**Process Documentation**

**Alteryx Workflows**

**Overview**

* A group of black text

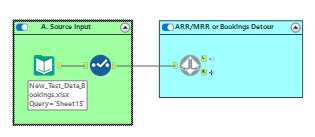
  Description automatically generatedWorkflows are built leveraging a combination of Desktop Alteryx and InDB Alteryx, powered by Snowflake.
  + **Step 1 – InDB Tech Enabler** leverages desktop tools and converts the deal/project data to the preset template format.
  + **InDB portion of workflow (Step 2 and 3 – InDB Tech Enabler)** calculates retention and formats outputs for the Databook.
  + Outputs are then stored on Snowflake and pulled into PBI and Excel using an ODBC connection.
* All 3 workflows have built-in logic to adjust the calculations based on project needs. Additionally, workflow structure allows for ad hoc updates to seamlessly process for specific situations that may not be available in the current version.
* The In-DB workflows incorporate two sets of outputs for PBI and Excel:
  + Step 2 – InDB Tech Enabler outputs to the PBI dashboard which includes Fact table and Retention table at three levels.
  + Step 3 – InDB Tech Enabler outputs to the Databook

**NOTE:** All steps of the workflow require an existing Snowflake database for the project, as well as designated ODBC connections. If you do not have these, please refer to the Snowflake & ODBC Process Set-Up guide.

