Download biodiversity data with galah: : **CHEAT SHEET** python™

galah is an interface to biodiversity data hosted by the [Atlas of Living Australia (ALA)](https://www.ala.org.au). It enables users to locate and download species occurrence records (observations, specimens, eDNA records, etc.), taxonomic information, or associated media such as images or sounds, and to restrict their queries to particular taxa or locations.

Build a query

No matter what kind of data you want to return in Python, **every download query** consists of the **same building blocks**.

1. Start a query with a **galah.atlas_ function**, which specifies the data your query will return.
2. **Modify a query** by adding filters, specifying taxa or other options
3. Output data to screen or store in variable

AN EXAMPLE QUERY:

```
data = galah.atlas_occurrences(
    taxa = “reptilia”,
    filters = [“year>2010”,”cl22=Victoria”],
)
data
```

Start a query

Modify a query

Data is contained in variable data

FOR MORE PYTHON INFORMATION:

Check out our ALA galah Python package & guides:

Download data

COUNTS

galah.atlas_counts()

Return the number of records that match a query

Return number of observations

coun t
40284

```
galah.atlas_counts(
    taxa=“reptilia”,
    filters=“year=2020”
)
```

Return number of species

count
537

```
galah.atlas_counts(
  taxa=“reptilia”,
  filters=“year=2020”,
  group_by= “species”,
  expand=False,
  total_group_by=True
)
```

Return grouped counts

```
galah.atlas_counts(
  taxa=“reptilia”,
  filters=“year=2020”,
  group_by=“order”,
  expand=False
)
```

order	count
Crocodylia	6388
Rhynchocephalia	1
Squamata	29679
Testudines	4304

OCCURRENCES

galah.atlas_occurrences()

Return species occurrence records that match a query

```
galah.galah_config(email=“your-email-here”)
galah.atlas_occurrences(
    taxa=“perameles”,
    filters=“year>2021”
)
```

Must use an ALA-registered email

Register at ala.org.au/

decimalLongitud e	decimalLatitude	eventDate	scientificName
-43.2	148.	2023-01-06 12:46:00	*Perameles gunnii*
-43.1	147.	2022-10-07 10:38:26	*Perameles gunnii*
-43.1	148.	2022-09-18 10:13:00	*Perameles gunnii*

… i = 1,658 more rows … j = 4 more columns

SPECIES LISTS

galah.atlas_species()

Return species information for each species that match a query

```
galah.galah_config(email=“your-email-here”)
galah.atlas_species(
    taxa=“perameles”),
    filters=“year>2021”
)
```

Must use an ALA-registered email

Register at ala.org.au/

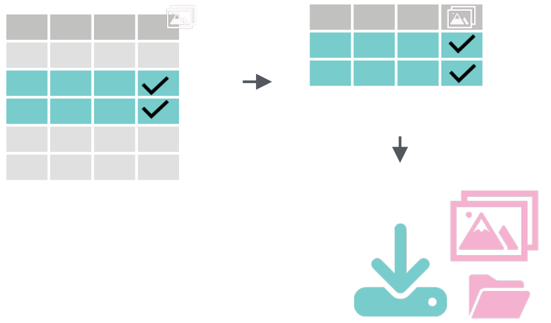
kingdom	phylum	class	order	family	genus	species
Animalia	Chordata	Mammalia	Peramelemorphia	Peramelidae	Perameles	*Perameles nasuta*
Animalia	Chordata	Mammalia	Peramelemorphia	Peramelidae	Perameles	*Parameles gunnii*
Animalia	Chordata	Mammalia	Peramelemorphia	Peramelidae	Perameles	*Perameles notina*
Animalia	Chordata	Mammalia	Peramelemorphia	Peramelidae	Perameles	*Perameles fasciata*

i = 1 more row

MEDIA & IMAGES

galah.atlas_media()

Return species occurrence records with associated images, sounds or media that match a query. Return matching records with atlas_media(), then add `collect=True` and specify a `path` to download locally



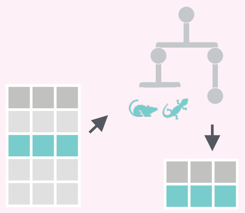
```
galah.atlas_media(
    taxa=“perameles”,
    filters=”year>2021”,
)
```

Returns output like atlas_occurrences() with media metadata

```
galah.atlas_media(
    taxa=“perameles”,
    filters=”year>2021”,
    collect=True,
    path = “path-to-folder”
)
```

Download media to folder

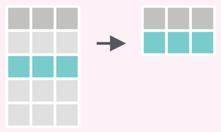
MODIFY QUERY ON THE SERVER BEFORE DOWNLOADING:



taxa=...

