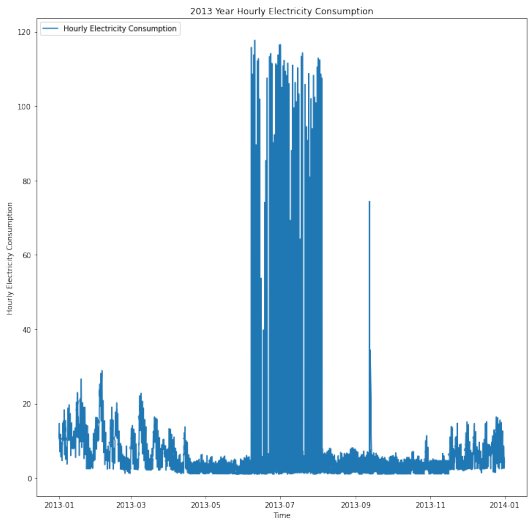
Guzman Homework Assignment Two

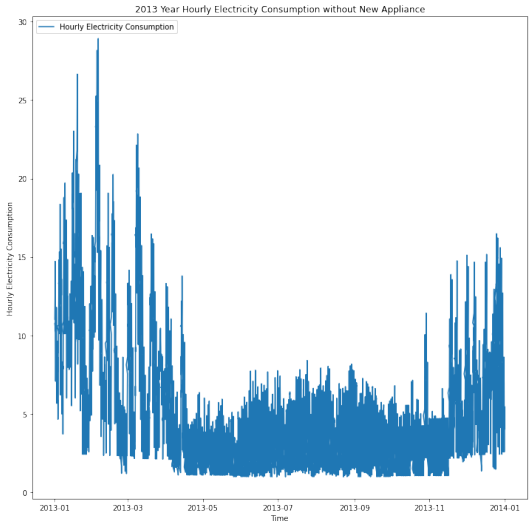
November 17, 2020

Assignment2 Meter Data formatting

1. **Data Merge**
   The goal of this task is to merge two data frames. One data frame is the time series electricity consumption for a factory, while another data frame is the electricity consumption of an new appliance that previous data frame does not capture. According to the requirement, I performed the data merge through the python package dfply. Besides, I also created a version of solution from pandas data frame for reference.I noticed that some columns in assignment two like "Gas:Facility [kW](Hourly), Heating:Gas [kW](Hourly)". These columns' names don't contain Electricity, but they are labeled with the unit of electricity kW. My first try is to summarize the columns containing Electricity. I sent an email to Brain for clarification, but I did not get response. I switched back to summarize all columns following the instruction in the GitHub. I kept the format of all columns unchanged for the second data frame which is the output data frame.
   First, I transfer the electricity column of the new appliance data frame with correct unit kW. Then, I transfer the format of the datetime to be consistent with the format of time column in the second data frame. I left join the second data frame with the new appliance data frame with identity key datetime. Group the merged data frame by its time column and summarize the electricity consumption in to a new column.

