



Understanding the OSCAR Data Release 2017

This document is to be used in conjunction with the release of the OSCAR raw data. Its purpose is to detail the data stored on OSCAR and its uses along with providing guidance to the detail of the data to aid understanding.

CONTENTS

CONTENTS.....	2
INTRODUCTION.....	3
What is OSCAR?	3
What does OSCAR do?.....	3
What information does the system hold?	3
THE OSCAR DATA MODEL	4
TME, DEL and AME.....	6
Resource, capital and administration budgets	6
Parliamentary Supply Estimates	6
Where does the data come from?.....	6
What the data can tell you.....	7
What the data can't tell you	7
Quality assurance.....	8
Whole of Government Accounts (WGA).....	8
The Data Files and Downloading.....	8
Files and fields in the data (Data structure/attributes)	9
Aggregation and Use of Data	10

INTRODUCTION

What is OSCAR?

OSCAR stands for Online System for Central Accounting and Reporting - a cross government project that replaced the COINS public spending database, which provides the Treasury with key management information and data for public reporting.

What does OSCAR do?

Provide a reliable and efficient management information system holding consistent, accurate and timely public spending data that enables the centre/Treasury to perform its key functions. This means the implementation of new data structures to simplify the system and aid understanding.

What information does the system hold?

Broadly speaking the same data COINS held i.e. public sector budgets, outturns and forecasts, with increased granularity of data in certain areas, for example data is now held at the level of each organisation, whereas on COINS all data was recorded at the parent Department level.

THE OSCAR DATA MODEL

The Data Model refers to the way in which we store data on OSCAR. The key elements are as follows:

We hold a number of key facts about public spending in order to uniquely define it. These facts (organisation, month, account, sub-segment etc) are held on OSCAR in what is known as the Fact Table.

For each fact we have a corresponding **dimension** to hold related information. E.g. there is an Organisation dimension that holds all the information about each organisation (such as which department group it belongs to, whether it is a core dept, arms length body, pension scheme, devolved administration or public corporation etc);

The information in these dimensions is held within **hierarchies**. I.e. information is held within nested parent/child structures to support more user-friendly aggregation and drilldown;

Each dimension comprises a **main hierarchy** and a series of **attribute hierarchies**. Using the Segment dimension as an example:

- The sub-segment rolls up through a main hierarchy comprising the segment, segment parent, segment grandparent and segment department;
- All the other information about the sub-segment (Estimates Row, Accounting Authority, Control Budget etc) is held within a series of attribute hierarchies.

