Example Project

Generated by Doxygen 1.9.1

1 Hierarchical Index 1

1.1 Class Hierarchy . . . 1

2 Class Index 3

2.1 Class List . . . 3

3 Class Documentation 5

3.1 project.crawler.base.BaseCrawler Class Reference . . . 5

3.2 project.convert.converter.Converter Class Reference . . . 6

3.2.1 Member Function Documentation . . . 7

3.2.1.1 to_df() . . . 7

3.3 project.crawler.manager.CrawlerManager Class Reference . . . 7

3.4 project.process.process.DataProcess Class Reference . . . 8

3.4.1 Detailed Description . . . 8

3.4.2 Member Function Documentation . . . 8

3.4.2.1 analyse() . . . 8

3.4.2.2 convert_to_csv() . . . 8

3.4.2.3 get() . . . 9

3.4.2.4 get_historical() . . . 9

3.4.2.5 get_recent() . . . 9

3.4.2.6 start_crawler() . . . 10

3.5 project.crawler.dwd.DWDCrawler Class Reference . . . 10

3.6 project.convert.dwd_converter.DWDJsonConverter Class Reference . . . 11

3.7 project.analysis.pipeline.JupyterPipeline Class Reference . . . 12

3.8 project.crawler.meteo_blue.MeteoBlueCrawler Class Reference . . . 13

3.9 project.utils.thread.SpinnerThread Class Reference . . . 14

3.10 project.utils.thread.ThreadWithReturnValue Class Reference . . . 15

Index 17

Chapter 1

Hierarchical Index

1.1 Class Hierarchy

This inheritance list is sorted roughly, but not completely, alphabetically:

project.crawler.manager.CrawlerManager 7
project.process.process.DataProcess 8
project.analysis.pipeline.JupyterPipeline 12
threading.Thread
    project.utils.thread.SpinnerThread 14
ABC
    project.convert.converter.Converter 6
        project.convert.dwd_converter.DWDJsonConverter 11
    project.crawler.base.BaseCrawler 5
        project.crawler.dwd.DWDCrawler 10
        project.crawler.meteo_blue.MeteoBlueCrawler 13
Thread
    project.utils.thread.ThreadWithReturnValue 15

Chapter 2

Class Index

2.1 Class List

Here are the classes, structs, unions and interfaces with brief descriptions:

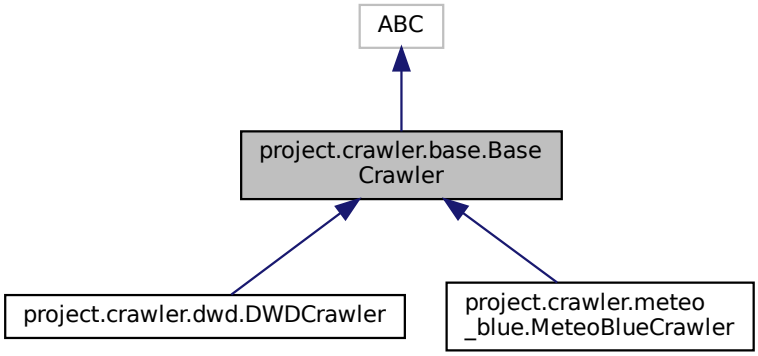
project.crawler.base.BaseCrawler 5
project.convert.converter.Converter 6
project.crawler.manager.CrawlerManager 7
project.process.process.DataProcess
Process for data literacy class 8
project.crawler.dwd.DWDCrawler 10
project.convert.dwd_converter.DWDJsonConverter 11
project.analysis.pipeline.JupyterPipeline 12
project.crawler.meteo_blue.MeteoBlueCrawler 13
project.utils.thread.SpinnerThread 14
project.utils.thread.ThreadWithReturnValue 15

Chapter 3

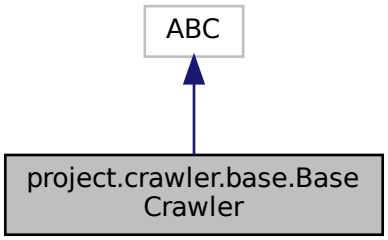
Class Documentation

3.1 project.crawler.base.BaseCrawler Class Reference

Inheritance diagram for project.crawler.base.BaseCrawler:



Collaboration diagram for project.crawler.base.BaseCrawler:



Public Member Functions

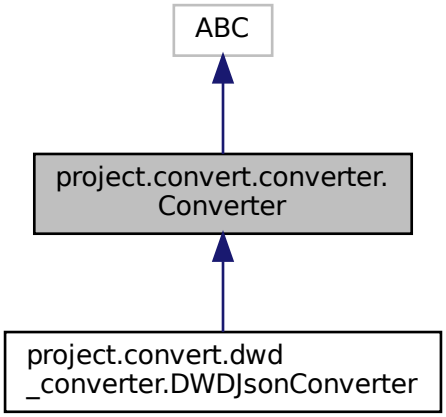
- None **__init__** (self)
- def **get** (self, bool save=False)

The documentation for this class was generated from the following file:

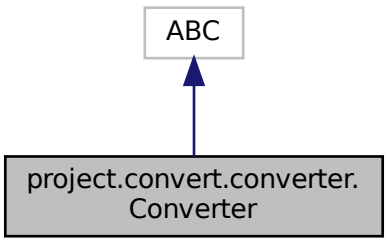
- docs/code/tmp/project/crawler/base.py

3.2 project.convert.converter.Converter Class Reference

Inheritance diagram for project.convert.converter.Converter:



Collaboration diagram for project.convert.converter.Converter:



Public Member Functions

- None **__init__** (self)
- List[DataFrame] to_df (self, object data)

  convert data to a DataFrame

3.2.1 Member Function Documentation

3.2.1.1 to_df()

```
List[DataFrame] project.convert.converter.Converter.to_df (
              self,
             object data )
```

convert data to a DataFrame

Parameters

data	*description*

Exceptions

NotImplementedError	if method is not implemented

Returns

DataFrame data in a DataFrame object

The documentation for this class was generated from the following file:

- docs/code/tmp/project/convert/converter.py

3.3 project.crawler.manager.CrawlerManager Class Reference

Public Member Functions

- None **__init__** (self, List[BaseCrawler] crawler, datetime pull_time)
- def **start** (self)
- def **process** (self)

The documentation for this class was generated from the following file:

- docs/code/tmp/project/crawler/manager.py

3.4 project.process.process.DataProcess Class Reference

process for data literacy class

Public Member Functions

- None __init__ (self)
  summary
- def start_crawler (self, str crawler_config_path="project/config/crawler.config.yaml")
  start crawler to collect weather data from APIs specified in crawler config
- def analyse (self, bool use_active_venv=False)
  start pipeline to analyse data.
- def get (self, bool save=True)
  send request to every embedded crawler and return pandas data frame heads onto terminal
- def get_historical (self, List[int] station_ids, str save_path)
  get historical (precipitation, pressure, air temperature) data from the dwd database
- def get_recent (self, List[int] station_ids, str save_path, List[str] features, bool unpack=True)
  get recent (precipitation, pressure, air temperature) data from the dwd database
- def convert_to_csv (self, str input="data/dwd/raw", str output="data/dwd/csv")
  convert the forecast data into csv format

3.4.1 Detailed Description

process for data literacy class

3.4.2 Member Function Documentation

3.4.2.1 analyse()

```
def project.process.process.DataProcess.analyse (
              self,
             bool  use_active_venv = False )
```

start pipeline to analyse data.

```
        use_active_venv (bool, optional): set this flag if you have poetry installed and would like to run the
```

3.4.2.2 convert_to_csv()

```
def project.process.process.DataProcess.convert_to_csv (
              self,
             str  input = "data/dwd/raw",
             str  output = "data/dwd/csv" )
```

convert the forecast data into csv format

Parameters

input	root folder of all forecast json files. Defaults to "data/dwd/json/raw @param output where to output the csv file structure. Defaults to "data/dwd/raw

3.4.2.3 get()

```
def project.process.process.DataProcess.get (
              self,
             bool  save = True )
```

send request to every embedded crawler and return pandas data frame heads onto terminal

```
        save (bool, optional): if set to true save results. Defaults to True.
```

3.4.2.4 get_historical()

```
def project.process.process.DataProcess.get_historical (
              self,
             List[int] station_ids,
             str save_path )
```

get historical (precipitation, pressure, air temperature) data from the dwd database

Parameters

station_ids	*description*
save_path	*description*

Exceptions

NotImplementedError	*description*

3.4.2.5 get_recent()

```
def project.process.process.DataProcess.get_recent (
              self,
             List[int] station_ids,
             str save_path,
             List[str] features,
             bool  unpack = True )
```

get recent (precipitation, pressure, air temperature) data from the dwd database

Parameters

station_ids	station ids from DWD
save_path	where you want to store the collected information. It will create this directory if it does not exist already.
unpack	if set to true we will also unpack the downloaded zips
features	features you want to extract from DWD API