**Step 1 – InDB Tech Enabler – Data Cleanse**

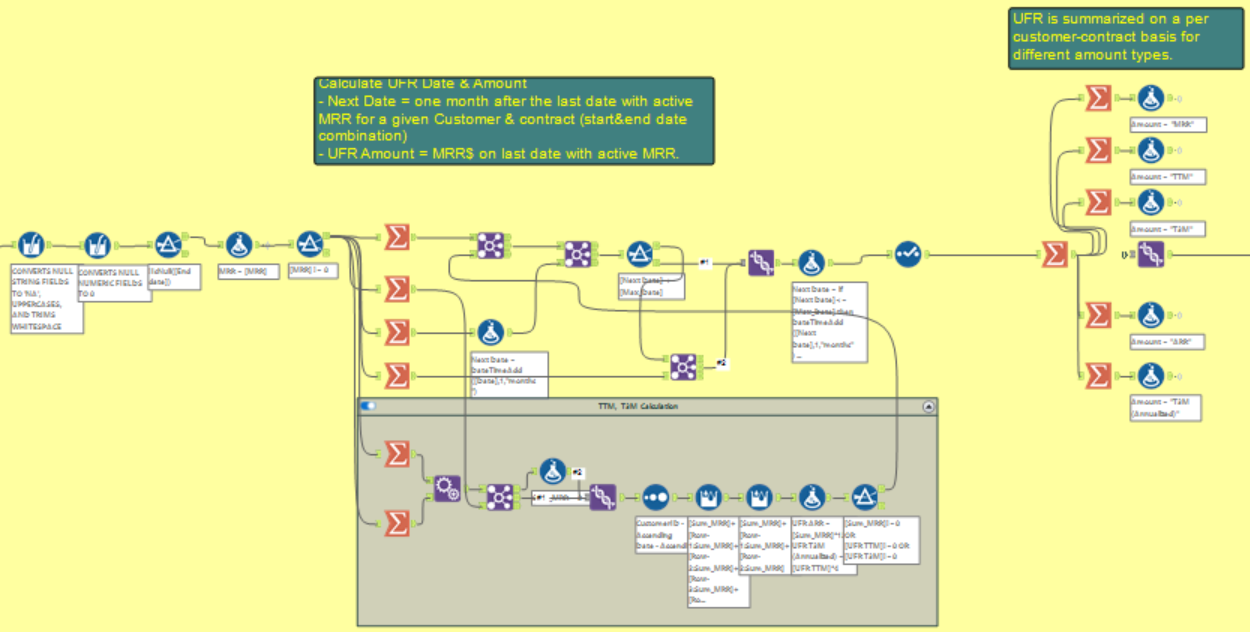
**NOTE:**

* It is recommended a user’s ETL process occurs prior to Step 1 – InDB Tech Enabler
* For performance optimization, it is recommended user utilizes the YXDB (Alteryx Database) file type as the input.
* Step 1 – InDB Tech Enabler allows for either bookings or revenue data. User may toggle the Detour tool depending on data format.
* Match the existing fields in data to the preset fields existing in the template. For example, although your field may not exactly be a product or region, you may still tag your data as such to minimize changes later in the workflows. These are the following relationships allowed per field:
  + **Product, Revenue Type** – M:1 - The process will calculate cross-sell and cohorts at both levels if input. In situations where there is no Product or Revenue Type data, the process will tag them as “NA.”
  + **Region** – 1:1 - These are considered attributes and will only work correctly if they are unique by customer. Step 1 is set up to force the relationship to 1:1 for existing fields. If the project data has other 1:1 that do not fit the description, it is possible to rename the data to the existing fields, and the relabel it in the outputs.
  + **Miscellaneous -** M:1 - It’s recommended this flow through concatenated with customer to make the process easiest. These may require more complex treatment depending on the project needs, including adjustments to customer counts, existing DAX measures, etc.
* Step 1 also creates the necessary tables for Renewals analysis. If your project is not using bookings data, feel free to close the relevant containers to improve performance.

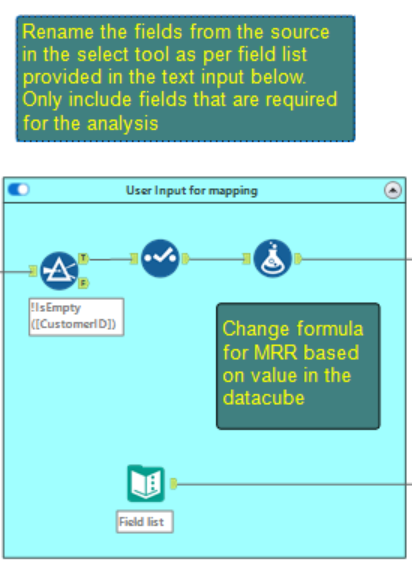
**Step 1 Workflow Summary:**

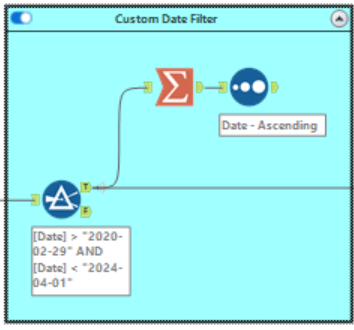
1. A table with numbers and letters

   Description automatically generatedTo get started, input your data into the workflow into container “A. Source Input”. It is recommended any major clean up occurs prior to this step in a separate workflow. We also recommend your input is in the YXDB format to improve performance. See right for a size comparison between CSV vs. YXDB files.
2. Light blue container “ARR/MRR or Bookings detour” allows for you to select the incoming data format. Click on the detour tool, and check the box if your data is bookings. Otherwise leave blank.
3. Bookings – follow the right output from Detour tool 834
   1. Container “**Billings Starting Point**” identifies and calculates bookings start and end dates, and filters records with invalid/null dates.
   2. Container “**Waterfall if source input is (Bookings/Billings Cube)**” calculates MRR basedd on the TCV and dates calculated in the prior container.
      1. **NOTE:** MRR is calculated based on years having 365 days. If your project requires a different approach, modify the MRR calculation in Formula 522. The container spits between two data streams: retention and renewals.
   3. Containers “**UFR Tagging**” calculate when contracts are up for renewal at all three levels (Customer, Customer-Product, Customer-Product-RevenueType). The Up for Renewal Date (UFR Date) is calculated as one month after the last date with active MRR for a givien customer (or cust-prod, cust-prod-revtype). The UFR amount is the MRR amout on the last date with active MRR
   4. Container 943 has all renewal table outputs, which are later mapped into the core retention calculation in Step 2. If you are not using bookings data/do not need renewal calculations, please deactivate this container.

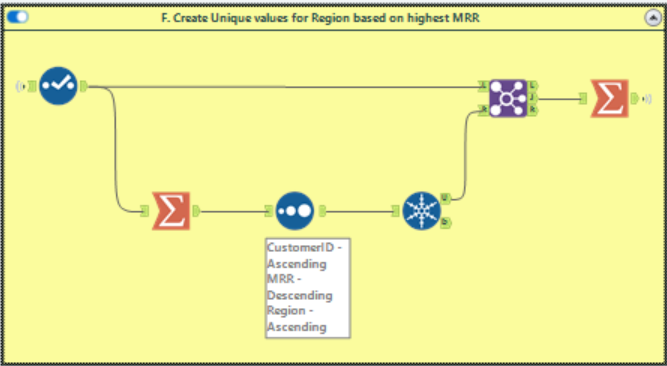
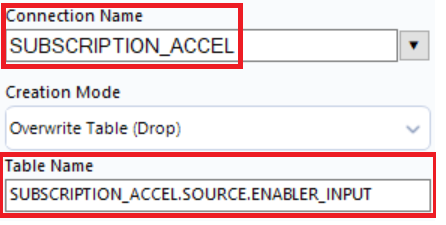


