DERIVED FORECAST Model

Reference Manual

Version 3 (January 2018)

edge analytics

0113 384 6087
popgroup@edgeanalytics.co.uk

Produced by:

Leeds Innovation Centre | 103 Clarendon Road | Leeds | LS2 9DF
0113 384 6087 | popgroup@edgeanalytics.co.uk | www.edgeanalytics.co.uk

Commissioned by:

Local Government Association

Local Government House | Smith Square | London SW1P 3HZ
Tel 020 7664 3000 | Fax 020 7664 3030 | info@local.gov.uk | www.local.gov.uk

With technical demographic advice from Ludi Simpson

Contents

1 Introduction 1
1.1. Document Structure 1
1.2. This Document 2
1.2.1. Assumptions 2
1.2.2. Document Conventions 3
2 Introduction to DF Model 4
2.1. What is the DF Model? 4
2.2. Key Features 5
3 Installation Guide 6
3.1. Installing DF 6
3.2. Macros and Trust Settings 7
3.3. Technical Support 8
4 DF Model Methodology & Setup Methods 9
4.1. DF Model Methodology 9
4.2. DF Setup Methods 10
4.3. Stages of Forecasting with DF 11
5 Working with DF 13
5.1. DF and Excel 13
5.2. DF Workbooks 13
5.3. DF Worksheets 13
5.4. Completing the Workbooks 14
5.4.1. Data Entry 14
5.4.2. Validation 14
5.4.3. Saving the Workbooks 15
6 DF Model Configuration 16
6.1. Introduction 16
6.2. Setting Up the Model 16
6.2.1. Step 1 – Initial Settings 16
6.2.2. Step 2 – Group Selection 18
6.2.3. Step 3 – Bespoke Settings (Optional) 19
6.2.4. Step 4 – Running DFSetup 20
7 Preparing Assumptions 22
7.1. Introduction 22
7.2. The ‘DF Pop’ Workbook: Population Forecast 23
7.2.1. Notes and Validation 24
7.3. The ‘DFPopAdjust’ Workbook: Population Adjustment 25
7.3.1. Notes and Validation 27
7.4. The ‘DFSubPop’ Workbook: Sub-Populations 28
7.4.1. Notes and Validation 29
7.5. The ‘DFRates’ Workbook: Derived Category Rates 31
7.5.1. Notes and Validation 32
7.6. The ‘DFFactor’ Workbook: Size Factors 34
7.6.1. Notes and Validation 35
7.7. The ‘DFCons’ Workbook: Constraints 36
7.7.1. Notes and Validation 38

8 Scenario Configuration 40
8.1. The 'DFScenario' Workbook 40
8.2. Running the Model 42

9 Model Results 43
9.1. Output File Summary 43
9.2. The 'DFForecastDetail' Workbook: Detailed Results 44
9.3. The 'DFForecastTotal' Workbook: Summary Results 44
9.4. The 'DFRatesOut' Workbook: Derived Category Rates 46
9.5. The 'DFRiskPop' Workbook: Population At Risk 47
9.6. The 'DFReporter' Workbook: Reports and Charts 48
9.6.1. Reporter 49
9.6.2. Charter 55
9.6.3. Creating Aggregates of Groups or Categories 59
9.6.4. The 'DFCompare' Workbook: Comparing Scenarios 60

10 DF Configuration – Advanced 63
10.1. Summary of Options 63
10.2. Modification of Age-Sex Groups 64
10.3. Population Adjustment Options 65
10.4. Sub-Populations Options 65
10.5. Modification of Derived Unit Categories 66
10.6. Validation Options 67
10.7. Calculations on Derived Units Options 67

11 Using DF within POPGROUP *(Including supply-led population projections)* 69
11.1. Summary 69
11.2. POPGROUP Constraints 69
11.2.1. Notes and Validation 70
11.3. The 'DFSupply' Workbook 71
11.3.1. Single Conversion Ratio 72
11.3.2. Households to Dwellings Ratio 73
11.3.3. Labour Force to Jobs 73
11.4. 'Constraints and Impacts' 74
11.5. Migration Weights 75

Appendix A DF Example 76

Appendix B Security Settings 78
For Users of Microsoft Excel 2007 Onward 78
For Users of Microsoft Excel 2003 79

Introduction

This reference manual is designed to enable a user to install and use the Derived Forecast (DF) model, which forms part of the POPGROUP family of models.

1.1. Document Structure

This manual has ten main chapters (Chapters 2-11) and two appendices (Appendix A & B):

Chapter 2: **Introduction to DF Model**
An overview of the DF model, its key features and recommended uses

Chapter 3: **Installation Guide**
How to install DF for use on a PC, including hardware and software requirements

Chapter 4: **DF Methodology & Setup Methods**
Outlines of DF methodology and setup methods

Chapter 5: **Working with DF**
Features of DF files and operation

Chapter 6: **DF Model Configuration**
Stage 1 of DF: Defining 'Groups' and setting up DF using default settings

Chapter 7: **Preparing Assumptions**
Stage 2 of DF: Entering data into the skeleton workbooks

Chapter 8: **Scenario Configuration**
Stage 3 of DF: How to produce a derived forecast

Chapter 9: **Model Results**
Stage 4 of DF: Examining and using the model output

Chapter 10: **DF Configuration – Advanced**
How to configure bespoke DF model setups

Chapter 11: **Using DF within POPGROUP**
Running POPGROUP scenarios with a derived forecast

Appendix A: **DF Example**

Appendix B: **Security Settings**

1.2. This Document

This manual can be used in a number of ways.

Much of the DF model is intuitive and the reference manual can be used to look up, for example, specific rules of data entry when they are needed.

However, if the user is beginning without the advantage of Data Modules for the UK local authority areas, then it may be useful to sequentially follow the chapters, learning about the DF model (Chapter 2), how to install it and its methodology (Chapter 3 and 4), working with the DF model (Chapter 5) and how to setup and produce forecasts and analyse results (Chapters 6 to 10).

Using the Data Modules, a user can replicate official forecasts for a chosen area or areas in the UK. The Data Modules are utilities that configure and run DF automatically. Data Modules are released periodically to coincide with the release of new data. Alternatives to the official assumptions can then be made by using the step-by-step guides, or by dipping into the relevant sections within Chapter 7.

Should the user wish to produce their own bespoke DF model setup, a step-by-step guide is provided in Chapter 10.

1.2.1. Assumptions

It is assumed that the user is familiar with Microsoft Excel workbooks, their feature and operation; this manual does not attempt to train the user in these.

It is also assumed that the user is familiar with or has the support for the acquisition of suitable demographic data with which to populate the model workbooks.

1.2.2. Document Conventions

In this document, the names of Excel workbooks are written without the .xls suffix, in bold grey font (e.g. **DF_Scenario**). Sheets within workbooks are written in bold grey italics surrounded by single quotation marks (e.g. ***'Default'***).

Folder names are written in uppercase, surrounded by singly quotation marks (e.g. '…DERIVED FORECAST\OUTPUT'). Users may name folders and workbooks using both lower and upper case, depending on preference.

Reference to cells within the worksheets, or buttons/options within the workbooks are in blue italics surrounded by single quotation marks, for example, *'Validate'*.

2 Introduction to DF Model

This section provides an overview of the DF model, its key features and some general points about its use.

2.1. What is the DF Model?

The DF model sits alongside the POPGROUP model. The DF model allows data to be entered for any variable that is closely related to the age-sex structure of the population as forecast by the POPGROUP model or independently, including household structure, economic activity and disability.

The DF model uses Excel to manage the data inputs and outputs and provide flexibility to enable users to experiment with and analyse alternative derived forecasts. Data extraction and chart routines provide users with easy access to both broad and detailed summary statistics.

DF is used for any characteristic related to the age and sex composition of the population, including household formation, economic activity, disability, illness and health conditions of different types, and many types of service demand. DF can be used to produce high-quality, defensible household, dwelling, labour force and employment demand growth statistics for use in strategic planning, service planning and policy development.

The DF model is used extensively by local authorities and local and regional organisations across the UK and beyond, providing a desktop utility for the evaluation of alternative growth scenarios to support local planning. In addition, it has the flexibility to accommodate the methods of household projection from each of the four UK statistical agencies.

2.2. Key Features

The DF model provides a single 'setup' procedure to define a model appropriate to the data available and their particular relationship to the age-sex structure of the population. It allows specification of different models, for example:

- One rate applied to the whole population for a standard set of five-year age-sex groups, as might be the case for some models of long-term limiting illness.
- Several rates adding to less than 100%, applied to a population modified by subtraction of those in institutions and excluding child age groups, as is the case for the current government sub-national household projections in England and Scotland.
- Several rates that sum to exactly 100%, to compute the population in each household type and further calculations involving household size to compute the number of households, as in the sub-national household projections for Wales and Northern Ireland.
- No disaggregation by sex or no disaggregation by age or sex, as in a very simple model of household formation by size.
- Application of rates to a population that is sub-divided into age-sex-student status, as is allowed for in labour force projections.

The DF model avoids complicating the use of the software for the majority of users who will probably not use its option flexibility, though each user may want a different subset of options. This is achieved through prepared standard model setups in which the options are pre-entered, such that the user need only enter the groups (usually areas) to be forecast, before creating the skeleton input files. The standard setup can be altered to create variant DF models.

3 Installation Guide

3.1. Installing DF

DF needs to be installed before a model can be setup. The DF model is installed from a CD or Dropbox link. All that is required prior to installation is for the user to give some thought to the folder structure that they wish to use for storing data input and model output workbooks, and to decide on a name for the folder that will hold DF and the `DFSetup` workbook. The `DFSetup` workbook is used to set up the initial details of models.

To install the system:

A. Decide on the name of the folder you want the system to create to hold DF. If you have POPGROUP, it should be the same folder that it is held in. The default is 'C:\FORECAST'.

B. Insert the CD into your disk drive or download the software from the Dropbox link.

C. Using Windows explorer, navigate to the CD drive or where the software from Dropbox has been downloaded to and double click on the 'Derived Forecast.exe' file. A dialogue box will ask you to confirm the folder in which the program will reside.

D. The installation will place the `DFSetup` workbook in your chosen folder and will also create a sub-folder called 'DFSYSFILES'.

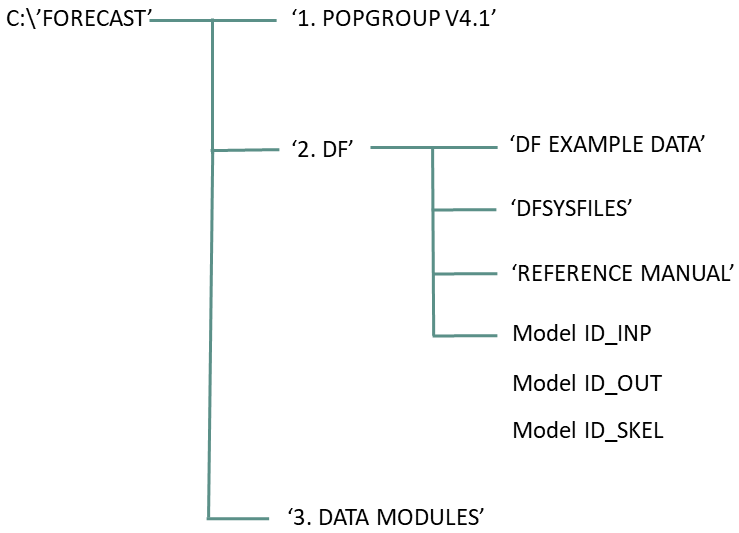
No other files from the CD or Dropbox link are needed for installation. A copy of this reference manual will also be provided with the software.

!	**Do not alter any files within the 'DFSYSFILES' sub-folder**; the user will never directly access these. Do not put any other files in the 'DFSYSFILES' folder.

After the system has been installed, the user will set up a model (Chapter 6), to store:

- Skeleton data input files
- Completed data input files and scenarios files
- Output files.

These files may be stored in any location, but for ease of documentation and for security, it is helpful to follow the standard folder structure, which is used in the DF documentation and encouraged by the DF model's default settings. If a model is set up, the folder structure might appear as follows:



The DF model has been developed in Excel VBA, specifically as an Excel 2003 application to run on a standard desktop PC. The DF model is supported for operation on versions of Excel from 2003 to 2016 and versions of Windows from XP onward.

!	Google Desktop may cause problems with Excel VBA when opening and closing workbooks that contain certain code modules. If you have Google desktop running, **please uninstall it.**
!	Any Microsoft Office 'Add-ins' which control access to or the organisation of files (e.g. 'document management' systems) should be disabled as they are likely to interfere with the proper operation of DF.

3.2. Macros and Trust Settings

Before DF model can be used for the first time, some additional settings are required to ensure that it is able to run. The DF model comprises a set of Excel workbooks whose operation is controlled by Excel VBA macros.

!	**Macros must be enabled when using POPGROUP**
On opening the `DFSetup` workbook, depending on your current Excel security settings, you may be asked whether to allow macros or not. **Always click to allow macros if asked when running POPGROUP**. For advice on enabling macros and the appropriate security settings in different versions of Excel, refer to Appendix B.	

Additionally, in order for the DF model to run correctly, the *'Trust access to Visual Basic Project'* setting must be enabled. The method for setting this depends on the version of Excel being used. Refer to Appendix B to determine the setting required.

If this setting is not configured correctly, DF will not be able to run. A warning is given when running `DFSetup` if this is the case.

3.3. Technical Support

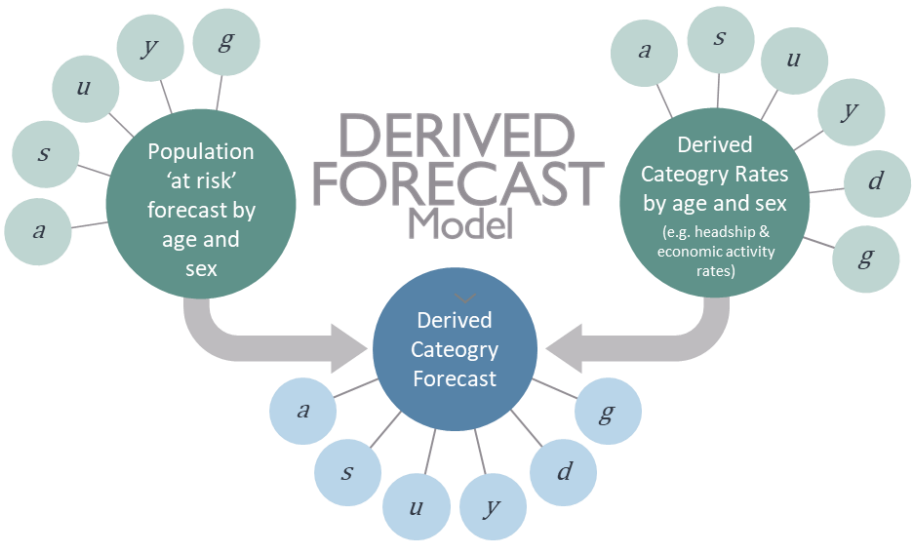
The software use is designed to be intuitive. This manual is designed to give detailed explanation when a user is in doubt. If a user still has problems, technical support is available from Edge Analytics using the following contact details:

Edge Analytics Ltd
Leeds Innovation Centre
103 Clarendon Road
Leeds, LS2 9DF
Tel: 0113 384 6087
Email: popgroup@edgeanalytics.co.uk

4 DF Model Methodology & Setup Methods

4.1. DF Model Methodology

The DF model produces its forecasts as a multiplication of two indicators: a forecast population 'at risk' (by age and sex) and a 'derived category' rate (by age and sex). Derived categories could, for example, be households (defined by headship rates) or the labour force (defined by economic activity rates).



$$D_{a,s,u,y,d,g} = \frac{P_{a,s,u,y,g}\ R_{a,s,u,y,d,g}}{100}$$

D	Derived Category Forecast	y	Year
P	Population 'at risk' Forecast	d	Derived category
R	Derived Category Rates	g	Group (usually an area, but can be an ethnic group or social group)
a	Age-group		
s	Sex		
u	Sub-population		

Figure 1: The DF model methodology

The population 'at risk' would typically be derived from a population forecast, making any adjustments necessary (e.g. removing population 'not-in-households') and sub-dividing into sub-populations where appropriate (e.g. student population).

The derived category forecast may be finally adjusted by 'factors', for example, to divide household population by household size (as in the Welsh household projection model).

The user controls the age-sex categories, the population adjustments and final adjustment, the number of derived categories, the number of groups (e.g. areas), and the labels used by the model.

4.2. DF Setup Methods

The DF model can be used in two main ways (Figure 2).

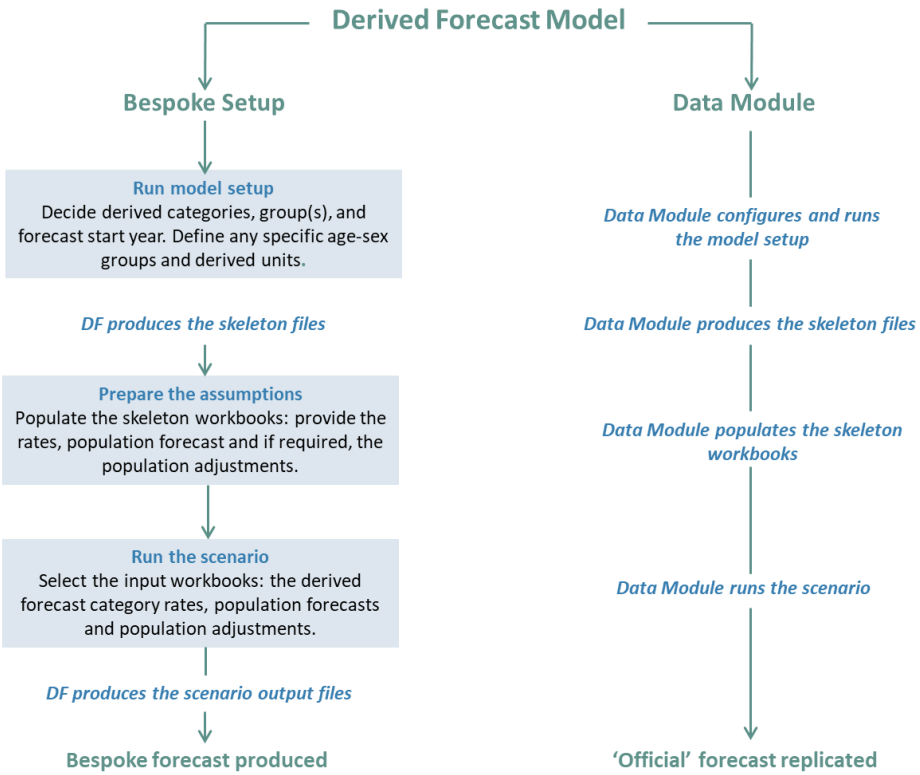


Figure 2: The DF model – setup methods

The bespoke setup involves the user defining the model and then entering inputs and assumptions manually. This is recommended when a forecast is required for areas that are not

forecast by the UK official agencies, or for a characteristic other than those for which Data Modules exist (at the time of writing, households and the labour force).

The second alternative, available to users in England, Wales and Scotland who are running forecasts of local authority districts, is to use a Data Module, which will set up and run DF automatically. This is recommended where the user wishes to replicate official projections (e.g. those produced by the Department for Communities and Local Government (DCLG), Office for National Statistics (ONS), Welsh Government (WG) and the Office for Budget Responsibility (OBR)) for a chosen area and explore alternative assumptions. The manual does not refer to these Data Modules, which are documented in separate User Guides.

DF is flexible, in that the model workbooks produced by a Data Module can then be modified by altering data or changing assumptions to develop bespoke variations of official projections. For example, the user might wish to examine the impact of altering economic activity rates or headship rates for their chosen area. Instructions how to alter assumptions are detailed in Chapter 10.

4.3. Stages of Forecasting with DF

Once the configuration options are complete, DF application will create a series of 'skeleton' files that reflect chosen options. To run the DF model these skeleton files must be populated with appropriate data to create the user's own 'input' files.

With a set of input files populated, the user runs the DF using a simple 'scenario' configuration process. A short run-time produces a series of 'output' files from which the user is able to examine data, charts and reports which summarise the chosen derived forecast. Different scenarios can be compared using the `DFCompare` utility.

The DF user is able to add extensive notes to describe the data inputs and assumptions used in each input file. This information is copied directly to output files to enable the most efficient scrutiny of results.

There are four main stages to DF configuration, each of which is covered by a separate chapter in this reference manual (see following page). When using a Data Module, Stages 1, 2 and 3 are made easy to replicate an official or standard forecast and do not require direct action by the

user. When making alternative assumptions, the user will make adjustments to the official forecast input files (Stage 2) and run the new forecast (Stage 3). The reporting of results (Stage 4) is helped by DF functions but is always under the control of the user.

STAGE 1 – SETTING UP THE MODEL BASICS (Chapter 6)

- Open the `DFSetup` workbook, which was created when the system was installed.
- Provide general information and labels for the headings, folder names and location and categories of DF to be produced.
- Provide the labels for the area setup or select the POPGROUP `MODEL_SETUP` to read the area labels from.
- Define base year, derived units and select age/sex groups.
- DF creates the skeleton data input workbooks in a folder nominated in the `DFSetup`.

↓

STAGE 2 – PREPARING THE ASSUMPTIONS (Chapter 7)

- In the input workbook `DFRates`, provide rates for each of the categories defined in Stage 1.
- In the input workbook `DFPop`, provide population forecast by age and sex for each population group defined in Stage 1.
- If applicable, in the `DFPopAdjust` workbook, provide population adjustments relevant to the derived forecast category (e.g. communal establishment population in the household projection model).
- If a fixed outcome is to be applied to a year or years of the forecast, provide details of these constraints in the `DFCons` workbook.
- Save the input workbooks with unique names in the input folder created in Stage 1. There can be many workbooks for each input, each one representing a different set of assumptions that has been prepared.

↓

STAGE 3 – RUNNING A FORECAST (Chapter 8)

- Open the workbook `DFScenario`.
- Give details of the final year for this forecast, the names and locations of the data input workbooks to be used in this model run.
- DF produces the forecasts in output workbooks saved with a scenario ID chosen by the user.

↓

STAGE 4 – USING THE OUTPUT: TABLES AND CHARTS (Chapter 9)

- In `DFReporter`, the report generator for the forecast, aggregate results to selected age bands and categories.
- Produce line charts for time series.
- Save the charts and tabulations in the report-generator itself, or copy them to your own workbooks and documents for further manipulation and dissemination.
- Print a summary of the forecast by derived category from the `DFForecastTotal` workbook.
- In the `DFRiskPop` workbook, view the population forecast and population adjustment.
- View detailed results in the `DFRatesOut` and `DFForecastDetail` workbooks.

