Unlocking Sales Insights: An Analysis of Sports Superstore Data Using MySQL

Overview: The Sports Store company is committed to increasing sales and profitability, leveraging our loyal customer base and dedication to customer satisfaction. To stay ahead of the competition and improve performance, they seek assistance from the data analytics team. As a data analyst, my role is crucial in identifying opportunities, targeting marketing efforts, and making data-driven decisions. The assigned project involves tasks such as analyzing profit margins by brand, popular subcategories within each brand, successful products, and preferred colors among customers, considering seasonality, delivery methods, and customer locations. The dataset, Sports Superstore, comprising sales, products, and customer tables, will be used for analysis. The final deliverable is a concise and accurate report summarizing findings and providing insights to enhance the company's bottom line.

1. **Color Prefernce Analysis for Women's Shoes:**

   The task is to discover which colors women are most interested in when selecting shoes. The goal is to identify the colors they rate highest so that we can apply a color filter whenever a customer searches for women's shoes.

```sql
SELECT color,
       avg(average_rating) AS rating
FROM products
WHERE sub_category = "Women/Shoes"
GROUP BY color
ORDER BY rating DESC;
```

2. **Color Selection for New Product Line:**

   The task is to provide a list of the highest-selling colors. We need to provide a list of five colors to forward to the supply chain department for procurement in those colors only. The company is planning to introduce several new items and these colors would be preferred while procuring new items.

```sql
SELECT p.color, SUM(s.items_sold) AS items_sold
FROM products AS p
INNER JOIN
      (
      SELECT MAX(product_id) AS product_id,
      SUM(quantity) AS items_sold
      FROM sales
      GROUP BY product_id
      ) AS s ON s.product_id=p.product_id
GROUP BY p.color
ORDER BY items_sold DESC
LIMIT 5
```

3. **Men's Product Analysis for Improving Sales:**

The task is to provide a list of 50 men's products that have high ratings and at least 500 reviews. The company is currently experiencing low sales of men's products. Our research shows that good ratings and a high number of reviews are the most important factors in the buying decision for men.

```sql
SELECT
    product_id,
    MAX(product_name) as product_name,
    AVG(average_rating) AS rating,
    SUM(reviews_count) AS reviews_count
FROM products
WHERE sub_category LIKE 'men%'
GROUP BY product_id
HAVING SUM(reviews_count) > 500
ORDER BY rating DESC
LIMIT 50
```

4. **Identifying Top Customers for Loyalty Rewards:**

You task is to provide a list of 100 customers who have bought the maximum number of products over time along with the number of products they have bought.

```sql
SELECT s.customer_id,
       c.customer_name,
       SUM(s.quantity) AS items_bought
FROM sales s
INNER JOIN customer c
ON s.customer_id = c.customer_id
GROUP BY s.customer_id, c.customer_name
ORDER BY items_bought DESC
LIMIT 100
```

5. **Brand Research for Men's Sports Shoes Category:**

   The task is to provide the names of two brands that are most preferred by men (in terms of quantity sold) when it comes to shoes.

```sql
SELECT p.brand_name,
       SUM(s.quantity) as total_items_sold
FROM products AS p
INNER JOIN sales as s ON p.product_id = s.product_id
WHERE p.sub_category = 'men/shoes'
GROUP BY p.brand_name
ORDER BY total_items_sold DESC
LIMIT 2;
```

6. **Research and Evaluation of Profitable Sub-Categories for New Product Introduction:**

   The task is to prepare a list of 10 sub_categories with the highest percentage of profit margin that can be introduced into our stores and help us increase our profit.

```sql
SELECT products.sub_category,
       (SUM(sales.profit) / SUM(sales.sales)) * 100 AS
profit_margin
FROM products
INNER JOIN sales
```

```
ON products.product_id = sales.product_id
GROUP BY products.sub_category
ORDER BY profit_margin DESC
LIMIT 10;
```

7. **Color Stocking Strategy for Optimal Sales and Positive Rewiews:**

   The task is to find the list of colors and the corresponding sum of sales for only those products which have at least 100 reviews and an average rating of 4+.

```
SELECT p.color,
        SUM(s.sales) AS sum_sales
FROM products p
INNER JOIN sales s
ON p.product_id = s.product_id
WHERE p.reviews_count >= 100
GROUP BY p.color
HAVING AVG(p.average_rating) > 4
ORDER BY sum_sales DESC;
```

8. **Analysis of Delivery Performance to Improve sales:**

   The task is to find the list of colors and the corresponding sum of sales for only those products which have at least 100 reviews and an average rating of 4+.

```
SELECT delivery_mode,
       SUM(sales) AS total_sales,
       SUM(profit) AS total_profit,
       COUNT(DISTINCT order_id) AS order_count
FROM sales
WHERE DATE(order_date) >= DATE_SUB(CURDATE(), INTERVAL 1 YEAR)
GROUP BY delivery_mode
ORDER BY total_sales DESC
LIMIT 10
```

