DEVELOPING A DATA MODEL FOR ARMENIAN REAL ESTATE MARKET USING BUSINESS INTELLIGENCE TOOLS

Hripsime Voskanyan
American University of Armenia
College of Science and Engineering
Yerevan, Armenia
hripsime_voskanyan@edu.aua.am

ABSTRACT

With the quick development of technology and the amount of data, one can claim that the digital revolution has spread its wings everywhere, including the financial sector, in this case, the real estate market. Real Estate Market has economic significance in each country. This project will connect the real estate industry with data science to generate insights for the market. It can be done in various ways, including creating Business Intelligence tools and making predictions with Time Series Forecasting and Machine Learning algorithms. However, this project will focus on the Data Engineering and Business Intelligence parts and use the data generated daily to create a data model. The data model will be an end-to-end pipeline with Business Intelligence tools to analyze Armenia's Real Estate Market with a big focus on apartments in Yerevan. It will provide its users with an opportunity to access the Real Estate data more insightfully, including better speed, readability, and reliability. The user segment of the project will include buyers, real estate agencies, investors, Government agencies, and anyone interested in gaining insights from Armenian Real Estate Market more comprehensively and efficiently.

***K*eywords** real estate · data warehouse · web scraping · data engineering · business intelligence

1 Introduction

There are various Key Performance Indicators: KPIs to evaluate the country's economy. For instance, the unemployment rate, inflation rate, gross domestic product, and Government debt can be considered KPIs. In the long list of these indicators, one can also add the effect of the real estate market.

The Armenian real estate market has been an exciting topic for a while because of country policies and regulations. The new tax law and the fact that foreigners can buy properties in Armenia without any obligations make the topic even more interesting. The following paragraphs will more precisely uncover why.

Based on the yearly research by Grant Thornton, the number of transactions done for real estate increased during the preceding three years [1]. The dominating city is Yerevan, Kotayk, and Armavir regions following. The transaction data was also used for price analysis, showing that apartments in Kentron, Arabkir, and Kanaker-Zeytun were the most expensive. The analysis shows that the market is growing, and there is a need to do an analysis and provide the users (real estate agencies, buyers, Government, and investors) with Business Intelligence tools to get the bigger picture of the Armenian real estate market in real-time.

In "Armenian Real Estate Market by Property and Business: Opportunity Analysis and Industry Forecast, 2019-2026," the authors mention that back then in 2018 Armenian real estate market was estimated to have a value of $880.4 million, which was pretty big taking into consideration that Armenia is a small country and the level of poverty is pretty high [2]. The forecasts show that in 2026 the market value would grow by 50% of what it had been in 2018, meaning it would be around $1,200.000 million.

The article mentions, and based on other research, the most critical factor in why the Armenian Real Estate Market is growing at the moment is the income tax return law. According to Armenian law, each working person in the Republic of Armenia pays different types of taxes, including pensions and payments to the army. One type of these taxes is the income tax, and according to the law, residents can take mortgage loans, and income tax, instead of going into the country budget, will pay for the loan's interest. In this way, many people can buy real estate in Armenia. Moreover, it is also stated in the article that mortgage loan interests are considered pretty low in the region, with an average of 12% for ten years, which compared to, for instance, ten years ago, makes taking loans more accessible. Moreover, the rentals segment of the local market is forecasted to dominate. It is due to two main factors: people use the rental segment for passive income, and since the Russia and Ukraine war, many Russians have come to Armenia, making the demand for the rental segment even go higher. Not only that, but an industrial segment is also going to dominate, and it is predicted that number of factories and workshops will increase.

With all this information that the market is growing, one can be interested in doing analysis and making getting the real estate data easier for agencies as it will benefit both business and academic perspectives. For that, one needs to have a fully developed pipeline to store the data and use BI tools with that. Therefore, the following paper will outline how data science tools can allow end-users to access information more quickly and gain more insights than just looking at listing websites. It can be a part of a bigger project to do deep analysis on the topic from a machine learning perspective and do various projects and predictions.

2 Literature review

Research is an essential part of any project, as it allows for a deeper and better understanding of the field, including what has been previously done, what has failed, what has succeeded, and what future advancements could be made. Therefore, in this project, research was done to better understand the study around the topic. As a result, one could notice that there is a lack of research done on the Armenian real estate market, making it harder to do a comparative analysis. However, from a methodological perspective, various studies have been conducted on the topic with different methods. Therefore, the following paragraphs will show the methods utilized by various authors in their research.