3.4.2.6 start_crawler()

```
def project.process.process.DataProcess.start_crawler (
              self,
             str  crawler_config_path = "project/config/crawler.config.yaml" )
```

start crawler to collect weather data from APIs specified in crawler config

Parameters

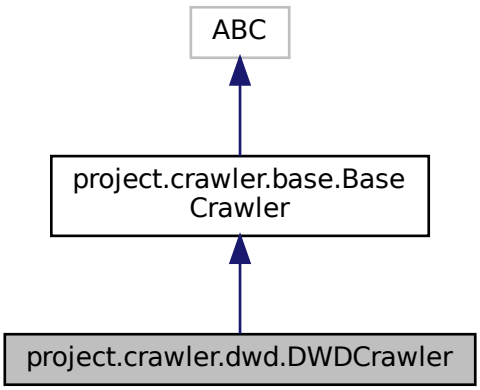
crawler_config_path	crawler config for individual apis

The documentation for this class was generated from the following file:

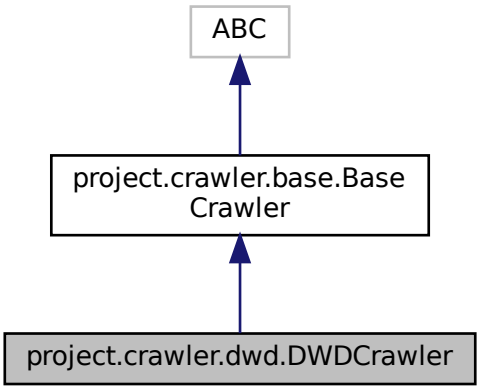
- docs/code/tmp/project/process/process.py

3.5 project.crawler.dwd.DWDCrawler Class Reference

Inheritance diagram for project.crawler.dwd.DWDCrawler:



Collaboration diagram for project.crawler.dwd.DWDCrawler:



Public Member Functions

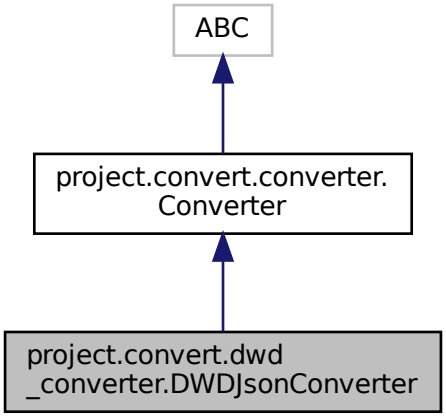
- None **__init__** (self, str api_identifier, List[int] station_ids)

The documentation for this class was generated from the following file:

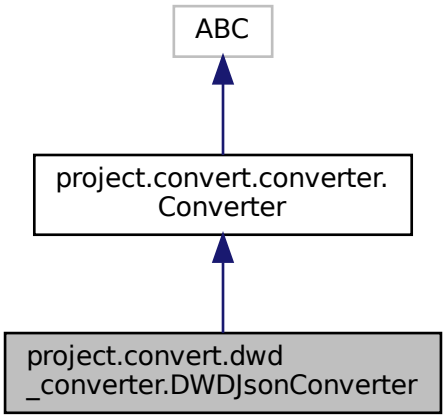
- docs/code/tmp/project/crawler/dwd.py

3.6 project.convert.dwd_converter.DWDJsonConverter Class Reference

Inheritance diagram for project.convert.dwd_converter.DWDJsonConverter:



Collaboration diagram for project.convert.dwd_converter.DWDJsonConverter:



Public Member Functions

- None **__init__** (self)
- List[DataFrame] **to_df** (self, Dict[str, Any] data)

The documentation for this class was generated from the following file:

- docs/code/tmp/project/convert/dwd_converter.py

3.7 project.analysis.pipeline.JupyterPipeline Class Reference

Public Member Functions

- None **__init__** (self, bool use_active_venv=False, str note_book_listing_path="project/config/analysation_↩ scripts.yaml")
- def run (self)

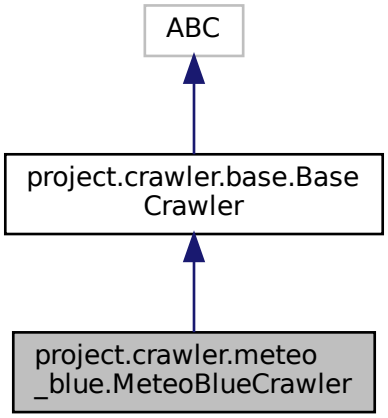
  Start pipeline.

