

# Penggunaan Decorator Python Terhadap Fungsi Validasi Tipe Data untuk Validasi Data dalam Analisis Statistik

Cyntia Kristina Sidauruk (122450023)<sup>1</sup>, Patricia Leondrea Diajeng Putri (122450050)<sup>2</sup>, Berliana Enda Putri (122450065)<sup>3</sup>, Feryadi Yulius (122450087)<sup>4</sup>, Naufal Fakhri (122450089)<sup>5</sup>

Program Studi Sains Data Institut Teknologi Sumatera  
Jl. Terusan Ryacudu, Way Huwi, Kec. Jatiagung, Kabupaten Lampung Selatan,  
Lampung 35365

Email: <sup>1</sup>[cyntia.122450023@student.itera.ac.id](mailto:cyntia.122450023@student.itera.ac.id), <sup>2</sup>[patricia.122450050@student.itera.ac.id](mailto:patricia.122450050@student.itera.ac.id),  
<sup>3</sup>[berliana.122450065@student.itera.ac.id](mailto:berliana.122450065@student.itera.ac.id), <sup>4</sup>[feryadi.122450087@student.itera.ac.id](mailto:feryadi.122450087@student.itera.ac.id), <sup>5</sup>[naufal.122450089@student.itera.ac.id](mailto:naufal.122450089@student.itera.ac.id)

## ABSTRAK

Penelitian ini bertujuan untuk menerapkan decorator dan closure dalam pengembangan program Python demi menciptakan fungsi utilitas bernama `describe()`, yang berfungsi untuk memvalidasi tipe data argumen yang diberikan dan menghasilkan statistik deskripsi. Prosedur ini membantu untuk memeriksa keakuratan informasi sebelum melakukan analisis lebih lanjut. Pada percobaan ini, pemanfaatan dekorator dan closure telah terbukti efektif dalam menciptakan fungsi `describe()` yang dapat digunakan untuk memvalidasi tiga jenis informasi universal, seperti list, array, dan dataframe. Percobaan ini menunjukkan bahwa pelaksanaan tersebut berhasil dalam memvalidasi ketiga jenis informasi tersebut. Meskipun begitu, masih ada kesempatan untuk memperluas fungsionalitas ini, contohnya dengan meningkatkan kemampuannya dalam menangani berbagai jenis informasi. Dengan memperluas cakupan validasi, kemampuan `describe()` menjadi lebih fleksibel dan dapat digunakan dalam berbagai situasi analisis data. Penelitian ini ikut memajukan teknik validasi data dalam analisis statistik menggunakan Python.

**Kata kunci :** *Array, Dataframe, Closure, Decorate, Deskripsi Statistik, List, Python, Validasi Data*

## PENDAHULUAN

Perkembangan metode validasi data dan analisis statistik dengan Python telah memberikan kontribusi yang signifikan bagi dunia ilmu data. Dalam penggunaan dekorator Python dan konsep penutupan menjadi bagian integral dalam meningkatkan proses validasi data. Bertujuan untuk menerapkan dekorator dan penutupan dalam bahasa pemrograman Python dengan fokus utama pada pembuatan fungsi deskripsi. Fungsi ini tidak hanya memvalidasi jenis informasi yang digunakan sebagai argumen, namun juga menghasilkan nilai deskriptif statistik yang relevan. Validasi data yang efektif membantu memastikan informasi akurat sebelum melakukan analisis lebih lanjut.

## TUJUAN

1. Menerapkan Decorator Python Fungsi Validasi Tipe data pada fungsi Deskripsi Statistik
2. Melakukan validasi tipe data dan mencari nilai hasil deskripsi statistik

## **METODE DASAR**

### **Closure**

Closure adalah sebuah konsep di dalam pemrograman yang digunakan untuk mengatur kondisi ketika suatu fungsi atau blok kode selesai dijalankan. Dalam python, sebuah closure adalah fungsi yang memiliki akses ke lingkup (scope) dari fungsi luar tempatnya didefinisikan, bahkan setelah fungsi luar tersebut selesai dieksekusi. Ini dapat diartikan sebagai fungsi yang dapat diakses dari dalam fungsi lain, atau fungsi yang dapat menyimpan dan mengelola data lokal, atau dengan pengertian sebuah closure dapat "mengingat" dan mengakses variabel-variabel dari lingkup luarnya bahkan setelah fungsi luarnya telah selesai dieksekusi.

