
Data Preparation

Relevant, accurate, up-to-date, data is the back bone of any Geographic Information System (GIS). This guide aims to define the data standard required for the supply of data to be hosted by and served from Spatial NI.

The following GIS data formats and projections will be accepted by Spatial NI:

- An Environmental Systems Research Institute (ESRI) geodatabase;
- ESRI Shapefiles;
- MapInfo formats;
- Comma Separated Values (CSV) with Eastings and Northings or Unique Property Reference Number (UPRN)
- Text File format

Projections:

- Irish Grid (IG) Transverse Mercator (TM) 65 - EPSG: 29902
- IG TM 75 - EPSG 29903
- Irish Transverse Mercator (ITM) - EPSG 2157

Spatial NI also require the following supporting information for each dataset:

- Details of the data structure (e.g. map document (MXD))
- Details on the relevant symbology (e.g. Layer Files)

Datasets should only contain data pertaining to the data topic. Spatial NI would request that all other irrelevant data be removed.

The Spatial NI Team will not provide any transformation or processing of data.

ESRI GDB and Shapefile Guidelines:



You must supply an MXD and Layer Files for each dataset.

In the MXD, please complete the following:

Map Document Properties

- Title
- Summary
- Description
- Description of data and all sub-layers
- Data owner details
- Scale thresholds
- Tags

Feature Dataset Properties (if applied)

- Feature Dataset name
- Description
- Credits

Feature Layer Properties

- Feature Layer name
- Description
- Credits
- Apply scale range (if required)
- Fields
 - Remove fields which are not required
 - Leave OBJECTID and Shape visible
 - Set field order with OBJECTID and Shape the last records in the list

Remove any references to data that is not included in the geodatabase or shapefiles.

MapInfo File Guidelines:



MapInfo File Guidelines:

You should supply your MapInfo data to Spatial NI in MIF/MID format. Ensure that you have included the following:

- MIF File Header
- Pen styles
- Brush styles
- Symbol styles
- Font styles
- Colours

Comma Separated Values (CSV) Guidelines:



General

- Ensure if you have more than one active worksheet in your workbook that each individual worksheet is given an appropriate name.
- Delete worksheets that contain no data.
- Ensure that you have included either Irish Grid Eastings and Northings or a UPRN for each record
- Ensure that all data is subdivided into the smallest units possible and added as individual columns within the worksheet. For example, if entering address information, ensure the house number, street name, town, county and postcode are all in separate columns.
- Ensure you have removed all formulas from your worksheet. If you have formulas in your worksheet, remove them by pasting the data into a new column using the paste special—values only option. Then delete the original column that contains the formulas from your worksheet.
- If you have formulas at the base of columns for example summing or totalling the column, remove these from your worksheet.
- Ensure there are no empty columns or rows in your worksheet.

Columns

- The first row in a worksheet should contain the column names. If the first row contains other content, you should modify the worksheet so that the first row contains the column names.
- Always input the column names with capital letters
- Column names must begin with a letter (no numbers or underscores).
- Column names must contain only letters, numbers or underscores (no spaces, hyphens or special characters).
- Column names must be less than 64 characters (these should be a maximum of 10 characters if possible).
- It is important in Excel that you format the cells in each column to be a particular data type. Columns should not contain multiple data types for various cells as ArcGIS scans the first eight rows in each column to determine the column data type.
- The building number (as part of an address) column should be formatted as text as this column may contain more than numeric data, e.g. 17A.
- If inputting co-ordinates, always use the column names, X_COORD and Y_COORD respectively.

Cell Data

- ArcGIS reads only the first 255 characters in a cell. It is important that you check your worksheet to look for cells with lots of data and truncate the data if necessary.
- Ensure you are consistent when entering data which may be repeated throughout the dataset. For example, if you are entering Newry and Mourne, you should always enter the data in this format. Entering Newry & Mourne would break these data integrity rules.
- Ensure any statistical/numerical data is held in a numerical field and that there are no gaps in the data. For any cell with no value, enter a zero.
- When entering numerical data with decimal places, ensure you setup the number of decimal places for the entire column to eliminate error and to maintain data integrity.
- When entering temporal data with date fields, ensure you set the columns in use to format as date and use the format 14/03/2001.
- When entering a postcode, please ensure you use the format with a space between the components, e.g. enter BT80 8PZ and not BT808PZ.
- When entering map co-ordinates, always use a 6 figure grid reference. For example, Belfast City Hall would be entered as 333829, 373988.

Example CSV:

REF_NO	ORGANISATION	BUILD_NAME	BUILD_NO	ADD_LINE1	ADD_LINE2	ADD_LINE3	TOWN	COUNTY	POSTCODE	UPRN	X_COORD	Y_COORD
1	Land & Property Services	Lanyon Plaza	7	Lanyon Place			Belfast	Antrim	BT1 3LP	187276084	334520	373964