COMPARATIVE ANALYSIS

Zenatix Business Analyst Intern Assignment/task

- Tools used: Jupyter Notebook & MS Excel
- Submitted by: **Renuka Meruva**
- Gmail: m.s.s.renuka@gmail.com
- Link to notebook: https://github.com/renukarm/Comparative-Analysis-of-different-Refrigiration-Systems

Problem Statement:

A Quick Service Restaurant has 4 walk-in refrigeration units. There is an energy meter installed at power supply of each of the 4 refrigeration units. There is also 1 temperature sensor installed inside each of these. Attached is data for power (kW) and temperature (degrees Celsius) for 4 energy meters and 4 temperature sensors for 10 days.

Do a comparative analysis of following parameters and prepare a presentation. You are allowed to use any tool for the analysis.

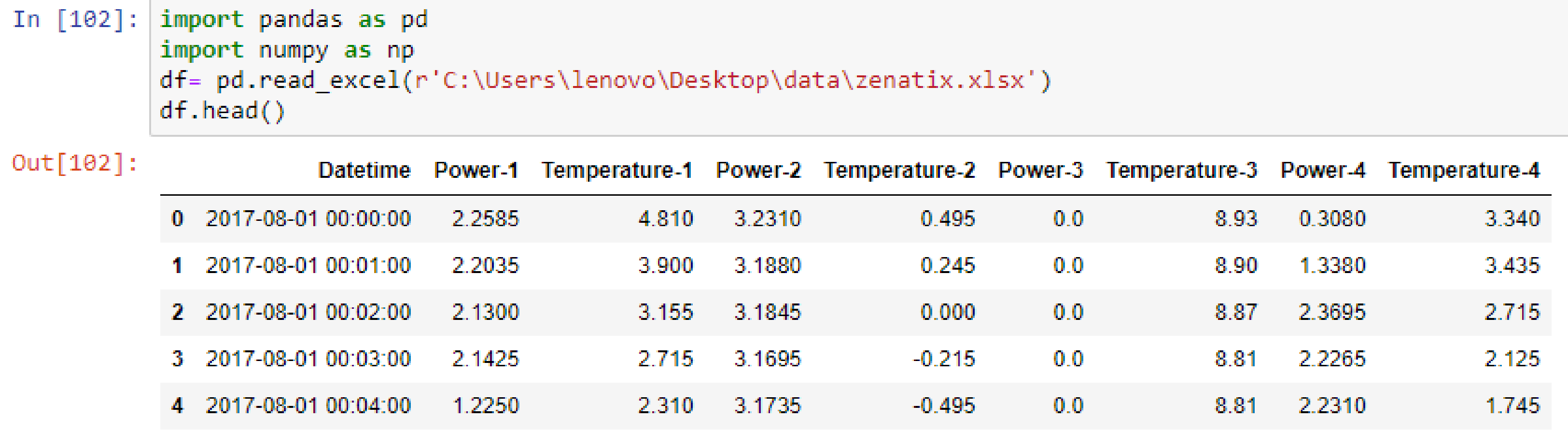
1. Duty cycle
2. Set temperature range (lower threshold and upper threshold)
3. Average daily energy consumption
4. Percentage of temperature compliance

Notes:

1. Refrigeration units usually have a range of set temperature instead of a single set point where lower and upper thresholds are defined.
2. Ideal temperature range is defined as 1-4 degree Celsius. Percentage of temperature compliance is percentage times temperature is within ideal range.

My Approach & Solution:

- Data used: An excel file, that was sent through mail.
- MS Excel: I used MS Excel for editing out a little bit on excel to get clear segregation of columns. i.e.. power-n, Temperature-n are characteristics of nth refrigerator respectively.
- Importing excel into jupyter notebook.



```
In [102]: import pandas as pd
          import numpy as np
          df= pd.read_excel(r'C:\Users\lenovo\Desktop\data\zenatix.xlsx')
          df.head()
```

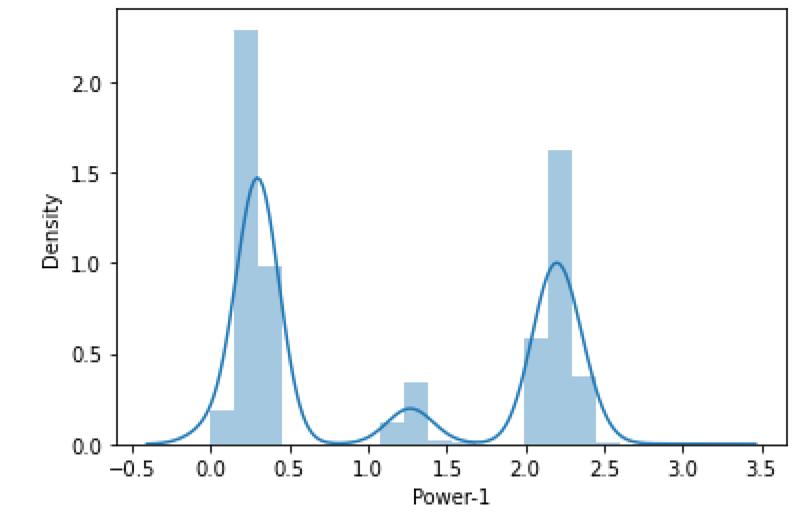
Out[102]:

