Karim Khoja – 3013798869
Bhakti Bhanushali - 301448639

CMPT 318 ASSIGNMENT 1 REPORT

We analyzed the values from 9th April 2007 to 16th April 2007 which is the 15th week of 2007.

1. Compute the arithmetic and the geometric mean and the standard deviation for features A, B, and C respectively.

	Arithmetic mean	Geometric mean	Standard Deviation	Mode
Global Active Power	0.6448586	0.5137321	0.5227447	0.216
Global Reactive Power	0.1140778	0	0.09481887	0
Voltage	240.5041	240.4985	1.641061	241.05

For features A and B compute the min and max values on weekdays and weekend days during day hours and night hours respectively.

	Weekday day min	Weekday _day max	Weekend day min	Weekend day max
Global Active Power	0.124	2.906	0.124	2.956
Global Reactive Power	0	0.462	0	0.518

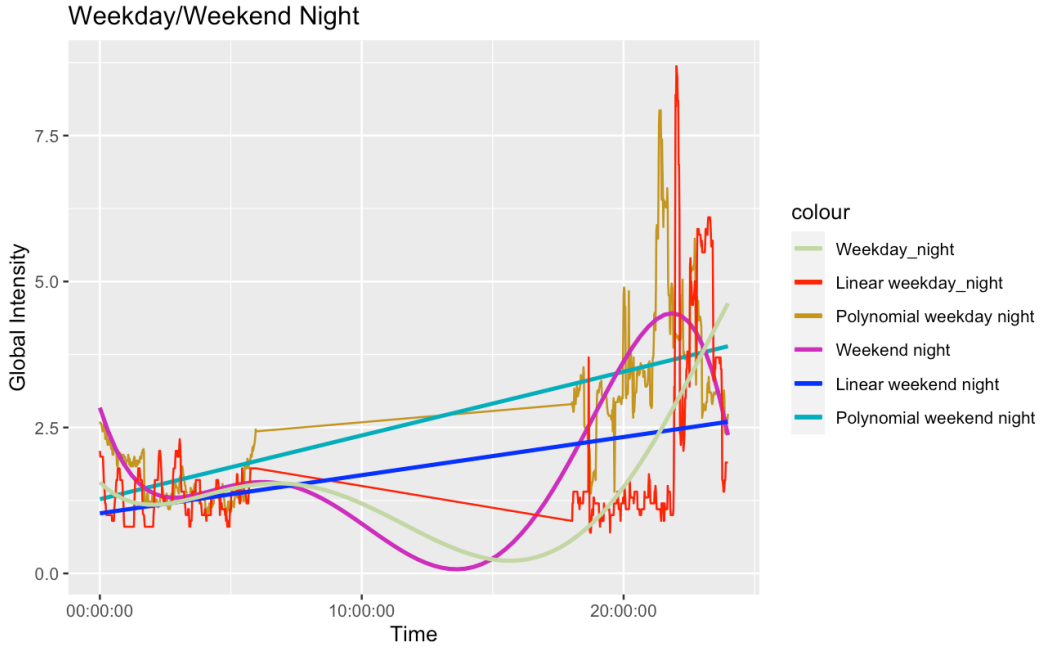
	Weekend night min	Weekend night max	Weekend night min	Weekend night max
Global Active Power	0.126	5.2	0.122	0
Global Reactive Power	0	0.548	3.906	0.36

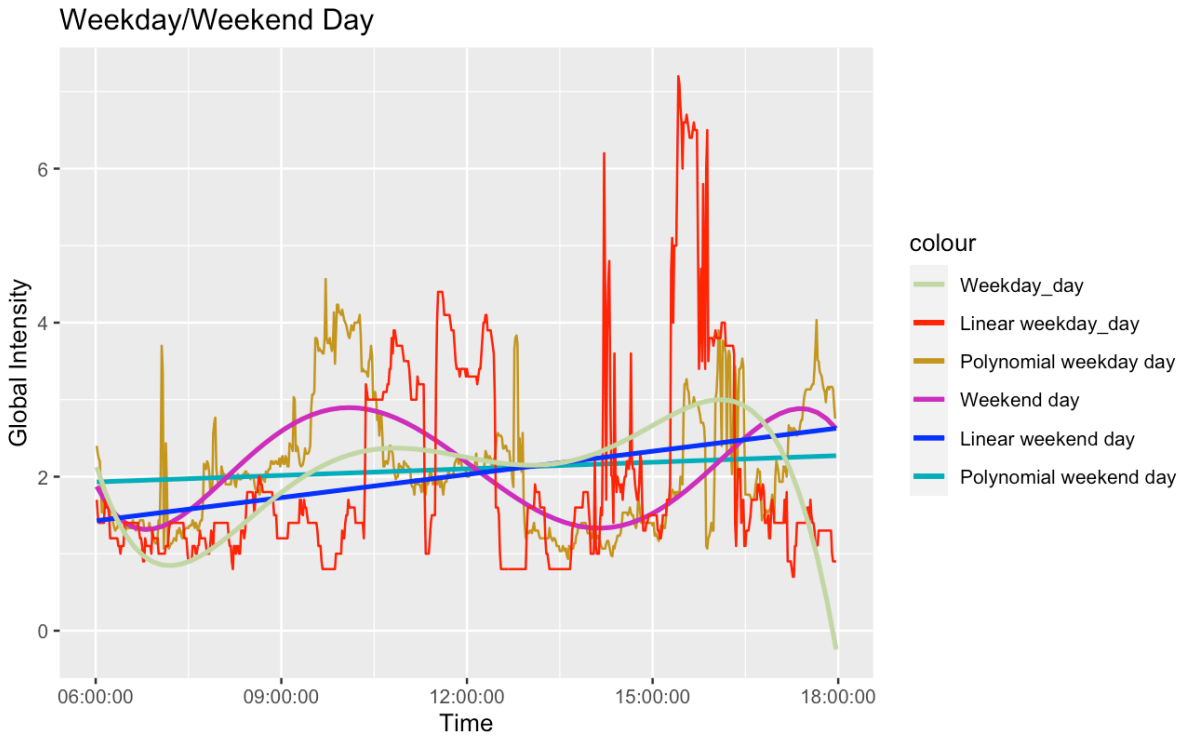
Compute the correlation for each disjoint pair of the responses, A, B, C, D, E, F and G, using Pearson's sample correlation coefficient as defined below. Represent the results of the correlation analysis in terms of a correlation matrix and visualize the relevant part of $_2$ the matrix using color-coding to show statistical significance.

	A	B	C	D	E	F	G
A							
B	0.1436						
C	-0.2474	-0.01684					
D	0.85177	0.24613	-0.313056				
E	0.31903	0.05752	-0.139875	0.43303			
F	0.23672	0.22222	-0.082509	0.35839	0.003689		
G	0.66253	0.02565	-0.209755	0.70150	0.031132	0.04876	

Legend:
A. Global active power
B. Global reactive power
C. Voltage
D. Global intensity
E. Sub_metering_1
F. Sub_metering_2
G. Sub_metering_3

Perform a linear regression based on the *least squares method* (LSM) and polynomial regression for each of the resulting four time windows and represent the results (four linear regression lines, four polynomial regression curves) graphically in two diagrams as an illustration of Global_intensity behaviour.
Our time frame for day is 6:00 AM to 6:00 PM and that for night time is 6:00 PM to 6:00 AM





We chose our polynomial to have a degree of 5 because that was the closest approximation of the actual trend.