Filter query to specific identified taxa

```
galah.atlas_counts(
    taxa=[“mammalia”, “reptilia”]
)
```



filters=...

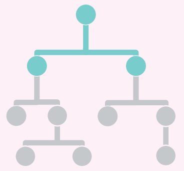
Filter query to rows that meet a logical criteria

```
galah.atlas_counts(
    filters=“year=2020”
)
```

group_by=...

Filter query to rows that meet a logical criteria

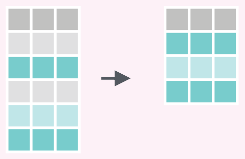
```
galah.atlas_counts(
    group_by=[“year”, “species”],
)
```



rank=...

Specify the lowest taxonomic level in a query for a species list

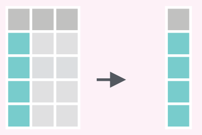
```
galah.atlas_species(
    taxa=“fungi”,
    rank=“phylum”
)
```



data_profile=...

Apply a set of data quality filters to narrow a query

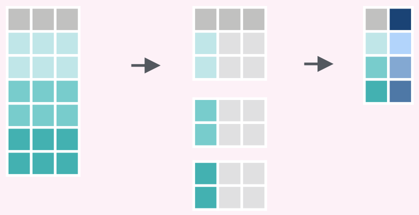
```
galah.galah_config(data_profile=”ALA”)
galah.atlas_counts(
    use_data_profile=True,
)
```

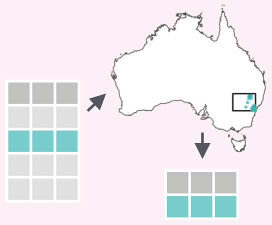


select=...

Filter query to return specified columns

```
galah.atlas_occurrences(
    select=[“scientificName”,
            “eventDate”]
)
```





polygon=... | bbox = ...

Specify the location of data returned with a polygon or bounding box

```
bbox = shapely.box(143,-29,148,-28)
galah.atlas_counts(
    bbox = bbox
)
```

Atlas of Living Australia
ala.org.au

Choose an atlas

The Global Biodiversity Information Facility (GBIF) network consists of a series of ‘nodes’ — **Living Atlases**—that collate biodiversity data from their own countries. GBIF acts as an umbrella organisation to store data from all nodes.

galah supports data downloads for 5 Living Atlases and GBIF.

See full list:
https://galah.ala.org.au/Python/galah_user_guide/Choosing_Atlas.html

```
galah.galah_config(atlas = “Austria”)

galah.galah_config(atlas = “GBIF”)
```

Species list in your area

If you want to know what species are in your area, you can use **atlas_species** to do this.

Species list for year 2022 in Victoria

For this example, we know that our filters are `cl22=Victoria` and `year=2022`. The Python code then looks like this:

```
galah.galah_config(email=“your-email—here”)
species_list = galah.atlas_species(
   filters=[“cl22=Victoria”,”year=2022”]
)
```

Finally, write data to a csv file.

```
species_list.to_csv(“NAME_OF_FILE.csv”)
```

Getting counts of species in your area

If you want to get the species name and counts, we will use the above example to get the initial species list. We will then get the `species` column, and use that as input for `atlas_counts`:

```
species_list_counts=galah.atlas_counts(
   taxa=list(species_list[“species”]),
   filters=[“cl22=Victoria”,”year=2022”],
   group_by=“species”,
   expand=False
)
species_list_counts
```

species	count
Aacanthocnema dobsoni	3
Aades cultratus	8

… i = 9,579 more rows

Finally, write data to a csv file.

```
species_list_counts.to_csv(“NAME_OF_FILE.csv”)
```

Atlas of Living Australia
ala.org.au

Lookup information — galah provides look-up functions to help users find ways to modify their queries

The living atlases store a huge amount of information, above and beyond only occurrence records. This information can be useful for modifying queries.

