كلية الحوسبة والمعلوماتية |
College of Computing and Informatics
الجامعة السعودية الإلكترونية
SAUDI ELECTRONIC UNIVERSITY
2011-1432

Semester 1 – 2021/2022

Course Code	DS620
Course Name	Data Visualization
Assignment type	Project
Module	ALL

Student ID	G200002614	G200007615
Student Name	Enad S. Alotaibi	Abdulaziz M. Alqumayzi

Real Estate

Enad Saud Alotaibi
G200002614@seu.edu.sa,
Saudi Electronic University,
Riyadh, Saudi Arabia

Abdulaziz Mohammed Alqumayzi
G200007615@seu.edu.sa,
Saudi Electronic University,
Riyadh, Saudi Arabia

ABSTRACT

The purpose of this project is to deliver a better perception of real-estate data. For this, we have taken a dataset of real estate. Power BI is used for the visualization of data. Along with the dataset, a few measures were also created to highlight the KPIs. By using relevant visualizations different stats are illustrated that can help buyers or sellers for property dealing.

Keywords: real-estate; visualization; KPIs; dealing; Power BI

1. Introduction

Data plays a powerful role in any business or process that requires prior interpretations. By using data, visualization tools like power bi or Tableau, data can be displayed in such a manner that can help in better decision-making.

In this project, a real-estate dataset was used to depict the market stats. The dataset is collected from a website and is clean, but a little transformation is required in data in order to analyze it properly. The dashboard will help the real-estate professionals to review their earnings by location and other KPIs. As a power bi dashboard is interactive, one can easily go through the data by selecting different filters and the whole dashboard will be updated according to it.

2. Body section

2.1 Data

There is a total of nine columns in the dataset. For better visualization, we have taken 28022 samples of the data. Each row is representing the data of a real estate deal done in a region. The first column identifies the deal number. It is of string data type and has a unique value. The second column represents the category of a property. It is also a string and has only two values either commercial or residential. The third column

contains the type of land or the purpose for which the land will be used. A total of ten types have a string value. A short description of each type is given below:

- **Agricultural land**: the land area that is either arable, under permanent crops, or under permanent pastures.
- **amenity**: Public facilities such as schools, mosques, parks, health centers, and other government departments
- **Apartment**: a place to live that is only part of a larger building.
- **Building**: an investment property or income property is real estate intended to generate a profit, from either capital gains or rental income.
- **Chalet**: a small house (sometimes big), especially used for gathering weekly rather than gathering in houses.
- **House**: a place built for people to live in.
- **Land**: the solid part of the surface of the earth.
- **Shopping center**: a group of retail stores and service establishments usually with ample parking facilities and usually designed to serve a community or neighborhood.
- **showroom\shop**: a space used to display products or show entertainment.
- **Villa:** a type of house, usually smaller than a house.

The fourth column contains the total transaction price of the deal in SAR. It is the product of price per square meter and total area in square meters. The fifth column has the area in square meters of the property. The sixth column represents the price of the unit square meter of the property. The data type of the fourth, fifth and sixth fields is numeric. The seventh field has the date on which the transaction was made. Its data type is a date. The eighth and ninth fields represent the region and city. The dataset contains the data of regions of Saudi Arabia.

Considering the data for a seller, the given dataset is helpful to get an insight into the property's types that are popular among buyers. Along with this, it can be used to analyze total deals done in a specific region or which duration has the greatest number of deals. If we consider the same data for the buyer, it can help him to propose a thoughtful offer or decide the final offer for any property.

2.2 Methods

The dataset was clean and does not need any changes in the column. A transformation is needed while importing the data on Power BI. The date in the dataset was in a format that was unparsable in power BI. To parse it, I changed the data type to text and then again changed it to DateTime with the right format.

There is only one table for data, so it is not very complicated to map. The transaction data is already present in the date hierarchy by which we can easily present the data in a year, quarter, month, or day. There are two measures that we have created i.e., total area and total deals. The purpose to create measures is to help in performing the calculations on data while one interacts with the dashboard.

2.3 Analysis

Figure 1 shows the dashboard that we have created using the provided dataset and performed analysis on it that can be helpful in decision-making and plans.

