REQUIREMENTS

1. How many records are there in the dataset?
   - Use COUNT(*) function
   - Select from the main table

2. How many unique cities are in the European dataset?
   - Use COUNT(DISTINCT) function
   - Apply it to the CITY column

3. What are the names of the cities in the dataset?
   - Use DISTINCT keyword
   - Select from the CITY column

4. How many bookings are there in each city?
   - Use COUNT(*) function
   - Group by CITY
   - Order results descending

5. What is the total booking revenue for each city?
   - Use SUM() function on the PRICE column
   - Group by CITY
   - Round the result
   - Order by total revenue descending

6. What is the average guest satisfaction score for each city?
   - Use AVG() function on GUEST_SATISFACTION column
   - Group by CITY

- Round the result
- Order by average score descending

7. What are the minimum, maximum, average, and median booking prices?
- Use MIN(), MAX(), AVG() functions on PRICE column
- Use PERCENTILE_CONT(0.5) for median
- Round results

8. How many outliers are there in the price field?
- Calculate Q1, Q3, and IQR using PERCENTILE_CONT()
- Define lower and upper bounds
- Count records outside these bounds

9. What are the characteristics of the outliers in terms of room type, number of bookings, and price?
- Create a view or CTE for outliers
- Group by ROOM_TYPE
- Use COUNT(), MIN(), MAX(), AVG() functions

10. How does the average price differ between the main dataset and the dataset with outliers removed?
- Create a view for cleaned data (without outliers)
- Calculate average price for both datasets
- Compare results

11. What is the average price for each room type?
- Use AVG() function on PRICE column
- Group by ROOM_TYPE

12. How do weekend and weekday bookings compare in terms of average price and number of bookings?

- Group by DAY column
- Use AVG() for price and COUNT() for bookings

13. What is the average distance from metro and city center for each city?

- Use AVG() on METRO_DISTANCE_KM and CITY_CENTER_KM columns
- Group by CITY

14. How many bookings are there for each room type on weekdays vs weekends?

- Use CASE statements to categorize room types
- Group by DAY and ROOM_TYPE

15. What is the booking revenue for each room type on weekdays vs weekends?

- Similar to previous question, but use SUM() on PRICE instead of COUNT()

16. What is the overall average, minimum, and maximum guest satisfaction score?

- Use AVG(), MIN(), MAX() functions on GUEST_SATISFACTION column

17. How does guest satisfaction score vary by city?

- Group by CITY
- Use AVG(), MIN(), MAX() on GUEST_SATISFACTION column

18. Is there a correlation between guest satisfaction and factors like cleanliness rating, price, or attraction index?

- Use CORR() function to calculate correlation coefficients

19. What is the average booking value across all cleaned data?

- Use AVG() function on PRICE column from cleaned data view

20. What is the average cleanliness score across all cleaned data?

- Use AVG() function on CLEANINGNESS_RATING column from cleaned data

21. How do cities rank in terms of total revenue?

- Use SUM() on PRICE column
- Group by CITY
- Use window function ROW_NUMBER() to assign ranks