



# ANALYSIS OF SALES DATA OF AMAZON



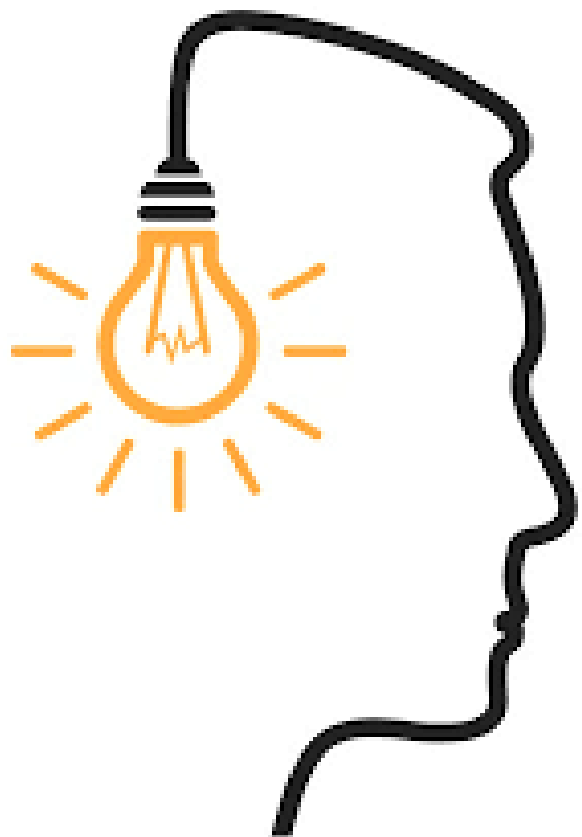
**Presented by Pankaj Kumar**





# Problem Statement

**As a data analyst at Amazon, a leading global e-commerce platform, you have been assigned the task of conducting a comprehensive analysis of sales data. The objective is to extract actionable insights by performing various SQL queries on the available datasets. This analysis will support data-driven decision-making processes and contribute to optimizing business strategies.**





# Methodology used in the analysis

**Data Source:** Amazon's internal sales database.  
**Tools:** MySQL for querying data from database.





**Problem:** List all products with a discounted price below ₹500?

**MySQL query for above problem:**



```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     discounted_price < 500;
```


The code block shows a MySQL query editor interface. On the left, a sidebar displays a database structure for 'amazondataset', including 'Tables' (with 'mytable' selected), 'Views', 'Stored Procedures', 'Functions', 'classicmodels', and 'newschema'. The main editor area shows a SQL query with line numbers 2 through 7. The query is: 'SELECT product\_name FROM mytable WHERE discounted\_price < 500;'. The 'WHERE' clause and its value are highlighted in a light gray background.



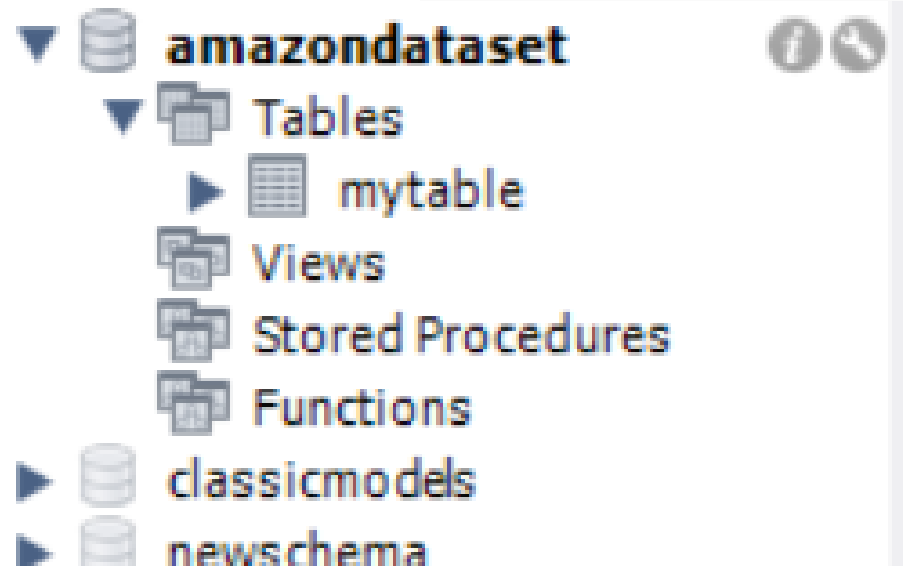


**Problem:** Find products with a discount percentage of 50% or more?

**MySQL query for above problem:**



```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     discount_percentage >= 0.50;
```



