

# Google AdWords Performance Analysis

**An End-to-End Marketing Analytics Pipeline for Keyword Performance, Cost Optimization, and Traffic Insights**

---

## 1. Abstract

This project analyses simulated Google AdWords data to evaluate keyword performance, traffic trends, and cost dynamics. A complete analytics pipeline was developed using Excel, Python, MySQL, and Power BI. Raw marketing data was cleaned, transformed, normalized into a relational database, and visualized through interactive dashboards. The project enables marketing teams to identify high-performing keywords, optimize advertising spend, and make data-driven campaign decisions.

---

## 2. Introduction

- Importance of digital advertising analytics
- Why Google AdWords data matters
- Role of analytics in marketing decision-making

### Example:

Digital advertising generates large volumes of performance data, but without structured analysis, marketing teams struggle to optimize budgets and keyword strategies. This project simulates a real-world advertising analytics workflow to demonstrate how raw Google AdWords data can be transformed into actionable insights using modern analytics tools.

---

## 3. Business Problem

Marketing teams face challenges such as:

- Identifying keywords that deliver high traffic at low cost
  - Detecting overspending on high-CPC, low-return keywords
  - Understanding how keyword difficulty affects performance
  - Monitoring seasonal and monthly traffic trends
  - Deciding which keywords to scale, optimize, or pause
- 

## 4. Project Objectives.

- Analyse keyword-level traffic, CPC, and competition
  - Identify high-impact and low-efficiency keywords
  - Track search demand and cost trends over time
  - Build a normalized relational database
  - Create interactive dashboards for decision-making
- 

## 5. Tools & Technologies Used

What this section is:

Tool	Purpose
Excel	Initial data cleaning, aggregation, lookups
Python (pandas)	Data preprocessing and ID generation
MySQL	Relational database modeling
Power BI	Data modeling, DAX, dashboards
GitHub	Version control and documentation

---

## 6. Dataset Description

- Source: Simulated Google AdWords export
  - File: `Raw_data.xlsx`
  - Key variables:
    - Keyword
    - CPC
    - Traffic
    - Search volume
    - Competition
    - SERP results
    - Keyword positions
- 

## 7. Data Cleaning & Preparation

### 7.1 Excel Processing

- Aggregated search volume (`SUMIF`)
- Calculated average difficulty (`AVERAGEIF`)
- Created difficulty categories (`IF`)
- Built lookup tables (`VLOOKUP`)

## 7.2 Python Processing

- Loaded Excel data using pandas
- Generated unique keyword ID values
- Ensured referential integrity across tables

```
df['keyword_ID'] = df['keyword'].astype('category').cat.codes
```

---

## 8. Exploratory Data Analysis (EDA)

### What to include:

- Distribution of traffic across keywords
  - CPC variation analysis
  - Relationship between keyword difficulty and traffic
  - Identification of outliers (high CPC, low traffic keywords)
- 

## 9. Data Modeling & SQL Schema

### Content:

- Star schema design
- Fact table: website\_traffic\_data
- Dimension tables:
  - keyword
  - keyword\_difficulty
  - search\_volume

Explain:

The star schema ensures optimized query performance, accurate filtering, and scalable reporting.

---

## 10. SQL Implementation

- Database creation
- Table creation
- Foreign key relationships

(Core SQL code)

```
CREATE DATABASE IF NOT EXISTS traffic_project;
```

```
USE traffic_project;
```

```
CREATE TABLE website_traffic_data (  
    title VARCHAR(250),  
    keyword VARCHAR(250),  
    keyword_ID INT,  
    positions INT,  
    previous_positions INT,  
    last_seen DATE,  
    Search_Volume INT,  
    CPC DECIMAL(10,2),  
    Traffic INT,  
    Traffic_Percent DECIMAL(10,2),  
    Traffic_Cost INT,  
    Traffic_Cost_Percent DECIMAL(10,2),  
    Competition DECIMAL(10,2),  
    Number_of_Results INT,  
    Keyword_difficulty INT  
);
```

```
CREATE TABLE keyword (  
    keyword_ID INT PRIMARY KEY,  
    keyword VARCHAR(250)  
);
```

```
CREATE TABLE keyword_difficulty (  
    keyword_ID INT PRIMARY KEY,  
    avg_difficulty DECIMAL(10,2),  
    difficulty_level VARCHAR(50)
```

);

