

Real Estate Database Management System

Project Proposal - Group 3

Introduction:

I'm sure everyone is aware of the housing crisis in the country today. The average cost of buying a house has almost doubled when compared to 1986. Even after all this the number of houses being sold across the country hasn't reduced. So let's analyse this huge market and find out how the major middlemen like realtor agencies and agents sell houses and make a gargantuan profit on it.

Data:

For this project we gathered data from the real trends website (<https://www.realtrends.com/>), and from Kaggle as well. These sites contain resources about the top real estate agents and real estate agencies across the country and also the properties that were listed for sale across the states in the last year.

The exact sources combined are below:

- Top 100 real estate agents across the country,
<https://www.realtrends.com/ranking/best-real-estate-agents-united-states/individuals-by-volume/>
- Top realtor agencies across the country,
<https://www.realtrends.com/500-by-volume/>
- Property listings across the country,
<https://www.kaggle.com/datasets/fratzcan/usa-house-prices>

For each table we combined the necessary data to facilitate future queries. We've illustrated the description of our data below.

Table 1:

<u>Field</u>	<u>Type</u>	<u>Description</u>
PropertyID	Numeric	Unique ID for House
Price	Numeric	Price of House
Bedrooms	Numeric	Number of Bedrooms
Bathrooms	Numeric	Number of Bathrooms
Sqft	Numeric	Size of house in sqft

Floors	Numeric	Number of floors
YearBuilt	Numeric	Year it was built
YearRenovated	Numeric	Year it was renovated
Street	AlphaNumeric	Locality of the House
City	Character	City of the house
ZipCode	Numeric	ZipCode of the house
AgentID	Numeric	UniqueID of Agent

Table 2:

<u>Field</u>	<u>Type</u>	<u>Description</u>
AgentID	Numeric	Unique ID of Agent
First	Character	First Name of Agent
Last	Character	Last Name of Agent
AgentCity	Character	City of Agent
AgentState	Character	State of Agent
Volume	Numeric	Volume of Sales of Agent
AgencyID	Numeric	UniqueID of Agency

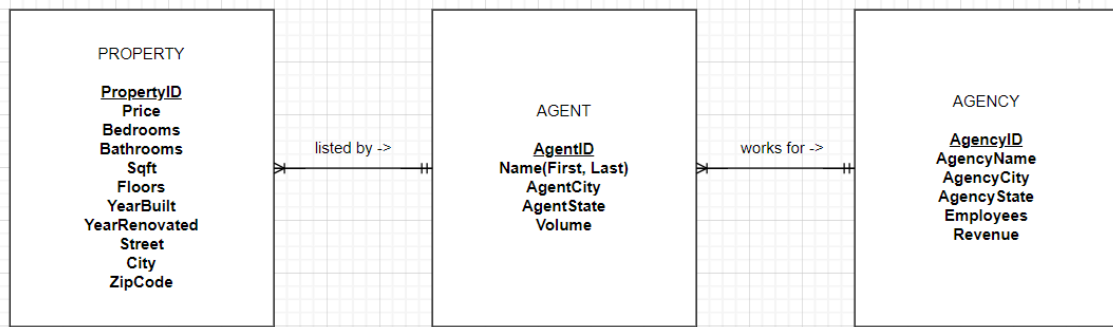
Table 3:

<u>Field</u>	<u>Type</u>	<u>Description</u>
AgencyID	Numeric	UniqueID of Agency
AgencyName	Character	Name of Agency
AgencyCity	Character	City of Agency
AgencyState	Character	State of Agency
Employees	Numeric	Number of Employees
Revenue	Numeric	Revenue of Agency

Database Design:

This database system has three entities- PROPERTY, AGENT and AGENCY. Each property can be listed by only one realtor but a realtor can show more than one listing with the minimum being one. So this is a one to many relationship. Similarly a realtor agent must belong to one agency only but the agency can have many realtors with the minimum being one. So this is again a one to many relationship.

The ERD model below displays the data.



All entities will have their own unique identifier in the ERD to eliminate weak entities. These 3 tables provide all the information and relationships needed to properly complete our data analysis within the Oracle APEX program.

Proposed Analysis:

Our analysis will focus on the following research questions.

1. Which three realtors have sold the most volume(total sales in dollars) of houses in the database? What percentage do they make up in the entire volume of all realtors?
2. Which two or three agencies(duopoly) have been the major market leader in the real estate business? What percentage of total share do they own in the entire revenue? Represent this in graphical form.
3. Which agency has the habit of creating the most successful realtors? This is calculated by dividing the total volume of an agency by the total number of realtors belonging to it.
4. Which states provide the most returns for the agents? How high are they when compared to the average price of properties across the board?

Team Members:

1. Gowri Shankar, Noodhan
2. Raghavendra, Nishanth Holla
3. Syed Abuthaheer, Sowfia
4. Singh, Saurabh Narendra
5. Rajamani Veni, Sushma Preethi