Assignment 4

Creating Reports and Dashboards for Predictive and Prescriptive Analysis

HETA CHAVDA (NF1014555)

❖ **Objective**

The intention of this assignment is to improve one's ability to produce Power BI reports and dashboards using predictive and prescriptive datasets. The objective is to view the data, extract insights and analyze predictive and prescriptive to obtain actionable insights.
Part 1: Predictive Analysis with Housing Prices Dataset

❖ **Introduction**

Housing prices are influenced by numerous factors, including property characteristics, location, and market conditions. The objective of this analysis is to build a predictive model that accurately estimates housing prices based on historical data. This report details the entire workflow, including data preprocessing, exploratory data analysis, predictive modeling, and key findings. The ultimate goal is to generate actionable insights that can be used for real estate market evaluation and price forecasting.

❖ **Dataset Overview**

The dataset used for this analysis includes various attributes influencing housing prices. The key attributes are categorized as follows:

Property Characteristics: Lot Area, Year Built, Overall Quality, Number of Rooms, Basement Size, Garage Availability, and Building Type.

Location Factors: Neighborhood and Geospatial Coordinates.

Market Conditions: Historical Sale Prices and Seasonal Trends.

Target Variable: Sale Price (Final transaction price of houses).

❖ **Data Preparation & Cleaning Steps**

1. Handling Missing Values: Checked for missing values and applied appropriate imputation techniques (e.g., mean, median, mode substitution).

Visuals Included:

House Prices Analysis Overview (KPI Cards) – Properties with above Avg. Price, Average Price Per Square Ft., Total Number of Properties, Average SalePrice
Sale Price Distribution (Histogram) – Displays the overall price distribution across properties.
Sale Price by Garage Area vs. Lot Area (Scatter Plot) – Visualize garage size vs. lot size impact on price.

Distribution of Properties by Overall Quality (Donut Chart) – A donut chart displaying the proportion of properties categorized by their overall quality rating.
SalePrice by Overall Quality & Total Square Ft. (Scatter Plot) – A scatter plot illustrating the relationship between SalePrice, Overall Quality, and Total Square Footage, highlighting how larger and higher-quality properties tend to have higher sale prices.
Sale-Price by Year Built (Scatter Plot) – Analyzes how property age affects pricing. Interactive Filters (Slicers) – Allows users to explore trends dynamically.

Feature Engineering:

Converted categorical fields (e.g., Neighborhood, Building Type) into numerical labels. Created additional time-based features to enhance seasonal trend analysis.

Predictive Analysis Considerations:

Identified variables affecting price fluctuations.

Analyzed factors that can be optimized to improve real estate investment decisions.

The dashboard of the housing price effectively shows several important relations of the real state dataset. A scatter plot showing the relationship between lot area and lot frontage is in the upper left quadrant. Most of the data points are concentrated in the lower ranges (less than 50K sq ft for lot area and less than 100 ft for frontage), showing that most of the lots in the dataset are standard residential lots (Hung et al. 2023). The few outlier properties with much larger dimensions likely represent rural or estate properties.

The first panel on the top right display's sale prices by year from 2006 to 2010 and shows considerable volatility in the marketplace. It peaks in 2008 and dramatically falls by 2010, which may capture the range of the housing market collapse with the 2008-2009 financial crisis. The temporal visualization offers important context of the study period to understand market conditions.

The last visualization in the bottom left explains the relationship between this sale price and lot area as generally positive but with a scattered correlation. Larger lots generally get higher prices, but the wide scatter shows that other variables also strongly dictate the values of property. The last chart plots sale price against construction year, with a distinct bit of a spike in prices for homes built 2000 and yet, there is an obvious premium for newer construction.

The second dashboard image gives additional housing market insights through summary statistics as well as relationship analyses. This dataset contains 1, 460 houses, the maximum sale price was $755K, it occured between 2006 and 2010, and the range of years was from 1872 to 2010 (Roy et al. 2023). The two scatter plots here examine different finished square feet areas of homes, and both exhibit positive correlations, suggesting finished space consistently increases value, regardless of which floor it is on.

The multi-dimensional approach followed by bike rental dashboard is used to analyze the rental patterns. Data is presented in the table above left in the details of rental counts by various categories and time periods accounting for its micro usage patterns. The other horizontal bar chart next to it visualises monthly rental volumes with very strong seasonal variation, peak season throughout summer months and low season in winter.
An analysis of the relationship between weather conditions, months, and rental counts is done on the scatter plot in the upper right corner. The presence of higher concentration of blue triangles (probably representing good weather), and their predominance displays strong evidence that it is weather which has a substantial impact on behavior of rental, with good weather enhancing usage.

❖ **Conclusion**

This analysis successfully developed a predictive model that estimates housing prices with high accuracy. The study highlights key determinants of property values, including location, house size, age, and overall quality. These insights provide a foundation for improving property valuation models and enhancing decision-making processes in real estate markets.Further improvements, such as incorporating additional data sources and refining machine learning models, can enhance accuracy and applicability. The findings from this study can be leveraged by investors, policymakers, and real estate professionals to make data-driven decisions, optimize pricing strategies, and identify emerging trends in the housing market