2. **Abnormality**



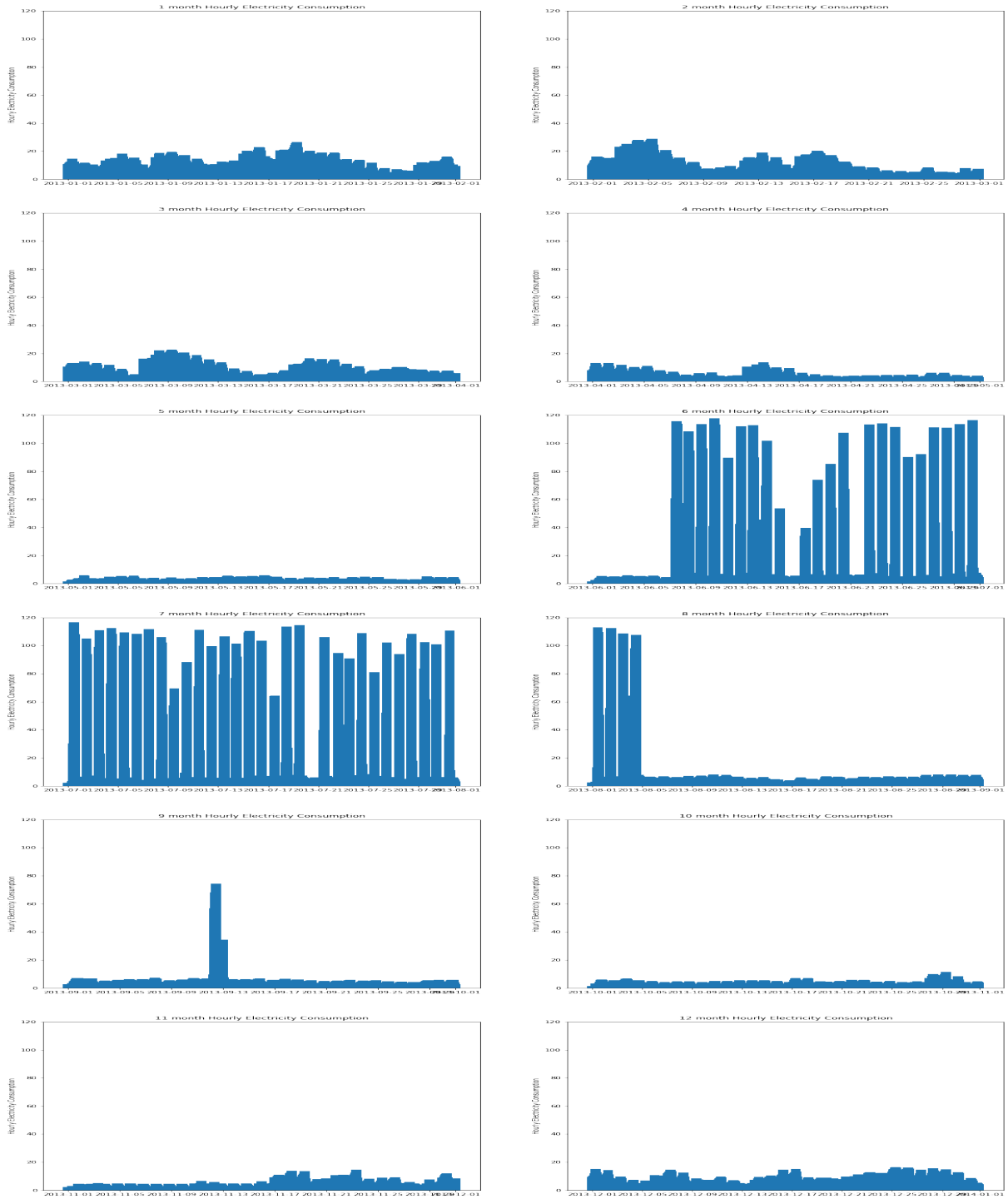


We can see some abnormalities in the time series total electricity, especially during the June, July, August, and September. Without counting the new appliance electricity consumption, we can see that it is flat from June to September. Thus, we can conclude that the new appliance bring up a lot of electricity consumption from June to September. From comparing the time indices of new appliance electricity consumption and the time indices where the total electricity consumption over 20 Kw, we

find a high overlapping. Filtering out one weird spike in September, we can see that there is an abnormal electricity consumption on Sept 12 2013 and Sept 13 2013.