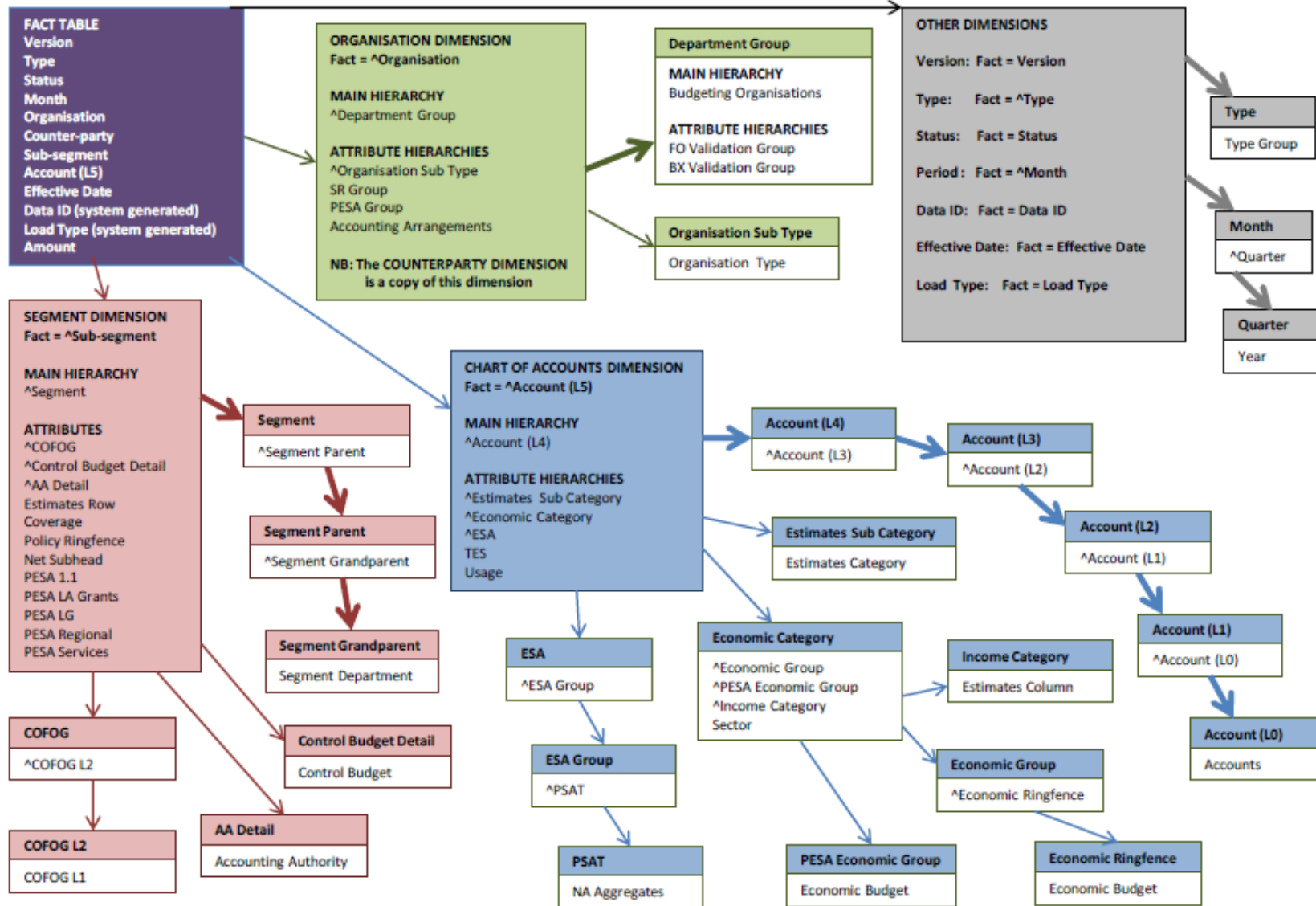
Attribute hierarchies are a new feature of OSCAR that did not exist on COINS - while COINS contained separate Budget Boundary and Category attributes, OSCAR has a Control Budget hierarchy that has 'DEL Admin' and 'DEL Programme' rolling up into a 'DEL' parent.

In addition to enabling more user-friendly functionality, the use of hierarchies also minimises the number of attributes that need to be assigned when new dimension members (e.g. sub-segments) are created. I.e. only the **base level attributes** (e.g. DEL Admin or DEL Programme in the above example) need to be assigned as these already roll up through pre-defined hierarchies.

The diagram on the next page shows a graphical representation of the OSCAR Data Model.

N.B. Some fields have been removed from the data release where these duplicate the information held in other fields. This has been done in order to reduce the overall size of the raw data file.

OSCAR Data Model - Dimensions & Hierarchies



Understanding The OSCAR Data Release

TME, DEL and AME

TME (Total Managed Expenditure) is a definition of aggregate public spending drawn from National Accounts. It is the consolidated sum of current and capital expenditure of central and local government, and public corporations. TME can also be presented as the sum of DEL (Departmental Expenditure Limit) and AME (Annually Managed Expenditure).

DELs are firm plans for up to four years (from 2011-12, for three years prior to that) for a specific part of a department's expenditure. DEL covers all administration budgets and all programme expenditure, except in certain cases where some programme spending is included in departmental AME because it cannot reasonably be subject to close control over a four-year period. Other AME covers spending that is not tied to a departmental budget (such as central government debt interest).

Resource, capital and administration budgets

Resource budgets control the current expenditure of a department and largely follow the contents of resource accounts. This comprises spending on items that are consumed in the process of providing public services or, in other words, is recurring spending. This includes, for example, wages and salaries, benefits, and purchasing goods and services. Resource budgets are set net of receipts from sales of goods and services.

Administration budgets cover the costs of running a central government department that do not relate to the delivery of frontline services; it includes the pay of most of the civil servants who work in the department, and associated expenditure such as pensions contributions, accommodation, travel and training. Administration budgets are a subset of resource DEL that are set net of income arising from departments' administrative activities.

