# **Scraping Data through Google Maps**

# **Introduction**

This project involves scraping data from **Google Maps** to extract details about bicycle shops and marketplaces in Copenhagen. The extracted data includes shop names, addresses, contact information, ratings, phone number, website, plus code and more. The script automates this process using Selenium and saves the results in a CSV file.

# **Objectives**

 Extract detailed information about bicycle-related businesses in Copenhagen.

 Automate data collection to avoid manual searches.

 Store the data in a structured format for further analysis.

# **Technologies Used**

 **Python**: Programming language for scripting.

 **Selenium**: Web automation tool for interacting with Google Maps.

 **Pandas**: Library for data storage and processing.

 **WebDriver Manager**: Manages Chrome Driver automatically.

# **Objectives**

**Challenge 1: Dynamic Content Loading**

 **Solution**: Implemented scrolling and waiting functions to ensure data loads before extraction.

**Challenge 2: Stale Element Errors**

 **Solution**: Used retries with StaleElementReferenceException handling.

**Challenge 3: CAPTCHA or Blocking**

 **Solution**: Added time delays between requests to mimic human-like browsing.

# **Challenges Faced After Scraping**

**Handling Inconsistent Data Formats** – Ratings and reviews were sometimes combined in different formats, requiring regex extraction.

**Duplicate Entries** – Some businesses appeared multiple times, requiring deduplication based on phone, address, and other fields.

**Missing or Incomplete Data** – Many entries had missing phone numbers, websites, or addresses, affecting data completeness.

**Combining Multi-Language Data** – Merging Danish and English datasets without losing critical information while maintaining data consistency.