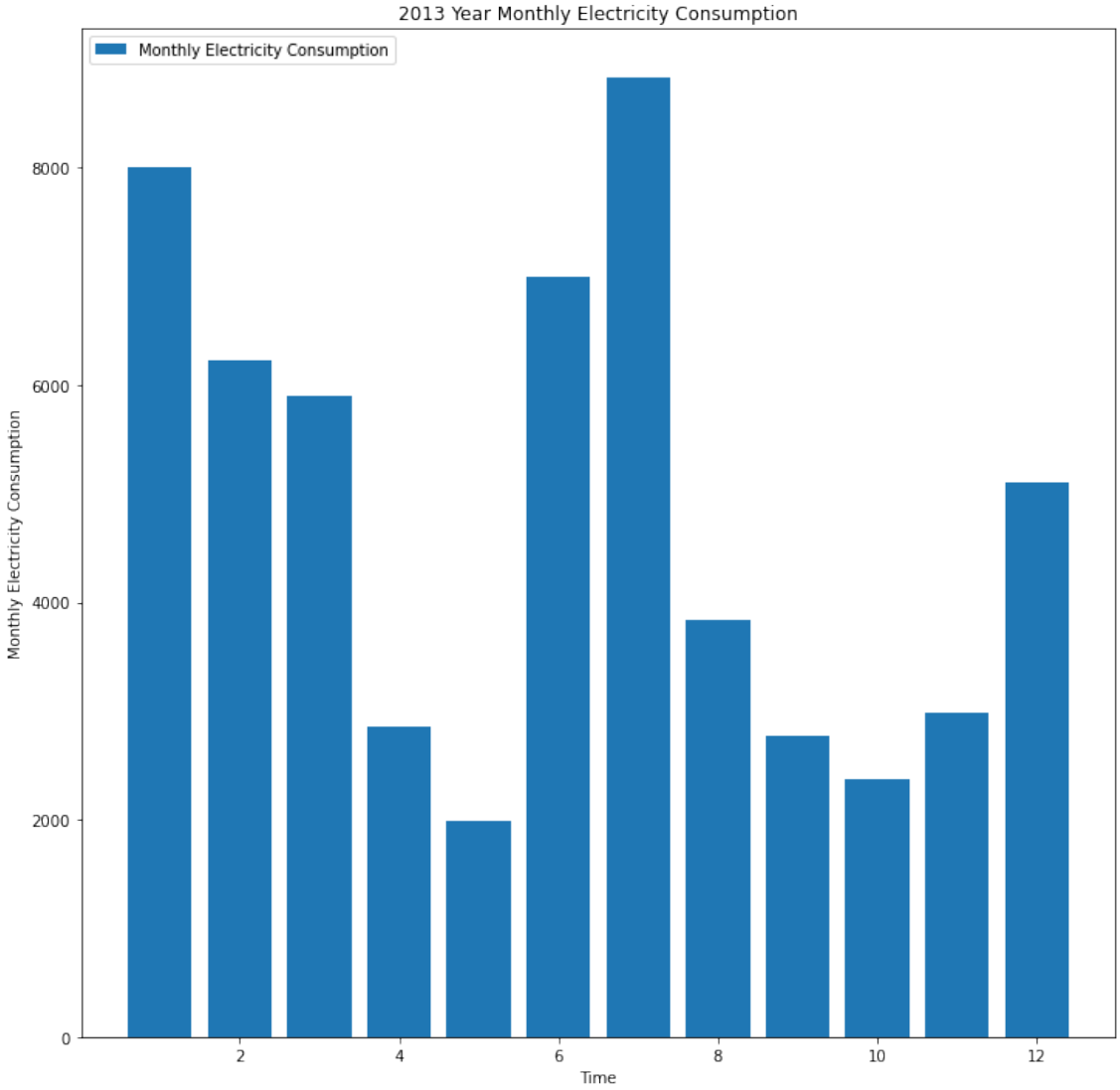
3. **Pattern**

(a) **Monthly**

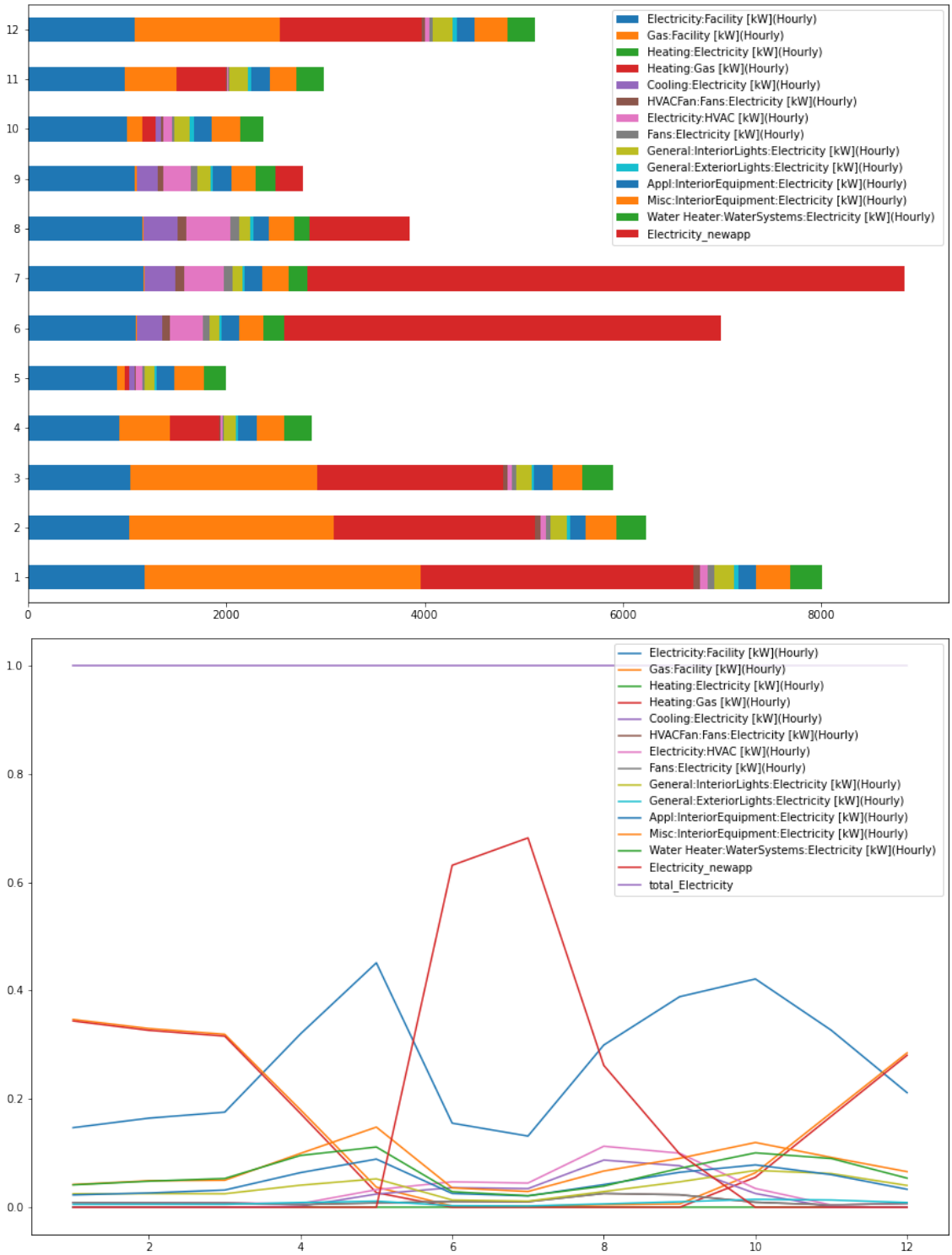


By plotting the monthly electricity consumption, we can gain some insight about the monthly

trend electricity consumption. We can see that electricity consumption stay at 120 kW from June to early August, but electricity consumption data sets are spare. It achieve 120 Kw at one day and stay at very low level where we can imply that the new appliance may need cooling period. There is a spike in September. In January, the electricity consumption is more stable around 15 Kw.