Taxonomic information

Look up taxonomic names before downloading data from the ALA using **atlas_** functions.

search_taxa(…**)** Search for valid taxonomic names, look up taxonomic information, disambiguate homonyms

```
galah.search_taxa(taxa=[“reptilia”, “mammalia”])
```

Specify taxonomic levels in a list using “specificEpithet”

```
galah.search_taxa(
   specific_epithet=[“class=aves”,
     ”family=pardalotidae”,”genus=pardalotus”,
     “specificEpithet=punctatus”]
)
```

Specify taxonomic levels in a dictionary using “scientificName”

```
galah.search_taxa(
   scientific_name={
     ”family”: [“pardalotidae”,”maluridae”],
     “scentificName”: [“pardolatus striatus”,
                          “malurus cyaneus”]}
)
```

Search for unique identifiers of a taxon. Identifiers are assigned by atlases to identify all taxonomic clades.

```
galah.search_taxa(
   identifiers=
   “https://id.biodiversity.org.au/node/apni/2914510”
)
```

galah.show_all(type=True**)**
Show all available options or categories for a specified type of information

Configuration

atlases

Show what atlases are available

```
galah.show_all(atlases=True)
galah.search_all(atlases=“Brazil”)
```

apis

Show what APIs & functions are available

```
galah.show_all(apis=True)
galah.search_all(apis=“counts”)
```

reasons

Show what values are acceptable as “download reasons” for a specified atlas

```
galah.show_all(reasons=True)
galah.search_all(reasons=“research”)
```

Filters

fields

Show fields that are stored in an atlas

```
galah.show_all(fields=True)
galah.search_all(fields=“state”)
```

assertions

Show data quality checks run by each atlas

```
galah.show_all(assertions=True)
galah.search_all(assertions=“longitude”)
```

licenses

Show what copyright licenses are applied to media

```
galah.show_all(licenses=True)
galah.search_all(licenses=“CC BY”)
```

galah.search_all(type=“query”**)**
Search for a specific option or category for a specified type of information

Taxonomy

ranks

Show valid taxonomic ranks (eg Class)

```
galah.show_all(ranks)
galah.search_all(ranks=“suborder”)
```

Group filters

profiles

Show what data profiles (sets of data quality filters) are available

```
galah.show_all(profiles=True)
galah.search_all(profiles=“ALA”)
```

lists

Show what species lists are available

```
galah.show_all(lists=True)
galah.search_all(lists=“EPBC”)
```

Data providers

providers

Show which institutions provide data

```
galah.show_all(providers)
galah.search_all(providers=“botanic”)
```

collections

Show specific collections within institutions

```
galah.show_all(providers=True)
galah.search_all(collections= “antarctic”)
```

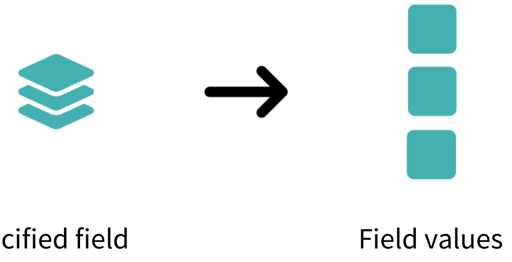
datasets

Show all data groupings collections

```
galah.show_all(datasets)
galah.search_all(datasets=“river”)
```

Values

Users may wish to see what values are **within** a chosen field, profile or list to modify a query or understand more about the information of interest. The **values functions** get their inputs from show_all(fields=True) or search_all(fields=“field-ID”). **Feed this field ID to show_values or search_values** to see this information.



show_values(field=“field-ID”**)**
Search for valid taxonomic names, look up taxonomic information, disambiguate homonyms

```
galah.show_values(field=“cl22”)
```

field	count
cl22	New South Wales
cl22	Victoria
cl22	Queensland
cl22	South Australia
cl22	Western Australia
cl22	Northern Territory
cl22	Australian Capital Territory
cl22	Tasmania

search_values(field=“field-ID”, value=“value”**)**
Search for valid taxonomic names, look up taxonomic information, disambiguate homonyms

```
galah.search_values(
    field=“cl22”,
    value=“territory”
)
```

field	count
cl22	Northern Territory
cl22	Australian Capital Territory

Want ideas on how to visualise your data?

For more information, including tutorials & examples, check out **ALA LABS.**

CC BY SA Atlas of Living Australia • support@ala.org.au • ala.org.au• Learn more at **galah.ala.org.au or galah.ala.org.au/Getting-Started** • package version 0.6.0 • Updated: 2023-10