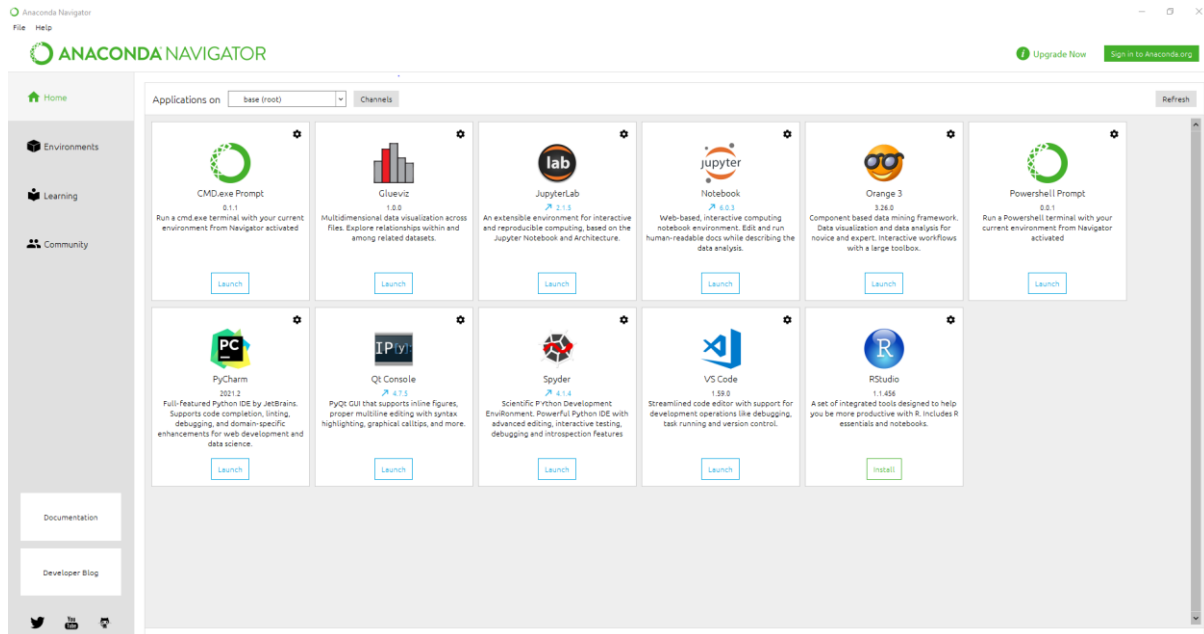


Figure 2 Hardware Requirements

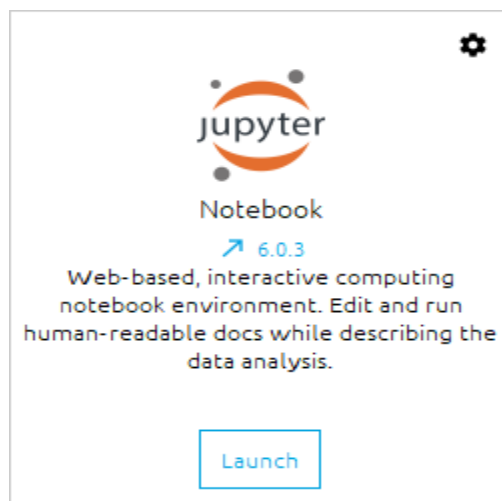


Figure 3 Jupyter Notebook

4 Dataset Source

Our dataset is an Historical dataset and it was downloaded from the Kaggle site which is available for public usage for free.