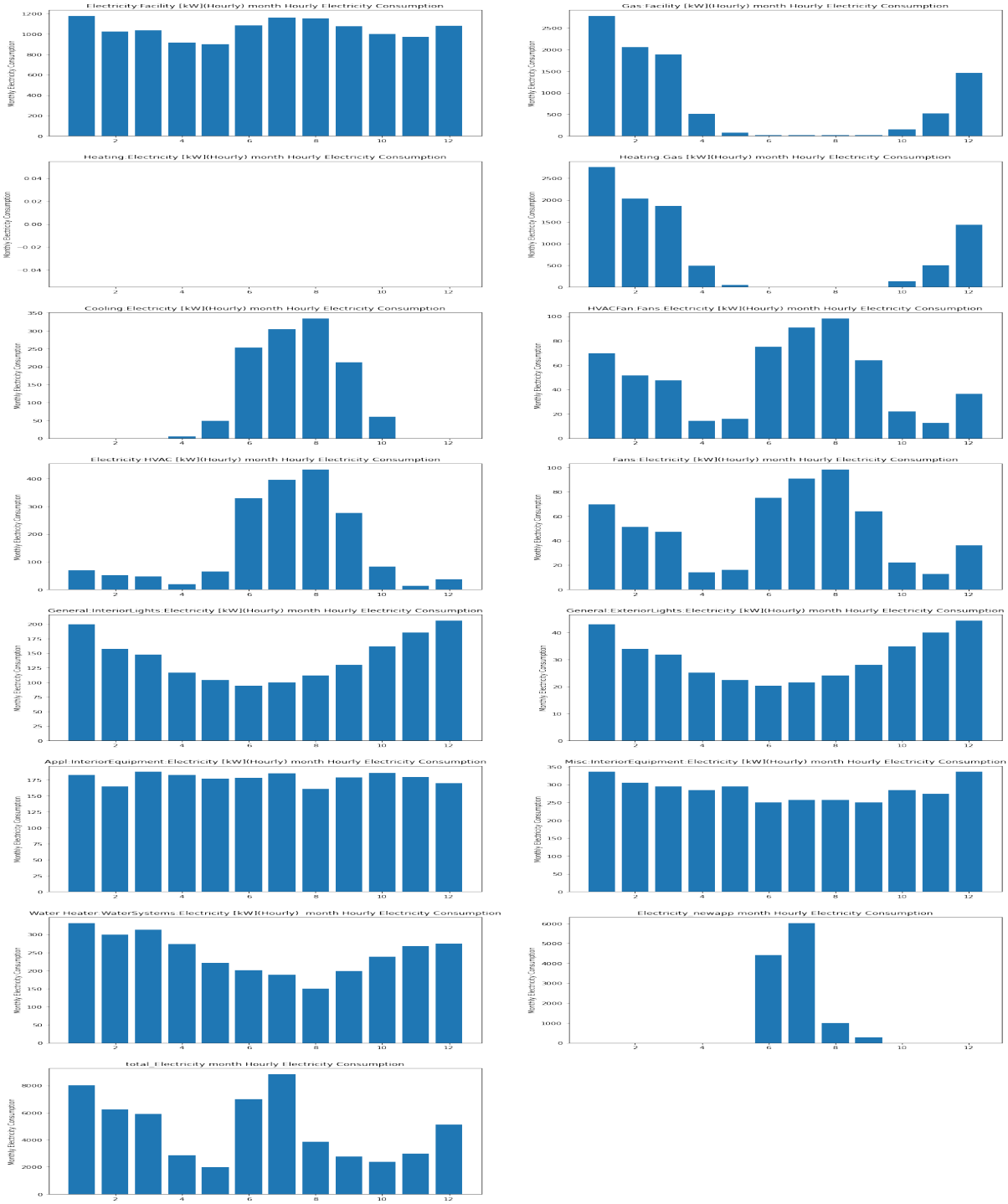


Summarizing monthly electricity consumption gives us some insights about how much electricity is consumed in a monthly basis. We can see that electricity consumption in July is the highest, followed by January and June. Electricity consumption in May and October is pretty low.



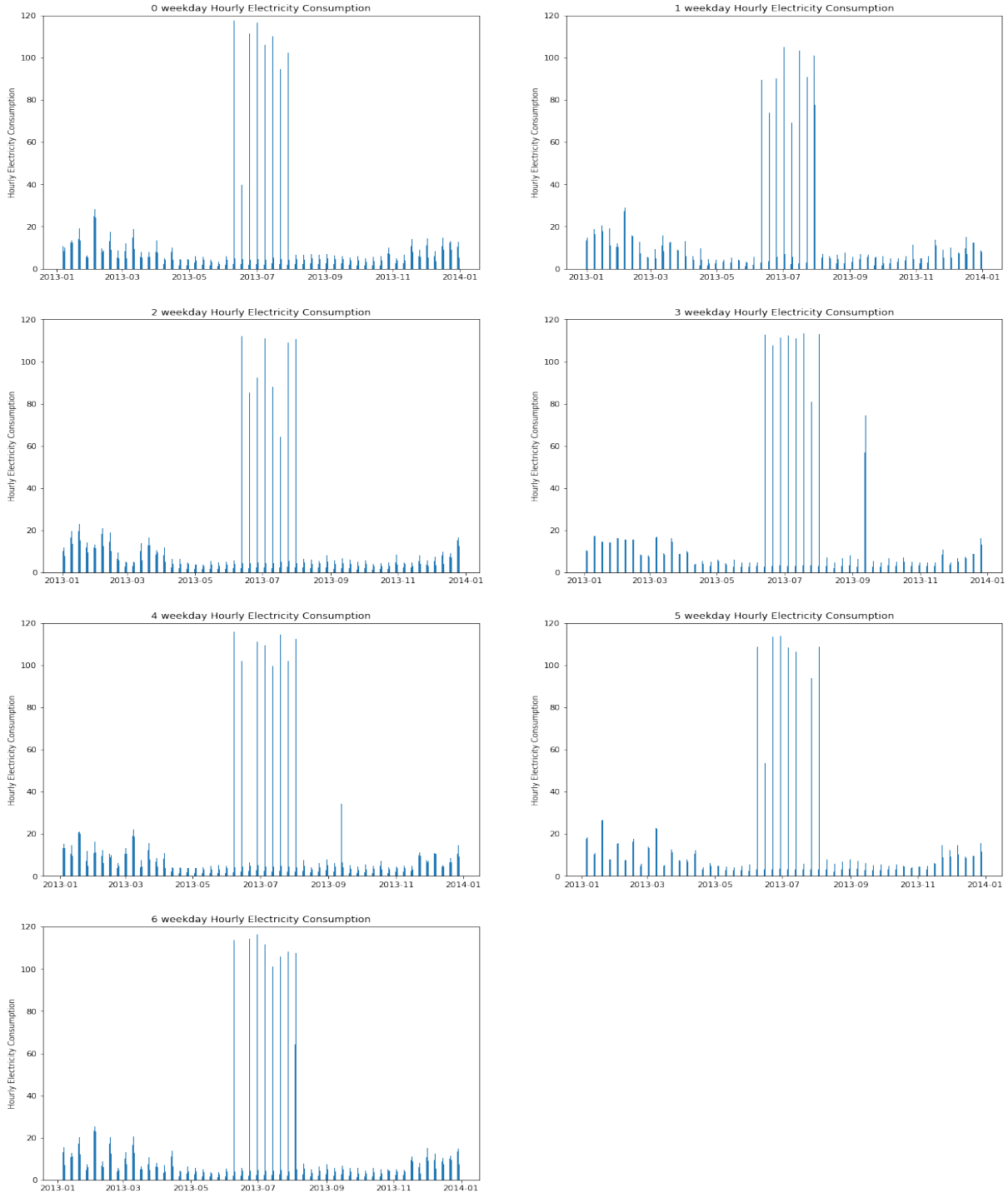
The above graphs can give us some insights about the contribution from each component to the total electricity by month. At the first glance, the contribution from Facility Electricity is fairly stable throughout the year. Electricity from Facility Gas and Heating increase from October, peak at January and fade at May. The electricity consumption from Cooling Electricity only occurs from May to October. What we can imply from this contribution graph is that weather at the place the factory locates. The winter looks long from October to May. The summer is a bit of short and hot. The factory may locate around Michigan or Chicago, assuming this factory is in

the United States.

