

# Circle K Loyalty Simulation Project

## Case-Based Data & Analytics Addendum

Abraham Borgi

### Table of Contents


- Project Objective ..... 2
- Project Process & Workflow ..... 2
  - Step 1: Scoping & Planning ..... 2
  - Step 2: Data Simulation & Logic Modeling (Python) ..... 2
  - Step 3: PostgreSQL ..... 3
  - Step 4: Power BI Dashboarding ..... 3
  - Step 5: Analysis & Question Answering ..... 3
  - Step 6: Finalization & Documentation ..... 3
- ETL Pipeline Overview ..... 4
- Data Model Overview ..... 5
- Power BI Report ..... 7
  - Page 1: Overview ..... 7
  - Page 2: Segmentation ..... 8
  - Page 3: Campaigns ..... 9
- Summary Outcome ..... 10
- Other relevant files ..... 10

## Project Objective

This project was developed with three key objectives:

1. **To demonstrate my capability** in modeling loyalty behavior, performing data-driven customer segmentation, and evaluating the impact of campaign mechanics using a robust simulation approach.
2. **To go beyond a traditional job application** by showcasing a concrete example of how I approach analytical challenges — not just describing what I can do but actively demonstrating it. **Show it, don't say it.**
3. **To challenge myself** by delivering a realistic, end-to-end simulation project under time constraints, without prior access to company-specific data or internal systems.

The result is a comprehensive simulation and reporting pipeline tailored to Circle K's European loyalty program. It includes synthetic data generation, campaign rule logic, behavior modeling by customer segment, and interactive dashboards built in Power BI.

 **Please note:** All transactional and customer data used in this simulation is fully synthetic. It is based on assumed customer types and statistical distributions designed to reflect plausible loyalty behavior. While inspired by real-world patterns, the data does not reflect actual Circle K operations or customers.

## Project Process & Workflow

This simulation project was built in phases, using an iterative approach that allowed testing, validation, and refinement at each step:

### Step 1: Scoping & Planning

- Defined key business questions inspired by the Circle K job posting and prior knowledge about the Circle K.
- Designed a scalable data model and identified core entities (customers, cards, stations, transactions, campaigns)

### Step 2: Data Simulation & Logic Modeling (Python)

Simulated realistic data for:

- Customers
- Segmentation groups
- Loyalty cards

- Cashiers (linked to stations)
- Transactions and transaction lines
- Campaign rules and reward logic

Incorporated external data from DAGI and DAR for:

- Stations (Circle K's P-units)
- Postcodes, municipalities, and regions
- Address structures

Stored all generated data in PostgreSQL for analysis and reporting

### Step 3: PostgreSQL

- Modeled the relational schema and created upsert logic for each entity
- Structured clean, loadable tables with primary/foreign keys and type constraints
- Created relevant Views (in SQL)

### Step 4: Power BI Dashboarding

- Connected PostgreSQL to Power BI
- Created four interactive report pages with clear KPIs and dynamic visuals
- Validated that simulated data supported relevant insights: customer behavior, campaign performance, segment LTV

### Step 5: Analysis & Question Answering

Used the dashboard to answer core business-facing questions:

- Which segments are most active?
- What is the average LTV per segment?
- What products drive loyalty engagement?

Each question was supported by visuals, KPIs, and written insight to demonstrate real-world decision-making potential.

