

End-to-End Big-Data Cloud-Computing Solution for Ecuadorian Grocery Retailer using AWS

Maroon Team 3

Current State



Ecuadorian grocery retailer Corporación Favorita is growing due to which the retailer is facing a huge data management and processing challenge.



Future State



Organizational transition from offline database to cloud based database platform supported by advanced analytical dashboard.



Data

- POS data containing transaction level information (4GB)
- Store and product information

Methodology

- Day level data from retailer's database is processed and transformed using **Hive** and stored in **S3**.
- Transformed data is appended to historical data using **PySpark**.
- Conducted Machine learning using **Mlib** on **Pyspark**
- Sales forecast for 6 months is done using **Sparklyr**.
- Results are visualized using **QuickSight**.

