



POWER CONSUMPTION

by GROUP5 —

MEMBERS

Dechatorn	Bumrungchoo	63070501207
Siravit	Tachaprasopchai	63070501220
Chansinee	Mueangnu	63070501221

ABOUT DATASET

“ This dataset is related to the power consumption of three different distribution networks of Tetouan city which is located in North Morocco. ”

ABOUT DATASET

	Datetime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.96386
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.08434
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.43373
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.27711
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.40964
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.31212
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.81152
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.48259
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.60504
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.49820

52416 rows × 9 columns

ABOUT DATASET

	DateTime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.96386
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.08434
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.43373
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.27711
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.40964
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.31212
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.81152
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.48259
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.60504
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.49820

52416 rows × 9 columns

**data is stored every 10 minutes
in 1 year (2017-01-01: 00:00:00
to 2017-12-31: 23:50:00.)**

ABOUT DATASET

	Datetime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.96386
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.08434
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.43373
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.27711
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.40964
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.31212
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.81152
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.48259
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.60504
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.49820

52416 rows × 9 columns

power consumption every 10 minutes

ABOUT DATASET

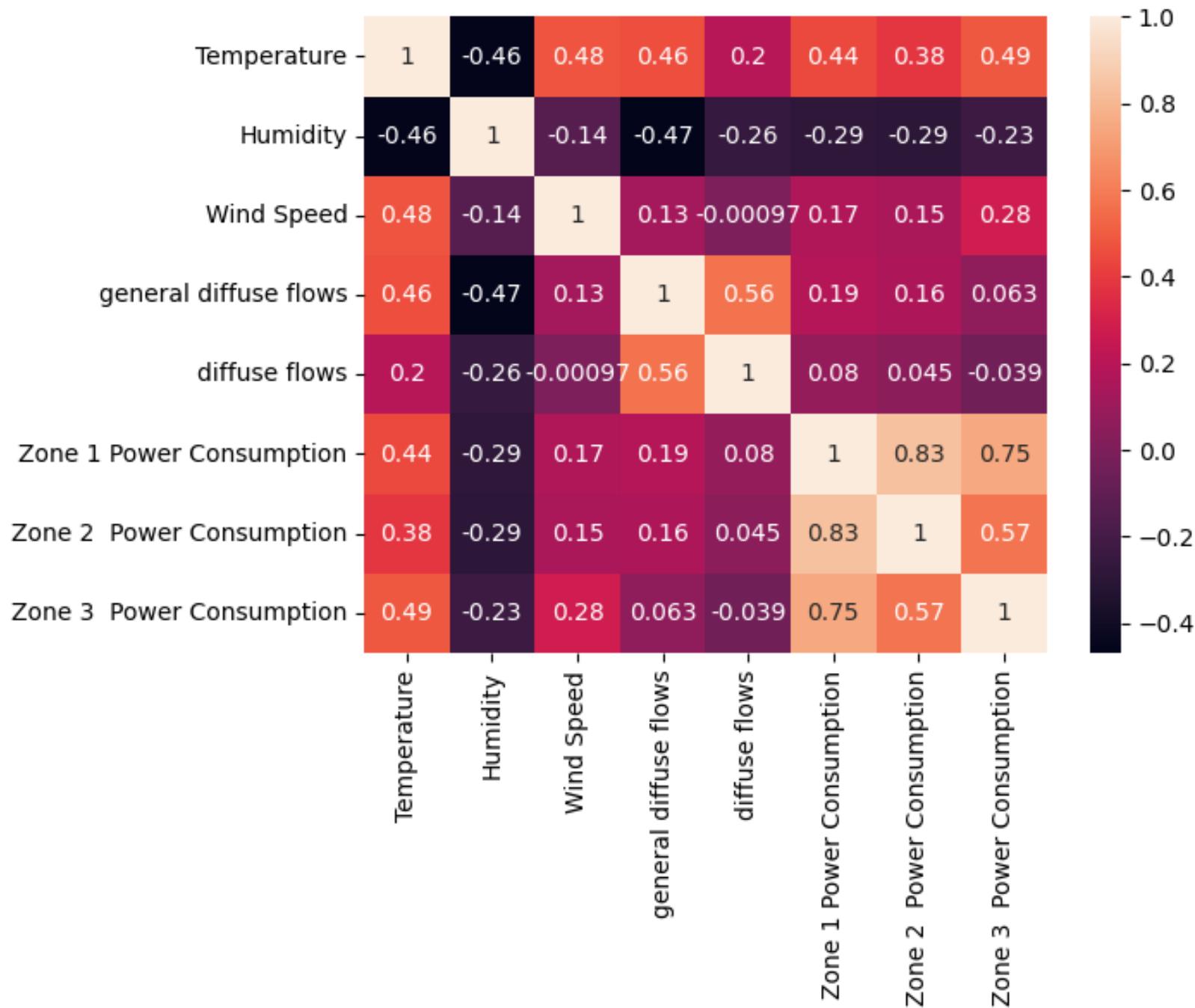
	Datetime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.96386
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.08434
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.43373
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.27711
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.40964
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.31212
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.81152
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.48259
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.60504
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.49820

52416 rows × 9 columns

other factors affecting the power consumption

EDA:

Fig1. the correlation of attributes in dataset



Label

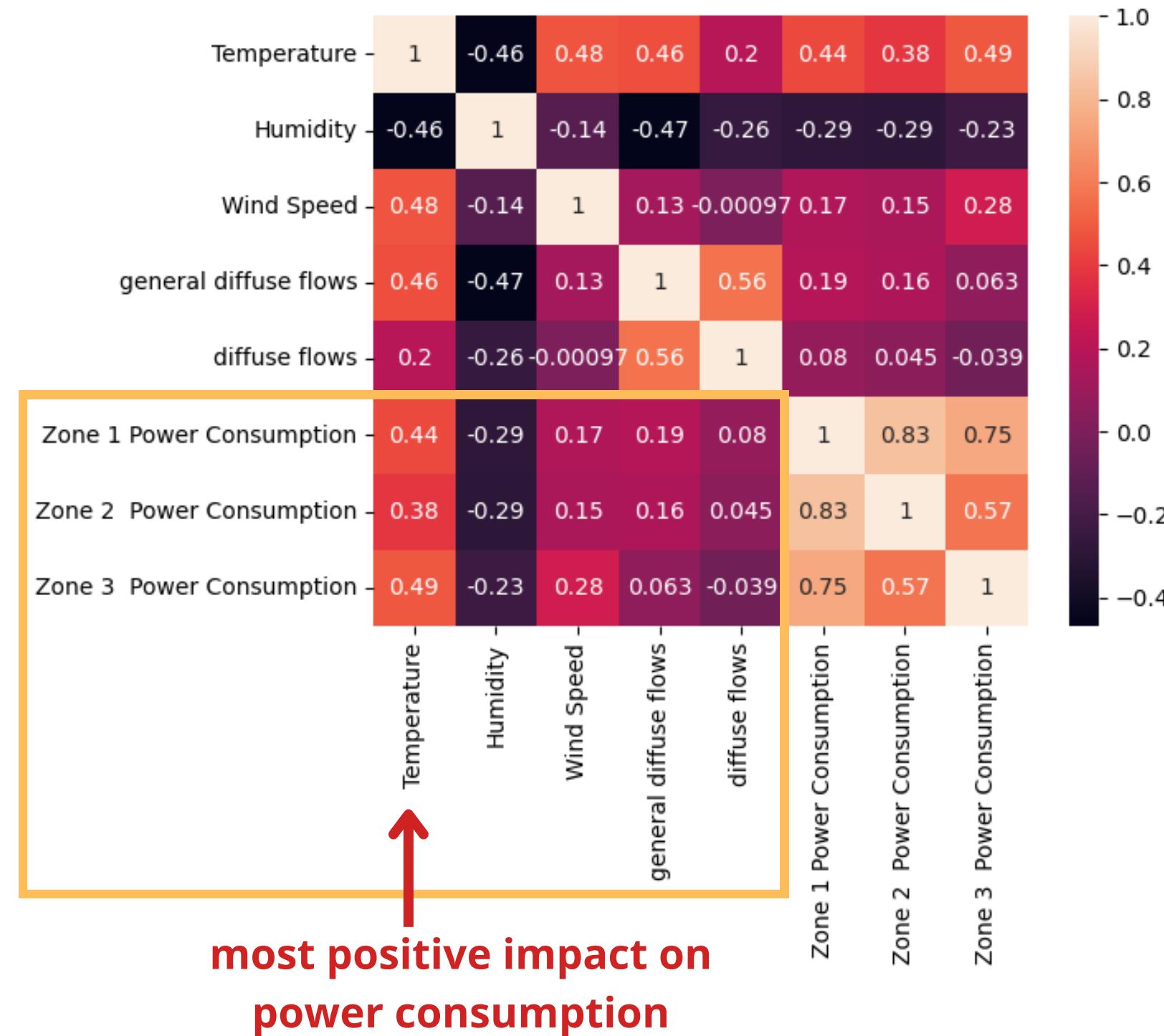
- Zone1 Power Consumption
- Zone2 Power Consumption
- Zone3 Power Consumption

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows

EDA:

Fig1. the correlation of attributes in dataset



Label

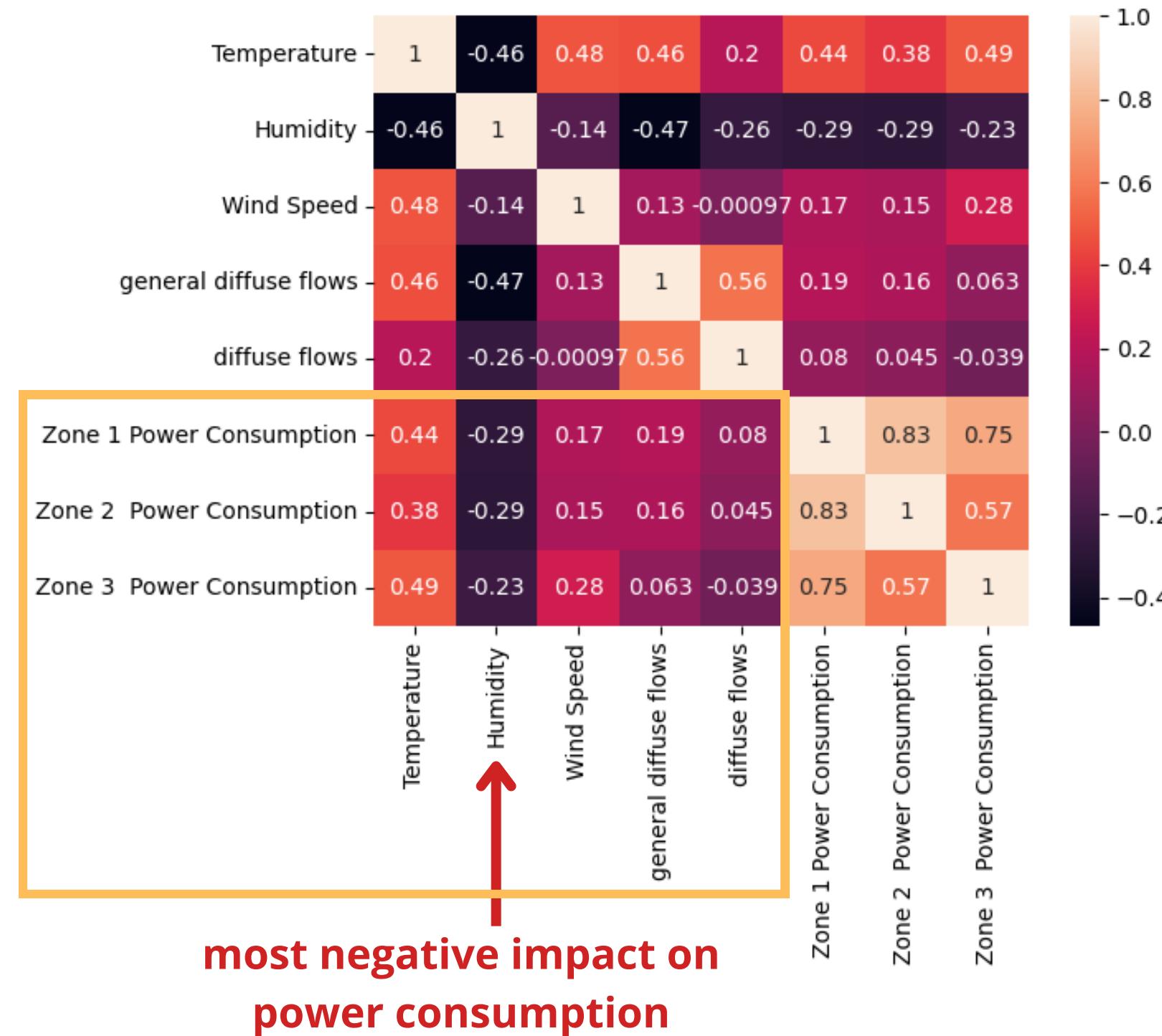
- Zone1 Power Consumption
- Zone2 Power Consumption
- Zone3 Power Consumption

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows

EDA:

Fig1. the correlation of attributes in dataset



Label

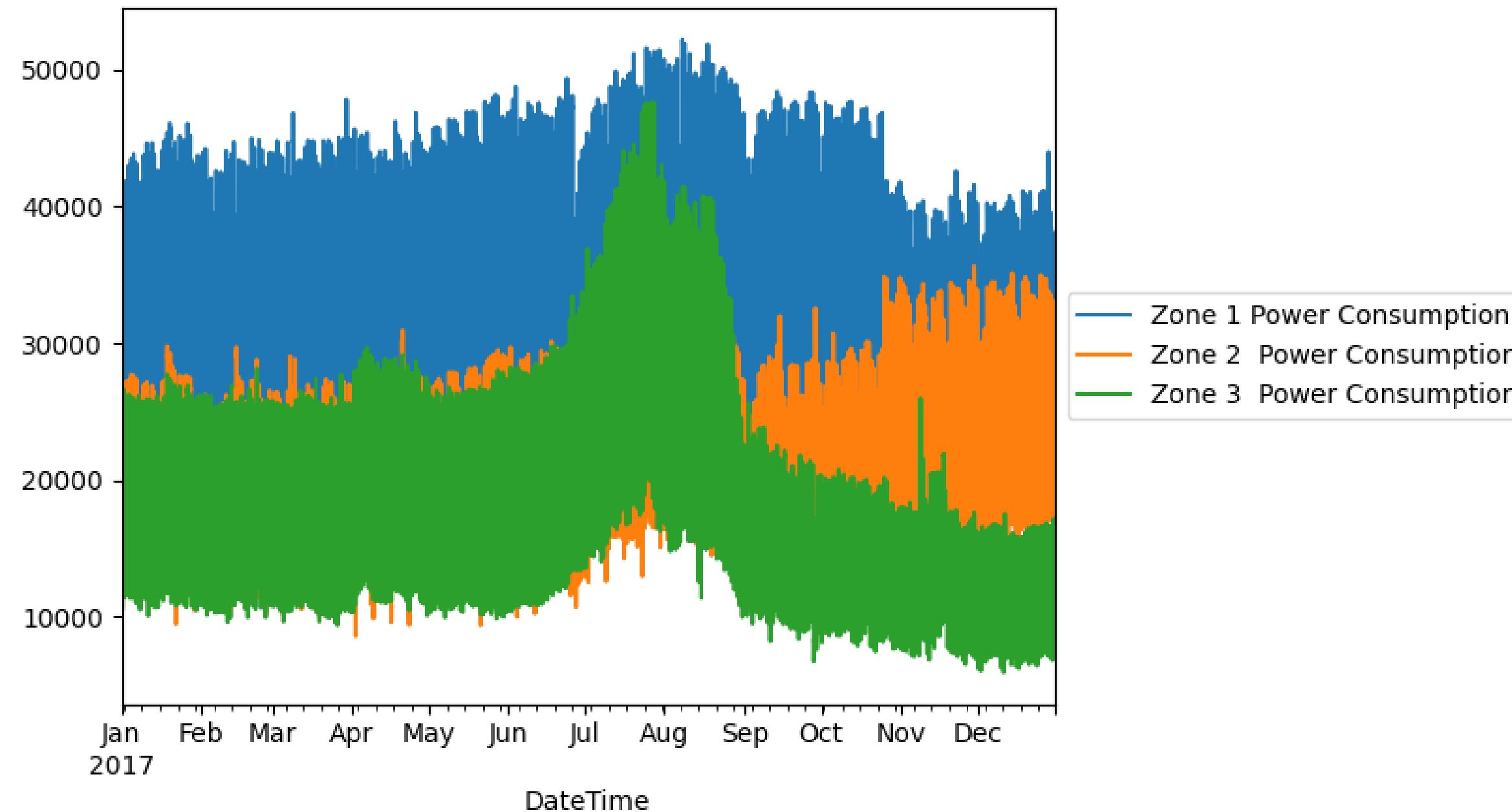
- Zone1 Power Consumption
- Zone2 Power Consumption
- Zone3 Power Consumption

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows

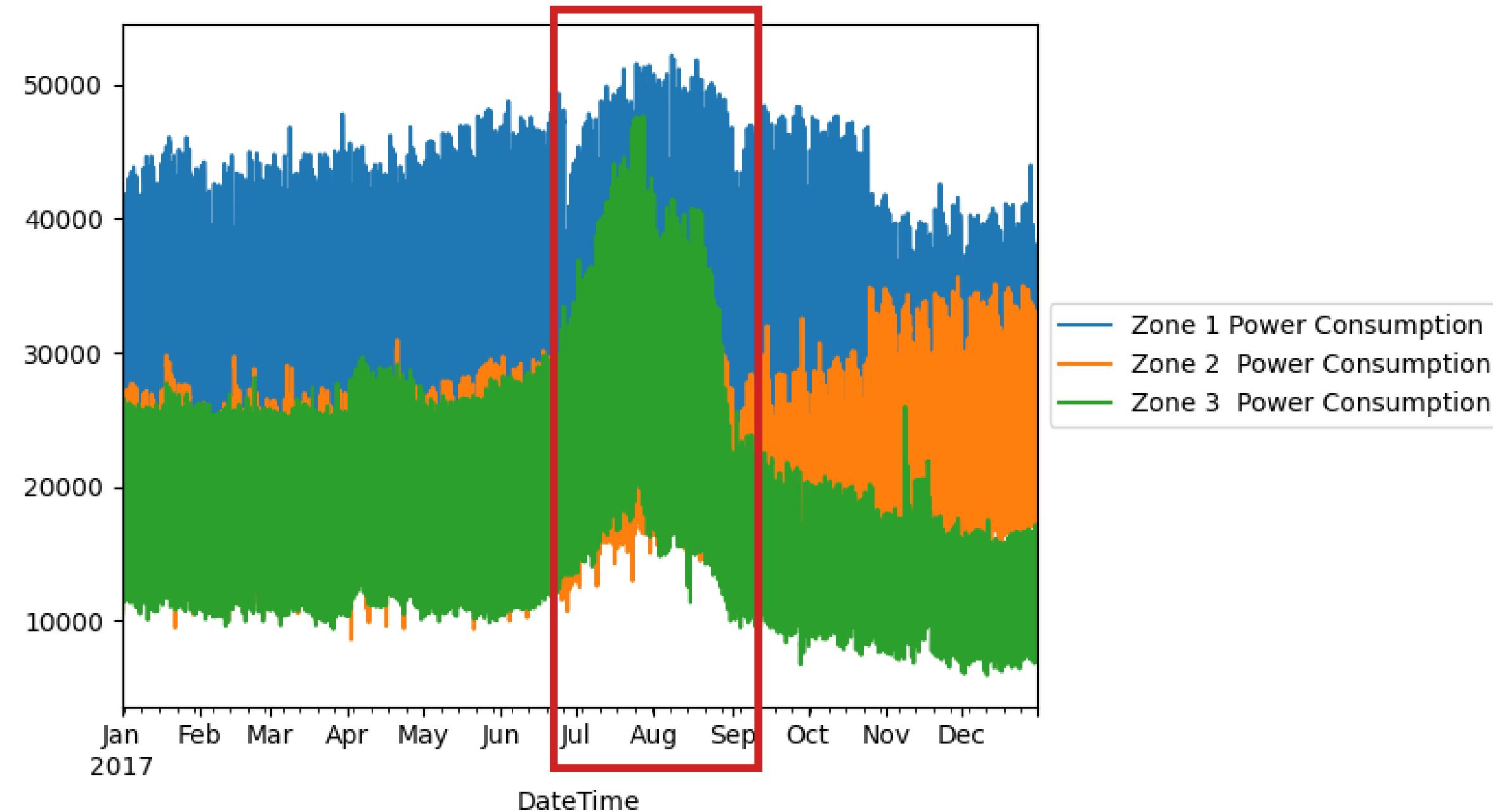
EDA:

Fig2. power consumption for the year of 2017 at each hour for 3 zones



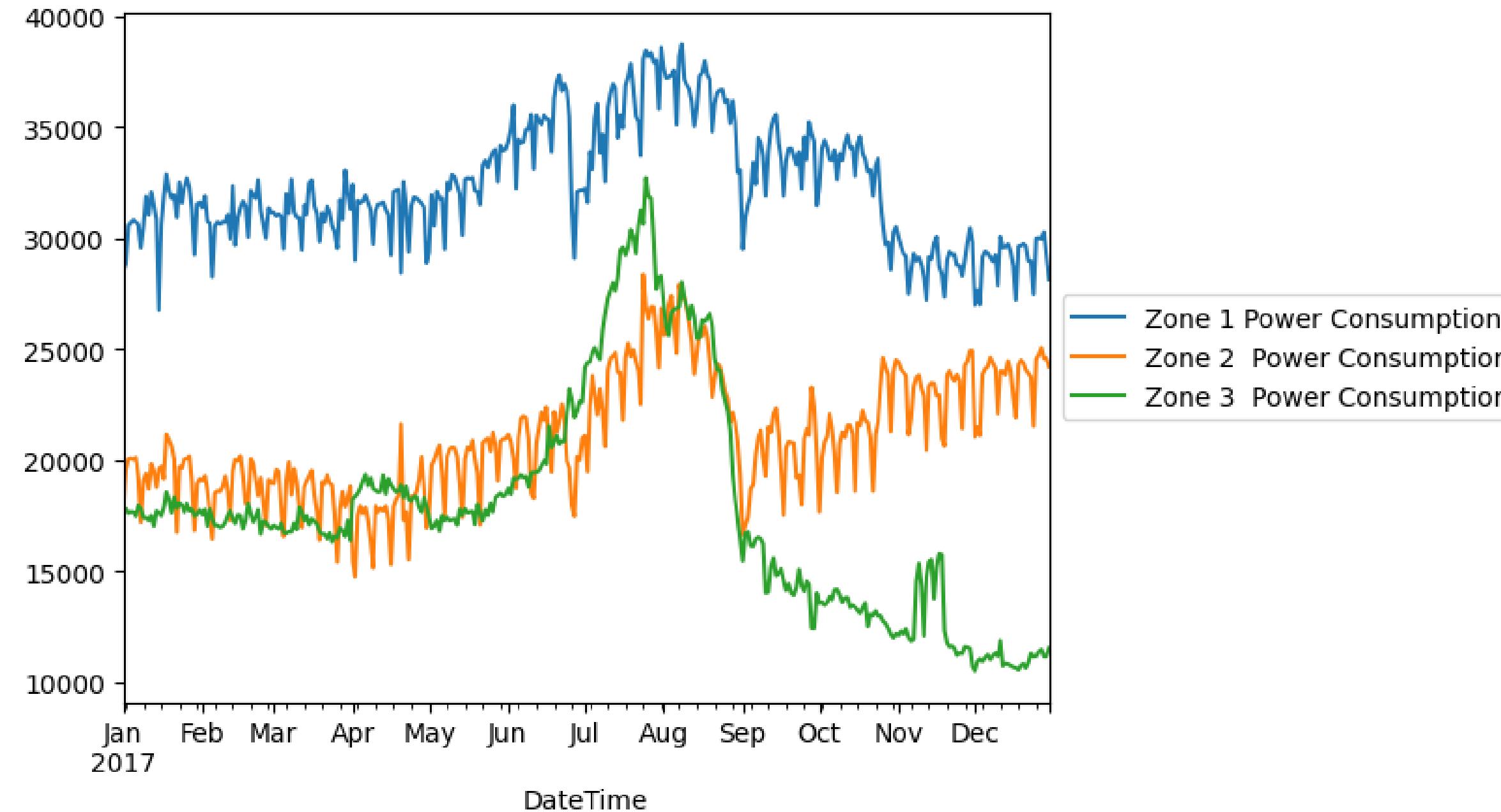
EDA:

Fig2. power consumption for the year of 2017 at each hour for 3 zones



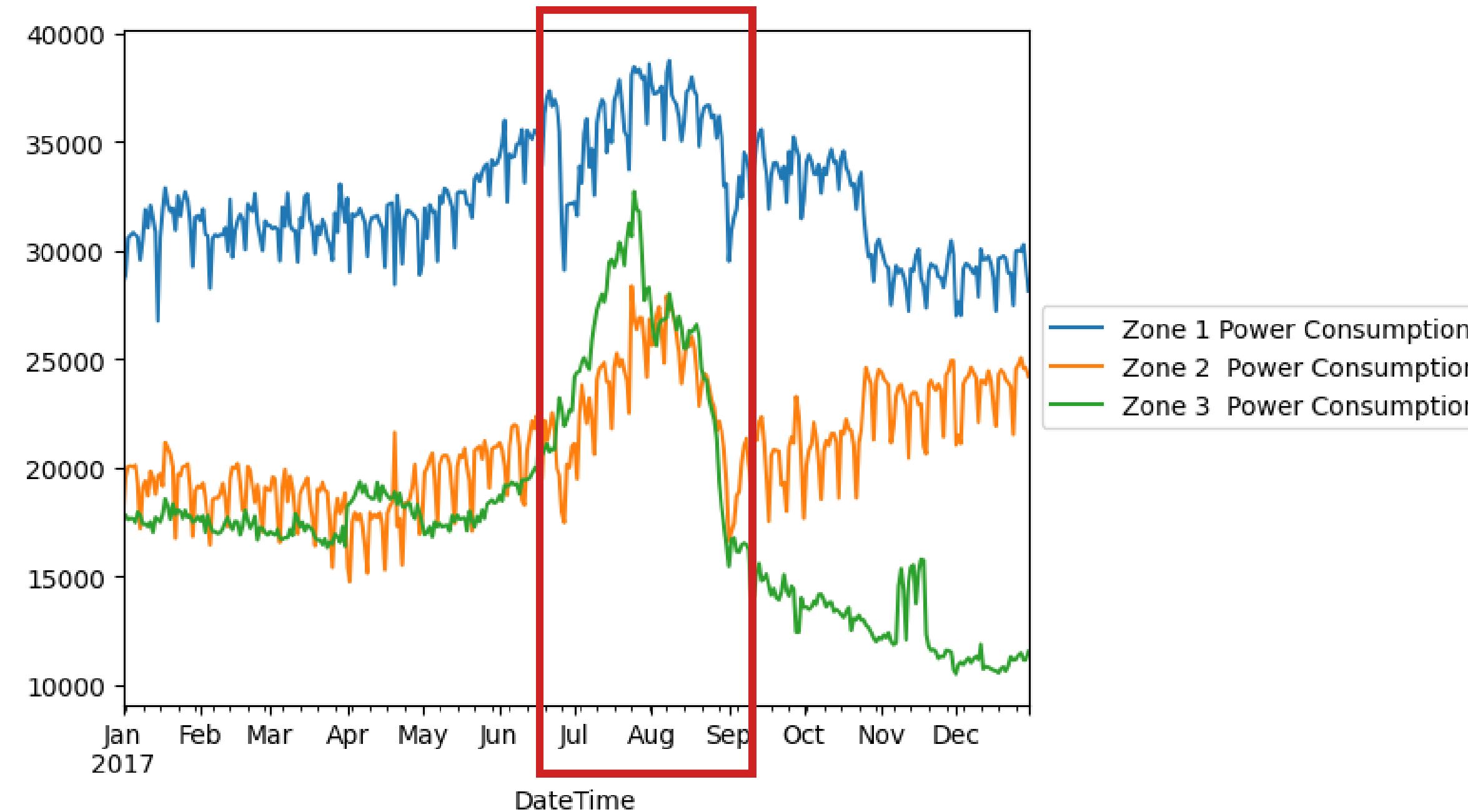
EDA:

Fig3. average daily power consumption at each hour for 3 zones



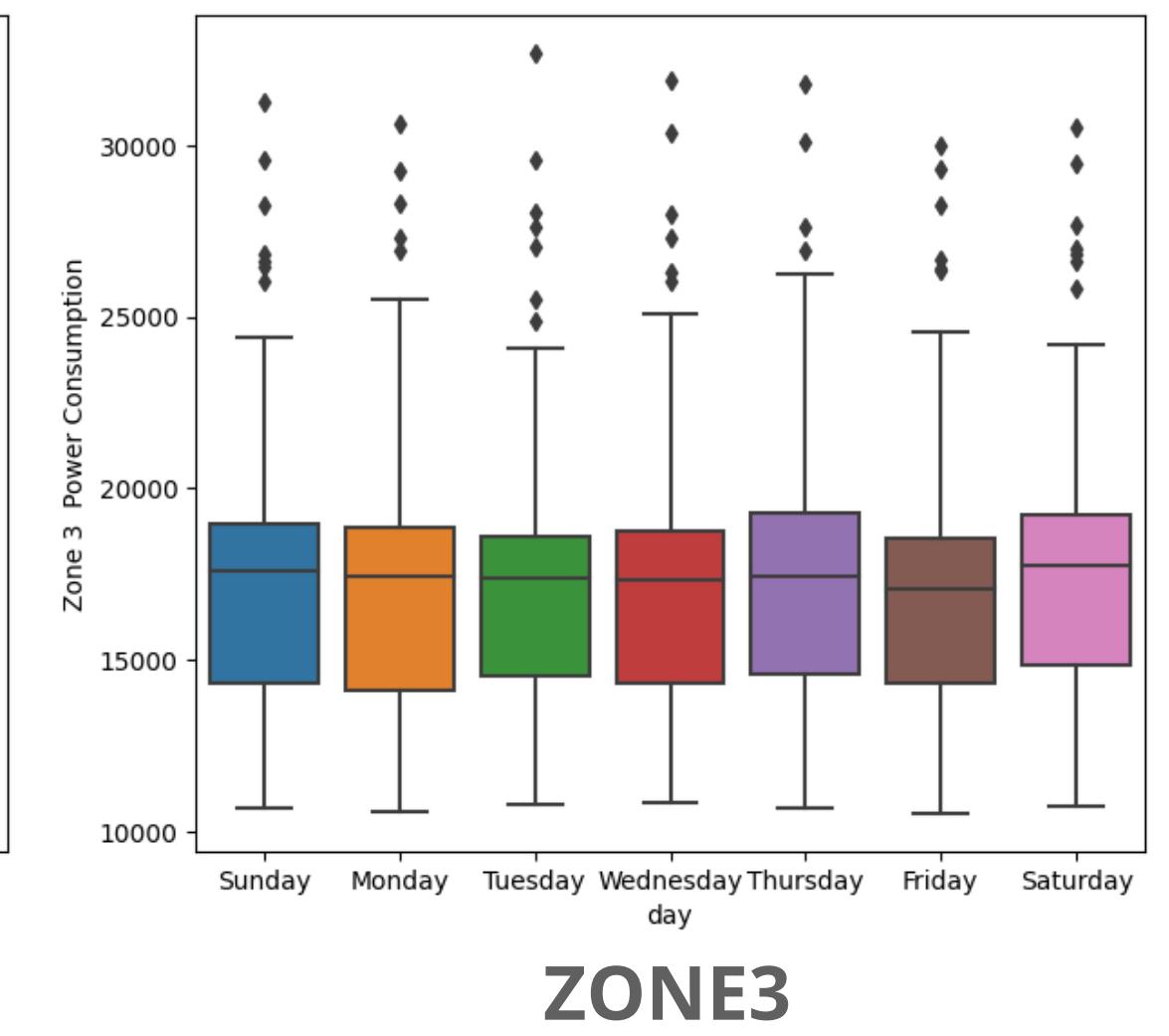
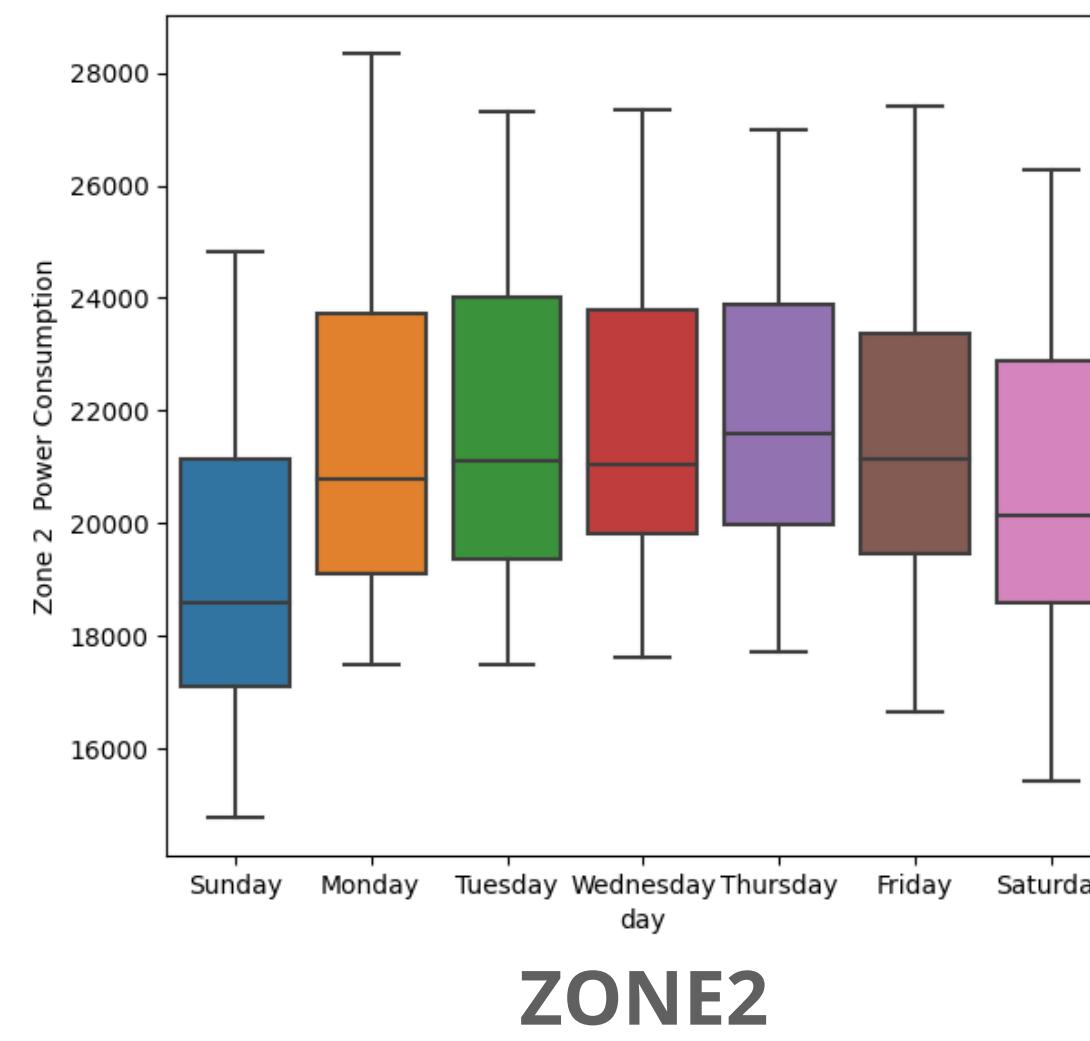
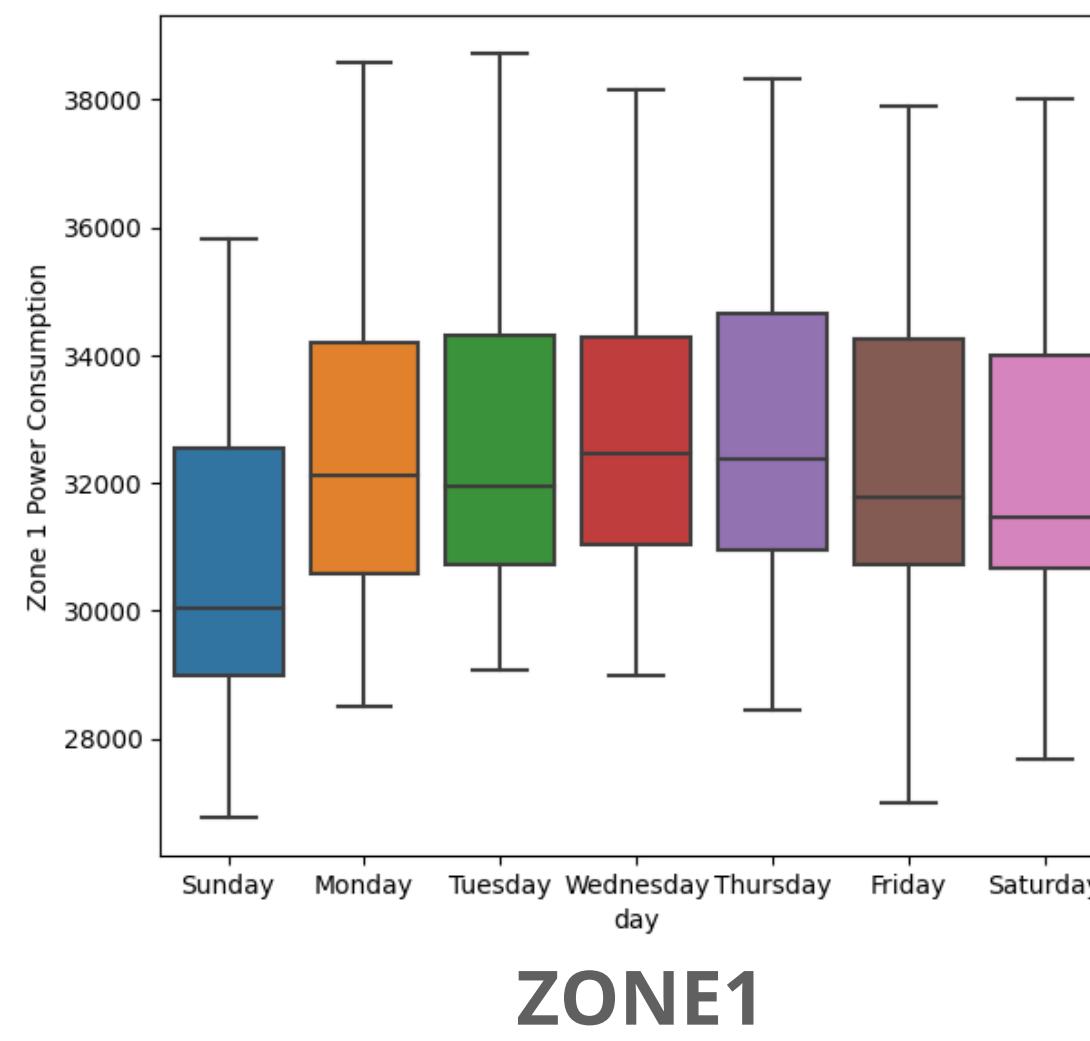
EDA:

Fig3. average daily power consumption at each hour for 3 zones



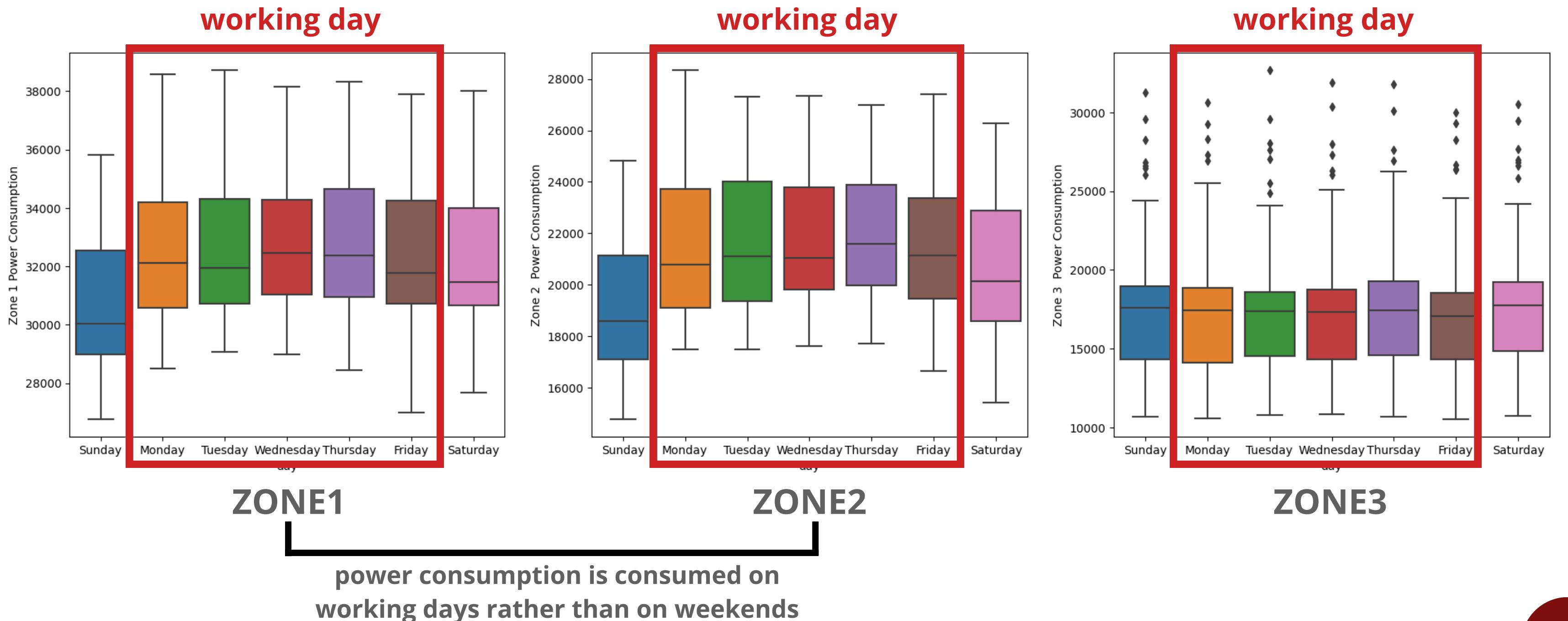
EDA:

Fig4. average power consumption per week in 1 year at each hour for 3 zones



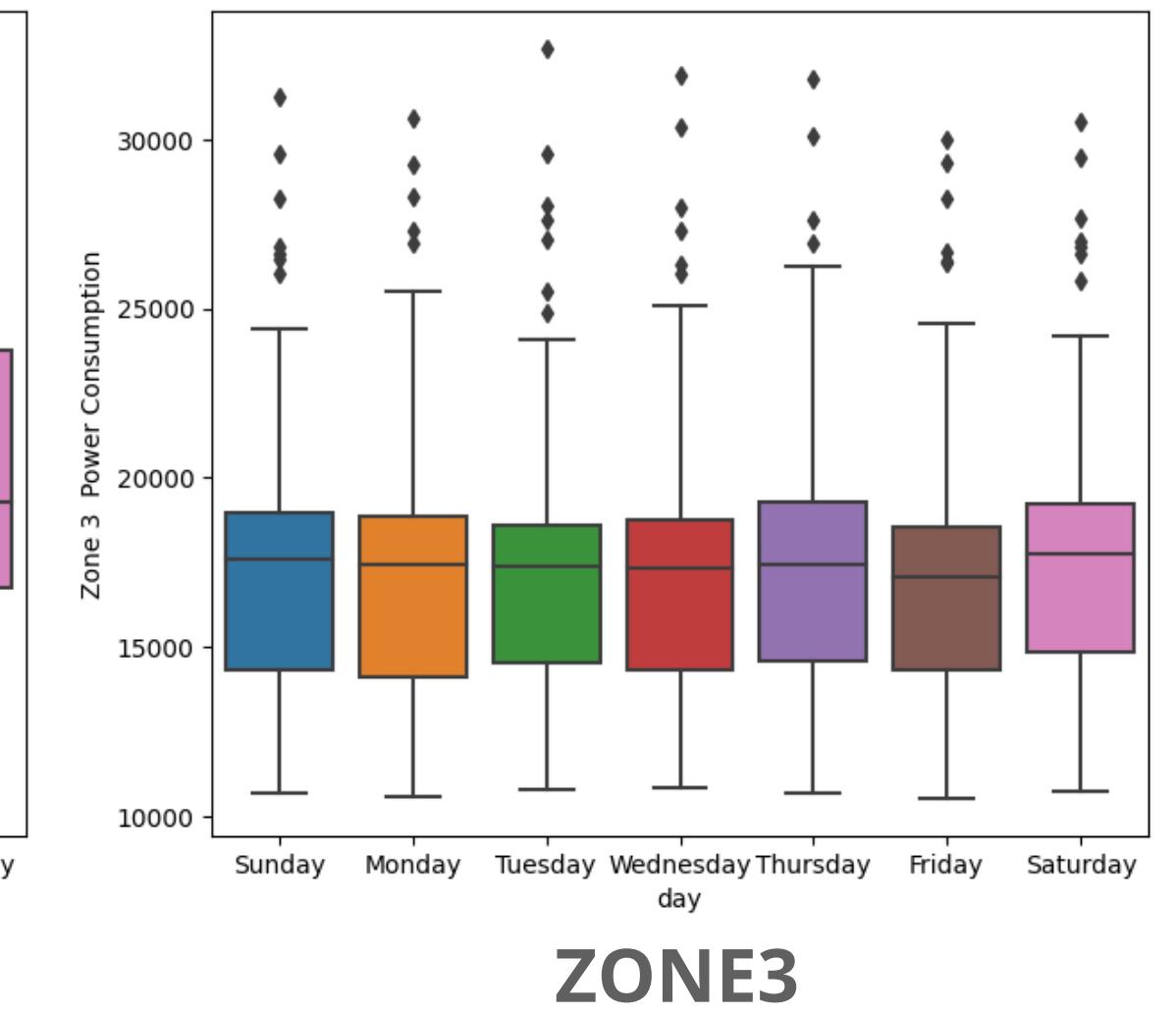
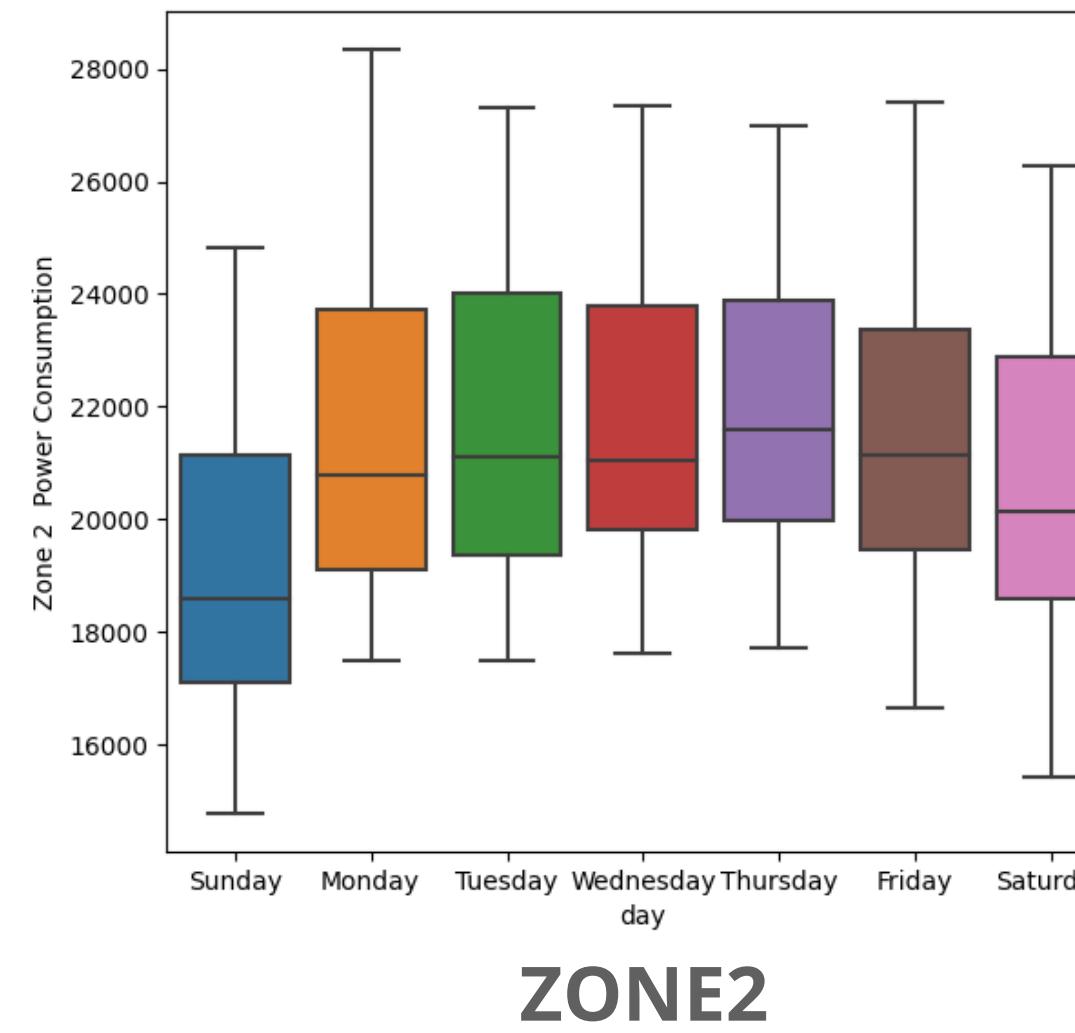
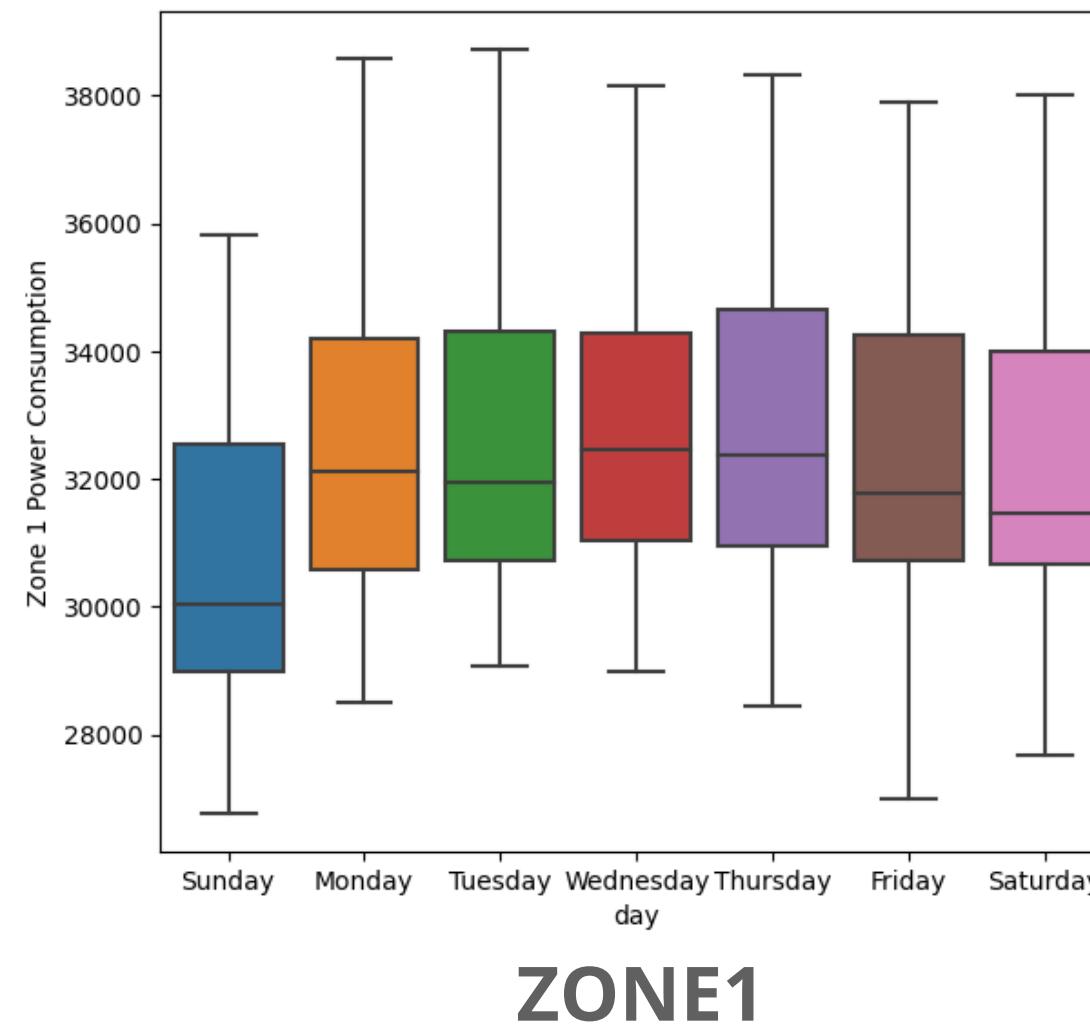
EDA:

Fig4. average power consumption per week in 1 year at each hour for 3 zones



EDA:

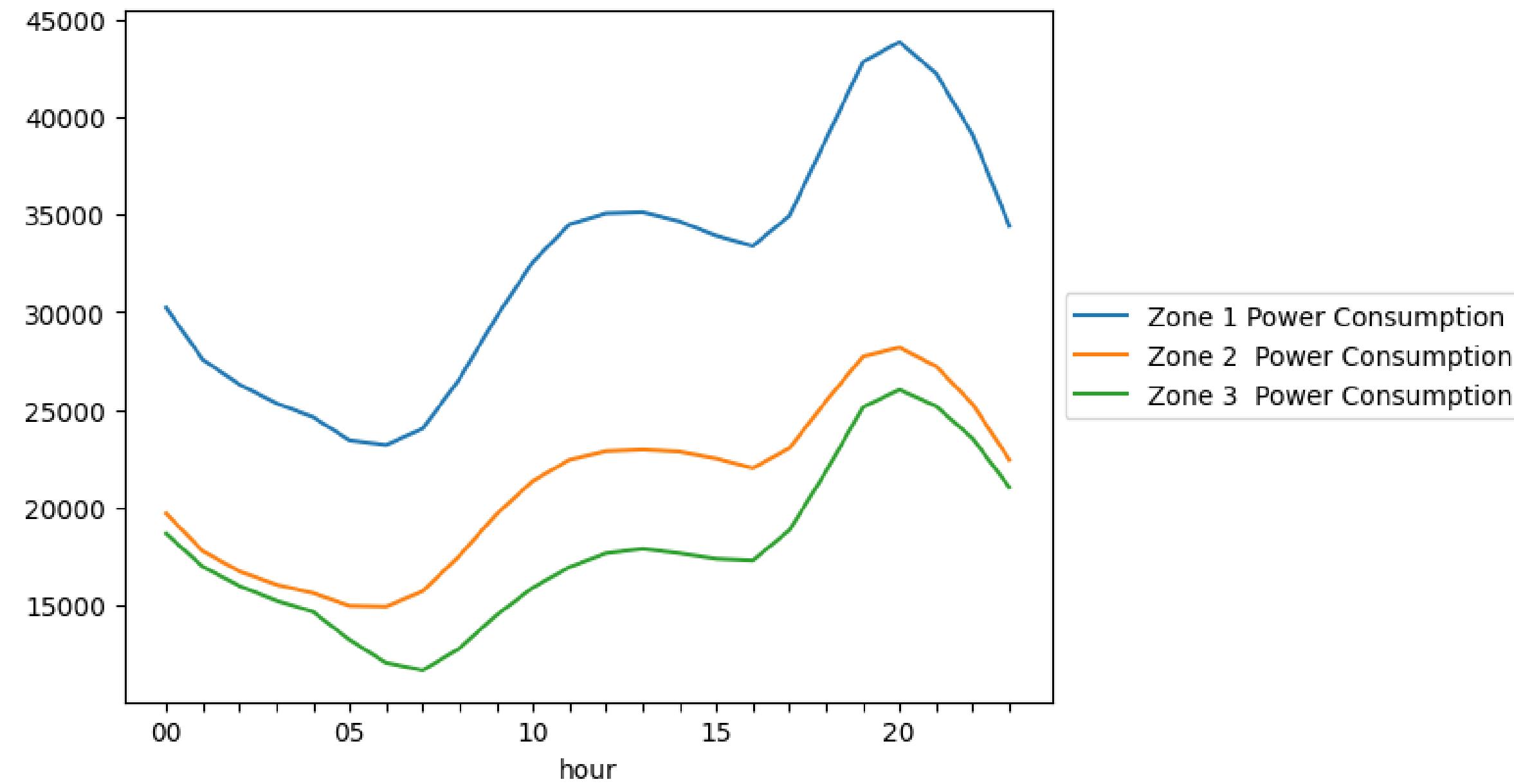
Fig4. average power consumption per week in 1 year at each hour for 3 zones



ASSUME: day of week (Sun-Sat) should affect power consumption

EDA:

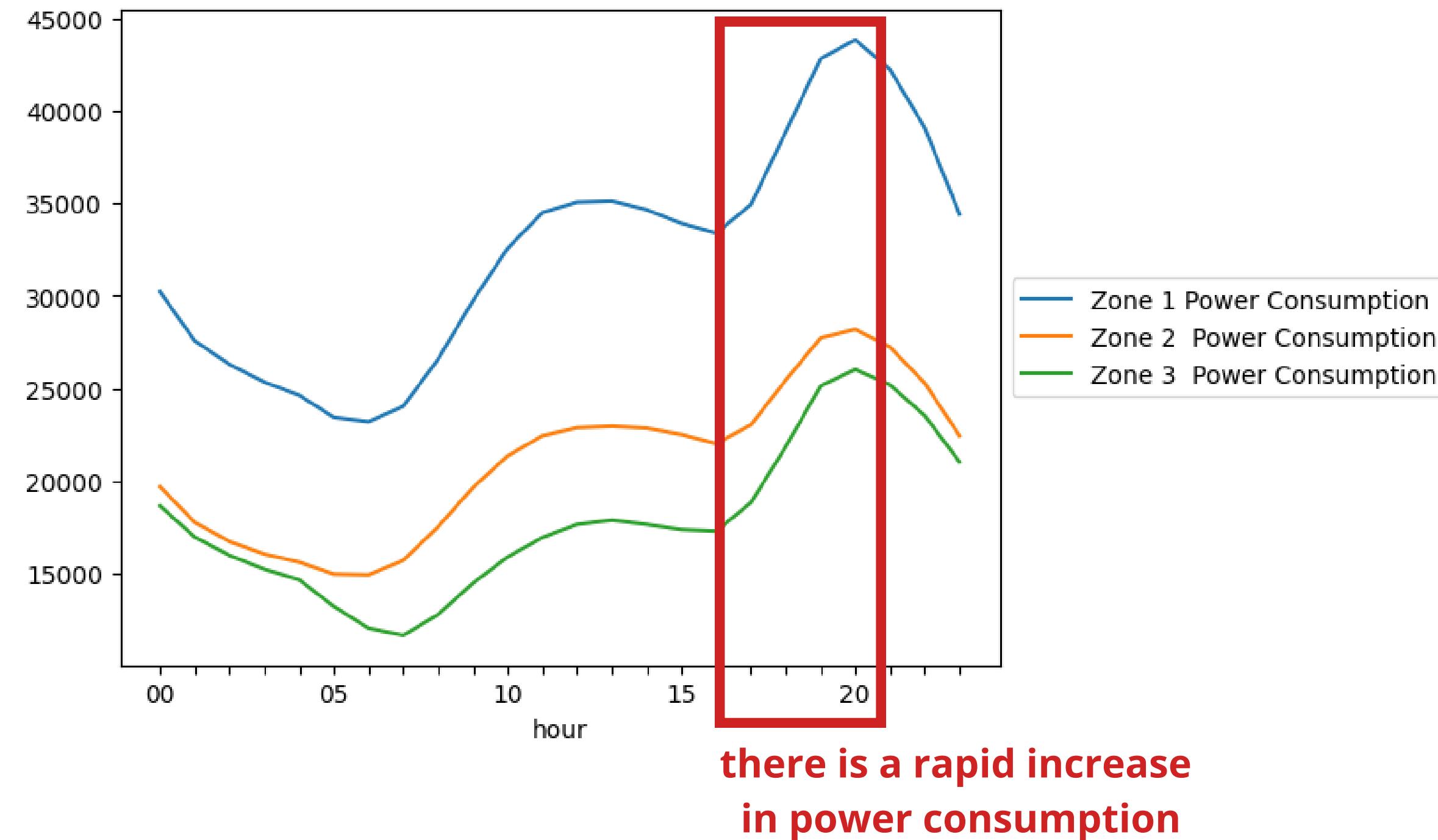
Fig5. average hourly power consumption per day in 1 year at each hour for 3 zones



"there is an increase in power consumption throughout the day"

