school_scraper

Release Version number declaration

Kevin S. Phillips

Sep 12, 2020

CONTENTS

1 school_scraper **3**
1.1 school_scraper package 3

2 Indices and tables **7**

Python Module Index **9**

Index **11**

Contents:

CHAPTER

ONE

SCHOOL_SCRAPER

1.1 school_scraper package

1.1.1 Submodules

school_scraper.fsddistrict module

Primitives for reading francophone school district information

class school_scraper.fsddistrict.**FSDDistrict**(*df*)
Bases: object

Abstraction around school district data parsed from the district website

Parameters **df** – Pandas data frame describing the district

DISTRICT_FIELD = 'Région'

get_school(*name*)
Gets a specific school from the district

Parameters **name** (*str*) – name of the school to get data for

Returns reference to the school with the given name, or None if no such school exists

Return type *FSDSchool*

property name
Name of the district

Type str

property school_names
list of school names associated with this district

Type list (str)

property schools
1 or more schools associated with this district

Type list (*FSDSchool*)

school_scraper.fsdschool module

Abstraction around a francophone school district school

class school_scraper.fsdschool.**FSDSchool**(*df*)
Bases: object

Abstraction around school-specific data parsed from the school district's website

Parameters **df** – Panda's dataframe containing school data parsed from the website

BUS_FIELD = 'Autobus'

MESSAGE_FIELD = 'Messages'

OPEN_FIELD = 'École'

SCHOOL_FIELD = "Nom de l'école"

property has_late_busses
True if 1 or more buses are late, False if all are running on time

Type bool

property is_open
True if school is open, False if not

Type bool

property messages
status messages associated with the school

Type str

property name
name of the school

Type str

school_scraper.fsdscraper module

Scraper for New Brunswick francophone school district website

class school_scraper.fsdscraper.**FSDScraper**(*html*)
Bases: object

Interface for parsing HTML data loaded from the New Brunswick francophone school district website

Parameters **html** (*str*) – HTML data loaded from the website. Is expected to contain a single HTML table containing rows describing each school in each school district

SCHEDULE_URL = 'https://bp.nbed.nb.ca/notices/BPRFtbl.aspx?dst=dsfs&vtbl=1'

property district_names
list of names of all districts parsed from the HTML

Type list (str)

property districts
0 or more districts parsed from the HTML content

Type list (*FSDDistrict*)

get_district(*name*)
Gets a specific district from the HTML content

Parameters **name** (*str*) – the name of the district to locate

Returns Reference to the district details for the named district, or None if no district with the given name exists

Return type *FSDDistrict*

property school_names
list of unique names of all schools in all districts

Type list (str)

static validate(*html*)
Checks to see if HTML loaded from the website is parseable

Parameters **html** (*str*) – HTML data loaded from the district website

Returns True if the HTML content was parseable, False if not. Details of any parsing errors are reported to the logger.

Return type bool

school_scraper.version module

Version number declaration

1.1.2 Module contents

Web scrapper for parsing school closure info from school websites

CHAPTER

TWO

INDICES AND TABLES

- genindex
- modindex
- search

PYTHON MODULE INDEX

S

school_scraper, 5
school_scraper.fsddistrict, 3
school_scraper.fsdschool, 4
school_scraper.fsdscraper, 4
school_scraper.version, 5

INDEX

B

BUS_FIELD (*school_scraper.fsdschool.FSDSchool attribute*), 4

D

DISTRICT_FIELD (*school_scraper.fsddistrict.FSDDistrict attribute*), 3
district_names() (*school_scraper.fsdscraper.FSDScraper property*), 4
districts() (*school_scraper.fsdscraper.FSDScraper property*), 4

F

FSDDistrict (*class in school_scraper.fsddistrict*), 3
FSDSchool (*class in school_scraper.fsdschool*), 4
FSDScraper (*class in school_scraper.fsdscraper*), 4

G

get_district() (*school_scraper.fsdscraper.FSDScraper method*), 4
get_school() (*school_scraper.fsddistrict.FSDDistrict method*), 3

H

has_late_busses() (*school_scraper.fsdschool.FSDSchool property*), 4

I

is_open() (*school_scraper.fsdschool.FSDSchool property*), 4

M

MESSAGE_FIELD (*school_scraper.fsdschool.FSDSchool attribute*), 4
messages() (*school_scraper.fsdschool.FSDSchool property*), 4
module
    school_scraper, 5
    school_scraper.fsddistrict, 3
    school_scraper.fsdschool, 4
    school_scraper.fsdscraper, 4
    school_scraper.version, 5

N

name() (*school_scraper.fsddistrict.FSDDistrict property*), 3
name() (*school_scraper.fsdschool.FSDSchool property*), 4

O

OPEN_FIELD (*school_scraper.fsdschool.FSDSchool attribute*), 4

S

SCHEDULE_URL (*school_scraper.fsdscraper.FSDScraper attribute*), 4
SCHOOL_FIELD (*school_scraper.fsdschool.FSDSchool attribute*), 4
school_names() (*school_scraper.fsddistrict.FSDDistrict property*), 3
school_names() (*school_scraper.fsdscraper.FSDScraper property*), 5
school_scraper
    module, 5
school_scraper.fsddistrict
    module, 3
school_scraper.fsdschool
    module, 4
school_scraper.fsdscraper
    module, 4
school_scraper.version
    module, 5
schools() (*school_scraper.fsddistrict.FSDDistrict property*), 3

V

validate() (*school_scraper.fsdscraper.FSDScraper static method*), 5