**CE101 Team Report Assignment**

**Team Number - A54**

**Team Members- Lewis Johnson, Arthur Fitzjohn, Arnav Ghosh, Yuhang Xin, Patrik Voicechovski and Bahir Usanmaz.**

### CSEEJira Project URL: <https://cseejira.essex.ac.uk/projects/CE101T54/summary>

Table of Contents

CSEEJira Project URL: Include link to your team’s jira project 1

Chapter 1 The Executive Summary (?? words) 3

Chapter 2 Team Working (?? words) 4

2.I An introduction to Team Working 4

2.II Team Activity Report 4

2.II.a Team's Scrum Meetings, Sprints 4

2.II.b Detailed report of each team members contribution to the project 4

Chapter 3 Product Development (?? words) 5

3.I An introduction to Product Development 5

3.II The Context 5

3.II.a The Customer/ End User 5

3.II.b Legal and Ethical Matters 5

3.II.c Sustainability 5

3.II.d Health and Safety Matters 5

3.III The Team Product 6

3.III.a Product Description 6

3.III.b Product Demonstration 6

Chapter 4: Project Management (?? words) 6

4.I An introduction to Project Management 6

4.II Project Management Report 6

4.II.a Summary of Teams Project management in Jira 6

4.II.b An evaluation of the Project management 6

Chapter 5: Conclusions (?? words) 7

Bibliography 7

# Chapter 1 The Executive Summary (?? words)

# Chapter 2 Team Working (?? words)

## 2.I An introduction to Team Working

## 2.II Team Activity Report

### 2.II.a Team's Scrum Meetings, Sprints

### 2.II.b Detailed report of each team members contribution to the project

# Chapter 3 Product Development (?? words)

## 3.I An introduction to Product Development

#### Product Development

“Product development” refers to the method you use when developing or modifying a product. Many different methodologies have been developed to aid companies when deciding how best to break down and organise the development of their product. Developing a product without first planning how will usually result in very inefficient use of time and resources as well as unhappy customers as the final product may not be exactly what they wanted.  
Methodologies break the problem down into smaller issues, allowing them to be resolved by individual teams.

#### Waterfall Methodology

The tasks within the methodology are typically only moved through one way, sequentially. The tasks are usually ordered similar to: “Requirements – Design – Implementation – Verification – Maintenance”. This sequential design usually makes things costly if a mistake was made at any stage. For example, if improper analysis was done in the requirements stage, the whole project may have to be scrapped and started anew because the product is not even attempting to do what the customer wants. Similarly, if the customer's needs change midway through the project, there is nothing in the waterfall methodology to facilitate this change and it may result in a complete restart. Therefore, this methodology should only be used when the development team is incredibly clear on what is required by the customer (usually if the developers are creating something for themselves) and if the sudden need for change is extremely unlikely (usually if the expected development time is very short).

#### Agile Methodology

Agile methodologies work more iteratively than the waterfall model. The problem is usually separated into the same tasks, but at least some of the tasks form a loop as the customer is shown prototypes of each aspect of the product at the end of each loop. If the customer is unhappy with the product, then another analysis will be made of the customer's exact requirements and the loop will begin again. This will be repeated until the customer is happy with the product.  
The tasks in agile methodologies are usually ordered similar to: “Requirements – Design – Implementation – Verification”, which will loop, with maintenance being a continual process after the customer has bought the product.

#### Scrum

“Scrum” is a framework that is commonly used with agile methodologies as it facilitates short bursts of production that are periodically monitored and can therefore be adjusted to facilitate the needs of the customer / the ideas of the product owner. For our project we have received information on which features are required from Kaggle, though we have our own supervisors and Vishuu for queries regarding the project.

The “product owner” is the person who knows how the final product should be and is therefore to be questioned about specific design choices, but does not usually have any direct interaction with product development (like programming or construction of physical components).

The “scrum master” is in charge of setting the workload for individual developers and scheduling meetups as well as making sure that everyone has a clear idea of what they have been instructed to do. The scrum master is often also a part of the scrum team. For our project the scrum master has been elected on a bi-weekly basis as instructed.

The scrum team is a small group of people who work together to complete their tasks within the time limit set by their scrum master. This is called a sprint. For our project all members of our group were set tasks during each scrum and worked together to complete them. We used JIRA for simple queries but the messenger app proved more useful for open discussions regarding work and meetups.

Our sprints will usually be two weeks long with a group session lasting 2 hours at the end of each sprint. During this time we will decide our scrum master and decide the workload for each member after reviewing the work completed during the last sprint.

## 3.II The Context

### 3.II.a The Customer/ End User

#### Supervisors

We need to make a product that the supervisors approve of and we are showing the work that we are doing by using Jira, Github and Kaggle.