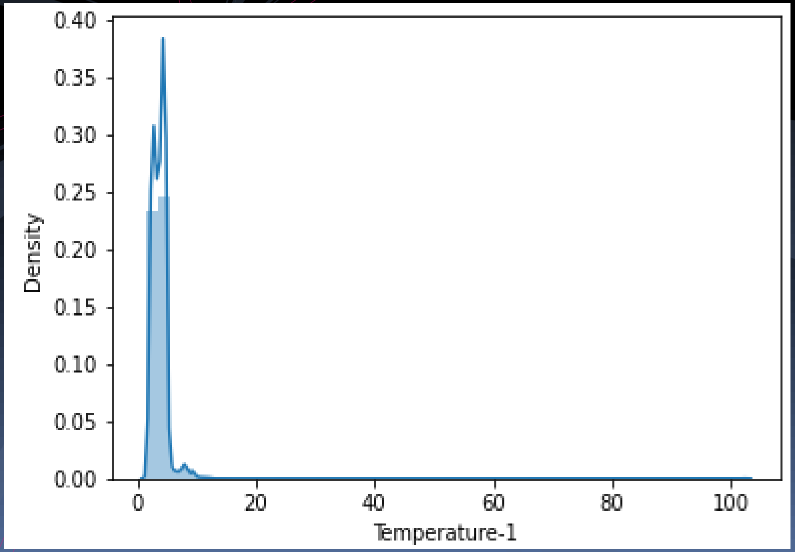
	Datetime	Power-1	Temperature-1	Power-2	Temperature-2	Power-3	Temperature-3	Power-4	Temperature-4
0	2017-08-01 00:00:00	2.2585	4.810	3.2310	0.495	0.0	8.93	0.3080	3.340
1	2017-08-01 00:01:00	2.2035	3.900	3.1880	0.245	0.0	8.90	1.3380	3.435
2	2017-08-01 00:02:00	2.1300	3.155	3.1845	0.000	0.0	8.87	2.3695	2.715
3	2017-08-01 00:03:00	2.1425	2.715	3.1695	-0.215	0.0	8.81	2.2265	2.125
4	2017-08-01 00:04:00	1.2250	2.310	3.1735	-0.495	0.0	8.81	2.2310	1.745

- Filled NA values with 0. (It was probably off then and power consumed will be 0)

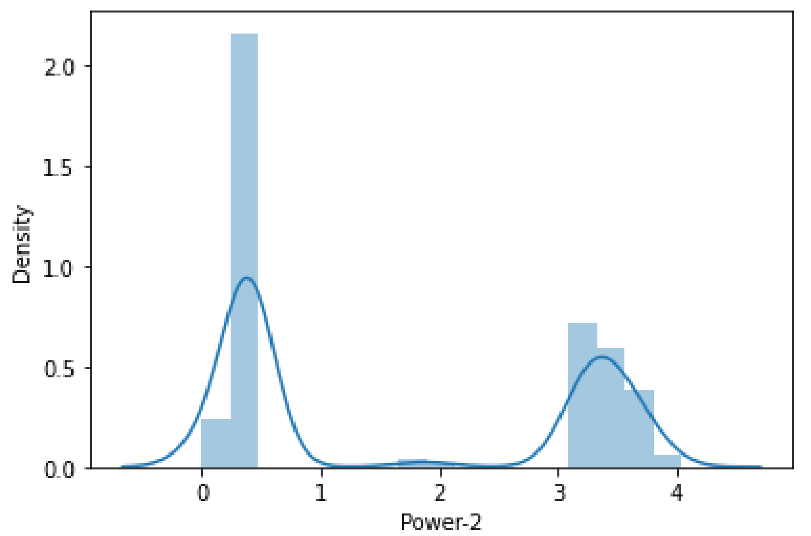
Visualizations (Distribution Plots, best for large data)