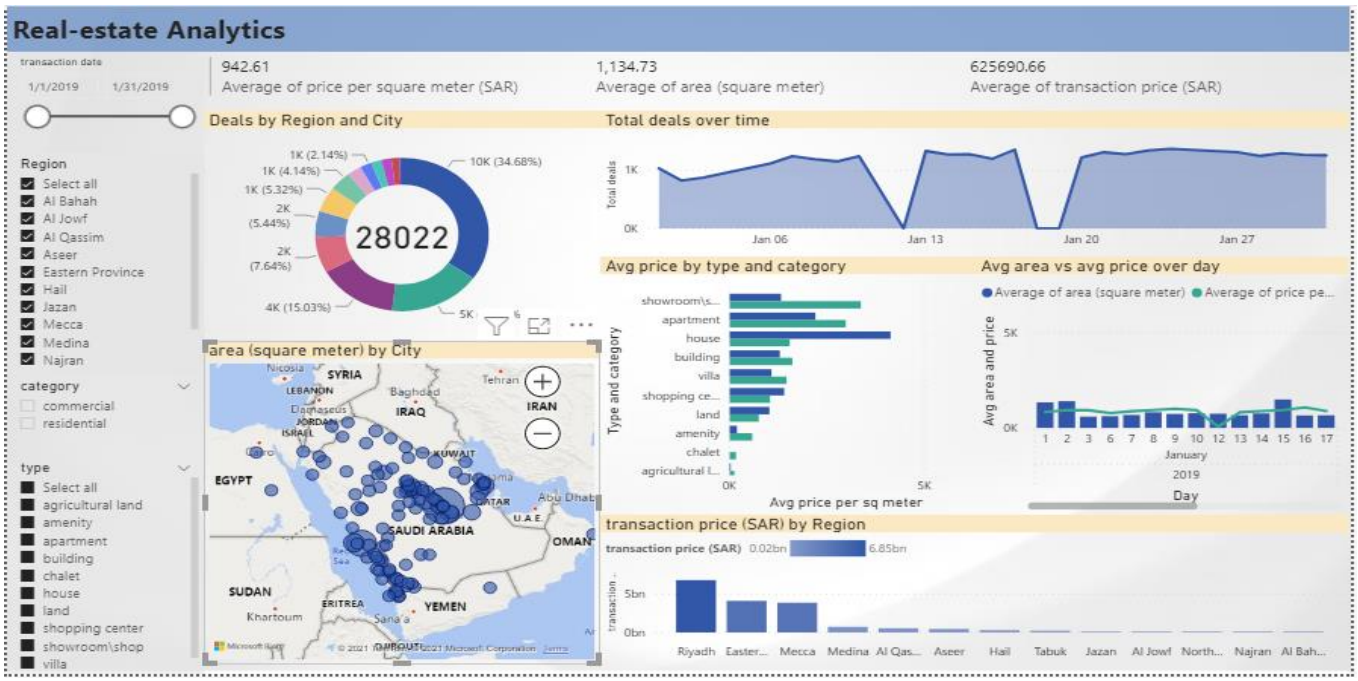

Figure 1: Real state dashboard

I have used transaction date, region, category, type as filters. The date slicer can be adjusted to view the data of specific time duration. All the other filters are multi-select.

Figure 2 shows a multi-row card that displays the average price of the land, average area, and average transaction price. This may help a seller to decide the price for a property for any region.

Figure 2: multi-row card of the average price

Considering the dataset for sellers, the key concern is how many deals he has done in a period and what are the most popular regions and cities for the deals. It shows the percentage of the deals that are done in a region. This can be further drilled down to get an insight into deals done in cities.