The image shows a MySQL database interface. On the left, a tree view displays the database structure: 'amazondataset' (expanded) contains 'Tables' (expanded) with 'mytable', 'Views', 'Stored Procedures', and 'Functions'. Below this are 'classicmodels' and 'newschema'. On the right, a SQL query editor shows a query with line numbers 2 through 7. The query is: 'SELECT product\_name FROM mytable WHERE discount\_percentage >= 0.50;'. The word 'FROM' is highlighted in a light gray box.





**Problem:** Retrieve all products where the name contains the word "Cable."?

MySQL query for above problem:



```
▼ amazondataset
  ▼ Tables
    ► mytable
  Views
  Stored Procedures
  Functions
  ► classicmodels
  ► newschema
```

```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     product_name LIKE '%cable%';
```





**Problem:** Display the difference between the average of the **actual price** and the **discounted price** for each product?

**MySQL query for above problem:**



```
▼ amazondataset
  ▼ Tables
    ▶ mytable
  Views
  Stored Procedures
  Functions
  ▶ classicmodels
  ▶ newschema
  ▶ sakila
  ▶ sys
```

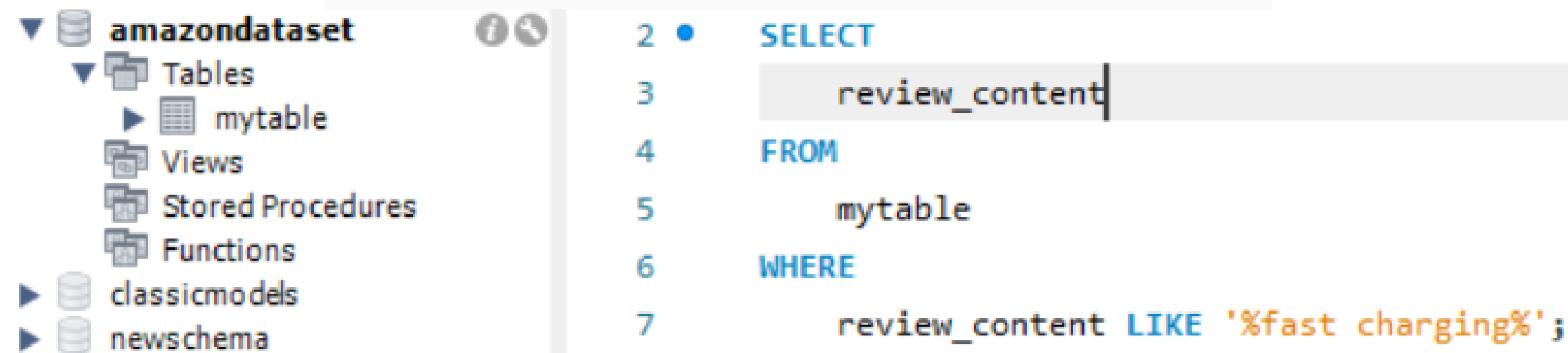
```
2 • SELECT
3     product_name,
4     AVG(actual_price) avg_actual_price,
5     AVG(discounted_price) avg_discounted_price,
6     AVG(actual_price) - AVG(discounted_price) avg_price_diff
7 FROM
8     mytable
9 GROUP BY product_name;
```





**Problem:** Query reviews that mention "fast charging" in their content?

**MySQL query for above problem:**

A screenshot of a MySQL query editor interface. On the left, a tree view shows the database structure: 'amazondataset' (expanded) contains 'Tables' (expanded) with 'mytable', 'Views', 'Stored Procedures', and 'Functions', and 'classicmodels' and 'newschema'. The main editor area shows a SQL query being typed: 'SELECT review\_content FROM mytable WHERE review\_content LIKE '%fast charging%';'. The text is color-coded: 'SELECT' is blue, 'review\_content' is black, 'FROM' is blue, 'mytable' is black, 'WHERE' is blue, and 'LIKE' is blue. The string '%fast charging%' is in orange. Line numbers 2 through 7 are visible on the left of the query text.

```
2 • SELECT
3   review_content
4 FROM
5   mytable
6 WHERE
7   review_content LIKE '%fast charging%';
```

