

PYTHON VISUALIZATION REPORT

PROJECT TITLE: STRAVA FITNESS DATA ANALYSIS

OBJECTIVE:

To utilise Python modules to visualise activity, calorie, and intensity trends in order to extract insights from Fitbit user data.

1. Dataset Overview: We used the following datasets for this visualisation phase:

- minuteIntensitiesNarrow_merged.csv
- minuteCaloriesWide_merged.csv
- minuteMETsNarrow_merged.csv
- minuteStepsWide_merged.csv

Each dataset contains minute-level data collected from Fitbit users, including intensity levels, calories burned, metabolic equivalents (METs), and step counts.

2. Tools and Libraries Used

- Python 3.x
- Pandas (for data handling)
- Matplotlib (for basic plotting)
- Seaborn (for aesthetically pleasing plots)
- Plotly (for interactive charts)
- Jupyter Notebook / Google Colab (for code execution and annotation)

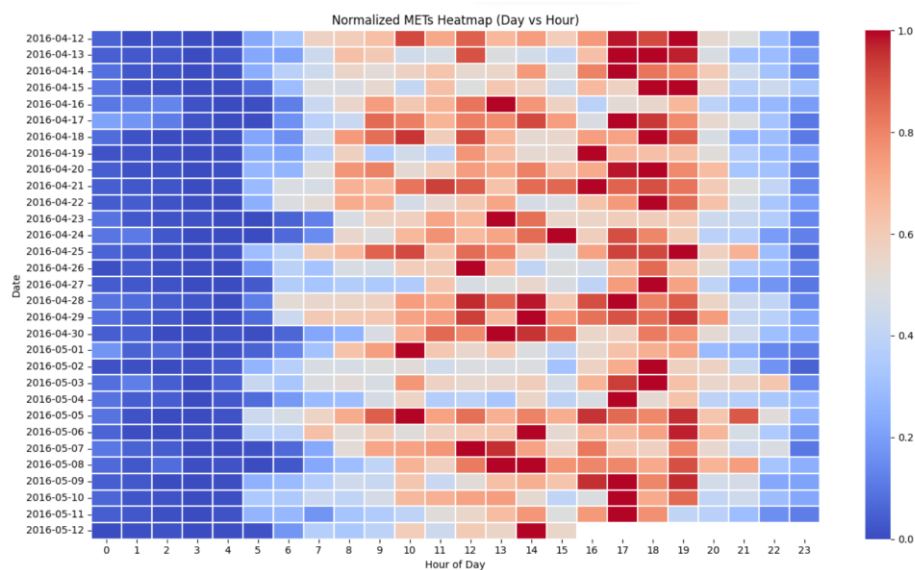
3. Data Cleaning and Preparation

- Converted ActivityMinute columns to datetime format for consistency
- Removed duplicates

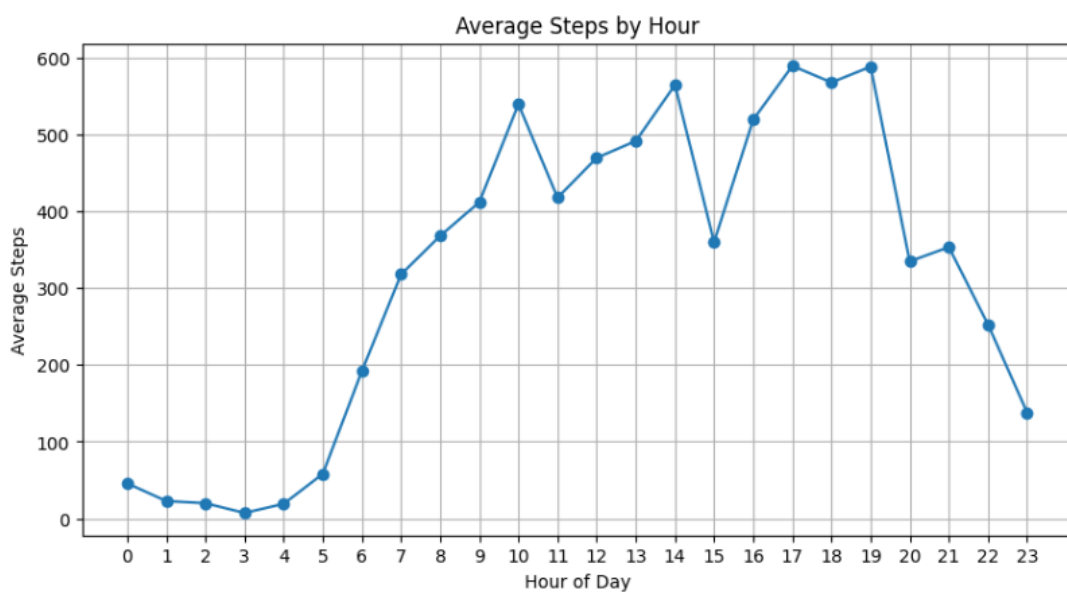
- Checked for missing/null values and handled appropriately
- Standardized date formats across datasets
- Merged datasets when necessary based on Id and ActivityMinute
-

