

# POSTER



## CONSUMPTION ENERGY



Konsumsi energi adalah jumlah energi yang digunakan oleh individu, rumah tangga, industri, atau negara untuk memenuhi suatu kebutuhan dalam melakukan berbagai aktivitas. Kemampuan memprediksi konsumsi tentunya membantu sekali dalam rangka penghematan energi untuk tetap menjaga bumi. Dengan menggunakan algoritma yang paling umum yaitu regresi.

### SOLUTION STATEMENTS

Untuk eksplorasi fitur dilakukan

- Analisis Univariat (data kategorik dan numerik)
- Analisis Multivariat.

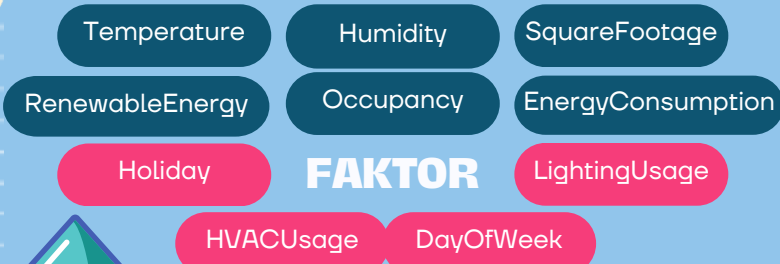
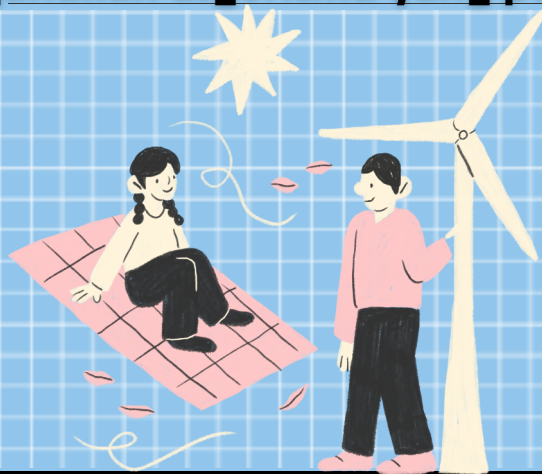
Teknik yang digunakan

- Catplot, pairplot, dan heatmap untuk melihat Correlation Matrix dari fitur-fitur yang dimiliki.

Agar didapatkan model prediksi yang baik maka dilakukan

- Proses Data Wrangling yang meliputi

Data Gathering, Data Assessing, dan Data Cleaning. Untuk mengetahui performa model dilakukan pengecekan performa dengan metrik evaluasi.



### Energy-consumption-prediction

Predict the Energy Consumption

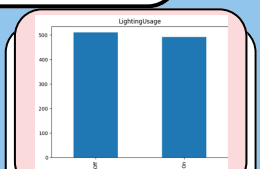
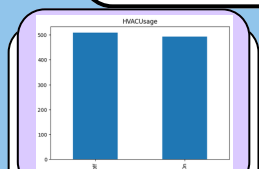
kaggle.com

- Dataset berupa CSV
- Dataset terdiri dari 1000 records dengan 10 buah fitur yang diukur.
- Dataset terdiri dari 4 data kategori dan 6 data numerik.
- Dataset memiliki missing value sejumlah 0 records

### Data Analysis

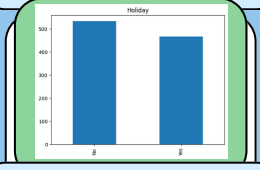
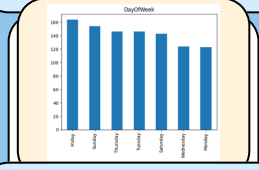
#### UNIVARIATE ANALYSIS

