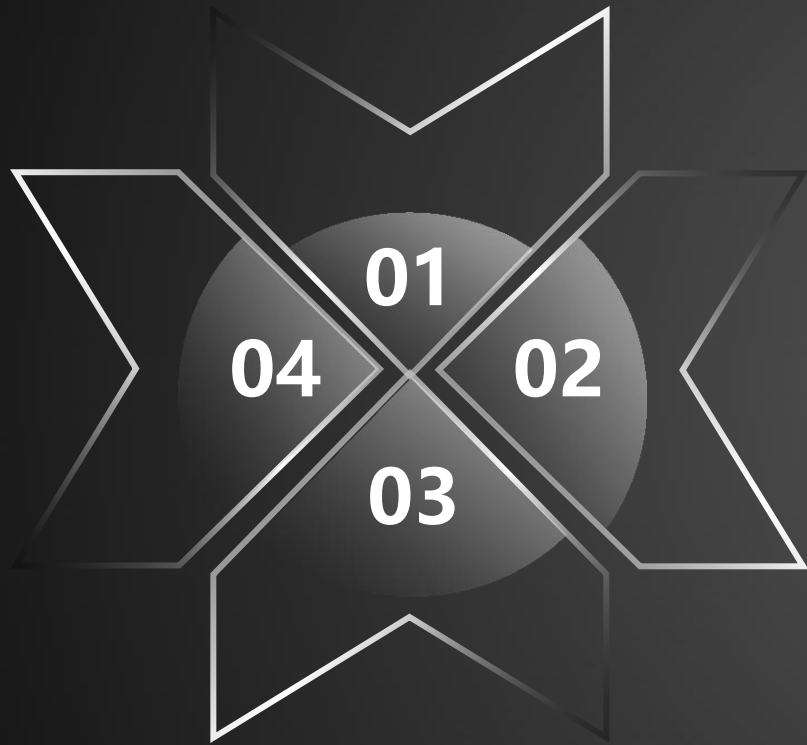# CAPSTONE PROJECT

Reporter: Joshua Kuriakose, 22BCAF31
Bebinto Thomas, 22BCAF14
Shivaani K, 22BCAF58
Efin Rijo, 22BCAF21

01. INTRODUCTION

The Woodcorp Order-to-Cash (O2C) datasets provide an in-depth view of the company's order processing, fulfillment, and financial transactions. The data is structured into two tables: the O2C Activity Table and the O2C Case Table, each serving a distinct purpose in tracking and analyzing the order lifecycle. The Woodcorp O2C dataset contains transactional records related to this process, making it valuable for analyzing business performance, supply chain efficiency, and customer satisfaction. However, before conducting any meaningful analysis, the dataset must undergo thorough data cleaning to ensure accuracy, consistency, and completeness.

01

04

02

03

02. Dataset Overview

Two Key Tables :

O2C Activity Table – Tracks order processing
events with timestamps.
O2C Case Table – Contains order details,
customer info, and financial data.

Common Fields :

CASE_KEY – Unique order identifier.
EVENTTIME – Timestamp of order-related
activities.
ORDER_VALUE, UNIT_PRICE – Financial data.
DELIVERED_DATE, PROMISED_DATE –
Delivery timeline details.

Key Improvements in the Cleaned Data:

1, Removal of Duplicate Entries.
2, Correction of Missing or Incomplete Data
3, Standardization of Formats
4, Resolving Order & Delivery Discrepancies
5, Improved Financial Accuracy

03.Problems Faced by Woodcorp Due to Data Issues in the O2C Process

‹Delays in Order Processing

◇ Problem:
Order processing delays vary across cities, with Aachen and Essen having the highest delays.
Some cities like Wuppertal have missing or zero data, which may indicate reporting issues.

◇ Impact:
Customer dissatisfaction → Leads to frustration and cancellations.
Operational inefficiencies → Bottlenecks in logistics or processing.
Revenue loss → Customers may switch to competitors.

◇ Example:
In Aachen, delays are around 35 days, causing potential cancellations. The company should identify and fix the issue.

