Bike Sales Data Analysis Report

Objective: To analyze the sales data of a bike retail company, focusing on customer demographics, store performance, sales trends, and product popularity.

The dataset contains 100,000 records with 11 columns. Here's a brief overview of the columns:

1. **Sale_ID**: Unique identifier for each sale.
2. **Date**: Date of the sale (format: DD-MM-YYYY).
3. **Customer_ID**: Unique identifier for each customer.
4. **Bike_Model**: The model of the bike sold.
5. **Price**: Price of each bike.
6. **Quantity**: Number of bikes sold in a single transaction.
7. **Store_Location**: Location of the store where the sale took place.
8. **Salesperson_ID**: Unique identifier for the salesperson handling the transaction.
9. **Payment_Method**: Method of payment used by the customer (e.g., Apple Pay, PayPal).
10. **Customer_Age**: Age of the customer.
11. **Customer_Gender**: Gender of the customer.

Basic Dataset Overview

```
-- Count total records in the dataset

SELECT COUNT(*) AS total_records FROM bike_sales_100k;

-- View a few sample records

SELECT * FROM bike_sales_100k LIMIT 10;
```

Sales Summary

-- Total sales amount (Price * Quantity)

SELECT SUM(Price * Quantity) AS total_sales FROM bike_sales_100k;

-- Average price of bikes sold

SELECT AVG(Price) AS average_price FROM bike_sales_100k;

Top 5 Selling Bike Models

-- Bike models with the highest number of sales

SELECT Bike_Model, SUM(Quantity) AS total_units_sold

FROM bike_sales_100k

GROUP BY Bike_Model

ORDER BY total_units_sold DESC

LIMIT 5;

Top 5 Customers by Spending

-- Find the top 5 customers who spent the most

SELECT Customer_ID, SUM(Price * Quantity) AS total_spent

FROM bike_sales_100k

GROUP BY Customer_ID

ORDER BY total_spent DESC LIMIT 5;

Sales by Store Location

```
SELECT Customer_ID, SUM(Price * Quantity) AS total_spent
FROM bike_sales_100k
GROUP BY Customer_ID
ORDER BY total_spent DESC
LIMIT 5;
```

Sales by Payment Method

```
SELECT Payment_Method, SUM(Price * Quantity) AS total_sales
FROM bike_sales_100k
GROUP BY Payment_Method
ORDER BY total_sales DESC;
```

Top 5 Product Categories by Revenue

```
SELECT Bike_Model, SUM(Price * Quantity) AS total_revenue
FROM bike_sales_100k
GROUP BY Bike_Model
ORDER BY total_revenue DESC
LIMIT 5;
```

Customer Segmentation by Purchase Amount

```
-- Segment customers based on total amount spent
SELECT Customer_ID, SUM(Price * Quantity) AS total_spent,
    CASE
        WHEN SUM(Price * Quantity) >= 10000 THEN 'High Spenders'
        WHEN SUM(Price * Quantity) BETWEEN 5000 AND 9999 THEN
'Medium Spenders'
        ELSE 'Low Spenders'
    END AS customer_segment
FROM bike_sales_100k
GROUP BY Customer_ID
ORDER BY total_spent DESC;
```

Sales by Customer Gender

```
SELECT Customer_Gender, SUM(Price * Quantity) AS total_sales
FROM bike_sales_100k
GROUP BY Customer_Gender
ORDER BY total_sales DESC;
```

Age Group Analysis

```
SELECT CASE
    WHEN Customer_Age BETWEEN 18 AND 30 THEN '18-30'
    WHEN Customer_Age BETWEEN 31 AND 45 THEN '31-45'
    WHEN Customer_Age BETWEEN 46 AND 60 THEN '46-60'
    ELSE '60+'
END AS age_group,
SUM(Price * Quantity) AS total_sales
FROM bike_sales_100k
GROUP BY age_group
ORDER BY total_sales DESC;
```

Most Popular Salesperson by Revenue

```
SELECT Salesperson_ID, SUM(Price * Quantity) AS total_revenue
FROM bike_sales_100k
GROUP BY Salesperson_ID
ORDER BY total_revenue DESC
LIMIT 5;
```

Top Payment Method by Age Group

```
SELECT
    CASE
        WHEN Customer_Age BETWEEN 18 AND 30 THEN '18-30'
        WHEN Customer_Age BETWEEN 31 AND 45 THEN '31-45'
        WHEN Customer_Age BETWEEN 46 AND 60 THEN '46-60'
        ELSE '60+'
    END AS age_group,
    Payment_Method,
    COUNT(*) AS method_usage_count
FROM bike_sales_100k
GROUP BY age_group, Payment_Method
ORDER BY age_group, method_usage_count DESC;
```

Highest Revenue Day of the Year

```
SELECT Date, SUM(Price * Quantity) AS total_revenue
FROM bike_sales_100k
GROUP BY Date
ORDER BY total_revenue DESC
LIMIT 1;
```

Insights:

1. The dataset consists of 100,000 records, including details like sale ID, date, customer age, gender, bike model, price, quantity, payment methods, and store location.
2. The total sales amount is $778,434,234.95, and the average bike price is $2,598.18.
3. The top 5 selling bike models by quantity are Cruiser, Hybrid Bike, BMX, Road Bike, and Folding Bike.
4. The top 5 customers by total spending spent between $225,651 and $265,308.
5. Sales by store location show New York leading with $113.6 million in total sales, followed by Phoenix and Philadelphia.
6. Debit Card is the most popular payment method, with Apple Pay and Credit Card following closely.
7. The top 5 bike models by revenue are Hybrid Bike, BMX, Cruiser, Road Bike, and Folding Bike.
8. Customer segmentation reveals that high spenders (those spending over $200,000) account for a large portion of revenue.
9. Sales by gender show that females contributed $392.3 million in total sales, while males contributed $386.1 million.
10. Customers aged 31-45 and 46-60 are the largest revenue contributors, with each group contributing over $220 million in sales.
11. The most successful salesperson (ID 794) generated $1,196,096 in sales.
12. Apple Pay is the most popular payment method for customers aged 18-30, while PayPal is favored by those aged 46-60.
13. The highest revenue day of the year was August 10, 2024, with $728,842.21 in sales.

Conclusion & Recommendations

- Top-Performing Models: Focus on promoting high-selling models like Cruiser and Hybrid Bikes.
- Store Expansion: Consider opening more stores in high-performing cities like New York and Phoenix.
- Payment Method Preferences: Ensure seamless payment processing for Debit Card and Apple Pay users.
- Customer Segmentation: Target high-spending customers with loyalty programs or special offers.
- Salesperson Training: Invest in training for salespeople to boost overall revenue.