5 Working with DF

5.1. DF and Excel

When running DF or the `DFSetup` program, do not attempt to use other 'instances' of Excel on the same machine – i.e. do not start up one or more copies of Excel.

5.2. DF Workbooks

DF operates using a series of 'program' files. These are stored in the 'DFSYSFILES' folder which is held within the folder nominated by the user on installation.

!	**The DFSYSFILES folder should not be altered in any way.**

5.3. DF Worksheets

The DF system comprises a set of Excel workbooks which are pre-formatted to allow data to be easily entered into the model.

!	**The formatting of the DF workbooks should not be changed.**

In particular, within the DF workbooks:

- Worksheets should not be deleted.
- Additional worksheets should not be inserted.
- The order of the worksheets should not be changed.
- Worksheet protection should not be altered.
- The extent and format of data entry areas should not be changed.
- The workbooks' named ranges should not be altered.

5.4. Completing the Workbooks

5.4.1. Data Entry

When entering data into DF files you will often have to paste data from other Excel files into your input workbooks. It is recommended that you use the 'paste values' function to do this. In this way you will not change the format of the DF files. Numbers entered by the user on the input workbooks are formatted in blue font and yellow shaded cells.

Extensive use is made of the Excel notes facility to provide guidance on the acceptable contents of data fields. To view the notes, rest the cursor over a cell containing a red triangle in its top right corner.

In some cases, users may wish to use formulae to provide an efficient audit of how data were copied from other files. However, under some circumstances, DF will fail when 'array formulae' are used. To avoid this, users should replace array formulae with values before running a scenario.

!	**Never drag cells from one place to another.** This will change DF's formulae and possibly prevent DF from operating correctly.

5.4.2. Validation

Each sheet of the input workbooks provides an opportunity to validate the user inputs. The validation process ensures that the model will work properly with the data items provided.

On validation, any errors or warnings are reported in the 'Diagnostics' table on the *'Notes'* worksheet. The user may carry out validation of the options chosen at any time while entering data into an input workbook. The errors must be corrected before the forecasting program can be run. Warnings alert the user to data entry that may be unusual.

When the DF scenario is run, the program separately validates the input files for rates, constraints, population forecast and population adjustment, applying rules specific to each workbook. If errors are found, the program will terminate without producing any forecasts. These validation checks are designed to ensure that:

- Individual data entry items are within acceptable and expected bounds.
- Data entry is complete.
- Data entry is consistent with the chosen DF configuration.

5.4.3. Saving the Workbooks

The skeleton workbooks produced are all read-only. Once data has been entered into them, they must be saved under a new name. It is recommended that the user stores the completed workbooks in the input folder created by the `DFSetup` routine.

!	To allow DF to run on different versions of Excel all workbooks are '97-2003' workbooks. If using versions of Excel after Excel 2003 (where a choice of workbook type is provided) **the workbooks must always be saved as this type – i.e. .xls** (Figure 3). The software will not operate correctly if the workbooks are saved as any other type. DF saves files as .xls automatically. The user does not need to change any settings, but should avoid saving any DF file as any type other than .xls.

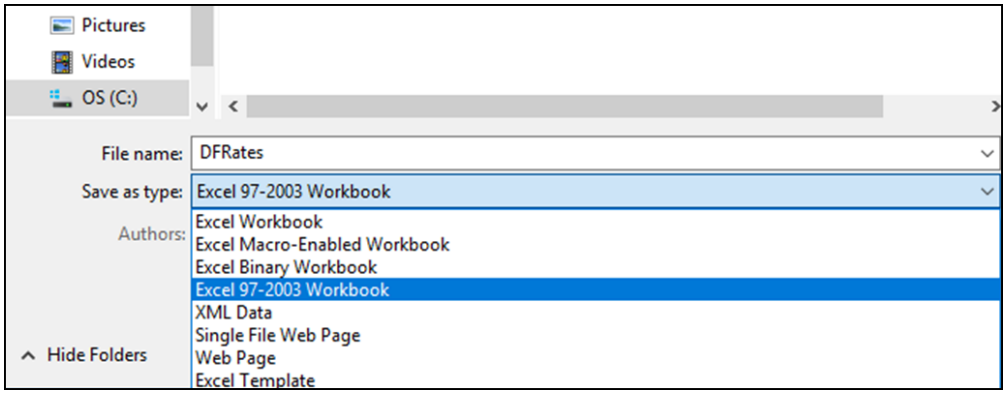


Figure 3: Workbooks saved as 97-2003 workbook type

6 DF Model Configuration

6.1. Introduction

This chapter describes how to configure the `DFSetup` workbook to set up a new DF model.

As part of the installation of DF, the `DFSetup` workbook will have been created in the main DF folder. This workbook is used each time a new model is set up.

Running the `DFSetup` process produces a number of 'skeleton' files. These workbooks are then populated by the user with assumptions in the form of data, from which scenario forecasts are generated (see Chapter 7).

!	**The entries given in the model setup cannot be changed later**. This includes the labels for each group and for the total of all of the groups ('all groups'). Choose them carefully before entering data.

6.2. Setting Up the Model

!	Model setup file (`DFSetup`) must be opened from the same folder as the 'DFSYSFILES' folder.

6.2.1. Step 1 – Initial Settings

Upon opening the `DFSetup` workbook, the user will be presented with an initial form (Figure 4) which requires the following inputs:

File Header

- In the *'File Header'* cell, write something simple that will describe the scope of the work. Usually, this will be the name of the overall geographical area, organisation, or project

title. Your input here will be reproduced at the top of each sheet on each of the input files and output files, appended to the words 'Derived Forecast - *Category*'.

Model ID, to name folders

- In the *'Model ID, to name folders'* cell, enter a name appropriate to the model you are setting up. Using this name, 3 sub-folders will be created in the folder that you select to hold them (see *'Location of Folders'* below). These three folders will be called <ID>_SKEL (to hold the skeleton files), <ID>_INP (to hold the input files) and <ID>_OUT (to hold the output files), where <ID> is the name you give in this cell.
- You may use this model to make a variety of projections. You might use a model ID to reflect your area(s), the type of forecast, and the base year, for example *Hampshire_Illness_2011*.

Location of Folders

- The skeleton, input and output files will be saved into sub-folders created in the folder named in the *'Location of Folders'* cell. The default is the folder where the Derived Forecast's `DFSetup` workbook is located, but you may change it, either by double clicking on the cell and then browsing for an alternative, or by typing in the desired location.
- If the recommended folder structure was used, you should not need to change the default location.

Type of model to be set up

- The *'Type of model to be set up'* can be defined using the dropdown list containing a number of pre-configured models based on official projection models, e.g. CLG 2014-based household model. Your selection here will be used to define the type of derived category. Alternatively, the user may choose the 'User Defined' option to specify a bespoke model type.

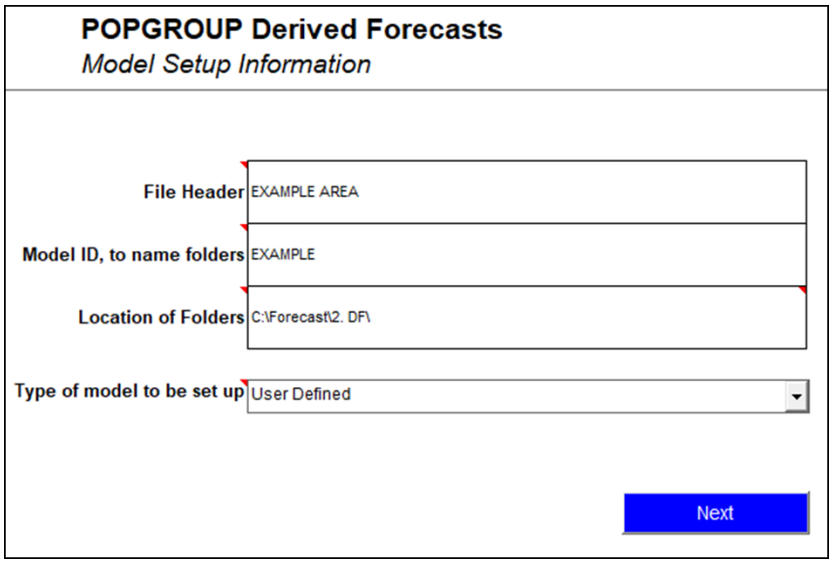


Figure 4: DF model setup: Step 1

6.2.2. Step 2 – Group Selection

When all selections have been made, use the *'Next'* button to progress to the next step of the setup process (Figure 5). This requires the user to identify the groups that are to be used by the model. Groups will most likely be a list of geographic areas, but may also be ethnic or social groups.

A 'short-cut' option is available to complete this step, whereby the user can access label information from an existing POPGROUP `MODEL_SETUP` file. Double-click in the box labelled *'Use Labels from a POPGROUP Model_Setup file'* to browse for the correct location and then click the *'Get Labels'* button. This will populate the remainder of the form with the required information on the number of groups, the 'all groups' short and long labels and each of the individual groups' short and long label.

Alternatively, the user must enter this information manually, using the following steps:

- Enter the number of required groups in the *'No. of Groups'* box. This will then prompt the user to enter the required information for each group.
- Enter a short label and a long label for the 'All' category. For example, if the groups consist of all local authority districts in the North West, the 'All' group short label might be 'NW' and the long label might be North West. The short label should be no more than 8 characters long.

- Enter a short label and a long label for each group.

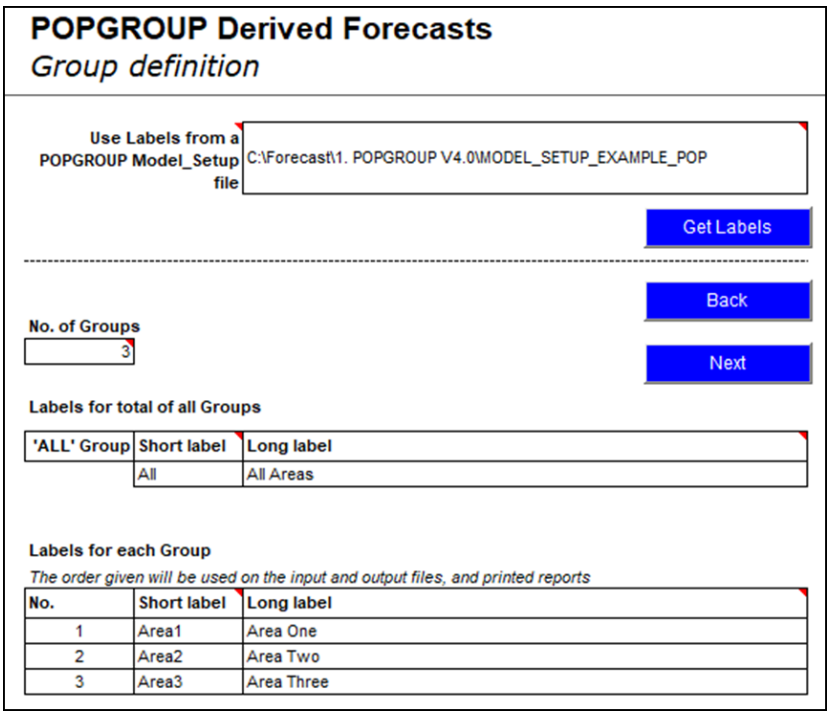


Figure 5: DF model setup: Step 2

This completes the 'mandatory' stages of the DF setup process. The user may use the *'Back'* button to return to Step 1 or, if they are satisfied with the details provided, use the *'Next'* button to progress to the optional stages of the setup process where the user has the opportunity to alter existing model options or define bespoke settings.

6.2.3. Step 3 – Bespoke Settings (Optional)

The DF model provides considerable flexibility in the setup process to enable the user to create a bespoke model configuration (e.g. disability model) or to modify the configuration of an existing model (e.g. 'CLG 2014-based Household').

If you have chosen one of the existing 'model types' in Step 1 of the `DFSetup` and do not require any changes to be made to the pre-defined settings, you can proceed and *'Run Setup'*.

However, if you have chosen a 'User Defined' model type (i.e. to run a bespoke model setup) or wish to modify any of the pre-defined settings (i.e. selected a model type); you can choose your bespoke settings here. Further detail on the available options and how to select them is described in Chapter 10 of this manual.

6.2.4. Step 4 – Running DFSetup

Once you have completed the setup and confirmed that the model type and 'Base Year' are correct, click the blue button *'Run Setup'* to run the setup process:

- The setup process will perform a number of checks to confirm the validity of the entries on the setup workbook. The process will warn of any changes that need to be made.
- The setup process will provide warnings where any existing files will be overwritten with the user given the choice of continuing or altering the entries in the `DFSetup` workbook.
- The `DFSetup` file will be saved automatically with the name `DFSetup_<Model ID>`, after a successful model setup.

The setup process creates folders and workbooks that the user will edit prior to running the model (see Chapter 7 on 'Preparing Assumptions'). When the `DFSetup` or model program is running, its progress is monitored in the bottom left-hand corner of the screen. In earlier versions of Excel, the progress may be monitored in a pop-up window in the centre of the screen.

The following table provides a summary of the structure and content of each of the skeleton files (in the folder location specified in Step 1, Figure 4) and describes the data that is required to transform each to an 'input' file. A distinction is made between the skeleton/input files that are mandatory and optional.

Table 1: Skeleton files

	Workbook Name	Purpose
Mandatory	DFPop	To store population forecast by age groups and sex for each group by year. Note that although this skeleton file is produced each time the model setup is run, it can be substituted in a DFScenario with a compatible POPGROUP forecast file (fore_) by single year of age and sex. To be compatible, the POPGROUP file must include worksheets in the same order and data for the same set of groups (areas) being modelled.
	DFRates	To store derived category rates by age group, sex, year and, optionally, sub-population type.
Optional	DFPopAdjust	To store population by age group and sex for each group by year. This population is removed or added prior to the application of rates (e.g. institutional population when applying household headship rates).
	DFSubPop	To store sub-populations. This allows age-group and sex categories by year to be sub-divided (e.g. into 'student' and 'non-student' categories).
	DFFactor	To specify a further size factor which can be applied to the derived forecasts disaggregated by category and, optionally, group (e.g. average household size in the WG household projections).
	DFCons	Allows the flexible application of constraints to the derived forecasts which can be provided for combinations of age group, sex, sub-population and category or for their total. For example, a model for small areas can be constrained to sum to an independent forecast for a larger region. The derived category rates are adjusted automatically to meet the constraint.
Scenario Setup	DFScenario	To record the location of all input workbooks required by the model program, and to run the model.

7 Preparing Assumptions

7.1. Introduction

Before running DF scenarios, the user must create a set of 'input' workbooks. Each skeleton workbook created during model setup (see Table 1) provides a template for the data that can be used within the model. The skeleton workbooks must be populated with a valid set of data (and then saved as input workbooks) before the forecast program can be run.

Some of the workbooks are optional, as are some of the data input options within individual workbooks. The strategy for preparing and developing assumptions for a derived category forecast will vary depending on the local circumstances, data availability and objectives of a particular forecast.

This chapter outlines some general information about the layout of the workbooks and provides notes on how they should be completed.

It is usual practice to use input file names that extend their skeleton name so that their content is clear. For example, once the `DFPop` skeleton file is populated with data, it is saved in the input folder as `DFPop_[Identifier]`.

!	**The *'Notes'* Worksheet**
Each of the input workbooks has a *'Notes'* worksheet onto which the user may optionally enter details of the assumptions used and data sources. When a scenario is run (see next chapter), the information provided in these worksheets is added to the *'Notes'* worksheet of the forecasts output workbooks. On validation (see section 5.4.2), any errors or warnings are reported on the *'Notes'* worksheet.	

7.2. The 'DF Pop' Workbook: Population Forecast

Status: *Mandatory* - unless the user decides to name an existing POPGROUP forecast output file (`fore_`) when listing the assumptions in DF scenario

The `DFPop` workbook is used to define a population forecast, disaggregated by age (and sex) to which a set of derived category rates can be applied (Figure 6). Population forecast data are required for each group (e.g. area) specified in the setup process, with each group represented as a separate worksheet in the skeleton/input file. There is no *'Default'* or *'All Groups'* worksheet in the `DFPop` file.

Derived Forecast - Households **EXAMPLE AREA**

Population forecasts on which the derived forecasts will be based

Area One

		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Male	0-4	20,793	20,191	19,860	20,059	20,466	20,995	21,752	22,461	23,239	24,108	24,513	25,295	25,855	26,250	26,188
Male	5-9	23,020	22,391	21,912	21,576	21,141	20,392	20,096	19,857	19,875	20,179	20,909	21,613	22,348	23,178	24,096
Male	10-14	24,492	23,593	22,948	22,490	22,497	22,334	21,979	21,637	21,346	20,967	20,367	20,040	19,778	19,884	20,120
Male	15-19	24,183	24,539	24,747	24,707	24,215	23,880	23,818	23,963	24,584	24,967	25,274	25,717	24,638	24,632	25,065
Male	20-24	28,846	30,276	32,014	34,869	37,682	37,364	37,295	36,782	35,416	35,054	35,320	36,574	37,179	37,554	37,848
Male	25-29	24,940	24,477	24,403	25,583	28,320	29,351	30,358	31,242	30,908	30,218	29,622	29,392	29,585	29,379	29,824
Male	30-34	26,697	26,454	25,950	25,992	26,420	26,025	25,455	25,385	26,024	26,820	27,089	27,397	27,744	27,701	27,581
Male	35-39	26,258	26,338	26,380	26,575	26,925	26,828	26,624	26,325	26,001	25,650	25,134	24,583	24,110	24,190	24,807
Male	40-44	23,626	23,957	24,198	24,542	25,586	25,838	26,121	26,251	26,261	26,019	26,112	25,647	25,441	25,039	24,463
Male	45-49	20,882	20,985	21,359	21,711	22,290	22,817	23,091	23,493	24,041	24,784	25,301	25,483	25,565	25,256	24,941
Male	50-54	22,299	21,239	20,402	20,025	19,946	19,910	20,194	20,535	20,794	21,274	21,773	22,235	22,741	23,270	23,986
Male	55-59	18,381	19,759	20,405	20,741	20,907	21,001	20,043	19,418	19,230	18,970	19,005	19,328	19,507	19,761	20,222
Male	60-64	15,893	15,595	15,580	15,755	16,096	16,639	17,872	18,594	19,100	19,375	19,540	18,654	18,234	17,946	17,718
Male	65-69	14,663	14,610	14,588	14,618	14,368	13,768	13,544	13,705	13,958	14,299	14,871	16,243	16,915	17,376	17,627
Male	70-74	11,974	11,978	12,039	12,036	12,148	12,257	12,250	12,354	12,398	12,281	11,912	11,717	11,902	12,252	12,545
Male	75-79	9,353	9,248	9,242	9,250	9,248	9,304	9,419	9,498	9,570	9,740	9,863	10,039	10,225	10,221	10,152
Male	80-84	5,869	6,109	6,308	6,403	6,242	6,181	6,256	6,344	6,399	6,513	6,615	6,655	6,749	6,864	7,038
Male	85+	2,713	2,665	2,583	2,541	2,820	3,036	3,116	3,211	3,262	3,313	3,453	3,589	3,601	3,678	3,770
Female	0-4	19,912	19,499	19,171	19,141	19,462	19,931	20,661	21,593	22,455	23,177	23,618	24,248	24,486	24,694	24,575
Female	5-9	21,906	21,331	20,942	20,580	20,173	19,617	19,263	19,039	19,013	19,427	19,936	20,572	21,451	22,222	22,914
Female	10-14	23,439	22,721	21,967	21,445	21,462	21,302	20,993	20,660	20,385	20,041	19,595	19,270	19,166	19,147	19,500
Female	15-19	24,437	24,543	24,812	24,492	24,153	23,699	23,817	23,943	24,792	25,354	25,644	26,174	25,029	25,050	25,694
Female	20-24	29,858	33,025	35,753	37,639	39,639	39,366	39,119	38,800	37,601	36,740	36,567	37,687	38,458	38,911	39,111
Female	25-29	25,588	24,842	24,758	25,494	27,539	29,099	30,259	31,282	31,059	30,753	29,982	29,464	29,085	28,669	28,590
Female	30-34	28,147	27,878	27,373	26,685	26,324	25,573	24,873	24,814	25,252	26,026	26,705	27,418	28,145	28,342	28,404
Female	35-39	27,680	27,884	27,499	27,347	27,213	27,119	26,895	26,551	25,794	25,308	24,646	24,083	23,869	24,111	24,753
Female	40-44	24,890	25,167	25,418	25,823	26,268	26,394	26,525	26,268	26,264	26,054	26,139	26,032	25,729	24,989	24,518
Female	45-49	21,026	21,420	21,926	22,400	23,081	23,587	23,914	24,381	24,792	25,167	25,344	25,413	25,417	25,486	25,287
Female	50-54	22,997	21,897	20,989	20,418	20,092	20,076	20,482	20,912	21,329	21,888	22,453	22,828	23,377	23,845	24,404
Female	55-59	18,629	20,140	20,779	21,268	21,530	21,663	20,577	19,855	19,603	19,305	19,352	19,728	20,147	20,637	21,251
Female	60-64	17,134	16,654	16,609	16,565	16,652	17,103	18,586	19,436	19,982	20,390	20,535	19,432	18,923	18,663	18,315
Female	65-69	15,833	15,948	16,026	16,120	16,023	15,584	15,178	15,146	15,282	15,410	15,907	17,500	18,257	18,778	19,167
Female	70-74	15,272	15,110	14,866	14,432	14,353	14,314	14,359	14,469	14,626	14,523	14,046	13,835	13,870	14,009	14,169
Female	75-79	13,226	12,934	12,859	12,841	12,795	12,835	12,767	12,657	12,365	12,394	12,499	12,534	12,712	12,892	12,781
Female	80-84	9,847	10,203	10,494	10,561	10,239	9,930	9,781	9,731	9,703	9,715	9,818	9,854	9,871	9,810	9,907
Female	85+	6,266	6,075	5,637	5,447	5,698	6,020	6,260	6,419	6,442	6,360	6,188	6,246	6,286	6,386	6,396

Figure 6: Example model: `DFPop`

The `DFPop` file may be populated with data directly from a previously created POPGROUP population forecast. The *'Notes'* worksheet of the `DFPop` file provides the user with the option to locate an appropriate POPGROUP file (`fore_`) and populate the individual *'Group'*

worksheets with population forecast data (Figure 6). Simply locate the POPGROUP file using browser, then select *'Load POPGROUP forecast'*.