First, it is essential to understand that integrating BI tools in real estate can be efficient. According to "BI tools for real estate companies: How to grow your business with data-driven decisions," it is time for the real estate market to make data-driven decisions as the market is growing every single day, and it is hard to keep the speed [3]. Data is being generated every day, and just scrolling through house listing websites and looking at prices, one cannot really make decisions that can improve profits. The author mentions that BI tools can give one the opportunity to see in real-time which markets are growing, in which markets prices are going down, and make a decision about where to buy a house and where to sell. It is an amazing opportunity for real estate agents to make informed decisions, gain a competitive advantage and grow their businesses. Not only that, but this can save their time and money, and one investment into BI integration can be efficient for their long-term profits. In short, the article summarizes the significance of using such tools in the market. As the Armenian real estate market is also growing, this can also be used in the Armenian market.

In " Business Intelligence Framework Design and Implementation: A Real-estate Market Case Study," the authors mention that the BI framework can provide investors with accurate information and analytics [4]. The article explores all stages of creating a Business Intelligence tool: designing and testing, in more detail: data collection, transformation, storage, analysis, and visualization, ending with machine learning algorithms for prediction analysis. The authors collected data from Jordanian websites for real estate listings. The dataset is big: about 11,000 listings with 27 attributes. The authors suggest that such data could be collected not only from websites but also from land registers and historical prices of markets. It is hard to do for the Armenian real estate market, as getting information from the land register in a short timeline is impossible. Selenium is used for web crawling and scraping, and NLTK for analysis.

After data collection, the authors cleaned and preprocessed the dataset to make it fit for further analyses. It is an important step that will also be used in this project. SSIS is used to transform data, and SSAS is used to implement data warehousing. These tools have widespread applications and are proven to be efficient. Similar tools will be used in this project as well. The XGBoost model is used to analyze which features are worth keeping. It is proven that XGBoost is one of the best to do feature extraction. After this, they were left with only four features which was a massive extraction. Following this, train/test separation and Linear Regression were done to predict house prices. As the paper'spaper's objective was to create business intelligence tools for investors, in the final stage PowerBI dashboard was created with visualizations and plots, including treemaps, bar charts, and maps. Overall, the authors concluded that this type of analysis is essential and shows practical price analysis and prediction.

Let one observe that researching the local market is also necessary to understand the local trends and important information. However, the problem is that the research about the Armenian market was a bit hard because of the lack of existing articles.

Despite the lack of research, "Automated Real Estate Valuation with Machine Learning: A Case Study on Apartment Sales in Yerevan" done by Henrik Sergoyan and Grigor Bezirganyan shows a real analysis of Yerevan's real estate market [5]. The authors write that of 2020, Armenia has a market capitalization of about 1 billion dollars. Even though during 2021, after the war and pandemic economic activity fell, only the real estate market rose. Therefore research was done, and data used in it was collected from three main websites, myrealty.am, list. am, and real-estate.am. Moreover, OpenStreetMap API was used to find geospatial features furtherly to be used in the ML model. The overall variables included information about the number of rooms, bathrooms, floor, the existence of certain attributes, and the variables collected with API.

Interestingly, after data exploration, the authors saw that the prices have a right-skewed normal distribution, having a mean between $800 and $1000 per square meter. Analysis of the number of rooms, square meters, apartment floor, and other things was also done to get the bigger picture better. As the objective of the research was value prediction, in this research as well, authors used XGBoost, linear regression, and geographically weighted regression, having XGBoost performing better than others. Overall, the research was done in a way to help the end-users with predictions and visualizations to get a bigger picture of the market.

Not only real estate agencies and investors but Government agencies can also benefit from using business intelligence tools to gain insights in "Mapping for the future: Business Intelligence tool to map regional housing stock," whole data integration and advanced analysis were done on existing housing stocks [6]. Data is collected from multiple agencies. The authors mention that creating a platform with BI tools can let Government agencies see the characterization of houses according to building attributes, such as ceiling insulations and rainwater tanks. In this way, policymakers can see the situation and act accordingly. A relational database management system was used for data warehousing and modeling. Overall, the model downloads raw data from websites, extracts data, cleans, loads to the database, and monitors ETL tasks. Such flow is efficient and will be used in this research as well. Map visualizations were done to show the insights to the end users.