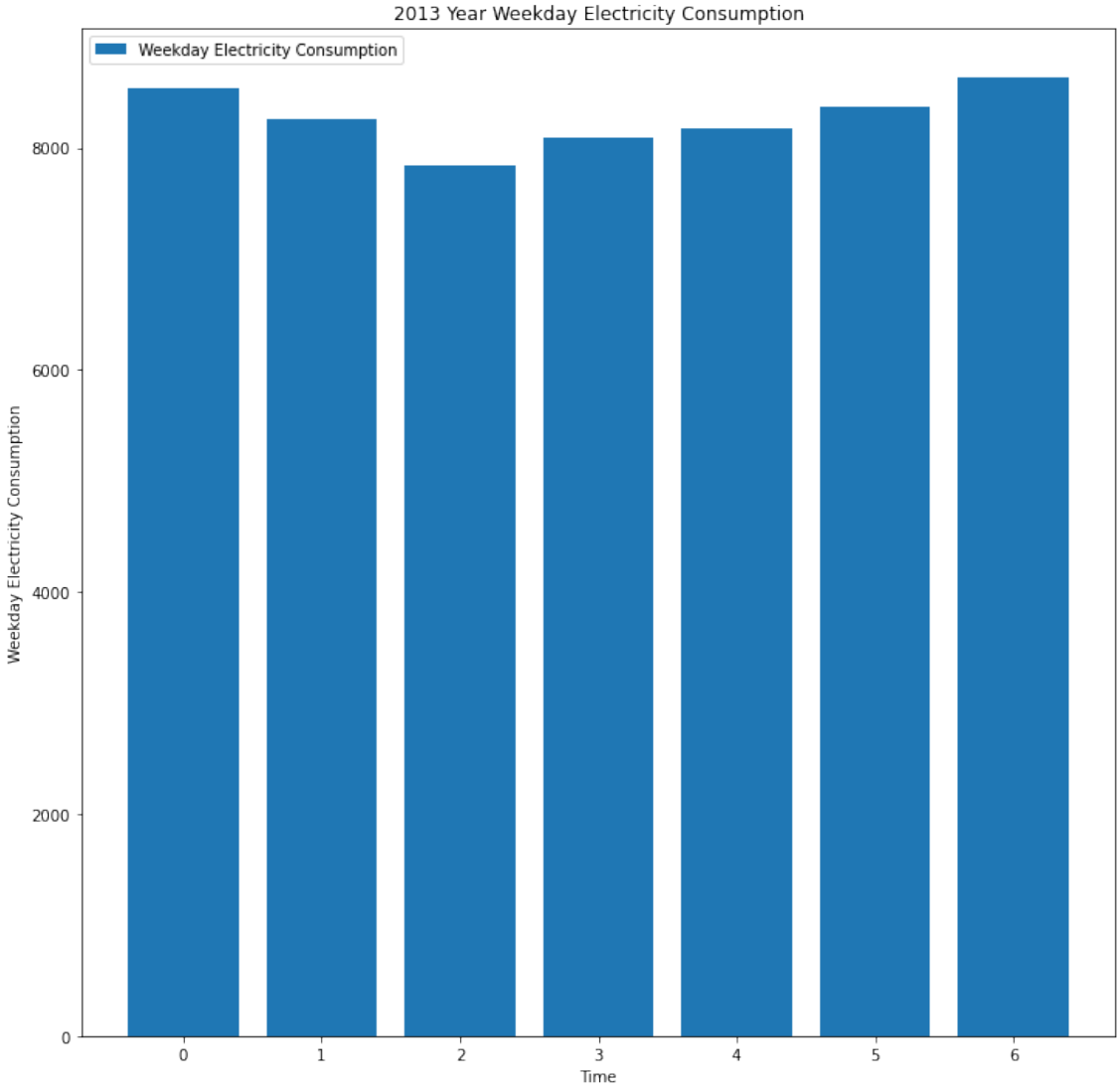


This graph provides a closer look at each component of the electricity consumption. Except for what we find in the previous graph, we can infer that the factory is located at the Northern hemisphere because the interior lights electricity and exterior light electricity consumption have a local minimum at June. The sunlight is strong enough so that the factory does not use a lot of interior and exterior lights during the summer. Moreover, the factory does not use electricity for heating.

