



CSCE 2501

Fundamentals of Database Systems

Project - Real-Estate Finder

Prof. Hossam Sharara
Dept. of Computer Science, AUC

Introduction

In this project, you will be building a database application for the real estate properties in Egypt, that allows users to get information about the real estate properties available for sale in the country, the different developments, agents, and locations.

The main site that provide such information is Property Finder website (<https://www.propertyfinder.eg/en/search?c=1&ob=nd>). However, the website has limited abilities for filtering or doing some higher level aggregations on the market size and state. Hence, your application will be used to provide more flexibility on that front.

Your task will to build a database system backend and an application that provide the real-estate market analytics and features described next.

Milestone I: Data Requirements

The database system you design should be based on the information provided on each property, its listing agent / owner and the other information provided on the property page (e.g. sample)

Your database design should store all the information about the real-estate **property**, including but not limited to: **area, location, development, type (apartment, villa, etc.), number of bedrooms and bathrooms, list of amenities, description, listing date, payment method (cash / installment), down payment, etc.** Many of the properties on the website is located in a **development project**, for such cases the development project needs to be recorded, including its **name and location**. In addition, each of the properties on the site is listed by an **agent**, for each of them you need to store their **name, contact number and email** if available. The agents listed sometimes work for a **broker** company, for such cases you need to store **the broker company's name, its address, and phone number**.

The system should also allow **users** to register on the system using their **email addresses, and pick a username, gender, age, birthdate, and their main areas of focus**. After registering, the users can add their **reviews** for any agent, providing their **rating (1-10) and a textual review**.

Milestone I: Database Design & Implementation

Tools: MySQL

Deadline: 2nd October, 2022

Description:

In this milestone, you are required to design the database ERD, create your database and schema in MySQL server for the data requirements presented on the previous slides.

Deliverables: (20 points)

- Entity-Relationship Diagram of the Real-Estate System
- Relational Model for your system
- A transcript for the SQL statements used to create your database and schema

Milestone II: Database Population

In this milestone, you are required to write a web crawler to crawl the pages from the property finder website for all the properties listed in the past 14 days priced up to 3 million EGP, parse the HTML you crawl and extract the relevant fields for populating the non-user tables in your schema. Note that you may need to extract some of the payment method (cash / installment and down payment) fields from the description text

At the end of this milestone, you should have all your non-user tables populated with the extracted data. For the user tables, use sample test information to populate the relevant tables.

Milestone II: Web Crawling and Data Population

Tools: Scrappy, pycharm, BeautifulSoup, selenium, python (or any other crawler library / language)

Deadline: 23rd October, 2022

Description:

In this milestone, you are required to implement your crawler, extract the data from the property finder website and populate your database

Deliverables: (50 points)

- Crawling script
- Populated MySQL database dump
- CSV files for the values in each table

Milestone III: Application Layer

In the final milestone, you are required to design a client application that is capable of connecting to the database hosted on remote MySQL server. The application should have the following functionalities:

- Register a user
- Add a new user review on an agent
- View existing reviews of a given agent
- View aggregated rating of a brokerage company
- Show the location of a given development, along with the average price / sqm and the number of listings for each unit type
- Show all the properties of in a certain city, along with the average price / sqm for each unit type
- Show all the properties in a certain city in a given price range, with a given **set of amenities**
- Show the top 10 areas in a given city by amount of inventory and price / sqm of a given unit type
- Show the top 5 brokerage companies by the amount of listings they have, along with their avg price / sqm, number of agents, and average listings per agent
- Show all the properties listed by a specific agent (given their first and last name and / or phone no)
- Estimate the cash discount per development based on the price differential between the cash price and the installment price for each unit type

Milestone III: Application Layer

Tools: Any language / platform to implement your application (Web-based, GUI, command-line are all acceptable)

Deadline: 13th November, 2022

Description:

In this milestone, you are required to implement an application to perform various transactions / queries on your backend database system. The database has to be hosted on an actual server (you can use any of the free DB hosting services available (e.g. www.db4free.net)) and make sure your application is communicating with this service not the localhost

Deliverables: (30 points + 5 bonus points for GUI/Web-based apps)

- Application Implementation / Source Code + Executable
- Latest Dump of your database
- Demo