### Step 6: Finalization & Documentation

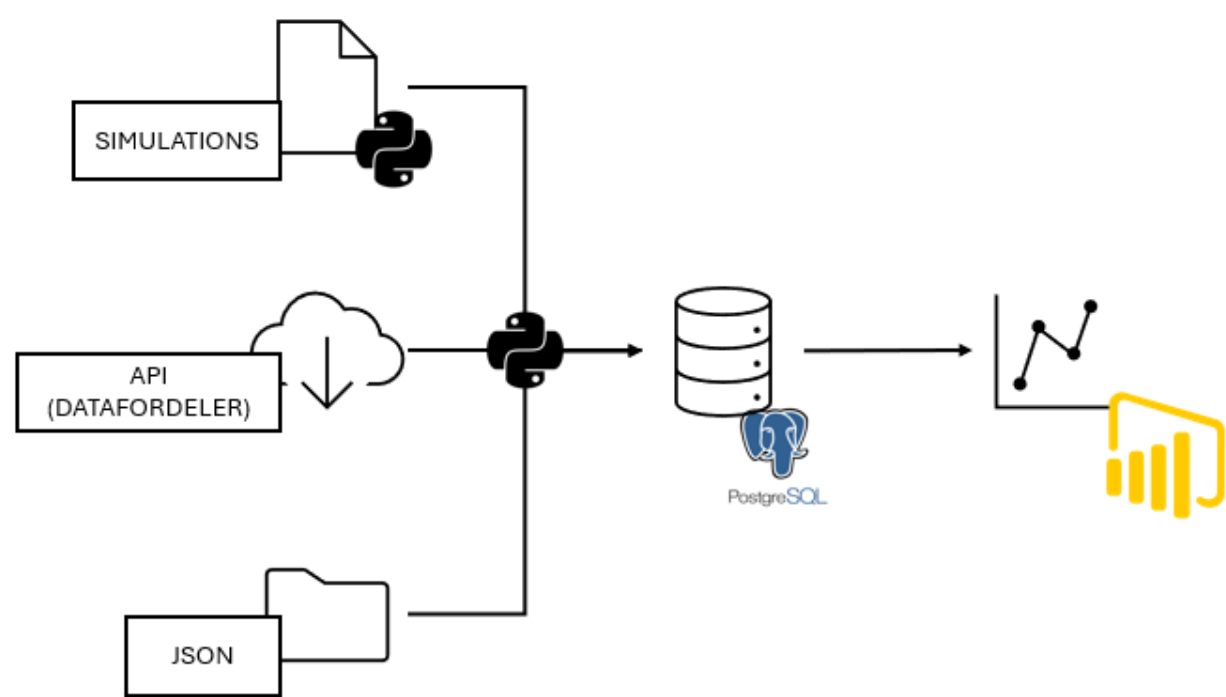
- Wrote structured documentation for scripts, data model, and logic
- Created dashboard screenshots and exportable summary for use in job application

ETL Pipeline Overview

The project was developed as an end-to-end simulation and analytics pipeline, designed to reflect realistic customer behavior, campaign mechanics, and geographic structure for Circle K.

The full implementation is included in the attached ZIP file (CircleK\_Loyalty\_Simulation.zip), which contains all simulation scripts, data loaders, and PostgreSQL integration logic.

The diagram and table below summarize how data flows from simulated and external sources into a PostgreSQL database and is then visualized in Power BI.



Stage	Description
Simulation (Python)	Simulated realistic customers, cashiers, products, and campaign behavior using probabilistic and rules-based models
External Data (DAR, DAGI)	Geographic and address data ingested from official Danish datasets.
Storage (PostgreSQL)	All data modeled and stored in relational form using PostgreSQL

Upserts (Python)	Data validated and loaded into database via modular upsert scripts
Visualization (Power BI)	Connected to PostgreSQL, built interactive dashboard with KPIs, segment breakdowns, campaign insights, and time series trends

ETL scripts were organized into api/, simulations/, and upserts/ folders, executed via orchestration scripts in pipelines/.