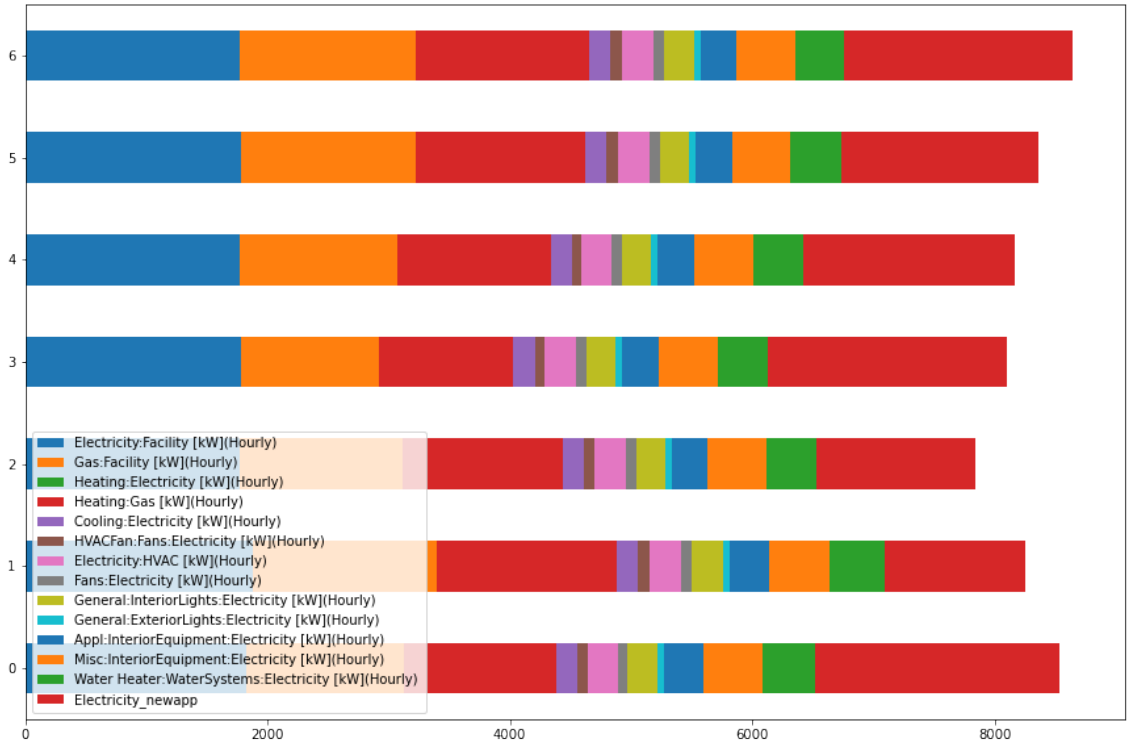
(a) **Weekday**

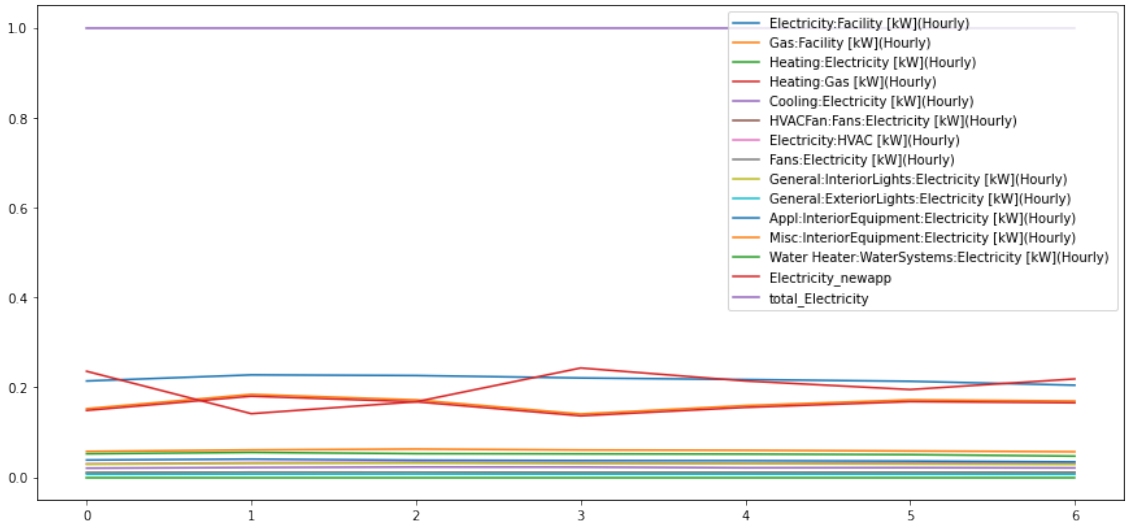


Given the weekday electricity consumption, what I can see from the graph is that the factory functions seven days a week even on Saturday and Sunday. Something we should notice is that the abnormality in September only occurs on Thursday and Friday.



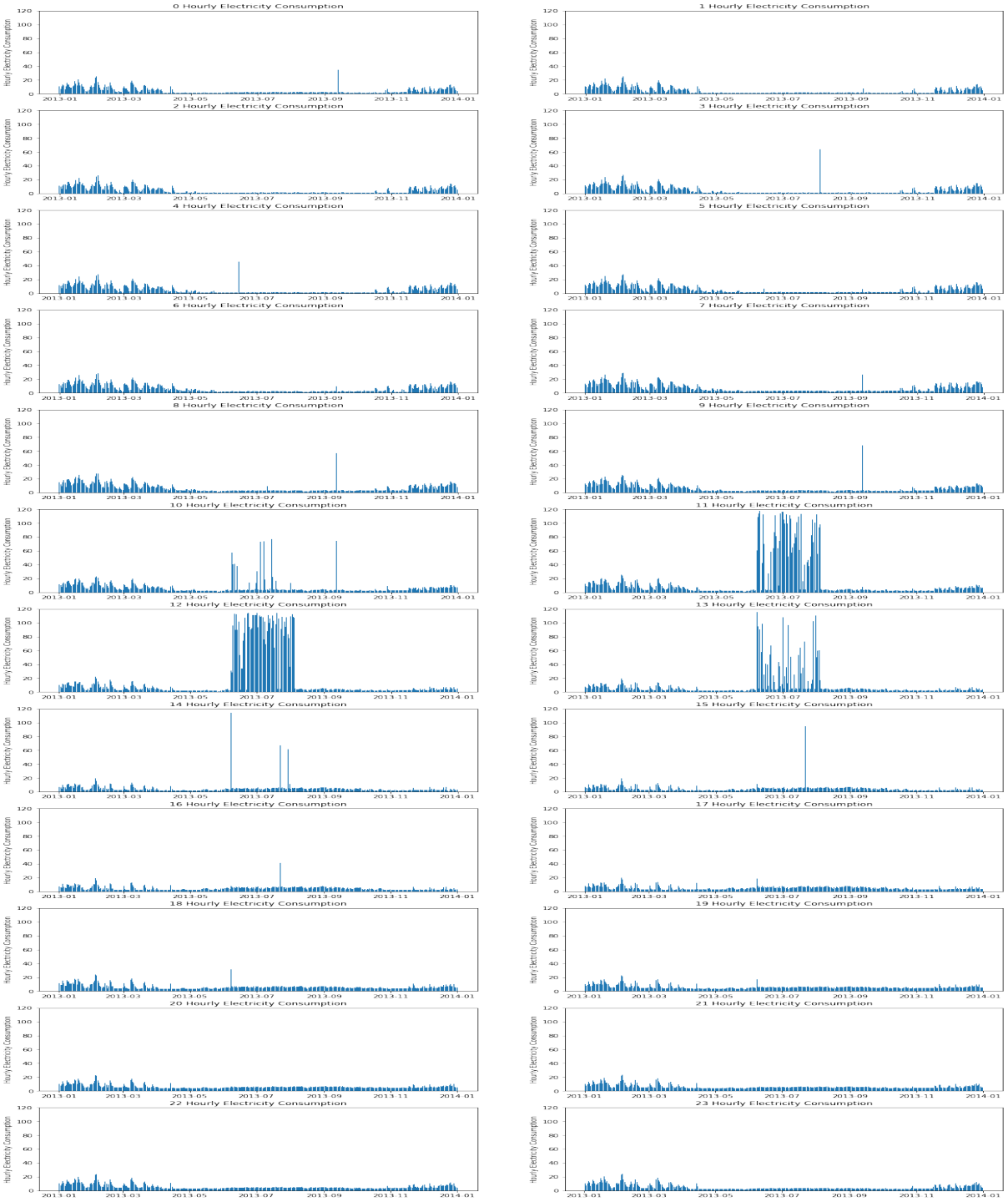
Wednesday has the lowest electricity consumption among these weekdays. Overall Monday and Sunday have the highest electricity consumption. Assuming more electricity consumption means more production, what industry is the company in so that the factory has higher production on Sunday than weekdays.


