**Marketing Analytics Project - Milestone 1: Final Deliverables**

### 1. SQL Script File - sql\_queries.sql

This file includes all the SQL queries used to analyze the customer and campaign data. Each query focuses on answering a specific business question based on customer behavior, acquisition performance, and churn risk.

-- Use the correct database  
USE marketing\_analytics;  
  
-- Q1: Show how many customers came from each acquisition channel  
SELECT acquisition\_channel, COUNT(\*) AS customer\_count  
FROM customers  
GROUP BY acquisition\_channel;  
  
-- Q2: Show average spending and number of orders for each customer segment  
SELECT customer\_segment,  
 ROUND(AVG(total\_spent), 2) AS avg\_spending,  
 ROUND(AVG(number\_of\_orders), 1) AS avg\_orders  
FROM purchase\_behavior  
GROUP BY customer\_segment;  
  
-- Q3: List customers who haven't purchased for over 90 days but are not marked as churned  
SELECT customer\_id  
FROM purchase\_behavior  
WHERE days\_since\_last\_purchase > 90 AND churned = 0;  
  
-- Q4: Combine email clicks and total spending to check for any correlation  
SELECT ci.customer\_id, ci.email\_clicks, pb.total\_spent  
FROM campaign\_interactions ci  
JOIN purchase\_behavior pb ON ci.customer\_id = pb.customer\_id;  
  
-- Q5: Show average spending grouped by gender  
SELECT c.gender, ROUND(AVG(pb.total\_spent), 2) AS avg\_spent  
FROM customers c  
JOIN purchase\_behavior pb ON c.customer\_id = pb.customer\_id  
GROUP BY c.gender;  
  
-- Q6: Estimate campaign ROI as total\_spent divided by total clicks (email + ad clicks)  
SELECT ci.customer\_id,  
 ROUND(pb.total\_spent / NULLIF((ci.email\_clicks + ci.ad\_clicks), 0), 2) AS estimated\_roi  
FROM campaign\_interactions ci  
JOIN purchase\_behavior pb ON ci.customer\_id = pb.customer\_id  
ORDER BY estimated\_roi DESC;

### 2. Data Quality Report - data\_quality\_report.docx

**Objective:** Make sure the data is clean, correct, and ready to be used for analysis.

**Steps Taken:**

* **Checked for Missing Values:**
  + All required columns were reviewed.
  + No missing values found in important columns like CustomerID, TotalSpent, and NumberOfOrders.
* **Found and Marked Outliers:**
  + Used IQR method to find unusual values in TotalSpent and NumberOfOrders.
  + Created two new flags: TotalSpent\_Outlier\_Flag and OrderCount\_Outlier\_Flag.
* **Made Sure Logic is Correct:**
  + Email clicks should never be more than email opens — this was validated.
  + Ad clicks should never be more than ad impressions — also validated.
* **Fixed Date Format Issues:**
  + The Registration Date had incorrect formats (like 15:20.8).
  + Converted all to proper date-time format: YYYY-MM-DD HH:MM:SS.
* **Rounded Numbers:**
  + Prices like total spent and average order value were rounded to 2 decimal places.
  + Click rates were rounded to 4 decimal places.
* **Churn Column Fix:**
  + Ensured the churned column only had 0 (not churned) or 1 (churned).

### 3. Initial Insights Document - initial\_insights.docx

**Goal:** Use the data to understand customer behavior and marketing performance.

### 🔍 Q1: Customers by Channel

* **Google Ads (3,044 customers)** brought in the most users.
* **Organic Search and Facebook Ads** were also strong.
* **Email and Direct** methods had fewer customers.

**Simple Takeaway:** Most customers came from online ads. Email and direct traffic need attention.

### 🔍 Q2: Average Spend and Orders by Customer Segment

| Segment | Avg Spend | Avg Orders |
| --- | --- | --- |
| High Value | ₹715.31 | 3.9 |
| Medium Value | ₹273.48 | 2.6 |
| Low Value | ₹74.01 | 2.2 |

**Simple Takeaway:**

* High Value customers spend much more and order more often.
* Focus on keeping High Value customers happy.

### 🔍 Q3: Customers Likely to Churn

* Found 130+ customers who haven’t purchased in over 90 days but are **not marked as churned**.

**Simple Takeaway:**

* These customers are at risk.
* Send offers or reminders to bring them back.

### 🔍 Q4: Email Clicks vs. Spending

* Some customers spent ₹500+ even without clicking emails.
* Some who clicked emails spent less.

**Simple Takeaway:**

* Clicking emails doesn’t always mean someone will spend money.
* Other things like loyalty or product need also matter.

### 🔍 Q5: Spending by Gender

| Gender | Avg Spending |
| --- | --- |
| Male | ₹198.59 |
| Female | ₹202.70 |
| Other | ₹219.82 |

**Simple Takeaway:**

* Spending is nearly equal across genders.
* Customers in “Other” gender category spend slightly more.

### 🔍 Q6: Campaign ROI (Return per Click)

* Some customers gave over ₹1,000 in revenue per click.
* Top customers: CUST\_04402, CUST\_00438, CUST\_02577, etc.

**Simple Takeaway:**

* Some users bring high returns with few clicks.
* Focus more effort on high-ROI customers.

### ✅ Summary:

* Ads are the top source of customers.
* High Value and High ROI customers deserve focus.
* Some customers may churn — need to re-engage them.
* Email clicks don’t always mean higher spending.

**Next Steps (Optional):**

* Create visuals in Power BI
* Build a churn prediction model in Python

✅ Milestone 1 is now complete. Ready for submission or to showcase on GitHub/LinkedIn!