To understand the data warehousing details, more research was done on that end as well. In "Mining Real Estate Listings Using ORACLE Data Warehousing and Predictive Regression," the authors used a multidimensional data model to create dim and fact tables [7]. Such a model will be used in this project as well. The dim tables used in the research are around office, agent, and area information, whereas the fact table includes all details with foreign keys to create the primary key. Afterward, linear regression is done to predict the price of the real estate. The authors conclude that further research can be done using more variables, e.g., whether the house has a pool or not.

A whole real estate market data model was created in "Real Estate: Developing a data model for the City of Lisbon to improve information sharing" for Lisbon, providing all the necessary tools for users to gain insights and generate knowledge [8]. The paper includes all the details from the start to the end, including all the methodology for the data collection, exploration, cleaning, data warehousing, and final dashboard creation. It shows how with those tools, they can improve the market situation and increase the efficiency of analyzing the real estate market of Lisbon.

Based on the research, one can conclude that BI tools are essential, and doing it for the Armenian real estate market, many institutions can benefit from them.

3 Methods

3.1 Model Proposal

The proposed model consists of four main stages:

- data collection
- ETL
- data design and warehousing
- providing a business intelligence (BI) solution.

The model aims to ensure an efficient flow of information within the framework of Business Intelligence. The model will be an end-to-end pipeline for collecting and storing data on the Armenian Real Estate Market while also providing user-friendly BI tools for market insights.

Overall, the proposed model aims to provide a comprehensive solution for storing and analyzing Armenian real estate data while also making it user-friendly and accessible to the public with BI tools. Its objective is to help stakeholders to understand market trends better, discover new opportunities, and make data-driven decisions to drive growth and success.

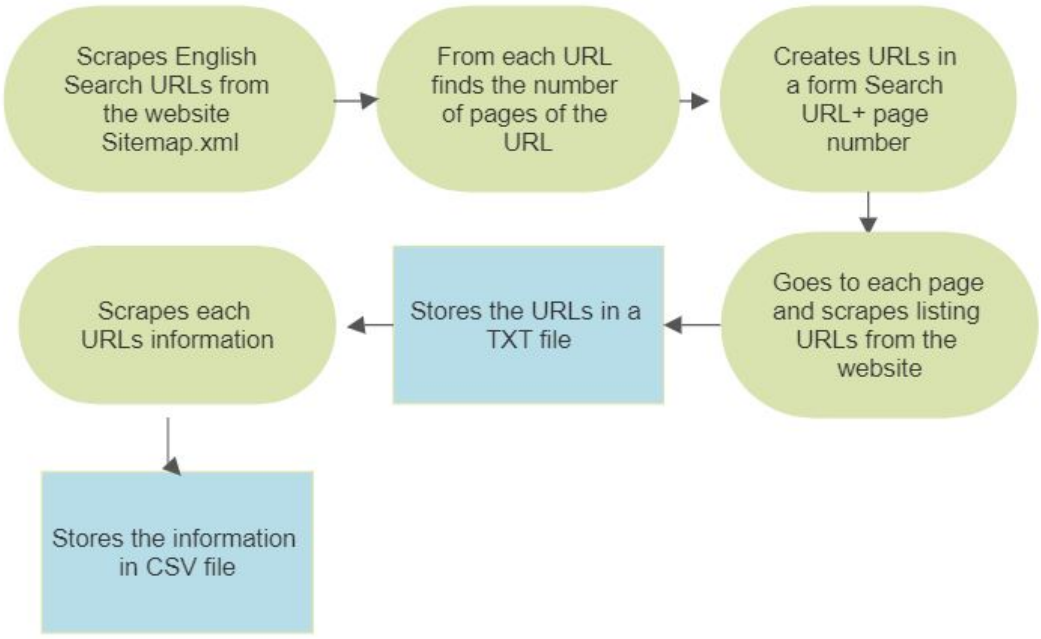

Figure 1: Flow of Web Scraping

3.2 Data Collection

There are various ways to collect data for real estate. One can get information from Government about transactions; however, sometimes the process is complicated, and the data needs to be more accurate regarding the prices. Another way can be getting already collected data; however, these are not always up to date and unsuitable for Business Intelligence live models. The last way is to do web scraping and write reusable code to scrape the data with the timeline.

