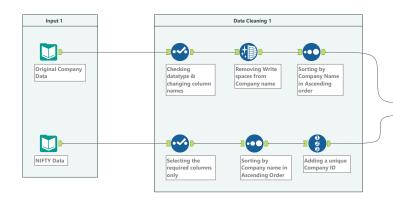
Project 1: Financial Analytics
Work flow 1: Data Cleaning
Dated: 13-05-2024
Created by: Arpita Deb
Methodologies used: Data Cleansing, Sorting, Joining, Filtering for duplicates, Union
Result: A cleaned data set with 5 columns - CompanyID, Company Name, Market Capitalization (in Crore),
Quarterly Sales (in Crore) and Industry



This output data set contains no industry data either because the company names were written differently than the NIFTY data set or there are no records in NIFTY data set. So I need to manually enter the industries and then join this data with the original data. Data Cleaning 2 Output 2 Output 1 Input 2 -100 Companies with industry after manually entering the industries **-**Changing Cleaned Finance Sorting the Adding a unique Company with no industries Company with industries Analytics data output industry case results by company and market

Final dataset with

industry column

Filtering to

duplicate entry

remove a

- Companies with industry column from JOIN

Ø.

Joining to add

industry column

capitalization