##### Data Kategori



Jumlah sampel	Persentase
Off	508 50.8
On	492 49.2

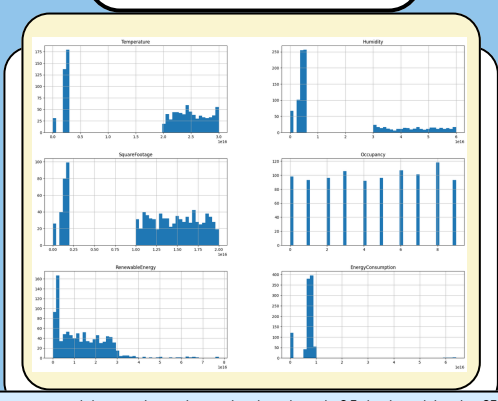
Jumlah sampel	Persentase
Off	509 50.9
On	491 49.1



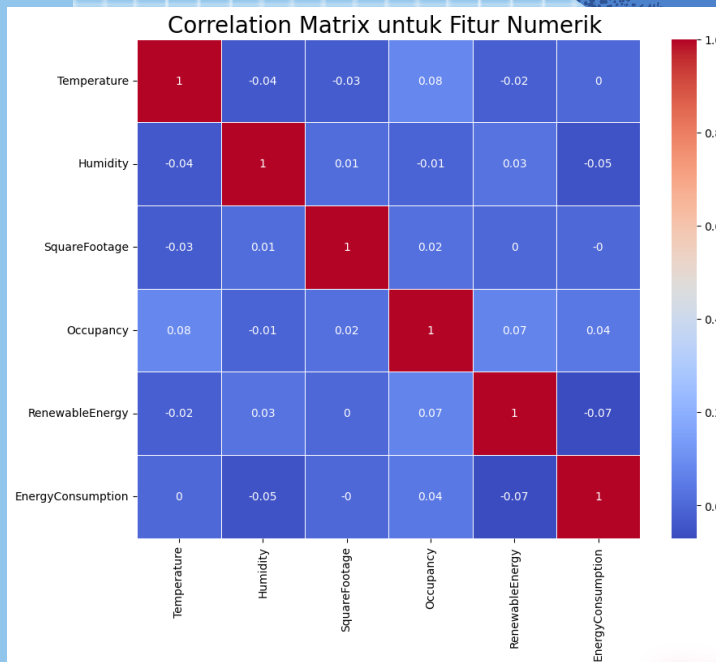
Jumlah sampel	Persentase
Monday	123 12.3
Tuesday	143 14.3
Wednesday	146 14.6
Thursday	146 14.6
Friday	164 16.4
Saturday	143 14.3
Sunday	164 16.4

Jumlah sampel	Persentase
No	533 53.3
Yes	467 46.7

##### Data Numerik

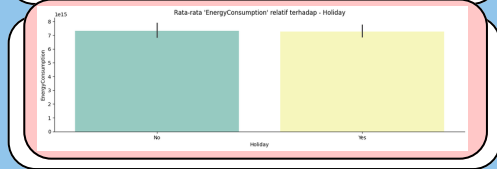
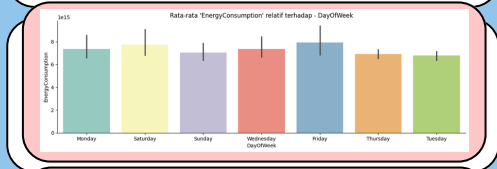
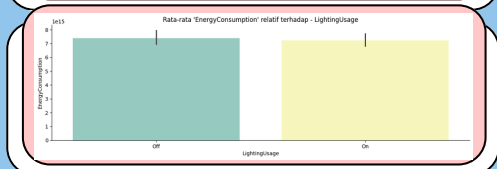
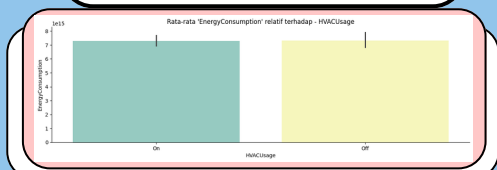


- Temperatur untuk konsumsi energi mayoritas berada pada 2.5 derajat celsius dan 25 derajat celsius.
- Humidity untuk konsumsi energi mayoritas berada pada 0.4% dan 0.3%.
- median dari Renewable Energy banyak terdistribusi pada rentang persentase 0%-3%, namun nilai terbanyak terdapat pada nilai 0.2%.
- rata-rata terbanyak untuk data total SquareFootage yaitu di antara angka 1%-2%.
- rata-rata terbanyak untuk data Occupansi berada di antara angka 6-8 orang.
- distribusi EnergyConsumption miring ke kanan (right-skewed). Hal ini akan berimplikasi pada model.

