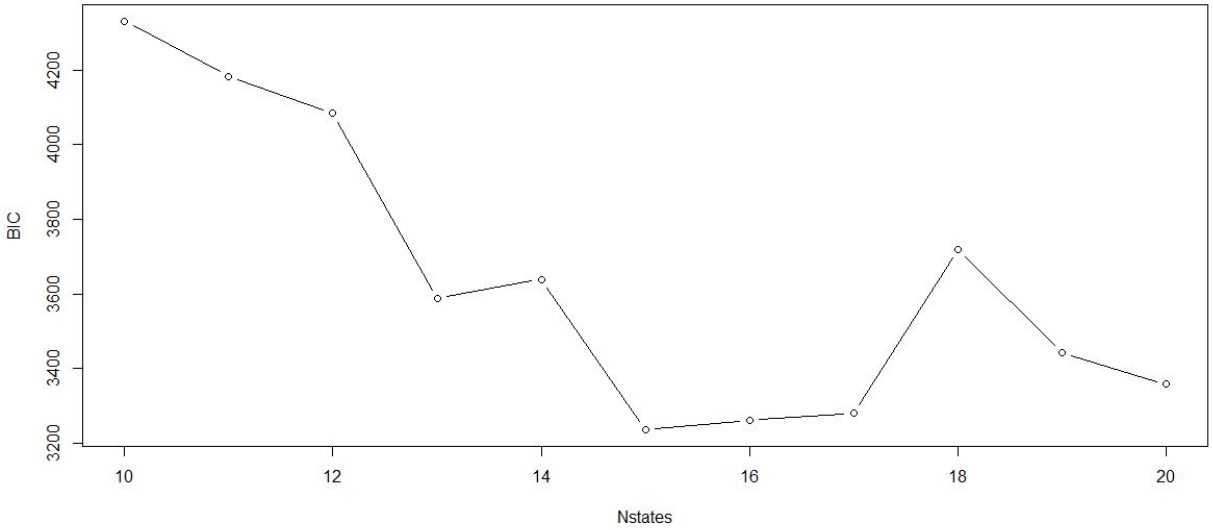
CMPT 318 - Assignment 2 - Group 15
Daniel Tavaszi, Stefano Macri, Khizr Pardhan, Darien Flamont, Andrei Hristea

Question 1:

After classifying our Morning Sunday data set to Sunday's between 8 and 11 am we decided to use the feature Global Active Power as a good representation of "normal electricity consumption" at these times. Running the HMM model for states between 2 and 20 we received a decreasing BIC for each increase in the amount of states (N) until we get to N=15, where we receive our lowest BIC, the BIC values start to increase. The transition matrix for N=15 we get high diagonal properties meaning states like to transition back to themselves frequently instead of other states. Using our 3 best BIC's our best 3 hidden markov models are:

N = 15, BIC = 3237 - Best Fit HMM
N = 16, BIC = 3263
N = 17, BIC = 3280



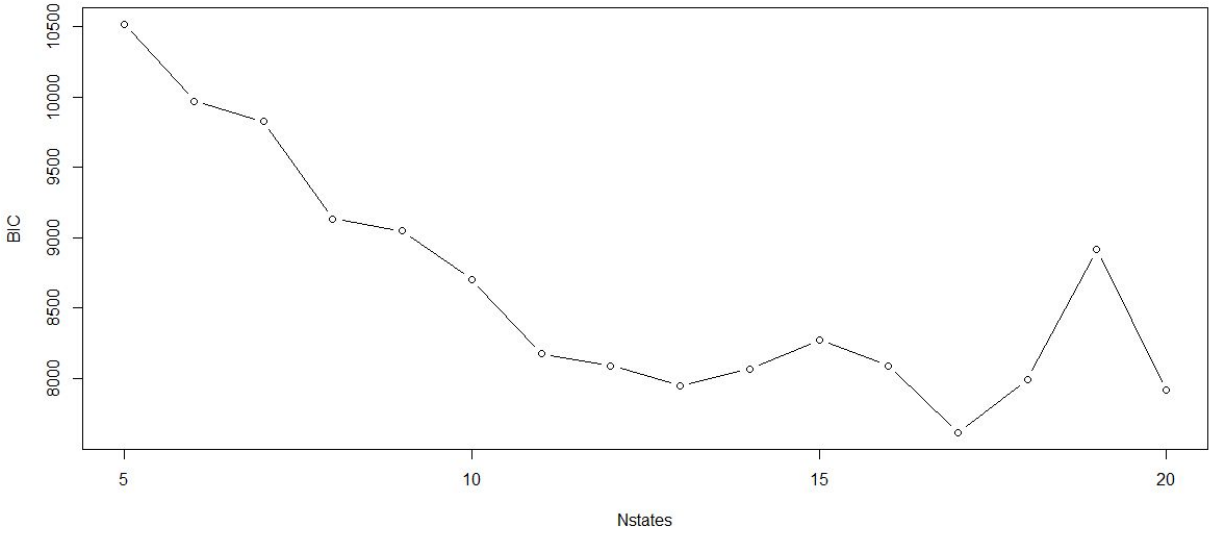
Question 2:

Repeating the process seen in Question 1 for a timeframe between 9pm and 12am and continuing to use Global Active Power as our feature set we received the best BICs at state 13, 17, and 20 where 17 states is our lowest BIC and best fit HMM. The plot below shows the BIC graph for the states chosen and our BICs received from R are as follows:

N = 13, BIC = 7951

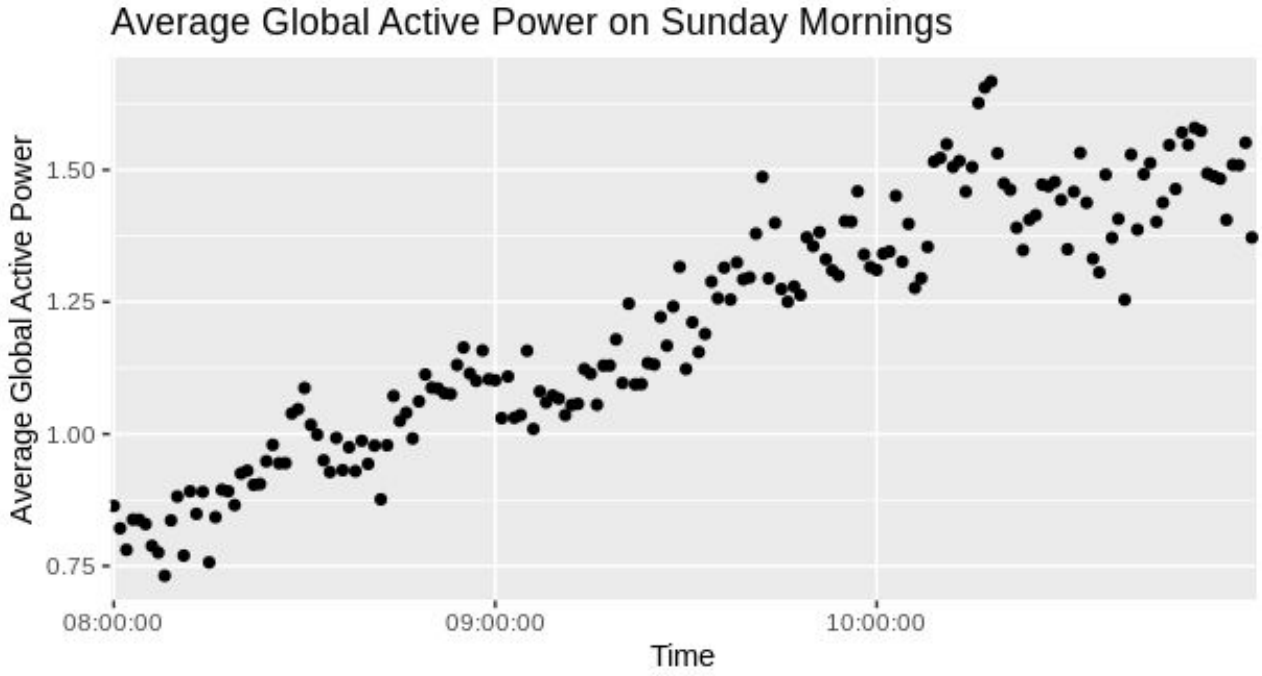
N = 17, BIC = 7617 - Best Fit HMM

N = 20, BIC = 7918

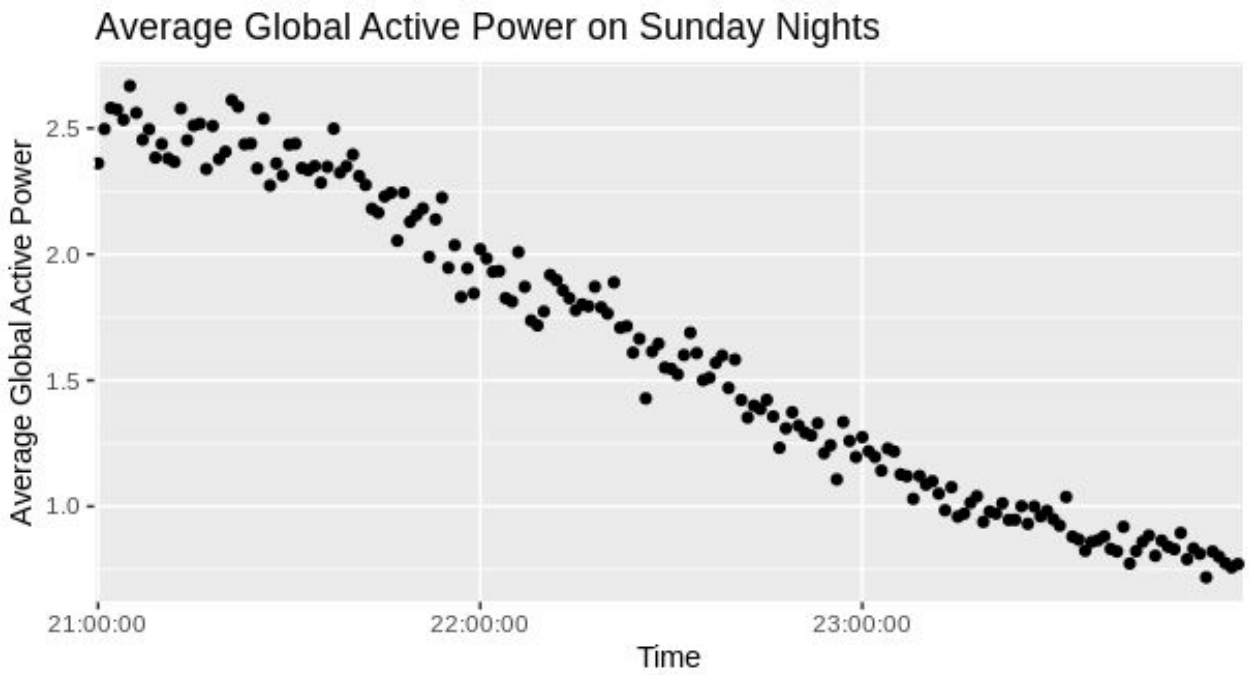


Question 3 A:

Using the Global Active Power as our feature we plotted the average Global Active Power for each minute between the 48 Sunday Mornings and received the following graph.



Plotting the average Global Active Power for each minute between the 48 Sunday Nights we received the following graph.