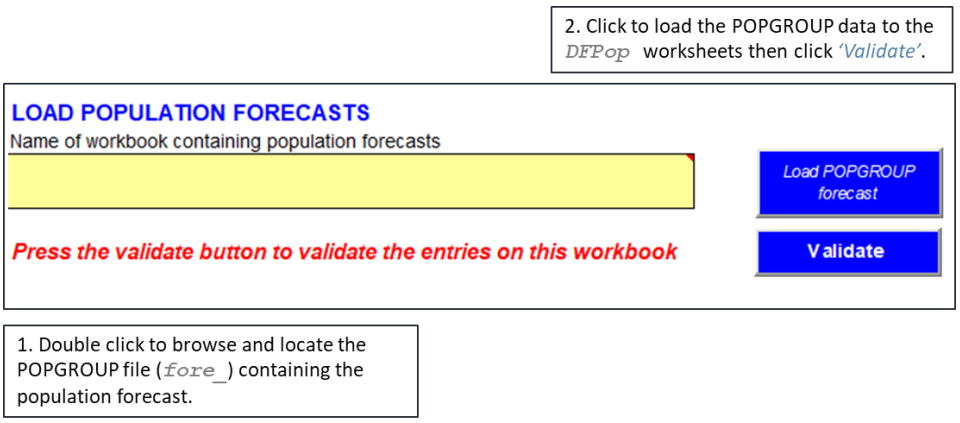


Figure 7: Loading a POPGROUP forecast to a DFPop file

If existing population forecast data from POPGROUP are not available to the user, then data must be entered manually in the format specified in the individual *'Group'* worksheets of DFPop.

Once data entry is complete, the skeleton DFPop should be saved as an input file with an appropriate name in the designated 'input' folder.

7.2.1. Notes and Validation

It is recommended that the source of the data is recorded on the *'Notes'* worksheet.

When all data has been entered into the *'Group'* worksheet, the user validates the data entry by clicking the *'Validate'* button. If there are any errors, these will be reported on the *'Notes'* worksheet and included in an accompanying diagnostics table (Figure 8). The user will be expected to correct errors prior to a scenario run of DF. The presence of warnings will not prevent the model from running but may affect the model outcome.

Diagnostics: summary of population options

	Data taken from:				
Group Name	Error (Base Year missing)	From Group	From Constant	From Interp	
Area1	0	1540	224	0	Area One
Area2	0	1540	224	0	Area Two
Area3	0	1540	224	0	Area Three

List of errors and warnings from validation

WARNING: One or more empty cells on Group sheet <Area1>; values will be held constant from 2018

WARNING: One or more empty cells on Group sheet <Area2>; values will be held constant from 2018

Figure 8: Example of validation checks on DFPop

Rules for data entry on `DFPop` workbook:

Rule	Comments
A population value is required in the base year for each age-sex combination, for each group.	If this rule is not followed, an **error** is returned and reported on the *'Notes'* sheet of the workbook.
If data are provided for the base year only, then the DF will assume all subsequent years have the same data values.	A note of this will be recorded on the *'Notes'* sheet.
If additional data are provided for later (but not contiguous) years, the DF will interpolate for intervening years.	A note of this will be recorded on the *'Notes'* sheet.
If data are provided up to but not beyond some years, then DF will assume that all subsequent years have the same data as the last year provided.	A note of this will be recorded on the *'Notes'* sheet.
Data values must be greater than or equal to zero.	If this rule is not followed, the user will be notified of the **error**.

'Error' indicates that the forecast program will terminate. If this happens, the error should be corrected and validation re-run before continuing.

The validation process is also repeated at run time, although checks will only be made on those years to be included in the forecast. A DF model will terminate if validation checks fail. The user will again be expected to correct any errors prior to a subsequent scenario run of the DF.

7.3. The 'DFPopAdjust' Workbook: Population Adjustment

Status: *Optional* – required only if specified at model setup

The `DFPopAdjust` file contains data on the population that is to be removed or added prior to the application of the derived category rates (e.g. population in communal households). The workbook contains a worksheet for each group plus a *'Default'* worksheet.

The user must enter population adjustments for each age and sex category. The age-sex groups for the population adjustment must be the same or more detailed than the age-sex groups for the derived units. In other words, the age-sex groups for the population adjustment must nest within the age-sex groups for the derived units.

Population adjustments may be actual values, percentage values or a mixture of both. They may also be either added to or subtracted from the population forecast or a mixture of both. The user is able to define these options at model setup, but may also change these choices for any age-sex category on `DFPopAdjust` using the *'Reformat'* option.

Derived Forecast - Households **EXAMPLE AREA**

Trend of population adjustments (Not in Households), from base year

Area One

Add or Subtract	Values (N or %)	Reformat		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Subtract	N	Male	0-4	10	11	12	13	14	16	17	19	21	23	25
Subtract	N	Male	5-9	10	10	10	10	10	10	10	11	11	12	12
Subtract	N	Male	10-14	115	105	97	92	88	85	80	75	70	65	59
Subtract	N	Male	15-19	1,346	1,585	1,805	2,001	2,165	2,327	2,530	2,744	3,036	3,322	3,606
Subtract	N	Male	20-24	789	1,011	1,271	1,625	1,997	2,219	2,433	2,611	2,690	2,820	2,989
Subtract	N	Male	25-29	341	360	376	420	490	534	576	626	646	654	656
Subtract	N	Male	30-34	255	261	261	270	283	287	287	295	311	329	339
Subtract	N	Male	35-39	226	233	241	250	259	264	268	271	273	276	276
Subtract	N	Male	40-44	182	187	191	196	208	211	216	220	223	224	226
Subtract	N	Male	45-49	119	122	126	131	137	143	146	151	158	165	171
Subtract	N	Male	50-54	108	103	100	99	99	100	101	104	105	107	111
Subtract	N	Male	55-59	98	103	103	103	101	100	93	88	84	80	78
Subtract	N	Male	60-64	78	75	75	75	75	75	79	82	82	81	80
Subtract	N	Male	65-69	70	70	72	74	74	72	72	74	78	81	86
Subtract	N	Male	70-74	131	127	124	120	117	115	111	109	105	100	95
Subtract	%	Male	75-79	2.0%	2.0%	1.9%	1.8%	1.7%	1.6%	1.6%	1.5%	1.4%	1.3%	1.3%
Subtract	%	Male	80-84	3.8%	3.6%	3.5%	3.4%	3.3%	3.1%	3.0%	2.9%	2.7%	2.6%	2.4%
Subtract	%	Male	85+	10.9%	10.5%	10.2%	9.7%	9.4%	9.0%	8.6%	8.2%	7.8%	7.4%	7.0%
Subtract	N	Female	0-4	9	10	11	12	13	14	16	18	20	22	23
Subtract	N	Female	5-9	10	10	10	10	10	10	10	10	10	11	11
Subtract	N	Female	10-14	64	61	58	57	56	55	54	52	51	49	48
Subtract	N	Female	15-19	1,193	1,534	1,898	2,221	2,551	2,863	3,220	3,584	4,021	4,408	4,792
Subtract	N	Female	20-24	508	753	1,016	1,283	1,573	1,793	2,015	2,234	2,390	2,546	2,737
Subtract	N	Female	25-29	88	96	107	122	144	166	188	210	225	237	246
Subtract	N	Female	30-34	74	73	71	68	66	63	60	60	60	62	63
Subtract	N	Female	35-39	69	68	65	63	59	57	55	53	49	47	44
Subtract	N	Female	40-44	72	70	69	66	65	62	59	55	51	47	43
Subtract	N	Female	45-49	70	70	68	68	67	65	63	59	57	53	49
Subtract	N	Female	50-54	53	51	48	47	47	47	48	48	50	51	51
Subtract	N	Female	55-59	77	81	82	80	77	75	68	62	58	52	48
Subtract	N	Female	60-64	70	66	63	62	60	59	62	62	60	58	56
Subtract	N	Female	65-69	96	94	92	89	87	82	78	75	74	72	73
Subtract	N	Female	70-74	180	174	167	157	152	147	144	140	137	132	124
Subtract	%	Female	75-79	2.6%	2.5%	2.4%	2.3%	2.2%	2.1%	2.1%	2.0%	1.9%	1.8%	1.7%
Subtract	%	Female	80-84	6.2%	6.1%	5.9%	5.8%	5.7%	5.5%	5.4%	5.3%	5.2%	5.1%	5.0%
Subtract	%	Female	85+	20.2%	19.7%	19.2%	18.7%	18.2%	17.7%	17.2%	16.7%	16.1%	15.6%	15.1%

Figure 9: Example model: `DFPopAdjust`

7.3.1. Notes and Validation

It is recommended that the source of the data is recorded on the `'Notes'` worksheet.

When all data has been entered into the `'Group'` worksheet(s), the user validates the data entry by clicking the *'Validate'* button. If there are any errors, these will be reported on the `'Notes'` worksheet and included in an accompanying diagnostics table (Figure 10). The user will be expected to correct any errors prior to a scenario run of DF. The presence of warnings will not prevent the model from running but may affect the model outcome.

Diagnostics: summary of adjustment options

	Base population adjustment from :					*Forecast popn adjustment from :*				
Group Name	Group, N	Group, %age	Default, N	Default, %age	Error: missing or < 0 or % >100	Group	Default	Constant	Interp	
Area1	30	6	0	0	0	612	0	1152	0	Area One
Area2	30	6	0	0	0	612	0	1152	0	Area Two
Area3	30	6	0	0	0	612	0	1152	0	Area Three

List of errors and warnings from validation

WARNING: One or more empty cells on Group sheet <Area1>; values will be held constant from 2018
WARNING: One or more empty cells on Group sheet <Area2>; values will be held constant from 2018
WARNING: One or more empty cells on Group sheet <Area3>; values will be held constant from 2018

Figure 10: Example of validation checks on `DFPopAdjust`

Rules for data entry on `DFPopAdjust` workbook:

Rule	Comments
A data value is required in the base year for each age-sex combination, either on each `'Group'` worksheet or on the `'Default'` worksheet.	If this rule is not followed, an **error** is returned and reported on the `'Notes'` sheet of the workbook.
If data are provided for the base year only, then the DF will assume all subsequent years have the same data values.	A note of this will be recorded on the `'Notes'` sheet.
If additional data are provided for later (but not contiguous) years, the DF will interpolate data for intervening years.	A note of this will be recorded on the `'Notes'` sheet.
If data are provided up to but not beyond a specific year, then DF will assume that all subsequent years have the same data as the last year provided.	A note of this will be recorded on the `'Notes'` sheet.

Rule	Comments
If the `'Group'` worksheet has no value (after interpolation), but a value on the `'Default'` sheet is present, then it will be used for the group. A default percentage value (*'%'*) is used for each group not specified on its own sheet, while counts (*'N'*) are shared to each group in proportion of its projected population for this age-sex combination.	The adjustments will be recorded in the `DFRiskPop` output file.
In the *'Values (N or %)'* column, the user may only choose to enter counts (*'N'*) or percentage values (*'%'*). The *'%'* refers to a percentage of the area's population of this age-sex combination.	Changing the *'Values (N or %)'* will require the user to select the *'Reformat'* option to change the data entry format.

The validation process is also repeated at run time, although checks will only be made on those years to be included in the forecast. An additional validation is performed, comparing population data from the `DFPop` file with population adjustment from the `DFPopAdjust` file to ensure that when the population adjustments is applied to the original population, it does not result in population less than zero in any age-sex category.

A DF model will terminate if validation checks fail. The user will be expected to correct any errors prior to a subsequent scenario run of the DF.

7.4. The 'DFSubPop' Workbook: Sub-Populations

Status: *Optional* – required only if specified at model setup

`DFSubPop` provides data on the population in the selected sub-population(s). It allows age-group and sex categories by year to be sub-divided, for example, into 'student' and 'non-student' categories or into marital status categories. The option is rarely used because these extra dimensions of a population forecast are rarely available, but it is available because government and academic work has sometimes made these distinctions.

The workbook contains a worksheet for each group plus a `'Default'` worksheet.

The user must enter a count (*'N'*) or percentage value (*'%'*) (as indicated in the *'Values (N or %)'* column) for each age, sex and sub-population. It is possible to enter data either for one sub-population (e.g. students) and to then assume that the remainder of the population constitutes the other sub-population (e.g. non-students).

The age categories match those of the `DFRates` workbook (Figure 11).

Derived Forecast - Households **EXAMPLE AREA**

Trend of sub-populations (Student Population), from base year

Area One

Reformat

Values (N or %)				2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
N	Persons	15-24	Students	17,148	18,138	19,478	19,836	20,025	19,538	19,049	18,743	19,085	19,517	18,882
N	Persons	15-24	Non-Students											
N	Persons	25-34	Students	11,432	6,046	6,493	6,612	6,675	6,513	6,350	6,248	6,362	6,506	6,294
N	Persons	25-34	Non-Students											
N	Persons	35-44	Students	1,143	3,023	3,246	3,306	3,338	3,256	3,175	3,124	3,181	3,253	3,147
N	Persons	35-44	Non-Students											

Figure 11: Example model: `DFSubPop`

7.4.1. Notes and Validation

It is recommended that the source of the data is recorded on the *'Notes'* worksheet.

When all data have been entered into the *'Group'* worksheet(s), the user validates the data entry by clicking the *'Validate'* button. If there are any errors, these will be reported on the *'Notes'* worksheet and included in an accompanying diagnostics table (Figure 12). The user will be expected to correct errors prior to a scenario run of DF. The presence of warnings will not prevent the model from running but may affect the model outcome.

Diagnostics: summary of Sub-Population options

	Base sub-pops data from :					Forecast sub-pops data from :				
Group Name	Group, N	Group, %age	Default, N	Default, %age	Error: missing or < 0 or % >100	Group	Default	Constant	Interp	
Area1	10	0	0	0	0	100	0	390	0	Area One
Area2	10	0	0	0	0	100	0	390	0	Area Two
Area3	10	0	0	0	0	100	0	390	0	Area Three

If leaving one Sub-Population blank - provide its name: Non-Students

Press the validate button to validate the entries on this workbook Validate

List of errors and warnings from validation

WARNING: Sub-population data for sub-population <Non-Students> has not been set and will be computed as the remaining population
WARNING: One or more empty cells on Group sheet <Area1>; values will be held constant from 2011
WARNING: One or more empty cells on Group sheet <Area2>; values will be held constant from 2011
WARNING: One or more empty cells on Group sheet <Area3>; values will be held constant from 2011

Figure 12: Example of validation checks: `DFSubPop`

Rules for data entry on `DFSubPop` workbook:

Rule	Comments
A data value is required in the base year for each age-sex combination, either in each *'Group'* worksheet or in the '*Default'* worksheet.	If this rule is not followed, an **error** is returned and reported on the *'Notes'* sheet of the workbook.
If data are provided for the base year only, then the DF will assume all subsequent years have the same data values.	A note of this will be recorded on the *'Notes'* sheet.
If additional data are provided for later (but not contiguous) years, the DF will interpolate data for intervening years.	A note of this will be recorded on the *'Notes'* sheet.
If data are not provided for a later year (after interpolation) but a value is available on the *'Default'* sheet, then it will be used, as follows: • If the value is provided as a percentage value (*'%'*), it is multiplied by the groups' adjusted population at that age-sex group. • If the value is provided as counts (*'N'*), it is shared to each group in proportion of its projected population for this age-sex combination.	The adjustments will be recorded in the `DFRiskPop` output file.
If data are provided up to but not beyond a specific year, then DF will assume that all subsequent years have the same data as the last year provided.	A note of this will be recorded on the *'Notes'* sheet.
In the *'Values (N or %)'* column, the user may only choose to enter counts (*'N'*) or percentage values (*'%'*). The *%* refers to a percentage of the area's population of this age-sex combination	Changing the *'Values (N or %)'* will require the user to select the *'Reformat'* option to change the data entry format.

The validation process is also repeated at run time, although checks will only be made on those years to be included in the forecast. An additional validation will take place which aligns the sub-populations with the population forecast in `DFPop` to check for consistency. Any inconsistencies will be reported at run-time and the user is given the choice to continue or abort the model run.

A DF model will terminate if validation checks fail. The user will be expected to correct any errors prior to a subsequent scenario run of the DF.

7.5. The 'DFRates' Workbook: Derived Category Rates

Status: *Mandatory*

`DFRates` contains data on the 'rates' which are to be applied to the population forecasts to produce the derived category forecasts. The workbook contains a worksheet for each group plus a `'Default'` worksheet.

The user must enter rates for each age and sex, for each derived category and for any sub-population that has been (optionally) defined. The age-sex categories are those defined in the model setup. They may be less detailed than the population adjustment. In the example used in this manual, the rates in Figure 13 are specified in the categories by broad age group and with no sex disaggregation, while the population adjustment (Figure 9) is specified by 5-year age group and sex.

Rates must be entered as percentage values. An optional 'annual increment' may be applied, either on a group worksheet or the `'Default'` worksheet (with the group worksheet taking precedence if values are found on both). This increment is applied to each year, either by addition or multiplication, as specified on the `'Notes'` sheet of the `DFRates` workbook.

Derived Forecast - Households **EXAMPLE AREA**

Trend of CLG Stage 2 HHld rep rates, from base year

Area One

	Category	Annual Increment	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons 0-14	One person male		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	One person female		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Couple no child		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Cple+adlts no child		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	One child		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Two children		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Three+ children		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Other households		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Persons 15-24	One person male		3.8%	3.8%	3.7%	3.7%	3.6%	3.6%	3.5%	3.4%	3.3%	3.2%	3.1%
	One person female		4.6%	4.4%	4.2%	4.1%	3.9%	3.7%	3.5%	3.4%	3.2%	3.0%	2.8%
	Couple no child		2.1%	2.1%	2.2%	2.2%	2.3%	2.3%	2.3%	2.3%	2.4%	2.4%	2.4%
	Cple+adlts no child		0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.4%	0.4%	0.4%	0.4%	0.4%
	One child		2.7%	2.7%	2.8%	2.8%	2.7%	2.8%	2.8%	2.9%	3.0%	3.1%	3.2%
	Two children		0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%	1.0%	1.0%	1.0%
	Three+ children		0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
	Other households		5.0%	5.3%	5.5%	5.7%	6.0%	6.2%	6.4%	6.6%	6.7%	6.8%	7.0%
Persons 25-34	One person male		7.8%	8.0%	8.1%	8.2%	8.4%	8.5%	8.5%	8.5%	8.5%	8.5%	8.6%
	One person female		5.3%	5.3%	5.4%	5.5%	5.6%	5.6%	5.6%	5.6%	5.6%	5.6%	5.6%
	Couple no child		11.6%	11.7%	11.8%	11.8%	12.0%	11.9%	11.9%	11.8%	11.7%	11.6%	11.5%
	Cple+adlts no child		0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%
	One child		9.6%	9.3%	9.1%	8.8%	8.5%	8.4%	8.3%	8.3%	8.3%	8.3%	8.3%
	Two children		8.4%	8.0%	7.7%	7.4%	7.0%	6.8%	6.6%	6.5%	6.4%	6.4%	6.3%
	Three+ children		4.4%	4.2%	4.1%	4.0%	3.8%	3.7%	3.7%	3.6%	3.6%	3.6%	3.6%
	Other households		3.7%	3.9%	4.0%	4.2%	4.3%	4.4%	4.5%	4.6%	4.7%	4.8%	4.8%

Figure 13: Example model: `DFRates`

7.5.1. Notes and Validation

It is recommended that the source of the data is recorded on the *'Notes'* worksheet.

When all data have been entered into the *'Group'* worksheet(s), the user validates the data entry by clicking the *'Validate'* button. If there are any errors, these will be reported on the *'Notes'* worksheet and included in an accompanying diagnostics table (Figure 14). The presence of warnings will not prevent the model from running but may affect the model outcome.

Value of each rate must be: ***From 0-1 inclusive***
Sum of rates across all categories: ***Sum to 1 or Less***

Diagnostics: summary of rates options and list of population group names used on the files of activity rates

	Base rates taken from:			*Forecast rates taken from:*				*Annual increment taken from:*		*Warnings*		
Group Name	Group	Default	Error (missing)	Group	Default	Constant	Interp	Group	Default	Rates, or rates sum, starting outside allowed range	Rates, or rates sum, outside allowed range after increments and trending	
Area1	80	0	0	3040	0	880	0	0	0	0	0	Area One
Area2	80	0	0	3040	0	880	0	0	0	0	0	Area Two
Area3	80	0	0	3040	0	880	0	0	0	0	0	Area Three

Press the validate button to validate the entries on this workbook.

Validate

List of errors and warnings from validation

WARNING: One or more empty cells on Group sheet <Area1>; values will be held constant from 2039
WARNING: One or more empty cells on Group sheet <Area2>; values will be held constant from 2039
WARNING: One or more empty cells on Group sheet <Area3>; values will be held constant from 2039

Figure 14: Example of validation checks on `DFRates`

Rules for data entry on `DFRates` workbook:

Rule	Comments
A data value is required in the base year for each age, sex, derived category and (optionally) sub-population combination, either in each group worksheet or on the *'Default'* worksheet. For each year after the base year: • If there is a value on the *'Group'* sheet, this will be used. • If additional data are provided for later (but not contiguous) years, the DF will interpolate data for intervening years. However, this only applies when all values in one category are not defined. • If only the *'Default'* sheet has values, then the group's previous year value will be incremented by applying the change from the previous year on the *'Default'* sheet. This will be applied by addition or multiplication according to the choice on the *'Notes'* sheet. If neither the *'Group'* sheet nor the *'Default'* sheet has values (after interpolation), the values on the *'Group'* sheet will be held constant from the previous year.	A note of this will be recorded on the *'Notes'* sheet.

Addition or multiplication is set on the *'Notes'* sheet. The annual increment and the *'Default'* rates are applied according to this setting.
• Annual increment (e.g. 1%) – either an addition of 1% each year, or a multiplication of 1.01 each year.
• *'Default'* sheet – either:
  • an *addition* to the previous year's group value of (*Default* value year y – *Default* value year y-1), or;
  • a *multiplication* of the previous year's group (*Default* value year y)/
	A note of this will be recorded on the *'Notes'* sheet.

Rule	Comments
(`Default` value year y-1). The `'Default'` sheet is often used to apply the trend of projected change in a reference area (national or regional) to a local area for which the base year value has been provided.	
If the annual increment or the application of `'Default'` values results in a rate value that is sat outside the bounds set in the original model setup process (e.g. from 0 to 1), then the rate is changed to the nearest valid value (0 or 1).	A warning is provided after validation and a note of this will be recorded on the `'Notes'` sheet.
If the sum of rates is outside the bounds set in the original model setup process (e.g. sum to 1), the rates are scaled to meet the bounds.	The rates are scaled to meet the bounds and a warning provided on the `'Notes'` sheet after validation.

An annual increment will be applied each year from the base year and will overwrite any values provided for later years. It is not used in combination with a trend.

If the user attempts to enter invalid data into individual cells they will be prompted by an error message.

The validation process is also repeated at run time, although checks will only be made on those years to be included in the forecast. A DF model will terminate if validation checks fail. The user will be expected to correct any errors prior to a subsequent scenario run of the DF.

