

Capstone Project - Fitbit Consumer Behavior Analysis

Report:

~ by Hariharan M

Content:

Respondents generated this dataset to a distributed survey via Amazon Mechanical Turk between 03.12.2016 and 05.12.2016. Thirty eligible Fitbit users consented to submit personal tracker data, including minute-level output for physical activity, heart rate, and sleep monitoring. Individual reports can be parsed by export session ID (column A) or timestamp (column B). Variation between output represents the use of different Fitbit trackers and individual tracking behaviors/preferences.

Business Task:

Analyze FitBit Fitness Tracker App data to gain insights into how consumers use the FitBit app and discover trends and insights for the marketing team.

Business Objectives:

- What are the trends identified?
- How could these trends apply to customers?
- How could these trends help influence marketing strategy?

1.What are the trends identified?

Physical Activity Patterns:

The dataset includes information about daily and hourly physical activity, including steps, calories burned, and intensity levels. By analyzing these patterns, you can identify individuals with consistent or irregular activity levels.

Sleep Patterns:

Sleep data is available, providing insights into total sleep records, total minutes asleep, and total time in bed. Patterns in sleep duration and quality can be identified, helping to understand individuals' sleep habits.

Weight and Body Metrics:

The dataset includes information on weight, BMI, and fat percentage. Trends in weight change, BMI variation, and fat percentage can be analyzed over time for each individual.

Heart Rate Analysis:

Heart rate data is available at a fine-grained level, capturing values at specific times. Patterns in heart rate during different activities or times of the day can be explored

2. How could these trends apply to customers?

Personalized Health Insights:

Customers can receive personalized insights into their physical activity, sleep, and weight trends. This information can help individuals understand their habits, set health goals, and make informed decisions about their lifestyle.

Wellness Recommendations:

Based on identified trends, customers can receive recommendations for improving their overall well-being. For example, if irregular sleep patterns are identified, suggestions for better sleep hygiene can be provided.

Targeted Fitness Plans:

Fitness plans can be tailored to individual preferences and patterns. Customers with specific activity preferences or challenges can receive targeted recommendations for exercise routines.

3. How could these trends help influence marketing strategy?

Segmented Marketing:

Identify customer segments based on activity levels, sleep patterns, and weight metrics. Tailor marketing campaigns to address the unique needs and interests of each segment.

Promotions for Health Products:

Launch promotions for health and fitness products based on identified trends. For instance, market sleep-related products if irregular sleep patterns are prevalent among the target audience.

Health and Wellness Programs:

Develop marketing strategies around comprehensive health and wellness programs. Highlight how products or services contribute to improved physical activity, better sleep, and overall well-being.

Customized Fitness Apps:

If irregular activity patterns are identified, marketing efforts can focus on promoting fitness apps or solutions that provide personalized workout plans.

Engagement and Rewards:

Create engagement strategies that reward customers for achieving specific health goals. Loyalty programs can be designed to encourage continued use of health and fitness-related products or services. In summary, understanding trends in physical activity, sleep, weight, and heart rate can provide valuable insights for creating personalized customer experiences and targeted marketing strategies in the health and wellness industry.

Deliverables:

1. Clear Summary of the Business Task:

Clearly articulate the primary business task or problem you are addressing. For example:

The business task is to analyze health and fitness data to derive actionable insights that can be used to improve customer experience, tailor marketing strategies, and enhance product offerings in the health and wellness industry.

2. Description of Data Sources:

Provide information about the data sources you utilized. This might include:

The dataset contains records of physical activity, sleep patterns, weight metrics, and heart rate collected from wearable devices and mobile apps. Data sources may include APIs, databases, or third-party platforms.

3. Documentation of Data Cleaning/Manipulation:

Detail the steps taken to clean and manipulate the data. This could involve:

Handling missing values: Explain how missing values were treated (removed, imputed, etc.).

Standardizing units or formats: Detail any efforts to ensure consistency in the dataset.

Removing outliers: Mention if outliers were identified and addressed.

4. Summary of the Analysis:

Provide a concise overview of the analysis conducted. For instance:

The analysis focused on identifying trends in physical activity, sleep, weight, and heart rate. Key metrics such as average steps per day, sleep duration, and heart rate patterns were examined.

5. Supporting Visualizations and Key Findings:

Include relevant visualizations (charts, graphs, etc.) and highlight key findings:

Visualizations might include time series plots of physical activity, histograms of sleep duration, or pie charts representing activity types. Key findings could include insights into common sleep patterns, peak activity times, or correlations between activity levels and heart rate.

6. High-level Content Recommendations:

Based on your analysis, provide high-level recommendations for the business:

Personalized Health Insights: Develop features or products that provide personalized health insights to customers. Targeted Marketing: Tailor marketing strategies based on identified customer segments. Wellness Programs: Consider launching wellness programs aligned with customer needs.

Exploratory Data Analysis:

