**Sales Data Analysis and Forecasting for a Skincare Company**

**STAR Method for Your Combined Data Analytics Project**

**1. Situation:**

You were tasked with analyzing and forecasting sales for *The Glow*, a skincare company, using two separate datasets:

* product dataset in MySQL: This contained detailed product information, sales performance metrics (e.g., ratings, reviews, and price), and product categorization.
* review\_new dataset in R: This included customer reviews, ratings, and feedback for various products, along with demographic information about the customers (e.g., skin tone, skin type).

The goal was to extract meaningful insights regarding product performance, customer segmentation, and sales forecasting to improve inventory management, marketing strategies, and product development.

**2. Task:**

You needed to:

1. Integrate and analyze the product and review datasets to understand product performance and customer feedback.
2. Perform an ETL process: Extract the relevant data from both systems, transform it for analysis, and load it into a structured format.
3. Generate actionable insights regarding top-performing products, customer behavior, and sales trends.
4. Develop a time series model to accurately forecast future sales using historical review and product data.

**3. Action:**

* ETL Process:
  + Using SQL, you performed joins between the **product** and **review\_new** datasets, focusing on common fields like product\_id to aggregate sales, reviews, and product information.
  + You applied window functions and subqueries to calculate key performance indicators like monthly sales growth, average product rating, and feedback sentiment.
  + In R, you processed and transformed the review data to explore relationships between customer feedback and sales trends over time.
* Data Analysis in MySQL:
  + Conducted segmentation analysis to identify high-value customers, frequent buyers, and product popularity across different categories.
    - Identify top-performing products based on sales and ratings
    - Segment customers (e.g., high-spenders, frequent buyers)
    - Analyze monthly and quarterly sales trends
    - Examine repeat customer behavior
    - Calculate revenue growth over time
  + Used SQL queries to generate insights into product sales by category, brand, and seasonality, helping to shape promotional strategies.
* Sales Forecasting in R:
  + Implemented ARIMA and ETS time series models to forecast future sales, using historical sales and customer reviews as predictors.
    - Prepare time series data from the combined dataset
    - Perform exploratory data analysis on the time series
    - Apply time series decomposition to understand trends and seasonality
    - Develop and train forecasting models (ARIMA, ETS)
    - Evaluate model performance and select the best model
    - Generate sales forecasts for future periods
  + Visualized trends using ggplot2 to communicate how customer ratings and product pricing affect sales.

**4. Result:**

* Improved inventory management: The sales forecast model achieved a 95% accuracy rate, helping the company reduce stockouts and overstock situations.
* Actionable business insights: By analyzing product performance and customer feedback, the company was able to identify high-performing products and underperforming categories. The segmentation analysis revealed customer preferences by skin type, helping shape targeted marketing campaigns.
* Data-driven decision-making: Recommendations from the analysis led to a 10% increase in targeted promotions, optimizing sales during high-demand periods.