The model will scrape two reliable real estate websites in Armenia. The code will be designed to be as efficient and reusable as possible to reduce the resources required.

3.3 ETL

ETL process is a popular way to use in data warehousing. ETL: Extract, Transform, and Load. During the Extract part, the process extracts data from its original source, then Transforms it, ensuring quality, cleaning it, and making it ready to use. During this part, duplicates are removed, and data is validated, sorted, standardized, and cleaned. Then the model loads to the database for users to get.

ETL method will be used for the project to process the data before it is sent to the data warehouse. The extract and transform stages will be scheduled to ensure the data remains consistent and up-to-date.

3.4 Data Design and Warehousing

Data warehousing is part of a data management system that is done to help the Business Intelligence activities and tools [9]. It supports the analytics process with efficient queries of large amounts of data. Moreover, it builds historical data to improve the decision-making process.

Data design will be created before warehousing to ensure that the necessary information from both websites is contained and the dimensional data model is in place. A dimensional data model (DDM) is a way to store the data in Facts and Dimensions. In this way, retrieval of big data is done more efficiently and faster. It ensures data quality and is easier to understand for the users.

Table 1: Possible Attributes to get from Websites

Attribute Names		
Address	Floor	Square meters
District	Description	Availability of gas, water, electricity
Coordinates	Price	Date the listing was added

Online tools will be used to make the design more visual and easily adaptable. Data will then be loaded into the cloud-based warehouse and loaded into dim and fact tables.

3.5 Dashboard and Analytics

A Business Intelligence dashboard is one of the best and most popular data visualization and analytics solutions. With the help of those dashboards, companies, and individuals make more data-driven decisions; as with filters and various other tools, users can access complex data, draw conclusions, and gain insights. Not only direct users but non-technical people can quickly get insights from BI dashboards if the dashboard is created correctly.

After the data warehousing dashboard will be created with Power BI tools. The dashboard will contain necessary visualizations, tables, and filters for users to gain insights about the real estate market. With the help of the dashboard, users will easily access data, analyze it and make data-based decisions. Overall, it is vital to notice that a dashboard is an important tool to monitor market trends and have it next to the data engineering pipeline.

4 Development

4.1 Web Scraping

Regarding the model, there was a need to gather data about the real estate market in Armenia. To achieve this, the research was conducted to identify appropriate websites that contained all the necessary information about apartment listings. Before selecting the websites, a list of required items was compiled, which one can find in Table 1.

After creating this list and reviewing the available websites that could be scraped, two main websites were chosen: myrealty.am and besthouse.am. These websites had the most items on the list and were similar in terms of the availability of information. It is crucial to note that in this stage, myrealty.am will be tested, and after success besthouse.am data will be loaded. A web scraping code was then developed in a generalized version using the BeautifulSoup library, which could be used for both websites. Functions were written to collect all available URLs from the sitemaps and, if not fully available, retrieve the remaining URLs from the website (See Figure 1). For example, in the case of myrealty.am, the sitemap includes all possible search URLs but not specifically listing URLs, so a code was written to go through these search URLs and retrieve the listing links from there. All available links were first saved in a data type that ensured no duplicates and then saved as TXT files. It was done to avoid scraping duplicate listing information in each period, which would be expensive and slow down the model's working process and the speed of loading jobs.

Only Python scripts were used for the web scraping part. Additionally, it is essential to note that constant scraping can lead to connection loss with the website or being blocked by them. To prevent such errors during scraping, a sleep tool was used. Furthermore, during the scraping process, various errors can occur. To avoid such incidents, error handling was used in the process with try and except blocks. However, it is important to note that these types of websites can change over time, such as changing tags or the representation of information. Thus, checking the data and results within a specific timeline is crucial to ensure that the model operates as intended.

4.2 Preprocessing

Preprocessing is an essential part of any data-related project, and given the large amount of data involved in this project, it is no exception. As the Armenian real estate market websites are not written comprehensively, the preprocessing stage is critical to ensure that the code is clean and can be loaded into the dim and fact tables.

One of the common ones was that numbers came with commas when extracting the price or other numerical information; therefore, those commas needed to be removed. Another example was that some information was written under the same HTML tag, meaning that we needed to separate the information from each to different columns. Some numeric information was written with unnecessary text, such as the square meters with the text 'sqm. Meter.'