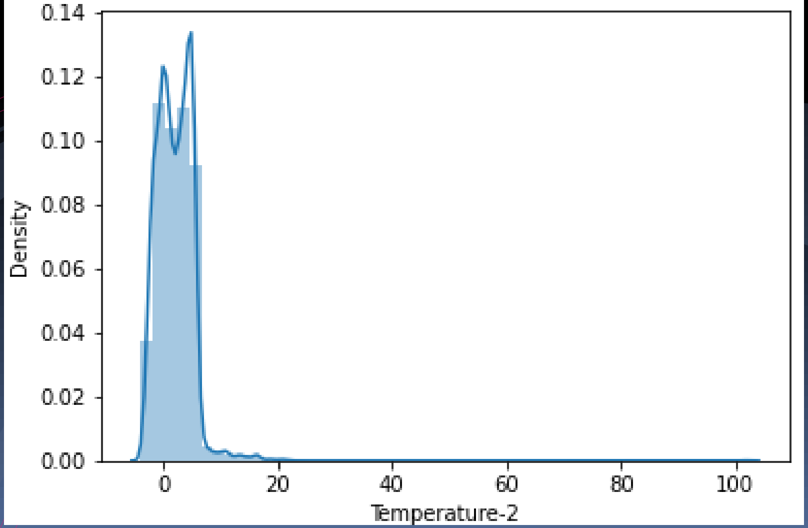
- **Refrigeration system -1**



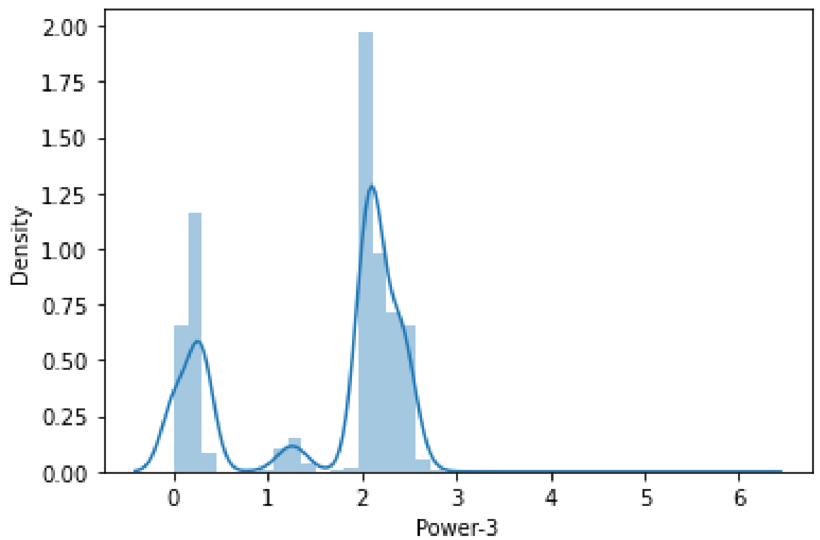


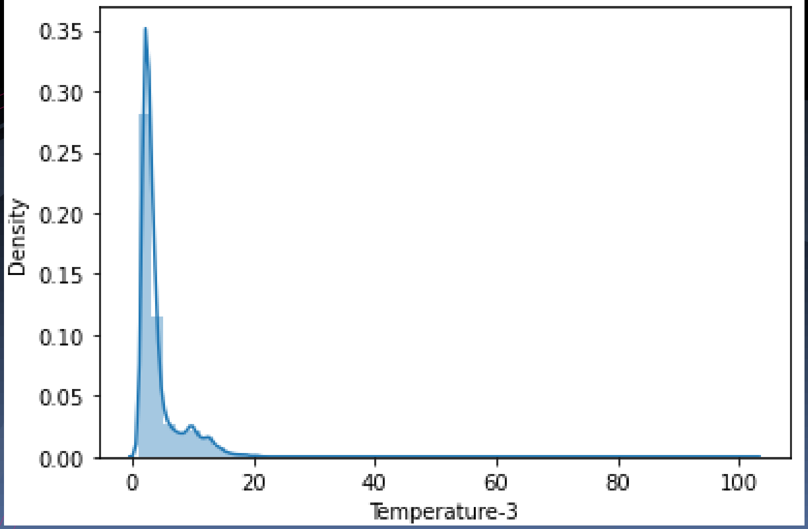
- Refrigeration system -2



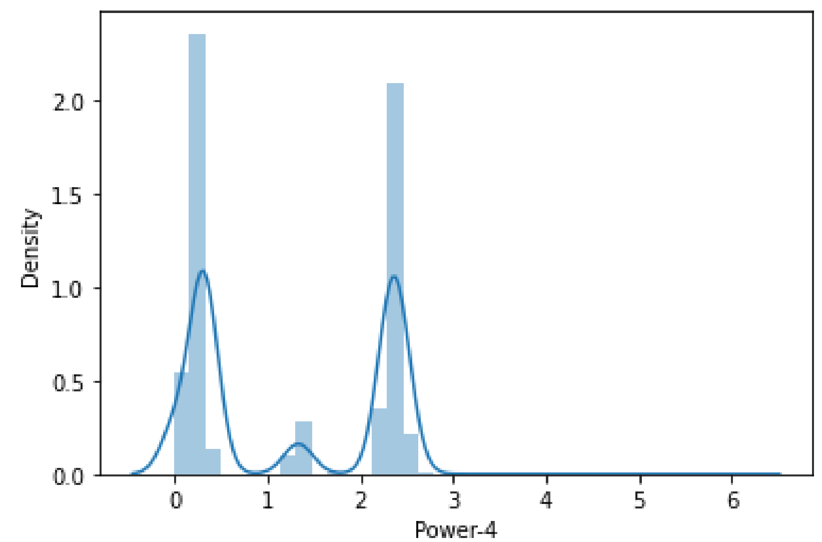


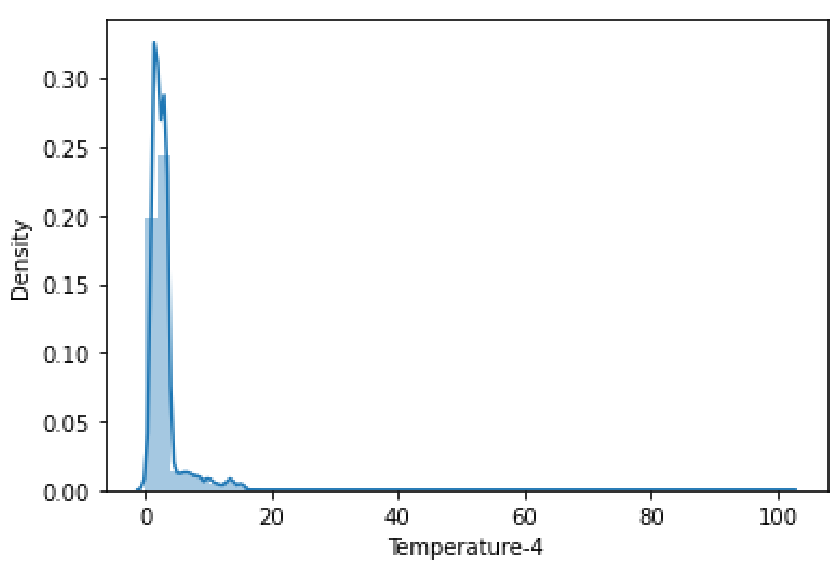
- # Refrigeration system -3





- Refrigeration system -4

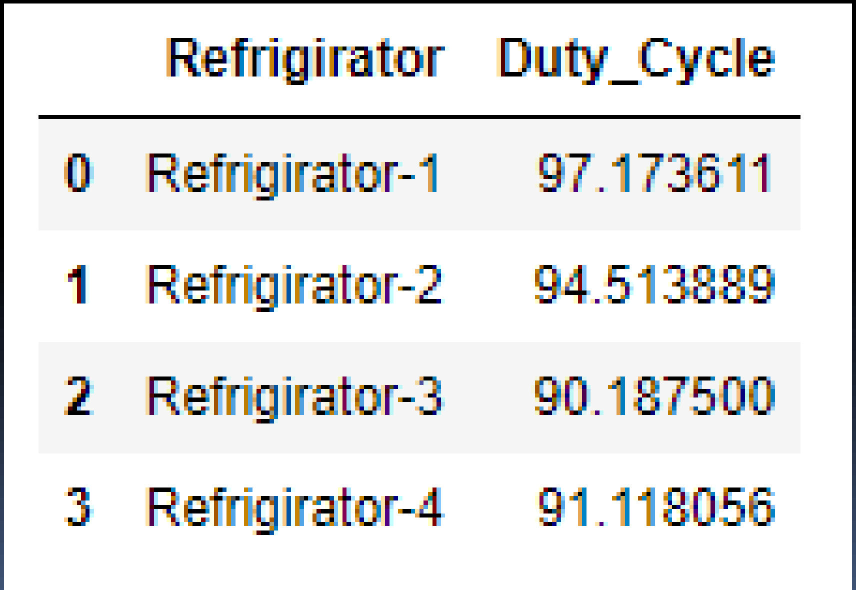




1. Duty Cycle

Formula= 100* (active time)/total time

