Marketing Data Analysis Project

This repository contains PostgreSQL scripts for analyzing marketing campaign effectiveness and customer demographics across different regions and channels.

Project Structure

```
├── sql/
│   ├── 01_create_tables.sql       # Database schema creation
│   ├── 02_combine_data.sql        # Joining tables and data transformations
│   ├── 03_demographics_view.sql   # Demographics analysis view
│   └── 04_ad_channel_analysis.sql # Ad channel effectiveness analysis
├── data/
│   ├── marketing_data.csv         # Customer demographic and purchase data
│   └── ad_data.csv                # Marketing channel conversion data
└── README.md                      # This file
```

Database Schema

The project uses the following database structure:

```
marketing_data (1) ---- (1) customer_data_combined (1) ---- (1)
demogs_by_country (view)
       |                              |
       |                              |---- (1) ad_channel_conversion_analysis
(view)
       |                              |---- (1) ad_channel_product_affinity
(view)
ad_data (1) ----------- (1) /         |---- (1) ad_channel_revenue_analysis
(view)
                                      └---- (1) ad_channel_behavior_analysis
(view)
```

Tables

1. **marketing_data**: Contains customer demographic information and purchase behavior
    - Primary Key: ID
    - Key fields: Year_Birth, Education, Marital_Status, Income, Country, Recency, Response, Complain
2. **ad_data**: Contains marketing channel conversion data
    - Primary Key: ID
    - Conversion fields: Bulkmail_ad, Twitter_ad, Instagram_ad, Facebook_ad, Brochure_ad

3. **customer_data_combined**: Integrated view of all customer data
    - Adds derived fields: Customer_Date, Income_Numeric
    - Contains all fields from both source tables
4. **demogs_by_country**: Analysis view that presents metrics by country
    - Rows: Various demographic metrics (age, income, spending, etc.)
    - Columns: Country codes (AUS, CA, GER, IND, ME, SA, SP, US)
5. **ad_channel_conversion_analysis**: Conversion rates and effectiveness by ad channel
    - Rows: Conversion metrics (global rates, channel share)
    - Columns: Each advertising channel plus aggregated metrics
6. **ad_channel_product_affinity**: Product preferences by marketing channel
    - Shows which products are most associated with each channel
    - Includes top three products by spending for each channel
7. **ad_channel_revenue_analysis**: Financial impact of each marketing channel
    - Total and average revenue metrics by channel
    - Percentage of total revenue attributable to each channel
8. **ad_channel_behavior_analysis**: Customer behavior patterns by channel
    - Engagement metrics (purchase frequency, website visits, etc.)
    - Relative performance compared to global averages

Setup Instructions

Prerequisites

- PostgreSQL 12 or higher
- psql command-line tool or PgAdmin 4

Database Creation

1. Create a new PostgreSQL database:

```
createdb marketing_analysis
```

2. Run the table creation script:

```
psql -d marketing_analysis -f sql/01_create_tables.sql
```

Data Import

Using psql:

```
\copy marketing_data FROM 'data/marketing_data.csv' WITH CSV HEADER DELIMITER
',';
\copy ad_data FROM 'data/ad_data.csv' WITH CSV HEADER DELIMITER ',';
```

Using PgAdmin:

1. Right-click on the table name in the object browser
2. Select "Import/Export Data"
3. Configure import settings:
    - Set filename to the appropriate CSV
    - Select "," as delimiter
    - Check "Header" option
    - Set Format to "CSV"
4. Click "OK" to import

Creating Analysis Views

Run the table combination and transformation script:

```
psql -d marketing_analysis -f sql/02_combine_data.sql
```

Run the demographics analysis view script:

```
psql -d marketing_analysis -f sql/03_demographics_view.sql
```

Run the ad channel analysis script:

```
psql -d marketing_analysis -f sql/04_ad_channel_analysis.sql
```

Using the Demographics View

The `demogs_by_country` view provides key metrics by country. Sample queries:

```
-- 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);
```

Using the Ad Channel Analysis Views

The ad channel analysis provides insights into marketing effectiveness through multiple views:

```
-- 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;
```

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. **Technical Features**
    - Null-safe calculations
    - Type-consistent operations
    - Formatted output with rounded values
    - String aggregation for complex data presentation
    - Relative performance metrics with control groups

Troubleshooting

If you encounter data type issues in the views, ensure all numeric values are being cast to text in the view definitions. The views use UNION ALL operations which require consistent data types across queries.

Next Steps

Planned analytical extensions include:

- Temporal analysis of purchasing behavior
- Customer segmentation by behavior patterns
- Correlation analysis between demographics and response rates
- ROI calculations for advertising channels
- Predictive modeling for customer lifetime value