1. Container “**User input for mapping**” is the beginning of the retention calculation, regardless of the type of input (ARR/MRR or Bookings). In Select tool 826, you must map the fields in your projects’ file to the template fields show in Text Input 367. Attributes must be 1:1 to customer, while Product and RevenueType can be M:1, and will have cross-sell and product specific calculations.
   1. **NOTE**: These fields do not need to be conceptually the same in your data, the most important part is that the relationships (1:1, M:1) align



1. Container “**Custom Date Filter**” allows you to limit the incoming data set to periods relevant to your analysis. If not necessary, feel free to remove it.
2. In container “**Account Size Addition**”, the workflow segments your data based the individual customers size based on two different methodologies:
   1. **Max MRR Bucket** – Calculates the maximum MRR at the customer level for all periods in the dataset. Formula tool 819 calculates the segments and divides it by Account Size – modify this tool if needed to adjust for your project dataset.
   2. **Quartile Base** – Segments the customers based on what quartile each customer is based on average monthly MRR. “1 Quartile” is the highest, while “4 Quartile" is the lowest.
3. The following section, “**Create Unique Values**”, forces attributes to have a 1:1 relationship with customer. Container J1 “Duplicate customer attribute test” identifies non 1:1 relationship through Test tool 503. Any M:1 relationship will be forced to be 1:1 based on MRR by attribute.
4. Lastly, the workflow outputs the data in the template format to table “**Enabler Input**”, under your project database. Replace the template output with your project database and ODBC connection.

Replace the Connection name and Table Name with your projects database and ODBC connection.

****

**Step 2 – InDB Tech Enabler – Retention Calcs**

**NOTE:**

* This step produces the bulk of the outputs and creates all the necessary tables required to refresh the PowerBI template.
* The output of Step 1 is the main input for Step 2. If you are calculating renewal rates, Step 2 also uses the respective UFR tables for the 3 levels.
  + **NOTE:** This step is all built leveraging InDB tools. If you need to adjust, you will need to do so using Snowflake syntax. (E.g. if you are adding an “IF” statement, you will need to do so using the “CASE” SQL function)
* It is recommended you turn off containers not relevant to your project. For example, if you do not need Quarter comparisons, or TTM calculated, turn those containers off so your workflow runs more efficiently.
* The temporary tables seen throughout improve performance. Add more if you have a large dataset.

A screenshot of a computer

Description automatically generated**Step 2 Workflow Summary**

* To get started, input the “Enabler Input” table into container “**A. Starting Point**”.
* The revenue calculations container has two different functions:
  + Ensures all periods of data are available for all customer, customer-product, customer-product-revenue type combinations.
  + Calculates the different variations of revenue – ARR, MRR, TTM, T3M
* Containers **A.2.2 through A.2.4** calculates the relevant cohort dates through levels.
* Container “**Add Boomerang Flag**” calculates the boomerang flag at the customer level. It is preset to a one-month lookback, but may adjusted by altering the LAG function in Formula 471
* Container “**Full Dataset**” calculates fields to be used in the Fact Table as part of the PowerBI logic.
  + NOTE: This is completed at this point to optimize performance vs. doing it on PBI later
* **(!!!)** The “**Fact Table**” is the first output of this workflow and is very important for the PBI dashboard as most pieces of data come from it.
* Container “**Cohorting**” calculates additional logic for cohorts at all relevant levels.
* **A screenshot of a computer

  Description automatically generated**At this point, the workflow divides into the three distinct levels (customer, customer-product, customer-product-revtype). It is recommended you close/delete the containers to the sections not relevant to your project, as they will make the workflow take longer.
  + Each section follows a similar structure calculating for the following retention categories - upsell, downsell, churn, etc. - at each respective level and pulls from the UFR tables generated in Step 1 for the renewal schedules.
  + Container “**Credit Treatment**” splits out negative revenue record and separates them from the rest into their own retention category of “Net Credit.”
  + Container “**Winback Tag**” records with 0 revenue for customers that present revenue in periods after
  + A screenshot of a computer

    Description automatically generatedContainer “**YoY retention**” calculates the retention categories with three different comparison levels (YoY or Year, QoQ or Quarter, and MoM or Month). These use function LAG to pull the revenue for the same customer combo 12, 3 or 1 months before
    - If you wish to change how a retention category is calculated, you will need to update it in all relevant levels with their respective amounts.
  + The UFR Tagging containers after calculating the categories relevant for the Renewals schedules at all 3 levels. Turn them off if you are not using bookings data.
  + Outputs for this workflow are pushed into schema “**OUTPUT**” with the full data. These are now the inputs for PBI.

