# Progress chase

This document records, in a not very systematic way, the changes introduced at each commit and the intentions for the next commit

### Improvements

* 1. Remove ‘pivoted sources’ table and unpivot all sources directly into the fact table
  2. Rewrite ‘Calculate Pivoted Additions’ to work on the Fact table instead of Factsource
  3. Years should go back earlier
  4. Rewrite the PENN calculations to combine in a single query
  5. Provide a DateTime field in place of the ‘Year’ field in the Fact file
     1. Remove the date Dimension – not needed if the fact file contains a properly managed Datetime field
     2. Overhaul the input packages to convert year to a datetime
  6. Remove the Definitions file – too cumbersome and confusion
  7. Add a ‘reduced view builder’ to construct specialised exports
     1. Note this can get quite messy as at present specified because we also have reduced versions of the dimension files

### Routine maintenance and debugging

* 1. Create space on server to ensure project will run on MSSQLExpress (10GB per database)
  2. Overhaul and validate ROLAP export
  3. Check the data
  4. Penn constant USD should be scaled up

### Data

* 1. Check and remove redundant duplicate datasets
     1. Check if WEO2020 includes the 1970-1980 data omitted from WEO2018
  2. Create reduced dataset for economic growth project
  3. Create reduced dataset for inequality project
  4. Create reduced dataset for learning

### Tests

* 1. Create basic excel spreadsheet and pivot table
  2. Create standardised offline dataset

### Tidying up

* 1. Change the name of the ‘GERGLE’ folder to something more meaningful
  2. Change all field names to be compliant with MariaDB/MySQL naming conventions (no spaces, all lower case – though the latter is not essential it eliminates risk of ambiguity)
  3. Delete all intermediate views and functions at the end so the database is tidier (and smaller)
  4. Name Space: more consistent naming of variables so that the SSIS mappings are obvious

### Visualisations

* 1. Create package to export to MariaDB
  2. Create tentative Power BI workspace

### Documentation

* 1. Update readme

#### Progress Chase

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Action | Date | Status | Commit Title | Tests |
| 1a | 21/8/20 | Unpivot directly done |  | Ran, data not checked |
| 1b | 21/8/20 | Remove pivoted tables and test |  | Ran, data not checked |
| 2a | 22/8/20 | Achieved by removing the extra index on FactSource |  | Ran, data not checked |
| 2b | 22/8/20 | Ran ROLAP export script through and debugged it | ROLAP overhaul and validate | Ran, data not checked |
| 5a | 21/8/20 | Done Test0822.xlsx |  | Checked Barro-Ursua |
| 5b | 22/8/20 | Done (NOTE names still contain mix of upper and lower but this should still be MariaDB compatible) | MariaDB Compatibility |  |
| 5c | 22/8/20 | Done – separate package | Tidy |  |
| 7a | 23/8/20 | Readme files all updated | Offline Dataset |  |
| 4b | 23/08/20 | Standard offline dataset and catalogue created | Offline Dataset |  |
| 1e | 25/9/20 | Datetime field created and tested | New Date Handling | Cursory check – constructed pivot table in ‘Test0925 New Date Handling’ |
| 1c | 25/09/20 | This would appear to be redundant with the new date handling |  |  |