View of cleaned and Manipulated of Minute activity dataset

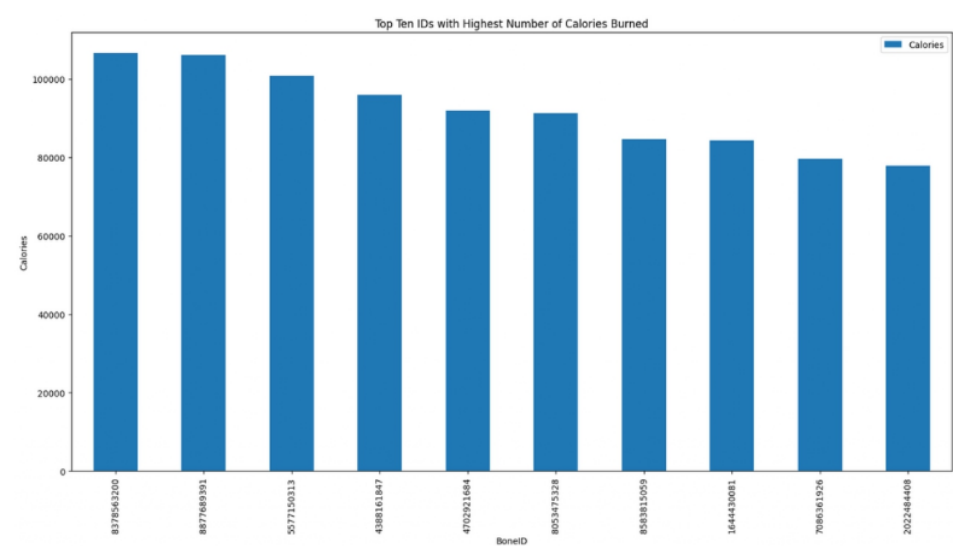
```
[ ] minutesactivity.head()
```

	Id	ActivityMinute	Intensity	Intensity00	Intensity01	Intensity02	Intensity03	Intensity04	Intensity05	Intensity06	...	Steps51	Steps52	Steps53	Steps54	Steps55	Steps56	Steps57	Steps58	Steps59	ActivityDate
0	1503960366	4-12-2016 12:00:00 AM	0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	...	9.0	8.0	0.0	20.0	1.0	0.0	0.0	0.0	0.0	4-12-2016 12:00:00 AM
1	1503960366	4-12-2016 12:01:00 AM	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4-12-2016 12:01:00 AM
2	1503960366	4-12-2016 12:02:00 AM	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4-12-2016 12:02:00 AM
3	1503960366	4-12-2016 12:03:00 AM	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4-12-2016 12:03:00 AM
4	1503960366	4-12-2016 12:04:00 AM	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4-12-2016 12:04:00 AM

5 rows x 186 columns

Minute activity merged

From the merged activity dataset, our objective is to identify the top 10 IDS with the highest calorie expenditure per minute.



Minute activity merged

```
[ ] # Identify id Has high number value in calories

calories_by_id = dailyactivity.groupby('Id')['Calories'].sum()
max_calories_id = calories_by_id.idxmax()
print(f"ID with the highest number of calories: {max_calories_id}")

ID with the highest number of calories: 8378563200
```

For ID 8378563200, this line graph represents the initial top calorie burn



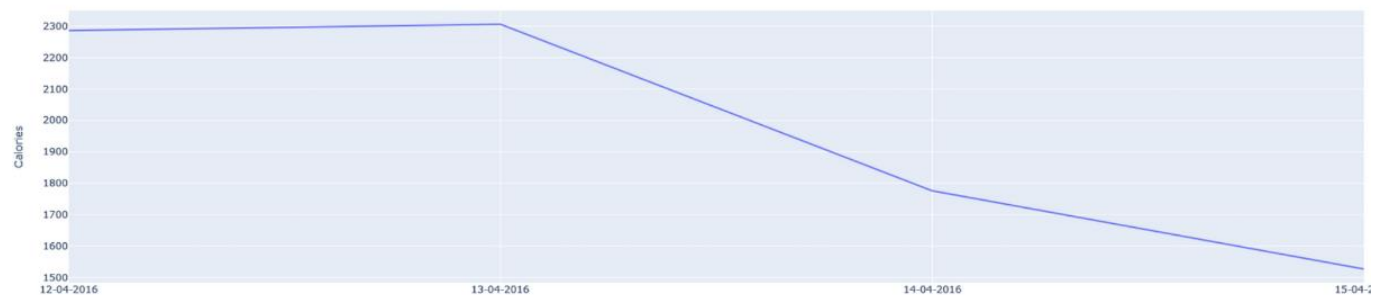
Minute activity merged

```
] # Identify Which id Has Low number value in calories

min_calories_id = calories_by_id.idxmin()
print(f"ID with the lowest number of calories: {min_calories_id}")

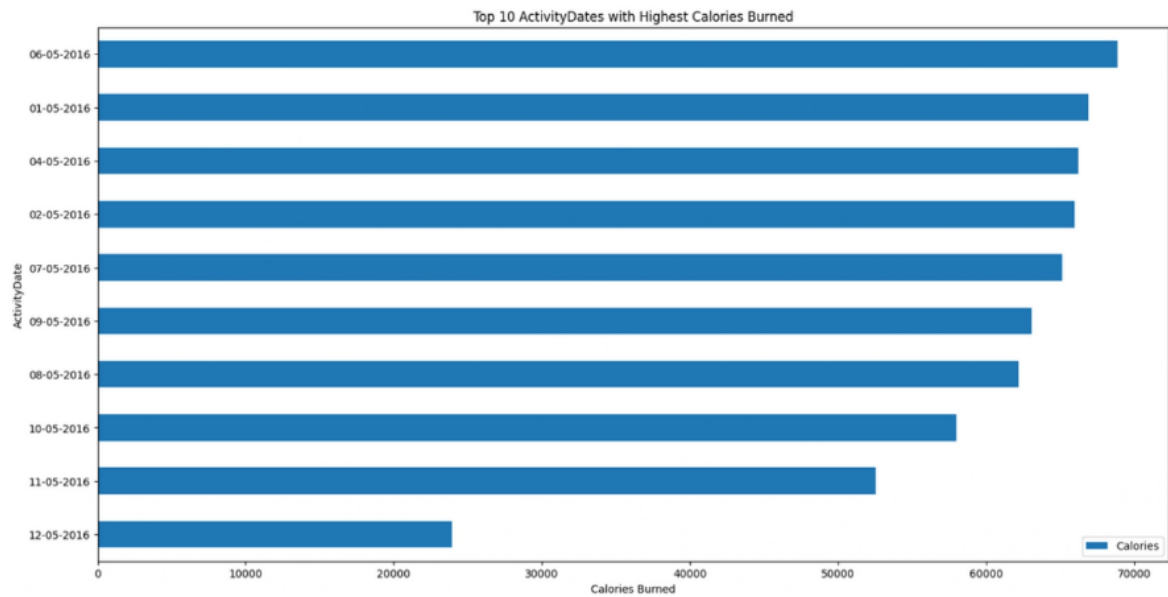
ID with the lowest number of calories: 4057192912
```

