📊 Retail Sales Analysis Dashboard

A comprehensive Streamlit-based web application for analyzing retail sales data stored in Google BigQuery. This dashboard provides interactive visualizations, SQL query execution, business intelligence insights, and data exploration capabilities.

🎯 Overview

This application serves as a complete business intelligence solution for retail sales data analysis, offering both technical and business users the ability to explore, analyze, and visualize sales data through an intuitive web interface.

✨ Features

📊 Data Analysis

- **Dataset Overview**: Complete schema analysis and metadata display
- **Data Quality Analysis**: Comprehensive null value detection and completeness metrics
- **Sample Data Display**: Interactive data preview with CSV download capabilities
- **Real-time Data Metrics**: Live connection status and data freshness indicators

🔍 SQL Query Interface

- **Pre-built Queries**: 6 ready-to-use analysis queries covering common business scenarios
- **Custom Query Execution**: Full SQL editor for advanced users
- **Query Results Export**: Download results as CSV files
- **Query Templates**: Categorized queries for different analysis types

📈 Interactive Visualizations

- **Revenue by Category**: Bar charts showing sales performance by product category
- **Customer Demographics**: Gender-based revenue analysis with pie charts
- **Monthly Trends**: Time-series analysis of sales patterns
- **Customer Spending**: Histogram distribution of customer spending
- **Age Group Analysis**: Revenue analysis by age demographics
- **Product Performance**: Top-performing product categories
- **Interactive Sales Over Time**: Advanced Altair-based charts with filtering capabilities

💡 Business Intelligence

- **Key Performance Indicators (KPIs)**: Real-time business metrics dashboard
- **Revenue Analysis**: Category-wise performance insights
- **Customer Segmentation**: High/Medium/Low value customer analysis
- **Business Recommendations**: Data-driven strategic suggestions
- **Trend Analysis**: Temporal patterns and seasonality insights

🛠️ Technical Architecture

Technology Stack

- **Frontend**: Streamlit (Python web framework)
- **Data Warehouse**: Google BigQuery
- **Visualizations**: Plotly Express, Plotly Graph Objects, Altair
- **Data Processing**: Pandas, NumPy
- **Authentication**: Google Cloud Service Account
- **Deployment**: Streamlit Cloud

Data Source

- **Dataset**: `moonlit-autumn-468306-p6.assignment_one_1.retail_sales`
- **Source**: Kaggle Retail Sales Dataset
- **Platform**: Google BigQuery
- **Data Types**: DATE, STRING, INTEGER, FLOAT

📊 Data Schema

Table: `retail_sales`

Column	Type	Description
Date	DATE	Transaction date
Product Category	STRING	Product category name
Customer ID	STRING	Unique customer identifier
Quantity	INTEGER	Number of items purchased
Total Amount	FLOAT	Total transaction value
Gender	STRING	Customer gender
Age	INTEGER	Customer age

🚀 Getting Started

Prerequisites

- Python 3.7 or higher
- Google Cloud Platform account
- BigQuery dataset access
- Streamlit Cloud account (for deployment)

Local Development Setup

1. **Clone the repository**

```
git clone <repository-url>
cd retail-sales-dashboard
```

2. **Install dependencies**

```
pip install streamlit pandas plotly altair google-cloud-bigquery
```

3. **Configure BigQuery credentials**
   - Create a Google Cloud Service Account
   - Download the JSON key file
   - Set up Streamlit secrets (see Configuration section)

4. **Run the application**

```
streamlit run streamlit_app.py
```

Streamlit Cloud Deployment

1. **Push to GitHub**

```
git add .
git commit -m "Initial commit"
git push origin main
```

2. **Deploy on Streamlit Cloud**
   - Go to share.streamlit.io
   - Connect your GitHub repository
   - Configure secrets (see Configuration section)
   - Deploy the application

⚙️ Configuration

BigQuery Credentials Setup

1. **Create Service Account**
   - Go to Google Cloud Console
   - Navigate to IAM & Admin > Service Accounts
   - Create a new service account
   - Grant BigQuery Data Viewer role

2. **Download Credentials**
   - Generate and download JSON key file
   - Keep this file secure

3. **Configure Streamlit Secrets**
    - In Streamlit Cloud, go to app settings
    - Navigate to Secrets section
    - Add the following configuration:

```
[gcp_service_account]
type = "service_account"
project_id = "your-project-id"
private_key_id = "your-private-key-id"
private_key = "-----BEGIN PRIVATE KEY-----\n...\n-----END PRIVATE KEY-----
\n"
client_email = "your-service-account@your-project.iam.gserviceaccount.com"
client_id = "your-client-id"
auth_uri = "https://accounts.google.com/o/oauth2/auth"
token_uri = "https://oauth2.googleapis.com/token"
auth_provider_x509_cert_url = "https://www.googleapis.com/oauth2/v1/certs"
client_x509_cert_url =
"https://www.googleapis.com/robot/v1/metadata/x509/your-service-
account%40your-project.iam.gserviceaccount.com"
```

Usage Guide

1. Home Page

- **Connection Status**: Verify BigQuery connection
- **Dataset Overview**: View metadata and data statistics
- **Analysis Objectives**: Understand the dashboard's capabilities

2. Dataset Analysis

- **Schema Exploration**: Examine table structure and data types
- **Data Quality**: Check for null values and completeness
- **Sample Data**: Preview actual data records

3. SQL Queries

- **Pre-built Queries**: Select from common analysis templates
- **Custom Queries**: Write and execute your own SQL
- **Results Export**: Download query results as CSV

4. Visualizations

- **Chart Selection**: Choose from 7 different visualization types
- **Interactive Features**: Filter data by date, category, and customer
- **Export Options**: Download charts and underlying data

5. Business Insights

- **KPI Dashboard**: View key performance indicators
- **Revenue Analysis**: Category-wise performance metrics
- **Customer Segmentation**: Value-based customer categorization
- **Recommendations**: Data-driven business suggestions

🔧 Code Structure

Main Components