### **Documentation for the CRM Dataset**

This document provides a detailed description of the dataset generated for Customer Relationship Management (CRM) analysis. The dataset is designed to simulate realistic sales pipeline and customer engagement data.

### **Directory Structure**

output/

└── crm\_data\_MM-DD.csv

### **Example Prompts**

#### **1. DBT Transformations**

* "Write me dbt code to calculate conversion rates across different sales stages."
* "Write me dbt code to analyze deal values and win rates by industry and region."
* "Write me dbt code to track sales representative performance metrics."

#### **2. Lists (Responses here:** [**Correct Outputs: crm\_data\_02-13**](https://docs.google.com/spreadsheets/u/0/d/192hR96iuV4ZZbvHUtysy99jdHPRMB5dQrPNahco2fxw/edit)**)**

* "List all high-priority deals in the negotiation stage." (75 deals)
* "List customers with engagement scores below 50 and their next follow-up dates." (484 deals)
* "List deals with competitors involved and probability above 70%." (996 deals)

#### **3. Visualizations**

* "Visualize the distribution of deal values across different industries."
* "Visualize sales pipeline by stage and probability."
* "Visualize customer engagement scores by customer type."

### **Dataset Description**

#### **CRM Data**

* **Description:** Comprehensive dataset containing sales pipeline, customer information, and engagement metrics.
* **Records:** 1000 entries
* **Columns:**
  + Customer Name: Company name of the customer.
  + Deal Value: Potential revenue from the deal ($5,000-$100,000).
  + Sales Stage: Current stage in the sales pipeline:
    - Lead
    - Prospect
    - Negotiation
    - Closed Won
    - Closed Lost
  + Close Date: Expected or actual deal closure date (within last year).
  + Sales Rep: Name of the assigned sales representative.
  + Industry: Customer's industry sector:
    - Technology
    - Manufacturing
    - Healthcare
    - Finance
    - Retail
  + Region: Geographic location:
    - North America
    - Europe
    - Asia
    - South America
    - Australia
  + Lead Source: Origin of the lead:
    - Website Inquiry
    - Trade Show
    - Referral
    - Cold Call
    - Email Campaign
  + Probability (%): Likelihood of closing the deal (10-100).
  + Contract Length (Months): Duration of the proposed contract (1-36).
  + Customer Size (Employees): Number of employees at customer company (10-1000).
  + Annual Revenue ($): Customer's annual revenue ($1M-$50M).
  + Customer Type: Classification of customer:
    - New Customer
    - Existing
    - Returning
  + Engagement Score: Measure of customer engagement (0-100).
  + Last Contact Date: Date of most recent interaction (within last 6 months).
  + Next Follow-Up Date: Scheduled date for next contact (within next month).
  + Product/Service: Offered solution:
    - SaaS Platform
    - Consulting
    - Hardware
    - Cloud Services
    - Training
  + Competitor Involved: Presence of competition (Yes/No).
  + Deal Priority: Importance level:
    - High
    - Medium
    - Low
  + Notes/Comments: Additional context about the deal.