Dictionary for Bike Station Data

Data Structure Documentation

October 15, 2024

1 Introduction

This document provides a dictionary for the bike station data collected from the JCDecaux API. The data includes various attributes related to each bike station, which are detailed in the following table.

2 Function Descriptions

- **clean_name(name):** This function takes a string input, cleans it by removing digits, hyphens, the substring "CB", and any extraneous spaces or parentheses. It returns a cleaned string.
- **nettoyer_donnees(stations):** This function takes a list of station dictionaries as input. It validates and cleans each station's data by checking for required keys and correct data types. Invalid stations are ignored. The cleaned list of stations is returned.
- **recup_stations():** This function retrieves the bike station data from the JCDecaux API. It sends a GET request and processes the JSON response to create a list of station dictionaries. Each dictionary is structured according to predefined keys. The function then cleans the data using `nettoyer_donnees()` and saves the cleaned data to a CSV file. Finally, it returns a DataFrame containing the cleaned data.

3 Data Dictionary

Key	Description
Address **Data Type:** String	The name or address of the bike station. This field may contain extra characters, which should be cleaned for clarity. **Example:** "Rue des Rives, 54000 Nancy"
Latitude **Data Type:** Float	The geographical latitude of the bike station, represented as a floating-point number. **Example:** 48.692054
Longitude **Data Type:** Float	The geographical longitude of the bike station, represented as a floating-point number. **Example:** 6.184417
Station **Data Type:** String	The name of the bike station, which may also be used for identification purposes. This field should be cleaned to remove unnecessary characters. **Example:** "Station Nancy"
CB **Data Type:** Null or Any	A placeholder field for any additional data related to "CB" (not defined in the current context). **Example:** None
Station_ID **Data Type:** Integer	The unique identifier for the bike station, typically a number assigned by the bike sharing service. **Example:** 23
CurNumberOfBikes **Data Type:** Integer	The current number of bikes available at the station. This is represented as an integer. **Example:** 5
MaxNumberOfBikes **Data Type:** Integer	The maximum capacity of bikes that the station can accommodate, represented as an integer. **Example:** 15

Table 1: Data Dictionary for Bike Station Data