7.6. The 'DFFactor' Workbook: Size Factors

Status: *Optional* – required only if specified at model setup

`DFFactor` contains a 'size factor' which can be applied to the derived forecasts disaggregated by category and group (area). For example, this would be applied to derive household numbers from household populations, as is the case with the WG household projection models.

This workbook contains a worksheet for each group plus a `'Default'` worksheet. The user must enter size factors for each derived category and group.

Derived Forecast - Households **EXAMPLE AREA**

Calculations on rates (Average Household Size), from base year

Area One

Category	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1 person	1.00											
2 person (No children)	2.00											
2 person (1 adult, 1 child)	2.00											
3 person (No children)	3.00											
3 person (2 adults, 1 child)	3.00											
3 person (1 adult, 2 children)	3.00											
4 person (No children)	4.00											
4 person (2+ adults, 1+ children)	4.00											
4 person (1 adult, 3 children)	4.00											
5+ person (No children)	5.46											
5+ person (2+ adults, 1+ children)	5.41											
5+ person (1 adult, 4+ children)	5.33											

Figure 15: Example model: `DFFactor`

7.6.1. Notes and Validation

It is recommended that the source of the data is recorded on the *'Notes'* worksheet.

When all data have been entered into the *'Group'* worksheet(s), the user validates the data entry by clicking the *'Validate'* button. The validation routine will check each year (column) of each *'Group'* worksheet and *'Default'* worksheet. If there are any errors, these will be reported on the *'Notes'* worksheet and included in an accompanying diagnostics table (Figure 16). The presence of warnings will not prevent the model from running but may affect the model outcome. The user will be expected to correct errors prior to a scenario run of DF.

Diagnostics: summary of factor options

	Base rates taken from:			*Forecast rates taken from:*				
Group Name	Group	Default	Error (missing or <0)	Group	Default	Constant	Interp	
Area1	12	0	0	0	0	588	0	Area One
Area2	12	0	0	0	0	588	0	Area Two
Area3	12	0	0	0	0	588	0	Area Three

Press the validate button to validate the entries on this workbook. Validate

List of errors and warnings from validation

WARNING: One or more empty cells on Group sheet <Area1>; values will be held constant from 2001
WARNING: One or more empty cells on Group sheet <Area2>; values will be held constant from 2001
WARNING: One or more empty cells on Group sheet <Area3>; values will be held constant from 2001

Figure 16: Example of validation checks on `DFFactor`

Rules for data entry on `DFFactor` workbook:

Rule	Comments
A data value is required in the base year for each derived category, either in each *'Group'* worksheet or the *'Default'* worksheet.	A note of this will be recorded on the *'Notes'* sheet.
If the data are provided for the base year only, DF will assume all subsequent years have the same data values.	A note of this will be recorded on the *'Notes'* sheet.
If the additional data are provided for later (but not contiguous) year, DF will interpolate data for intervening years.	A note of this will be recorded on the *'Notes'* sheet.
If data are provided up to but not beyond some years, then the DF will assume that all subsequent years have the same data as the last year provided.	A note of this will be recorded on the *'Notes'* sheet.
Values on the *'Default'* sheet are only used if the *'Group'* sheet has no values in the base year.	A note of this will be recorded on the *'Notes'* sheet.

The validation process is also repeated at run time, although checks will only be made on those years to be included in the forecast. A DF model will terminate if validation checks fail. The user will be expected to correct any errors prior to a subsequent scenario run of the DF.

7.7. The 'DFCons' Workbook: Constraints

Status: *Optional*

`DFCons` allows users to enter data which will constrain the derived forecasts. For **any single year** these constraints can either be:

- Total constraints
- Category constraints
- Sex-age-category constraints
- Sub-population constraint if the data are available.

The purpose of the **DFCons** file is to force the projection to agree with an independent projection that has already been generated but which has less detail than the user's model. For example, a projection for smaller areas may be forced to sum to a projection for a larger area.

There will be a worksheet for each group and an *'All Groups'* worksheet. It is possible to choose any constraint option on the *'All Groups'* worksheet or a *'Group'* worksheet **but not on both.**

The user specifies which constraint options are required in each year. The worksheet will then prompt for data to be entered into the appropriate cells in the worksheet.

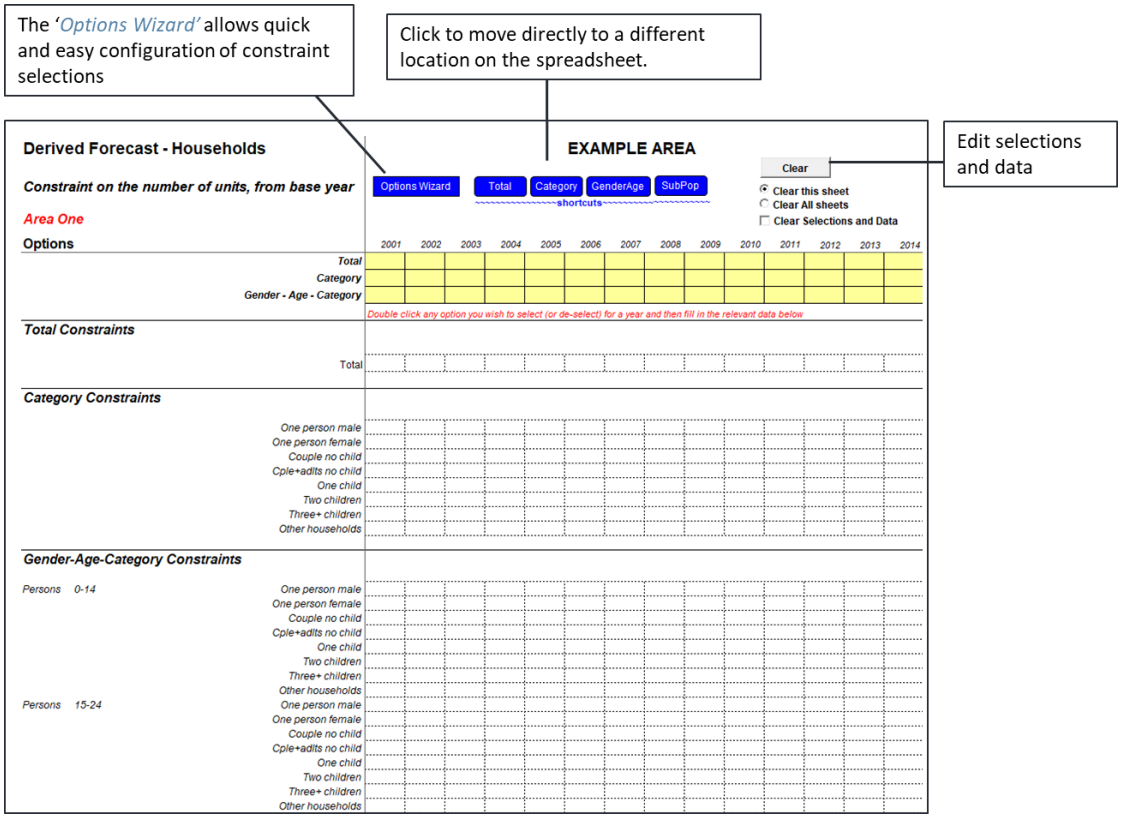


Figure 17: Example model: **DFCons**

The *'Options Wizard'* button provides the user with a flexible way of quickly entering constraint selections into the relevant parts of the worksheet (Figure 18). Data must be entered when selections have been made. Alternatively, the user can simply 'double-click' on the relevant cell to select an option and to enter data appropriately.

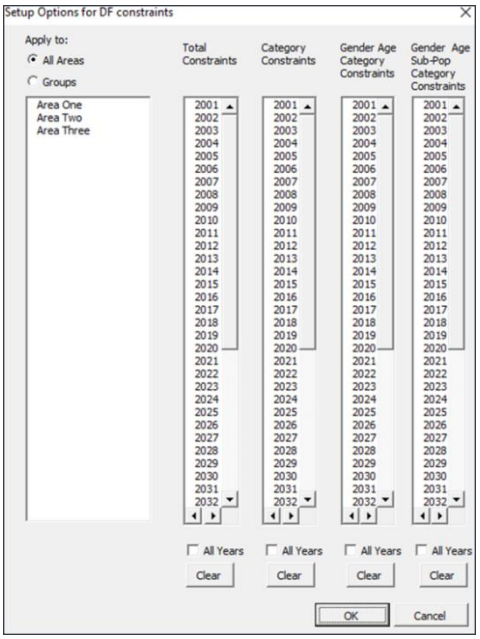


Figure 18: *'Options Wizard'* in `DFCons`

To assist with editing, a number of *'Clear'* functions are provided, allowing selections on a sheet or all sheets to be removed and to also remove both data and selections simultaneously.

Initial results for the derived units will be adjusted by scaling to the values given on the `DFCons` workbook. The user's forecast will be adjusted to agree in total or in detail with the constraints applied. The DF model will also rescale the derived unit rates to be consistent with the constrained forecast. The adjusted rates are reported on the `DFRatesOut` output file. Note that it is possible for this rescaling to result in rates that no longer adhere to the rules for derived unit rates specified on `DFSetup`.

7.7.1. Notes and Validation

It is recommended that the source of the data is recorded on the *'Notes'* worksheet.

When all data have been entered into the *'Group'* worksheet(s), the user validates the data entry by clicking the *'Validate'* button. The validation routine will check each year (column) of each group worksheet and *'All Groups'* worksheet. If there are any errors, these will be reported on the *'Notes'* worksheet and included in an accompanying diagnostics table (Figure 19). The user will be expected to correct errors prior to a scenario run of DF. If the user attempts to enter invalid data into individual cells they will be prompted with an error message.

When removing an option for which data has been provided, any data entry will remain but will be formatted grey and not used in the forecast. If the option is reinstated, the data will again appear in blue and the cell will be highlighted in yellow.

Summary of constraints chosen

	Number of years				
Group name	Total	Category	Gender - age - Category	Gender - age - sub-pop - Category	
Area1	0	0	0	0	Area One
Area2	0	0	0	0	Area Two
Area3	0	0	0	0	Area Three
All	0	0	0	0	All Areas

Press the validate button to validate the entries on this workbook. Validate

There are no errors or warnings from validation

Figure 19: Example of validation checks on DFCons

Rules for data entry on DFCons workbook:

Rule	Comments
In any year only one constraint option may be chosen on any sheet.	The user will be unable to select more than one option.
In any year a constraint may only be chosen on either the *'Group'* sheet or *'All Groups'* sheet – not both.	If the constraint is selected on the *'All Groups'* sheet, the options in the *'Group'* sheet will be formatted in grey and not used in the forecast.
Non-contiguous values for the same type of constraint will be interpolated where no other constraint types are used in the intervening years.	A note of this will be recorded on the *'Notes'* sheet.
Values will be held constant where contiguous blank cells for a selected constraint type follow one or more defined values and where no other constraint types are used in intervening years.	A note of this will be recorded on the *'Notes'* sheet.

The validation process is repeated at scenario run-time, although checks will only be made on those years to be included in the forecast. A DF model will terminate if validation checks fail with errors. The user will be expected to correct any errors prior to a subsequent scenario run of the DF.

8 Scenario Configuration

8.1. The 'DFScenario' Workbook

The user runs the derived forecasting model from the `DFScenario` workbook, created during model setup and found in the skeleton folder (Figure 20). The `DFScenario` workbook specifies the input files that are to be used.

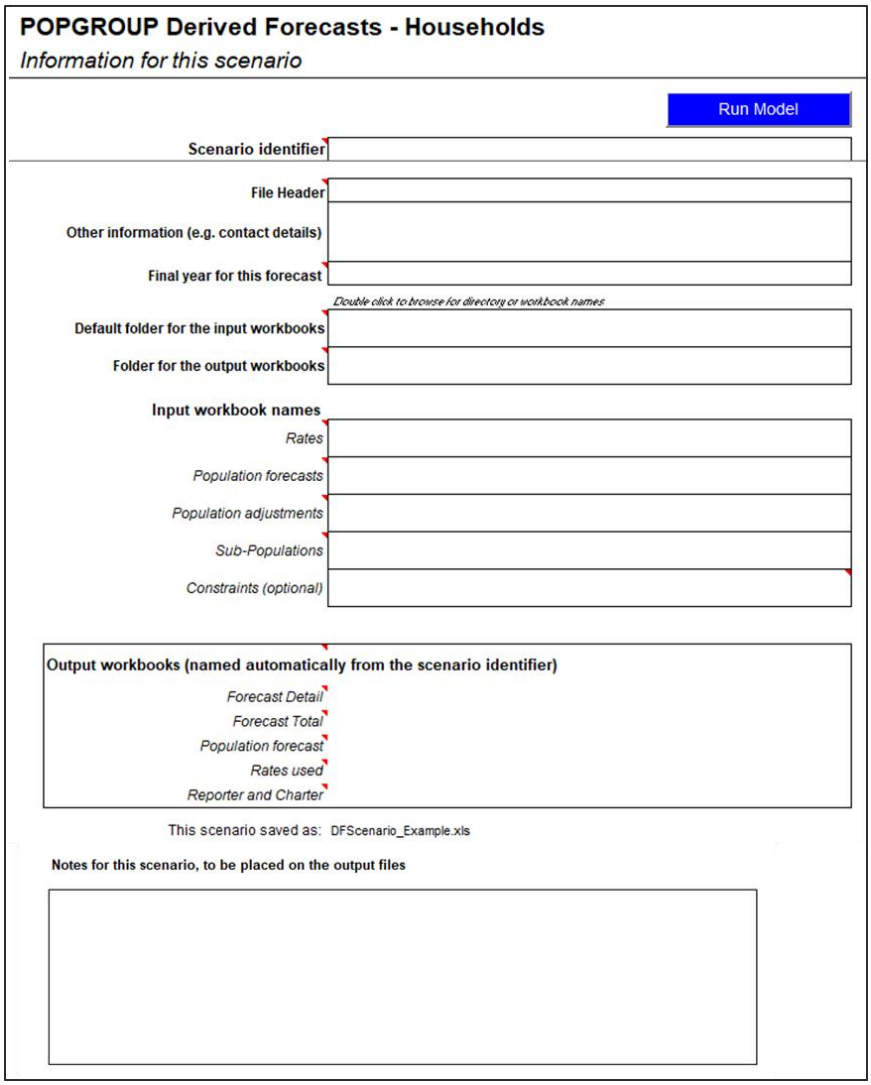


Figure 20: Example model: `DFScenario`

Scenario Identifier

A name for the scenario must be given, as the scenario ID defines a forecast and allows it to be replicated in the future. The completed version of the scenario workbook will be saved in the DF 'input' folder with this name if the scenario is run successfully. This scenario ID will also be appended as a suffix in the naming of the output files.

File Header

The file header will appear on each worksheet of the output files. This could be something that helps the user identify the project or the scenario.

Other Information

Additional information that will be printed on the output reports (optional).

Final Year for This Forecast

The forecasts will be provided up until and including the year specified here. The year given must be at least one greater than the base year and no further ahead than the final year of the population forecast.

Default Folder for the Input and Output Workbooks

Input and output workbooks will be assumed to be in the default folder, if there is no explicit path given with their name (see next item). These are prefilled with the folders that were named in the model setup. An entry is mandatory in these cells.

Input Workbook Names

The user must enter the names of the input files to be used in the scenario. If a workbook is not in the default input folder specified above, its name must include its full file path.

Notes

The user can include a note to summarise the assumptions, etc. that lie behind their scenario forecast run. The note will be reproduced on the output files.

8.2. Running the Model

The forecast model program is started by clicking the *'Run Model'* button at the top of the *'ScenarioSetup'* worksheet. The DF carries out the following checks before it produces the forecasts. If all required workbooks are present and contain no errors, the forecast is generated and a series of output workbooks is created in the specified output folder.

Rules for data entry on **DFScenario** workbook:

Rule	Comments
The final year chosen for the forecast needs to be within allowable bounds (i.e. at least one greater than the base year and no further ahead than the final year of the population forecast).	If NO, the program provides an on-screen message and then terminates.
The default input and output folders must exist in the specified locations.	If NO, the program provides an on-screen message and then terminates.
Each named workbook must exist in the specified location.	If NO, the program provides an on-screen message and then terminates.
All files used in the scenario must be consistent with each other, i.e. all files used in the scenario must have been created from the same DFSetup. Input files from different models cannot be mixed. The setting up of a DF model is very flexible; this restriction prevents mismatches in the dimensions of age-sex, categories, sub-populations and so on.	If NO, the program provides an on-screen message and then terminates.
All files are validated for content, regardless of whether the user has validated previously.	If any errors are found, the program provides an on-screen message and then terminates.

If any of the output workbooks with the same scenario ID already exist, the user will be asked whether they are to be overwritten.

The **DFScenario** workbook is automatically saved in the input folder when a forecast has run successfully. The scenario workbook will be saved using the scenario ID. This allows the user to run a forecast again, after amending the input workbooks, without repeating the entry of workbook names and documentation.

9 Model Results

9.1. Output File Summary

The output files are located in the output folder specified in the `DFScenario` file. Each file will have the scenario ID suffix that was previously defined:

Output workbook	Description
`DFForecastDetail_[Scenario ID]`	Full detail on the derived forecast counts by group and year – disaggregated by age group, sex, category and (optionally) sub-population.
`DFForecastTotal_[Scenario ID]`	Total derived forecasts by group, year and category.
`DFRatesOut_[Scenario ID]`	Derived forecast category rates used in the forecast by group and year – disaggregated by age group, sex, category and (optionally) sub-population.
`DFRiskPop_[Scenario ID]`	Population 'at risk' details by group and year – disaggregated by age group and sex and (optionally) sub-population.
`DFReporter_[Scenario ID]`	A utility for producing reports and charts from the scenario results, to user specification.

In addition, a utility within the `DFReporter_[Scenario ID]` file allows all output data to be written to a single Excel sheet. This `FlatDump` file is suitable for export to other software for further analysis.

A separate utility called `DFCompare` is provided that enables comparison of results from two or more scenarios. It can be found in the same folder as the `DFSetup` file.

A more detailed description of each of these files and their content is provided below.

9.2. The 'DFForecastDetail' Workbook: Detailed Results

The `DFForecastDetail` workbook provides detailed output for each group, including a disaggregation of derived category counts by age and sex for each year of the forecast (Figure 21). If sub-populations have been defined, these will be included as an additional dimension to the output.

An *'All Groups'* worksheet is included, which is a summation of the individual *'Group'* worksheets.

Derived Forecast - Households EXAMPLE

Households by gender, age group, sub-population and category

Area One

				2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	15-24	Students	One person male	0	0	0	0	0	0	0	0	0	0	0
			One person female	0	0	0	0	0	0	0	0	0	0	0
			Couple no child	0	0	0	0	0	0	0	0	0	0	0
			Cple+adlts no child	0	0	0	0	0	0	0	0	0	0	0
			One child	0	0	0	0	0	0	0	0	0	0	0
			Two children	0	0	0	0	0	0	0	0	0	0	0
			Three+ children	0	0	0	0	0	0	0	0	0	0	0
			Other households	181	198	229	261	276	304	332	332	311	317	342
Persons	15-24	Non-Students	One person male	95	101	107	106	116	121	123	119	119	126	125
			One person female	132	137	142	138	148	150	149	140	138	142	138
			Couple no child	140	141	141	133	139	136	132	121	114	115	108
			Cple+adlts no child	56	58	60	59	63	64	64	60	59	61	59
			One child	187	188	190	180	188	186	182	172	172	177	170
			Two children	49	51	52	50	53	54	53	52	52	55	54
			Three+ children	16	16	16	14	14	14	13	11	11	11	10
			Other households	417	451	484	491	548	578	596	584	596	638	642
Persons	25-34	Students	One person male	104	103	100	107	118	126	131	134	131	125	124
			One person female	75	74	72	77	84	90	94	96	94	90	89
			Couple no child	192	190	187	200	221	236	247	254	250	240	239
			Cple+adlts no child	16	17	18	21	24	28	31	34	35	36	38
			One child	161	159	157	168	186	199	210	223	229	226	229
			Two children	144	139	133	139	149	156	160	165	165	159	157
			Three+ children	63	61	59	62	67	70	73	75	76	74	73
			Other households	54	57	60	68	79	89	98	106	109	110	114

Figure 21: Example model output: `DFForecastDetail_[Scenario ID]`

9.3. The 'DFForecastTotal' Workbook: Summary Results

The `DFForecastTotal` workbook provides a more aggregate version of the previous output, presenting derived category counts for each year of the forecast, summed over all age-sex categories (Figure 22). Again, if sub-populations have been defined, these will be included as an additional dimension to the output.

An *'All Groups'* worksheet is included, which is a summation of the individual *'Group'* worksheets.

Derived Forecast - Households **EXAMPLE**

Households by category

Area One

	Households											
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
One person male	4,684	4,760	4,842	4,970	5,116	5,243	5,357	5,436	5,495	5,546	5,637	5,701
One person female	6,936	6,914	6,905	6,927	6,973	6,979	6,994	6,950	6,910	6,861	6,842	6,833
Couple no child	10,866	10,819	10,789	10,834	10,888	10,928	10,955	10,913	10,850	10,764	10,756	10,725
Cple+adlts no child	3,007	3,021	3,051	3,098	3,147	3,193	3,222	3,241	3,251	3,269	3,307	3,313
One child	4,535	4,549	4,592	4,660	4,803	4,919	5,039	5,224	5,443	5,582	5,704	5,876
Two children	4,816	4,777	4,754	4,761	4,824	4,866	4,917	5,030	5,162	5,207	5,244	5,336
Three+ children	1,961	1,959	1,938	1,928	1,942	1,950	1,959	1,992	2,032	2,034	2,037	2,061
Other households	2,752	2,848	2,968	3,091	3,256	3,403	3,535	3,593	3,639	3,738	3,841	3,905

Figure 22: Example model output: `DFForecastTotal_[Scenario ID]`

If the size factors have been used from a `DFFactor` input file, then the `DFForecastTotal_[Scenario ID]` contains the final derived units *after* application of the factors, and the factors themselves, while `DFForecastDetail_[Scenario ID]` contains the derived units *before* application of the factors.

Note that in some rare cases it may be impossible for the model to calculate the 'factor' values. This may occur on the `'All Groups'` sheet where a zero derived forecast would lead to a division by zero and therefore a factor value cannot be calculated – even though there may have been a value for each individual `'Group'` sheet. These instances are marked in the output by an asterisk (*) accompanied by an explanatory note at the bottom of the data panel.