For ID 4057192912, this line graph represents the low calorie burn

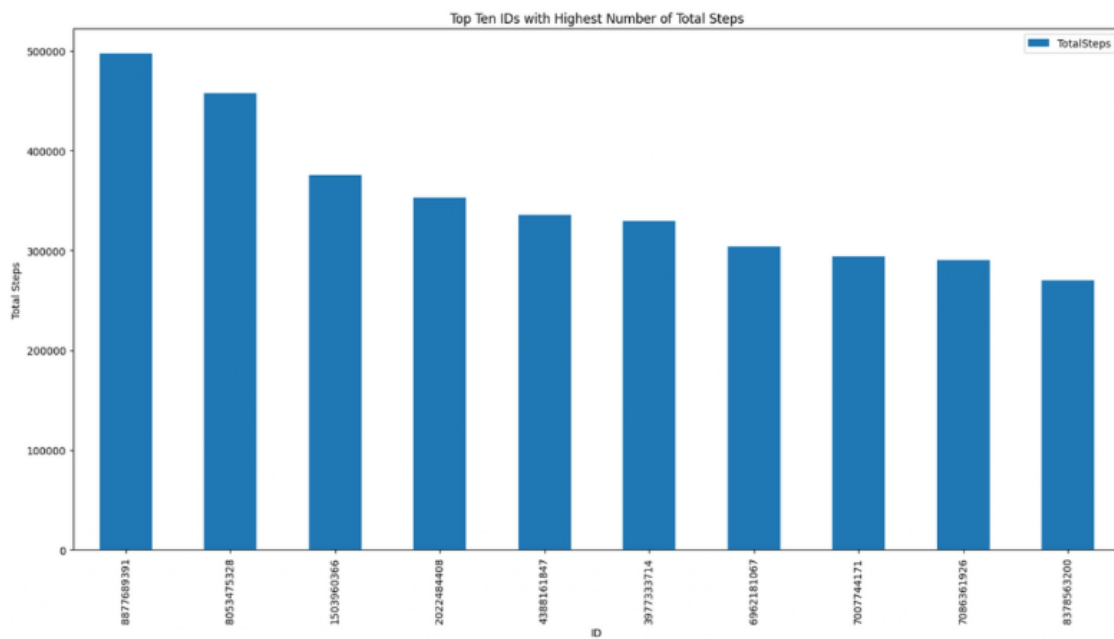


Minute activity merged

From the merged activity dataset, our objective is to identify the top 10 IDS with the highest Total steps.