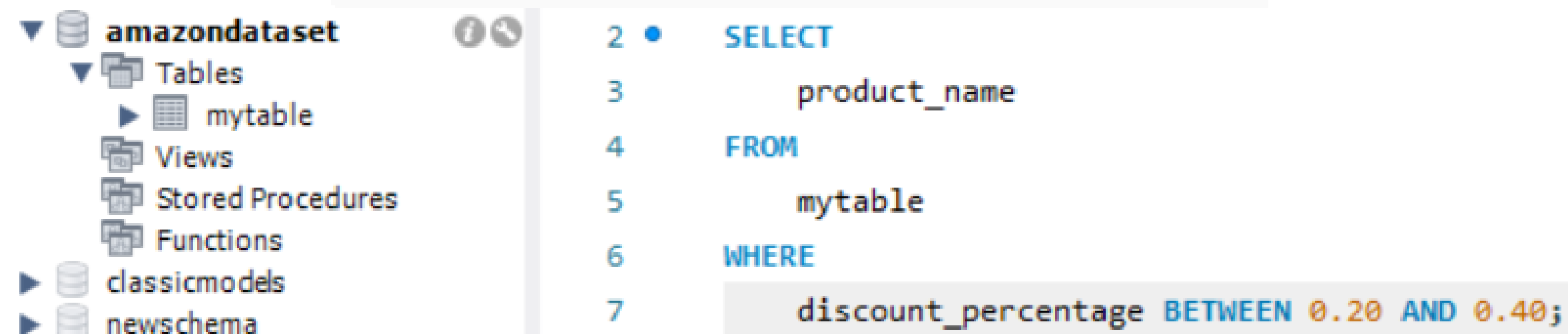






**Problem:** Identify products with a discount percentage between 20% and 40%?

MySQL query for above problem:

A screenshot of a MySQL database management tool interface. On the left, a tree view shows the database structure: 'amazondataset' is expanded, showing 'Tables' (with 'mytable' selected), 'Views', 'Stored Procedures', and 'Functions'. Below this are 'classicmodels' and 'newschema'. On the right, a SQL editor shows a query with line numbers 2 through 7. The query is: 'SELECT product\_name FROM mytable WHERE discount\_percentage BETWEEN 0.20 AND 0.40;'. The 'WHERE' clause is highlighted with a grey background.

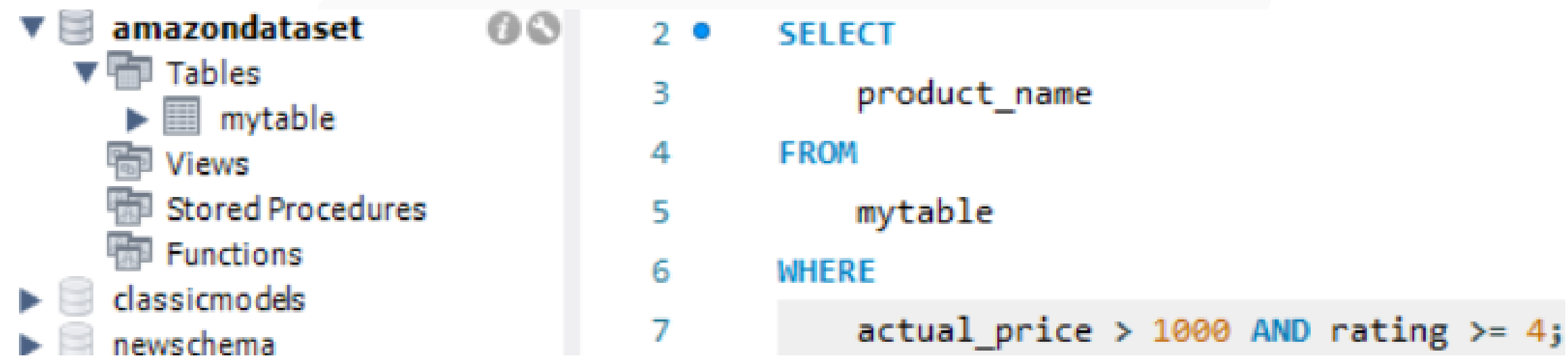
```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     discount_percentage BETWEEN 0.20 AND 0.40;
```





**Problem:** Find products that have an actual price above ₹1,000 and are rated 4 stars or above?

**MySQL query for above problem:**


A screenshot of a MySQL query editor interface. On the left, a tree view shows the database structure: 'amazondataset' (expanded) containing 'Tables' (expanded) with 'mytable', 'Views', 'Stored Procedures', and 'Functions', and other schemas like 'classicmodels' and 'newschema'. The main editor area shows a SQL query with line numbers 2 through 7. The query is: 'SELECT product\_name FROM mytable WHERE actual\_price > 1000 AND rating >= 4;'. The 'WHERE' clause is highlighted in a light gray box.

```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     actual_price > 1000 AND rating >= 4;
```



1. **Problem:** Find products where the discounted price ends with a 9?

