Predicting and tracking the sold quantity (in t) of plastic packaging materials

Ersan Kucukoglu

Background Summary

The retailer company is desperate to lower the negative effect on the environment. In order to do that the business goal is to lower the sold plastic packaging with our product by 20% till 2025. By developing a data product which calculates the KPI of sold plastic packaging by the HU-contracted suppliers in tones and compare it to the goals an instant intervention would be possible in case of a risk.

Key problems:

- evaluating the business goals is done quarterly
- there is only a high effort possibility to follow the measures within the year, no forecast is available

Objectives:

- building a data product, which allows the user to see the sold plastic packaging in tones and to compare it with the business goals also by seeing the predicted outcome till end of business year

Solution type:

- dashboarding + prediction

What questions do I need to answer?

- What is the sold plastic packaging quantity in tones?
- What is the estimation of sold plastic packaging till the end of business year?
- Which items are significantly high and should be optimized (packaging change) based on the past / predicted values?

Available data:

- Receipt data (long term history)

- Store main data, item main data

Expected output:

- Dashboard (incl. automatic update) with option for import main data file

Methods of approach

Data-science methods to be used:

- ➔ Model based prediction, data-cleaning, dashboarding. (Sales and goods-in based.)
- ➔ Receipt data analysis
- ➔ Predicting yearly outcome
- ➔ Importing main item data
- ➔ Dashboarding, creating visualization

Software:

- ➔ Databricks platform
- ➔ Databricks SQL-Dashboard

Outcomes

A dashboard with the following functions:

- allows the user to see the sold plastic packaging quantity in tones
- allows the user to compare it to the business goals
- uses two calculation method: sales and goods in based
- sales figures can be loaded from database, goods in should be imported manually (csv)
- give estimation till the end of business year for the run-out
- marks top items, which past / predicted values are significantly high and should be optimized (recipe change)

Notebooks:

- possible to re-run every month
- possible upload (csv) for:
  - -main item data
  - -targeted values