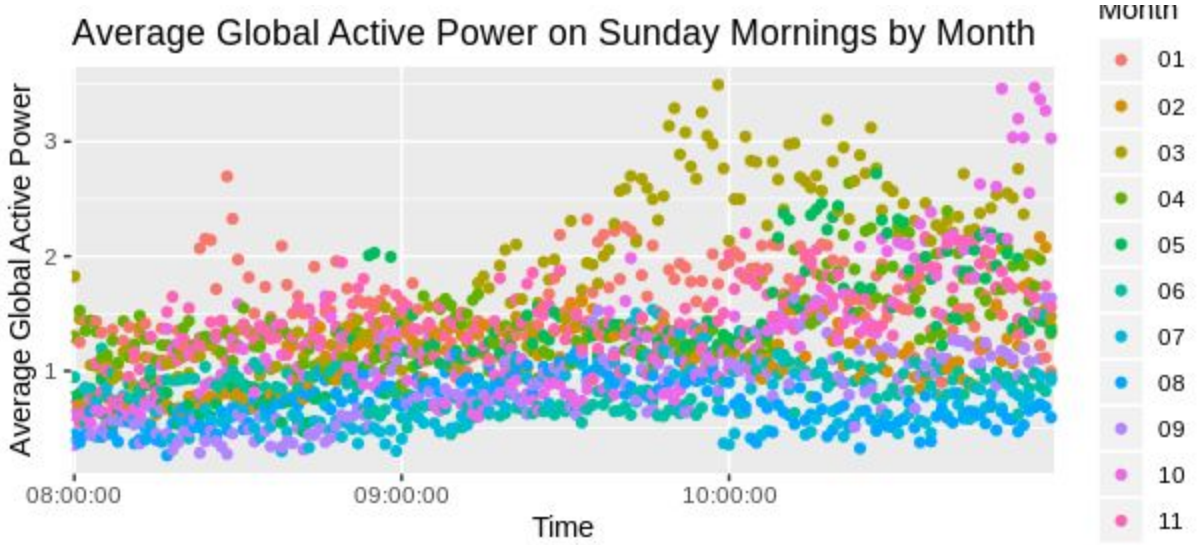


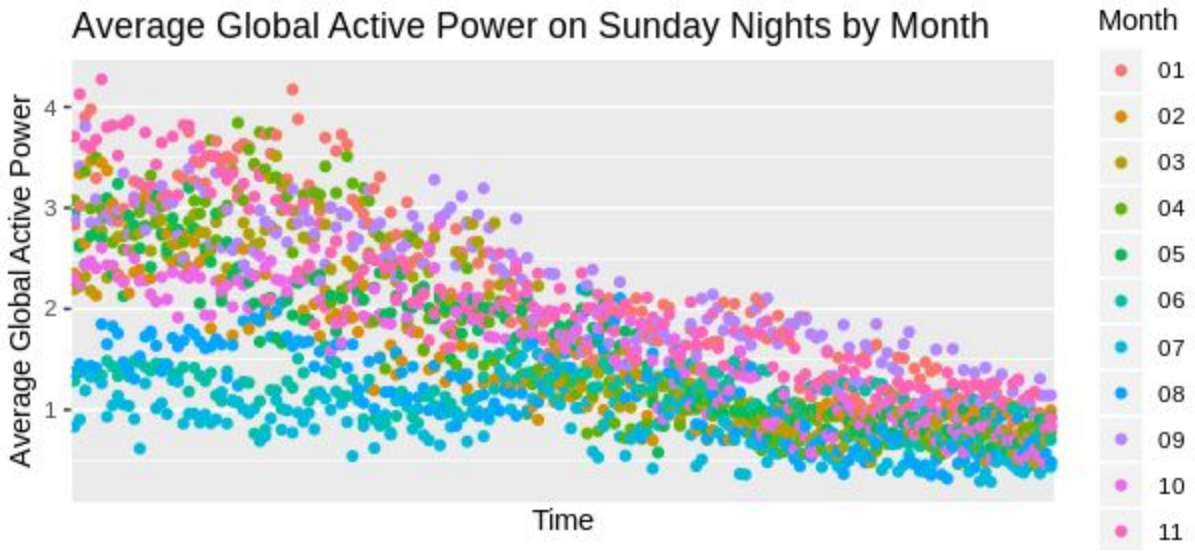
We can interpret that throughout all Sunday's in our 48 week block in the morning Average Global Active Power increases linearly for each minute between 8:00 AM leading up until 11:00

