**Milestone 1: Data Sources & Predictive Models**

**1. Data Sources & Acquisition Methods**

**1.1 Government & Industry Reports**

**Examples:**

* **Statistics Denmark:** <https://www.dst.dk/en>
* **Copenhagen City Council:** <https://urbandevelopmentcph.kk.dk/mobility-cycling/copenhagen-the-best-cycling-city-in-the-world>​
* **Bike Europe:** <https://www.bike-eu.com/>
* **Eurostat:** <https://ec.europa.eu/eurostat>

**Acquisition Method:**

* Download reports from official websites.
* Extract relevant statistics using data processing tools.

**Storage Strategy:**

* Relational database (MySQL/PostgreSQL) for structured data.

**Reliability & Completeness:**

* These are official and regularly updated sources, ensuring data accuracy and consistency.

**1.2 Business Directories**

**Examples:**

* **Google Maps:** <https://www.google.com/maps>​
* **Yelp:** <https://www.yelp.com/>​
* **LinkedIn:** <https://www.linkedin.com/>​
* **Yellow Pages:** <https://www.yellowpages.com/>

**Acquisition Method:**

* Web scraping using BeautifulSoup/Scrapy.
* API access where available.

**Storage Strategy:**

* NoSQL database (MongoDB) to handle semi-structured data.

**Reliability & Completeness:**

* Business directories are continuously updated, ensuring comprehensive market data.

**1.3 Trade & Customs Data**

**Examples:**

* **ITC Trade Map:** <https://www.trademap.org/>​
* **Eurostat:** <https://ec.europa.eu/eurostat>

**Acquisition Method:**

* Download structured reports from official portals.

**Storage Strategy:**

* Relational database (MySQL/PostgreSQL) for structured trade data.

**Reliability & Completeness:**

* Government and trade organizations provide verified, official trade statistics.

**1.4 E-commerce & Marketplaces**

**Examples:**

* **DBA.dk:** <https://www.dba.dk/>​
* **Cykelpartner.dk:** <https://www.cykelpartner.dk/>​
* **Amazon:** <https://www.amazon.com/>

**Acquisition Method:**

* Web scraping (Scrapy) or API access.

**Storage Strategy:**

* NoSQL (MongoDB) for varying data structures.
* Cloud storage (AWS S3) for scalability.

**Reliability & Completeness:**

* Real-time market insights from e-commerce platforms ensure up-to-date consumer trends.

**1.5 Market Research Reports**

**Examples:**

* **IndustryARC:** <https://www.industryarc.com/>​
* **McKinsey:** <https://www.mckinsey.com/>​
* **IBISWorld:** <https://www.ibisworld.com/denmark/industry-statistics/>​

**Acquisition Method:**

* Purchased reports or accessed via institutional subscriptions.

**Storage Strategy:**

* Cloud-based document storage (Google Drive/AWS S3).

**Reliability & Completeness:**

* Professional industry reports provide in-depth market analysis, though access may be restricted.

**2. Predictive Models & Trends**

**2.1 Chosen Predictive Models**

**Time Series Forecasting (ARIMA, Prophet):**

* **Why?** Best suited for predicting bicycle sales and usage trends over time.
* **Trends Predicted:** Monthly and yearly cycling adoption rates in Copenhagen.

**Regression Models (Linear Regression, Random Forest Regression):**

* **Why?** Effective for analyzing the impact of economic factors on bike sales.
* **Trends Predicted:** Price elasticity and consumer spending on bicycles.

**Clustering (K-Means, DBSCAN):**

* **Why?** Useful for segmenting consumer preferences.
* **Trends Predicted:** Identifying distinct cyclist demographics and behaviors.

**Sentiment Analysis (Natural Language Processing - NLP):**

* **Why?** Helps understand consumer sentiment from reviews and social media.
* **Trends Predicted:** Customer satisfaction and emerging market trends.