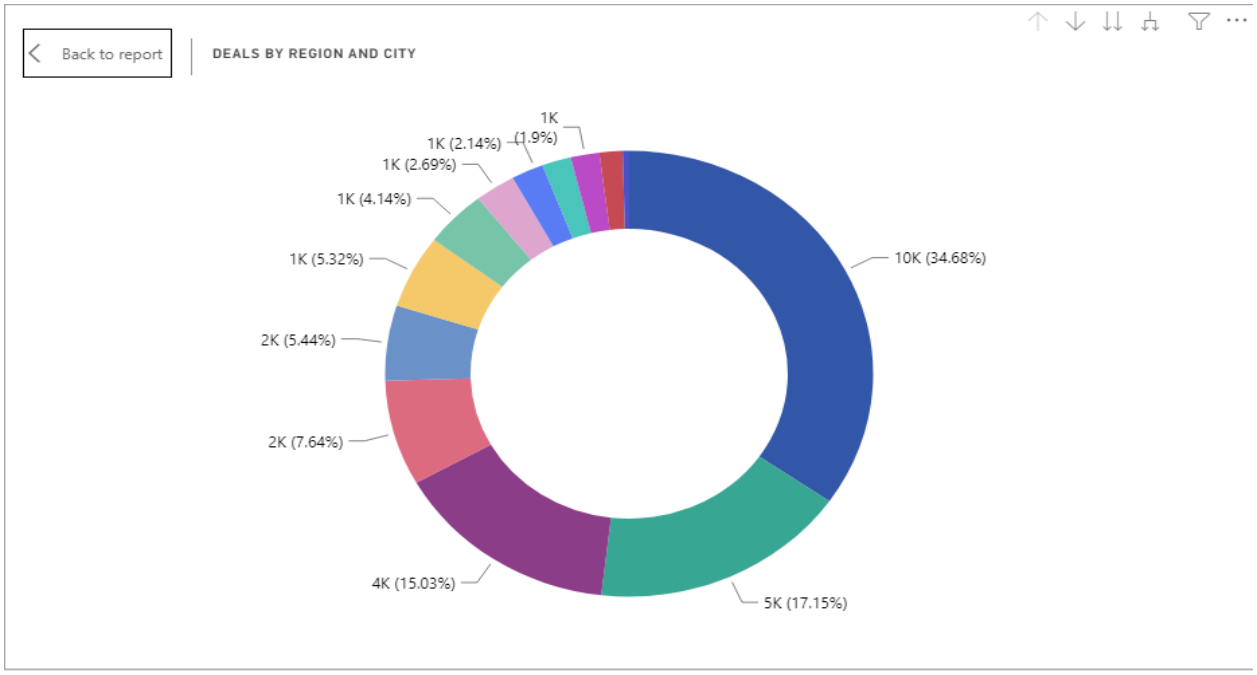

Figure 3: Deals by region and city

Figure 4 shows the breakout of deals by day. It is showing a trend about the deals over a period.

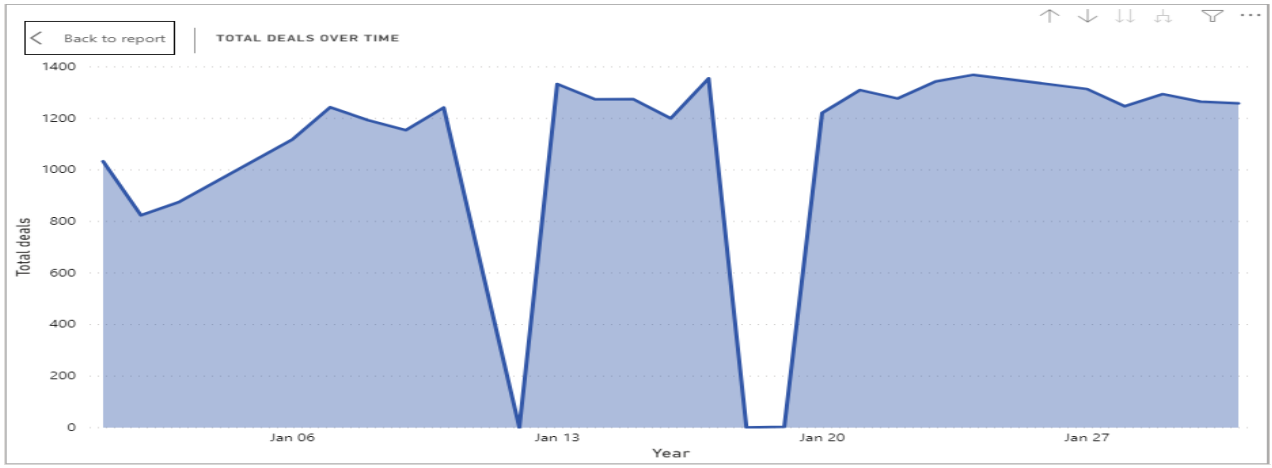

Figure 4: total deals over time

The next visual contains a pin map that represents the total area of a city that is involved in transactions.

