

Strava Fitness Data Analytics Project

Project Type

Exploratory Data Analysis (EDA) & Business Intelligence Dashboard

Tools & Technologies Used

- **Python** (Pandas, Matplotlib, Seaborn)
 - **SQL**
 - **Power BI**
 - **Jupyter Notebook**
 - **Microsoft Excel / CSV Dataset**
-

Dataset Description

The dataset used for this project is **Strava / Smart Device Fitness Activity Data**, which contains daily activity records of users captured through wearable fitness devices.

It includes information such as:

- User ID
- Activity Date
- Total Steps
- Total Distance
- Calories Burned
- Active Minutes (Very Active, Fairly Active, Lightly Active)
- Sedentary Minutes

The dataset helps analyze user fitness behavior, activity intensity, and engagement patterns over time.

Project Objective

The objective of this project is to analyze smart fitness device data to:

- Understand user activity patterns over time
 - Identify trends in steps, calories burned, and active minutes
 - Analyze daily and weekly behavior variations
 - Provide data-driven insights that can help improve fitness engagement and wellness strategies
-

Methodology

1. Data Collection

The dataset was collected in CSV format and imported into:

- Python for Exploratory Data Analysis
 - SQL for structured querying and validation
 - Power BI for dashboard creation
-

2. Data Cleaning & Preprocessing

- Converted date columns into proper datetime format
 - Checked for missing and null values
 - Verified data consistency across steps, calories, and distance
 - Created additional columns such as:
 - **Day Name**
 - **Day Number (Weekday Index)**
-

3. Exploratory Data Analysis (Python)

Using Python, the following analyses were performed:

- Overview of dataset structure and data types
- Descriptive statistics for numerical columns
- Line charts for:
 - Daily steps over time
 - Calories burned over time
- Identification of activity fluctuations and patterns

Each visualization was followed by a short **insight explanation** describing user behavior and trends.

4. SQL Analysis

SQL was used to:

- Validate Python findings
- Perform aggregations such as:
 - Total steps
 - Average steps per day
 - Total calories burned
 - Total active minutes
- Group data by day and user to understand engagement patterns

The SQL queries were documented and included in the SQL report PDF.

5. Power BI Dashboard Development

A professional interactive dashboard was created in Power BI with:

Key KPI Cards

- Total Steps
- Average Steps per Day

- Total Calories Burned
- Average Calories per Day
- Total Active Minutes

Visualizations

- Line chart: Steps trend over time
- Bar chart: Calories burned by day
- Column chart: Active minutes distribution
- Slicers:
 - Date
 - Day Name
 - User ID

The dashboard enables dynamic filtering and quick insight extraction.

Key Insights

- User activity levels fluctuate significantly across different days
 - Certain days show higher engagement in steps and calories burned
 - Active minutes are unevenly distributed, with sedentary time dominating
 - Consistent activity correlates with higher calorie burn and distance covered
-

Business Recommendations

- Encourage users to maintain consistent daily activity through reminders
 - Promote engagement on low-activity days using personalized challenges
 - Use insights to design fitness programs targeting sedentary behavior reduction
-

Project Deliverables

- ✓ SQL Script File
 - ✓ SQL Analysis PDF
 - ✓ Python Jupyter Notebook
 - ✓ Python EDA PDF
 - ✓ Power BI Dashboard (.pbix)
 - ✓ Power BI Dashboard PDF
 - ✓ Final Project Report
-

Conclusion

This project successfully demonstrates the use of **Python, SQL, and Power BI** to analyze fitness data and extract meaningful insights. The combined approach of data analysis and visualization provides a clear understanding of user behavior and supports data-driven decision-making in the fitness and wellness domain.

Future Scope

- Integrate real-time fitness data using APIs
 - Add predictive modeling for activity forecasting
 - Include sleep and heart rate analysis for deeper insights
-

End of Report