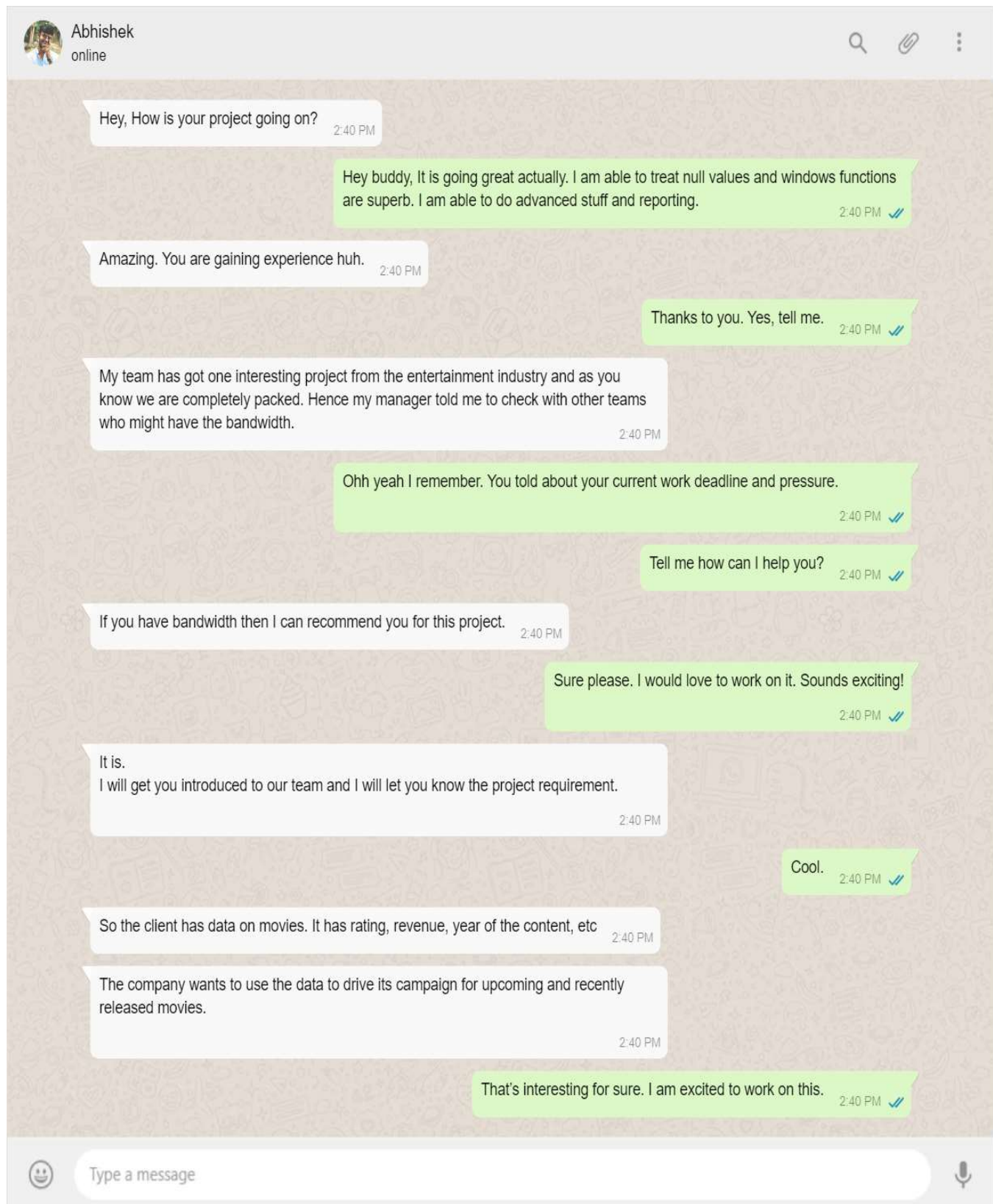




[Course](#) > [SQL Mo...](#) > [Part 6....](#) > Unguid...

Unguided Exercise



Dataset explanation:

IMDB is one of the most sought-after rating platforms for movies. The data also contains the movie titles, rating, revenue, year of the content, etc. The data is very helpful in studying the impact of rating on the revenue of the movie. The data is also very helpful in finding correlations between genres, rating, and revenue.

Below is the data dictionary for the dataset.

Dataset: imdb_movies

Column name	Sample value	Description
ID	1	ID of the movie
Title	Guardians of the Galaxy	Name of the movie
Genre	Action	Genre of the movie
Director	James Gunn	Director of the movie
Year	2014	Year in which movie was released
Runtime_minutes	121	Duration of movie
Rating	8.1	Rating of the movie
Votes	757074	Number of people rated the movie
Revenue_millions	333.13	Revenue in millions (Dollar)

Assignment 1

1/1 point (graded)

Create a report where all the movies are ranked based on revenue in their respective genre. What is the rank of 'Inception' movie in it's genre?