Data Model Overview

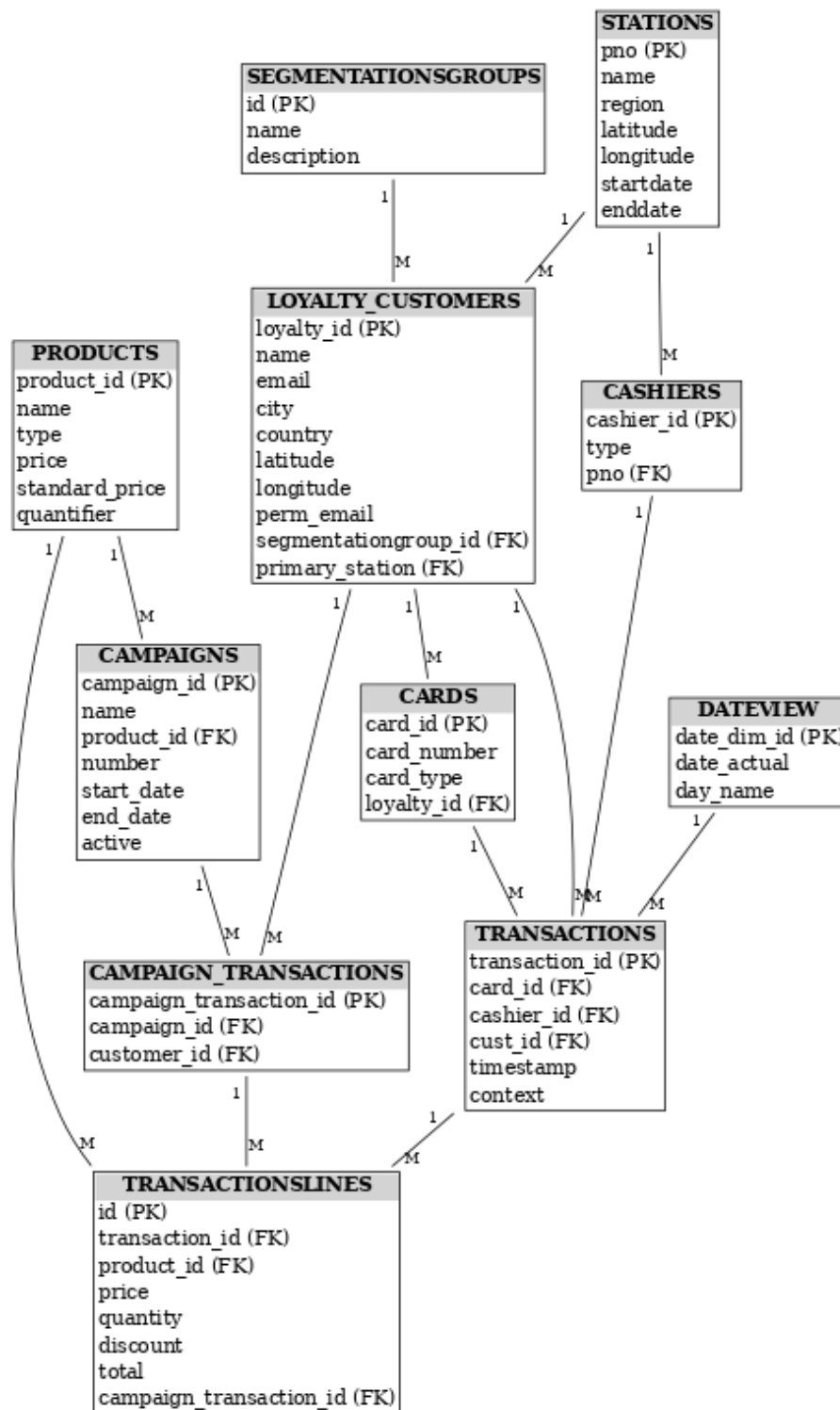
The simulated data model was designed to reflect a simplified yet realistic structure of Circle K’s loyalty ecosystem.

This model is a **simplification** of what I think the actual Circle K loyalty architecture is based on my prior knowledge.

While it omits components like notifications or CRM integration, it captures the core structure needed to analyze customer behavior, transactions, and campaign performance.

