

This document outlines a detailed workflow for data processing and analysis, likely related to warranty claims or product quality, defining various data schemas and providing step-by-step instructions for eleven key processes including raw data and VIN processing, product description tagging, L1 and L2 layer tagging, root cause tagging and quality assurance, TSB validation, SQIDS root cause mapping, DTC mapping, free text translation, share rate calculation, and parts return mapping.

This also highlights the list of schema where each file needs to be uploaded for refresh

Schema List

Raw preprocessed data – Raw_V12

VIN – VIN_extract

Raw TSB- under tsb free text

tsb pdf link - tsb pdf link

Added_share_rate_v1 - tagged

RC & SC - Root and supplier comment – Tagged data


8646C - Product Description Output – Product Description – Tagged data

Share rate – Share Rate Updated – Tagged Data

Translation – All Translated – Raw Data

Process:

1. Raw Data and VIN Processing:

1. Download the DOMO/ALL Claims data that is given by PL
2. Create a new folder in format “YYYYMMDD - All Claims DOMO” in common drive folder -
Link:  08. Warranty Chargeback
3. Upload the downloaded csv to newly created folder
4. Use this file as the input to the script available in this folder:
https://6af301396b18f0c3-dot-us-west1.notebooks.googleusercontent.com/lab/tree/02%20PO%20Folders/Praneeth/Raw_VIN_Refresh_Codes
and make sure to make no changes to the Master_Records file that is available in this folder
5. Make sure to run the VIN script in the offline notebook as it needs connection with SSMS
6. Once code runs successfully, VIN and Raw Data Refresh files will be ready to upload to the app
7. Schema to be used for these two files are:
 - a. For Raw refresh : Raw_V12
 - b. For VIN processing refresh : VIN_extract

2. Product Description Tagging:

1. Use the respective bakery models that are available for each commodity to find the main problem based on the free text

3. L1 Layer Tagging:

1. For each commodity get the list of L1 problem and L1 condition for the respective commodity
2. Make sure to use only the tagging that is already present in the app for each commodity don't tag new categories
3. L1 Problem is tagged using failure mode and L1 Condition using symptom mode
4. If symptom and failure modes seem ambiguous try to use the free text for tagging the L1 layers

4. Root Cause Tagging and QA:

1. Use the respective bakery models, for each commodity
2. For some models we are using tagging as Root Cause A, Root Cause B,... just make sure to map it to the correct tags
3. After getting the backfill of RC tagging, it is important to do 100% manual QA of these tags as this is the important parameter to do the share rate calculation

5. L2 Layer Tagging: (L2 Problem and L2 Condition)

1. This tagging or buckets is something which we have received from the client
2. Client has given only the buckets or tags for only few commodities, just tag only those commodities and same can be known based on the model that is available in bakeray

6. TSB Validation:

1. First export data from app for current commodities, then filter TSB Extracted column and check unique TSB for this week data
2. Compare it with TSb Mapping – Master sheet new TSB file, to check if any new TSB codes are there to process, if so download all those pdfs, only from NHTSA website
3. Change the TSB's to standard format of ['T-SB-XXXX-XX'] in the app data
4. Create a new folder in Pradheep folder and upload downloaded pdf's this folder name goes as input to "TSB_PDF Content extraction-2.ipynb" code
5. We get this kind of file "data_tsb_extracted_22_aug.csv", paste the contents from this into TSb Mapping – Master sheet new
6. Go to prashanth folder – denso - root_cause_112124 – Ask for api key and run "02 TSB_extraction.ipynb" - we get "TSB_extracted_PDFs_22_aug.csv" file
7. Give this output csv file to "03 TSB_Cleaning.ipynb" file , we get this csv as output "TSB_extracted_PDFs_22aug_Cleaned.csv"
8. Copy all this content and paste it into TSb Mapping – Master sheet new file, download this file now and send this as input to "04 TSB_Validation.ipynb" and also give app data export(Change the TSB's to standard format of ['T-SB-XXXX-XX'] in the app data) for this week as well to this script, will get output csv file as "Output_Final__22_aug.csv"
9. Pass this file as input to the "TSB_Output_Processing.ipynb" script and get final TSB mapping sheet – paste it into your current week or combined processing sheet