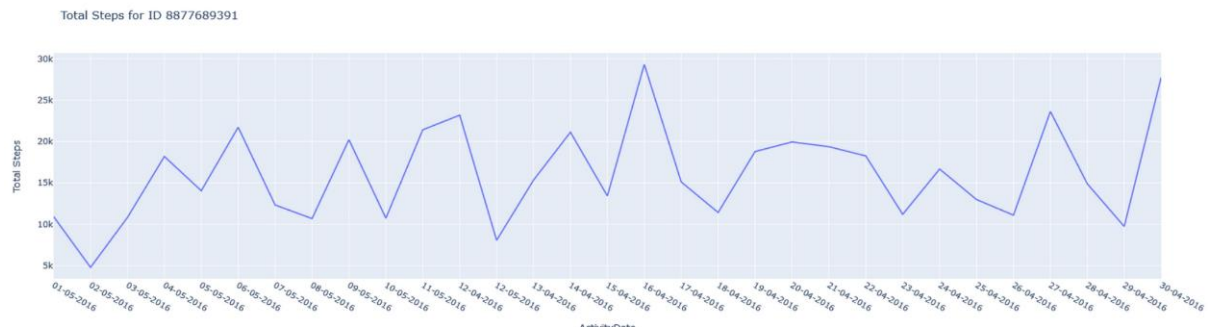


Minute activity merged



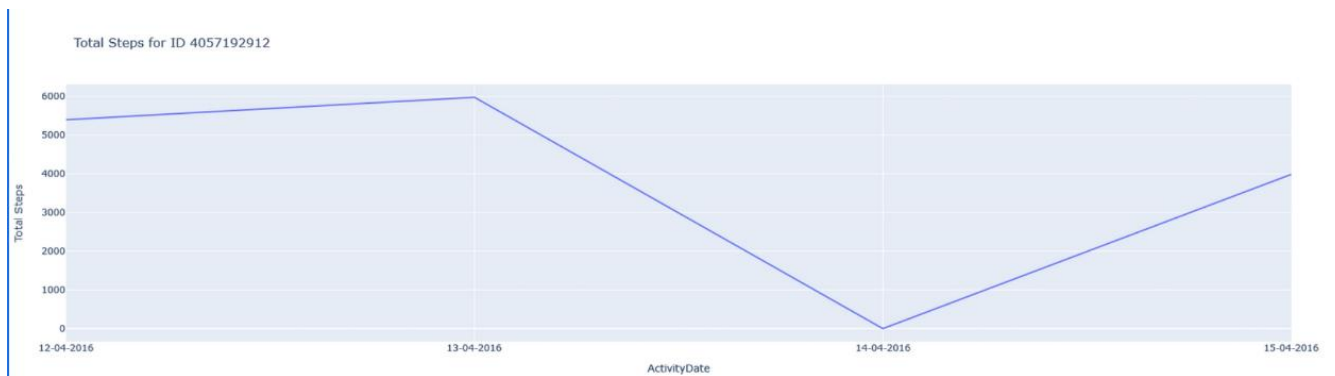
Minute activity merged

For ID 8877689391, this line graph represents the initial top Total steps



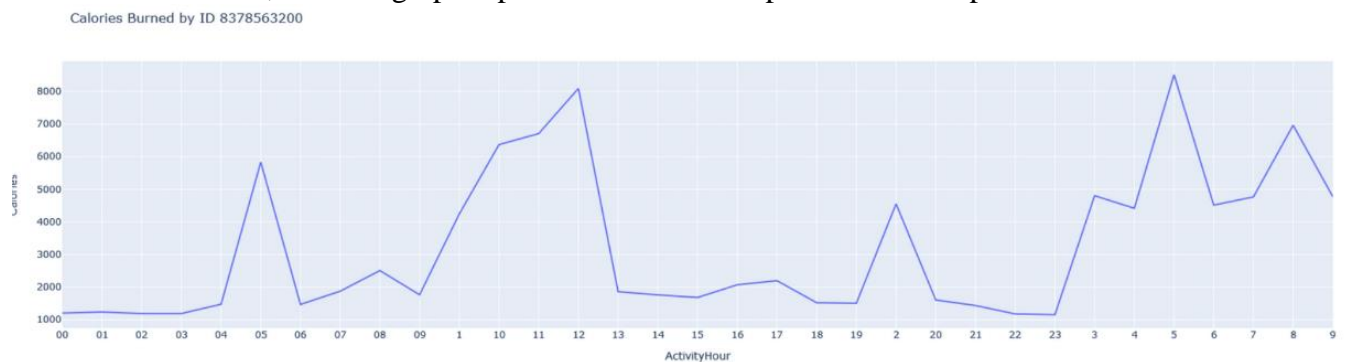
Minute activity merged

For ID 4057192912, this line graph represents the lowTotal steps.



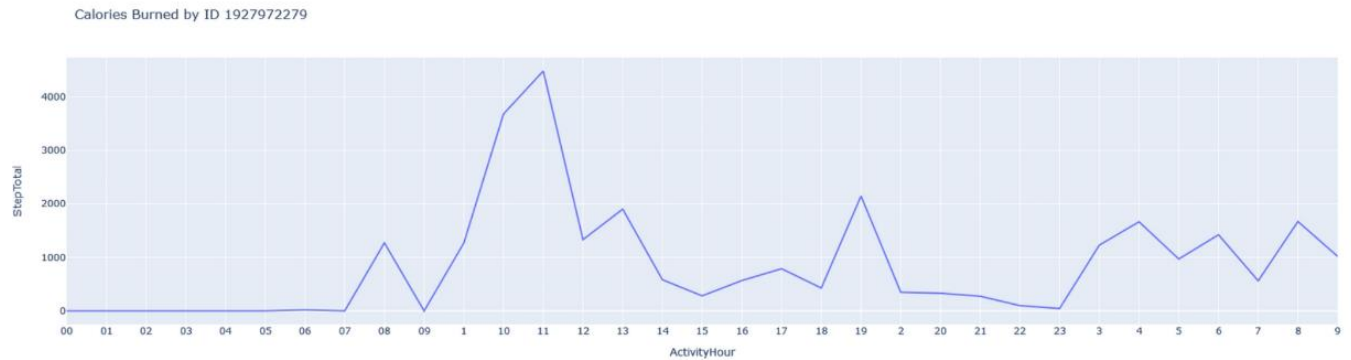
Hourly activity merged

For ID 8378563200, this line graph represents the initial top Calories burn per Hour



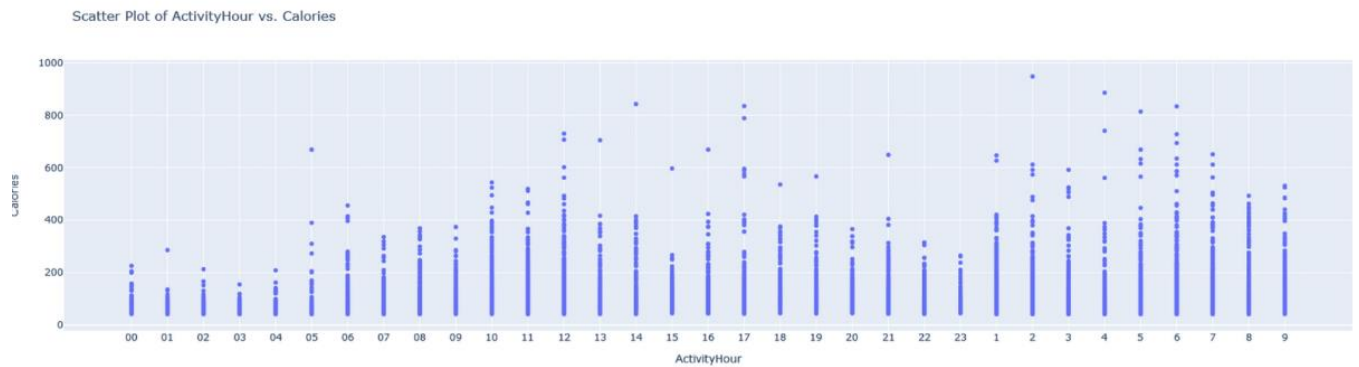
Hourly activity merged

For ID 1927972279, this line graph represents the initial low Calories burn per Hour



