

# Australian datasets and Django

Malcolm Tredinnick  
June, 2010

Importing data is a puzzle to solve

Motivation...



**“The dataset creator hates you!”**

# General Principles

- Allow restarting
- Assert for all assumptions
- Use brain and eyes before computer
- Use all available tools, not just Python
  - grep
  - cut
  - head / tail
  - ...

# Case study #1





“Powerhouse Museum” by Beth Kanter, CC Attribution 2.0 licensed.



# Collection Data

- 75,814 items in a text file
- Creative Commons license
- Appears “mostly sane”



**Worked examples...**

## Case study #2



# Australian River Basins



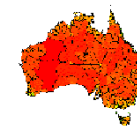
[Home](#) | [About Us](#) | [News and Media](#) | [Products](#) | [Education](#) | [Jobs](#) | [Contact Us](#)

You are here: [Home](#) > [Free Data Downloads](#)

Updated: 16 November 2009

## Australia's River Basins 1997

**Catalogue Number :** 42343  
**Release Date :** 8 October 2007  
**Senior Author :** null  
**Product Type :** GIS Dataset-Thematic  
**Volume/Map ID :** null  
**Scale :**  
**Medium :** WEB  
**Theme(s) :** GIS, boundaries, environmental, hydrology, mapping



Shows the boundaries of the Australian basins as defined by the Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ). This data shows boundary and attribute information for 12 divisions, 77 regions and 245 basins. It also contains, for each basin, information relating to its individual basin/region/division name and number. State borders are also included in the data. Data for Division XIII Distant Islands Division is not included.

These basins are the primary building block for the collection of national hydrologic data and the assessment of water resources. Data are suitable for GIS applications.

Free online download. Available in ArcView Shapefile and MapInfo mid/mif.

### Product Specifications

**Coverage:** Australia

**Currency:** June 1997

**Coordinates:** Geographical

**Datum:** AGD66

**Format:** ArcInfo Export, ArcView Shapefile and MapInfo mid/mif

**Medium:** Free online and CD-ROM (fee applies)

**Spatial extent :** **N Lat :**9.0 **W Long :**112.0  
**S Lat :**44.0 **E Long :**154.0

**Comments :** Data re-released in 2004 with minor changes.

**Free Resources :** [Get PDF Reader](#)  
[Free Download: Australia's River Basins 1997](#)  
[User Guide \[PDF 147KB\]](#)

**Extended Metadata :** [Australia's River Basins 1997](#)

For further information on Geoscience Australia's products please contact the [Sales Centre](#).

Unless otherwise noted, all Geoscience Australia material on this website is licensed under the [Creative Commons Attribution 3.0 Australia Licence](#).

[feedback](#) | [security & privacy](#) | [glossary](#) | [disclaimer](#) | | © Commonwealth of Australia 2010

# GIS Dataset

- Creative Commons Attribution license
- Polygons, lines, points (labels)
- Well documented
- Not entirely trivial on the surface



# GeoDjango overview

**Worked examples...**



<http://github.com/malcolmt/pycon-au-2010>