

## Advanced Analytics & Dashboard Design

### 6.1: Sourcing Open Data

#### Bikes Sales in Europe

##### About this dataset

This dataset provides a comprehensive look for bikes and accessories across country. The dataset also includes information about the date of sale, the store location, the cost price, the selling price, and the profit made on each sale. To extract insights from the sales data, we can analyze the sales figures across different dimensions such as time, product category, store location, and customer demographics.

##### Data Source

Bikes Sales in Europe is external data.

Source: <https://www.kaggle.com/datasets/sadiqshah/bike-sales-in-europe>

##### OBJECTIVE

The objective of this dataset is to analyze the sales figures of bikes and accessories across the country and extract insights that can help improve sales and profitability. By analyzing the sales data across different dimensions such as time, product category, store location, and customer demographics, we can identify the top-selling products, the most profitable stores, the most lucrative customer segments, and the trends in sales over time. Overall, the objective of this dataset is to leverage the power of data analytics to drive business growth and profitability by making informed decisions based on the insights extracted from the sales data. The analysis is based on the following key questions.

##### KEY QUESTIONS:

- What is the distribution of customer age and gender in our sales data?
- Which product categories and sub-categories are the most profitable for us?
- What is the average order quantity and revenue per sale? How have these metrics changed over time?
- Which states and countries are our best and worst performing markets in terms of sales and profitability?
- Is there a relationship between unit cost and unit price for our products? How does this relationship vary across different product categories?

## **Data Collection**

This dataset provides a comprehensive look cost, price and profit for each product within dataset.

## **Data contents**

The dataset includes information about the date of sale, the store location, the cost price, the selling price, and the profit made on each sale.

## **Limitations of Data Set:**

In this data set we don't have data for some country e.g., The Netherlands and we don't have the possibility to compare sales between country across Europe.

## **Data Quality**

Ensuring data quality for the data set is crucial before conducting further analysis. The data set contains 112036 rows and 17 columns. The initial step involves reviewing all the information contained within it and identifying the specific columns that we will use to gain insights. Performing in Python the data consistency and wrangling.