SFWRENG/COMPSCI 2XB3: SE/CS Practice and Experience: Binding Theory to Practice

Requirements Specifications

Project Title: Fresh House of Bel-Air

Lab Section Number: L02

Student Names: Gilbert Cherrie, Ethan Dhanraj, Christian Hsu, Hasan Kibria, Nicholas Lobo

Student McMaster Emails: cherrieg@mcmaster.ca, dhanraje@mcmaster.ca, hsuc13@mcmaster.ca, kibriah@mcmaster.ca, lobon3@mcmaster.ca

By virtue of submitting this document we electronically sign and date that the work being submitted is our own individual work.

**Table of Contents:**

# 1. Domain

### 1.1 Stakeholders:

- The product development team
- Individuals and families relocating and investing in property
- Real estate and financial advisors guiding home buyers
- Housing development companies considering areas in which to expand

### 1.2 Expectations/Usage of the System

- The expectations of the product development team are that a variety of users will use the app to help them find good areas to buy and build houses in the United States based on house price index and pollution levels
- The individuals and families that are using this app expect it to help them find the optimal destination to move or buy a home in the United States based on various factors such as distance, house price index and the pollution levels in the area
- Real estate and financial advisors can use the app to compare house prices to find relative home values for their clients and help clients make a good financial investment
- Housing development companies expect the app to show them the best possible states to build and develop houses in terms of increasing house values

What are the main entities that characterize the domain?

- States in the U.S.A
- Each state's respective Housing Price Index
- Each state's respective air pollution levels

What are their main relationships?

- How each state's Housing Price Index corresponds with its air pollution levels, and how this dictates the cumulative value of a state's housing with respect to these two factors along with its proximity and ease of travel to the client's current place of residence

How are they affected by the system we will develop?

- The values of housing price index and air pollution levels are not affected by the system, instead the system uses the values to create an arbitrary value based on both values
- Individuals and families will be able to make better investments when home buying
- Real estate and financial advisors could potentially be affected negatively, as now a client can skip their expert opinion and solely use the system's algorithmic comparison system to decide which state suits them best, if they prioritize pricing, proximity, and pollution.
- Housing development companies can use the information as a tool to plan future profitable projects

## 2. Functional Requirements

The main program is a GUI interface made in Eclipse that has various features or outputs.The various features or outputs are:

- A heatmap of the United States with coloured to circles to show the state's relative home value based on either house price index, pollution data or both and the user can also select for year.
    - This involves using a sorted data set.
- The optimal state for a user to move to based on house price index, air pollution and distance.
    - This feature involves implementing a graphing algorithm.
- The user can search for a specific state and display that state's values.
    - This feature involves implementing a search algorithm on a sorted dataset.

Other than a GUI and its various features that it implements the program also requires a clean data set to work with. This clean dataset will be obtained by combining a dataset on house price index and a dataset on air pollution into a single dataset and only extracting values that are required.

## 3. Product Backlog

| **Precedence** | **Item #** | **Description** | **Contributors** |
|---|---|---|---|
| Very High | 1 | Implement sorting module | ED |
| | 2 | Implement data class | HK |
| | 3 | Implement searching module | CH |
| | 4 | Implement graph analysis module | GC |
| High | 5 | User interface, heat map | NL |
| Medium | 6 | Clean data sets | GC, ED, CH, HK, NL |
| | 7 | Average out data, merge sets | GC, ED, CH, HK, NL |

## 4. Non-Functional Requirements:

### 4.1 Reliability:

- System always returns the correct heatmaps with the given nodes from the dataset.
- If the dataset is updated it will be cleaned for any missing info before updating the software it will also be properly vetted to make sure there are no inconsistencies.
- When the user inputs a specified location or state in the U.S the output should consistently find the best state to move to based on HPI and air pollution percent.
- If the user does not input a valid location the software will return an error messaging informing the user to input a proper value in the correct format.
- The software should be available throughout the year being taken down every 3 or 4 months to be updated with better sorting and searching algorithms or an updated dataset.

### 4.2 Accuracy of Results:

- Correctness of results will heavily depend on the success of sorting, searching, and graph analysis algorithms of which testing procedures are specified in the subsequent section of this document.
- The service should always be returning the appropriate results based on the available data, however it is constrained by several factors.
- The latest pollution data is from 2016, thus in order to preserve relevance and accuracy of results, maintenance to update data sources is required.
- The service is constrained to value average pollution levels and housing price index per state, thus accuracy is being reduced significantly.

### 4.3 Performance:

- The two factors affecting performance of the product service are memory usage and run time, both of which are impacted by the large data sets.
- Data for housing price index and mean pollution levels will be stored in a database server to be accessed by the application.
- Run time for the application will also depend on the size of the data as well as the chosen algorithms and processes.
- Lexicographic sorting is constrained to a run time of O(n lg n) using a stable sorting algorithm.
- After an initial sorting of data, only search and graph process algorithms will be performed, the former running in O(lg n) and the latter depending on implementation choices to be determined.

### 4.4 Human Computer Interface Issues:

- Technical constraints and issues of the user interface are outlined in operating constraints.

- In terms of behavioural aspects, the application provides the user with both visual and numerical information in order to present data in an easily digestible manner.
    - The generated heat map layers air pollution and HPI over one another.
    - As low pollution and high HPI is generally desirable for a long term investment, design of the heat map to indicate these relative regions is integral.
- Potential issues arise from incorrect interpretation of the presented information due to design choices.

**4.5 Operating Constraints:**

- Constraints on storing and accessing data for the final version of our program.
- The project will be using a database to store and access both datasets so it can display on the java applet using heatmaps.
- There will be a cost to buying and maintaining a server which will have to be taken into consideration when implementing the final version of the project.
- Future versions of the product will have the applet added on to a webpage making it easier for users to access the software.
- To have the website available it will also need to be hosted on a server.
- These servers for the database and web page will be our main operating constraints for this project.

**4.6 Portability Issues:**

- Project will be first implemented using a java applet.
- Code can be easily ported to websites that can be displayed on a variety of platforms.
- Implementation of code using java makes it easy to port to an android or apple app.
- Since the interface for the application is also created in java there will be major changes to the implementation when designing it to be used for phones.

## 5. Development/Maintenance Requirements

**5.1 Development Process:**

- Analyze data sets to remove incorrect data or data that causes an error
- Several modules are used to read, edit and rewrite data sets
- Data is sorted using an algorithm for functionality of final product
- Searching and graphing algorithms will be implemented to create desired output such as a heatmap
- An external interface created to receive input and provide external output
- Multiple modules are created for each function without or with little integration of others
- Large emphasis on editing data and sorting for final product to have correct output
- Time constraints allow for sorting, searching and graphing to be prioritized
- Algorithms can be tested using fake data in similar form to data sets
- Modules to edit data are next in priority so algorithms can be tested with real data

**5.2 Testing Process:**

- System testing on algorithms performed in multiple stages
- First stage is testing with mimic data from the data set for basic function
- Allows simple and edge tests to be created
- This stage allows algorithms to be improved as errors in algorithms could be easily located
- Second set of tests would consist of real data to ensure algorithms function with same test but different data

**5.3 Maintenance Process:**

- Regular maintenance includes updates to data set and external interface
- Data needs to be edited and integrated to be resorted
- External interface can be constantly updated without interacting with other modules as it is contained in its own module
- If algorithms require maintenance other modules are not affected as they are contained in its own module