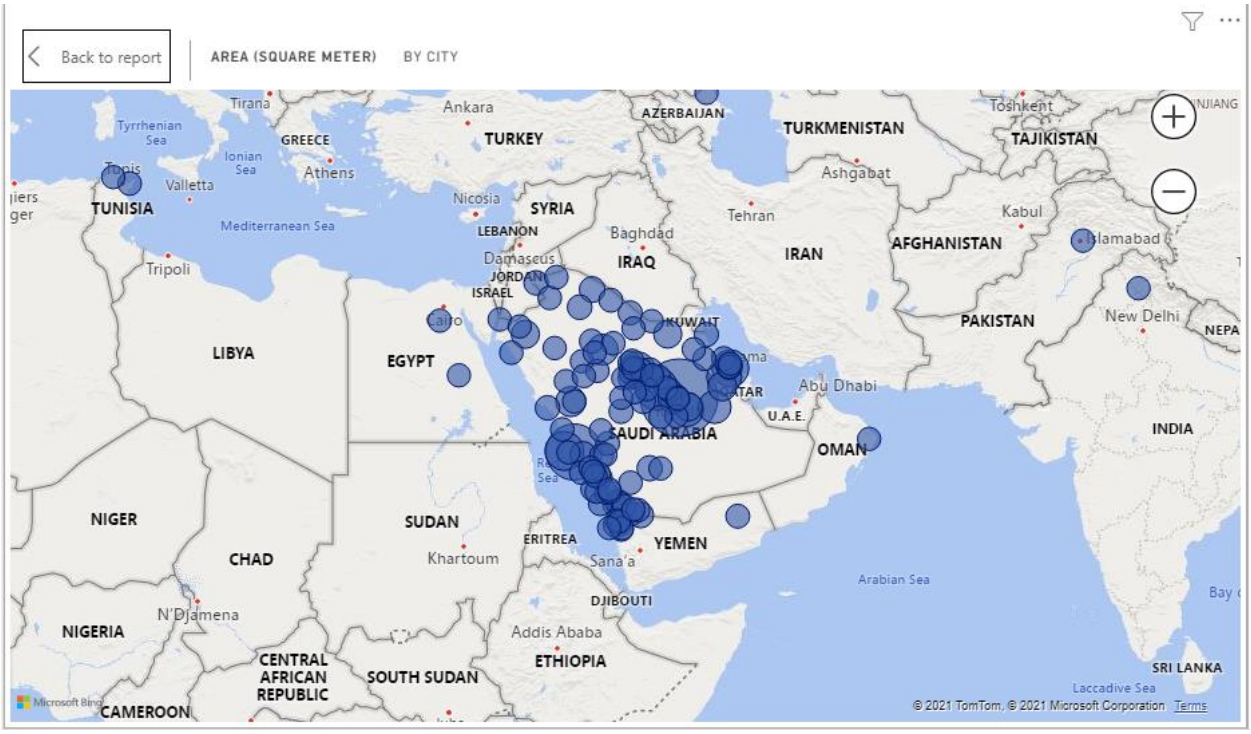

Figure 5: pin map of total area of a city

Figure 6 shows a comparison between the average price with types and categories. This helps in deciding the prices for a specific type and specific category. For example, in the given visual we can easily perceive that the house type of commercial category has the highest average price that is 4137.44 SAR.