### **Decorator**

Python decorator adalah sebuah fungsi yang menambah atau mengubah perilaku fungsi atau metode lain tanpa mengubah kode sumber aslinya. Decorator bertindak sebagai kelas tambahan yang menggabungkan fungsi yang ada, memungkinkan juga untuk menambahkan fungsionalitas tambahan sebelum atau sesudah fungsi dijalankan. Decorator dapat digunakan untuk memodifikasi fungsi atau metode dengan cara yang rapi dan elegan. Juga berguna dalam berbagai kasus penggunaan, seperti memvalidasi fungsi atau parameter masukan, operasi pencatatan, mengukur waktu eksekusi fungsi, mengelola otorisasi atau autentikasi, serta menangani pemrosesan dan manipulasi hasil fungsi. Decorator dapat didefinisikan dengan kata kunci `@` diikuti dengan nama fungsi dekorator.

### **Validasi data**

Validasi data merupakan bagian yang tak terpisahkan dari proses pengendalian kualitas yang bertujuan untuk memastikan konsistensi data sebelum disajikan kepada pengguna. Proses ini umumnya melibatkan perbandingan nilai pengujian dengan data input, verifikasi data, dan kontrol mutu. Validasi data mencakup pemeriksaan sederhana dan kompleks serta dapat diotomatisasi pada beberapa tahap pemrosesan dan penyimpanan data. Teknik validasi dirancang untuk mendeteksi kesalahan umum yang mungkin terjadi. Biasanya, hasil validasi dimaksudkan untuk memberikan penjelasan mengenai nilai data yang dilaporkan. Saat menentukan prosedur validasi yang tepat untuk suatu variabel tertentu, penting untuk selalu mempertimbangkan ketepatan variabel yang diamati serta kemampuannya untuk memperbaiki kesalahan yang terdeteksi.

### **Statistika deskriptif**

Statistik deskriptif, atau yang juga disebut sebagai statistik deskripsi, merujuk pada bidang statistik yang berkaitan dengan pengumpulan, penyusunan, pengolahan, penyajian, dan analisis data. Statistik deskriptif bertujuan untuk menggambarkan atau menjelaskan ciri-ciri suatu kelompok atau rangkaian data (baik data sampel maupun data populasi) tanpa melakukan generalisasi, yaitu penarikan kesimpulan yang biasanya didasarkan pada informasi dari sampel dan diterapkan pada populasi yang sama.

## PEMBAHASAN

### Pseudocode

Pertama dirancang pseudocode dari fungsi yang akan dibuat. Terdapat dua buah fungsi utama, yaitu fungsi dekorator untuk melakukan validasi tipe data, yang berisi fungsi dalam bernama wrapper yang bekerja sebagai pengecek tipe data yang diinputkan dengan python closure. Kemudian didefinisikan fungsi kedua, yaitu fungsi untuk melihat deskripsi statistik dari data seperti nilai rata-rata, nilai tengah, standar deviasi, kuartil, dan lain-lain. Didapatlah pseudocode sebagai berikut:

Unset

1. Import pandas as pd
2. Define a decorator function named `validate_data_type` which takes a function `func` as input:
3. Define an inner function named `wrapper` that takes `data` as input.
4. Check the type of data:
  5. If data is an instance of list:
    6. Call `func` with `data` as argument and return the result.
  7. If data is an instance of `pd.DataFrame`:
    8. Convert data to list using `data.values.tolist()` and call `func` with the list as argument.
- Return the result.
  9. If none of the above conditions are met, print an error message indicating invalid data type.
10. Return the wrapper function.
11. Define a function `describe(data)`:
  12. If data is an instance of list:
    13. Convert data to `DataFrame`.
  14. Return `data.describe()`.
15. Apply the `validate_data_type` decorator to the `describe` function.

### Kode

Setelah dirancang pseudocode untuk kedua fungsi, selanjutnya dibuat kedua fungsi tersebut pada bahasa pemrograman python dengan menerapkan python closure dan python decorator.

Python

```
import pandas as pd

def validate_data_type(func):
    def wrapper(data):
        if isinstance(data, list):
```

```

        return func(data)
    elif isinstance(data, pd.DataFrame):
        return func(data.values.tolist())
    else:
        print("Error! Invalid data type, expected list or DataFrame")
    return wrapper

@validate_data_type
def describe(data):
    if isinstance(data, list):
        data = pd.DataFrame(data)
    return data.describe()

```

## Penerapan

Untuk melihat apakah fungsi yang telah dibuat bekerja dengan baik, bisa diterapkan pada beberapa tipe data, yaitu list, array, dan dataframe. Kemudian dilakukan percobaan pada tipe data yang invalid dan menghasilkan keluaran berupa error.

## Penerapan pada List

```

Python
# Contoh penggunaan pada tipe data list
data_list = [1, 2, 3, 4, 5]
print("Describe of list:")
print(describe(data_list))

```

```

Describe of list:
      0
count  5.000000
mean   3.000000
std    1.581139
min    1.000000
25%    2.000000
50%    3.000000
75%    4.000000
max    5.000000

```

**Gambar 1 :** Keluaran Penerapan fungsi pada tipe data list

Pada Gambar 1 di atas didapat nilai deskripsi statistik dari data bertipe list, yaitu `data_list = [1, 2, 3, 4, 5]` berupa nilai count, mean, standar deviasi, Q1 (25%), Q2 (50%), Q3 (75%), minimal dan maksimal.