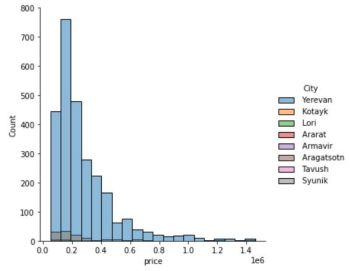

Figure 2: Apartment Price Distribution by Region

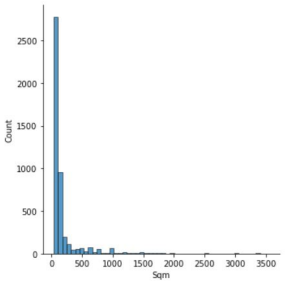

Figure 3: Apartment Surface Area Distribution

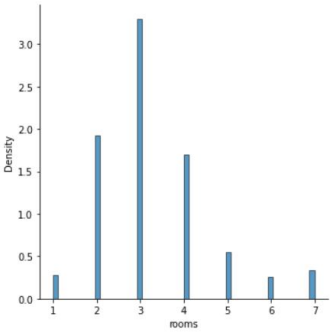

Figure 4: Apartment Room Numbers

The websites that were being scraped had information about facilities. The name was mentioned on the listing page if the house had gas, storage room, or furniture. This information was scraped as a whole string and then separated into different columns. The column names are the name of the facilities, and then values are binary: 0 and 1, 1 showing the house has that facility and 0 the absence of it.

Moreover, the address region, street, and district were extracted as those would become dimensional tables. Additionally, year, month, quarter, and day were created from the date, as those will also be included in the dimensional tables and the dashboard.

Python script was used with Pandas module for all these processes to do the necessary modifications: Transform part and to be uploaded to the cloud storage.

4.3 Data Exploration

After cleaning the data, it is a good practice to go over it and explore it using various methods and visuals to see what, in general, one has in the data. This step is crucial because it allows one to evaluate the data quality and see whether it can be put into production.

In Figure 2, one can see the price distribution of apartments. Before making the visual (this refers to all the visuals mentioned in this section), it was cleaned from outliers using 99% quantile to get an understandable result. The plot of price distribution by region shows that most of the selling apartments are in Yerevan, and the prices are right-skewed. From Cadastre's last trimester's report, one can observe that, in fact, Yerevan has the most transactions, and about 33% of all transactions are from there [10]. In more detail, one can also see the prices distribution, and as already claimed, the highest prices are in Kentron and Arabkir Districts. It can also be claimed from Figure 5, where one can find the distribution of prices by districts. The surface area distribution plot in Figure 3 shows that the vast majority of the houses are up to 200 square meters, from which up to 100 square meters is the most common. From the density plot of the number of rooms for Apartments (Figure 4), one can observe that selling apartments mostly have two or three rooms, which is the number of rooms new families are looking for to start a family. Other visuals were also created to go deeper into data.

4.4 Database Design

A correct database design is necessary to create an effective working data warehouse, as it would only be possible to proceed with it. As mentioned, our model scrapes information from two websites, which are stored together, although the end-user can differentiate which listing information is from which website. After scraping, the following attributes need to be stored: address, city, district, street, rooms, floor, description, price, views, latitude, longitude, added date, edited date, type, platform, square meters, price per square meter, and availability of various attributes, such as iron door, gas, water, storage room, and parking, among others. At this stage, these attributes are saved in a CSV file. Several assumptions were made before starting the database design.

- The price is assumed to be in USD currency.
- The availability of attributes should have only 0 and 1 values.
- The floor is written in the $5/6$ format, showing both the house's floor and the number of floors the building has.
- The type shows whether the house is an apartment or a private house.

Figure 5: Listing Houses Star Schema diagram

After the database design, the star schema methodology is used. This approach is popular in data modeling, making data querying more efficient and easy [11]. Dim tables are constructed around the fact table. The fact table contains one record for distinct, discrete measurements, whereas dim tables show the details of that measurement. Star schema reduces the complexity of the process. Therefore, in Figure 5, a star schema of the initial relational database is shown, with one fact table and various dim tables for attributes such as date, district, region, platform, and type.

Each dim table has a primary surrogate key, which means that the key is created artificially. Thus, the fact table's primary key consists of all foreign keys from those dim tables.

