SMARTER DECISIONS, **DELIVERED.**

Case Study

Data Quality Checks – Use the data set and rules set attached

31 July 2024

Case Study:

Create a logic module that takes an excel as input and performs below quality checks and send a notification with consolidated data quality issues

Dataset

The dataset for this case study will be shared separately over email.

Analysis Requirements

- Certain sheets should be present in the workbook.
- All column names should be checked and there shouldn't be any special characters, spaces in the column names.
- Predefined datatype for each column should be maintained to check if that column contains data that respects the datatype defined.
- All date columns should contain the same date formats
- Whether the columns with non-nullable status have any null
- Make sure that Your logic consolidates the quality issues identified in a given file and send an alert to a mailbox

Cleaning the data – Use the same dataset that is created as part of above step

- Convert the columns to appropriate data types by checking the data in it
- Cleanse the data so that datatype conversion should not fail (removing leading or trailing spaces, making sure all entries in a date column are following same date format YYYYMMDD, removing special characters from column names, removing commas from cost/money columns)

Audience

The audience for the presentation is the Director of Sales and the Sales Operations Manager, along with their team members. The purpose of the meeting is to review the analysis and recommendations in detail and discuss implementation plans.

Audience