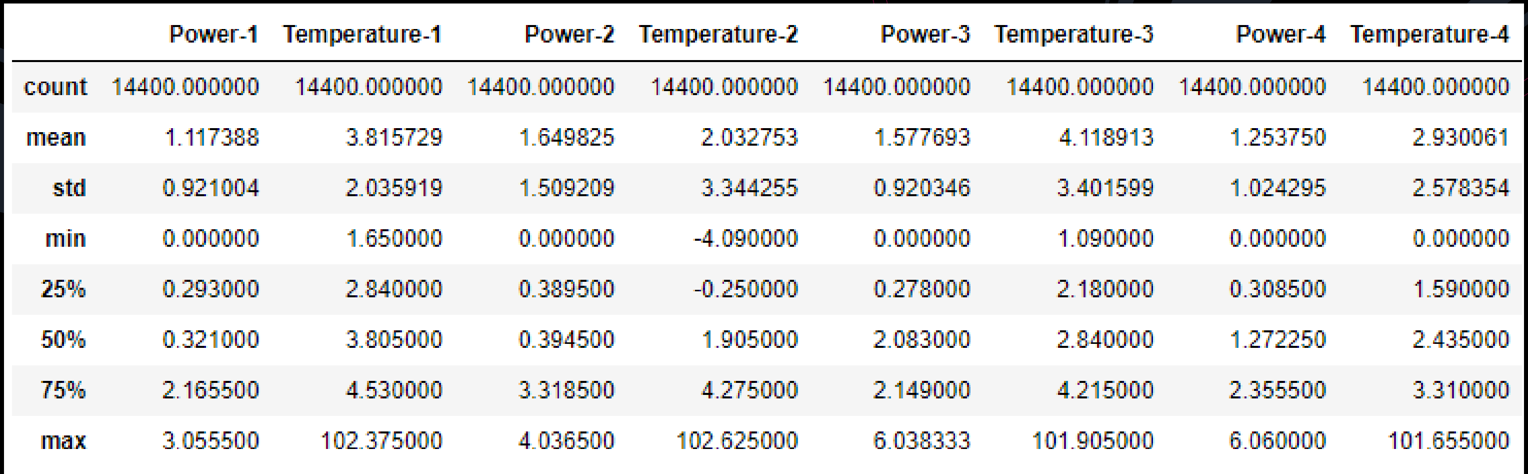
- Since the data is given for every single minute for 10 days, we'll be considering active time as, when power consumption is not zero. And calculate active mins for each system respectively.
- Total time= 10 days = 14,400 mins
- Solved and analyzed in Jupyter notebook, the comparison table:



	Refrigirator	Duty_Cycle
0	Refrigirator-1	97.173611
1	Refrigirator-2	94.513889
2	Refrigirator-3	90.187500
3	Refrigirator-4	91.118056

