Introduction

The purpose of this document is to outline the individual project specification that you will be working on during the training. This project will involve concepts from all core training modules; more specifically, this will involve:

* Agile
* Databases
* Python
* R
* Hadoop
* Linux

The individual project will use the Iowa homes dataset. The project will require you to build an interactive UI which allows data to be entered about a user’s home from which a prediction should be made on how much the house is worth.

This is purposefully open as we want to endorse creativity and logical thinking in terms of how to present the data.

Objective

The overall objective of the project is the following:

* To create an application which utilises tools, methodologies and technologies that encapsulate all core modules covered during training.

Scope

The requirements set for the project are below. Note that these are a minimum set of requirements and can be added onto during the duration of the project.

The requirements of the project are as follows:

* A Trello board (or equivalent Kanban board tech) with full expansion on user stories, use cases and tasks needed to complete the project. It should also provide a record of any issues or risks that you faced creating your project.
* A relational database used to store the house pricing data for the project.
* An interactive user interface in which data about a house can be entered.
* A predicted sales price of the house from the entered house attributes should appear on screen.
* Full documentation written for all functions made. Uploaded to GitHub with appropriate README.md.

**Extension 1:**

* A prediction to be made when only some of the house attributes are available.
* 80% of functions are tested.

**Extension 2:**

* From an entered house value, predict the attributes of the house.

You should consider the concept of MVP (Minimum Viable Product) as you plan your project, complete requirements before you add extra functionality.

Constraints

The time constraint of this application will be discussed when the specification is given out, as this can fluctuate based on several factors.

The other constraint for this is certain technology that needs to be used. The application needs to utilise the technology discussed during the training modules. The tech stack required would be the following:

* Kanban Board: Trello or an equivalent Kanban Board
* Database: MySQL
* Programming language: Python or R

Deliverable

The final deliverable for this project is the completed application with full documentation around utilisation of supporting tools. This will require a fully functional application based on the domain that you have chosen.

A presentation of work may also be required towards the end of the deadline. However, you will be required to produce weekly reports of any designs and work created throughout the duration of the project.

Marking Scheme

Below are the skills that we will be evaluating for this Databases assessment. These skills are as described in the SFIA 7 framework; please see below if you wish to have more information:

<https://www.sfia-online.org/en/framework>

The skills this assessment will discussed are the following:

* Programming/software development
* Systems integration and build
* Data analysis
* Database design

**Programming/software development**

Designs, codes, verifies, tests, documents, amends and refactors simple programs/scripts. Applies agreed standards and tools, to achieve a well-engineered result. Reviews own work.

**Systems integration and build**

Produces software builds from software source code. Conducts tests as defined in an integration test specification, records the details of any failures. Analyses and reports on integration test activities and results. Identifies and reports issues and risks.

**Data Analysis**

Applies data analysis and data modelling techniques to establish, modify or maintain a data structure and its associated components (entity descriptions, relationship descriptions, attribute definitions).

**Database Design**

Translates and implements simple development project requirements into physical database structures. Assesses proposed changes to object and data structures and implements these changes in physical databases. Assists in database management system support activities for operational database systems.

Programming/Software Development – marking scheme

Below is the list of criteria that will be assessed from your deliverable:

|  |  |  |
| --- | --- | --- |
| **SFIA Skill** | **Rating** | **Details** |
| Designs, codes, tests, verifies, documents, amends and refactors simple programs/scripts. | 1 | Software is missing functionality in major areas for creating, reading, updating and deleting. No tests implemented and designs documented were not implemented. Best practices not adhered to in the project. |
| 2 | Software functionality is working in areas but is not a fully working product. Tests were not implemented; code was commented in small areas. Parts of the project adhered to best practices but not consistently throughout the software. |
| 3 | Software is functional in all major areas but still has small bugs and/or errors. Tests were beginning to be implemented, with basic functions being tested. Best practices were adhered to for most areas of the project. |
| 4 | Software is fully functional and has been tested in all relevant areas. Best practices were consistently adhered to throughout the project. |
| 5 | Software is fully functional and has been tested in all areas, with best practices and refactoring adhered to and implemented throughout the project. Software implements concepts outside of the brief specified at a good level. |

Systems integration and build – marking scheme

Below is the list of criteria that will be assessed from your deliverable:

|  |  |  |
| --- | --- | --- |
| **SFIA Skill** | **Rating** | **Details** |
| Produces software builds from software source code | 1 | VCS implementation was non-existent. Build server was not installed and therefore no builds of software were created. |
| 2 | VCS implementation was attempted but structure was poor and/or content in VCS was irrelevant. Build server was installed but software did not build successfully from repository. |
| 3 | VCS was implemented and code was stored in a structured manner. Some of the content in the VCS could have been omitted but the majority was relevant. Build server installed and successfully built software manually. |
| 4 | VCS was implemented and code was stored in a structured manner with branches. All VCS content was relevant, no unnecessary files. Build server installed and successfully built software after a push. |
| 5 | VCS was implemented and code was stored in a structured manner with branches All files relevant to the repo, nothing unnecessary. Build server installed and successfully built software after a push with artefact produced for successful builds. |

Data Analysis – marking scheme

|  |  |  |
| --- | --- | --- |
| **SFIA Skill** | **Rating** | **Details** |
| Applies data analysis and data modelling techniques to establish, modify or maintain a data structure and its associated components (entity descriptions, relationship descriptions, attribute definitions). | 1 | No data analysis techniques implemented. Best practices not adhered to |
| 2 | Basic data analysis techniques implemented poorly. Some best practice adhered to. |
| 3 | Data analysis techniques used but not implemeneted correctly. Some best practice adhered to. |
| 4 | Data analysis techniques implemented correctly with best practice adhered to. |
| 5 | Data analysis techniques implemented correctly with best practice adhered to. With documentation and descriptions |

Database Design

|  |  |  |
| --- | --- | --- |
| **SFIA Skill** | **Rating** | **Details** |
| Translates and implements simple development project requirements into physical database structures. Assesses proposed changes to object and data structures and implements these changes in physical databases. Assists in database management system support activities for operational database systems. | 1 | Poor database design which does not suit the business needs |
| 2 | Database designed to include tables which works to some degree but would not suit all business needs. |
| 3 | Database designed which would suit all business needs with relationships between tables |
| 4 | Database designed to suit all business needs with documentation |
| 5 | Database designed to suit all business needs with documentation and best practices adhered too. |