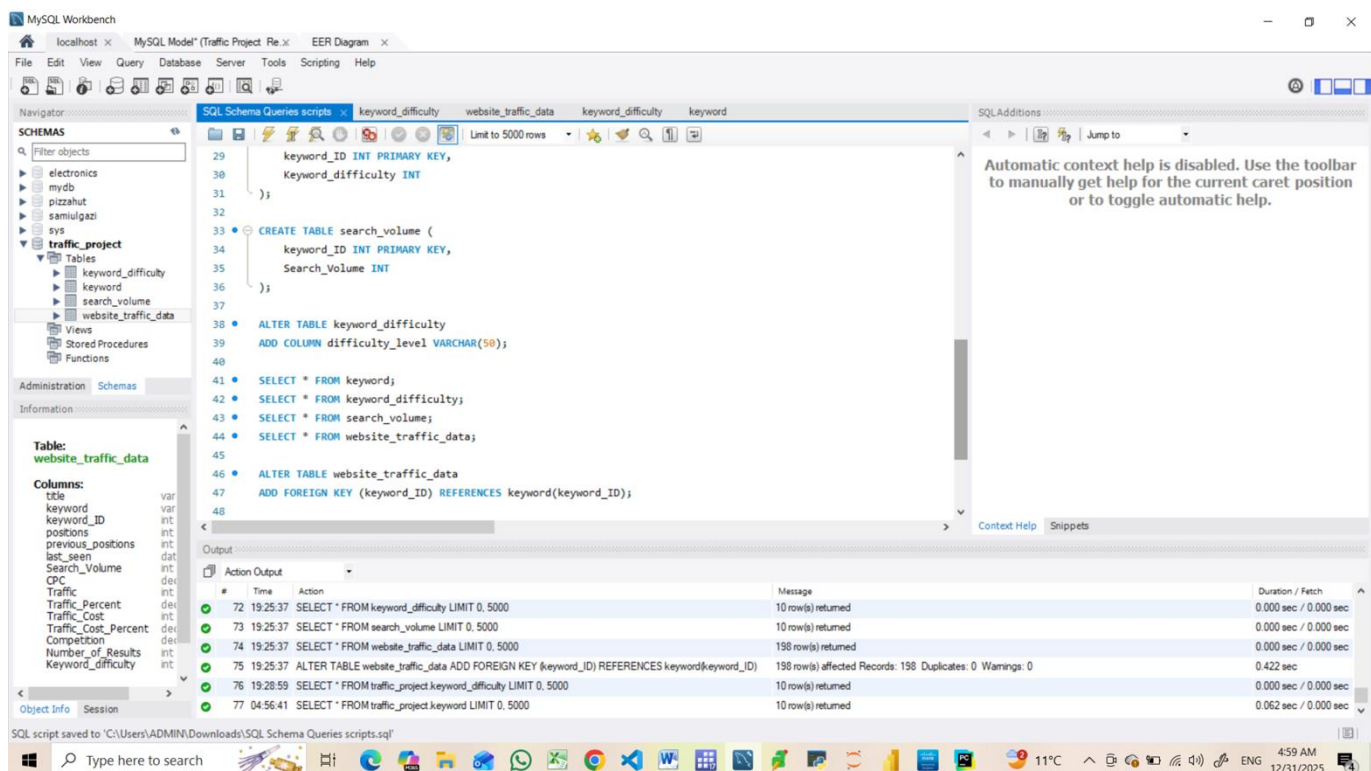
CREATE TABLE search\_volume (

keyword\_ID INT PRIMARY KEY,

Search\_Volume INT);

ALTER TABLE website\_traffic\_data

ADD FOREIGN KEY (keyword\_ID) REFERENCES keyword(keyword\_ID);