4.4.1 Data Warehousing

Data Warehouse is designed to supply the needs of the model. As the listings are added to the websites daily, data will be added to the warehouse, meaning the ETL process works periodically. That is why the incremental method is used to supply the warehouse with data. To ensure that it is possible to track the data, the back ingestion date and natural key are available in the data. The model is created so that only new information will be added to the database or if the old information was updated. It is possible to check with the natural key available on the website.

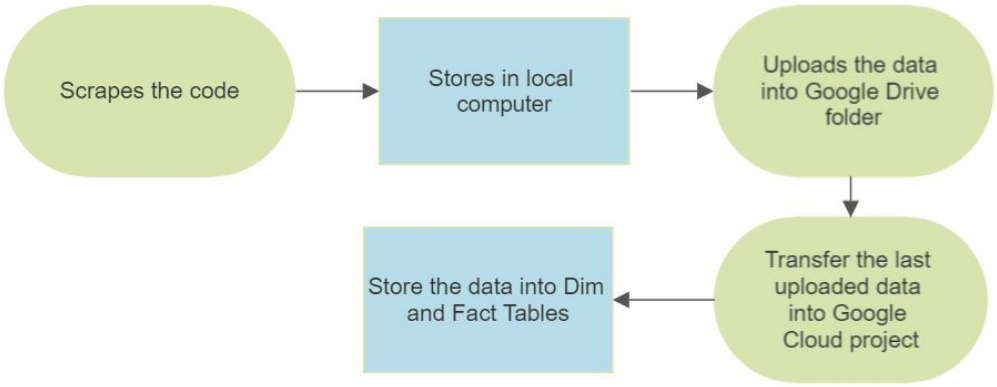

Figure 6: Data Warehousing Flow

As mentioned, the database has several dimensional tables and one fact table. In Table 2, one can see the descriptions of each dimensional table. After scraping and preprocessing, data needs to be loaded into the database; however, it is much more complex than just uploading. The whole flow can be found in Figure 6. Uploading from a local computer and querying on it might be problematic because of the lack of resources. Another problem can be more technical and cause the quality of information to decrease. Thus, a cloud-based solution is used.

When the data is ready to go to the database, a Python code is written that accesses the Google Drive folder to upload the file there. These special credentials access the folder and modules such as PyDrive and Google.Cloud. When the process takes the CSV file from the local computer, it adds the ingestion date and uploads it to the Google Drive-specific folder as a staging table. Staging tables are used to decrease table fragmentation and are suitable for temporary storage. They are only used for interim results; the permanent storage will be in the warehouse.

After that, another function finds the folder files, chooses the last updated one, and uploads it to Google Cloud. Here already created BigQuery project and dataset are there, as the data should go there. BigQuery is a cloud-based data analytics web service allowing users to access very large sets of data more efficiently and faster. Before uploading the data, infrastructure initialization is done, meaning empty dimensional and fact tables are created. Dimensional tables use Slowly Changing Dimension Type 1, meaning that when the new data loads, it checks for updates and overwrites the existing data with new data and keys. Each dimensional table creates its surrogate key, which is loaded into the fact house table with all the other information from the staging table. The primary key consists of surrogate keys coming from dimensional tables. After that, a complete cycle of tasks is run:

- Uploading staging table to Google Drive
- Uploading staging table to Google Cloud Project
- Loading data into dimensional and fact tables

It is important to notice that ingestion date and reload parameters are added to the updating flow code. An ingestion date is added to ensure that no duplicates are added when the data is being loaded the second time during one day. Reload parameter ensures that if the data is being loaded a second time during the same day, all data previously uploaded during the day is removed again to improve data quality and avoid duplicates.

4.4.2 Dashboard

As already mentioned in the Methodology section, the dashboard is an important part of the data model as it is the product that the end-users get to generate knowledge about the market. PowerBI was used in this project to create an interactive dashboard for users. Daily and Monthly division is used in the dashboard, meaning that the users will have two opportunities to look at the visuals, choosing the days of the last two months or the months they want to look at. If the user is interested in the current situation, they will prefer the daily mode; if they want to get a bigger picture of the market, monthly will be their preferred choice.

