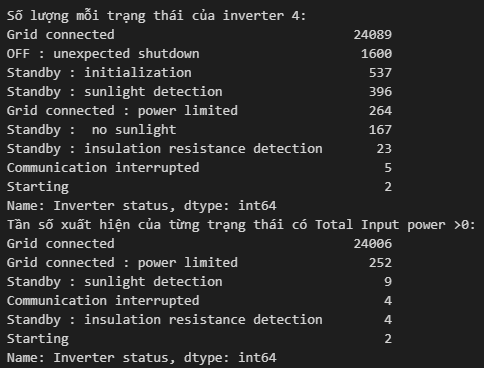
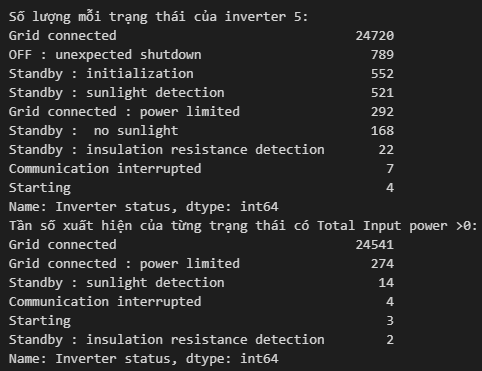
BÁO CÁO PHƯƠNG PHÁP CHẮT LỌC DATA

# Khảo sát trạng thái inverter (inverter status)

Trong nghiên cứu này thì chúng ta sẽ xét đến Total Input Power trong các trạng thái của inverter status (em đang nghiên cứu inverter 4-7):



Hình 1: Liệt kê các giá trị của inverter 4



Hình 2: Liệt kê các giá trị của inverter 5

Text

Description automatically generated

Hình 3: Liệt kê các giá trị của inverter 6

Text

Description automatically generated

Hình 4: Liệt kê các giá trị của inverter 7

Nhận xét:

* Qua các bảng liệt kê lên ta có thể thấy giá trị của trạng thái “Grid connected” là nhiều nhất sau đó đến trạng thái ‘OFF: unexpected shutdown’.
* Sau khi so sánh ta có có 9 trạng thái của inverter nhưng chỉ có 6 trạng thái của inverter là có “Total Input power” , chúng ta phải chọn ra những giá trị có tần suất xuất hiện lớn và có dung lượng lớn như trạng thái ‘Grid connected’ và ‘Grid connected : power limited’ với hơn 90% giá trị >0

# Xử lý số liệu để vẽ ra đường hiệu suất và đường tuyến tính hiệu suất

df=df[(df['Inverter status']=='Grid connected')|(df['Inverter status']=='Grid connected : power limited')]

Sau khi khai báo chúng ta chọn giá trị trạng thái của inverter như trên

df1=df.groupby(['month','date','hour'])['Total input power(kW)','Ambient temperature(℃)','Irradiance(W/㎡)'].mean()

df1=df1.reset\_index()

df1=df1[df1.groupby(['month','date'])['Irradiance(W/㎡)'].transform(max)==df1['Irradiance(W/㎡)']]

df1=df1.reset\_index()

df1['Tcell']=df1['Ambient temperature(℃)']+df1['Irradiance(W/㎡)']/800\*(25-20)

df1['Pm']=2.172\*1.303\*220\*2\*590\*df1['Irradiance(W/㎡)']/1000\*(1+0.0045\*(df1['Tcell']-25))\*20.5/100

df1['Hieu suat']=df1['Total input power(kW)']\*1000/df1['Pm']\*100

print(df1)

fig = go.Figure()

fig.add\_trace(go.Scatter(

    x=df1['date'],

    y=df1['Total input power(kW)'],

    name="Sản lượng điện trung bình cao nhất theo khung giờ mỗi ngày inverter 4"))

fig.add\_trace(go.Scatter(

    x=df1['date'],

    y=df1['Hieu suat'],

    name="Hiệu suất trung bình cao nhất theo khung giờ mỗi ngày inverter 4"))

fig . update\_traces (

    showlegend = True

    )

fig.update\_xaxes(

    rangeslider\_visible=True,

)

fig.show()

* Chúng ta lọc theo các giá trị 'Total input power(kW)','Ambient temperature(℃)','Irradiance(W/㎡)’ theo trung bình
* Tiếp theo chúng ta lấy Giá trị max của 'Irradiance(W/㎡)' rồi tính toán hiệu suất theo công thức trên
* Vẽ thử đường giá trị của Total Input Power và Hieu suat



Graphical user interface

Description automatically generated with medium confidence

Đường màu đỏ là đường hiệu suất, nhưng vẫn còn gập gềnh nên chúng ta sẽ thử vẽ đường hiệu suất tuyến tính cho nó.

# Vẽ đường tuyến tính

Graphical user interface, application, Word

Description automatically generated

A picture containing chart

Description automatically generated

y= ax+b

b= 69.91785474460849

a= -0.04751006627209331

Làm mịn data với filter: (lọc ra dữ liệu, tính trung bình Irra và tính hiệu suất)

df=df[df['Ambient temperature(℃)']>6]

df=df[df['Irradiance(W/㎡)']>300]

df=df[df['Total input power(kW)']>83]

df=df[df['Total input power(kW)']<110]

của dữ liệu gốc

y= ax+b

b = 67.14568858006089

a = -0.04135657863106545