

|  |
| --- |
| Business Template  **ChicaGo CICLISTIC 2023** |
|  |



Contents

[1 Business Description 3](#_Toc155614186)

[1.1 Business background 3](#_Toc155614187)

[1.2 Problems because of poor data management 3](#_Toc155614188)

[1.3 Benefits from implementing a Data Warehouse 3](#_Toc155614189)

[1.4 DATASETS DESCRIPTION 3](#_Toc155614190)

[1.5 GRAIN / DIM / FACT 4](#_Toc155614191)

[2 Business Layer 3NF 4](#_Toc155614192)

[3 Business Layer Dimensional Model 4](#_Toc155614193)

[4 Logical Scheme 4](#_Toc155614194)

[5 Data Flow 4](#_Toc155614195)

[6 Fact Table Partitioning Strategy 4](#_Toc155614196)

# 

# Business Description

## Business background

Cyclistic 2023 is a dynamic and innovative bike-sharing service operating in the vibrant city of Chicago. In collaboration with Divvy Bikes, the service provides convenient and eco-friendly transportation options to residents and visitors alike. The business thrives on the idea of making urban mobility seamless, efficient, and sustainable.

## Problems because of poor data management

Ineffective data management poses a significant hurdle for Cyclistic 2023. Without proper tools for analysis, the business risks insufficient insights into user behavior, popular routes, and market trends. This hampers strategic decision-making, puts the company at a competitive disadvantage, and limits the formulation of targeted business strategies. To thrive in the bike-sharing market, Cyclistic 2023 must prioritize data utilization and invest in advanced analytical tools for informed and proactive operations.

## Benefits from implementing a Data Warehouse

The integration of a data warehouse at Cyclistic 2023 unlocks key advantages, addressing prior data management challenges and enhancing strategic decision-making. Through this implementation, Cyclistic 2023 gains the ability to:

* **Optimize Bike Distribution:**

Identify demand patterns for electric and classic bikes, ensuring strategic placement for optimal user accessibility.Which ones have the widest distribution of prices?

* **User Behavior Patterns:**

Correlate user behaviors with ride patterns, enabling targeted initiatives to enhance user satisfaction and improve service offerings.

* **Diversity in Bike Usage:**

Understand user preferences between electric and classic bikes, facilitating adjustments to the fleet for more effective meeting of user demands.

* **Strategic Marketing Decisions:**

Utilize data insights to inform marketing strategies, targeting peak usage times, popular routes, and preferred bike types

* **And many other.**

## DATASETS DESCRIPTION

The first dataset contains the following information about rents on Chicagos market.

Cycle Information:

Cycle type: The type of cycle (electric or classic).

Location Information:

Start station: In which station ride started.

End station: In which station ride ended.

Start lat: Start coordinates.

Start lng long: Start coordinates.

End lat: End coordinates.

End lng: End coordinates.

Customer Information:

Name: Name of the customer

Last\_name: Last name of the customer

Gender: The gender who rented the bike (men or women).

Age: The age range of the customer.

Member: The customer registered in our system or not.

Time:

Start\_date: The date when ride started.

Start\_time: The time when ride started.

End\_date: The date when ride ended.

End\_time: The time when ride ended.

Ride\_time\_mins: Duration of the ride.

Price:

Price\_for\_min: The price per minute, it’s different for the electric and classic cycles

Dicount: Dicount by percent depending on rent duration

The second dataset is for online paid rides; it mostly contains the same dates, and additionally, it has the following information...

Payment:

Bank\_account:

Bank\_name:

Payment\_date:

Payment\_amount:

## GRAIN / DIM / FACT

# Business Layer 3NF

# Business Layer Dimensional Model

# Logical Scheme

# Data Flow

# Fact Table Partitioning Strategy