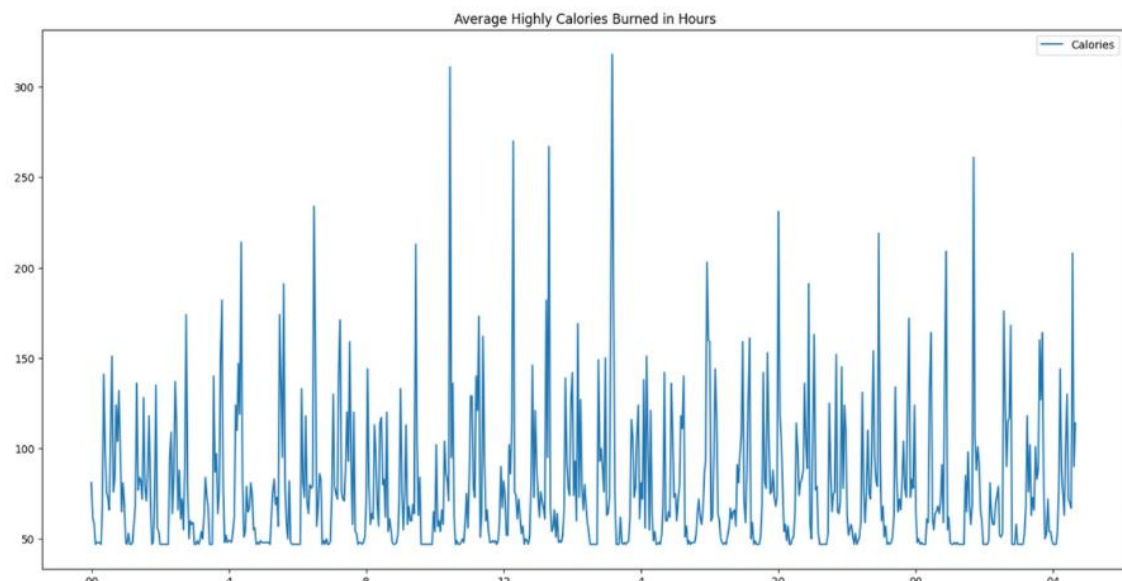
Hourly activity merged

This scatter plot graph illustrates The calories verses activityhours



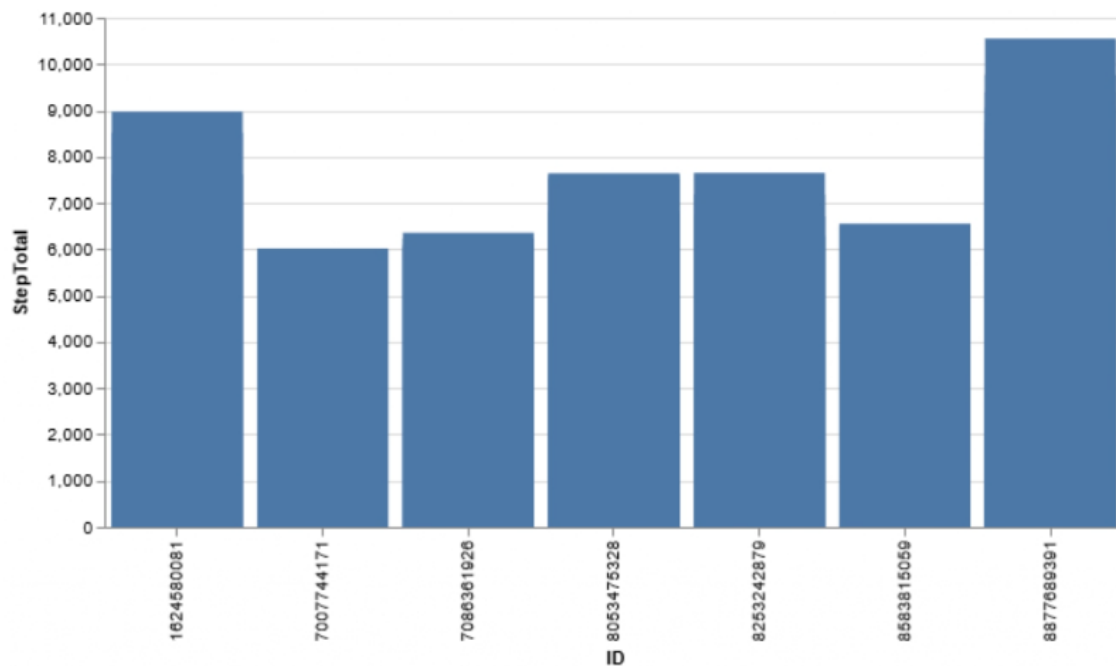
Hourly activity merged

This graph illustrates the average calories burned per hour



Hourly activity merged

Top 7 IDs with a high total number of steps



Hourly activity merged

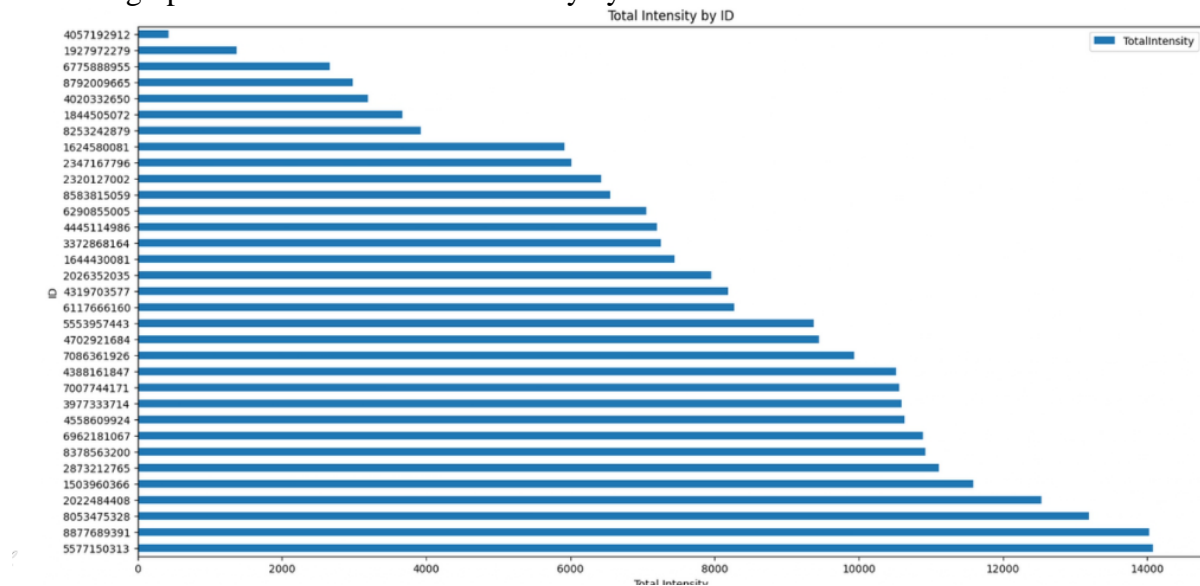
Create a tree map to represent the 33 unique Id's and its Step Total.