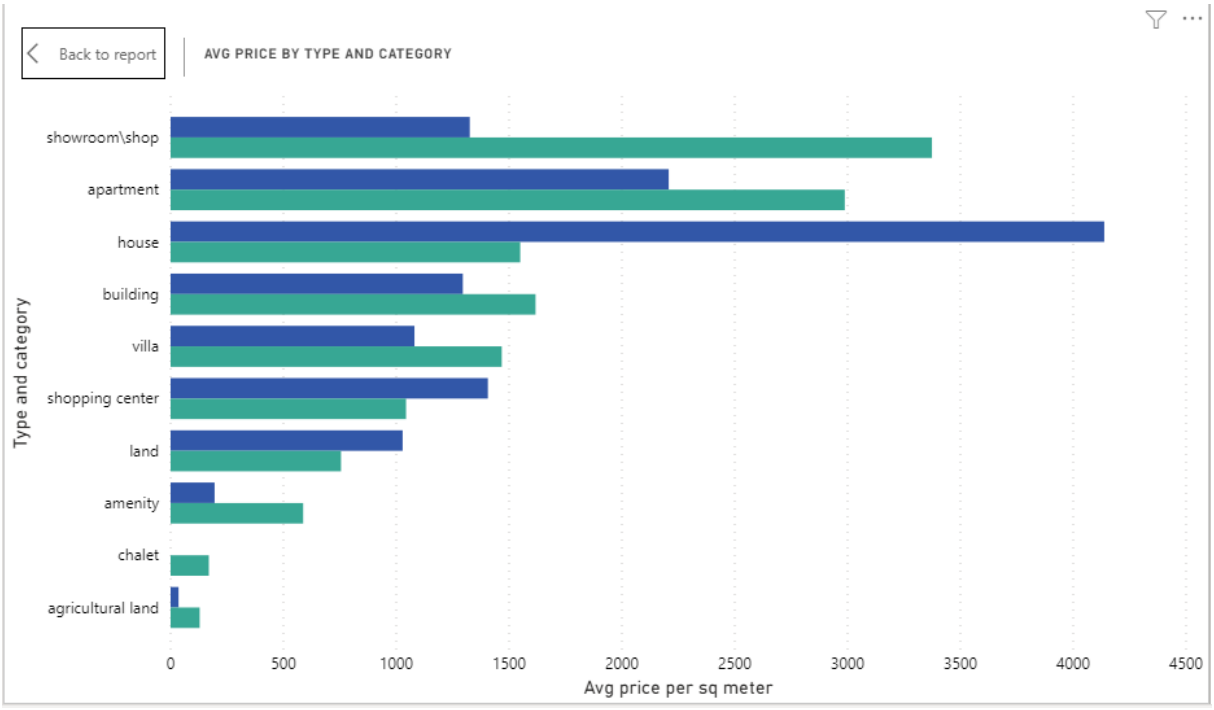

Figure 6: Average price by type and category

Figure 7 depicts the overall transaction price in SAR by region which can be further drilled down to cities.