Core Tables

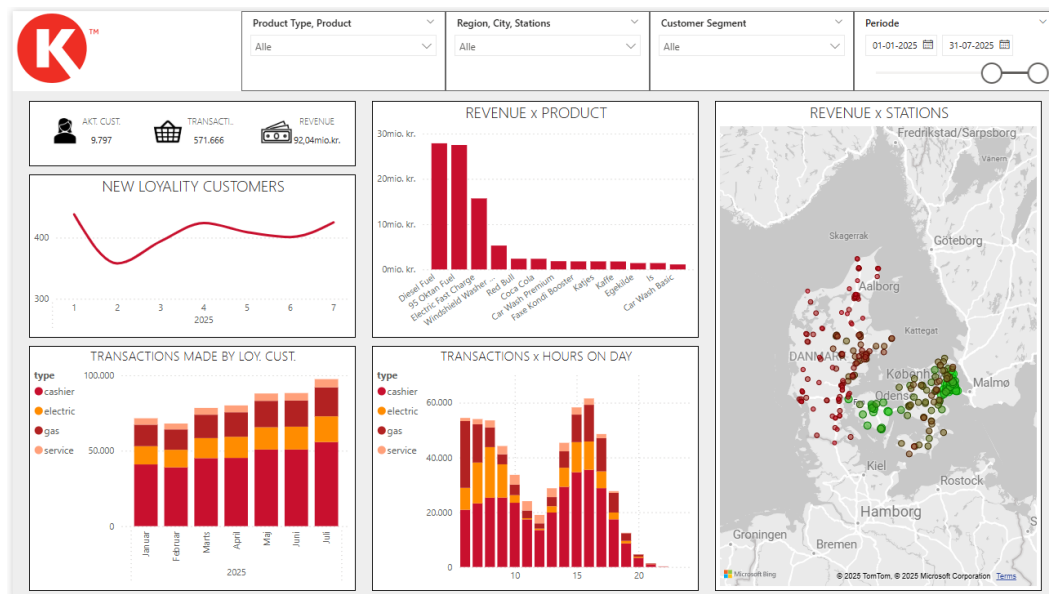
- **LOYALTY\_CUSTOMERS:** Represents unique loyalty customer.
- **SEGMENTATIONSGROUPS:** Defines customer behavioral profiles (e.g. EV Commuter, Heavy Shopper) used for simulation and analysis.
- **CARDS:** Each customer can have multiple loyalty-linked cards
- **TRANSACTIONS:** A transaction links a card and a cashier.
- **TRANSACTIONSLINES:** Contains detailed line items per transaction, including price, quantity, product, and optional campaign linkage
- **PRODUCTS:** Master data for all purchasable items (fuel, beverages, services), including pricing and type
- **CAMPAIGNS:** Modeled after real Circle K promotions (e.g. buy 5 get 1), linked to product and timeframe
- **CAMPAIGN\_TRANSACTIONS:** Tracks when a customer qualifies for and redeems a campaign reward
- **CASHIERS:** Virtual representations of POS by type (e.g. gas, electric, service) at specific stations
- **STATIONS:** Station-level metadata including region, location, and operational period



## Power BI Report

The Power BI dashboard was structured around three analytical pages — **Overview**, **Segmentation**, and **Campaigns** — each designed to answer specific business questions aligned with Circle K's loyalty strategy.

### Page 1: Overview

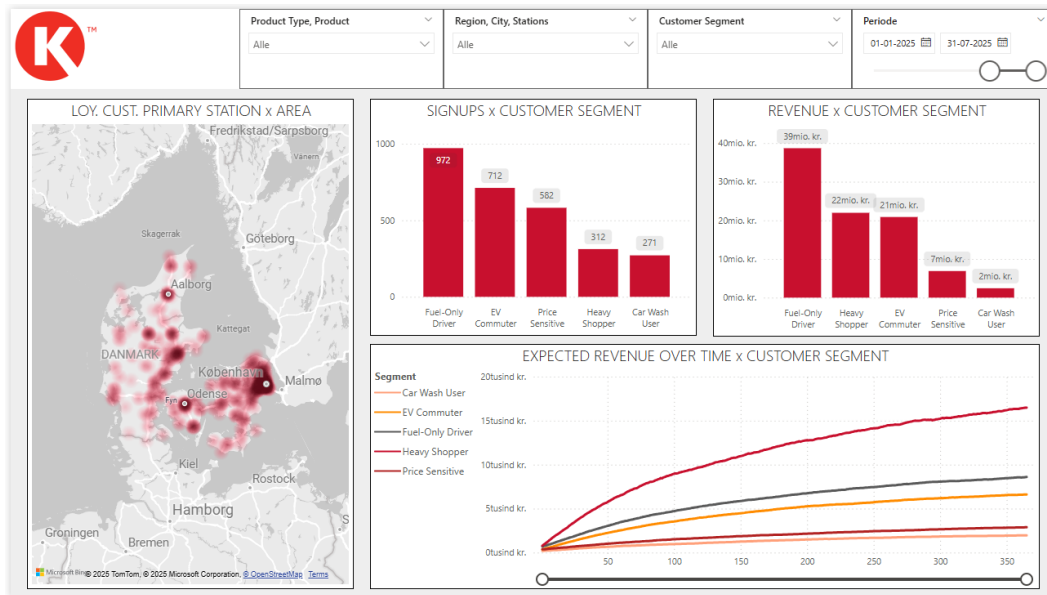


Focused on volume, revenue, and temporal activity patterns.