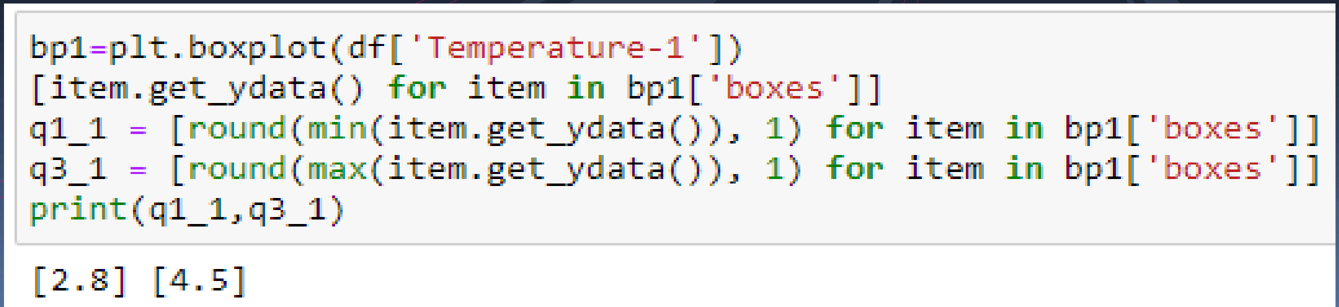
2. Set temperature range (lower threshold and upper threshold)

- Now, from observing general stats, we can see that Temp range is very varied. Max temp is almost >100 degree Celsius, which is not always the case. So we have to opt for another method to find min and max.



	Power-1	Temperature-1	Power-2	Temperature-2	Power-3	Temperature-3	Power-4	Temperature-4
count	14400.000000	14400.000000	14400.000000	14400.000000	14400.000000	14400.000000	14400.000000	14400.000000
mean	1.117388	3.815729	1.649825	2.032753	1.577693	4.118913	1.253750	2.930061
std	0.921004	2.035919	1.509209	3.344255	0.920346	3.401599	1.024295	2.578354
min	0.000000	1.650000	0.000000	-4.090000	0.000000	1.090000	0.000000	0.000000
25%	0.293000	2.840000	0.389500	-0.250000	0.278000	2.180000	0.308500	1.590000
50%	0.321000	3.805000	0.394500	1.905000	2.083000	2.840000	1.272250	2.435000
75%	2.165500	4.530000	3.318500	4.275000	2.149000	4.215000	2.355500	3.310000
max	3.055500	102.375000	4.036500	102.625000	6.038333	101.905000	6.060000	101.655000

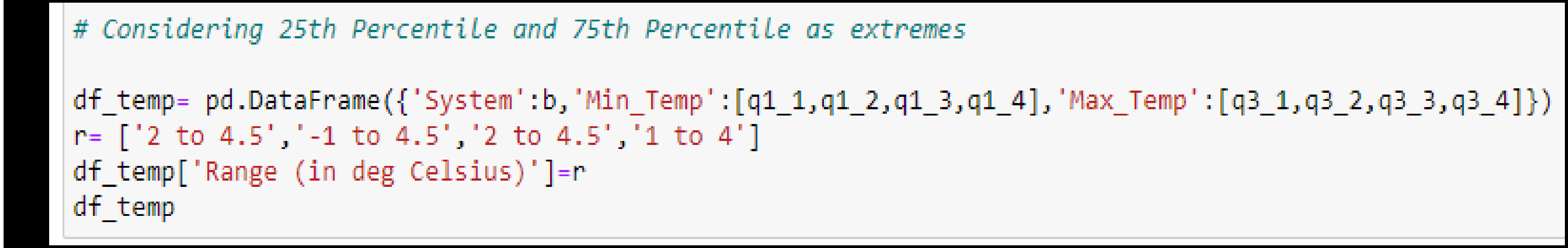
- Considering Inter-quartile range (IQR) for min and max temperatures. Here's a small code snippet to find min and max from IQR for Refrigeration system-1.



```
bp1=plt.boxplot(df['Temperature-1'])
[item.get_ydata() for item in bp1['boxes']]
q1_1 = [round(min(item.get_ydata()), 1) for item in bp1['boxes']]
q3_1 = [round(max(item.get_ydata()), 1) for item in bp1['boxes']]
print(q1_1,q3_1)
```

```
[2.8] [4.5]
```

