**FitBit Fitness Tracker Data**

Pattern recognition with tracker data: Improve Your Overall Health

Link: <https://www.kaggle.com/datasets/arashnic/fitbit>

This dataset was initially collected from 30 female users of the Fitbit smartwatch. Thirty eligible Fitbit users consented to the submission of 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 use of different types of Fitbit trackers and individual tracking behaviors / preferences.

**Analysis Questions:**

1. What are some trends in smart device usage?
2. How could these trends apply to healthcare customers?
3. How could these trends help influence healthcare companies’ marketing strategies?

**Overview of Problem to be Solved:**

We are looking to unlock new growth opportunities that could be revealed by gaining insight on how smart wearable devices are used which will help guide the marketing strategy for health-focused companies.

**Directions to Consider:**

1. Studying the relationship between calories burned and the total steps walked in a day
2. The average total number of active minutes in a day
3. Analyze the very active, fairly active, and lightly active minutes on each day of the week
4. The number of inactive minutes on each day of the week
5. The number of calories burned on each day of the week
6. Predicting the number of calories burned in a day

**Hypothesis to test:**

1. People who have more intense activity burn more calories.
2. People who have more intense activity sleep better.

**Test Statistic used: Correlation Coefficient & Linear Regression**

**Correlation coefficient:** A correlation coefficient is a measure of the strength and direction of the linear relationship between two variables. It ranges from -1 to +1, where -1 indicates a perfect negative correlation, 0 indicates no correlation, and +1 indicates a perfect positive correlation. A correlation coefficient can be used to determine the degree to which two variables are related to each other.

**Linear regression:** A linear regression is a statistical method used to model the relationship between two variables by fitting a linear equation to the observed data. The equation takes the form of y = mx + b, where y is the dependent variable, x is the independent variable, m is the slope, and b is the y-intercept. A linear regression line can be used to make predictions about the value of y for a given value of x.

# How do I test a hypothesis using the critical value of t?

To [test a hypothesis](https://www.scribbr.com/statistics/hypothesis-testing/) using the critical value of t, follow these four steps:

1. Calculate the t value for your sample.
2. Find the critical value of t in the [t table](https://www.scribbr.com/statistics/students-t-table/).
3. Determine if the (absolute) t value is greater than the critical value of t.
4. Reject the null hypothesis if the sample’s t value is greater than the critical value of t. Otherwise, don’t reject the [null hypothesis](https://www.scribbr.com/statistics/null-and-alternative-hypotheses/).