

Step-by-Step Process for Associations & Aggregated Job Titles Automation

Introduction

Interloop receives two .csv files for each run to their SFTP website. Interloop's automation is triggered at 9pm nightly. If Interloop detects both files are uploaded to their site and both files follow the proper naming convention, then the automation is triggered. Below is a high-level overview of the steps taken to get the original uploaded files to the final output table.

Step-by-Step Process

Step 1: Receiving the Excel Files

Upon receiving the Excel files, the schemas of the files are verified and errors out if they are not the expected schema.

Step 2: Initial Data Cleaning

Begin with basic data cleaning. This includes:

- Removing any empty rows or columns
- Standardizing field names (snake case)

This initial step helps ensure that the data is ready for deeper analysis and transformation.

Step 3: Update HubSpot Data

Update HubSpot data in the data lake. This includes:

- Extracting the latest HubSpot changes.
- Cleaning the latest HubSpot changes.

This initial step helps ensure that the data is ready for deeper analysis and transformation.

Step 4: Data Transformation

Transform the data as needed to fit the requirements of the final table. This includes:

1. Define job titles based on rules set by Mythic. These original fields include contact_role and contact_title).
2. Combine both ICRM and DCRM tables into one table.
3. Filter table to contain only active HubSpot contacts.
4. Associations
 - a. Pre-filter HubSpot Companies.
 - b. Select contacts that have a HubSpot company match.
 - c. Select only association labels relevant for contact-company.
 - d. Create associations table.
5. Aggregated Job Titles
 - a. Split contact_role and dedupe so that there is one row for each individual email/job_title.
 - b. Split contact_title and dedupe so that there is one row for each individual email/job_title.
 - c. Union to make one table as well as remove special character and nulls.
 - d. Group by email.
 - e. Pre-filter HubSpot contacts.
 - f. Create table joining pre-filtered HubSpot contacts and transformed aggregated job titles.

These transformations convert the organized data into a table ready for updates in HubSpot.

Conclusion

By following these steps, the raw data from the Excel files is polished to a table that accurately and effectively presents the necessary information. This systematic approach ensures consistency and accuracy in the data transformation process.