Table 2: Dimensional Tables with Descriptions

Table Names	Description
DimRegion	The City/Region name of the listing
DimStreet	The Street name of the listing
DimPlatform	The platform name of the listing
DimType	The type(rent, sale) of the listing
DimHouseType	The house type(apartment, house) of the listing
DimAddDate	The date listing was added
DimEditDate	The date listing was edited

On the first pages of the dashboard, one can see the overall number of listings available at that moment; the average price is in United States Dollars, the average number of rooms being sold, and the average surface area. Moreover, the users will be able to see the price and surface area relationship, as well as the number of houses over time to see how it is being changed. As in Armenia and everywhere else, three main facilities are important to have: gas, electricity, and water; they are also pictured visually. The user should note that these facilities are considered to be available if the seller mentions them on the listing website. If the seller mentioned anything about facilities but not about this, we considered they are not available, and Blank just shows that overall information was not available.

Further details are shown listing by listing in the further pages with the whole information about a specific house. Here users can find various filters, including the district, street, house type, number of rooms, and floor number. In this way, the user can find the house they are looking for, what that house has, the price, and other details.

Most importantly, an interactive map (See Figure 7) is created for all the houses being sold in the current year. Each district has its color on the map, as shown in the picture. The houses are shown with bubbles, where the size of the bubble is directly proportional to the surface area. Hovering over the bubbles, users can get detailed information about the listings. Additionally, here as well, users have the opportunity to filter a number of rooms, timeline, floor number, type (rent, sale), and house type. In future versions, the published dashboard will be refreshed automatically over time, so the users can get real-time information about the real estate market, which was the project's main objective.

5 Results

The data model was developed based on the literature research and the available resources. The solution provided in this project tries its best to provide a Business Intelligence solution as efficiently as possible with fewer resources. It ensures that large amounts of data can be supported and stored in a data warehouse.

The main problem of the data model is the bad structure of the Armenian Real Estate websites, which lowers the quality of the information. However, with good data quality specialists that can periodically check the data quality and accuracy, the project can also be used in practice. The dashboard is built to be user-friendly, including only easy-understanding visuals and maps, so access to the information would be as smooth as possible. It includes all the important information that is needed. However, further feedback can be collected from users to add or remove other parts of information as well.

Another good aspect of this solution is that it is not a one-time analysis; it constantly adds information and changes its visual outputs according to those. However, it takes into account storing the data in a literate way so the information does not repeat itself, and it uses the tools it needs to store the data correctly.

6 Further Advancements

In every research, one should be able to analyze the results and highlight the future advancements that can be done to improve the results. This project had two big constraints, without which the project would have given even more professional results. The first is the time and resource constraints. The project had a deadline; however, due to a lack of resources, some processes were slower than expected, which resulted in getting less data than expected. The second one was the lack of good structure that Armenian real estate websites have, which made the scraping process harder or when scraped, the preprocessing more challenging and time-consuming.

In case of more time and resources, more websites could be scraped, so the database would be bigger, giving the end-users more accurate information. More analyses could be done, such as using Machine Learning algorithms. For

