# List of database tasks 3 July

This document is to aid teamwork. It also introduces a systematic (‘Agile’) approach to updating and testing.

All these to be done locally first. Then, when tested and working, implemented on the remote database

## Tests

The tests to be completed for each section are:

1. Delete the databases from the studio
2. Create two new, empty databases
3. Run SSIS in Visual Studio to populate the OLTP data base
4. Verify the contents of the tables using SSMS
5. Create or re-use spreadsheet named ‘TEST\_ddmm\_nn’ in ‘RESULTS’ folder, with worksheet called test\_xx where xx is the section number, ddmm is the date and nn starts at 1 and increments each time there is a failure (this way, we keep track of the failures)
6. Create a pivot table in TEST\_xx containing the Main Industry data by year and Main industry, filtered for province, and test on ‘Canada’ and ‘Manitoba’
7. Ditto for the Creative Industry data

## Documentation

1. Document the ‘Mapping.xlsx’ spreadsheet so we are clear what the purpose of each sheet is.
2. Document the SSIS process which is implemented by Visual studio so it’s clear how the data gets in
3. Document the OLTP and ROLAP scripts so it’s clear what each item in each script does
4. Check the documentation in the ‘META’ folder
5. Update ‘Readme’

## Name space

### Database

Rename CANADA-CI to CANADA\_CI\_OLTP

Rename ROLAP to CANADA\_CI\_ROLAP

### Table names

After item 2a, rationalise the names of the OLTP tables

### Mapping source names

After item 3a, change the worksheet names to coincide with the renamed OLTP tables

### Table fields

All Primary keys should have PK\_ prefix

All Foreign keys should have FK\_ prefix

## Folder rationalisation

Move ‘Mapping.xlsx’ to ‘Meta’

## Date extension

The intermediate unpivoted tables need to be extended to 2020

## Simplify Data Import

1. Import LFS data directly into the CI Fact table, unpivoting it at the same time as it is imported
2. Import nominal data directly into CI Fact
3. Import real GDP data directly into CI Fact
4. Import P & H data directly into CI Fact
5. Amalgamate the CI Fact and MI Fact tables

## Licensing

We should identify a suitable Copyright license

This should be identified in all documents of the project

## Bugs

1. ‘Unpivot LFS source’ did not run initially

## New Changes

1. OLTP: rename CI NAICS table to dimCreativeIndustry
2. Mappings: rename CI NAICS to dimCreativeIndustry
3. OLTP: rename CNOCS to dimCcreativeOccupations
4. Mappings: rename CO NOCS to dimCcreativeOccupations
5. OLTP: MI NAICS: Change name to dimMainIndustries.
6. Mappings: MI NAICS: Change name to dimMainIndustries.
7. Mappings: create a single dimGeography table
8. Remove data viewer
9. Remove GeoName from MainIndustries query
10. Number the main industries

# Progress chasing

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item | Date | Actions | Tests | Notes |
| 3a | 3/7 | Created tables and views | Tests 1-4 | Create tables and views edited to change ‘USE’  Had to change most of the dataflow operations to use the designated excel connection managers |
| 3a | 3/7 | Imported Source Data | Tests 1-4 |  |
| 3a | 3/7 | Create Fact File | Tests 1-4 | ‘Unpivot LFS source’ did not run initially. Inserted Data viewer. Then it ran. See ‘bugs’ |
|  |  |  |  |  |
| 2a | 3/7 | Documented CI NAICS |  | Changed the name of the SSIS source to include the worksheet name – check for side-effects but should have none. There are redundant fields. |
| 2a | 3/7 | Documented Main Industries Master |  | Changed the name of the SSIS source to include the worksheet name– check for side-effects but should have none.  There are redundant fields |
| 2a | 3/7 | Documented IOICC Standardiser |  | Changed the name of the SSIS source to include the worksheet name |
| 2a | 3/7 | Documented NOCS standardiser |  | Changed the name of the SSIS source to include the worksheet name |
| 2a | 3/7 | Documented Geo Standardiser |  | Changed the name of the SSIS source to include the worksheet name |
| 2a | 3/7 | Documented Geography |  | Changed the name of the SSIS source to include the worksheet name |
| 2a | 3/7 | Did not fully document the remainder | Tests 1-4 | Tests 1-4 all ran |
| 8 | 3/7 | Check that bug 7a has gone | Tests 1-4 | All tests ran; now we should remove the data viewer |
| 9a | 3/7 |  | Tests 1-2 | Test ran |
| 9b | 4/7 |  | Test 2 | Test ran |
| 9c,d | 4/7 |  |  |  |
| 9e,f | 4/7 |  | Tests 1-4 | Ran, but Import Data wouldn’t start the first time through.  Removed the Data Viewers from ‘Create Fact File’ |
| Unlisted | 4/7 | Removed all the ‘Excel File Manager’ connection managers.  Removed the redundant LFS connection parameter from ‘Project Params’ | Tests 1-4 | Ran. Peculiar error is reported about a null connection string, but it still works OK |
| 6a | 4/7 | Removed the LFS Source table and unpivot directly into CI Fact | Tests 1-3 | Ran |
| 6a | 4/7 | See above | Test 4-5 | Ran – recreated CI for Manitoba which seemed to be about right |
| 6b,c,d | 4/7 | Unpivot all data directly into CI Fact | Tests 1-3 | Some debugging needed but ran. Main change is to convert Year to text string. |
| 6 a,b,c,d | 4/7 | More testing | Tests 1-5 | Only for LFS jobs hence not really complete. |
| 9g | 4/7 | Create a single dimGeography table | Tests 1-4 |  |
| 6e | 4/7 | Amalgamate CI and MI facdt | Tests 1-4 | New field for MI sector added to dimCreativeIndustries  New Query ‘Main Industries’ created |
| 6c | 4/7 | More tests | Test 5,6,7 | Ran after correcting bugs in CreativeIndustries view |