Capital budgets include expenditure on fixed capital assets, capital grants and the acquisition of certain financial assets acquired or sold for policy reasons. Capital budgets include capital expenditure financed by finance leases and on balance sheet Private Finance Initiative transactions. Capital budgets are set net of the sale value of receipts from the sale of capital assets.

Parliamentary Supply Estimates

The Main and Supplementary Estimates are the means by which Parliament approves Government spending. Parliament approves both the use of cash and resources (which additionally contains non-cash items such as depreciation and impairments). Estimates follow resource accounting definitions as well as the budgeting framework and so sometimes contain additional transactions that are not within DEL or AME - these are described as 'non-budget' spending.

Where does the data come from?

Data are provided by central government departments, who retain ownership of their data on OSCAR. They do not include individual transactions, such as invoice payments, but aggregate transaction for reporting to Treasury. The way in which individual transactions are aggregated is largely a matter for each department, as long as they meet the minimum requirements defined by the Treasury. The level of data can, therefore, vary between departments. Departments will

generally provide the level of detail that map most conveniently from their accounting systems, and those used by other members of the departmental family.

Plans, Outturn and Forecast Outturn (referred to as **In-Year Returns** in the OSCAR **Version** dimension) data are predominantly concerned with income and expenditure with some limited balance sheet data.

What the data can tell you

All of the data contained in OSCAR are recorded against multiple coding frameworks, with the same piece of data being used in combination with various other pieces of data to produce the documents that Treasury publishes. The way in which these data are combined is complex, and may not be immediately obvious.

Updates to Plans and Outturn information are recorded in the “Fact Table”, as noted above. These can initially be entered as “draft” by a department. “Draft” entries are then updated to “proposed” before being “confirmed”.

Forecast Outturn is supplied on a monthly basis throughout the year. Each return comprises data for all twelve months of the current financial year, the initial return in June comprising two outturn months and ten forecast months. Each subsequent return will contain an additional outturn month and one fewer forecast month. Certain Forecast Outturn returns are copied into the Outturn data stream where they are used to publish estimates of annual spending during the year (for example in the Budget). Once the final return is copied into the Outturn data stream, departments load adjustments to align firstly with their draft resource accounts and then to audited accounts.

To publish consistent historical time series data departments are required to maintain OSCAR data consistent with previous years. This means that, for example, departments adjust all live years to reflect Machinery of Government changes (decision to change the responsibilities of departments) or other classification changes (such as changes to the budgeting framework or ONS decisions on the classification of organisations). It should be noted, however, that Plans data for past years are not maintained along these lines – these are fixed at the end of each year. Therefore, Plans and Outturn data for historic years will not necessarily be comparable. This maintenance excludes Forecast and Audited Outturns which relate to a single year only at a point in time, and are never changed.

A word of caution: Within the Treasury pre-defined aggregates, departments can choose the level of granularity at which they record spending data. Departmental data are therefore recorded on OSCAR at different levels of detail. Moreover, the same department may choose to record data on OSCAR at different levels of granularity in different years and/or for different entities within the departmental group.

What the data can't tell you

Departments may record data for Plans at a less detailed level than for Outturn and Forecast Outturn. As a consequence, comparisons of detailed Plans and Outturn data will not necessarily reflect genuine over- or under-spends. In addition, you may not be able to make comparisons between departments at a detailed level, as departments can, to a large extent, choose the detail at which they report spending.

OSCAR does not necessarily hold spending data against individual programmes or projects. Departments will have determined their aggregate level structure for reporting in OSCAR in consultation with their relevant Select Committee.

It is possible that you won't be able to recreate the numbers that Treasury or ONS have published. This may be because:

- OSCAR has a single structure that is updated to reflect the latest classification of spend and organisation of government. The snapshots taken by the Treasury at defined points (e.g. to enable reports to be published) contain certain key fields, which then reference the latest structure. Any changes to the structure since the report was published may mean that it is not possible to recreate published figures;
- The time difference between the publication of aggregated information and this release of data, with the potential that data have been updated between the differing points in time.
- Not all the data used to calculate the published numbers are sourced from OSCAR

Quality assurance

Budgeting (Plans and Outturn) and Forecast Outturn Data

The data on OSCAR are quality-assured and complete at the level at which they are required for the following purposes:

- fiscal management;
- operational publications (e.g. Main and Supplementary Estimates); and
- Statistical publications (e.g. Public Expenditure Statistical Analyses, the joint ONS/Treasury Public Sector Finances statistical bulletin and the National Accounts).