**Step 3 – InDB Tech Enabler - Databook**

**NOTE:**

* This workflow converts the Step 2 outputs into the lighter format required for the Excel Databook
* You may customize what data you want to bring into the databook.
  + For example, you may choose to only bring March, June, September and December data instead of every month.

**Step 3 Workflow Summary**

* Table **Date\_Dim\_Databook** brings in the periods for the analysis and is used in all pivot table schedules (E.g. Top Movers, Rollforwards) to maintain structure.
* Table **Dimension\_Dim\_Databook** is used to carry all 1:1 field by customer. These are used to slice views throughout the file.
* **Section B** brings in the three retention tables, labeled accordingly. In instances where you do not need any given level, deactivate/delete the logic associated with that level.

**A screenshot of a computer

Description automatically generated**

* **A computer screen shot of a diagram

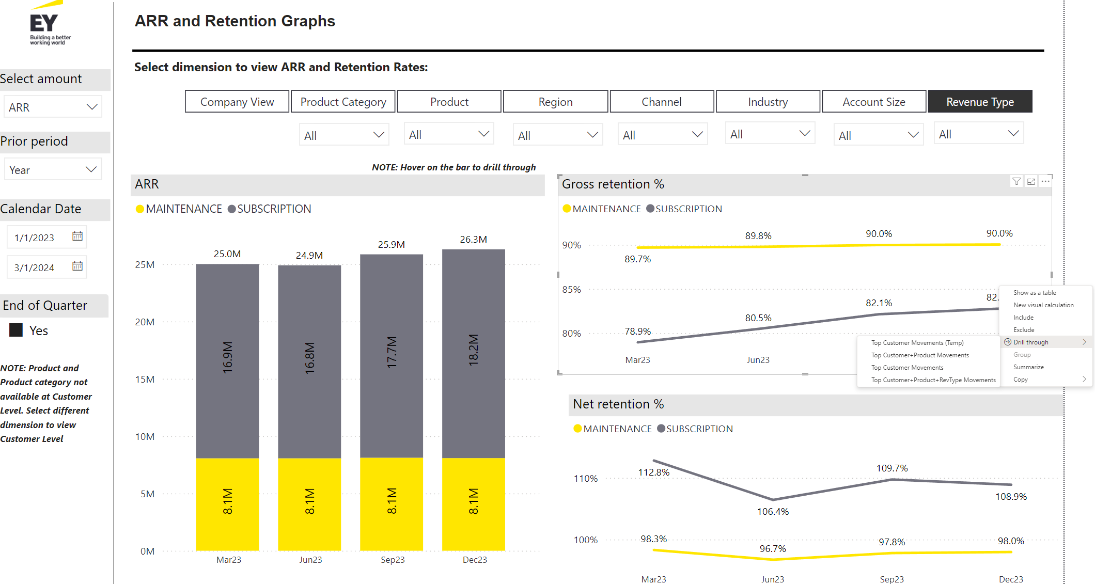
  Description automatically generatedSection C** brings in the Product\_Analysis\_Databook table, which is used in product specific schedules such as bundling or cross sell. We recommend bringing in any custom analysis through this table when possible.
* After the retention table outputs, the workflow produces the Top Mover tables based on the number of Top Movers you want. This is materially more efficient that loading a full customer detail pivot.

**PowerBI**

* To refresh the PBI, input your project database into the “Snowflake Project Name” parameter. All inputs/measures should automatically populate.
* The PBI is useful for both gathering insights and questions for management, as well as to present the results of the analysis.
* Most calculations are measure driven, meaning they can be modified easily to fit each project’s nuances.
* The template has a large number of different views/dashboards that may be deleted I not relevant to your specific project.

**PowerBI Walkthrough**

* You may input your data by changing the “Snowflake Project Name” parameter to your project’s Snowflake database name. Note you will need to have the ODBC connection properly set up for it to work. For further details please review the connection documentation file
* The connection pulls data from 7 tables, all generated in Step 2 of the workflows:
  + Date\_Dim: Summary of all dates in the analysis. Each record is a date. This is connected to the Fact\_Table and related calculations.
  + Customer\_Dim: Identifies all attributes and products per customer, with no date.
  + Fact\_Table: Identifies all attributes and products per customer, as well as all different revenue amounts (TTM, ARR, MRR, etc.).
  + Cust\_Retention: Customer level retention calculations.
  + Cust\_Prod\_Retention: Customer-Product level retention calculations.
  + Cust\_Prod\_RevType\_Retention: Customer-Product level retention calculations
  + Retention\_Date\_Dim: Summary of all dates in the analysis. Each record is a date. This is connected to the Fact\_Table and related calculations.
* All schedules are measure driven to an extent – can be adjusted as needed.
* Use Case – creating questions for an agenda:
  + You are curious why the gross retention for Revenue Type “Subscription” is improving.
  + The template can easily show you the main drivers behind the improvement.
    - To do so, right click on the metric and period you care about 🡪 click on the GRR line for Subscription, click drill through and go to Top Customer+Product+RevType Movements