# Power BI Data Model & DAX

## 11.1 Data Model

- Star schema in Power BI
- Fact-dimension relationships

## 11.2 DAX Measures

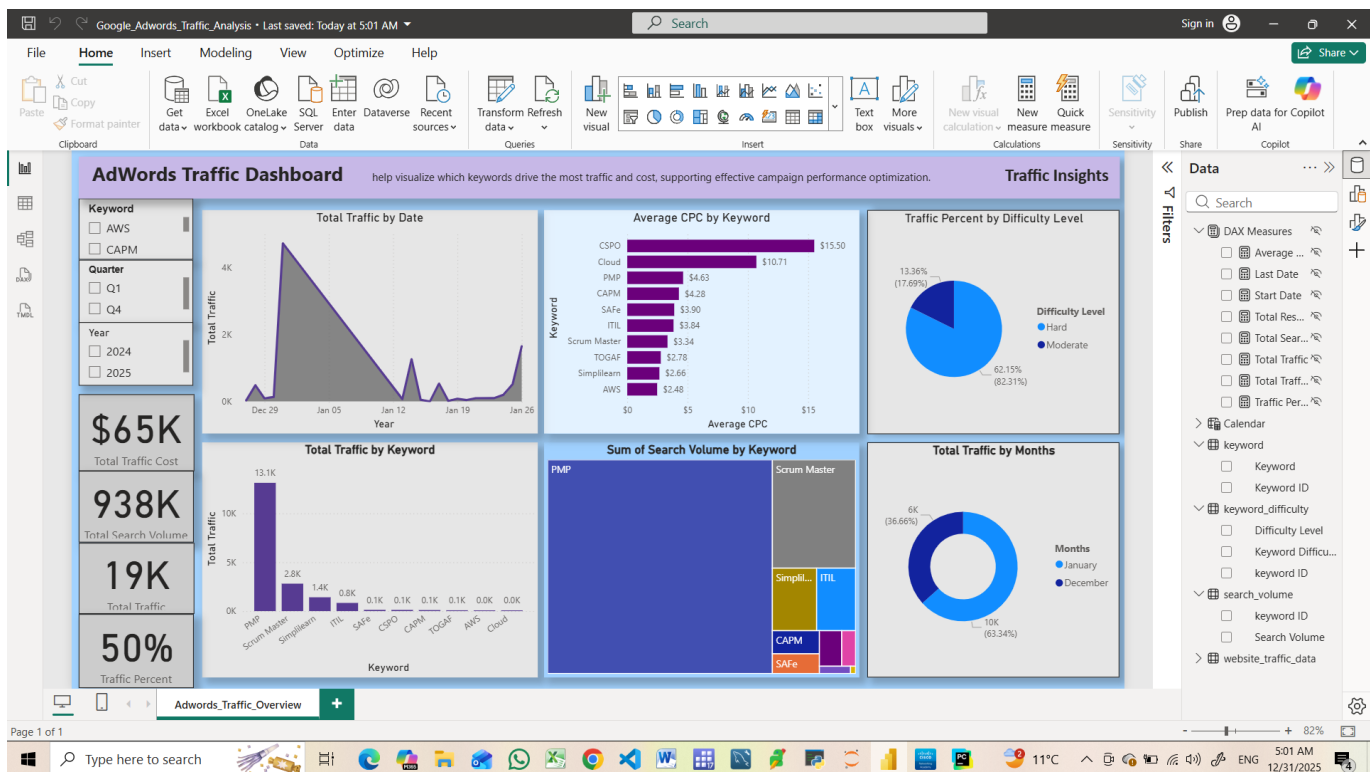
Explain key KPIs:

- Total Traffic
- Average CPC
- Total Cost
- Time intelligence (Month, Quarter, Year)

## 12. Dashboard Design & Insights

### KPI cards overview

- Traffic trend analysis
- Keyword performance ranking
- Difficulty-wise traffic distribution



Interactives slicers allow users to filter by time, keyword, and difficulty.

## 13. Key Findings

### Results.

- High-traffic keywords with low CPC identified
- Several high-CPC keywords showed low traffic efficiency

- Moderate-difficulty keywords contributed most traffic
  - Clear seasonal traffic patterns observed
- 

## 14. Future Improvements

- Google Ads API automation
  - Conversion & ROAS metrics
  - Forecasting models
  - Campaign-level analysis
  - Cloud deployment
- 

## 15. Conclusion

This project demonstrates a complete marketing analytics pipeline, integrating data cleaning, relational modelling, SQL analysis, and business intelligence visualization. The approach reflects real-world industry practices and highlights the value of structured analytics in digital marketing decision-making.

---

## 16. Author & Contact

### **Samiul Gazi**

M.Sc. Economics – University of Calcutta  
Kolkata, West Bengal

Email: [sgaziamumh@gmail.com](mailto:sgaziamumh@gmail.com)

LinkedIn: [linkedin.com/in/samiul-gazi](https://www.linkedin.com/in/samiul-gazi)

GitHub: [github.com/Samiul1947](https://github.com/Samiul1947)