



## **Annex D - DATA QUALITY ASSURANCE PROCESS**

The following sections highlight the quality assurance process that is carried out on the administrative data that is used in the **UK House Price Index (HPI)**.

### **1. Summary of Process**

- Extract house sale data from the Land Register
- Run Macro to highlight possible errors or issues
- Do QA – identify duplicates, price paid errors, blank fields, date of entry errors, etc
- Issue data amendments/errors to quality team
- Export data for loading into database
- Run new build and cash sales reports and add values to cash and new build fields
- Load data into database
- Run local authority, cancellation and data amendments reports and update database.

### **2. QA Investigation**

Each week the previous week's application data is extracted from the Land Register. The data contains all applications where there has been a transfer of title to property.

Applications containing transfers of title are identified using deed codes that are allocated by RoS when an application for registration is received.

The data is run through a macro to highlight any possible errors or issues. The entries that are highlighted by the macro are shown below:

#### Value Check

- This highlights entries where the consideration (the price paid) and the value do not match. Such entries are likely to relate to property sales where the full market value is not being paid. If that is the case then these entries are coded so that they are not used within any house price statistical analysis.

#### Consideration Check

- Converts a text number entry to a number. Where the consideration field contains text, the sale is unlikely to be a full market value sale, so the entry is coded so that it is not included in the house price statistics.
- Highlights where consideration = zero. These will not be full market value sales so entries are coded so they are not included in the house price statistics.
- Highlights entries that include text as well as a numerical value. Entries can include text for many reasons, for example where the deed has stated no VAT is payable. The text is deleted from the entry.
- Highlights £750,000 and above entries. The deed is checked to ensure that the figures are accurate, and the land class is correct.
- Highlights £20,000 and below entries. The deed is checked to ensure that the figures are accurate and to check the correct property class has been recorded. If the sale is not for full market value, it is coded so that it is not included in the house price statistics.



### Land Class / Address Check

- Highlights blank entries. The deed is checked and the land class is populated.
- Highlights Commercial entries with "flat" or "dwellinghouse" in suffix/prefix. The deed is checked to ensure the correct land class is entered.
- Highlights Residential entries with "shop" or "hotel" in suffix/prefix. The deed is checked to ensure the correct land class is entered.
- Highlights Residential entries with hectares, metres or industrial. The deed is checked to ensure the correct land class is entered.
- Highlights Residential entries with "*pro indiviso*" mentioned in suffix/prefix (*pro indiviso* sale is a term used for shares of a property). The deed is checked to see if this is a sale for the whole property. If it is just for part, the entry is coded so that it is not included in the house price statistics.
- Highlights Residential entries containing areas of ground/land. The deed is checked to ensure the correct land class is entered.
- Highlights Land and Other entries with shop, house or flat. The deed is checked to ensure the correct land class is entered.

### Duplicates Check

- Highlights duplicate application and title numbers. There are a number of reasons why duplicate application and title numbers might appear within the data. They are usually for internal administrative reasons. Any duplicate application numbers are removed so that a single sale is not counted more than once.

### Post Town Check

- Highlights blank post town entries. All post towns are checked against the official [post town list](#) and all incorrect post towns are updated and noted (errors are recorded for quality purposes). The deed is checked and the correct post town added to the entry. This is to support the quality of the addressing within the data.

## **3. Final Checks and Additions**

Once all issues and errors have been investigated, a final eye check is done over the lines of data to ensure nothing has been missed. This includes checking missing or incorrect dates of entry (date that new owner took possession of the title).

A list of any errors or omissions is compiled for quality purposes. The quality team will also provide details of any corrections or rectifications they have made to titles or pending applications so that these can be accounted for within the data.



**Registers  
of Scotland**  
ros.gov.uk

The data file is then matched with a builders' title list to identify any possible 'new-build'<sup>1</sup> transactions. These are checked to ascertain if the property is part of a new development and, if so, they are coded to allow them to be identified within the data.

RoS house class methodology<sup>2</sup> is used to identify house classes for residential properties (i.e. terraced, detached, semi-detached, flat). The house class is added to the data.

Cash sales are also identified and coded. Sales are considered to be cash sales if they are submitted without a mortgage transaction and if a mortgage transaction is not submitted in the same month of registration as the sale or in the two months following.

The data files are then loaded into the database.

A local authority report is run to identify any missing local authorities, if missing these are added in and updated on the database.

A regular report is run to identify any applications that have been cancelled or amended following their receipt and they are coded within the database so that they are not used within any house price statistics.

---

<sup>1</sup> The new-build methodology only relates to single properties sold out of a major builder's development title and not to smaller developments or single new build properties by private or small-scale builders.

<sup>2</sup> Further information about this methodology is available in our [guidance notes](#).