Lower levels of data are not quality assured by the Treasury. Individual departments can to some extent choose the level of granularity that they use within pre-defined aggregates set by the Treasury. Lower level detailed data may therefore appear incomplete and be inconsistent across departments.

Whole of Government Accounts (WGA)

WGA data is released on a separate cycle, due to the requirement that the WGA account is published before we release the underlying raw data. Hence there will be no WGA data from OSCAR included in the October 2017 data files. Raw data for WGA will form a separate release, available at <https://www.gov.uk/government/collections/whole-of-government-accounts>.

The Data Files and Downloading

The data released in October 2017 are in a number of compressed text type files containing "raw" delimited data. The information released covers historic years only, from 2012-13 to 2016-17. Certain limited data, for example that relating to national security, has not been included in the data released.

The data files were generated from the live OSCAR system on Wednesday 18th October 2017 and there is a file for data relating to each year in scope, e.g. 2017 Oscar Extract 20xx-xx.txt

Each file contains a string of characters that represent a row of data delimited by the “pipe” symbol, ‘|’ (normally obtained with the ‘SHIFT+\’ keystroke).

Due to the size of the raw files they are provided as compressed zip files. Please note that when uncompressed they may be twenty or more times the size.

Once downloaded and unzipped the files will need to be imported into an appropriate medium, such as an SQL database or an Excel worksheet, to allow them to be properly accessed and manipulated. The first line of each file lists the column headers. Some files may be too large to import into certain versions of Excel.

OSCAR is a live system and as such is updated frequently, often on a daily basis, by departments and by Treasury. Therefore, some of the data in this release may have been subsequently adjusted after being extracted. These adjusted figures will be available in the next annual data release.

Files and fields in the data (Data structure/attributes)

As previously described, there is a file for each year called **2017 Oscar Extract 20xx-xx.csv**. Data aggregates may be derived from each of them.

The following data fields are contained within the files:

Version – This comprises **Plans** and each in-year return. In-year returns start with **Return0** (a forward profile loaded in April) followed by **Return1** (actuals for April, forecasts for May-Mar), **R2** (actuals for Apr-May, forecasts for Jun-Mar) etc all the way through to **Return13/Outturn**, which will hold actuals for Apr-Mar and will be loaded in May after the end of the financial year.

Type – This is a ‘reason for change’ field.

Status – Data is held as **Draft**, **Proposed** or **Confirmed**. Data is loaded onto OSCAR as **Draft**. When valid data is submitted, it will either be moved to **Proposed** status (where Treasury review is required) or **Confirmed** status (where no Treasury review is required). Once reviewed by Treasury, data in **Proposed** status will then be either returned to **Draft** or moved to **Confirmed**.

Month – This is the period to which the spending relates. Plans are annual and so are held against **Month 0**. Outturn is held against the relevant month within each financial year, which includes a **Period 13** (for year-end adjustments) and a **Period 14** (for subsequent classification and machinery of government changes) in addition to **Apr-Mar**.

Account Level 5 – This reflects the detailed accounting treatment of the spending.

Sub-segment – Departments and the Treasury have agreed segmental reporting structures for OSCAR. The sub-segment is the most detailed meaningful breakdown of departments’ spending.

Organisation – The organisation is the body responsible for the spending, and includes all arms length bodies that are identified in Note E of the Estimate i.e. it separately identifies spending by executive non-departmental public bodies (**NDPBs**) as well as by **core departments**, **pension schemes**, **devolved administrations** and **public corporations**.

Counter-party – This is the ‘other’ organisation involved in a **Budget Cover Transfer** or **Machinery of Government** change.

Effective date – This is the point in time at which data feeds into outputs. In most cases the effective date will be the date on which the data was loaded, but sometimes data is loaded with an effective date in the past (Treasury teams making retrospective changes to outputs) or in the future (enabling Departments to load data that will impact future publications e.g. an adjustment to the next Main Estimates that should not impact the upcoming Supplementary Estimates).

Data ID (system generated) – OSCAR generates an ID when data is transferred from a data entry grid into the data warehouse. A single ID is applied to all data within each data entry grid at the point at which the data is transferred.