9. **Identifying Brands with High Average Profit Per Order:**

The task is to provide the names of brands along with the average average profit per order during the past year sorted by average profit per order from maximum to minimum.

```
SELECT brand_name,
       (total_profit / total_orders) AS avg_profit_per_order
FROM (
SELECT p.brand_name, SUM(s.profit) AS total_profit,
COUNT(DISTINCT s.order_id) AS total_orders
FROM sales s
JOIN products p
ON s.product_id = p.product_id
WHERE s.order_date >= DATE_SUB(CURDATE(), INTERVAL 1 YEAR)
GROUP BY p.brand_name
) AS c
ORDER BY avg_profit_per_order DESC;
```

10. **Allocation for Men's Shoes Inventory for Three States:**

The task is to find the percentage of shoes sold last year in these three states so that we can provide that percentage from our production to each of these three states. we need to carefully allocate our inventory to ensure that we can meet the demand in all of our locations. Some of the key markets we are focusing on are California, New York, and Texas.

```
SELECT (SUM(s.quantity) / 1365) * 100 AS
percentage_of_products_sold,
        c.state,
        p.sub_category
FROM sales s
JOIN customer c ON s.customer_id = c.customer_id
JOIN products p ON p.product_id = s.product_id
WHERE p.sub_category = 'Men/Shoes'
AND DATE(s.order_date) >= DATE_SUB(CURDATE(), INTERVAL 1 YEAR)
AND c.state IN ('California', 'New York', 'Texas')
GROUP BY c.state, p.sub_category
```

```
ORDER BY percentage_of_products_sold DESC;
```

11. Personalized Email Marketing to Drive Sales and Build Customer Relationships:

The task is to provide a list of our five best-selling (in terms of quantity in the state of California) shoes that are white in color. This is to drive sales and provide a personalized experience to customers, and send them a customized email offering a promotion.

```
SELECT p.product_name,
       MAX(p.brand_name) AS brand_name,
       MAX(p.color) As color,
       MAX(c.state) AS state,
       MAX(p.sub_category) AS sub_category,
      SUM(s.quantity) AS total_products_sold
FROM products p
JOIN sales s ON p.product_id = s.product_id
JOIN customer c ON s.customer_id = c.customer_id
WHERE p.color = 'White' AND c.state = 'California' AND
p.sub_category = 'Men/Shoes'
GROUP BY p.product_name
ORDER BY total_products_sold DESC
LIMIT 5;
```

12. Optimizing Stock for New Store Opening in New York based on Subcategory Profitability:

The task is to provide a list of four subcategories from each category, ranked based on the total profit they are generating in New York.

```
SELECT category,
       sub_category,
       total_quantity
FROM (SELECT *, RANK() OVER (PARTITION BY category ORDER BY
total_profit DESC) AS rank_
```

```
FROM (SELECT max(p.category) as category,
p.sub_category,
sum(s.profit) total_profit,
sum(s.quantity) AS total_quantity,
c.state
FROM products AS p
JOIN sales AS s ON p.product_id = s.product_id
JOIN customer AS c ON s.customer_id = c.customer_id
WHERE c.state = "New York"
GROUP BY p.sub_category
ORDER BY total_profit DESC) as subcat_profit
) as subcat_ranked
WHERE rank_ <= 4
```

13. Brand Performace Analysis for Summer 2019:

The task is to find out which brand had the highest profit during the summer season (the months of June, July and August) of 2019.

```
SELECT brand_name,
       SUM(profit) AS total_profit
FROM sales
JOIN products ON sales.product_id = products.product_id
WHERE MONTH(order_date) BETWEEN 6 AND 8 AND YEAR(order_date) =
2019
GROUP BY brand_name
ORDER BY total_profit DESC
LIMIT 1;
```

14. Brand Performace Analysis for Summer 2019:

The task is to determine the top 2 brands with the highest total sales value for the products belonging to the most profitable category during the specified time period of January 1, 2019, to June 30, 2019. It's essential that the delivery mode used is "Standard Delivery".

```sql
SELECT brand_name,
       SUM(sales) AS total_sales
FROM sales
INNER JOIN products ON sales.product_id = products.product_id
WHERE order_date BETWEEN '2019-01-01' AND '2019-06-30'
AND delivery_mode = 'Standard Delivery'
AND category = (
SELECT category
FROM (
SELECT category,
       SUM(profit) AS total_profit_per_category
FROM sales
INNER JOIN products ON sales.product_id = products.product_id
WHERE order_date BETWEEN '2019-01-01' AND '2019-06-30'
AND delivery_mode = 'Standard Delivery'
GROUP BY category
ORDER BY total_profit_per_category DESC
LIMIT 1
) AS z
)
GROUP BY brand_name
ORDER BY total_sales DESC;
```