Electricity:Facility [kW](Hourly)
Gas:Facility [kW](Hourly)
Heating:Electricity [kW](Hourly)
Heating:Gas [kW](Hourly)
Cooling:Electricity [kW](Hourly)
HVACFan:Fans:Electricity [kW](Hourly)
Electricity:HVAC [kW](Hourly)
Fans:Electricity [kW](Hourly)
General:InteriorLights:Electricity [kW](Hourly)
General:ExteriorLights:Electricity [kW](Hourly)
Appl:InteriorEquipment:Electricity [kW](Hourly)
Misc:InteriorEquipment:Electricity [kW](Hourly)
Water Heater:WaterSystems:Electricity [kW](Hourly)
Electricity_newapp
0
1
2
3
4
5
6
0
2000
4000
6000
8000

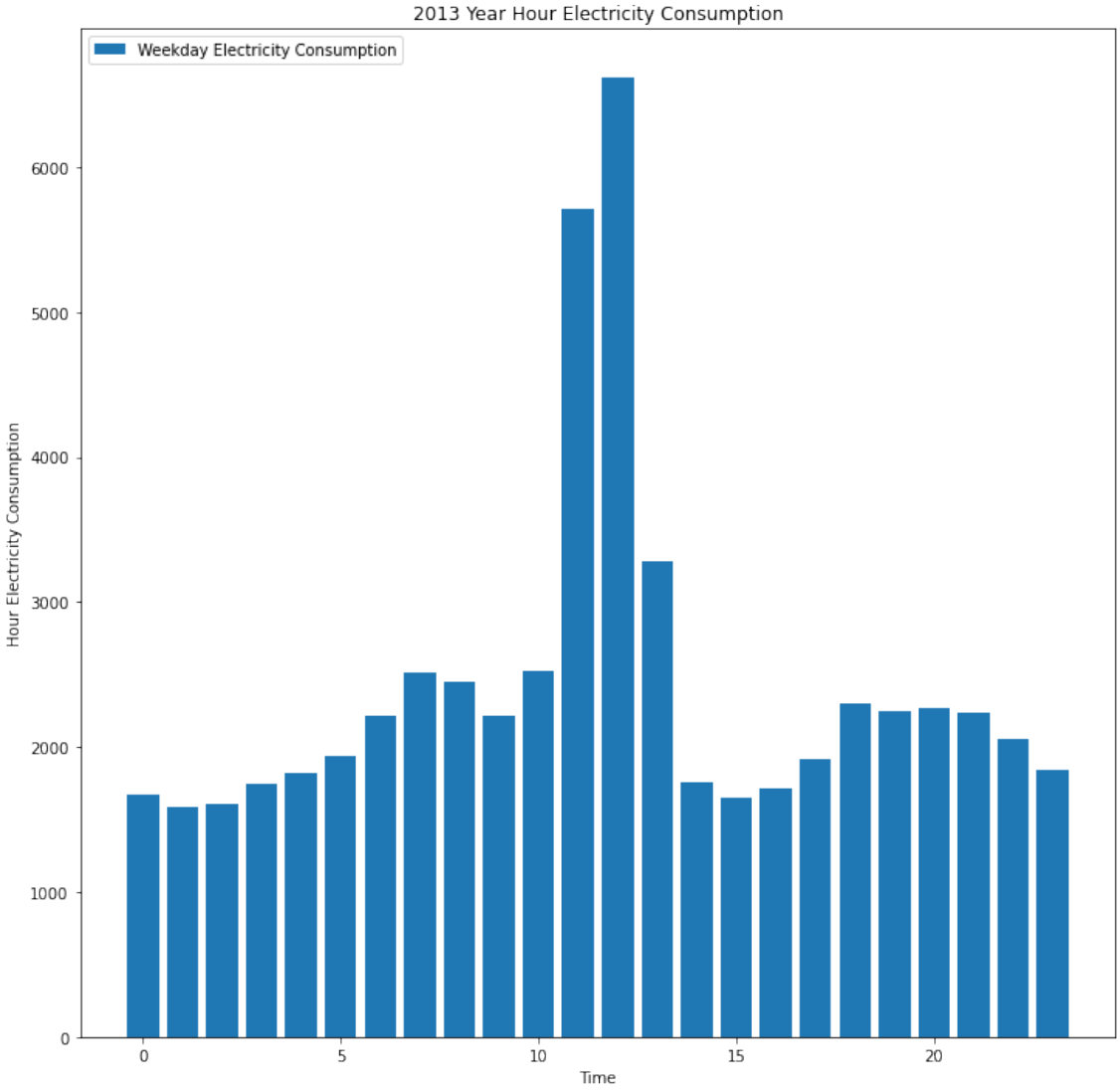

Electricity:Facility [kW](Hourly)
Gas:Facility [kW](Hourly)
Heating:Electricity [kW](Hourly)
Heating:Gas [kW](Hourly)
Cooling:Electricity [kW](Hourly)
HVACFan:Fans:Electricity [kW](Hourly)
Electricity:HVAC [kW](Hourly)
Fans:Electricity [kW](Hourly)
General:InteriorLights:Electricity [kW](Hourly)
General:ExteriorLights:Electricity [kW](Hourly)
Appl:InteriorEquipment:Electricity [kW](Hourly)
Misc:InteriorEquipment:Electricity [kW](Hourly)
Water Heater:WaterSystems:Electricity [kW](Hourly)
Electricity_newapp
total_Electricity
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6

All component seems to contribute evenly throughout the week. Except for New Appliance Electricity consumption, it looks like electricity consumption on Tuesday achieves the highest among most of electricity consumption components. Solely focusing on the new appliance consumption, we may find that the new appliance electricity consumption on Monday, Thursday and Sunday are relatively high compared to Tuesday, Wednesday, Friday and Saturday.

4. (a) **Hour**

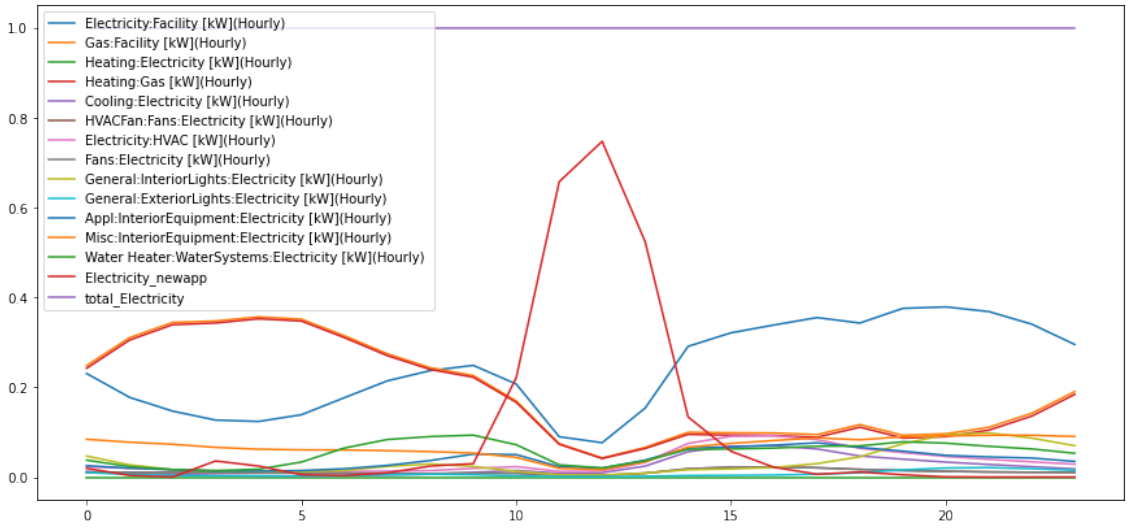


Given the weekday electricity consumption, what I can see from the graph is that the factory functions seven days a week even on Saturday and Sunday. Something we should notice is that the abnormality in September only occurs on Thursday and Friday.