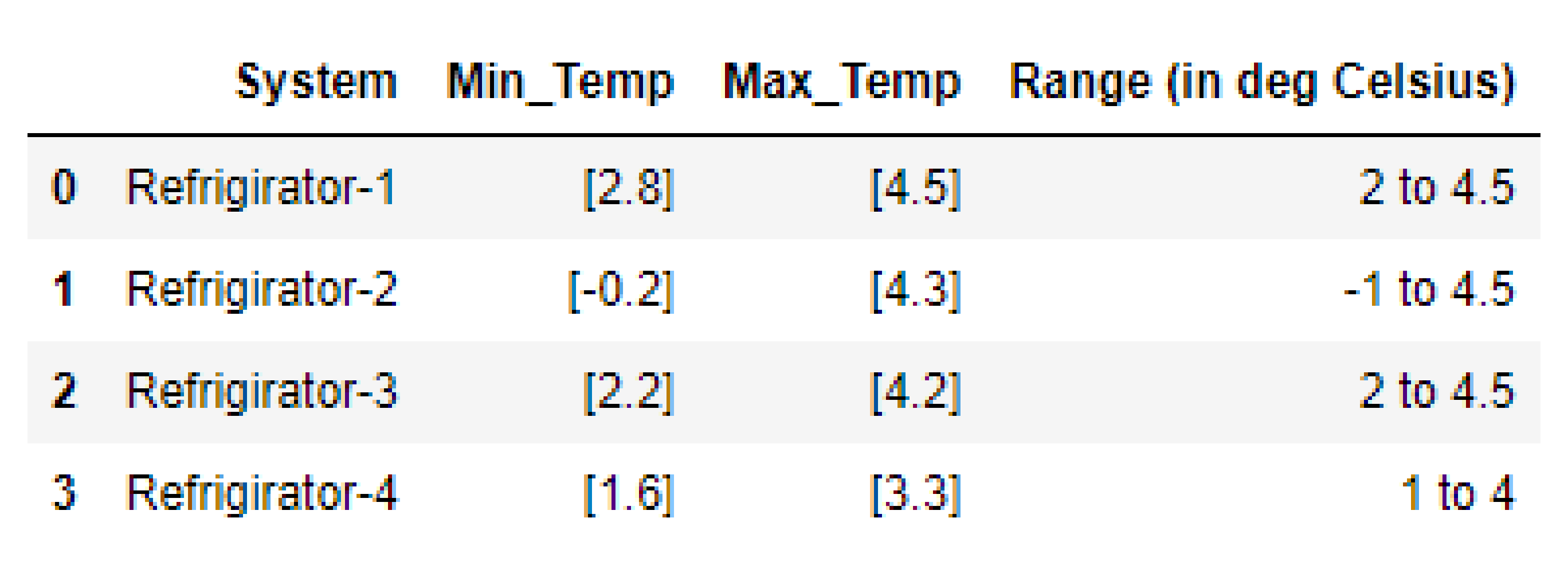
- Similarly, I computed min, max temperatures for all the refrigeration system using box plot IQR method. And made it into a data frame for easy inference.



```
# Considering 25th Percentile and 75th Percentile as extremes

df_temp= pd.DataFrame({'System':b,'Min_Temp':[q1_1,q1_2,q1_3,q1_4],'Max_Temp':[q3_1,q3_2,q3_3,q3_4]})
r= ['2 to 4.5','-1 to 4.5','2 to 4.5','1 to 4']
df_temp['Range (in deg Celsius)']=r
df_temp
```

- **Final data frame with min, max temps and the practical set range of 4 refrigeration systems**:

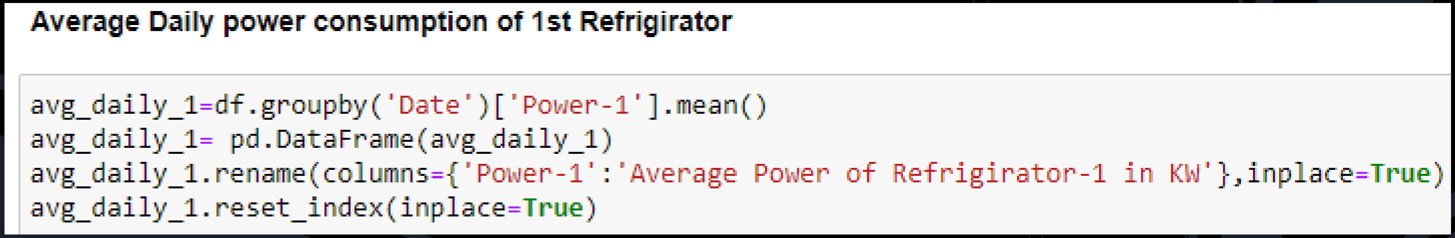


	System	Min_Temp	Max_Temp	Range (in deg Celsius)
0	Refrigirator-1	[2.8]	[4.5]	2 to 4.5
1	Refrigirator-2	[-0.2]	[4.3]	-1 to 4.5
2	Refrigirator-3	[2.2]	[4.2]	2 to 4.5
3	Refrigirator-4	[1.6]	[3.3]	1 to 4

3. Average daily energy consumption

- For this I created a new column consisting on date only, to group and calculate average power consumption daily for the 4 systems.