Average Step Total by ID



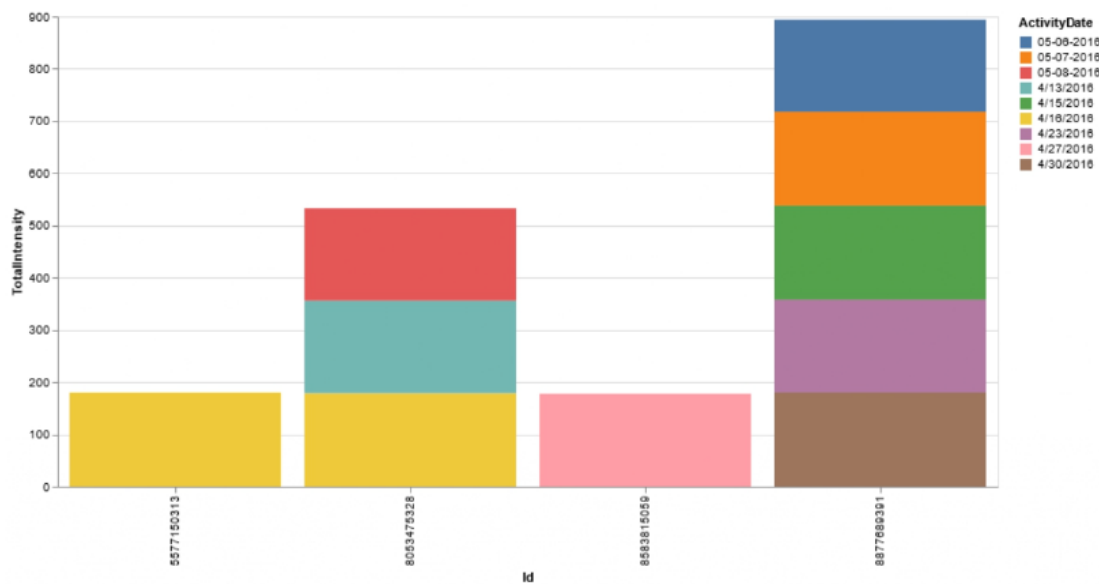
Hourly activity merged

This bar graph illustrates The Total intensity by ID



Hourly activity merged

This Stacked Bar Chart illustrates The Total intensity by Top 4 ID According to activity date.



Daily activity merged

View of cleaned and Manipulated of Daily activity dataset

```
dailyactivity.head()
```

	Id	ActivityDate	TotalSteps	TotalDistance	TrackerDistance	LoggedActivitiesDistance	VeryActiveDistance	ModeratelyActiveDistance	LightActiveDistance	SedentaryActiveDistance	...	SedentaryMinutes	Calories
0	1503960366	12-04-2016	13162	8.50	8.50	0.0	1.88	0.55	6.06	0.0	...	728	1965
1	1503960366	13-04-2016	10735	6.97	6.97	0.0	1.57	0.69	4.71	0.0	...	776	1797
2	1503960366	14-04-2016	10460	6.74	6.74	0.0	2.44	0.40	3.91	0.0	...	1218	1776
3	1503960366	15-04-2016	9762	6.28	6.28	0.0	2.14	1.26	2.83	0.0	...	726	1745
4	1503960366	16-04-2016	12669	8.16	8.16	0.0	2.71	0.41	5.04	0.0	...	773	1863

5 rows x 23 columns

Daily activity merged

Identify the number of unique id's in the daily activity dataset

```
# find the no of Unique ID's

unique_ids = dailyactivity['Id'].nunique()
print(f"Number of unique IDs: {unique_ids}")
```

Number of unique IDs: 33

Daily activity merged

Bar chart to represent the 33 unique ids and Overall Total activity date.

```
# Using dataframe dailyactivity: Calculate how many days ActivityDate with respect to Id

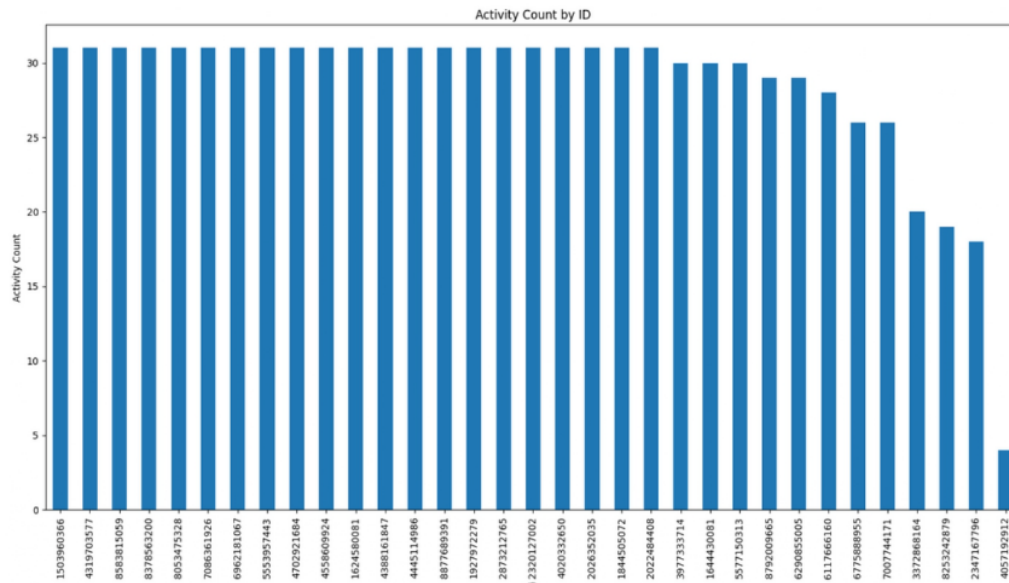
dailyactivity.groupby('Id')['ActivityDate'].count()
```