Wednesday has the lowest electricity consumption among these weekdays. Overall Monday and Sunday have the highest electricity consumption. Assuming more electricity consumption means more production, what industry is the company in so that the factory has higher production on Sunday than weekdays.

Electricity:Facility [kW](Hourly)
Gas:Facility [kW](Hourly)
Heating:Electricity [kW](Hourly)
Heating:Gas [kW](Hourly)
Cooling:Electricity [kW](Hourly)
HVACFan:Fans:Electricity [kW](Hourly)
Electricity:HVAC [kW](Hourly)
Fans:Electricity [kW](Hourly)
General:InteriorLights:Electricity [kW](Hourly)
General:ExteriorLights:Electricity [kW](Hourly)
Appl:InteriorEquipment:Electricity [kW](Hourly)
Misc:InteriorEquipment:Electricity [kW](Hourly)
Water Heater:WaterSystems:Electricity [kW](Hourly)
Electricity_newapp
0 1000 2000 3000 4000 5000 6000


Electricity:Facility [kW](Hourly)
Gas:Facility [kW](Hourly)
Heating:Electricity [kW](Hourly)
Heating:Gas [kW](Hourly)
Cooling:Electricity [kW](Hourly)
HVACFan:Fans:Electricity [kW](Hourly)
Electricity:HVAC [kW](Hourly)
Fans:Electricity [kW](Hourly)
General:InteriorLights:Electricity [kW](Hourly)
General:ExteriorLights:Electricity [kW](Hourly)
Appl:InteriorEquipment:Electricity [kW](Hourly)
Misc:InteriorEquipment:Electricity [kW](Hourly)
Water Heater:WaterSystems:Electricity [kW](Hourly)
Electricity_newapp
total_Electricity
0.0 0.2 0.4 0.6 0.8 1.0
0 5 10 15 20

All component seems to contribute evenly throughout the week. Except for New Appliance Electricity consumption, it looks like electricity consumption on Tuesday achieves the highest among most of electricity consumption components. Solely focusing on the new appliance consumption, we may find that the new appliance electricity consumption on Monday, Thursday and Sunday are relatively high compared to Tuesday, Wednesday, Friday and Saturday.

```
1 import numpy as np
```

5. mmbtus is million British thermal units