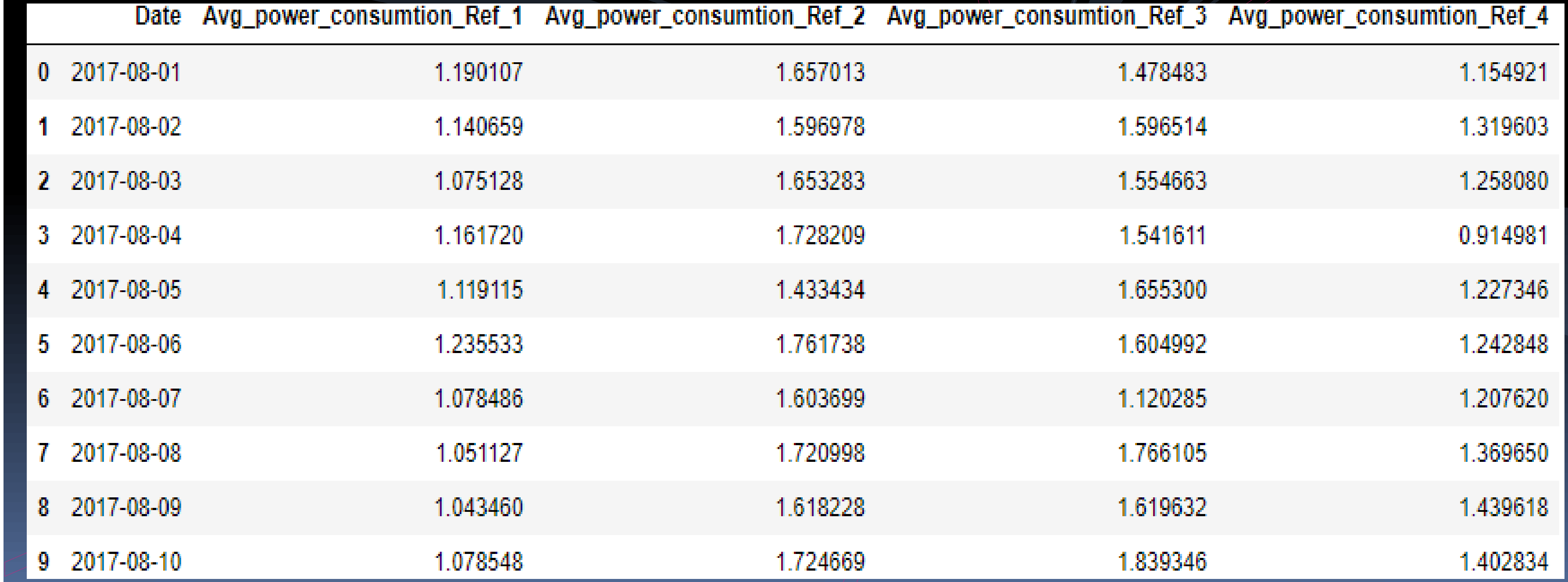
A small code snippet to calculate grouped averages of 1st system:



Average Daily power consumption of 1st Refrigirator

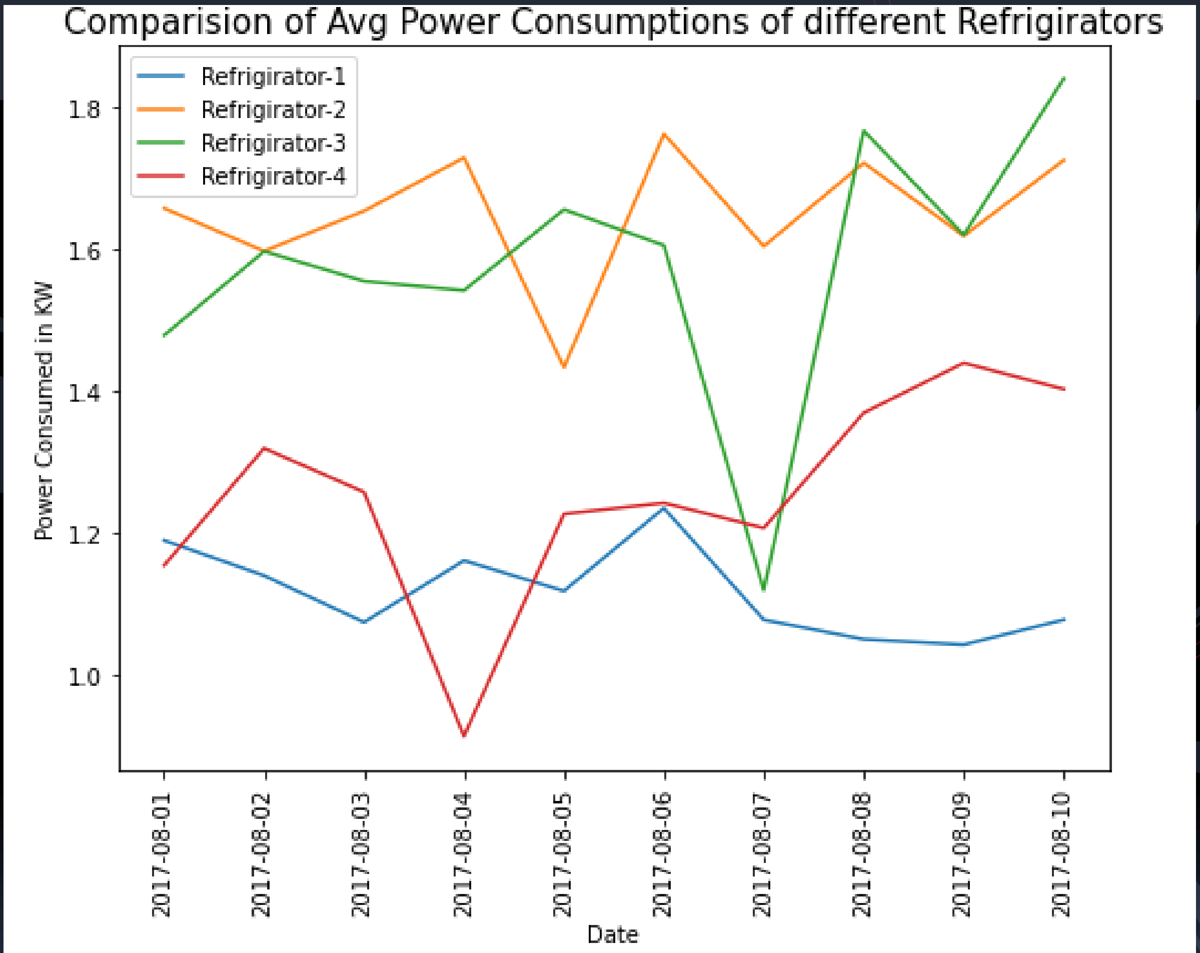
```
avg_daily_1=df.groupby('Date')['Power-1'].mean()
avg_daily_1= pd.DataFrame(avg_daily_1)
avg_daily_1.rename(columns={'Power-1':'Average Power of Refrigirator-1 in KW'},inplace=True)
avg_daily_1.reset_index(inplace=True)
```