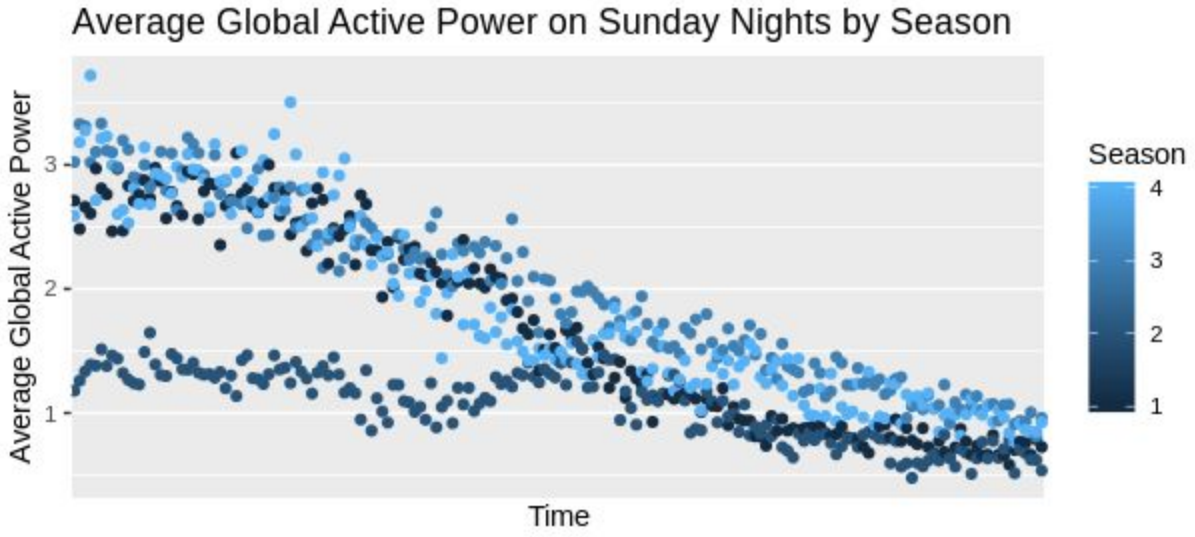
AM receiving the highest Average Global Active Power around 10:15 AM. Likewise for the night Average Global Active Power across the 48 weeks we receive a decreasing linear function between 9:00 PM and 12:00 AM. We interpret these both as normal behaviour and what we would expect from a 48 week average as people wake-up in the morning their power consumption will increase as they turn on lights, make coffee, and watch TV on a Sunday morning and will decrease as people turn off lights and the TV on a Sunday night as most people would work on Monday morning.

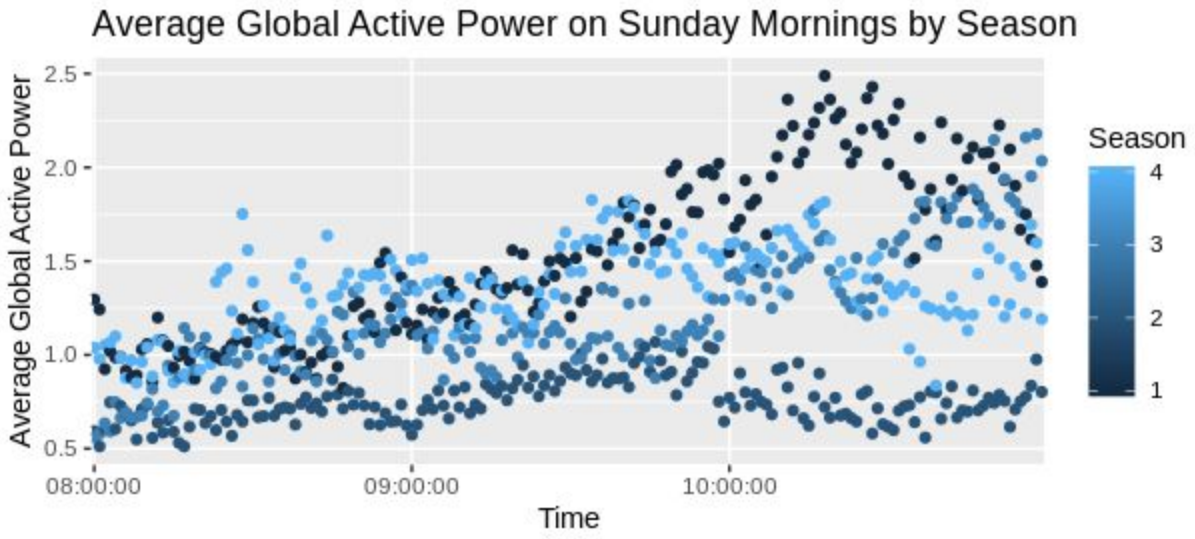
Question 3 B:

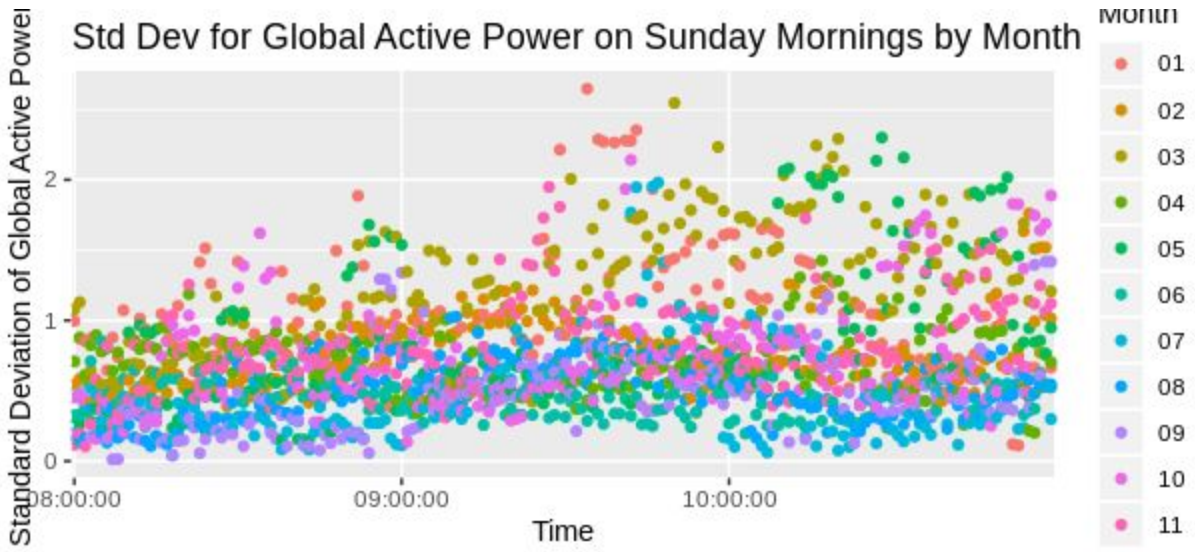
Seasonality we used a function found online that split our seasons into 4 categories. Season 1 = Spring, Season 2 = Summer, Season 3 = Fall, Season 4 = Winter. We used a function called aggregate that provided our different seasons and months we were able to get the mean and standard deviation for each different category. One caveat is that we only have data for between January 1 and December 1 and December 1st is not a Sunday so we only were able to split our data into 11 months as we have no power consumption data for December Sunday's. We have titled and linked 8 different plots, Monthly Averages for Sunday Morning and Night, Seasonal Averages for Sunday Morning and Night, Monthly Standard Deviation for Sunday Morning and Night, and Seasonal Standard Deviation for Sunday Morning and Night. Some apparent features we can see right away is that for the colder **months** January, February, October, November we see higher Global Active Power Averages during both Morning and Night than the warmer **months** assuming more power is used for heating the house. **Seasonally** we see during the night time the winter and fall seasons have higher Global Active Power Averages for the same apparent reason as above keeping the house warm at night uses power, but during the Mornings it seems that Spring has the highest **seasonal** averages especially around 10:15 AM as seen previously in our analysis for 3A this may be noise or an anomaly in our data but we cannot be certain for sure as we are lacking Ground Truth. **Seasonally** during the morning summer has the lowest power average indicating people are outside enjoying summer and not using household electricity.

