1. Query 1 - Find the top 10 cities from lowest to highest median house value

SELECT a.city, AVG(h.median\_house\_value) as median\_house\_value

FROM HouseOwnership ho

JOIN DimAddress a ON ho.add\_id = a.add\_id

JOIN DimHousing h ON ho.housing\_id = h.housing\_id

GROUP BY a.city

ORDER BY median\_house\_value ASC

LIMIT 10;

This query returns the top 10 cities from lowest to highest median house value by joining the HouseOwnership, DimAddress, and DimHousing tables. The result of this query can provide insights into the least expensive cities to buy a house in.

1. Query 2 - Find the distribution of housing median age by ocean proximity value

SELECT o.ocean\_proximity\_value,

CASE

WHEN h.housing\_median\_age BETWEEN 1 AND 20 THEN '1-20'

WHEN h.housing\_median\_age BETWEEN 21 AND 40 THEN '20-40'

WHEN h.housing\_median\_age BETWEEN 41 AND 62 THEN '40-62'

END AS age\_range,

COUNT(\*) as count

FROM HouseOwnership ho

JOIN DimHousing h ON ho.housing\_id = h.housing\_id

JOIN DimOceanProximity o ON ho.proximity\_id = o.proximity\_id

WHERE o.ocean\_proximity\_value = '{}'

AND h.housing\_median\_age IS NOT NULL

GROUP BY

CASE

WHEN h.housing\_median\_age BETWEEN 1 AND 20 THEN '1-20'

WHEN h.housing\_median\_age BETWEEN 21 AND 40 THEN '20-40'

WHEN h.housing\_median\_age BETWEEN 41 AND 62 THEN '40-62'

END

ORDER BY

CASE

WHEN h.housing\_median\_age BETWEEN 1 AND 20 THEN '1-20'

WHEN h.housing\_median\_age BETWEEN 21 AND 40 THEN '20-40'

WHEN h.housing\_median\_age BETWEEN 41 AND 62 THEN '40-62'

END;

This query returns the distribution of housing median age by ocean proximity value by joining the HouseOwnership, DimHousing, and DimOceanProximity tables. The result of this query can provide insights into the relationship between the age of houses and their proximity to the ocean.

1. Query 3 - Find the average loan amount by month and year

SELECT l.month, l.year, AVG(l.loan\_amount) as avg\_loan\_amount

FROM LoanSanction ls

JOIN DimLoan l ON ls.loan\_id = l.loan\_id

GROUP BY l.month, l.year

ORDER BY l.year, FIELD(l.month, 'January', 'February', 'March', 'April', 'May', 'June', 'July', 'August', 'September', 'October', 'November', 'December');

This query returns the average loan amount by month and year by joining the LoanSanction and DimLoan tables. The result of this query can provide insights into the seasonality of loan approvals and loan amounts.

1. Find the top 10 most common ocean proximity values for houses with a median house value greater than $500,000.

SELECT

dp.ocean\_proximity\_value,

COUNT(\*) AS num\_houses

FROM

DimHousing AS dh

JOIN HouseOwnership AS ho ON dh.housing\_id = ho.housing\_id

JOIN DimOceanProximity AS dp ON ho.proximity\_id = dp.proximity\_id

WHERE

dh.median\_house\_value > 500000

GROUP BY

dp.ocean\_proximity\_value

ORDER BY

num\_houses DESC

LIMIT 10;

Description: This query joins the DimHousing, HouseOwnership, and DimOceanProximity tables to find the top 10 most common ocean proximity values for houses with a median house value greater than $500,000. The WHERE clause filters the houses based on median house value, and the GROUP BY clause groups the houses by ocean proximity value. The ORDER BY clause sorts the results by the number of houses in descending order, and the LIMIT clause limits the results to the top 10.

Analytical operations: This query uses a join to combine data from multiple tables and filters the results based on a specific condition. It also groups the results and sorts them in descending order to find the top 10 values. This query provides insights into the most desirable ocean proximity values for expensive houses, which can help in identifying potential locations for future real estate developments.

Once upon a time, in a bustling city, a group of real estate analysts set out on a mission to uncover the housing market trends and patterns. Armed with their expertise and a vast database of housing information, they embarked on an exciting journey.

Their first query was focused on finding the cities with the most affordable housing options. By joining the tables of HouseOwnership, DimAddress, and DimHousing, they were able to calculate the average median house value for each city. Curious to explore the cities where housing was relatively inexpensive, they sorted the results in ascending order and limited the output to the top ten cities. This information would be useful for people looking for budget-friendly housing options.

Continuing their exploration, the analysts shifted their attention to understanding the relationship between the age of houses and their proximity to the ocean. By joining the tables of HouseOwnership, DimHousing, and DimOceanProximity, they were able to categorize houses based on their ocean proximity values and housing median ages. They created age ranges, such as '1-20,' '20-40,' and '40-62,' to group the houses accordingly. They then counted the number of houses in each age range for a specific ocean proximity value. The results provided valuable insights into how different age groups of houses were distributed across ocean proximity categories.

Intrigued by the time aspect of housing data, the analysts decided to investigate the average loan amounts sanctioned over time. They joined the tables of LoanSanction and DimLoan to gather information about loan amounts and their associated months and years. By grouping the data based on the month and year, they were able to calculate the average loan amount for each time period. To make the results more intuitive, they sorted the data chronologically by year and month names, following the sequence from January to December.

For their final query, the analysts aimed to identify the ocean proximity areas with the highest number of expensive houses. They joined the tables of DimHousing, HouseOwnership, and DimOceanProximity, filtering the houses based on a median house value threshold of $500,000. By grouping the houses according to ocean proximity values, they were able to count the number of houses in each proximity category. To reveal the most common ocean proximity areas for expensive houses, they sorted the results in descending order and limited the output to the top ten categories.

Armed with these valuable insights and data-driven conclusions, the real estate analysts were well-equipped to guide potential homebuyers, investors, and policymakers in making informed decisions about housing options, age considerations, loan amounts, and desirable locations.

And so, their journey through the world of housing data came to an end, leaving behind a wealth of knowledge and a path for future exploration in the ever-evolving realm of real estate.