Load Type (system generated) – OSCAR identifies from which data entry grid the data was loaded into the data warehouse. This is largely designed to enable users to identify whether data was loaded by departments or the Treasury.

A full list of the data fields is provided in the accompanying spreadsheet:

OSCAR_Fields_In_Data_Release

Aggregation and Use of Data

Please note that the information that follows is not comprehensive but provides a guide to some aspects and examples of aggregation that can be performed.

Data are recorded on OSCAR in £k multiples. A figure of 1 therefore represents £1,000, a figure of 2000 represents £2,000,000 and so on.

To arrive at meaningful numbers, various elements of the data need to be combined in differing combinations depending on the aggregate required. It is essential that when aggregating data to ensure that the data for inclusion are specified so that only right data items are included. For example if all data for a given year for a specific account code were to be aggregated, it could potentially bring through the following values:

- Forecast outturn for each of twelve months in the year (a return is submitted monthly);
- Plans data, representing at least the three main status levels for an adjustment, e.g. draft , proposed and confirmed which are all separately held records, and included in these data releases;
- Outturn data – as above for Plans information, which applies also to Outturn data (held against Version = R13).

If all of the data items listed above were brought through for a specific account, the numbers returned could potentially be several times higher than the real figure. You need to decide which stream you are looking at and select carefully.

Also note that within each data stream there can also be legitimately differing figures for a given account and/or sub-segment and chart of account code combination, given the different points in time data are recorded.

To help you choose which data you want to look at, the following may be of assistance.

Plans data starts with the settlement following a spending review announcement. It can be continually updated up to the point the Supplementary Estimates are completed for the year to which the plan relates. The Supplementary Estimate that takes place during a given financial year is the latest point a department can update their planned expenditure for that year. Plans data can therefore reflect both data recorded prior to the plan year and during the plan year. To view Plans data you would filter the Version field to only include 'Plans'.

Outturn represents data recorded after the end of the financial year, consistent with resource accounts. As with all the data and data types, changes may be made to reflect the impact of Machinery of Government, budget regime and other classification changes, which impact on the data for all maintained years. To view annual Outturn data you would filter the Version field to only include 'R13'.

Forecast Outturn data represents monthly returns of estimated actual data for past months and forecast data for future months, reported monthly during the current year. Figures can change month on month, for both actual and forecasts, throughout the year. To view the monthly forecast outturn data, care is required over the selection of both the Version and the Period dimensions. For each Version, there will be twelve Periods.

It is important to bear in mind that some departments, by arrangement with HM Treasury, supply monthly data either one or two month in arrears.

The accompanying spreadsheet titled "**Understanding_Period_Version_Relationship**" will help on this.

Defining the data required for a given aggregate can be very complex. The example below shows the "fields" and settings needed for those fields to produce a DEL resource plans number for a given department and year:

Version = 'Plans';

Organisation_long_name = set to desired organisation(s);

YearNo = set to desired year(s);

Month = 'Period 0 - xx-xx' where xx-xx is the financial year being queried;

Control_Budget_LO_long_name = 'Departmental Expenditure Limits'

Economic_Budget_Code = 'RESOURCE'

Other Sources of Information

A number of information documents are available on the Treasury public website that may assist in understanding COINS/OSCAR data and their recording, and how that data may be used and aggregated. These include:

How to understand public sector spending, available here:

<https://www.gov.uk/government/publications/how-to-understand-public-sector-spending>

Consolidated Budgeting Guidance, available here:

<https://www.gov.uk/government/publications/consolidated-budgeting-guidance-2014-to-2015>

Introduction to Classification, available here:

<https://www.gov.uk/government/publications/introduction-to-classification>

Supply Estimates guidance manual, available here:

<https://www.gov.uk/government/publications/supply-estimates-guidance-manual>

Slightly differing approaches are taken to some data recording and outputs for the Devolved Administrations. Some information on that, by means of concordats, is available here:

http://webarchive.nationalarchives.gov.uk/20130129110402/http://www.hm-treasury.gov.uk/psr_devolved_assemblies.htm

Treasury publishes a number of outputs using OSCAR data. Public Spending Statistics are available here:

<https://www.gov.uk/government/collections/public-expenditure-statistical-analyses-pesa>