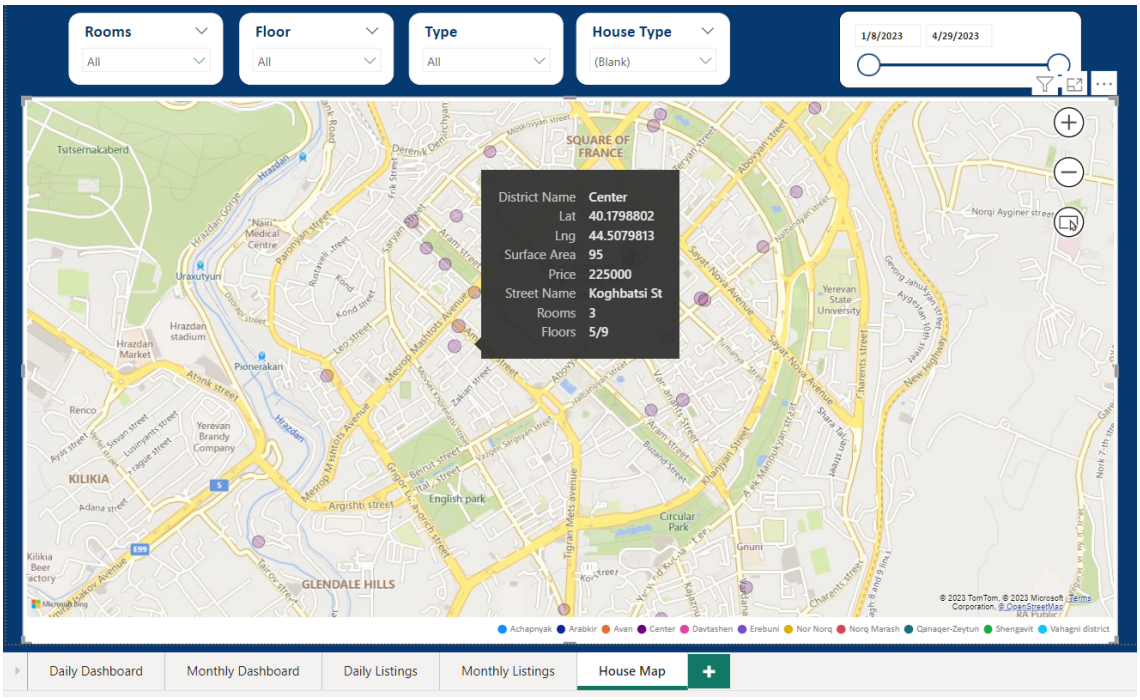

Figure 7: Yerevan Map of Selling Houses

instance, Linear Regression or XGBoost could be used to make price predictions. The project would include both data engineering and data science components. More than that, to have more attributes for running Machine Learning algorithms, Open StreetMap API could be used. It could add variables such as whether the house has close educational, financial, healthcare, and other important facilities near it.

7 Conclusion

Overall, the solution provided in this paper is part of a bigger problem, meaning that forecasting algorithms can be added to this model with the automotive job scheduler to prepare the data model for production fully. The model provided here does its best to increase information quality, user-friendliness of how data is shown to the end-users, efficiency, and interactiveness. It organizes the listings information in a literate way, so there would be no duplicates or information would not be lost.

This tool allows users to store the data, gain knowledge of the market, look for houses, filter through listings with interesting attributes, see price trends, and see detailed information. Even though the project targets only two website agencies, it can be adjusted to get more data and improve quality.

This tool proves that modern technology and Business Intelligence tools are in constant need for all markets, including the most conservative ones, because making data-driven decisions is very important now.

References

[1] Grant Thornton Armenia. Real Estate Market Overview: Armenia 2021, April 2021. Accessed: May 2, 2023.

[2] Rohit S and Onkar S. Armenia Real Estate Market by Property Type (Residential, Commercial, and Industrial), and Key Players: Global Opportunity Analysis and Industry Forecast, 2020-2027, March 2020. Accessed: May 2, 2023.

[3] Domo resource - bi tools for real estate companies: How to grow your business with data-driven decisions.

[4] Salam Fraihat, Walid A Salameh, Ammar Elhassan, Bushra Abu Tahoun, and Maisa Asasfeh. Business intelligence framework design and implementation: A real-estate market case study. *ACM Journal of Data and Information Quality (JDIQ)*, 13(2):1–16, 2021.

[5] Henrik Tigran Sergoyan and Grigor Vahan Bezirganyan. Automated real estate valuation with machine learning: A case study on apartment sales in yerevan. *Journal of Architectural and Engineering Research*, 2:83–91, 2022.

[6] Murad Safadi, Jun Ma, Rohan Wickramasuriya, Daniel Daly, Pascal Perez, and Georgios Kokogiannakis. Mapping for the future: Business intelligence tool to map regional housing stock. *Procedia engineering*, 180:1684–1694, 2017.

[7] Wuri Wedyawati and Meiliu Lu. Mining real estate listings using oracle data warehousing and predictive regression. In *Proceedings of the 2004 IEEE International Conference on Information Reuse and Integration, 2004. IRI 2004.*, pages 296–301. IEEE, 2004.

[8] André Sevinate Sousa Rendas Pereira. *Real Estate: Developing a Data Model for the City of Lisbon to Improve Information Sharing*. PhD thesis, Universidade NOVA de Lisboa (Portugal), 2021.

[9] Oracle Corporation. What is a data warehouse? June 2021.

[10] State Committee of Real Property Cadastre of the Republic of Armenia. Land Register of the Republic of Armenia, 2023, 2023. Accessed: May 2, 2023.

[11] Hui Li. Star schema. *Towards Data Science*, June 2021.