The documentation for this class was generated from the following file:

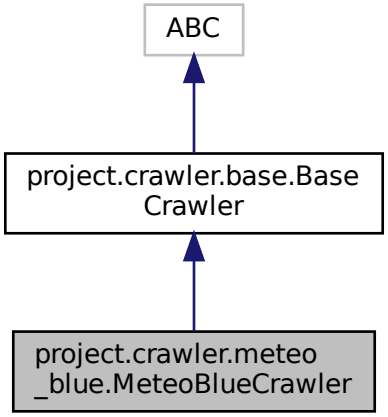
- docs/code/tmp/project/analysis/pipeline.py

3.8 project.crawler.meteo_blue.MeteoBlueCrawler Class Reference

Inheritance diagram for project.crawler.meteo_blue.MeteoBlueCrawler:



Collaboration diagram for project.crawler.meteo_blue.MeteoBlueCrawler:



Public Member Functions

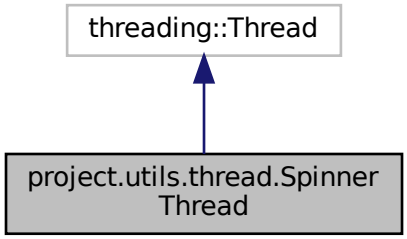
- None **__init__** (self, str key, str package, str format="json")

The documentation for this class was generated from the following file:

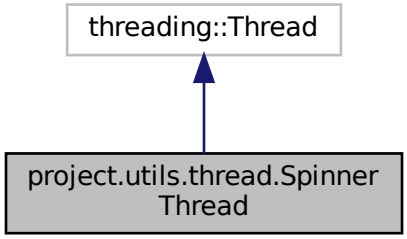
- docs/code/tmp/project/crawler/meteo_blue.py

3.9 project.utils.thread.SpinnerThread Class Reference

Inheritance diagram for project.utils.thread.SpinnerThread:



Collaboration diagram for project.utils.thread.SpinnerThread:



Public Member Functions

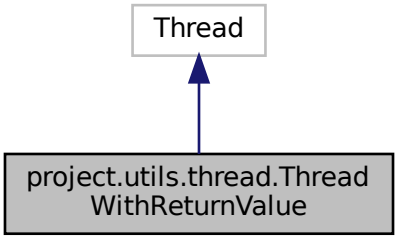
- def **__init__** (self)
- def **stop** (self)

The documentation for this class was generated from the following file:

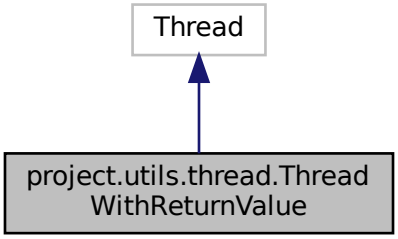
- docs/code/tmp/project/utils/thread.py

3.10 project.utils.thread.ThreadWithReturnValue Class Reference

Inheritance diagram for project.utils.thread.ThreadWithReturnValue:



Collaboration diagram for project.utils.thread.ThreadWithReturnValue:



Public Member Functions

- def **__init__** (self, group=None, target=None, name=None, args=(), kwargs={}, Verbose=None)
- def **run** (self)
- def **join** (self, ∗args)

The documentation for this class was generated from the following file:

- docs/code/tmp/project/utils/thread.py

Index

analyse
    project.process.process.DataProcess, 8

convert_to_csv
    project.process.process.DataProcess, 8

get
    project.process.process.DataProcess, 9
get_historical
    project.process.process.DataProcess, 9
get_recent
    project.process.process.DataProcess, 9

project.analysis.pipeline.JupyterPipeline, 12
project.convert.converter.Converter, 6
    to_df, 7
project.convert.dwd_converter.DWDJsonConverter, 11
project.crawler.base.BaseCrawler, 5
project.crawler.dwd.DWDCrawler, 10
project.crawler.manager.CrawlerManager, 7
project.crawler.meteo_blue.MeteoBlueCrawler, 13
project.process.process.DataProcess, 8
    analyse, 8
    convert_to_csv, 8
    get, 9
    get_historical, 9
    get_recent, 9
    start_crawler, 10
project.utils.thread.SpinnerThread, 14
project.utils.thread.ThreadWithReturnValue, 15

start_crawler
    project.process.process.DataProcess, 10

to_df
    project.convert.converter.Converter, 7