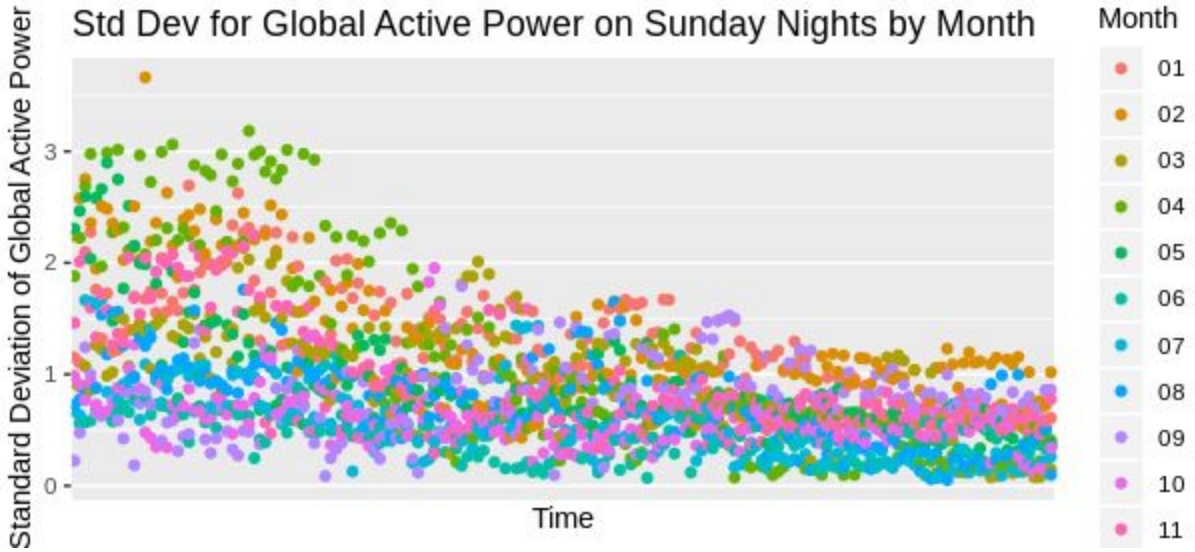


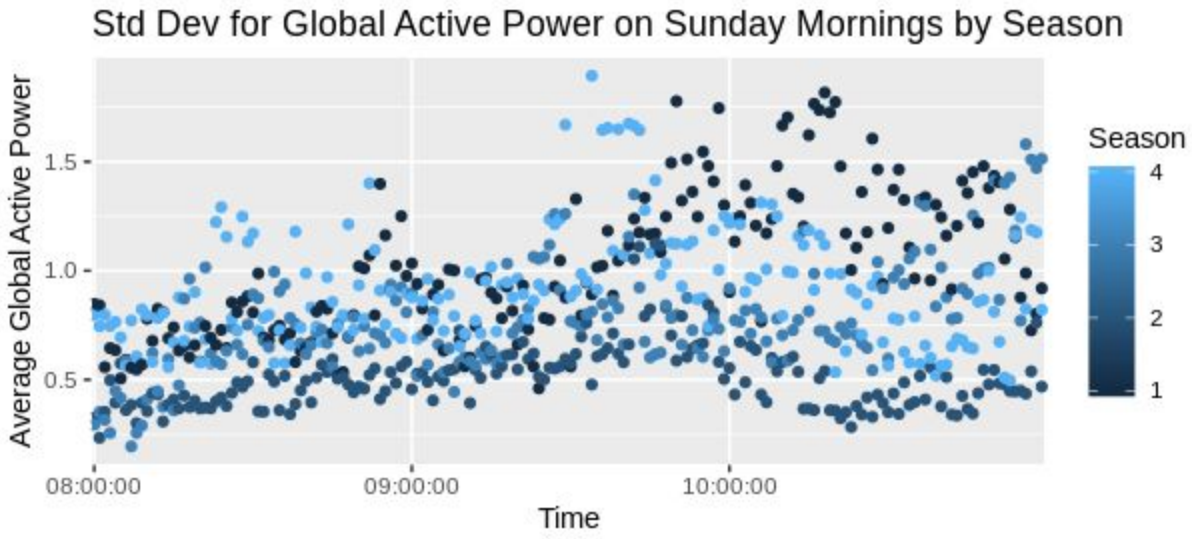

Average Global Active Power on Sunday Nights by Month
Average Global Active Power
4
3
2
1
Time
Month
01
02
03
04
05
06
07
08
09
10
11

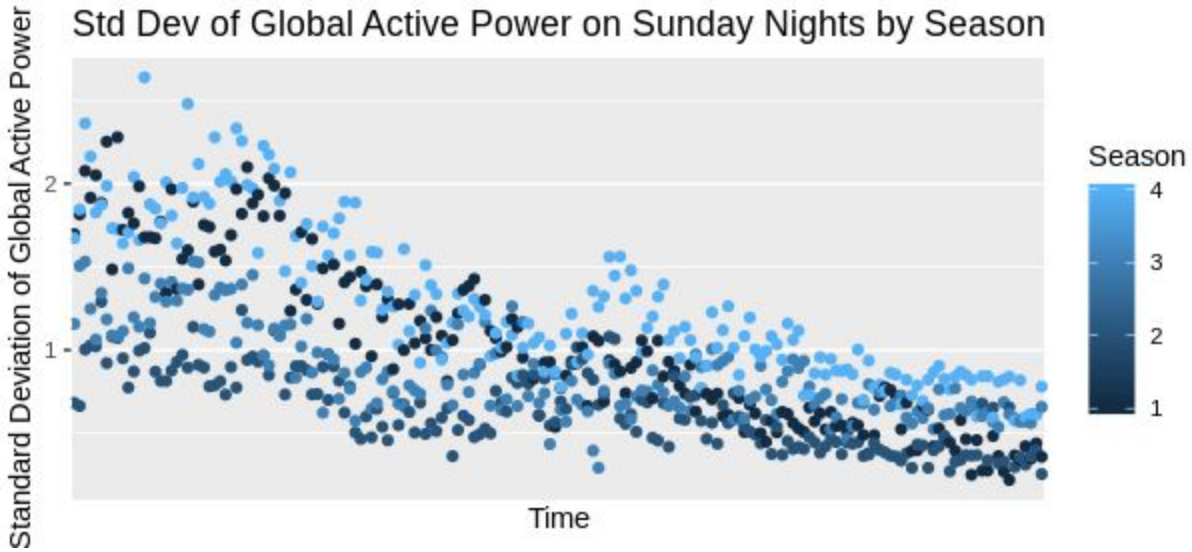

Average Global Active Power on Sunday Nights by Season
Average Global Active Power
3
2
1
Time
Season
4
3
2
1


Average Global Active Power on Sunday Mornings by Season
Average Global Active Power
2.5
2.0
1.5
1.0
0.5
08:00:00
09:00:00
10:00:00
Time
Season
4
3
2
1


Std Dev for Global Active Power on Sunday Mornings by Month
Standard Deviation of Global Active Power
0
1
2
08:00:00
09:00:00
10:00:00
Time
Month
01
02
03
04
05
06
07
08
09
10
11


Std Dev for Global Active Power on Sunday Nights by Month
Standard Deviation of Global Active Power
0
1
2
3
Time
Month
01
02
03
04
05
06
07
08
09
10
11


Std Dev for Global Active Power on Sunday Mornings by Season
Average Global Active Power
1.5
1.0
0.5
08:00:00
09:00:00
10:00:00
Time
Season
4
3
2
1


Std Dev of Global Active Power on Sunday Nights by Season
Standard Deviation of Global Active Power
2
1
Time
Season
4
3
2
1

Sources:

Get Seasons Function:

https://stackoverflow.com/questions/9500114/find-which-season-a-particular-date-belongs-to