

Tugas Besar IF2220 Probabilitas dan Statistika: Penarikan Kesimpulan dan Pengujian Hipotesis

Nomor 1

Menulis deskripsi statistika (Descriptive Statistics) dari semua kolom pada data yang bersifat numerik, terdiri dari mean, median, modus, standar deviasi, variansi, range, nilai minimum, maksimum, kuartil, IQR, skewness dan kurtosis. Boleh juga ditambahkan deskripsi lain

```
# Mengimport Library
```

```
import pandas as pd
```

```
# Membaca data dan melihat beberapa data pertama
```

```
df = pd.read_csv('water_potability.csv', index_col=0)
```

```
df.head()
```

| | pH | Hardness | Solids | Chloramines | Sulfate \ |
|----|-----------|------------|--------------|-------------|------------|
| id | | | | | |
| 1 | 8.316766 | 214.373394 | 22018.417441 | 8.059332 | 356.886136 |
| 2 | 9.092223 | 181.101509 | 17978.986339 | 6.546600 | 310.135738 |
| 3 | 5.584087 | 188.313324 | 28748.687739 | 7.544869 | 326.678363 |
| 4 | 10.223862 | 248.071735 | 28749.716544 | 7.513408 | 393.663396 |
| 5 | 8.635849 | 203.361523 | 13672.091764 | 4.563009 | 303.309771 |

| | Conductivity | OrganicCarbon | Trihalomethanes | Turbidity |
|------------|--------------|---------------|-----------------|-----------|
| Potability | | | | |
| id | | | | |

| | | | | |
|---|------------|-----------|------------|----------|
| 1 | 363.266516 | 18.436524 | 100.341674 | 4.628771 |
| 0 | | | | |
| 2 | 398.410813 | 11.558279 | 31.997993 | 4.075075 |
| 0 | | | | |
| 3 | 280.467916 | 8.399735 | 54.917862 | 2.559708 |
| 0 | | | | |
| 4 | 283.651634 | 13.789695 | 84.603556 | 2.672989 |
| 0 | | | | |
| 5 | 474.607645 | 12.363817 | 62.798309 | 4.401425 |
| 0 | | | | |

```
# Menyatakan kolom-kolom numerik
```

```
col_numeric = [  
    'pH',  
    'Hardness',  
    'Solids',  
    'Chloramines',  
    'Sulfate',  
    'Conductivity',
```


| | | | | |
|----------|-------------|-------------|-------------|-------------|
| count | 2010.000000 | 2010.000000 | 2010.000000 | 2010.000000 |
| mean | 426.476708 | 14.357940 | 66.400717 | 3.969497 |
| std | 80.701872 | 3.325770 | 16.081109 | 0.780471 |
| min | 201.619737 | 2.200000 | 8.577013 | 1.450000 |
| 25% | 366.619219 | 12.122530 | 55.949993 | 3.442882 |
| 50% | 423.438372 | 14.323286 | 66.482041 | 3.967374 |
| 75% | 482.209772 | 16.683562 | 77.294613 | 4.514663 |
| max | 753.342620 | 27.006707 | 124.000000 | 6.494749 |
| variance | 6512.792113 | 11.060746 | 258.602066 | 0.609135 |
| IQR | 115.590553 | 4.561031 | 21.344620 | 1.071781 |
| range | 551.722883 | 24.806707 | 115.422987 | 5.044749 |
| skew | 0.268012 | -0.020220 | -0.051383 | -0.032266 |
| kurtosis | -0.237206 | 0.031018 | 0.223017 | -0.049831 |