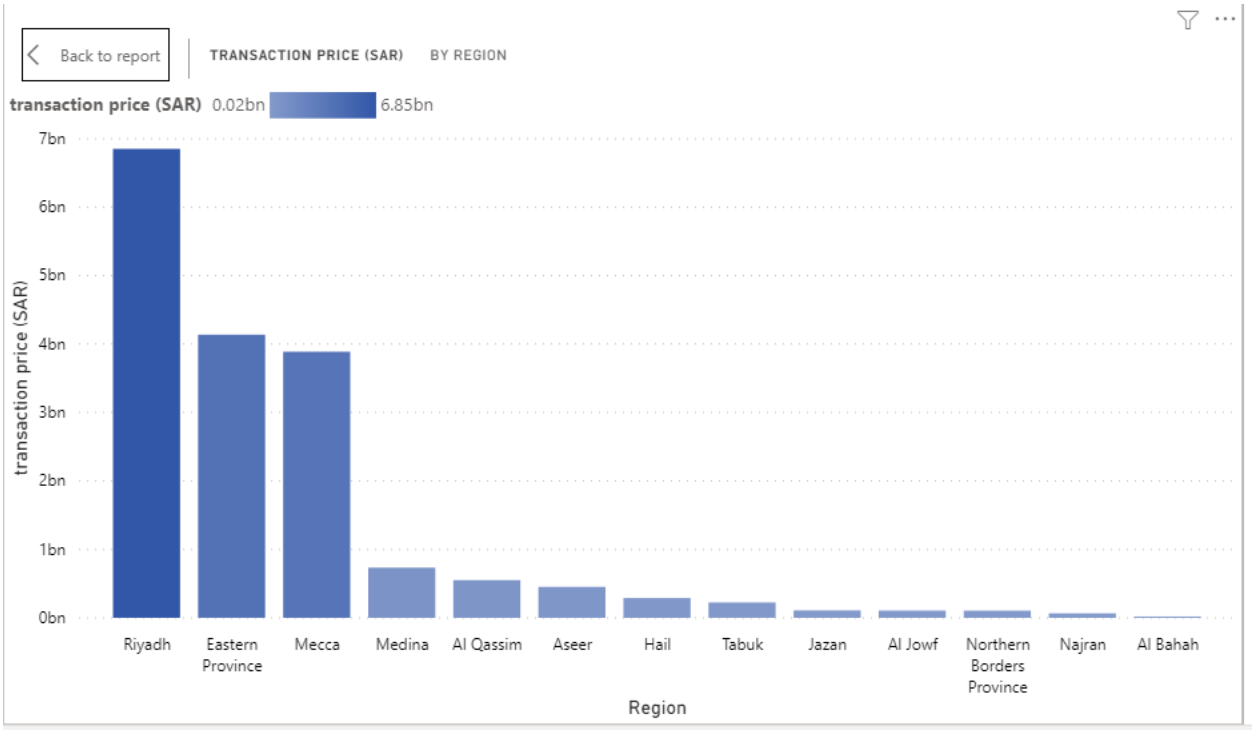

Figure 7: Overall transaction price (SAR)

2.4 Results

By using the data in the above dashboard, one can simply decide the prices for any area in any region or city. Along with this, it can be easy for any buyer to go through the visualizations and decide a decent amount of offer for any property. The seller can decide which the strongest region is for him in terms of deals and which region has what type of land in demand. The dashboard is published to Power BI services on this link.

3. Conclusion

Data science is playing an important role in decision-making. In this project, the dataset of real estate was used to display. An interactive dashboard involves users digging out the story of the data. Several analyses were made in the dashboard by using measures and data that are given.

There can be more to add to the dataset for a better picture of the data. For example, we can add the number of floors for houses and apartments that will help in deciding the price of it depending on the floors. Qualitative data can also be added to the dataset for better visualizations.

Acknowledgments

This work would not have been possible without the extraordinary support of our supervisor Dr. Mohamed Tahir. His knowledge, interests and attention to detail have made the work of this project go in the right direction from the beginning.

References

[1] Knight, D., Pearson, M., Schacht, B., & Ostrowsky, E. (2020). Microsoft Power BI Quick Start Guide. Packt. ISBN 978-1-80056-157-1

[2] Reani, M., Peek, N., & Jay, C. (2019). How different visualizations affect human reasoning about uncertainty: An analysis of visual behavior. Computers in Human Behavior, 92, 55–64. https://doi-org.sdl.idm.oclc.org/10.1016/j.chb.2018.10.033*Examples:*

[3] Liu, D., Bober, M., & Kittler, J. (2021). Visual Semantic Information Pursuit: A Survey. IEEE Transactions on Pattern Analysis & Machine Intelligence, 43(4), 1404–1422. https://doi-org.sdl.idm.oclc.org/10.1109/TPAMI.2019.2950025

[4] Liu, Z., Crouser, R. J., & Ottley, A. (2020). Survey on Individual Differences in Visualization. Computer Graphics Forum, 39(3), 693–712. https://doi-org.sdl.idm.oclc.org/10.1111/cgf.14033

[5] C. Capodiferro, & M. Mazzei. (2020). An Approach Adopted for Smart Data Generation and Visualization Problems. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, VI-4-W2-2020, 41–45. https://doi-org.sdl.idm.oclc.org/10.5194/isprs-annals-VI-4-W2-2020-41-2020

[6] Precious, C. (2021). Fruit sales data analysis and visualization with Power BI. Retrived from: https://medium.com/@precadeton/fruit-sales-data-analysis-and-visualization-with-power-bi-oftentimes-after-carrying-out-analysis-93ae90433483

[7] Chowdhury, K. (2021). Map-based visualizations in Power BI. Retrived from: https://www.analyticsvidhya.com/blog/2021/04/map-based-visualizations-in-power-bi/

[8] Lachev, T., & Price, E. (2018). *Applied Microsoft Power BI Bring your data to life!*. Prologika Press.

[9] Amrapali, B., & Upadhyay, A. (2017). Microsoft Power BI. *International Journal of Soft Computing and Engineering*, *7*(3), 1-94.

[10] Export transaction data. aqarsas. (2021). Retrieved December 9, 2021, from https://www.aqarsas.com/export/