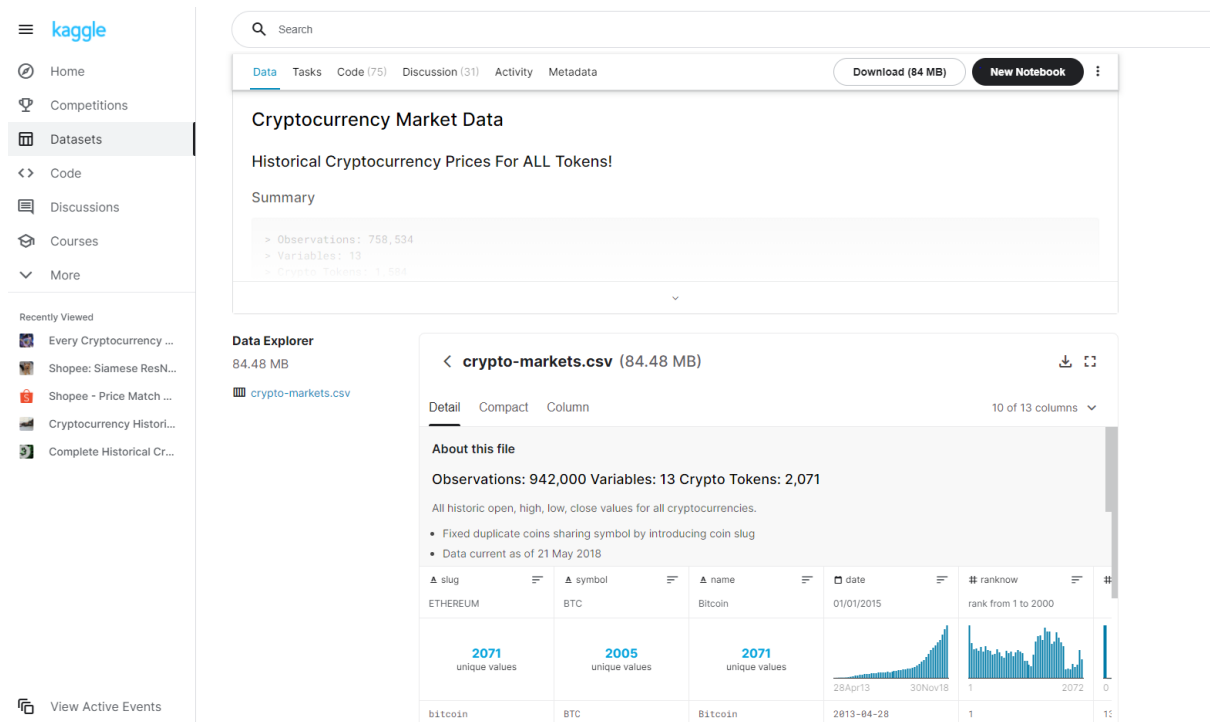


Figure 4 Dataset Source

5 Library installation

There are 4 python libraries which are required for this research.

Required library packages

- pandas
- numpy
- apriori-python
- Matplotlib

These can be installed using the following commands in jupyter notebook.

```
pip install pandas
pip install numpy
pip install apriori-python
pip install matplotlib
```

6 EDA

Exploratory data Analysis very important before starting any project to know well about the data set we have. The following are the analysis we did with our dataset.

Number of columns are displayed and their data type

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 942297 entries, 0 to 942296
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   symbol          942297 non-null  object
1   date            942297 non-null  object
2   open            942297 non-null  float64
3   high            942297 non-null  float64
4   low             942297 non-null  float64
5   close           942297 non-null  float64
6   volume          942297 non-null  float64
7   market          942297 non-null  float64
8   close_ratio     942297 non-null  float64
9   spread          942297 non-null  float64
dtypes: float64(8), object(2)
memory usage: 71.9+ MB
```

Min and Max of columns where calculated.

```
print('Min, Max of cols')
for col in data_columns:
    if data[col].dtype != 'object':
        print(
            col,
            '\tMin',
            np.min(data[col].values),
            'Max',
            np.max(data[col].values)
        )

Min, Max of cols
open      Min 2.5e-09 Max 2298390.0
high      Min 3.2e-09 Max 2926100.0
low       Min 2.5e-10 Max 2030590.0
close     Min 2e-10 Max 2300740.0
volume    Min 0.0 Max 23840899072.0
market    Min 0.0 Max 326502000000.0
close_ratio  Min -1.0 Max 1.0
spread    Min 0.0 Max 1770563.0
```

Correlation matrix for columns were produced.

```
print('Corelation matrix')
print(data.corr())
```

```
Corelation matrix
```

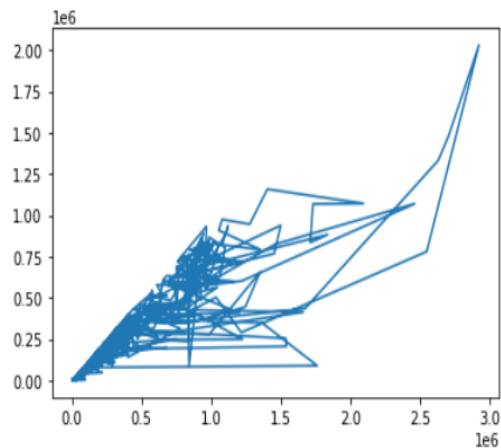
	open	high	low	close	volume	market
open	1.000000	0.968123	0.971673	0.957178	0.011051	0.012543
high	0.968123	1.000000	0.947305	0.971748	0.009241	0.010461
low	0.971673	0.947305	1.000000	0.972246	0.012803	0.014673
close	0.957178	0.971748	0.972246	1.000000	0.011124	0.012659
volume	0.011051	0.009241	0.012803	0.011124	1.000000	0.880793
market	0.012543	0.010461	0.014673	0.012659	0.880793	1.000000
close_ratio	-0.000631	0.000256	0.000875	0.003756	0.009710	0.010468
spread	0.741319	0.856532	0.646075	0.749033	0.001392	0.001285

	close_ratio	spread
open	-0.000631	0.741319
high	0.000256	0.856532
low	0.000875	0.646075
close	0.003756	0.749033
volume	0.009710	0.001392
market	0.010468	0.001285
close_ratio	1.000000	-0.000800
spread	-0.000800	1.000000

Graph plotting done for the high and low price rates available in the dataset.

```
plt.plot(data["high"], data["low"])
```

```
[<matplotlib.lines.Line2D at 0x133dd79a040>]
```



7 References

Python Release Python 3.8.5 (no date) Python.org weblink :

<https://www.python.org/downloads/release/python-385/>

Anaconda | Individual Edition (no date) Anaconda weblink:

<https://www.anaconda.com/products/individual>