4. Key Visualizations and Insights

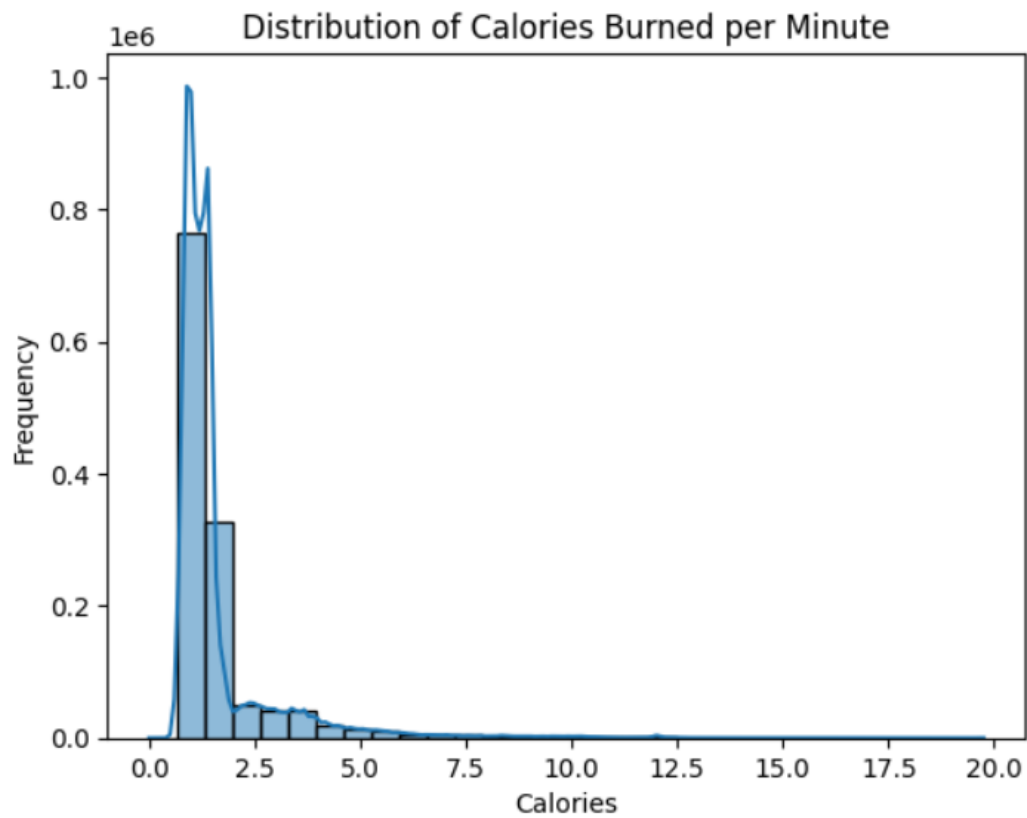
a. Heatmap - Normalized METs Heatmap



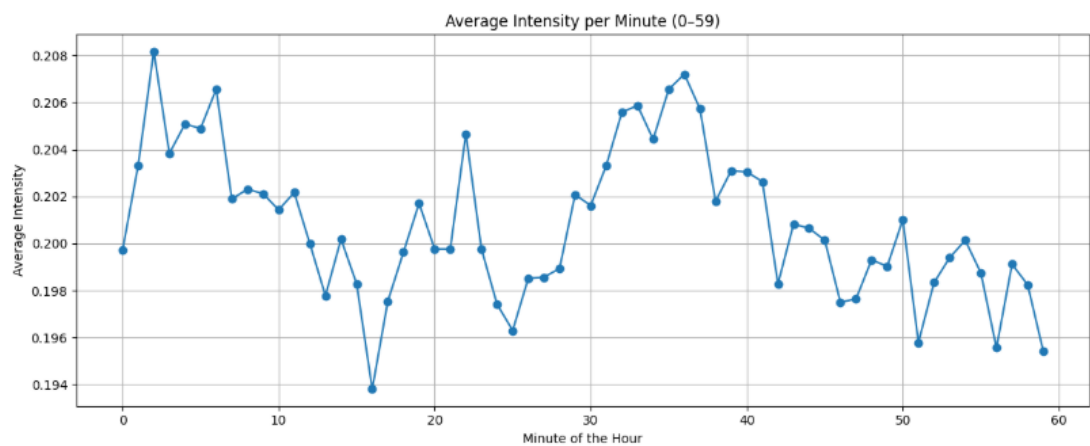
b. Line graph - Average Steps by Hour



c. bar and line chart - Distribution of Calories Burned per Minute



d. Line chart - Average Intensity per Minute



5. Challenges Faced:

- Handling large minute-level datasets efficiently
- Unifying date formats across different files
- Ensuring visual clarity in multi-line or overlaid plots

6. Conclusion:

Python visualisations helped uncover valuable activity patterns and behaviour trends at a granular level. These insights are useful for health researchers or fitness product analysts to make data-driven decisions or recommend personalised goals. Combined with Power BI dashboards and SQL-based cleaning, this analysis completes a comprehensive data storytelling process.