Similarly, calculating for others and consolidating, we get:



	Date	Avg_power_consumtion_Ref_1	Avg_power_consumtion_Ref_2	Avg_power_consumtion_Ref_3	Avg_power_consumtion_Ref_4
0	2017-08-01	1.190107	1.657013	1.478483	1.154921
1	2017-08-02	1.140659	1.596978	1.596514	1.319603
2	2017-08-03	1.075128	1.653283	1.554663	1.258080
3	2017-08-04	1.161720	1.728209	1.541611	0.914981
4	2017-08-05	1.119115	1.433434	1.655300	1.227346
5	2017-08-06	1.235533	1.761738	1.604992	1.242848
6	2017-08-07	1.078486	1.603699	1.120285	1.207620
7	2017-08-08	1.051127	1.720998	1.766105	1.369650
8	2017-08-09	1.043460	1.618228	1.619632	1.439618
9	2017-08-10	1.078548	1.724669	1.839346	1.402834

Visualization of comparing avg daily power consumptions:



We can clearly infer that power consumption is least for Refrigeration system-1

And the order goes, **2>3>4>1 (In order of avg power consumption)**

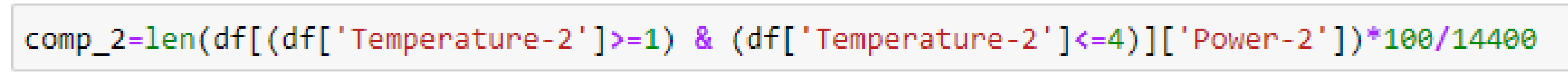
4. Percentage of temperature compliance

Compliance percentage is the percent of items with in the given range, here Temperature Compliance is no of values lying between the ideal range 1-4 deg Celsius would be compliant.

Formula = no.of compliant values (i.e. temps lying between 1-4)*100/ total values recorded.

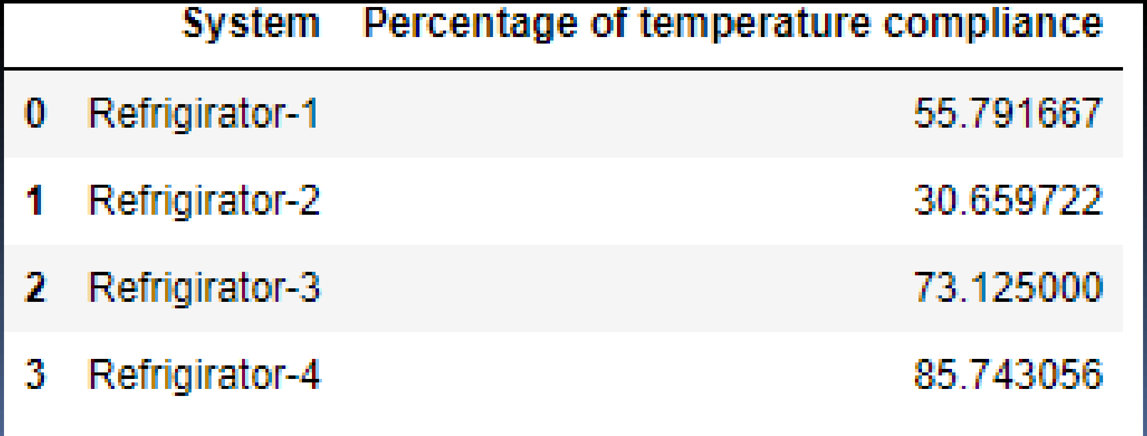
- So first we'll count no.of values lying between 1-4 degree Celsius in each temp profile. And then find their percentage. Here's the code snippet .

Ex: for system 2. we'll calculate similarly for rest of the systems as well.



```
comp_2=len(df[(df['Temperature-2']>=1) & (df['Temperature-2']<=4)]['Power-2'])*100/14400
```

- Final table with all consolidated values:



	System	Percentage of temperature compliance
0	Refrigirator-1	55.791667
1	Refrigirator-2	30.659722
2	Refrigirator-3	73.125000
3	Refrigirator-4	85.743056

5. Final Observations & Thoughts

After careful comparative analysis of several parameters, we can conclude that we calculated and compared given parameters like duty cycle, temperature range, avg power consumption and percentage of temperature compliance.

- Comparing duty cycle, **1>2>4>3** (in terms of active time)
- Comparing temperature ranges, **system 4 is close to the ideal ra**nge and system 2 is far deviated.
- Comparing average daily power consumption (from the graph) , order goes, **2>3>4>1.**
- Coming to compliance, order goes **4>3>1>2**

Finally inferring from above statements we can say that Refrigeration system-4 performs well than the rest.

THANK YOU!