2Market Data Analysis Project

Project Overview

This PostgreSQL project processes and analyzes marketing and ad conversion data for 2Market. The analysis helps the company understand:

1. The demographics of their customers
2. Which advertising channels are most effective
3. Which products sell best and how sales vary based on demographics

Database Setup Process

```
1. Create database on pgAdmin
2. Run tableMaking.sql to create tables
3. Run tableDataCleaning.sql to create a single joined table
4. Run customerDemographics.sql to analyze customer demographics
5. Run adChannelAnalysis.sql to evaluate advertising channel effectiveness
6. Run productDemographics.sql to analyze product preferences by demographic
```

Project Structure

```
├── sql/
│   ├── tableMaking.sql         # Creates initial database tables
│   ├── tableDataCleaning.sql   # Joins and cleans the data
│   ├── customerDemographics.sql # Customer demographic analysis
│   ├── adChannelAnalysis.sql   # Ad channel effectiveness analysis
│   └── productDemographics.sql # Product preferences by demographic
├── data/
│   ├── marketing_data.csv      # Customer demographic and purchase data
│   └── ad_data.csv             # Marketing channel conversion data
└── README.md                   # This file
```

Database Schema and ERD Diagrams

Initial Data Tables (tableMaking.sql)

The initial schema creates two tables to match the structure of the CSV files:

```
┌──────────────────────┐         ┌──────────────────────┐
│   marketing_data     │         │      ad_data         │
├──────────────────────┤         ├──────────────────────┤
│ ID (PK)              │         │ ID (PK)              │
│ Year_Birth           │         │ Bulkmail_ad          │
```

```
| Education        |    | Twitter_ad    |    |
| Marital_Status   |    | Instagram_ad  |    |
| Income           |    | Facebook_ad   |    |
| Kidhome          |    | Brochure_ad   |    |
| Teenhome         |    |               |    |
| Dt_Customer      |    |               |    |
| Recency          |    |               |    |
| AmtLiq           |    |     2 / 6     |    |
| AmtVege          |    |               |    |
| AmtNonVeg        |    |               |    |
| AmtPes           |    |               |    |
| AmtChocolates    |    |               |    |
| AmtComm          |    |               |    |
| NumDeals         |    |               |    |
| NumWebBuy        |    |               |    |
| NumWalkinPur     |    |               |    |
| NumVisits        |    |               |    |
| Response         |    |               |    |
| Complain         |    |               |    |
| Country          |    |               |    |
| Count_success    |    |               |    |
```

Combined Data (tableDataCleaning.sql)

The script creates a single table joining the data and adding derived columns:

```
|    customer_data_combined    |
|------------------------------|
| ID (PK)           |
| Year_Birth        |
| Education         |
| Marital_Status    |
| Income            |
| Kidhome           |
| Teenhome          |
| Dt_Customer       |
| Recency           |
| AmtLiq            |
| AmtVege           |
| AmtNonVeg         |
| AmtPes            |
| AmtChocolates     |
| AmtComm           |
| NumDeals          |
| NumWebBuy         |
| NumWalkinPur      |
| NumVisits         |
| Response          |
```

```
| Complain                     |
| Country                      |
| Count_success                |
| Customer_Date (derived)      |
| Income_Numeric (derived)     |
| Bulkmail_ad                  |
| Twitter_ad                   |
| Instagram_ad                 |
| Facebook_ad                  |
| Brochure_ad                  |
```
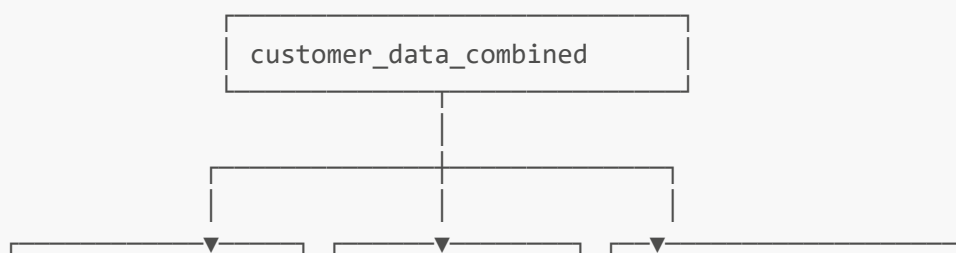
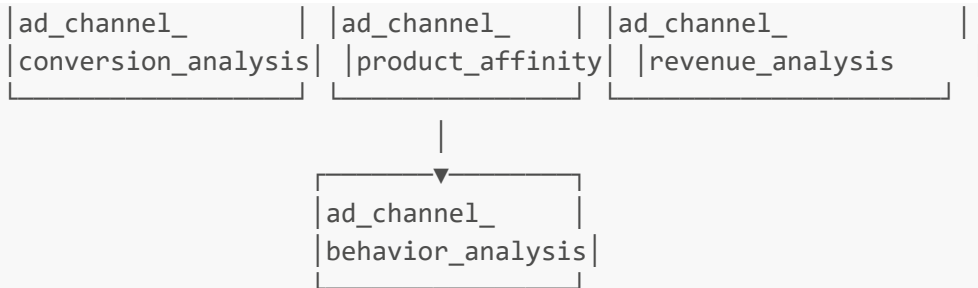## Customer Demographics Analysis (customerDemographics.sql)