Derived Forecast - Households **Example**

Households by category

All Areas

	Households									
	Average Household Size									
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
1 person	69,181	70,567	72,098	73,693	75,259	76,866	78,497	80,188	82,043	83,660
2 person (No children)	72,150	72,510	73,010	73,563	73,988	74,377	74,944	75,596	76,166	76,558
2 person (1 adult, 1 child)	9,975	10,248	10,510	10,830	11,072	11,366	11,665	11,962	12,233	12,520
3 person (No children)	18,362	18,413	18,524	18,630	18,689	18,718	18,746	18,757	18,739	18,733
3 person (2 adults, 1 child)	19,554	19,577	19,609	19,727	19,733	19,812	19,928	20,014	20,081	20,177
3 person (1 adult, 2 children)	6,724	6,693	6,662	6,665	6,635	6,631	6,630	6,639	6,637	6,654
4 person (No children)	6,226	6,344	6,501	6,652	6,780	6,891	6,996	7,082	7,147	7,230
4 person (2+ adults, 1+ children)	28,185	28,156	28,123	28,166	28,058	28,047	28,064	28,064	28,067	28,114
4 person (1 adult, 3 children)	2,204	2,186	2,169	2,163	2,147	2,139	2,132	2,131	2,127	2,129
5+ person (No children)	2,445	2,568	2,732	2,880	3,006	3,115	3,219	3,307	3,374	3,465
5+ person (2+ adults, 1+ children)	14,431	14,385	14,332	14,314	14,215	14,162	14,127	14,085	14,044	14,036
5+ person (1 adult, 4+ children)	813	792	773	757	739	724	709	697	685	676
1 person	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2 person (No children)	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
2 person (1 adult, 1 child)	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
3 person (No children)	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
3 person (2 adults, 1 child)	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
3 person (1 adult, 2 children)	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
4 person (No children)	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
4 person (2+ adults, 1+ children)	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
4 person (1 adult, 3 children)	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
5+ person (No children)	*	*	*	*	*	*	*	*	*	*
5+ person (2+ adults, 1+ children)	*	*	*	*	*	*	*	*	*	*
5+ person (1 adult, 4+ children)	*	*	*	*	*	*	*	*	*	*

Figure 23: Asterisk where factor value cannot be calculated

9.4. The 'DFRatesOut' Workbook: Derived Category Rates

The `DFRatesOut` workbook provides the rates calculated by the DF scenario for each age, sex and derived category combination for each year of the forecast (Figure 24). If sub-populations have been defined, these will be included as an additional dimension to the output.

An *'All Groups'* worksheet is included. These data will differ from those on the *'Default'* sheet of the input `DFRates` file, as they will be calculated as an average of the rates used for individual *'Groups'*.

The rates presented in `DFRatesOut_`*[Scenario ID]* will be the same as the original data in `DFRates` unless the following changes have been made:

- Constraints have been applied to the scenario output in `DFCons` input file
- An annual increment has been specified in the `DFRates` input file
- Infill and interpolation has been required to produce the complete set of rates specified in the `DFRates` input file.

If any of these conditions are met, the DF recalculates derived category rates after scenario results have been derived.

Derived Forecast - Households **EXAMPLE**

CLG Stage 2 HHld rep rates by gender, age group, sub-population and category

Area One

				2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	15-24	Students	One person male	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			One person female	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			Couple no child	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			Cple+adlts no child	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			One child	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			Two children	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			Three+ children	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
			Other households	4.8%	5.0%	5.1%	5.2%	5.3%	5.5%	5.6%	5.7%	5.7%	5.8%	5.9%
Persons	15-24	Non-Students	One person male	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%
			One person female	1.5%	1.5%	1.5%	1.5%	1.4%	1.4%	1.4%	1.4%	1.3%	1.3%	1.3%
			Couple no child	1.6%	1.5%	1.5%	1.4%	1.4%	1.3%	1.2%	1.2%	1.1%	1.0%	1.0%
			Cple+adlts no child	0.7%	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%	0.5%	0.5%
			One child	2.2%	2.1%	2.0%	1.9%	1.8%	1.8%	1.7%	1.7%	1.6%	1.6%	1.6%
			Two children	0.6%	0.6%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%
			Three+ children	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
			Other households	4.8%	5.0%	5.1%	5.2%	5.3%	5.5%	5.6%	5.7%	5.7%	5.8%	5.9%
Persons	25-34	Students	One person male	6.1%	6.1%	6.0%	5.9%	5.9%	5.8%	5.7%	5.6%	5.5%	5.3%	5.3%
			One person female	4.4%	4.3%	4.3%	4.3%	4.2%	4.2%	4.1%	4.0%	3.9%	3.8%	3.8%
			Couple no child	11.3%	11.2%	11.1%	11.1%	11.0%	10.9%	10.8%	10.7%	10.4%	10.2%	10.1%
			Cple+adlts no child	0.9%	1.0%	1.1%	1.1%	1.2%	1.3%	1.4%	1.4%	1.5%	1.5%	1.6%
			One child	9.4%	9.4%	9.4%	9.3%	9.2%	9.2%	9.2%	9.3%	9.5%	9.7%	9.7%
			Two children	8.4%	8.2%	7.9%	7.7%	7.4%	7.2%	7.0%	6.9%	6.9%	6.8%	6.6%
			Three+ children	3.7%	3.6%	3.5%	3.4%	3.3%	3.3%	3.2%	3.2%	3.2%	3.1%	3.1%
			Other households	3.2%	3.4%	3.6%	3.8%	3.9%	4.1%	4.3%	4.4%	4.6%	4.7%	4.8%
Persons	25-34	Non-Students	One person male	6.1%	6.1%	6.0%	5.9%	5.9%	5.8%	5.7%	5.6%	5.5%	5.3%	5.3%
			One person female	4.4%	4.3%	4.3%	4.3%	4.2%	4.2%	4.1%	4.0%	3.9%	3.8%	3.8%
			Couple no child	11.3%	11.2%	11.1%	11.1%	11.0%	10.9%	10.8%	10.7%	10.4%	10.2%	10.1%
			Cple+adlts no child	0.9%	1.0%	1.1%	1.1%	1.2%	1.3%	1.4%	1.4%	1.5%	1.5%	1.6%
			One child	9.4%	9.4%	9.4%	9.3%	9.2%	9.2%	9.2%	9.3%	9.5%	9.7%	9.7%
			Two children	8.4%	8.2%	7.9%	7.7%	7.4%	7.2%	7.0%	6.9%	6.9%	6.8%	6.6%
			Three+ children	3.7%	3.6%	3.5%	3.4%	3.3%	3.3%	3.2%	3.2%	3.2%	3.1%	3.1%
			Other households	3.2%	3.4%	3.6%	3.8%	3.9%	4.1%	4.3%	4.4%	4.6%	4.7%	4.8%

Figure 24: Example model output: `DFRatesOut_`*[Scenario ID]*

Note that in some rare cases it may be impossible for the model to calculate the 'rates' values. This may occur on the `'All Groups'` sheet where a zero population would lead to a division by zero and therefore a rate value cannot be calculated – even though there may have been a value for each individual `'Group'` sheet. These instances are marked in the output by an asterisk (*) accompanied by an explanatory note at the bottom of the data panel.

Derived Forecast - Households **EXAMPLE**

CLG Stage 2 HHld rep rates by gender, age group, sub-population and category

All Areas

				2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	85+	Students	One person male	*	*	*	*	*	*	*	*	*	*	*
			One person female	*	*	*	*	*	*	*	*	*	*	*
			Couple no child	*	*	*	*	*	*	*	*	*	*	*
			Cple+adlts no child	*	*	*	*	*	*	*	*	*	*	*
			One child	*	*	*	*	*	*	*	*	*	*	*
			Two children	*	*	*	*	*	*	*	*	*	*	*
			Three+ children	*	*	*	*	*	*	*	*	*	*	*
			Other households	*	*	*	*	*	*	*	*	*	*	*

Figure 25: Asterisk where rate value cannot be calculated

9.5. The 'DFRiskPop' Workbook: Population At Risk

This `DFRiskPop` output file contains the adjusted population, disaggregated by age and sex, and the adjustments made to reach the adjusted population (Figure 26). If sub-populations have been defined, these will be included as an additional dimension to the output.

An `'All Groups'` worksheet is included, which is a summation of the individual `'Group'` worksheets.

Population at Risk
Population Adjustment - Not in Households
Original Population

Area One

			2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Students	Persons	15-24	3,761	4,006	4,509	4,997	5,165	5,557	5,945	5,862	5,451	5,494	5,830
Students	Persons	25-34	1,705	1,695	1,676	1,808	2,010	2,165	2,282	2,382	2,401	2,340	2,361
Students	Persons	35-44	666	645	639	630	631	627	639	646	678	706	718
Students	Persons	45-54	383	384	378	378	370	363	356	360	356	348	343
Non-Students	Persons	15-24	8,638	9,099	9,536	9,415	10,265	10,581	10,674	10,333	10,467	11,068	10,949
Non-Students	Persons	25-34	10,416	10,194	10,089	10,305	10,771	11,080	11,460	11,757	12,116	12,264	12,430
Non-Students	Persons	35-44	14,084	14,311	14,303	14,359	14,415	14,423	14,345	14,263	14,113	13,752	13,494
Non-Students	Persons	45-54	11,882	11,596	11,702	11,957	12,300	12,621	12,928	13,374	13,639	14,002	14,409
	Male	15-19	-221	-257	-283	-275	-348	-400	-442	-472	-546	-622	-618
	Male	20-24	-325	-365	-411	-456	-467	-526	-583	-596	-570	-585	-639
	Male	25-29	-73	-76	-78	-83	-95	-104	-111	-115	-116	-114	-117
	Male	30-34	-22	-22	-22	-22	-22	-23	-25	-26	-28	-30	-31
	Male	35-39	-25	-25	-24	-23	-23	-22	-21	-20	-20	-19	-18
	Male	40-44	-20	-21	-21	-21	-21	-22	-22	-22	-23	-24	-24
	Male	45-49	-11	-11	-12	-13	-14	-16	-16	-17	-18	-19	-20
	Female	15-19	-386	-435	-475	-480	-577	-620	-633	-647	-735	-838	-855
	Female	20-24	-269	-291	-329	-376	-389	-440	-511	-532	-540	-558	-606
	Female	25-29	-85	-82	-77	-83	-85	-90	-95	-98	-100	-96	-96
	Female	30-34	-37	-34	-32	-29	-27	-25	-24	-22	-20	-18	-17
	Female	35-39	-15	-15	-14	-14	-13	-12	-12	-12	-11	-10	-10
	Female	40-44	-17	-17	-17	-17	-17	-17	-16	-15	-13	-12	-11
	Female	45-49	-11	-11	-11	-11	-11	-11	-11	-11	-11	-11	-11
	Male	15-19	3,015	3,186	3,157	2,785	3,214	3,381	3,440	3,397	3,753	4,110	3,862
	Male	20-24	3,899	4,246	4,687	5,122	5,312	5,704	6,052	5,952	5,443	5,541	5,998
	Male	25-29	2,733	2,796	2,839	3,014	3,459	3,752	3,936	4,071	4,061	3,939	4,016
	Male	30-34	3,203	3,092	3,027	3,025	2,995	2,989	3,100	3,203	3,374	3,575	3,632
	Male	35-39	3,661	3,695	3,632	3,610	3,625	3,615	3,557	3,563	3,529	3,392	3,347
	Male	40-44	3,670	3,757	3,731	3,699	3,748	3,772	3,862	3,797	3,808	3,827	3,775
	Male	45-49	3,215	3,177	3,277	3,384	3,511	3,634	3,669	3,717	3,709	3,726	3,767
	Male	50-54	3,042	2,917	2,906	2,984	3,011	3,072	3,111	3,226	3,319	3,407	3,530
	Female	15-19	3,063	3,255	3,368	3,220	3,667	3,629	3,458	3,321	3,654	4,067	3,975
	Female	20-24	3,623	3,765	4,330	4,872	5,018	5,409	5,837	5,771	5,458	5,447	5,662
	Female	25-29	2,950	2,854	2,749	3,011	3,242	3,463	3,669	3,869	3,942	3,860	3,855
	Female	30-34	3,452	3,361	3,359	3,279	3,314	3,282	3,291	3,257	3,404	3,488	3,549
	Female	35-39	3,990	3,983	3,924	3,951	3,766	3,636	3,565	3,644	3,585	3,561	3,504
	Female	40-44	3,505	3,599	3,731	3,804	3,981	4,100	4,071	3,974	3,936	3,743	3,649
	Female	45-49	3,020	3,119	3,216	3,292	3,375	3,447	3,586	3,743	3,801	3,953	4,074
	Female	50-54	3,044	2,820	2,732	2,726	2,825	2,885	2,970	3,100	3,218	3,317	3,433

Figure 26: Example model output: `DFRiskPop_[Scenario_ID]`

9.6. The 'DFReporter' Workbook: Reports and Charts

The `DFReporter` provides the flexibility to produce reports and charts, with data aggregated across user-defined categories.

9.6.1. Reporter

The `'Reporter'` worksheet presents the following options:

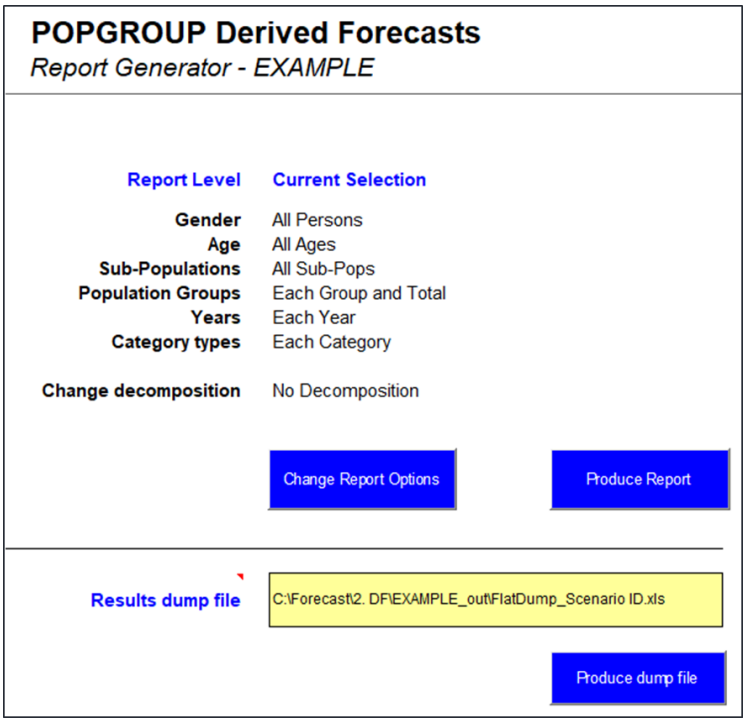


Figure 27: `DFReporter`: Report Generator

If the user wishes to accept the default selections, they can proceed to generate report. Once the *'Produce Report'* option is clicked, the user is prompted to give a label to the report worksheet that is subsequently produced (Figure 28).

Derived Forecasts Summary Report **EXAMPLE**

Households - All Persons - All Ages - All Sub-Pops

All Areas

Category of Households	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*
One person male	122,798	126,466	129,850	134,436	140,818	143,957	146,419	148,817
One person female	168,815	169,403	169,554	170,729	173,637	173,260	172,228	171,100
Couple no child	229,187	230,921	232,009	234,753	239,833	240,270	239,777	239,562
Cple+adlts no child	62,586	62,575	62,359	62,553	63,242	63,049	62,352	61,755
One child	108,812	107,630	107,026	106,892	107,554	108,181	109,912	112,179
Two children	100,979	98,324	96,147	94,514	93,704	92,899	93,073	93,537
Three+ children	45,267	44,020	42,975	42,216	41,879	41,479	41,506	41,658
Other households	66,942	69,214	71,389	74,148	77,810	79,207	80,304	81,369
Total	**905,386**	**908,554**	**911,307**	**920,242**	**938,477**	**942,302**	**945,572**	**949,978**

Area One

Category of Households	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*
One person male	40,933	42,155	43,283	44,812	46,939	47,986	48,806	49,606
One person female	56,272	56,468	56,518	56,910	57,879	57,753	57,409	57,033
Couple no child	76,396	76,974	77,336	78,251	79,944	80,090	79,926	79,854
Cple+adlts no child	20,862	20,858	20,786	20,851	21,081	21,016	20,784	20,585
One child	36,271	35,877	35,675	35,631	35,851	36,060	36,637	37,393
Two children	33,660	32,775	32,049	31,505	31,235	30,966	31,024	31,179
Three+ children	15,089	14,673	14,325	14,072	13,960	13,826	13,835	13,886
Other households	22,314	23,071	23,796	24,716	25,937	26,402	26,768	27,123
Total	**301,795**	**302,851**	**303,769**	**306,747**	**312,826**	**314,101**	**315,191**	**316,659**

Figure 28: Example summary report

!	**Important Note for those using Size Factors (e.g. average household size)**

Note that depending on the selections made, the reporting and charting routines will pick up data from different model output files. If the model setup included the use of size factors (i.e. perhaps converting between 'household population' and 'households'), this can mean that different units will be displayed in the outputs.

Where user choices are *'All Genders'*, *'All Ages'* and *'All Sub-Populations'* (where there are some), the model output data will be retrieved from the `DFForecastTotal` file; otherwise data will be retrieved from the `DFForecastDetail` file. Where size factors have been used, this will result in different output units being used and displayed for different requests.

For example, where factors have been used to convert from 'household population' to 'households', a user selection involving *'All Genders'*, *'All Ages'* and *'All Sub-Populations'* will result in 'household' output being displayed. Any more detailed selection – i.e. where perhaps individual age groups or genders have been selected – will, in this case, result in 'household population' data being displayed. Where factors have not been used, a request involving an age-sex group gives the households headed by that age-sex group.

The units being displayed will be indicated at the top of the report or chart.

The format and content of the reports can be controlled using a series of options which are accessed from the *'Change Report Options'* button. The following options may be selected:

- Gender & Age
- Sub-Populations (if defined)
- Population Groups
- Years
- Categories
- Decomposition.

Gender & Age

The default selection is *'All Persons'* and *'All Ages'* (Figure 29). Users have the option to choose *'All Persons'* or to select either *'Male'* or *'Female'*. If the data are available, single age group or user defined age ranges can be selected.

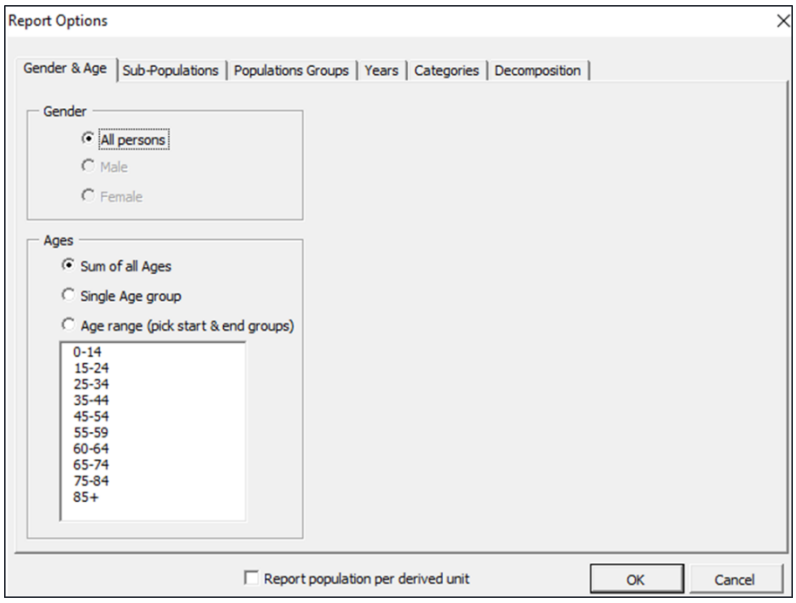


Figure 29: Report options – Gender & Age

Sub-Populations

The default selection is *'Sum of all Sub-Populations'* (Figure 30). Users have the option to choose *'Single Sub-Populations'* instead.

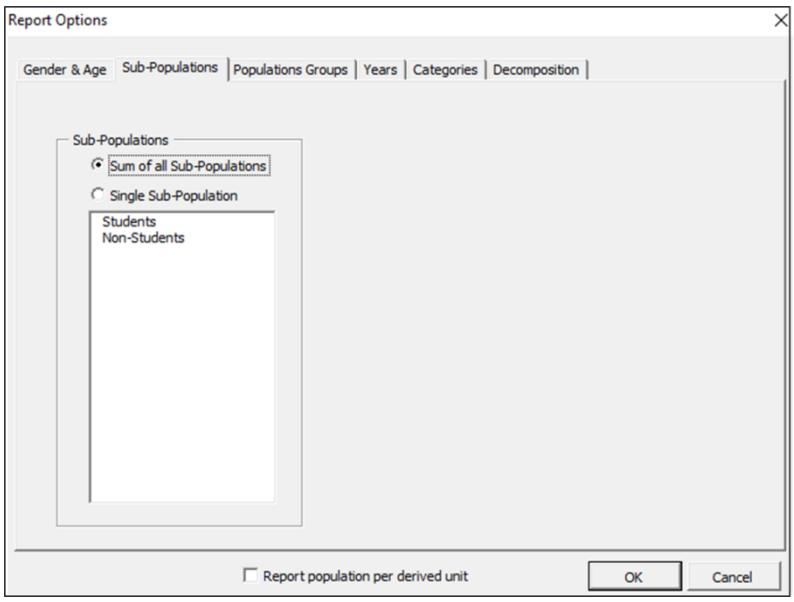


Figure 30: Report options – Sub-Populations

Population Groups

By default each population group and total will be displayed in the report (Figure 31).

The user can choose to display just the *'Sum of all Groups'*, *'Individual Groups'* or can create an aggregate as the summation of groups (see guidance on creation of aggregates on page 60).

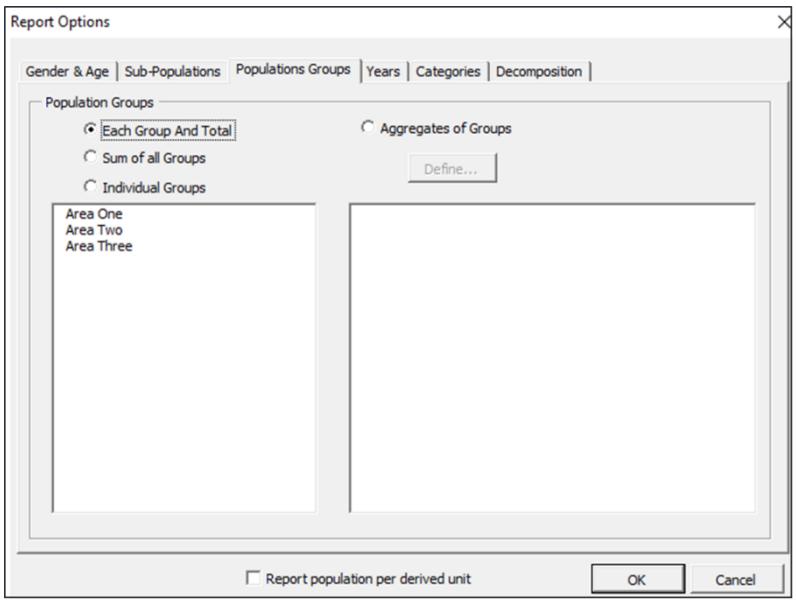


Figure 31: Report options – Population Groups

Years

Each year is displayed by default but the user may choose to display *'Every five years'* or *'Selected years'* within the forecast period (Figure 32). For *'Selected years'*, use the *Control* and *Shift* keys to select more than one year.