Id	ActivityDate
1503960366	31
1624580081	31
1644430081	30
1844505072	31
1927972279	31
2022484408	31
2026352035	31
2320127002	31
2347167796	18
2873212765	31
3372868164	20
3977333714	30
4020332650	31
4057192912	4
4319703577	31
4388161847	31
4445114986	31
4558089924	31
4702921684	31
5553957443	31
5577150313	30
6117666160	28
6290855005	29
6775088955	26
6962181067	31
7007744171	26
7086361926	31
8053475328	31
8253242879	19
8378563200	31
8583815059	31
8792009665	29
8877689391	31

Name: ActivityDate, dtype: int64

Daily activity merged

The information above should be transposed into a bar chart



Daily activity merged

This pie chart illustrates the distribution of activity distance by type

Activity Distance by Type



Daily activity merged

This pie chart depicts the distribution of activities based on distance by ID

Activity Distance by Type for ID 1503960366



Daily activity merged

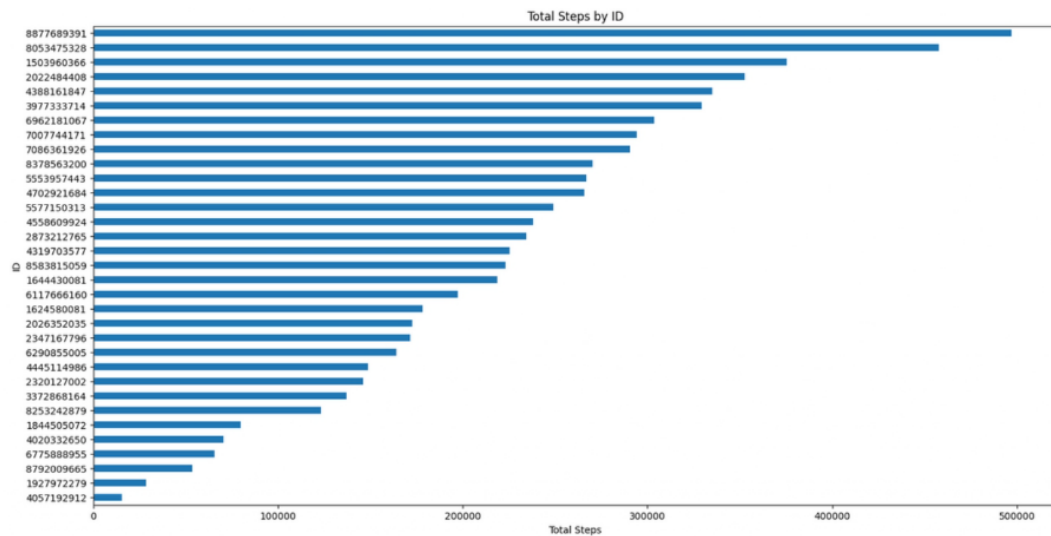
Create a tree map to represent the 33 unique Id's and its Total distance

Total Distance by ID



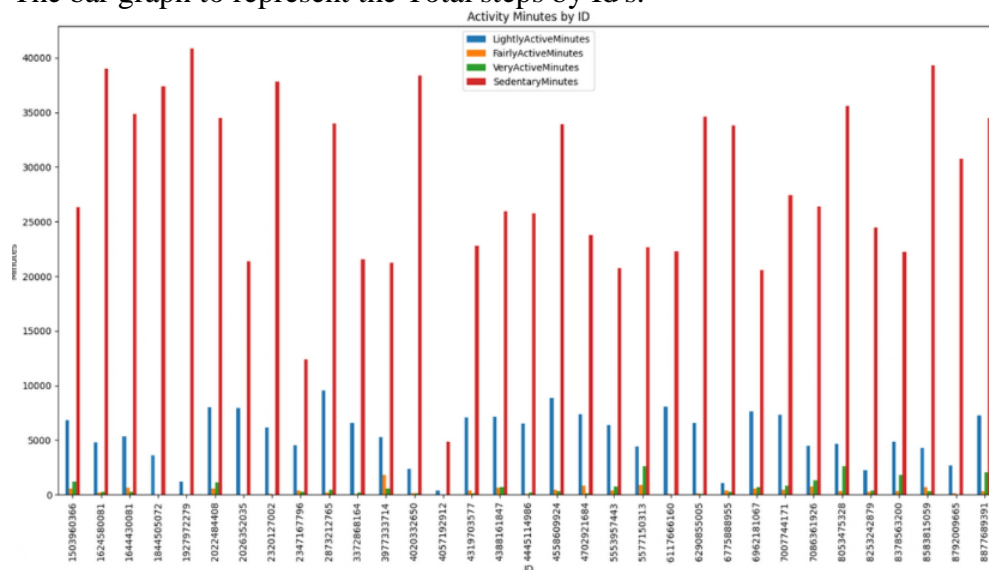
Daily activity merged

The bar graph to represent the Total steps by Id's



Daily activity merged

The bar graph to represent the Total steps by Id's.



