ragnar

Release MIT

Aug 20, 2022

Contents

Ragnar is a lightweight Extract-Transform-Load (ETL) framework for Python 3.5+.

- Free software: MIT license
- Documentation: https://ragnar.readthedocs.io.

CHAPTER 1

Features

- Keeps a functional programming philosophy.
- Code reuse instead of “re-inventing the wheel” in each script.
- Customizable for your organization’s particular tasks.

CHAPTER 2

Example

A pipeline that applies capital letters to the list and then filters through the one starting with "B":

```
>>> from ragnar.stream import Stream
>>> st = Stream(["apple", "banana", "cherry"])
>>> st.do(lambda x: x.upper())
<ragnar.stream.Stream object at 0x7fbe8e3509d0>
>>> st.filter(lambda x:x.startswith("B"))
<ragnar.stream.Stream object at 0x7fbe8e3509d0>
>>> for row in st:
...     print(row)
BANANA
```

2.1 Contents

2.1.1 Installation and Dependencies

Ragnar is pure Python and so is easily installable by the standard dependency manager `pip`:

```
pip install ragnar
```

Ragnar endeavors to be a very light dependency. It accomplishes this in three ways:

1. Ragnar is pure Python
2. Ragnar relies only on the standard library
3. Ragnar simultaneously supports Python versions 3.5+ and PyPy

2.1.2 Ragnar API

This page contains a comprehensive list of all functions within `ragnar`. Docstrings should provide sufficient understanding for any individual function.

Stream

class ragnar.stream.**Stream**(*value*, ***kwargs*)
Bases: object

This object brings together the advantages of generators and the functional programming paradigm. Stream is an object that allows accumulating actions to be applied to a dataset.

Parameters **repeatable** (*bool, optional*) – This parameter is used to enable repeating iterations if *'True'* allows to iterate as many times as required

do(*func*, *chain=False*)
This method adds a function to apply to the execution stack.

Parameters

- **func** – method to be included in the execution stack
- **chain** (*bool, optional*) – The results are merged into a single dataset. For example if you read multiple files the results are merged to loop like a single list.

filter(*func*)
This method adds a filter to apply to the execution stack.

Parameters **func** – method to be included in the execution stack. It must be a function that returns a boolean value, otherwise the filter is not applied.

ragnar.stream.**func_cat**(*value*, **forms*)

Objstream

class ragnar.objstream.**ObjStream**(*value*, ***kwargs*)
Bases: *ragnar.stream.Stream*

This object inherits from the Stream object, specific for record-based data.

Parameters

- **columns** – Object headers.
- **skip_first** – 0mits the first row (usually headers)

applyto(*fields: Union[str, list], func, entire_object=False*)
This method adds a function to apply to the execution stack.

Parameters

- **fields** (*typing.Union[list, str]*) – fields where to apply the function.
- **func** – method to be included in the execution stack
- **entire_object** (*bool, optional*) – passing the field or the entire record as a dictionary to the function

Python Module Index

r

Index