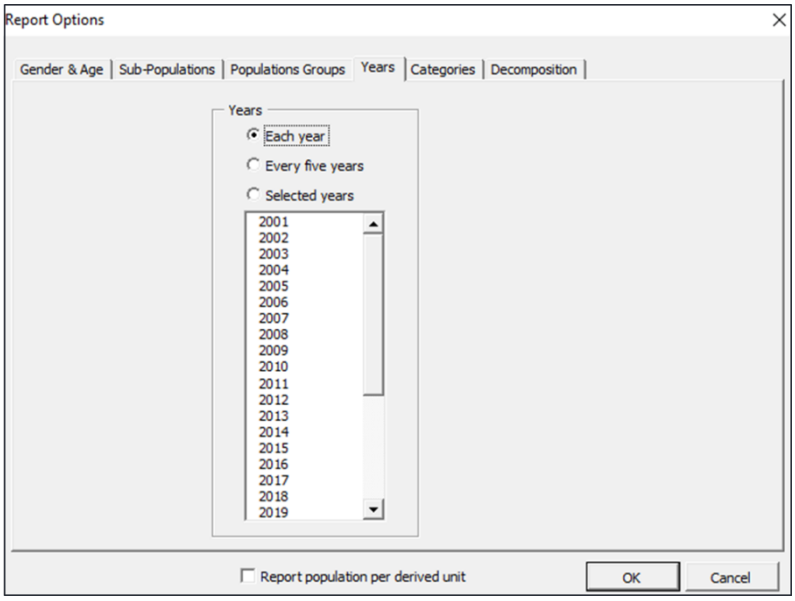


Figure 32: Report options – Years

Categories

By default the report will display results for each derived category (Figure 33).

If *'Default Category Groups'* have been specified during model setup, then the option to display these will be available. Users may also create aggregates of categories (see guidance on creation of aggregates on page 59).

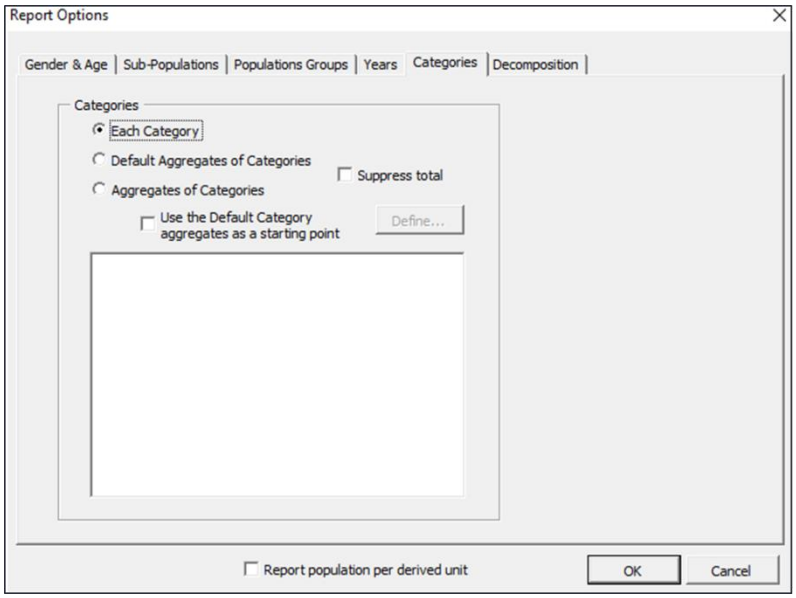


Figure 33: Report options – Categories

Decomposition

Decomposition illustrates the relative impact of population change and change in the derived category rates (e.g. headship rates) upon the final derived category numbers (e.g. households).

The default setting is for no decomposition to be displayed. If the user wishes to see decomposition statistics, then a start and end year must be selected (Figure 34).

The 'population effect' is the change in derived units if the derived category rates were to remain constant. The 'rates effect' is the remainder of the projected change in derived units.

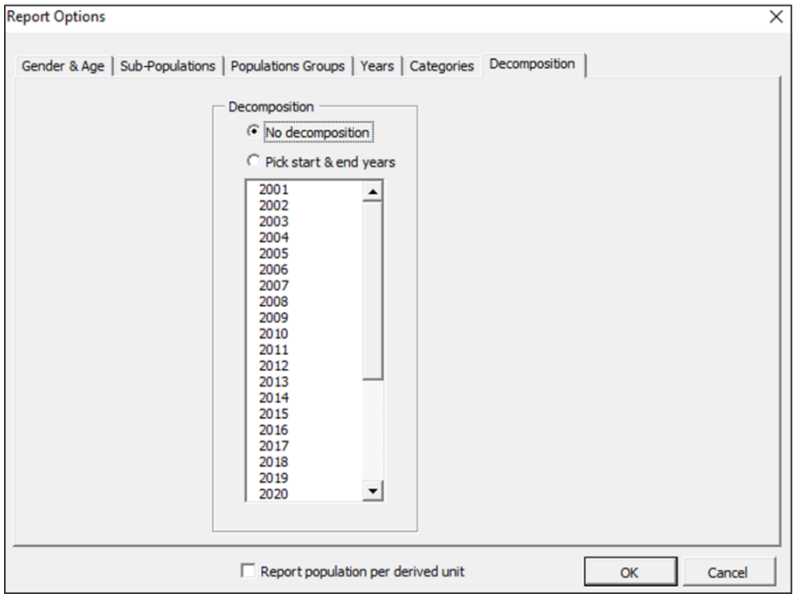


Figure 34: Report options – Decomposition

Population per Derived Unit

A tick box *'Report population per derived unit'* is included at the bottom of the form.

Where age groups used within the model sum to the 'total population' (i.e. includes child age groups), the reporter can produce an additional indicator which is 'total adjusted population' / 'derived units'. For example, in household projections this indicator is the average household size.

Dump File

The reporter provides the facility to produce a `FlatDump` file. This takes output from each of the model files and writes it to a 'flat file' format where it may be easily imported into other applications for further analysis and/or processing.

To produce `FlatDump` file, use the *'Produce dump file'* option on the *`Reporter`* sheet (Figure 27).

9.6.2. Charter

The *'Charter'* worksheet presents the following options:

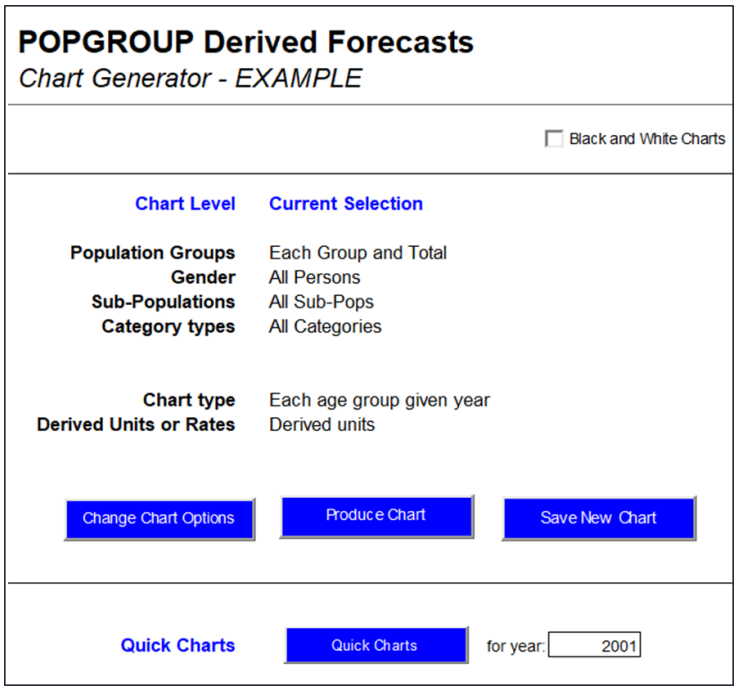


Figure 35: Chart Generator

The option for 'Sub-Populations' will only appear if sub-populations have been defined in the model setup.

Again, the user may accept the default selections and choose to *'Produce Chart'* without making any changes. This will prompt the user to give a label to the resulting chart worksheet (Figure 36).

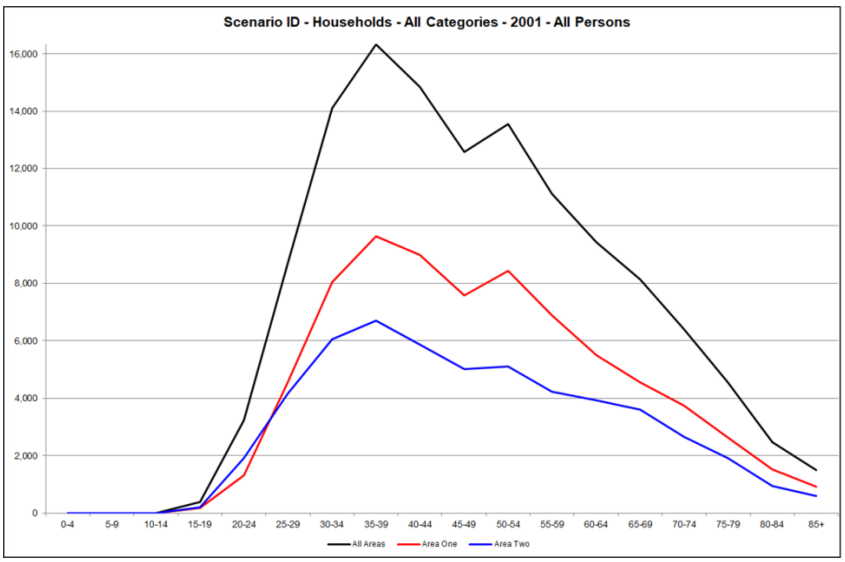


Figure 36: Example summary chart

The content of the charts can be controlled by the user using a series of options which are accessed from the *'Change Chart Options'* button. The following options may be selected:

- Chart Type
- Sub-Populations (optional)
- Populations Groups
- Categories.

Chart Type

The default selection for the chart type is to display a graph of the number of *'Derived Units'* for each age group in a given year (Figure 37).

The user can configure this to change the specific year and/or to choose to display *'Rates'* rather than *'Derived Units'*.

In addition, the user can modify the chart type so that it displays a graph of the derived units (or rates) for a selected age-group across each year, to display a time series.

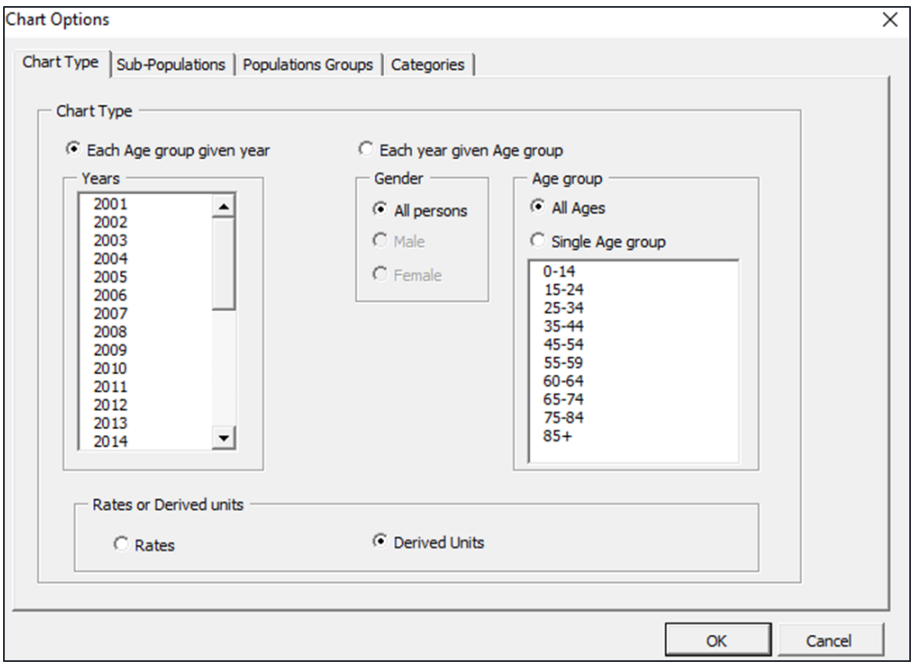


Figure 37: Chart options - Chart Type

Sub-Populations

The default selection for the sub-populations is to display the *'Sum of all Sub-Populations'*. Additional option allows the user to display a *'Single Sub-Population'* (Figure 38).

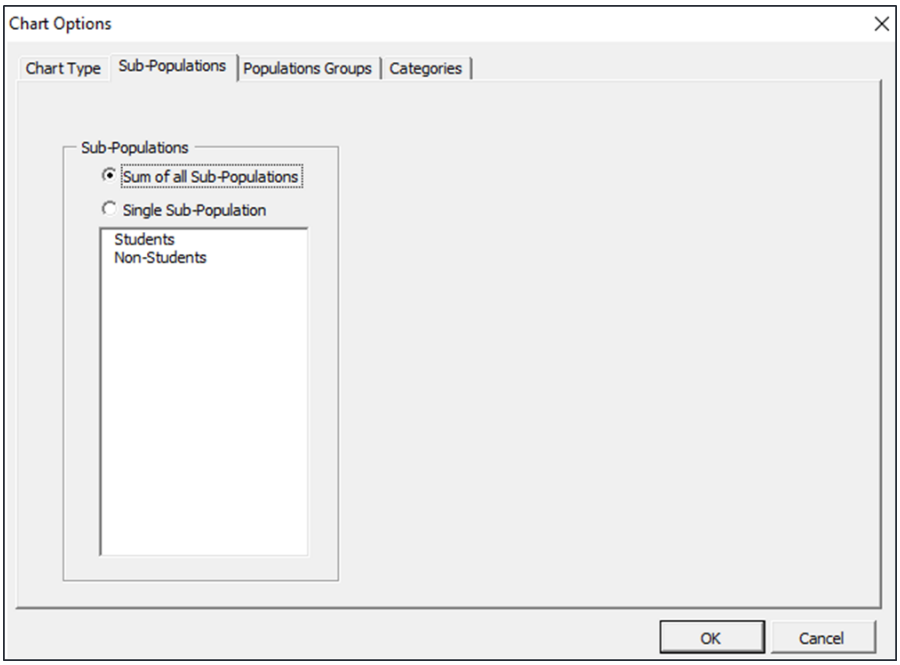


Figure 38: Chart options - Sub-Populations

Population Groups

The default selection for the population groups is to display *'Each Group and Total'*. Additional options allow the user to display just the *'Sum of all Groups'* or *'Individual Groups'* (Figure 39).

In addition, there is an option to create *'Aggregates of groups'* (see guidance on creation of aggregates on page 60).

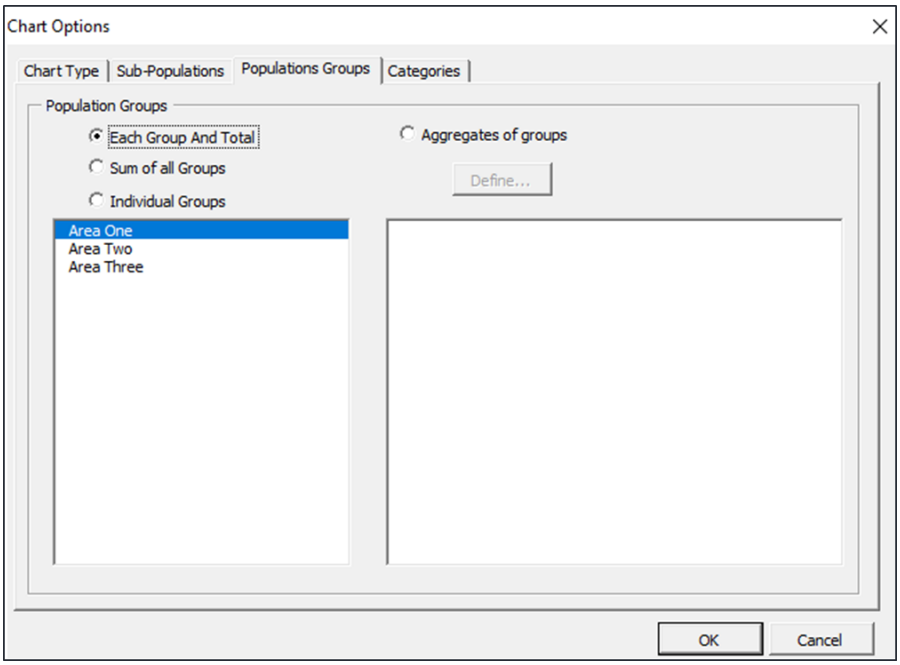


Figure 39: Chart options - Population Groups

Categories

The default setting for the derived categories is to display *'Sum of all categories'* (Figure 40). This can be modified to display an *'Individual category'* or the user may create *'Aggregates of categories'* (see guidance on creation of aggregates on page 60).

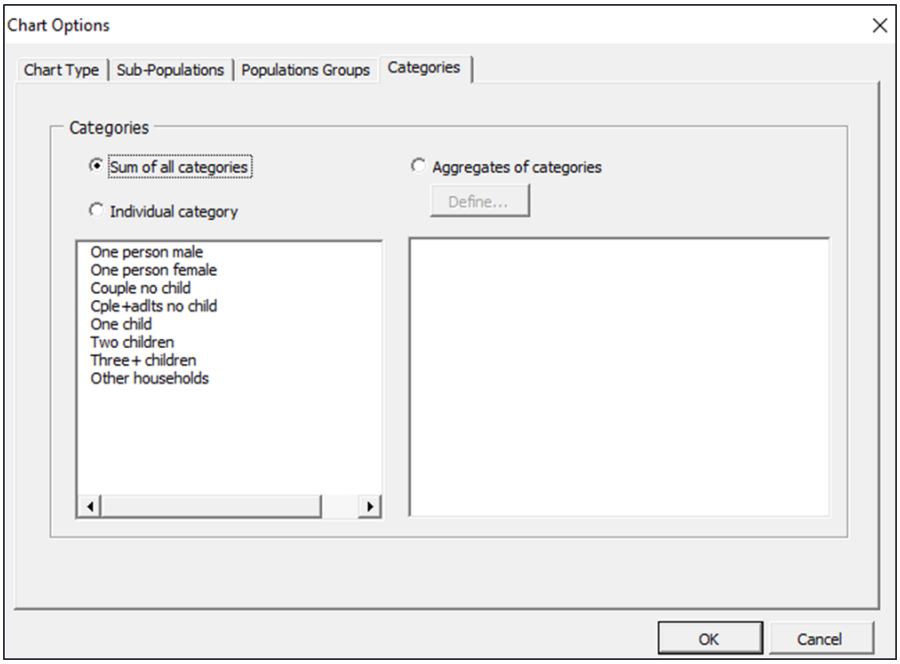


Figure 40: Chart options - Categories

Quick Charts

A *'Quick Charts'* option is also provided from the `'Charter'` worksheet. This produces a set of 'rate' graphs for a specified year for each of the derived categories (Figure 41).

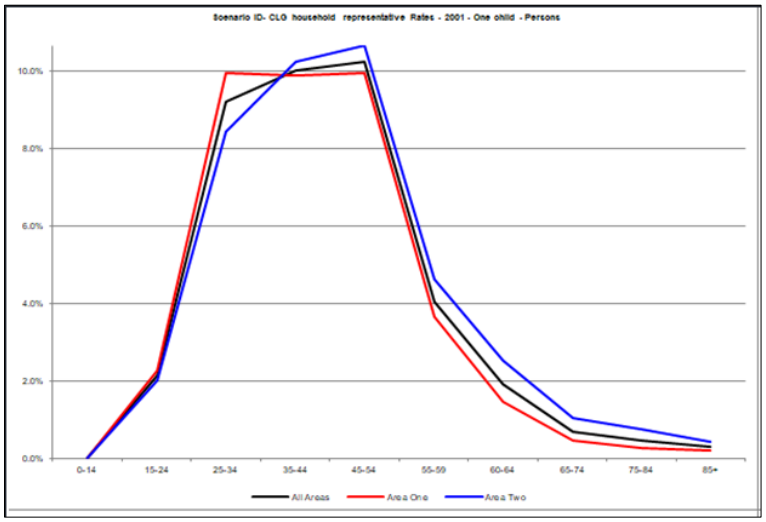


Figure 41: Example 'Quick Chart' output

9.6.3. Creating Aggregates of Groups or Categories

A number of the `'Reporter'` and `'Charter'` options enable the user to *'Create an Aggregate'* (e.g. of two or more population groups or of two or more categories). The functionality for achieving this is identical for each option and guidance is provided in the illustration below:

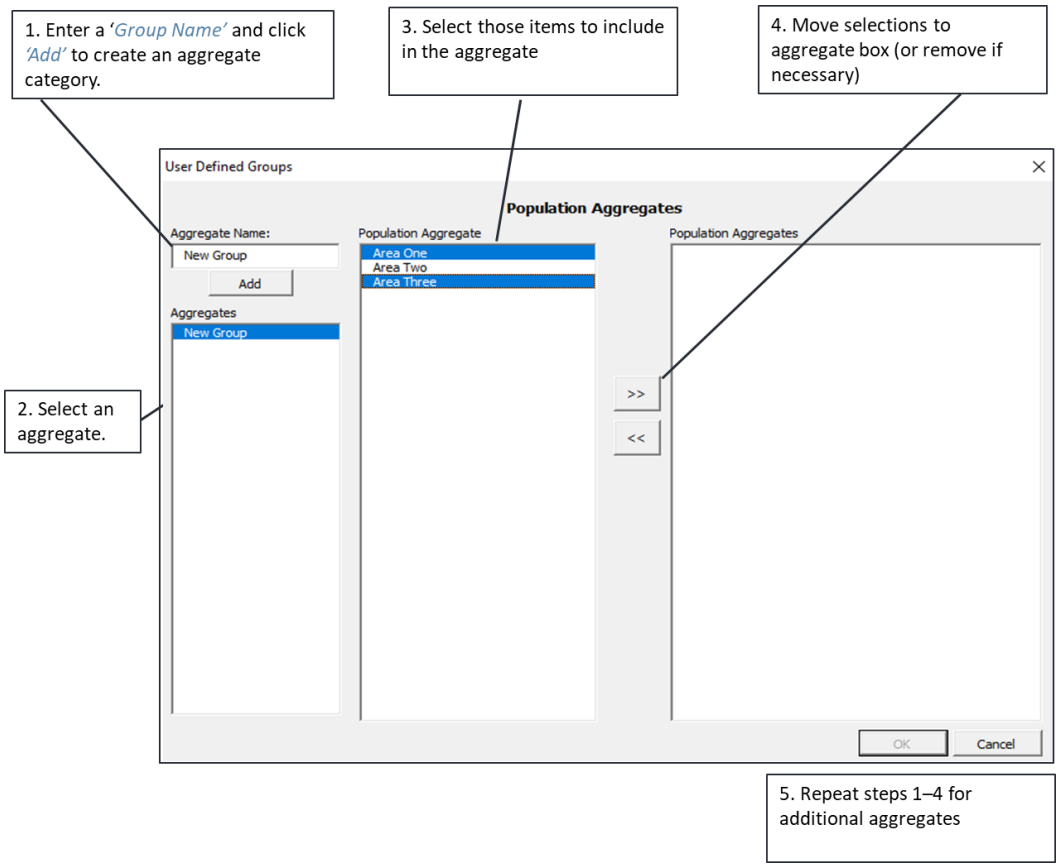


Figure 42: Creating a new group in Report and Chart Options

9.6.4. The 'DFCompare' Workbook: Comparing Scenarios

An additional 'comparison' utility (`DFCompare`) is provided to enable DF users to compare output from alternative scenarios. This utility can be found in the folder where the `DFSetup` file resides. It uses selected `DFForecastDetail` workbooks to read the data from. The interface to this utility is illustrated in Figure 43.