#### Key questions answered:

- **What is the total transactional footprint of the loyalty program?**
- **Which product types and products generate the most revenue?**
- **How do transactions vary by time of day and customer type?**
- **Where are the most active stations geographically located?**

## Page 2: Segmentation



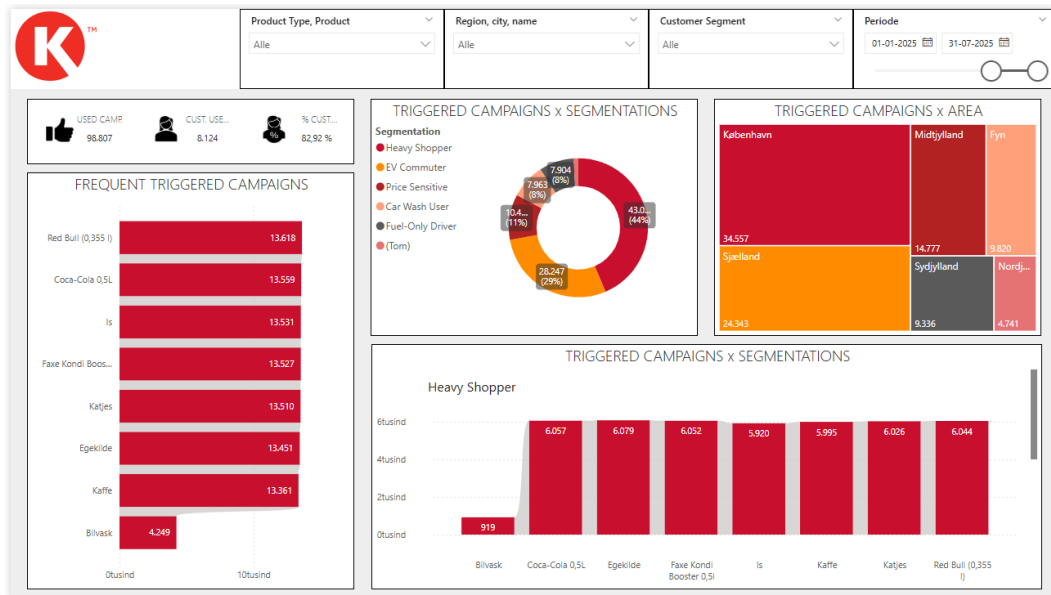
Explored behavioral and value differences across customer groups.

### Key questions answered:

- Which customer segments sign up most frequently?
- How does revenue differ by segment?
- Which segment delivers the highest long-term value?
- How geographically distributed are customers by primary station?



## Page 3: Campaigns



Analyzed reward redemptions and campaign performance.

### Key questions answered:

- Which campaigns are triggered most often?
- Which segments engage most actively with campaigns?
- How does campaign performance vary by region?
- What product types most frequently drive redemptions?

Each question was supported by relevant metrics, visuals, and filters — enabling dynamic exploration by **time period**, **region**, **station**, and **customer segment**.

This structure enabled the simulation to not only showcase data modeling capabilities but also extract **strategic insights** applicable to real-world loyalty optimization and personalization efforts.

## Summary Outcome

This project reflects my effort to independently design and implement a complete analytics pipeline — from data modeling and simulation to dashboarding and insight generation.

Without access to internal data or systems, I aimed to create a realistic loyalty simulation inspired by Circle K's business context. While simplified, the result is a functional and testable model that supports analysis of customer segments, transactions, campaign redemptions, and lifetime value.

The structure and visuals are built to explore meaningful business questions and demonstrate the kind of analytical thinking I hope to contribute in a Senior Data & Analytics role.

## Other relevant files

- CV
- Application
- Case Summary (this document)
- Case Questions & Answers (separate document highlighting key business insights from the simulation)
- Github project (<https://github.com/AbrahamBorgiPrivat/Circle-K>) containing the full simulation project, including:
  - All Python scripts for data generation, upserts, and orchestration
  - PostgreSQL schema and sample data structure
  - External data ingestion (DAR/DAGI)
  - Power BI theme and PBIP file
  - README files for setup and project context

These materials are intended to give a clear view of how I approach data problems from end to end — from technical modeling to business-facing insight.