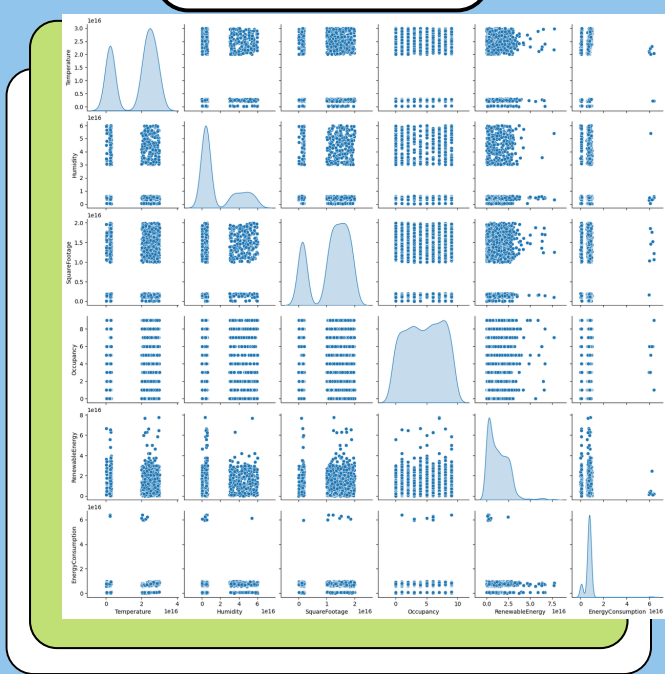


#### UNIVARIATE MULTIVARIATE

##### Data Kategori



##### Data Numerik

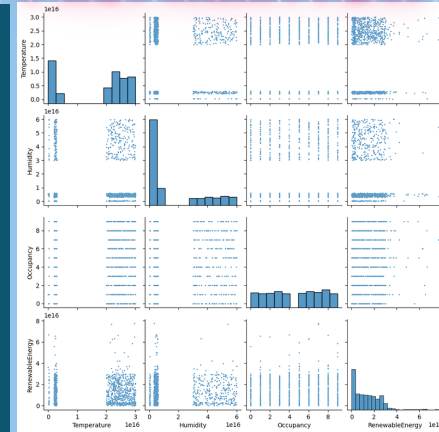


### DATA PREPARATION

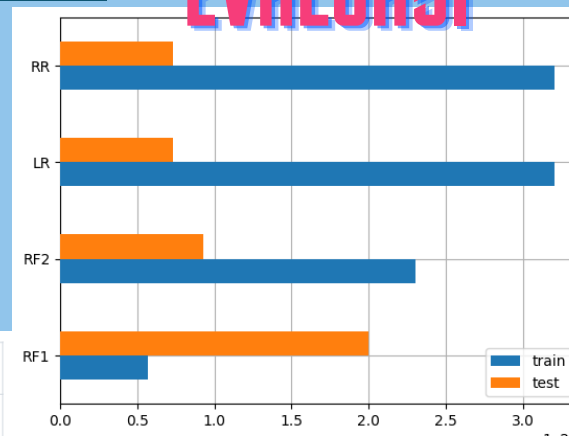
#### MODELING

```
GridSearchCV
estimator: RandomForestRegressor
RandomForestRegressor(random_state=100)
RandomForestRegressor
RandomForestRegressor(random_state=100)
```

```
{'max_depth': 3,
'min_samples_leaf': 4,
'min_samples_split': 2,
'n_estimators': 80}
```



### EVALUASI



Ketahui Lebih lanjut dan Lengkap pada Link!

[MACHINE-LEARNING](#)

BY. FADILA NUR AZIZAH

	y_true	prediksi_LR	prediksi_RR	prediksi_RF1	prediksi_RF2
131	7.142759e+14	7.633840e+15	7.633540e+15	4.995583e+15	5.906515e+15