****

* + The dashboard will show you all the top moving customer, as well as their respective ARR change for that period (you may change the Amount and Periods being used)

**A screenshot of a computer

Description automatically generated**

* + Dig in deeper by clicking on drill through again and reviewing the customer details.
  + Repeat this through attributes/products to develop your agenda faster.

**Excel Databook**

* The current template has two separate versions of the databook:
  + Lite – This version includes Customer-Product and Customer level views, and is a good fit for introductory analysis, or projects where there is simple/few special data considerations. If it is your first time using the template, use this version to get familiar with the structure.
  + Advance – This version includes all views for all levels of analysis, as well as ODBC connections to all three retention tables. Use this version if you need specific schedules/levels not available in the Lite version. Please review the “Advanced Databook Best Practices” section below to understand how to optimize the file.
* Available views/analysis calculated in the
  + Retention at 3 levels (Customer, Customer-Product, Customer-Product-RevenueType)
  + Renewals at 3 levels (Customer, Customer-Product, Customer-Product-RevenueType)
  + Cohort Analysis
  + Customer Concentration summary
  + Top Movers Analysis
  + Cross Sell
  + Product Bundle
  + Product Penetration
* All levels are fully disconnected and have their own measures and slicers. That allows users to delete entire levels when they are not needed.
* The databook is connected through the ODBC connection set up before Step 1

**Excel Databook Guide**

* A black text on a white background

  Description automatically generatedTo get started, replace the template database with your own. Please review the Snowflake/ODBC connection guide for details
* Go to the Index sheet, and input your analysis start and end dates as well as the periods you are including. This will dynamically update all databook schedules to use your dates.
* The template will come with 36 or 60 months per schedule. Delete all columns that are not in use.
* If your data has more than 60 periods, it’s recommended you review all schedules to confirm there is enough columns for the data to flow through
* Some of the new schedules highlighted below:
  + KPI Matrix
    - Each individual attribute/product section is pulled from individual pivot tables organized vertically by measure.
    - To add a new attribute section, create a new sheet and copy the two pivot tables you see in any KPI matrix pivot sheet. Make sure columns and rows are the same as in the original. In the KPI Matrix tab, copy and past any of the existing sections, and using Ctrl + F and Replace, change the sheet reference to reference your newly create sheet.

Each section is pulling from its own pivot. Easy to copy and paste/rename to fit your deal.

* + - A screenshot of a graph

      Description automatically generatedTo add a new KPI, create it as a measure in PowerPivot, and bring into each KPI Matrix pivot. Then, copy the columns of another KPI, and replace the measure being called for all attributes/products.

Amounts may be measure driven (EoP$ and BoP$ aside). Adjust as needed.

* + Top Movers
    - Schedules come from separate Snowflake tables that improve performance materially, especially for large customer datasets

A white and black chart with black and yellow text

Description automatically generated with medium confidence

* + - A screenshot of a computer

      Description automatically generatedSelect a retention category and a period combination using the slicers. The Change Section will reflect the- Top N of the category you selected for that period.
    - The ARR section will reflect ARR per period. Use the slicer to change your revenue figure (MRR, TTM, etc.)

**Databook Best Practices:**

* All dates are formula driven. Before updating the file with your data, go to the Index sheet and input your dates, as well as the periods you are currently inputting (monthly, quarterly, yearly, etc.)
* Attributes, Product, and RevType fields are are all pulled through pivot tables into formatted schedules. It is recommended you adjust the formatted schedules before doing a full refresh of the data.
* Delete any levels you may not need. All are independent and have their own measures so nothing should break when deleting. Sheets are color coded.
  + Light Blue – Cust-Prod-RevType Level
  + Orange – Cust-Prod Level
  + Pink – Cust Level
* Most views summarize all relevant KPIs through all attributes and levels. Most data is pulled using INDEX MATCH formulas. If your data is particularly large, confirm the ranges in the file are big enough. Additionally, you may need to add additional rows to fit all categories. You should be able to do so by simply dragging formulas down.
* Non-dynamic components of schedules will require your input/modification – these will be highlighted throughout the file.