7. SQIDS Root Cause Mapping:

1. From bakeray download the backfill of warranty chargeback model, from this copy this data to new sheet called "Tagged" and copy remaining columns from previous week sheet
2. Replaced parts we will get from running app data and uploading to replaced parts.ipynb script, which is available in 89467 folder
3. Root cause column you need to populate based on the RC_SC sheet we filled
4. In the SA Folder – denso - 15. Denso – 89467 folder upload this week's data to untitled folder 1 – but before uploading check the correct positions of TWC_No, Diagnosis codes and Replaced parts code, if not move them to correct position by checking combined output file earlier got downloaded
5. DTC code in TSB validation sheet is the DTC [TSB] column to be mapped in Tagged sheet
6. In the same folder, is "SQIDS.ipynb" code,
7. TSB consolidated column we push as empty list
8. L2 root cause push the TSB numbers only, otherwise blank
9. If some TSB is mentioned but not available online then L2 Root cause for that record should be "TSB Mentioned but not found online" and TSB reason will also be "TSB Mentioned but not found online"
10. If there are multiple TSB's mentioned for that event, then TSB reason will be "Pending Feedback - Multiple TSBs" and the same to be mentioned for L2 Root Cause
11. Run the Replaced parts.ipynb file that is available in the same folder, for this we just need to give input as "All Claims" data, this will give the replaced parts list and that can be mapped to replaced parts[Claims] column in the tagging sheet

8. DTC Mapping:

1. Based on the DTC code, we will give its short and full description of the DTC codes
2. Download the data where axion short description is null
3. Keep only TWC_No, DTC Code [AX], DTC Code [SQIDS] [AX]
4. Upload it to DTC Short Description Process.ipynb code
5. Rename Combined codes and description to DTC Code [Claims+SQIDS] and Axion Short Description and TWC_No, remove remaining
6. Upload to schema DTC Short + Full Description

9. Free text Translation:

1. From this week's **raw data** – keep TWC_No and CCR columns
2. Using detect try to determine the language of columns
3. Filter out non English and check whether it detected properly or not
4. Convert them to English using googletranslate formula
5. Keep only TWC_No, Condition Translated, Cause Translated, Remedy Translated
6. Let other records be empty only

10. Share Rate Calculation:

1. Ask for export of whole app data with commodity,twc(eventid),condition(denso),problem(denso),rootcause(denso)(1) and rootcause(denso)(2), Parts Returned – all are custom attributes
2. Append them with current week data

3. While appending make sure to add it in non list format
4. Merge root cause 1 and root cause 2 in list format
5. Upload this file to share rate calculation(All commodities) code
6. Commodity wise summation should come out to be ~100%
7. Change the U,V,W columns to Supplier [IxL], Toyota Mfg / Sales [JxL], Toyota Design [KxL]
8. Change U,V,W cols to number type and keep three decimal places
9. Rename remaining cols this way:
Supplier % - Supplier
Toyota Mfg / Sales % - Toyota Mfg / Sales
Toyota Design % - Toyota Design
10. Ensure that for all commodities the summation of share rate should be approximately 100%, if not there is some error in the processing
11. Keep TWC_No and remove all other and push to Share Rate Updated(Tagged Data type) Schema

11. Parts Return Mapping:

1. In this folder "16. Denso - 861C0" folder – use code "part_returns_mapping.ipynb"
2. Download app data for this week or for which ever commodities parts return file has been sent and upload to the above script, also upload the parts return file we got
3. Ensure to rename OFP number in parts return file to 'Customer_Product_No' and VIN to 'vin'