

ProfileManager

Table of contents

- Introduction
- Documentation
- Usage

Introduction

ProfileManager is a class for helping manage a database of business profiles. It uses the SEC's (United States Securities and Exchange Commission) CIK codes (Central Index Key) to act as identifiers and allows the user to compile a variety of information on corporate entities in an easy to use and query format. Assisting the accessibility of information, ProfileManager supports using a series of mappings from CIK codes to names and back, names to aliases, and mappings from industry code standards and descriptions of them. The hope to provide for a flexible data solution for complex business oriented applications.

Documentation

`__init__(rel_path=None)`

- **rel_path** is the relative path from the script to the parent folder of where your data will be housed. ProfileManager always assumes that data will be held in "data/profilemanager/data" and profiles will be held in "data/profilemanager/profiles" so this allows you to orient your ProfileManager instance to your data source.

Sets the **rel_path** variable and reads in various mappings including:

- CIK (Central Index Key) to name
- Name to CIK (Central Index Key)
- Name to (a list of) aliases
- NAICS (North American Industry Classification System) to description of classification
- SIC (Standard Industrial Classification) to description of Classification
- NAICS to SIC
- SIC to NAICS

`__contains__(key)`

- **key** can be a CIK (Central Index Key) code, name, or alias

Returns true if the **key** is found in the ProfileManager instance.

`__getitem__(key)` and `get(key)`

- **key** can be a CIK (Central Index Key) code, name, or alias

Returns the profile identified by the key if it is found, returns None else.

`__iter__()`

A generator that yields the profiles of the contained corporate entities.

`__len__()`

Returns the number of CIK (Central Index Key) codes in the instance.

`__repr__()` and `__str__()`

Returns a sorted and indented dictionary of CIK (Central Index Key) codes to names of the profiles contained.

cik_to_alias(cik)

- **cik** is the CIK (Central Index Key) code of a business

Returns a list of aliases of the business entity identified by the CIK (Central Index Key) code.

cik_to_description(cik)

- **cik** is the CIK (Central Index Key) code of a business

Returns a list of descriptions of the industries of the NAICS (North American Industry Classification System) and SIC (Standard Industrial Classification) codes of the business entity identified by the CIK (Central Index Key) code.

cik_to_naics(cik)

- **cik** is the CIK (Central Index Key) code of a business

Returns the NAICS (North American Industry Classification System) codes of the business entity identified by the CIK (Central Index Key) code.

cik_to_name(cik)

- **cik** is the CIK (Central Index Key) code of a business

Returns the name of the business entity identified by the CIK (Central Index Key) code.

cik_to_sic(cik)

- **cik** is the CIK (Central Index Key) code of a business

Returns the SIC (Standard Industrial Classification) codes of the business entity identified by the CIK (Central Index Key) code.

generate_profiles()

Currently the code uses three additional JSON files to generate profiles:

- a map from CIK (Central Index Key) to SIC (Standard Industrial Classification)
- a map from SIC (Standard Industrial Classification) to NAICS (North American Industrial Classification System)
- a dictionary that maps from CIK (Central Index Key) to a dictionary containing **10K**, **8K**, and **EX21** as keys and their values are a list of dictionaries of SEC filings with **url** and **time_of_filing** fields for each type of document respectively.

The code parses the documents and uses the various mappings to aggregate the information into one profile.

naics_to_description(naics)

- **naics** is a NAICS (North American Industry Classification System) code.

Returns a list of descriptions of the industrial code.

naics_to_sic(naics)

- **naics** is a NAICS (North American Industry Classification System) code.

Returns the SIC (Standard Industrial Classification) most closely associated with **naics**

name_to_aliases(name)

- **name** is the name of a business entity

Returns a list of aliases of the business entity.

name_to_cik(name)

- **name** is the name of a business entity

Returns the CIK (Central Index Key) of the named business entity.

name_to_description(name)

- **name** is the name of a business entity

Returns a list of descriptions of the industries of the named business entity.

update_profiles()

The code uses a JSON mapping CIK (Central Index Key) to a dictionary that maps from CIK (Central Index Key) to a dictionary containing **10K**, **8K**, and **EX21** as keys and their values are a list of dictionaries of SEC filings with **url** and **time_of_filing** fields for each type of document respectively. The code parses the filings which haven't already been parsed.

Usage

Coming Soon!