To generate comparison of forecasts, follow the steps detailed on the worksheet:

1. The *'Refresh list'* option identifies all the output files available for comparison from the folder specified in the *'Output folder to search'* field. Delete as appropriate.
2. The *'Get data from workbooks'* option reads each of the selected files and creates the data necessary to produce the comparison.
3. The *'Produce Summaries'* option produces the pivot tables and charts which compare scenario output.
4. An option is provided which allows the user to specify which group (area) they wish to be the default in the chart illustrations.

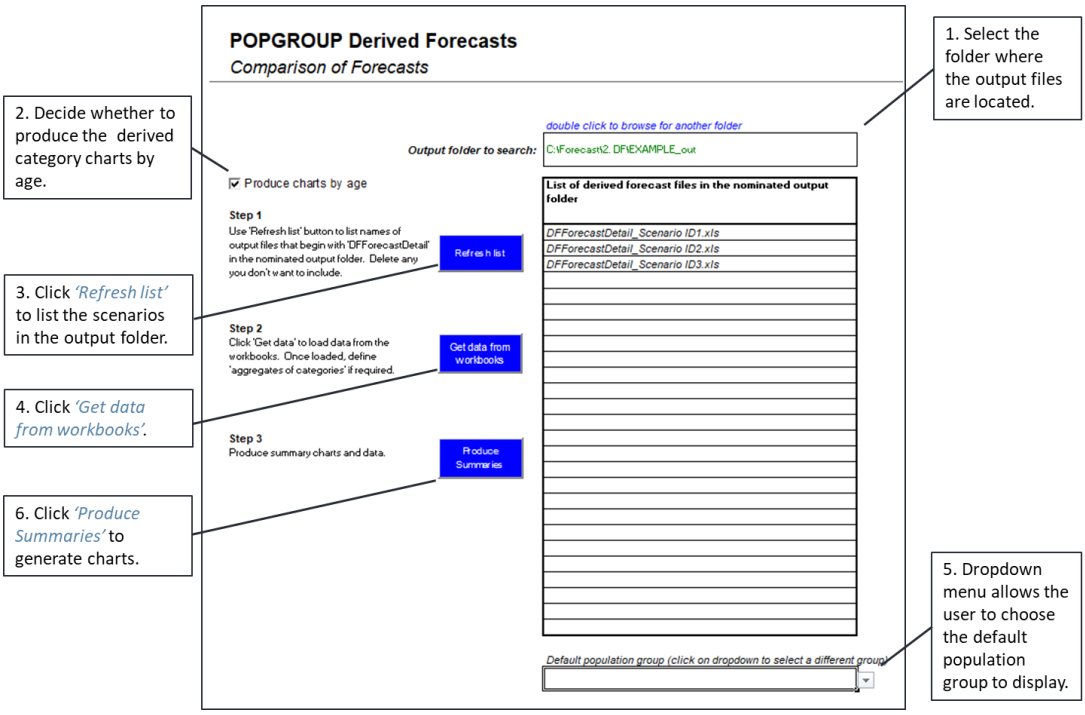


Figure 43: DFCompare

An additional option allows the user to specify alternative aggregates of the derived categories for the comparison worksheets. The simple method is illustrated below, with a three-step procedure:

1. Select the number of aggregates.
2. Give a name to the aggregates.
3. Allocate each derived category to one of the aggregate groups.

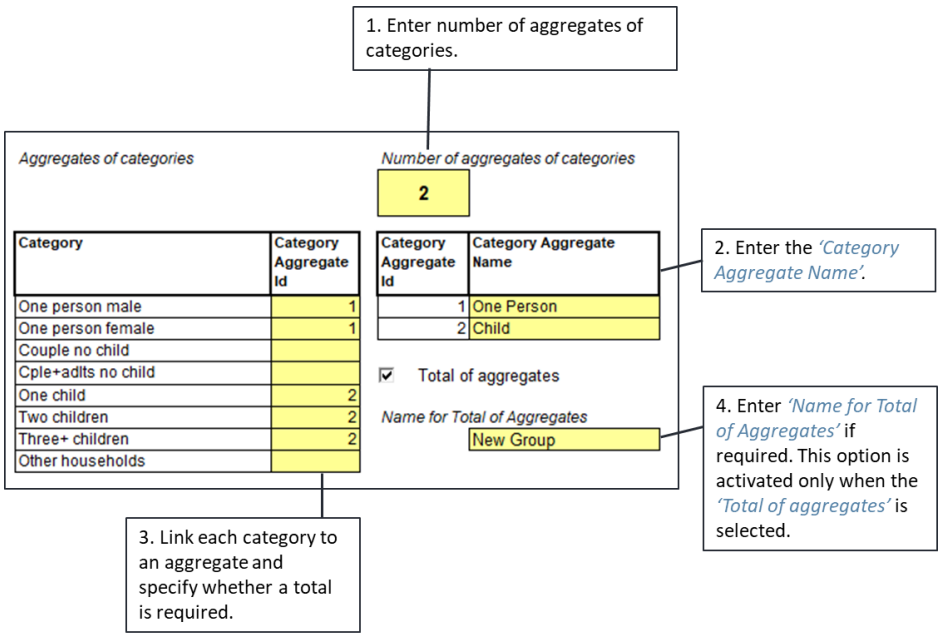


Figure 44: Aggregates of categories in DFCompare

The user may also specify whether a total of the aggregates is required. If this is the case, the user may then enter a name for the *'Total of Aggregates'*. If the aggregates do not include all categories, their total will not be the same as the overall total of categories. All selections will appear as extra comparison sheets.

Based on the options selected, **DFCompare** will compile a worksheet containing a pivot table of the data and will provide a series of charts (Figure 45) which compare the results of the selected scenarios.

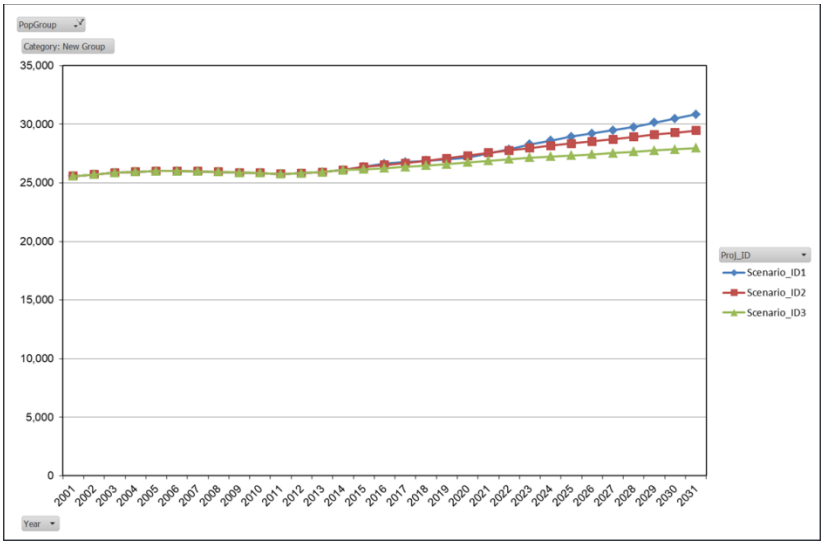


Figure 45: **DFCompare** - comparison charts

10 DF Configuration – Advanced

10.1. Summary of Options

The `DFSetup` process provides a range of options which are likely to appeal to more advanced users of the model. These options allow a user to modify the configuration of an existing model type (e.g. DCLG 2014-based household) or to create a new model type to a precise specification, for example for health, disability or any other characteristic. A modified or new configuration will result in a set of skeleton files which the user will populate with appropriate data.

The options available to the user are contained within a single form presented in Figure 46. On the left hand side, the user can modify the *'Age/Sex groups'*, *'Population Adjustment'* and *'Sub-Populations'* options. On the right hand side, the user can modify the *'Derived Units'*, *'Validation'* and *'Calculations on Derived Units'* options.

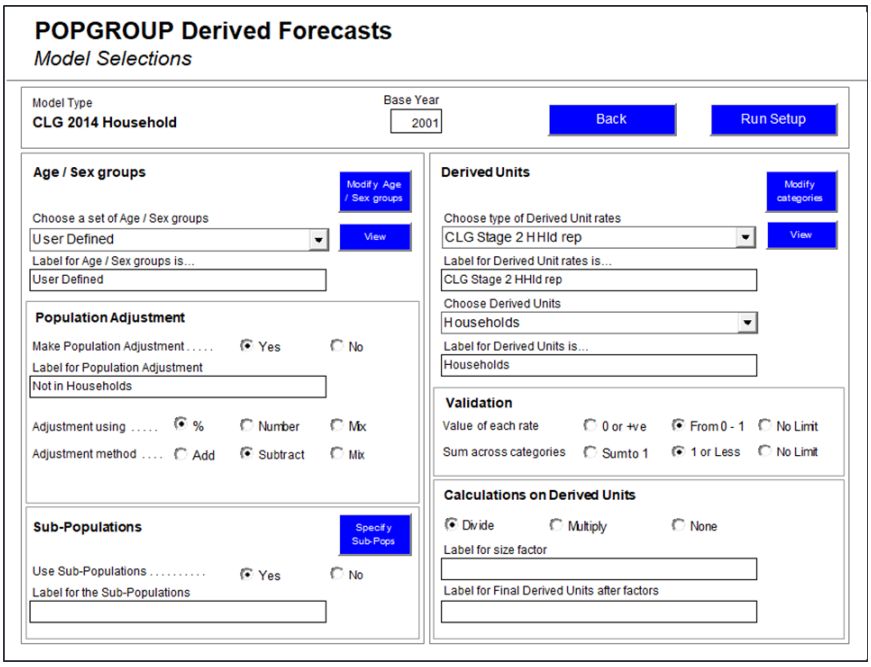


Figure 46: `DFSetup`: model selections

The remainder of this section provides a summary of the functionality that is available from each of these options.

10.2. Modification of Age-Sex Groups

The first of the advanced options available in `DFSetup` is the ability to specify which age groups are required for both the derived units and/or for the population adjustment (the latter is specifically chosen on the model setup). Select the *'Modify Age / Sex groups'* button to edit existing categories.

There must be female, male, or person age categories with the necessary requirement that the final age (e.g. 85) must end with a plus (+) sign (Figure 47). This rule applies to both: *'Age Groups in Derived Units'* and *'Age Groups in Pop Adjustment'*.

The user has a number of options to *'Clone'* data from one selection to another to simplify the configuration process. In order to verify if the age categories are correct, there is an option to *'View'* existing age-sex categories.

Skeleton files will be configured accordingly and data will be expected for the specified age and sex categories.

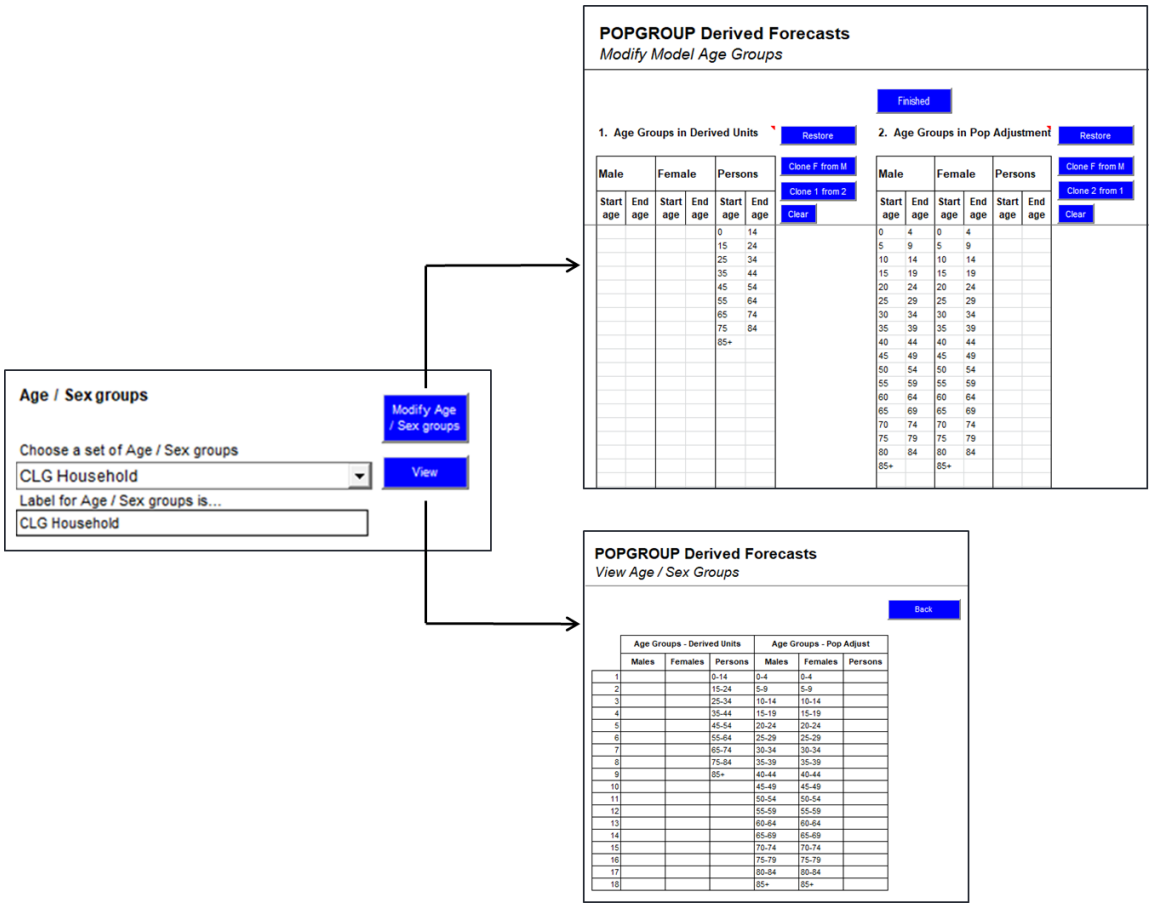


Figure 47: Modification of age-sex groups

10.3. Population Adjustment Options

The *'Population Adjustment'* options (Figure 48) allow the user to specify the rules which dictate how the final population 'at risk' is determined. The user can choose whether or not to make an adjustment to the population (removing people not in households, or adding visitors, for example).

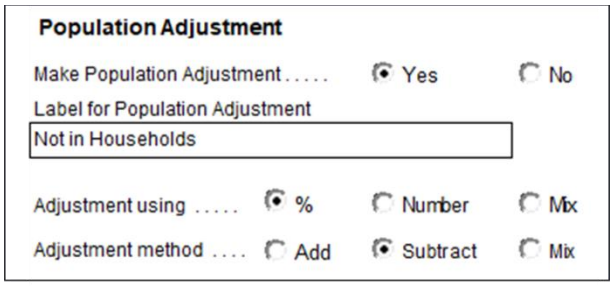


Figure 48: Population adjustment parameters

If an adjustment is chosen, then the user must supply an appropriate label. In addition, the type of adjustment and a method by which the adjustment is to be made must be specified: *'%'*, *'Number'* or *'Mix'*; and whether the value needs to be *'Added'* or *'Subtracted'* from the base population, or a *'Mix'* of the two. These will be defaults which the user can also alter on the input file for one or more age-sex groups.

The skeleton file will be configured accordingly and data will be expected to enable the population adjustment to be applied.

10.4. Sub-Populations Options

Introducing sub-populations allows a specific focus on different sections of the population. An example of this would be the ability to divide the population into students and non-students, and specify different population adjustments and rates for each sub-population. In the 2000s this was the official ONS approach to modelling economic activity, because students have much lower economic activity than others of the same age. However, with a model of sub-populations, the data requirement is more extensive. The user must supply a population forecast and a set of rates for each of the sub-populations.

If the option to *'Use Sub-Populations'* is chosen (Figure 49), then the user must click *'Specify Sub-Pops'* to list at least two individual sub-populations (students and non-students, for example, or different marital statuses). Finally, a label is required that provides a general description of the type of chosen sub-populations, such as *'Student Status'* or *'Marital Status'*.

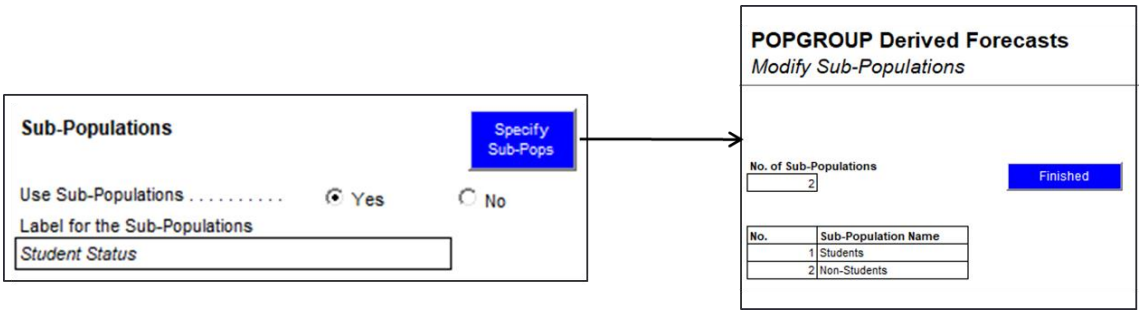


Figure 49: Specification of sub-populations

Skeleton files will be configured accordingly and data will be expected for the specified sub-population.

10.5. Modification of Derived Unit Categories

The user can choose a pre-defined type of derived unit rates and the derived unit category from a drop-down list (Figure 50). Once the user chooses one of the pre-defined or 'User Defined' options, the *'View'* button allows the user to review the derived units. If any amendments are required, *'Modify categories'* button allows the user to enter a new set of derived units or edit an existing one.

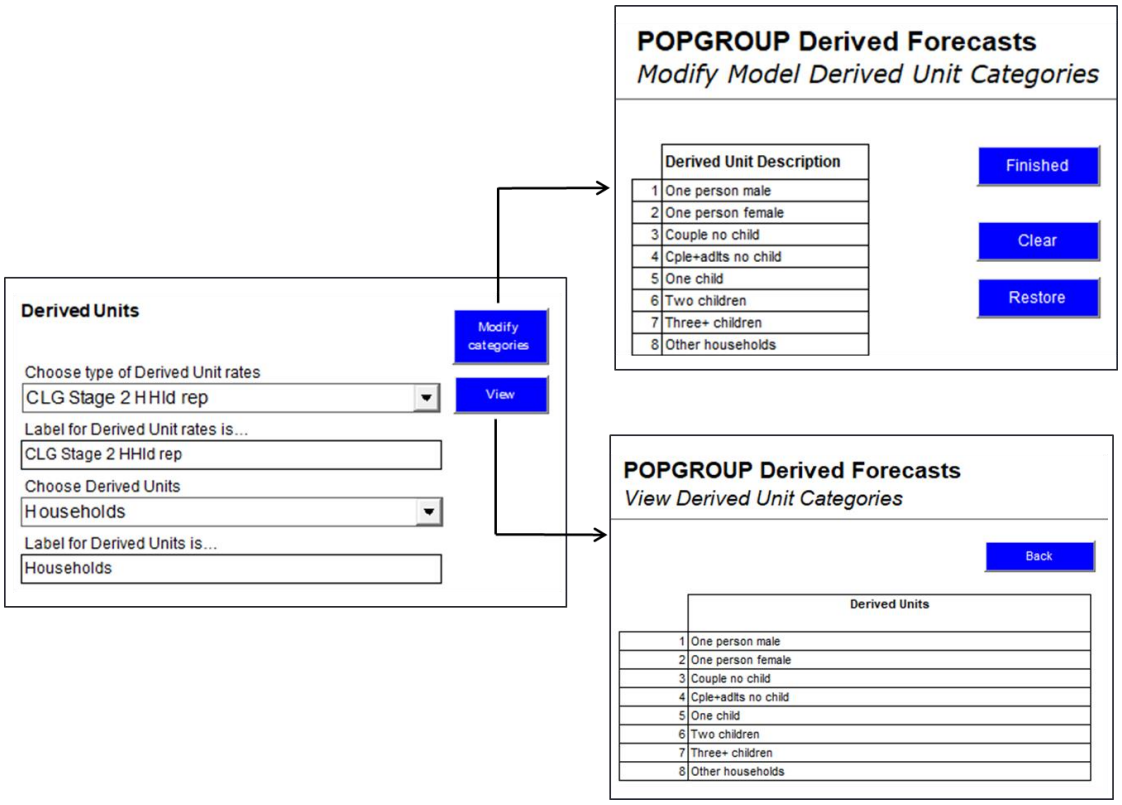


Figure 50: Modification of derived unit categories

10.6. Validation Options

The user must also select the *'Validation'* options that are to be applied to the 'rates' that will be input for each derived unit. This validation specifies the *'Value of each rate'* and the expected *'Sum [of rates] across categories'*. The DF will use these parameters to guide its validation, both during data input and at scenario run-time.

Figure 51: Validation options

10.7. Calculations on Derived Units Options

In some cases it will be necessary for an additional 'size factor' to be applied to calculate the final derived units total. This is the case in the WG household model, for example, where an 'average

household size' is required to convert the calculated household population into a number of households.