MySQL query for above problem:



```
▼ amazondataset
  ▼ Tables
    ▶ mytable
  Views
  Stored Procedures
  Functions
  ▶ classicmodels
  ▶ newschema
```

```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     discounted_price LIKE '%9';
```





**Problem:** Display **review contents** that contains words like **worst**, **waste**, **poor**, or **not good**?

**MySQL query for above problem:**

A screenshot of a MySQL database management tool interface. On the left, a tree view shows the 'amazondataset' database selected, with sub-items for Tables, Views, Stored Procedures, and Functions. Below this, several other databases are listed: classicmodels, newschema, sakila, sys, test, and world. The main area on the right displays a SQL query in a text editor. The query is as follows:

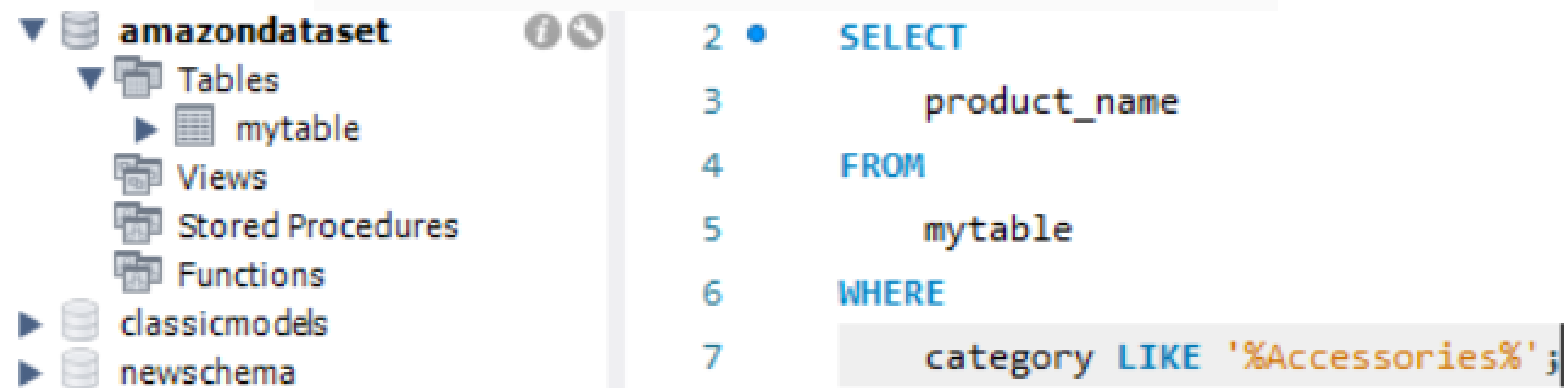
```
2 • SELECT
3     review_content
4 FROM
5     mytable
6 WHERE
7     review_content LIKE '%worst%'
8     OR review_content LIKE '%waste%'
9     OR review_content LIKE '%poor%'
10    OR review_content LIKE '%not good%';
```

The line numbers 2 through 10 are visible on the left side of the query editor. The last line of the query is highlighted with a light gray background.



**Problem:** List all products where the **category** includes "Accessories."?

**MySQL query for above problem:**

A screenshot of a MySQL query editor interface. On the left, a tree view shows the database structure: 'amazondataset' (expanded) contains 'Tables' (expanded) with 'mytable', 'Views', 'Stored Procedures', and 'Functions'; 'classicmodels'; and 'newschema'. The main area shows a SQL query with line numbers 2 through 7. The query is: 'SELECT product\_name FROM mytable WHERE category LIKE '%Accessories%';'. The 'WHERE' clause is highlighted with a grey background.

```
2 • SELECT
3     product_name
4 FROM
5     mytable
6 WHERE
7     category LIKE '%Accessories%';
```





Connect for more insights related to data analytics



[www.linkedin.com/in/pankaj-kumar-da](https://www.linkedin.com/in/pankaj-kumar-da)