#### Team members

We need to make a product that helps us learn how to use object oriented programming in data science. This will help us in our future study/ careers to analyse datasets and write reports about them.

#### Estate agent professionals

Estate agents can use the algorithms from this project to predict future house prices in individual areas whilst also showing which features are most likely to influence house price negotiations. This allows them to make the best decisions on which houses to buy and sell as well as which features to showcase when marketing their buildings.

As the training data used for the project is from Iowa only, the outputs of the algorithm submitted for the assignment may not be hugely applicable anywhere else as the general opinions of the public when it comes to housing may vary greatly from area to area or country to country. However, the algorithms to create these predictions should be usable in any area as long as a large enough quantity of accurate data has been entered.

### 3.II.b Legal and Ethical Matters

The technologies that we are using in this project are the programming language Python 3.7 and Jupyter Notebook.

#### Python

Python is an open source technology which means that we are able to use it for educational purposes without needing to pay for it [1]. It is a programming language and we use its syntax to control the flow of operations of our program.

#### Jupyter Notebook

Jupyter notebook is programming software which allows us to write markup between blocks of code that can be run individually. It is an open source project [2].

#### Ethical issues

We do not believe that our project poses any risk to privacy or other sensitive information as data used in our algorithms has no need of any details connecting the houses to individual people so they should not be collected and will not be used to create the ouput. We also do not think that our project could affect society (presently or in the future) in a negative way as it only empowers estate agents which has potential to improve the economy.

### 3.II.c Sustainability

Having read through the principles of DEFRA, we would like to declare that this project does not have significant effect on the environment, justice, health, science or governance. We do, however, predict this project to have a positive effect on the economy because more informed trade will be occurring in the housing market.  
Our product should not consume a significant amount of energy unless a substantial amount of data is entered as it should only be run once for each set of data and has only taken a matter of seconds to complete its predictions in previous tests.

### 3.II.d Health and Safety Matters

As the project is purely software and has no physical components, there should not be any physical health and safety concerns for the user other than those to be expected from using a computer. We also do not believe that the gathering and use of housing data to predict prices poses any significant risk to mental health.

## 3.III The Team Product

### 3.III.a Product Description

We have created an algorithm that estimates the price of housing when given a large quantity of accurate data.

This will take appropriate data in the form of a CSV file and will output the estimate for the house prices along with their relevant Ids in another CSV file.

We used the “Random Forests” model amongst other things to transform the given data into an accurate estimate.

The product is limited in that it needs a large amount of accurate input data in order to output an accurate result. The algorithm could make better estimates for missing data in each column to deal with this issue.

test.csv

train.csv

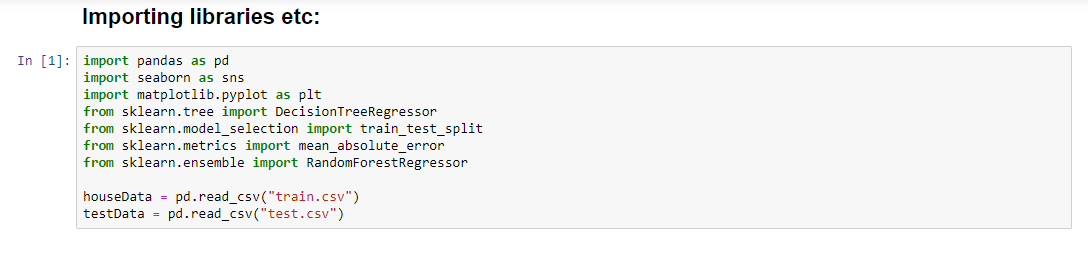
Python application running on Jupyter.