## Penerapan pada Array

Python

# Contoh penerapan pada tipe data array

```
data_array = [1, 6, 9, 24, 0]
```

```
print("\nDescribe of array:")
```

```
print(describe(data_array))
```

Describe of array:

	0
count	5.00000
mean	8.00000
std	9.66954
min	0.00000
25%	1.00000
50%	6.00000
75%	9.00000
max	24.00000

**Gambar 2 :** Keluaran Penerapan fungsi pada tipe data array

Pada Gambar 2 di atas didapat nilai deskripsi statistik dari data bertipe array, yaitu `data_array = [1, 6, 9, 24, 0]` berupa nilai count, mean, standar deviasi, Q1 (25%), Q2 (50%), Q3 (75%), minimal dan maksimal.

## Penerapan pada Dataframe

Python

# Contoh penerapan pada tipe dataframe

```
data_df = pd.DataFrame({'A': [1, 2, 3, 4, 5],
```

```
                        'B': [1, 4, 6, 2, 6],
```

```
                        'C': ['a', 'b', 'c', 'd', 'e']})
```

```
print("\nDescribe of DataFrame:")
```

```
print(describe(data_df))
```

Describe of DataFrame:

	0	1
count	5.000000	5.000000
mean	3.000000	3.800000
std	1.581139	2.280351
min	1.000000	1.000000
25%	2.000000	2.000000
50%	3.000000	4.000000
75%	4.000000	6.000000
max	5.000000	6.000000

**Gambar 3 :** Keluaran Penerapan fungsi pada tipe dataframe

Pada Gambar 3 di atas didapat nilai deskripsi statistik dari data frame, yaitu data\_df = [1, 2, 3, 4, 5] berupa nilai count, mean, standar deviasi, Q1 (25%), Q2 (50%), Q3 (75%), minimal dan maksimal.

### Penerapan pada tipe data invalid

Python

# Contoh penerapan pada data invalid

```
print("\nInvalid Data type Example:")
```

```
print(describe((1, 2, 3, 4, 5)))
```

```
Invalid Data type Example:  
Error! Invalid data type, expected list, array or DataFrame  
None
```

**Gambar 4 :** Keluaran Penerapan fungsi pada tipe yang invalid

Pada Gambar 4 di atas didapat keluaran sebuah pernyataan bahwa terjadi Error, dengan tipe error ialah tipe data yang invalid, yang diharapkan sebuah list, array, atau dataframe.

## KESIMPULAN

### 1. Pentingnya validasi data dan analisa statistik dengan Python

Implementasi validasi data dan metode analisis statistik dengan Python memberi pengaruh yang signifikan terhadap dunia data science - Python memungkinkan pengguna untuk memvalidasi suatu data secara efisien untuk memastikan keakuratan informasi. Pemanfaatan teknik validasi data dan analisis statistik dengan Python memiliki pengaruh yang besar dalam bidang ilmu data. Python membantu pengguna memverifikasi data dengan cepat untuk menjamin keakuratan.

### 2. Dekorator dan Closure Python

Dekorator dan penutupan telah menjadi kunci dalam meningkatkan proses validasi data dengan Python. Dekorator menyederhanakan penambahan atau modifikasi fungsi tanpa mengganti kode asli, sehingga meningkatkan fleksibilitas. Demikian pula, penutupan memfasilitasi akses ke cakupan fungsi di luar tempat fungsi tersebut didefinisikan, yang berfungsi sebagai alat berharga untuk validasi data.

### 3. Pembuatan Deskripsi Fungsi

Saat membuat fungsi untuk deskripsi statistik, dekorator dan enclosure digunakan untuk mengkonfirmasi tipe data masukan dan menghasilkan nilai statistik deskriptif yang sesuai.

## REFERENSI

Auffarth, B. (2021). *Machine Learning for Time-Series with Python: Forecast, predict, and detect anomalies with state-of-the-art machine learning methods*. Packt Publishing Ltd.

Husnul, N. R. I., Prasetya, E. R., Ajimat, A., & Purnomo, L. I. (2020). Statistik deskriptif. *Universitas Pamulang: Banten*.

Syaferi, F. (2023). *Memahami Decorators di Python*. Asia Career Blog. Retrieved April 24, 2024, from <https://blog.unmaha.ac.id/memahami-decorators-di-python>

Windatiningsih, D. n.d. “Uji Validasi Data Debit untuk Deteksi Penyimpangan Data Studi Kasus: Das Citarum Hulu.” *Jurnal Sumber Daya Air* 15 (2): 121-136.