‹ Factory Order Distribution

◇ Problem:

Orders are concentrated in Aachen and Essen, while other factories, like Duisburg and Wuppertal, receive very few.

This creates overburdened factories and underutilized facilities, leading to inefficiencies.

◇ Impact:

Production Delays – High-volume factories face backlogs & errors.

Customer Delays – Late deliveries affect customer satisfaction.

Financial Loss – Increased costs & reduced profitability.

◇ Example:

Aachen receives 8,000 orders, while Duisburg gets only 200.

Aachen struggles with delays, while Duisburg's resources go unused.

Result: Higher costs, late deliveries, and unhappy customers.

‹ Order Completion Time per Factory

◇ Problem:
Order completion time varies significantly across factories, with Wuppertal having the highest time.
Crefeld shows the lowest order completion time, indicating inconsistencies in processing efficiency.

◇ Impact:
Delays in high-time factories → Slower order fulfillment and customer dissatisfaction.
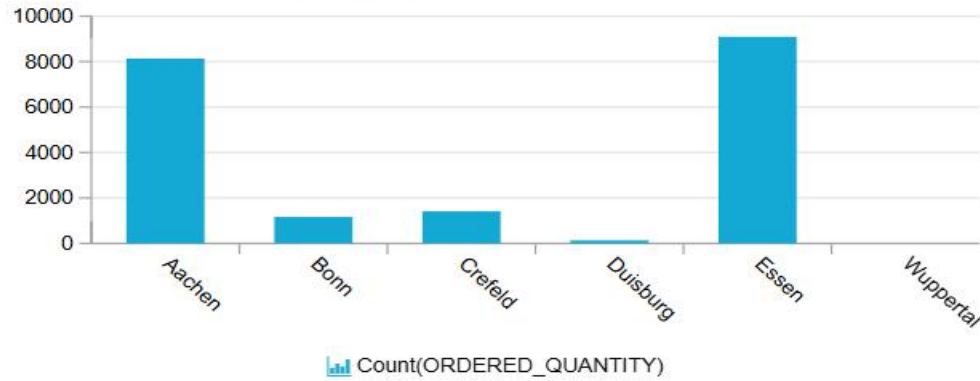Operational inefficiencies → Possible resource misallocation affecting productivity.
Increased costs → Longer processing times may lead to higher operational expenses.

◇ Example:
Wuppertal has the highest order completion time (~15,000+), while Crefeld has the lowest. The company should investigate reasons (e.g., staff shortages, machine inefficiencies) and optimize processes.
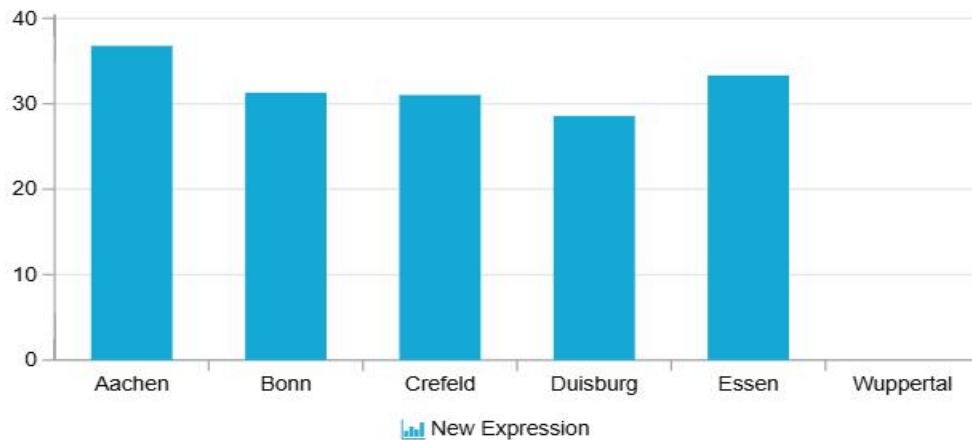
PROGRESS