This information will be held within the `DFFactor` file and must be specified at setup as either a *'Divide'* or *'Multiply'* calculation. Labels for size factor and final derived units after factors should also be specified (Figure 52).

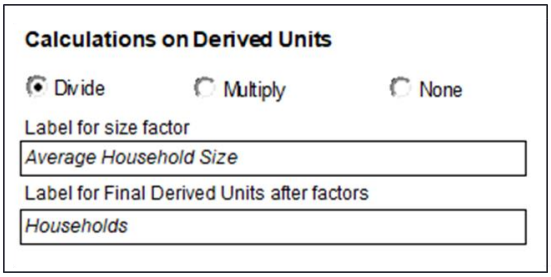


Figure 52: Options for calculations on derived units

The `DFFactor` skeleton file will be configured according to the selection and data will be expected to enable the DF scenarios to be run successfully. Validation checks on data input and at run-time will again assist the user in the process.

11 Using DF within POPGROUP

(Including supply-led population projections)

11.1. Summary

The DF model has the flexibility that allows it to be used within POPGROUP for the production of population forecasts which are constrained by derived forecasts. This means that POPGROUP can be used to produce housing- or jobs-led forecasts, for example – where migration is changed to meet a specific target number of dwellings or jobs. It can also be used to simply report the implications for the derived units of the current population forecast.

11.2. POPGROUP Constraints

In deriving a population forecast constrained by either population or a derived forecast, POPGROUP will adjust the migration flows until the constraints are met. In addition to the application of population constraints (covered in the POPGROUP Reference Manual), the following constraints may be used within a POPGROUP forecast (Figure 53):

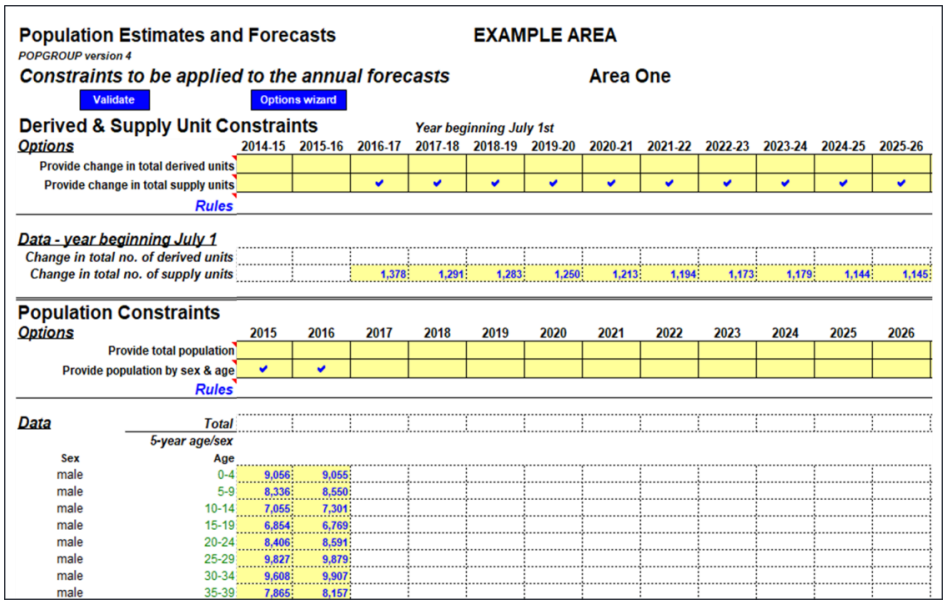
- **Derived Unit Constraint** - Provide change in total derived units, e.g. households or labour force
- **Supply Unit Constraint** - Provide change in total supply units, e.g. dwellings or jobs.

It is possible to specify *'Population'* and *'Derived/Supply Unit'* constraints in the same file (but not in the same years). This might be done if a constraint to ONS or other mid-year estimated populations is required for early years, and a constraint to housing or jobs growth targets (for example) is required for later years.

It is possible to specify both a *'Population'* constraint for each `'Group'` and for `'All Groups'`. This might be done to use ward totals with a district age-sex structure, each

estimated for recent years. If the two constraints are inconsistent, the *'All Groups'* constraints take precedence, i.e. the *'Group'* constraints are scaled to agree with them.

The calculations to adjust the migration flows to be consistent with the constraints are described in detail in Chapter 6 of the POPGROUP Reference Manual. By default, the first migration flow (the UK in-migration unless specified otherwise by the user) is adjusted. Alternatively, all four flows, or any weighted combination of them, can be specified on the POPGROUP_Scenario file at the time of running a forecast.



Population Estimates and Forecasts — **EXAMPLE AREA**

POPGROUP version 4

Constraints to be applied to the annual forecasts — **Area One**

Validate | Options wizard

Derived & Supply Unit Constraints — *Year beginning July 1st*

Options	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	2023-24	2024-25	2025-26
Provide change in total derived units												
Provide change in total supply units			✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Rules												
Data - year beginning July 1												
Change in total no. of derived units												
Change in total no. of supply units			1,378	1,291	1,283	1,250	1,213	1,194	1,173	1,179	1,144	1,145

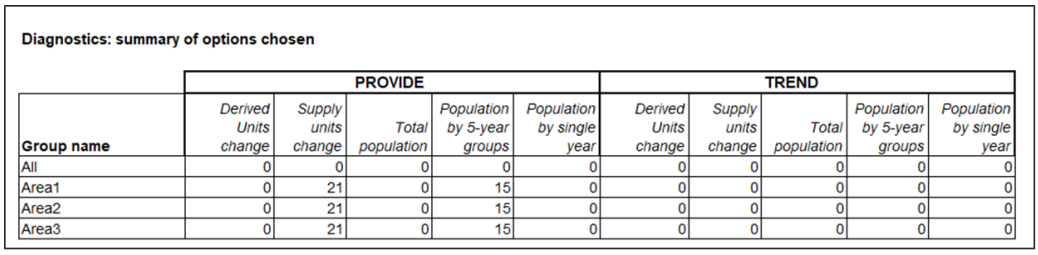
Population Constraints

Options	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026
Provide total population												
Provide population by sex & age	✔	✔										
Rules												
Data — *Total*												
5-year age/sex												
Sex / Age												
male 0-4	9,056	9,055										
male 5-9	8,336	8,550										
male 10-14	7,055	7,301										
male 15-19	6,854	6,769										
male 20-24	8,406	8,591										
male 25-29	9,827	9,879										
male 30-34	9,608	9,907										
male 35-39	7,865	8,157										

Figure 53: POPGROUP cons file

11.2.1. Notes and Validation

On validation, a 'diagnostics table' is written to the *'Notes'* sheet, summarising the options chosen in the Cons workbook (Figure 54).



Diagnostics: summary of options chosen

	PROVIDE					TREND				
Group name	*Derived Units change*	*Supply units change*	*Total population*	*Population by 5-year groups*	*Population by single year*	*Derived Units change*	*Supply units change*	*Total population*	*Population by 5-year groups*	*Population by single year*
All	0	0	0	0	0	0	0	0	0	0
Area1	0	21	0	15	0	0	0	0	0	0
Area2	0	21	0	15	0	0	0	0	0	0
Area3	0	21	0	15	0	0	0	0	0	0

Figure 54: Example of validation checks: POPGROUP cons file

Rules for data entry on POPGROUP Cons workbook:

- At most, one of the *'Options'* may be chosen on any sheet.

- Each group must have the same option (or none) chosen.
- Housing or employment constraints for *'All Groups'* and *'Group'* cannot both be chosen.
- If *'All Groups'* and *'Group'* population constraints are chosen, the *'All Groups'* constraints must be more detailed than the ones for individual groups.
- The choice of single year of age population or population by 5-year age-sex group option will be deduced from the panel of data that has entries. The 'single year of age' panel is below that for the 5-year age-sex groups.
- A trend is deduced if non-adjacent years have an option of the same type chosen. The intervening years are linearly interpolated between the two years given. Note that this is different from other input files where a trend must be explicitly requested. The user may specify population constraints as a linear trend from the base year to the first provided population constraint, by ticking the box provided on the *'Notes'* worksheet.
- The results of a population constraint using information known for recent years gives new estimates for migration, which may be used in assumptions about the future.

If any of these rules are not followed, a note of the error is provided on the *'Notes'* sheet. The forecast program will terminate if an error is detected on the constraints workbook.

11.3. The 'DFSupply' Workbook

The `DFSupply` workbook is used to define the conversion between the derived units and the supply units. The `DFSupply` workbook is used when the user wants to:

- Apply a non-population constraint to a population forecast (e.g. dwellings or jobs)
- Model the impact of a population forecast on a derived forecast.

On the *'Notes'* worksheet, the user specifies the choice of conversion units (Figure 55). The three options are:

- **Single conversion ratio** derived units/supply units

The DFSupply workbook allows a general conversion factor to be used where the detail of the separate housing and labour force rates are not required or where the derived units are neither households nor labour force.

- **Household to dwellings:** separate rates for vacancy, holiday home and sharing.
- **Labour force to jobs:** separate rates for unemployment and commuting.

The user should also specify an appropriate title for the supply unit, for example, dwellings or jobs on the *'Notes'* worksheet. This will be used in the POPGROUP comp output file.

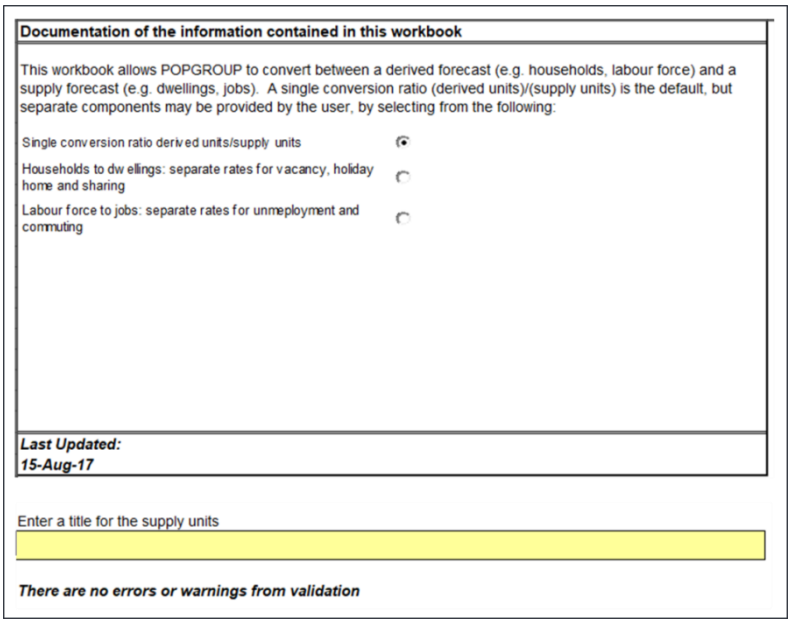


Figure 55: The *'Notes'* worksheet in the DFSupply workbook

Depending on the option selected on the *'Notes'* worksheet (Figure 55), the format of the *'Group'* worksheets and the *'Default'* worksheet will vary.

11.3.1. Single Conversion Ratio

If a *'Single conversion ratio'* option is used, only one conversion factor needs to be specified (Figure 56). This is the default setting in the DFSupply workbook. Any given value must lie between 0% and 200%.

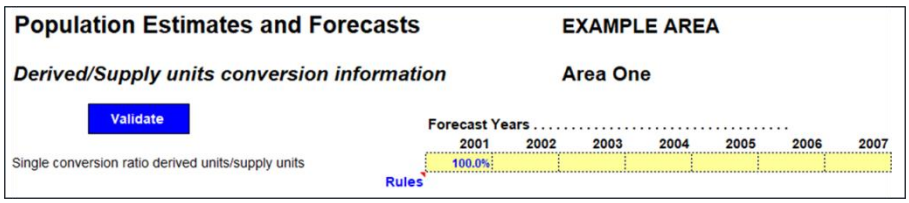


Figure 56: DFSupply - single conversion ratio

11.3.2. Households to Dwellings Ratio

If the *'Households to dwellings'* option is chosen on the `'Notes'` worksheet, three different conversion factors can be specified (Figure 57). Note that if this option is specified, data must be entered in each of the three conversion factor categories:

- **Dwellings vacant rate**

  The percentage of all dwellings which are vacant

- **Dwellings holidays/second homes rate**

  The percentage of all dwellings which are holiday homes or second homes

- **Household sharing rate**

  The proportion of households which are extra to the number of dwellings occupied by households; for example, if 5,000 households live in 4,000 occupied dwellings, then the households sharing rate is 1,000/5,000 = 20%.

The values for the three housing factors must each lie between 0% and 100%. They are converted to an overall conversion factor as follows:

Households/dwellings = (100 – vacancy rate – holiday home rates)/(100 - sharing rate)

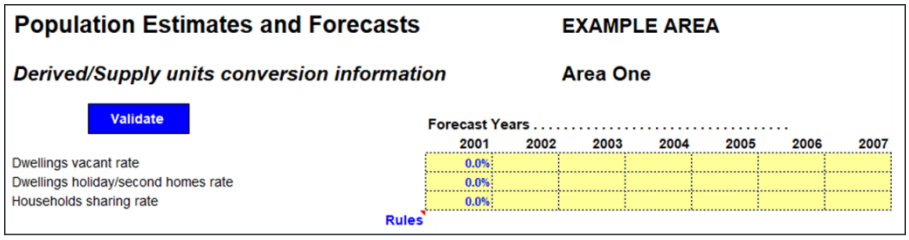


Figure 57: `DFSupply` – households to dwellings ratio

11.3.3. Labour Force to Jobs

If the *'Labour force to jobs'* option is selected in the `'Notes'` worksheet, two conversion factors can be specified (Figure 58):

- **Unemployment Rate**

  The unemployment rate is the ratio of (unemployed residents in the area)/(employed + unemployed residents in the area). If an unemployment rate is specified, it must be greater than or equal to 0% and less than 100%.

- **Commuting Ratio**

  The commuting ratio describes the impact of commuting. It is the ratio of (employed residents in the area)/(jobs in the area). Where a commuting ratio is specified, the value must be between 0 and 100.

They are converted to an overall conversion factor as follows:

Labour force/jobs = (commuting ratio*100) / (100 – unemployment rate)

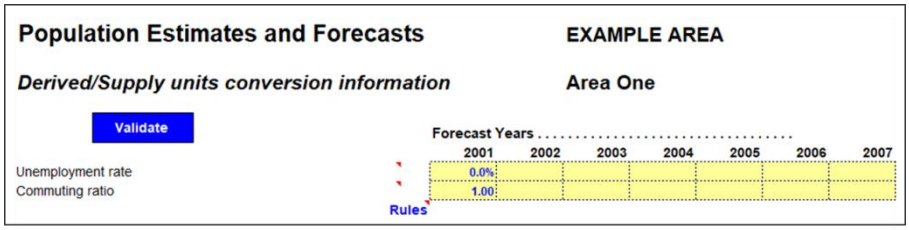


Figure 58: `DFSupply` – labour force to jobs

11.4. 'Constraints and Impacts'

Setting up a POPGROUP model is covered in the POPGROUP Reference Manual (refer to Section 7). This section describes the additional inputs required when using non-population constraints or reporting the impact of a population forecast on derived units. These additional inputs are entered on the *'Constraints_and_impacts'* worksheet of the `POPGROUP_Scenario` file (Figure 59).

In addition to the specification of the POPGROUP constraints file, there are two sections on this worksheet - the *'First Derived Forecast'* and the *'Second Derived Forecast'*. The first derived forecast files are used if you specify a `cons` file containing derived or supply units. The second derived forecast files are used only to provide the impact of the population forecast. For example, it might provide the total labour force and jobs implied by a population forecast constrained by housing targets. The housing targets would be entered in the POPGROUP constraints file, and the household files entered on the scenario as the first derived forecast.

If the user specifies a constraints file with only population constraints, or specifies no constraints at all, the first and second derived forecast sections are used to model the impact of population forecast.

In each derived forecast section, the files relevant to a particular DF model are listed. The *'Check Files'* button validates that the files listed are compatible, that all files belonging to a particular setup are listed and that they are all in the folders stated. The model will terminate if any of these criteria are not met.

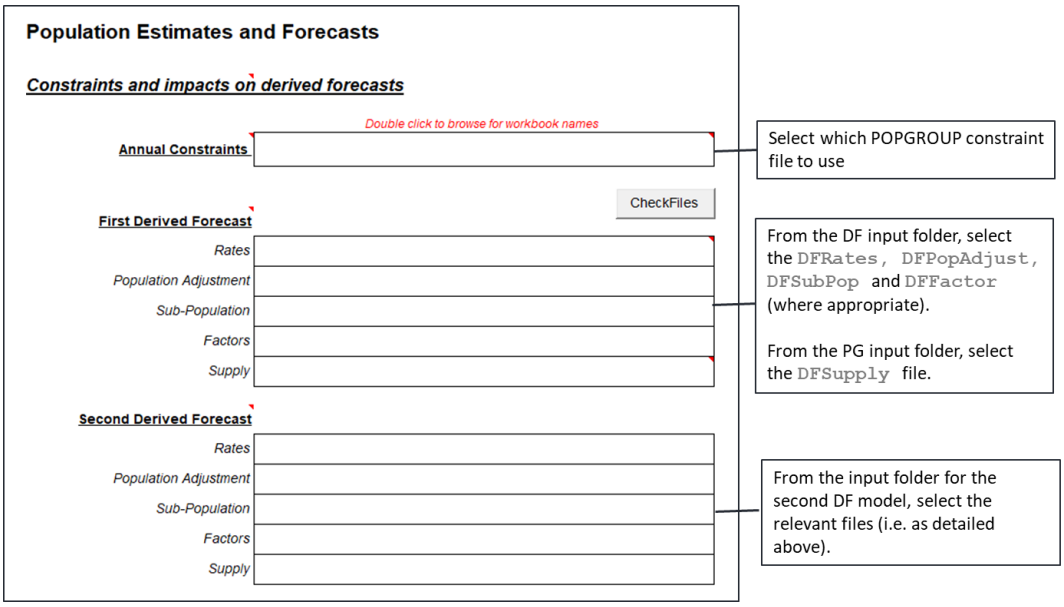


Figure 59: `POPGROUP_Scenario`: *'Constraints and Impacts'* sheet

11.5. Migration Weights

Migration weights are used only where constraints have been specified. The weights are used for each of the four migrant flows when adjusting migration to meet a constraint, defined on the *'Run_Details'* worksheet of the `POPGROUP_Scenario` file. POPGROUP has two sets of migration weights, one for population constraints, and the other for derived units such as households and labour force. Depending on which of the constraints is being applied for any year, the appropriate set of weights is used by the forecasting program to decide which migration flows to adjust to meet the constraint. Detail on options and how migration weights are implemented are outlined in the POPGROUP Reference Manual (Section 7 and Appendix B).

Appendix A
DF Example

To provide further guidance in the application of DF, an 'example' model has been configured and the input files, scenario files and output files have been provided with the main DF installation.

The example is based upon a configuration of the DF using the DCLG 2014-based household model setup (using headship rates to derive household numbers). The example data are presented for a two-zone study area (Area 1 and Area 2) and their aggregate (District). Two alternative scenarios have been run, to generate household numbers from different population forecasts.

The example data (all .xls files) are provided using the standard folder structure as follows:

File Name	Description
C:\Forecast\2. DF\...	
`DFSetup`	This is used to generate the skeleton data files.
C:\Forecast\2. DF\2. DF Example Data\...	
`DFCompare`	
C:\Forecast\2. DF\2. DF Example Data\DF Example_skel...	

`DFCons`
`DFPop`
`DFPopAdjust`
`DFRates`
`DFScenario`
	Skeleton files produced by the `DFSetup` process. See Table 1 in the reference manual for a definition of each file.

C:\Forecast\2. DF\2. DF Example Data\DF Example_inp...	

File Name	Description
DFPopAdjust_*Example*	Population adjustment, removing communal household population from resident population to create population 'at risk'.
DFRates_*Example*	Household formation rates by age, sex and household category.
DFPop_*Example_1* DFPop_*Example_2*	Population forecast for scenario 1 and 2.
DFScenario_*Example_1* DFScenario_*Example_2*	Scenario setup 1 and 2 (based on different population forecasts).
C:\Forecast\2. DF\2. DF Example Data\DF Example_out...	
DFForecastTotal_*Example_1* DFForecastTotal_*Example_2*	Household forecast by household category and year.
DFForecastDetail_*Example_1* DFForecastDetail_*Example_2*	Households by age group and household category.
DFRatesOut_*Example_1* DFRatesOut_*Example_2*	Household formation rates used in scenario.
DFReporter_*Example_1* DFReporter_*Example_2*	Utility for producing reports and charts of scenario results.
DFRiskPop_*Example_1* DFRiskPop_*Example_2*	Population 'at risk' used in household forecasts. This file includes population 'not-in-households', household population and total population.

Appendix B
Security Settings

When using any of the POPGROUP family of products, **macros must always be enabled**. Any security setting that disables macros completely will prevent the POPGROUP and Derived Forecast models from running correctly.

You may need to modify your Excel macro and security settings, which will differ depending on the version of Excel being used. The choice of these settings will also depend on the IT environment in your organisation and what it allows.

For Users of Microsoft Excel 2007 Onward

To modify the security settings in Excel 2007 onward:

1. In **Excel 2007**, click the round 'Office' button in the top left hand corner of the Excel window. Then select *'Excel Options'* at the bottom of the drop-down menu.
   In **Excel 2010 onward**, select *File > Options*.
2. Click on the *'Trust Center'* button on the left hand side of the dialog box and click *'Trust Center Settings'* on the right hand side.
3. Click on the *'ActiveX Settings'* button on the left hand side of the *'Trust Center Settings'* dialog. You can choose either of the following options: *'Prompt me before enabling all controls with minimal restrictions'* or *'Enable all controls without restrictions and without prompting'*.
4. Click on the *'Macro Settings'* button on the left hand side of the *'Trust Center Settings'* dialog. Enable *'Trust access to the VBA project object model'* and then choose one of the following options: *'Disable macros with notification'* or *'Enable all macros'*. If the former option is chosen, a security warning will prompt the user to *'Enable Content'* on the opening of the files unless all of the files (the software itself and any of the files created using the software) are stored in a 'trusted location'. Trusted locations can be specified in the *'Trust Center'*.

For Users of Microsoft Excel 2003

To modify the security settings in Excel 2003:

1. Click on *Tools > Macro > Security*.
2. On the *'Security Level'* tab, set the macro security setting to *'Medium'* or *'Low'*. If the former option is chosen, security warning will prompt you to *'Enable Macros'* on opening of the files.
3. On the *'Trusted Publishers'* tab, enable *'Trust access to Visual Basic Project'*.