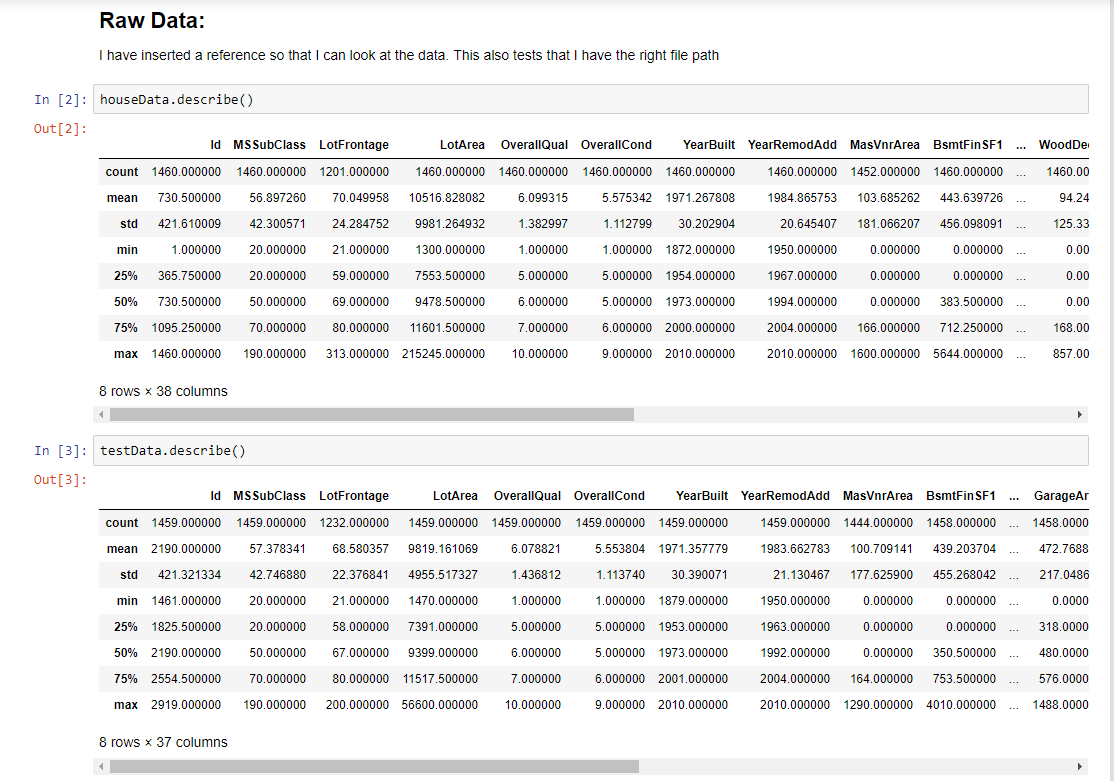
OutputCSV.csv

### 3.III.b Product Demonstration

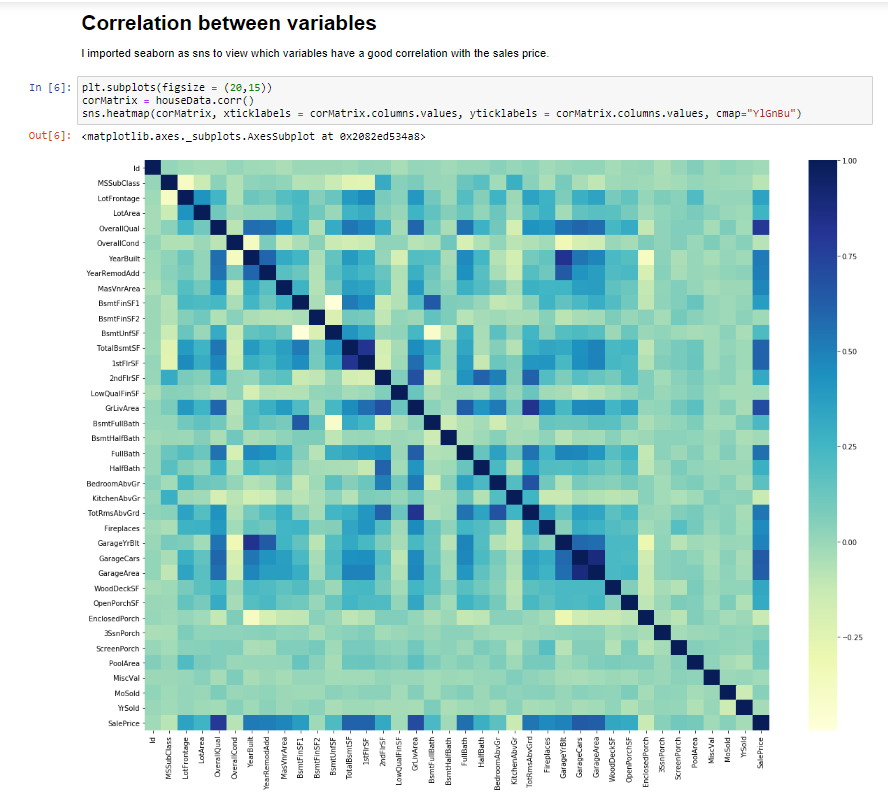
The algorithm was tested using data from Iowa and the output was submitted to Kaggle to receive a rank. Many columns with different kinds of data were also used, meaning that our product should deal with any type of data given.

**Here are some of our Code Blocks:**









**Outputting Data to a CSV:**



# Chapter 4: Project Management (?? words)

## 4.I An introduction to Project Management

## 4.II Project Management Report

### 4.II.a Summary of Teams Project management in Jira

### 4.II.b An evaluation of the Project management

# Chapter 5: Conclusions (?? words)

# Bibliography

Jupyter Steering Council. (2019). *About Us*. Retrieved February 22, 2019, from https://jupyter.org/about

Python Software Foundation. (2019). *History and License*. Retrieved February 22, 2019, from https://docs.python.org/3/license.html