# VALIDATING DATA FOR MAP

To facilitate a more effective way of validating our data as well as to set a standard on all our checks, please follow the following guidelines in Validating Tables we are creating or modifying.

For every table we bring in, we want to be able to validate the following:

* 1. We are getting the correct count of records per table
  2. There are no duplicates by the table’s keys
  3. Transformations are coded for correctly

**General Guidelines**

Creating Test Cases

1. The goal is to be able to validate our loaded data by comparing with data from our source tables by way of the EngA ( \_S tables)
2. Create A Test Script document for the Subject Area you are testing. Below is a template.



1. Create the applicable Test Cases. Check below for the recommended Test Cases
2. Use the XLS Tool Below to help create (Expected and Actual result) queries for the test cases



1. If your Target table is a result of joins of tables, you can reference the Source-To-Target mapping document for the appropriate joins for the “Expected Results” query

|  |
| --- |
| ACTUAL RESULTS:  sel SUM(ValidSrvyCnt) from  **L2\_SelfServAnlysEngADev\_T.MbrSrvyCnt**  where MthNbr = 201408;  EXPECTED RESULTS:  sel count(distinct O.RespnId) from  **L2\_SelfServAnlysEngADev\_S.InStrSrvyRespnDtl O**  **INNER JOIN L2\_SelfServAnlysEngADev.FiscCalndr F**  **ON O.IntrvwEndDt = F.DayNbr**  **LEFT OUTER JOIN L2\_SelfServAnlysEngADev\_S.SrvyDtl D**  **ON D.SrvyId = O.SrvyId**  **LEFT OUTER JOIN**  **L2\_SelfServAnlysEngADev\_S.SrvyCatg C**  **ON C. SrvyCatgId = D.SrvyCatgId**  where FbkStat = 1  and **MthNbr = 201408;** |

1. Ensure that you are applying a valid time period for testing for all queries

|  |  |
| --- | --- |
| sel SUM(ValidSrvyCnt) from L2\_SelfServAnlysEngADev\_T.MbrSrvyCnt  where **MthNbr = 201408**;    sel count(distinct O.RespnId) from  L2\_SelfServAnlysEngADev\_S.InStrSrvyRespnDtl O  INNER JOIN L2\_SelfServAnlysEngADev.FiscCalndr F  ON O.IntrvwEndDt = F.DayNbr  LEFT OUTER JOIN L2\_SelfServAnlysEngADev\_S.SrvyDtl D  ON D.SrvyId = O.SrvyId  LEFT OUTER JOIN  L2\_SelfServAnlysEngADev\_S.SrvyCatg C  ON C. SrvyCatgId = D.SrvyCatgId  where FbkStat = 1  and **MthNbr = 201408;** | sel MIN(OrdDt), MAX(OrdDt) from L2\_SelfServAnlysEngADev\_T.MbrSrvyRespnInStr  where **OrdDt <= CURRENT\_DATE and IntrvwEndDt >= CAST( '2014-08-31' as DATE) and IntrvwEndDt <=CAST( '2014-10-04' as DATE;**  sel MIN(OrdDt), MAX(OrdDt) from L2\_SelfServAnlysEngADev\_S.InStrSrvyRespnDtl where **OrdDt <= CURRENT\_DATE and IntrvwEndDt >= CAST( '2014-08-31' as DATE) and IntrvwEndDt <=CAST( '2014-10-04' as DATE);** |

Executing Test Cases

1. When running test cases, please fill out Date Of Execution, Validated By and A short description of results especially for failures
2. Create A Duplicate Tab for each Iteration of the testing. Mark the Iter (1 for First Pass, 2 for 2nd) on the Tab Name.
3. You can add new worksheets to paste results. Mark them as “Iter N results” so we can link the results to the Pass

**Recommended Test Cases**

Test Case 1: Counts of Records by Table

* Please create 1 sub test case per table

Test Case 2: Check Duplicates Records

* If multiple columns make up a key, create a test case to count that the unique key combinations between source and target are the same. If there is only one Test Case 3 should handle this.
* Please create 1 sub test case per table that has multiple columns as key

Test Case 3: Distinct Counts of attributes

* Please create 1 sub test case per attribute

Test Case 4: Sum of Amounts, Scores and other Numbers

* Please create 1 sub test case per attribute

Test Case 5: Counts Distribution of attributes with a finite possible values

* Please create 1 sub test case per attribute
* Put the distribution results in a separate worksheet