Sleep Data

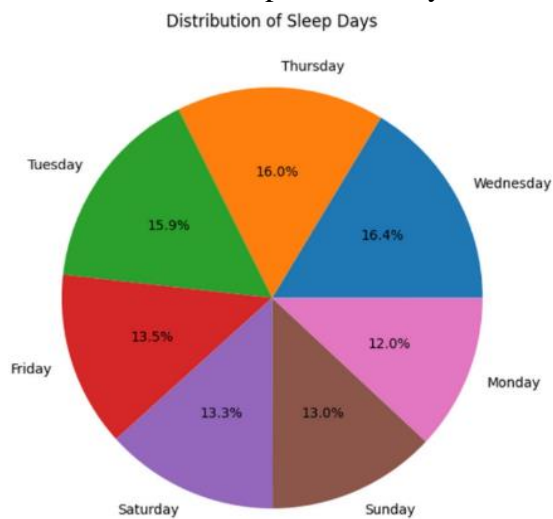
View of cleaned and Manipulated of Sleep dataset

SleepMerged.head()

	Id	SleepTime	SleepDay	SleepDate	TotalTimeInBed	TotalMinutesAsleep	TotalSleepRecords	value	logId
0	1503960366	2:47:30	Tuesday	4/12/2016	346	327	1	3	11380564589
1	1503960366	2:48:30	Tuesday	4/12/2016	346	327	1	2	11380564589
2	1503960366	2:49:30	Tuesday	4/12/2016	346	327	1	1	11380564589
3	1503960366	2:50:30	Tuesday	4/12/2016	346	327	1	1	11380564589
4	1503960366	2:51:30	Tuesday	4/12/2016	346	327	1	1	11380564589

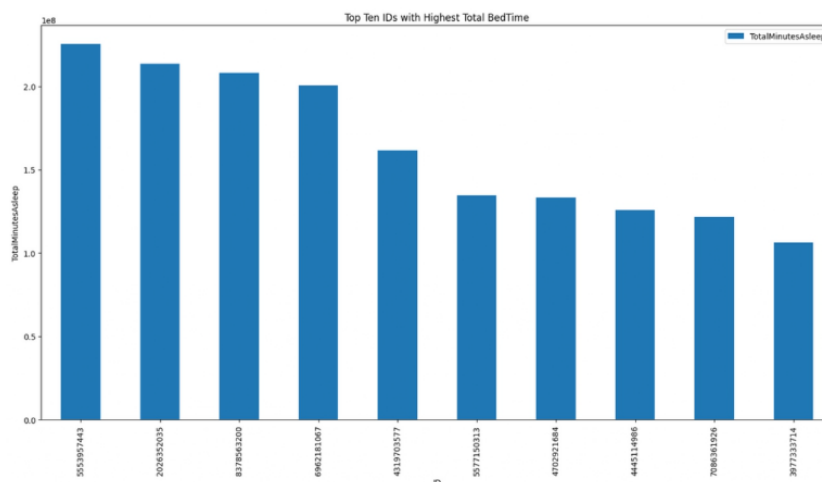
Sleep Data

The distribution of sleep on weekdays



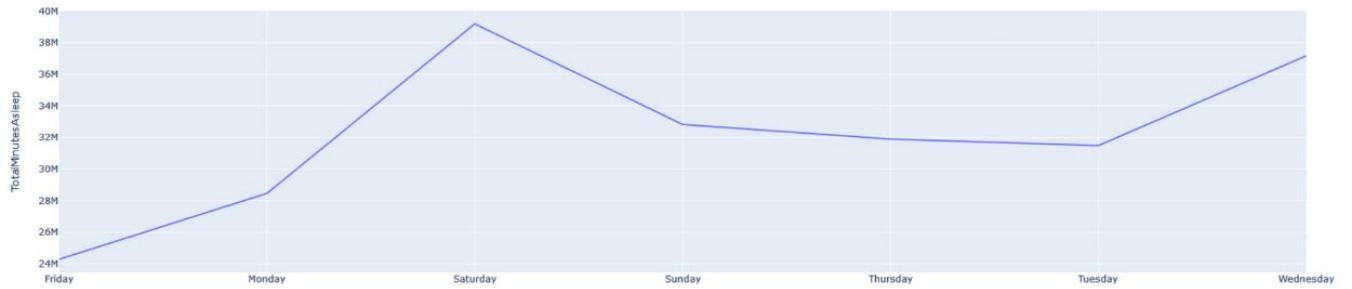
Sleep Data

From the Sleep Dataset, our objective is to identify the top 10 IDS with the Highest total bedtime.



Sleep Data

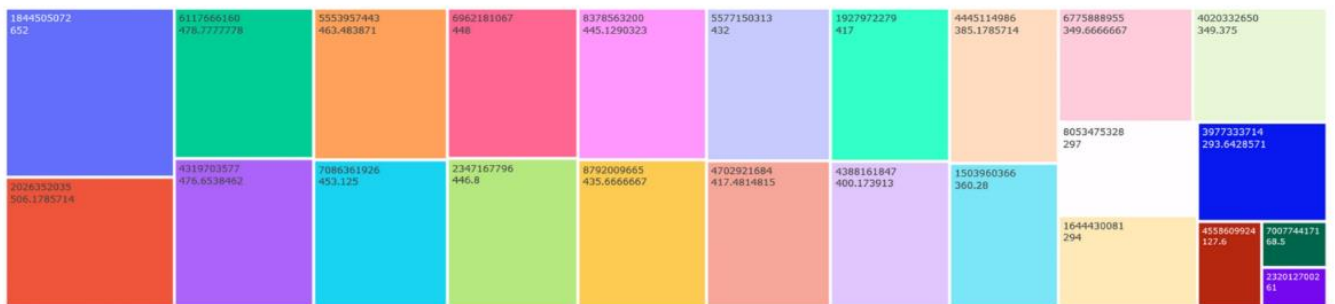
For ID 5553957443, this line graph represents the Total steps Time by day.



Sleep Data

Create a tree map to represent the 33 unique ids and its average minute asleep

Average Minutes Asleep by ID



Sleep Data

To represent Time in bed vs total minutes asleep in a scatter plot graph.

