

ETL Sprint 1 Project Case Study

‘Grocery & More’

The Problem Statement

An online grocery chain 'Grocery & More' has multiple stores in different cities in India. Customers can either buy directly from the store or order their products online. 'Grocery & More' procures the products from different vendors.

'Grocery & More' wants to analyze their profit & margins for each store from data available for the last 3 years. They also want to analyze customer buying trends.

To get that insight, 'Grocery & More' wants to create a Data Warehouse solution on which they can do data analytics.

Recommended Project Implementation Steps

- Construct an ER-Diagram for the above-mentioned requirement
- Construct Tables as per the ER-Diagram
- Identify the relationships between tables and use appropriate standards where applicable
- Insert the appropriate data into the identified tables
- Configure the source and target connections (Data Integration)
- Implement Rank & Router transformations in single pipeline using ETL
- Implement Filter & Aggregation transformations using ETL
- Implement Flat file transferring to target using Aggregation
- Create different sources & Targets integrate with ETL

Task outcomes

As part of the solution design, you are required to come up with the following outcomes:

- Design ER Model
- Sample Data
- Table Scripts
- Informatica ETL Code

Technology / Software

- Oracle 11g
- Informatica Power Center v10.x
- Device configuration – Min 8GB RAM, Min 100GB Disk, Windows 10

Project Document outcomes

As part of the project submission, following documents needs to be provided:

- Informatica Project Export (XML file)
- ER Model, Data scripts and Sample data
- Unit Testing document with outcomes
- Integration Testing document with outcomes
- Regression Testing document with outcomes
- Project presentation (preferably in a PowerPoint format)

Project evaluation parameters

The project solution would be evaluated on the following parameters

- Conceptual design and implementation
- URS / System Requirement Understanding
- Unit Testing
- Exception Handling / Data Validation
- Coding Standards (Naming Conventions, Comments, and Indentation)
- Source Code Optimization
- Ability to have handle complexity / changes
- Code Reusability
- Timeliness of Solution design and completion as per requirements
- Working demonstration of the solution, as per requirements
- Presentation skills during the Viva
- Communication & Coordination during the project