Creates a view-based analysis of customer demographics:

```
    ┌──────────────────────────┐
    │   demogs_by_country      │
    ├──────────────────────────┤
    │ Metric                   │
    │ AUS                      │
    │ CA                       │
    │ GER                      │
    │ IND                      │
    │ ME                       │
    │ SA                       │
    │ SP                       │
    │ US                       │
    └──────────────────────────┘
          ▲
          │
          │  derives from
          │
    ┌──────────────────────────┐
    │ customer_data_combined   │
    └──────────────────────────┘
```
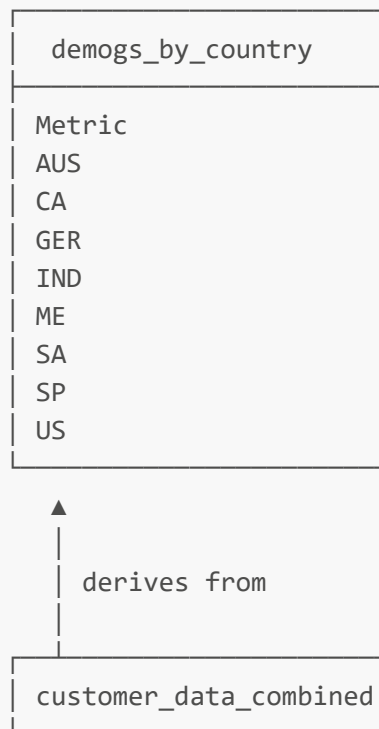
Ad Channel Analysis (adChannelAnalysis.sql)

Creates multiple analysis views for marketing channel effectiveness:

```
              ┌──────────────────────────────┐
              │   customer_data_combined     │
              └──────────────────────────────┘
                            │
            ┌───────────────┼───────────────┐
            │               │               │
    ┌───────▼────┐   ┌──────▼─────┐   ┌──────▼──────
    │            │   │            │   │
```

```
        |ad_channel_       | |ad_channel_      | |ad_channel_           |
        |conversion_analysis| |product_affinity| |revenue_analysis     |
        └──────────────────┘ └────────┬───────┘ └──────────────────────┘
                                       │
                             ┌─────────▼────────┐
                             |ad_channel_       |
                             |behavior_analysis |
                             └──────────────────┘
```

Product Demographics Analysis (productDemographics.sql)