Assignment 2

1/1 point (graded)

Create a report where all the movies are ranked based on rating in their respective genre. What is the rank of 'Star Trek' movies in it's genre?

Note: In case of the same rating, movies with higher number of votes should be ranked higher.



Assignment 3

1/1 point (graded)

Compute the revenue contribution of movies towards the total revenue of respective genres. What is the revenue contribution of the 'Gone Girl' towards its genre?



Assignment 4

1/1 point (graded)

Compute the percentage rating in terms of top rated movies in their respective genre. For example, a movie in 'Action' genre has maximum rating of 9, then percentage rating for all the movies in 'Action' genre would be $(\text{rating} \times 100 / 9)$. What is the percentage rating of the movie 'Blood Diamond'?



Assignment 5

1/1 point (graded)

Segment the movies based on the below criteria using CASE WHEN

Revenue greater than or equal to 300 million : Blockbuster

Revenue between 200 and 299.99 million : Super hit

Revenue between 100 and 199.99 million: Hit

Else: Normal

How many movies are 'Super hit' in the year 2014?



Assignment 6

1/1 point (graded)

Segment the movies based on the below criteria using CASE WHEN

Revenue greater than or equal to 300 million : Blockbuster

Revenue between 200 and 299.99 million : Super hit

Revenue between 100 and 199.99 million: Hit

Else: Normal

What is the total revenue of the Blockbuster movies in the year 2015?



Assignment 7

1/1 point (graded)

Segment the movies based on ratings using IF statement as below:

Rating greater than or equal to 8: 'Must Watch'

Rating between 6 and 7.9: 'Can Watch'

Rating below 6: 'Avoid'

How many movies are a 'Must Watch' in the entire dataset?



Assignment 8

1/1 point (graded)

Compute total revenue by year and genre for all the movies with ratings less than 7. What is the total revenue of movies with ratings less than 7 in the year 2016 for 'Drama'? Use IF statement.



Assignment 9

1/1 point (graded)

How many movies have NULL values in the field revenue?



Assignment 10

1/1 point (graded)

What is the average rating of the movies which has revenue as NULL?



Assignment 11

1/1 point (graded)

If the revenue field for a movie is NULL, then assign the revenue as $\text{rating} \times \text{revenue factor}$.
Revenue factor = total revenue / total rating (where revenue is not NULL)

What is the average revenue of all the movies for the year 2016 considering imputed values?



Assignment 12

1/1 point (graded)

If the revenue field is NULL, consider it 0 using coalesce function.

What is the average revenue of all the movies for the year 2016 considering imputed values?



Assignment 13

1/1 point (graded)

A function that has no partial functional dependencies is in _____ form

☐ 1 NF☒ 2 NF☐ 3 NF☐ 4 NF

Assignment 14

1/1 point (graded)

Below table can be categorized in which normal form?

EMP_ID	EMP_NAME	EMP_PHONE	EMP_STATE
14	John	7272826385	UP
14	John	9064738238	UP
20	Harry	8574783832	Bihar
12	Sam	7390372389	Punjab
12	Sam	8589830302	Punjab

☒ 1 NF

☐ 2 NF

☐ 3 NF

☐ 4 NF



Submit

Assignment 15

1/1 point (graded)

The third normal form is based on the concept of ____

☐ Closure dependency

☒ Transitive Dependency

☐ Normal Dependency

☐ Functional Dependency

Assignment 16

1/1 point (graded)

Denormalization is useful in less execution time for

☐ Root cause analysis☐ Dashboards☐ Daily and weekly reports☒ Both B and C

Assignment 17

1/1 point (graded)

A table is no longer required. But we want to retain the table schema, so that we can use it to create similar tables again. Which of the following statements should be used?

☐ DROP☒ DELETE

Submit

Assignment 18

1/1 point (graded)

Below are the two representative queries on table1 (80 rows) and table2 (40 rows).

All the 40 rows in table 2 matches with table1.

```
Query 1:
SELECT *
FROM   table1 a
      LEFT JOIN table2 b
            ON a.column1 = b.column
```

```
Query 2:
SELECT *
FROM   table1 a
LEFT JOIN table2 b
where  a.column1=b.column1
```

- ☐ Both the queries will have same output (80 rows)
- ☐ Both the queries will have same output (40 rows)
- ☒ Query 1 will give 80 rows while Query 2 will give 40 rows
- ☐ Query 1 will give 40 rows while Query 2 will give 80 rows



Submit