Creates multiple views for product preference analysis across demographic segments:

```
                          ┌──────────────────────┐
                          | customer_data_combined|
                          └───────────┬───────────┘
                                      │
        ┌──────────┬─────────┬────────┼────────┬──────────┐
        ▼          ▼         ▼        ▼        ▼
   ┌──────────┐ ┌──────────┐ ┌──────────┐ ┌──────────┐ ┌──────────┐
   |product_  | |product_  | |product_  | |product_  | |product_  |
   |analysis_by_| |analysis_by_| |analysis_ | |analysis_by_| |analysis_by_|
   |country   | |age_group | |by_family_| |income    | |education |
   └──────────┘ └──────────┘ |size      | └──────────┘ └──────────┘
                             └────┬─────┘
                                  │
                             ┌────▼─────┐
                             |product_  |
                             |analysis_by_|
                             |tenure    |
                             └──────────┘
```
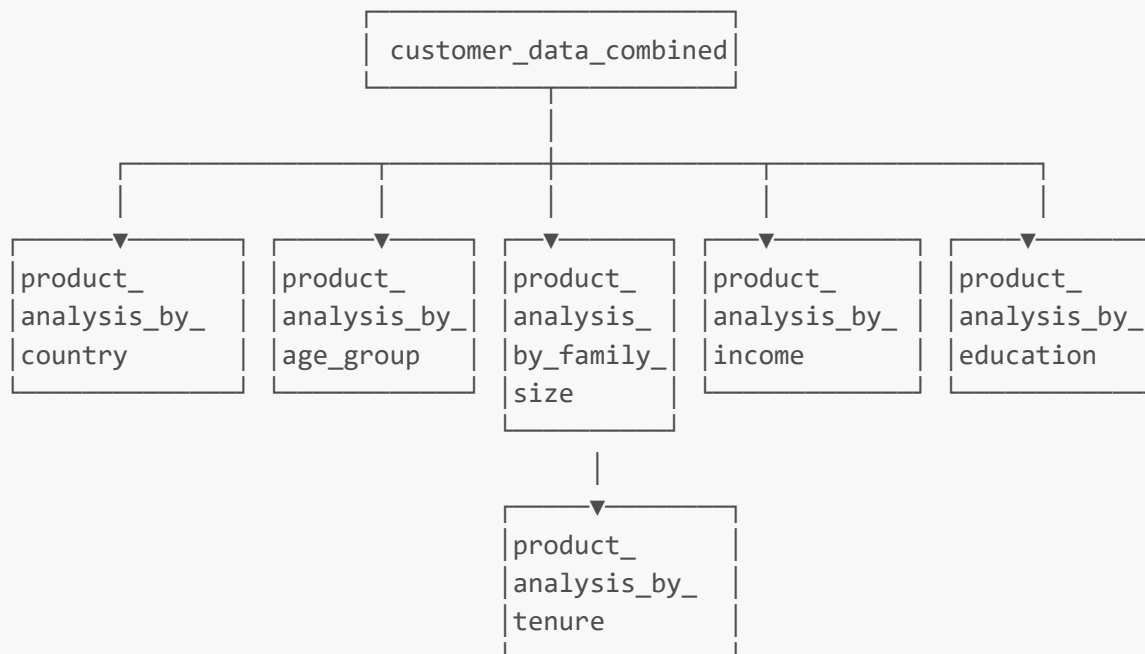
Using the Analysis Views

Customer Demographics View

```sql
-- View all metrics for all countries
SELECT * FROM demogs_by_country;

-- View specific metrics for all countries
SELECT * FROM demogs_by_country
WHERE Metric IN ('Avg_Age', 'Avg_Income', 'Response_Percentage');

-- Compare specific countries
SELECT * FROM demogs_by_country
WHERE Metric = 'Top_Three_Channels'
  AND (US IS NOT NULL OR SP IS NOT NULL);
```

Ad Channel Analysis Views

```sql
-- View conversion rates across all channels
SELECT * FROM ad_channel_conversion_analysis;

-- Compare product preferences by channel
SELECT Channel, Top_Three_Products FROM ad_channel_product_affinity;

-- Analyze revenue contribution by channel
SELECT Channel, Channel_Total_Revenue, Pct_of_Total_Revenue
FROM ad_channel_revenue_analysis
ORDER BY Pct_of_Total_Revenue DESC;

-- Examine engagement metrics by channel
SELECT Channel, Avg_Purchase_Frequency, Avg_Response, Avg_NumVisits
FROM ad_channel_behavior_analysis
WHERE Channel NOT IN ('All Customers', 'All Ad Channels', 'No Channel')
ORDER BY Avg_Response DESC;
```

Product Demographics Analysis Views

```sql
-- View product preferences by country
SELECT * FROM product_analysis_by_country;

-- View product preferences by age group
SELECT * FROM product_analysis_by_age_group;

-- View product preferences by family size
SELECT * FROM product_analysis_by_family_size;

-- View product preferences by income bracket
SELECT * FROM product_analysis_by_income;

-- View product preferences by education level
SELECT * FROM product_analysis_by_education;

-- View product preferences by customer tenure
SELECT * FROM product_analysis_by_tenure;
```

Key Features

1. **Data Transformation**

   - Date standardization
   - Income normalization
   - Derived metrics calculation

---

2. **Demographic Analysis**

   - Age distribution by country
   - Income comparison
   - Family structure analysis
   - Spending patterns across product categories

3. **Marketing Channel Analysis**

   - Channel effectiveness by country
   - Top performing channels ranking
   - Response and complaint analysis
   - Product affinities by channel
   - Revenue attribution by channel
   - Customer behavior patterns by channel

4. **Product Preference Analysis**

   - Product performance by country
   - Age-based product preferences
   - Family size impact on purchases
   - Income-based purchasing patterns
   - Education level correlations
   - Customer loyalty effects