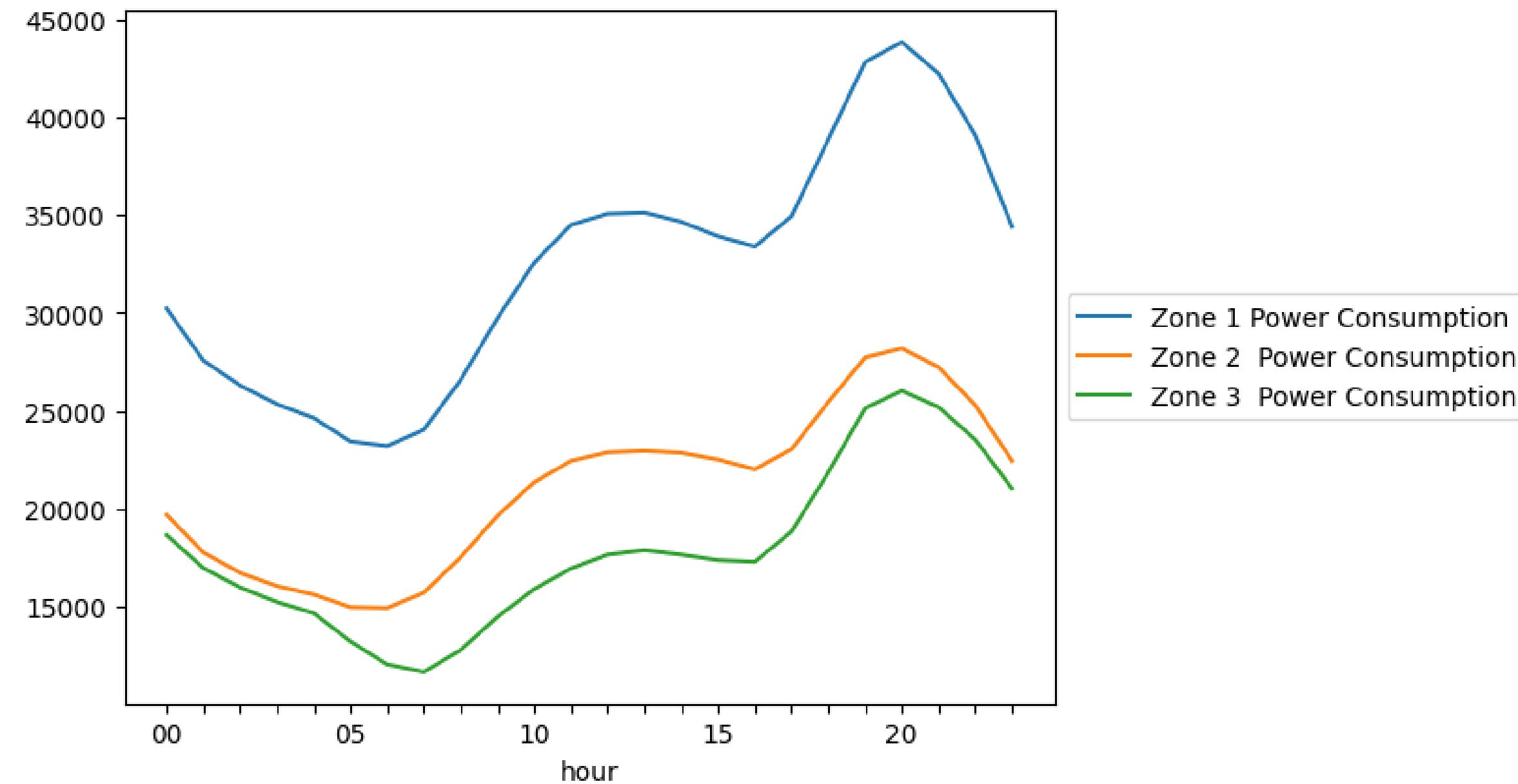
EDA:

Fig5. average hourly power consumption per day in 1 year at each hour for 3 zones



EDA:

Fig5. average hourly power consumption per day in 1 year at each hour for 3 zones



ASSUME: hour (00-23) should affect power consumption

ABOUT DATASET

	Datetime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption	hour	day_of_week
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.96386	0	0
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.08434	0	0
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.43373	0	0
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.27711	0	0
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.40964	0	0
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.31212	23	6
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.81152	23	6
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.48259	23	6
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.60504	23	6
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.49820	23	6

52416 rows × 11 columns

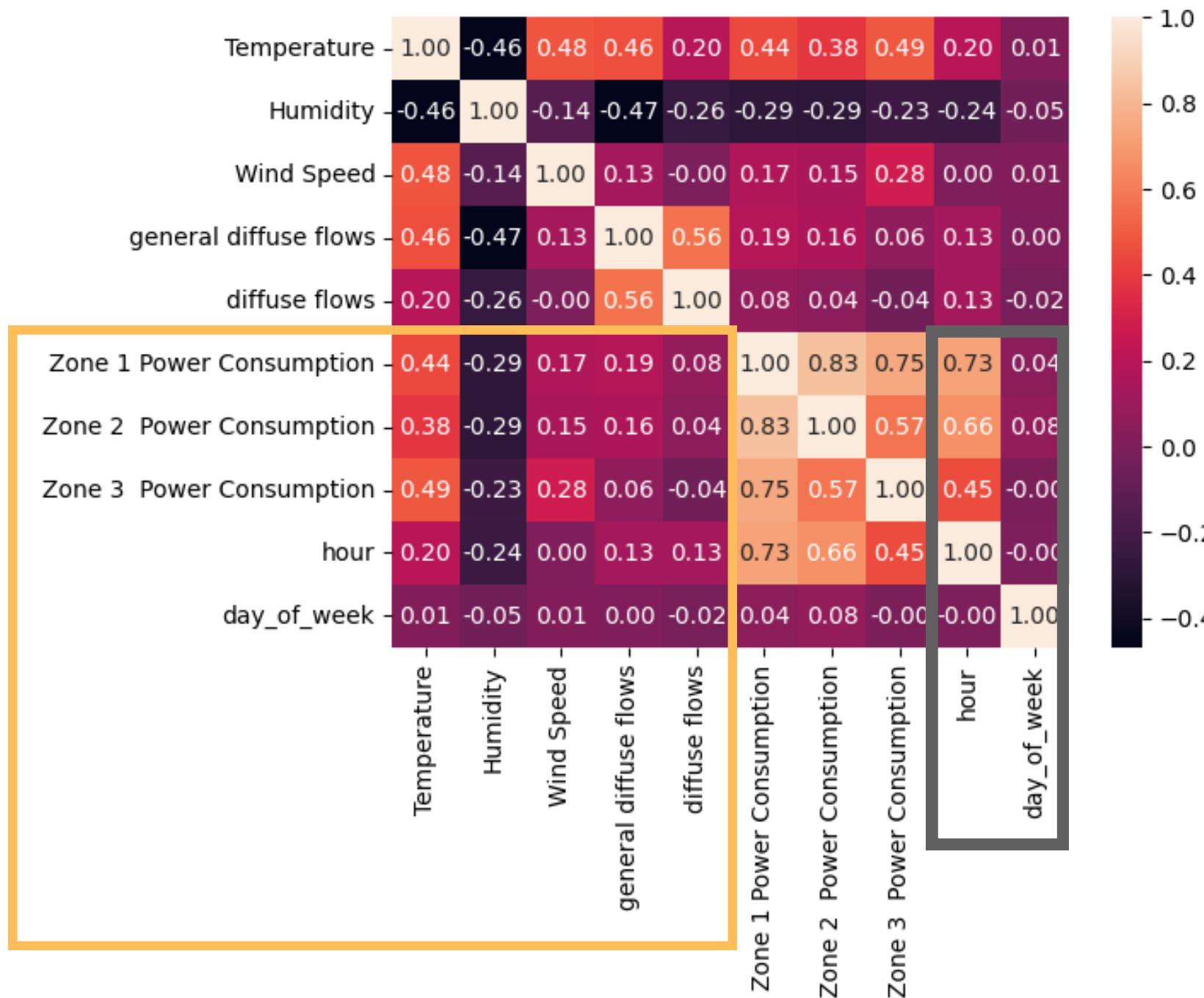
ABOUT DATASET

	Datetime	Temperature	Humidity	Wind Speed	general diffuse flows	diffuse flows	Zone 1 Power Consumption	Zone 2 Power Consumption	Zone 3 Power Consumption	hour	day_of_week
0	2017-01-01 00:00:00	6.559	73.8	0.083	0.051	0.119	34055.69620	16128.87538	20240.9638	0	0
1	2017-01-01 00:10:00	6.414	74.5	0.083	0.070	0.085	29814.68354	19375.07599	20131.0843	0	0
2	2017-01-01 00:20:00	6.313	74.5	0.080	0.062	0.100	29128.10127	19006.68693	19668.4337	0	0
3	2017-01-01 00:30:00	6.121	75.0	0.083	0.091	0.096	28228.86076	18361.09422	18899.2771	0	0
4	2017-01-01 00:40:00	5.921	75.7	0.081	0.048	0.085	27335.69620	17872.34043	18442.4096	0	0
...
52411	2017-12-30 23:10:00	7.010	72.4	0.080	0.040	0.096	31160.45627	26857.31820	14780.3121	23	6
52412	2017-12-30 23:20:00	6.947	72.6	0.082	0.051	0.093	30430.41825	26124.57809	14428.8115	23	6
52413	2017-12-30 23:30:00	6.900	72.8	0.086	0.084	0.074	29590.87452	25277.69254	13806.4825	23	6
52414	2017-12-30 23:40:00	6.758	73.0	0.080	0.066	0.089	28958.17490	24692.23688	13512.6050	23	6
52415	2017-12-30 23:50:00	6.580	74.1	0.081	0.062	0.111	28349.80989	24055.23167	13345.4982	23	6

52416 rows × 11 columns

EDA:

Fig6. the correlation of attributes in dataset



Label

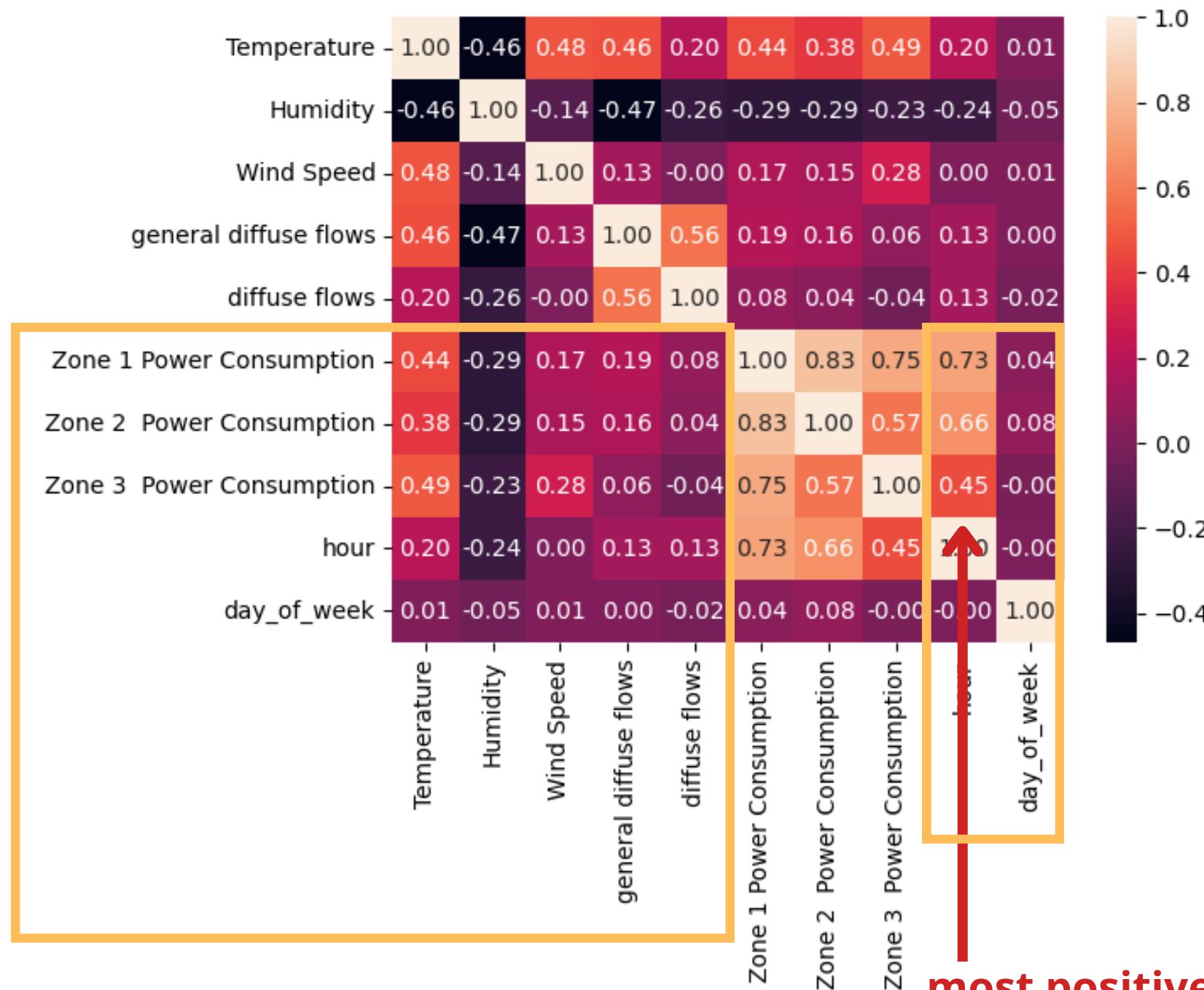
- Zone1 Power Consumption
- Zone2 Power Consumption
- Zone3 Power Consumption

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows
- Hour ←
- Day of Week ←

EDA:

Fig6. the correlation of attributes in dataset



most positive impact on
power consumption

Label

- Zone1 Power Consumption
- Zone2 Power Consumption
- Zone3 Power Consumption

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows
- Hour 
- Day of Week

ML MODEL

- Linear / Non linear Regression
- Support Vector Machine for Regression
- Decision Tree
- Random Forest

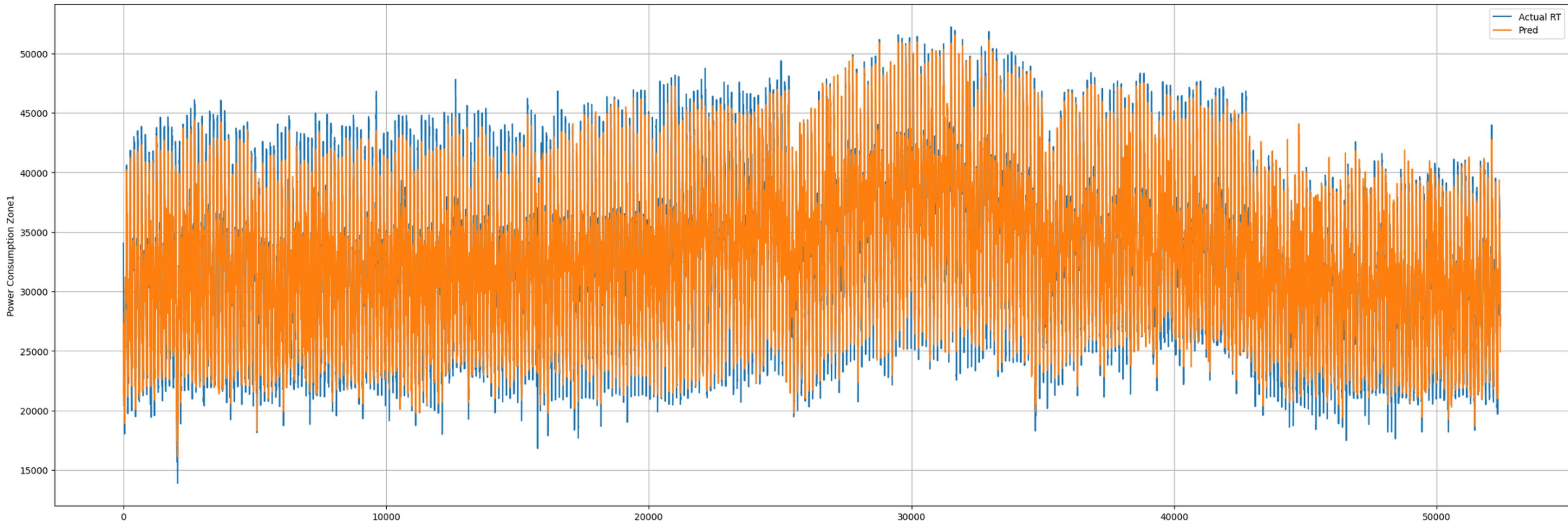
Desired:

- Able to handle complex data.
- Able to predict in case of multiple parameters.
- Have high performance than the decision tree model.
- Reduce the risk of overfitting by using a subset of features to build each decision tree.

SELECTED MODEL: RANDOM FOREST

RF MODEL1 (Whether)

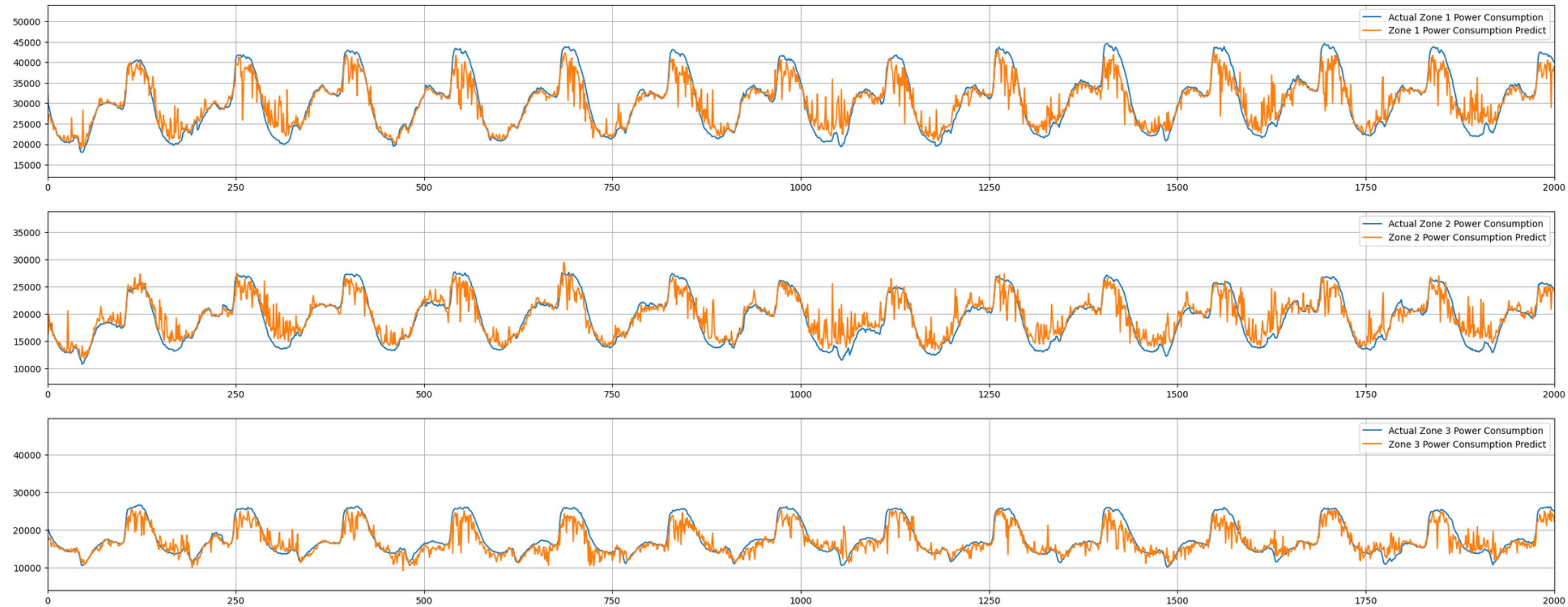
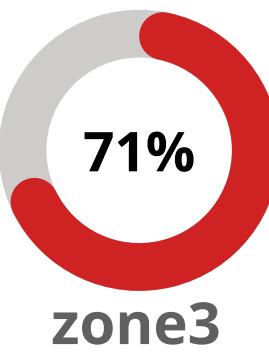
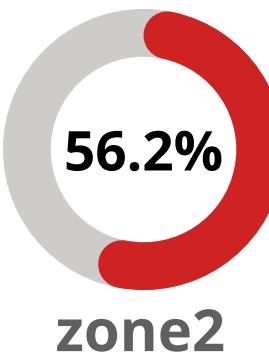
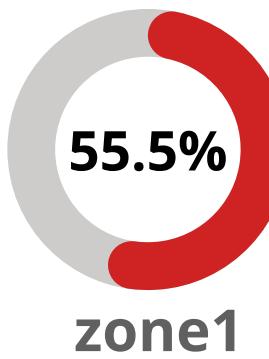
- Features**
- Temperature
 - Humidity
 - Wind Speed
 - General Diffuse Flows
 - Diffuse Flows



RF MODEL1 (Whether)

- Features
 - Temperature
 - Humidity
 - Wind Speed
 - General Diffuse Flows
 - Diffuse Flows

Evaluate: R Square

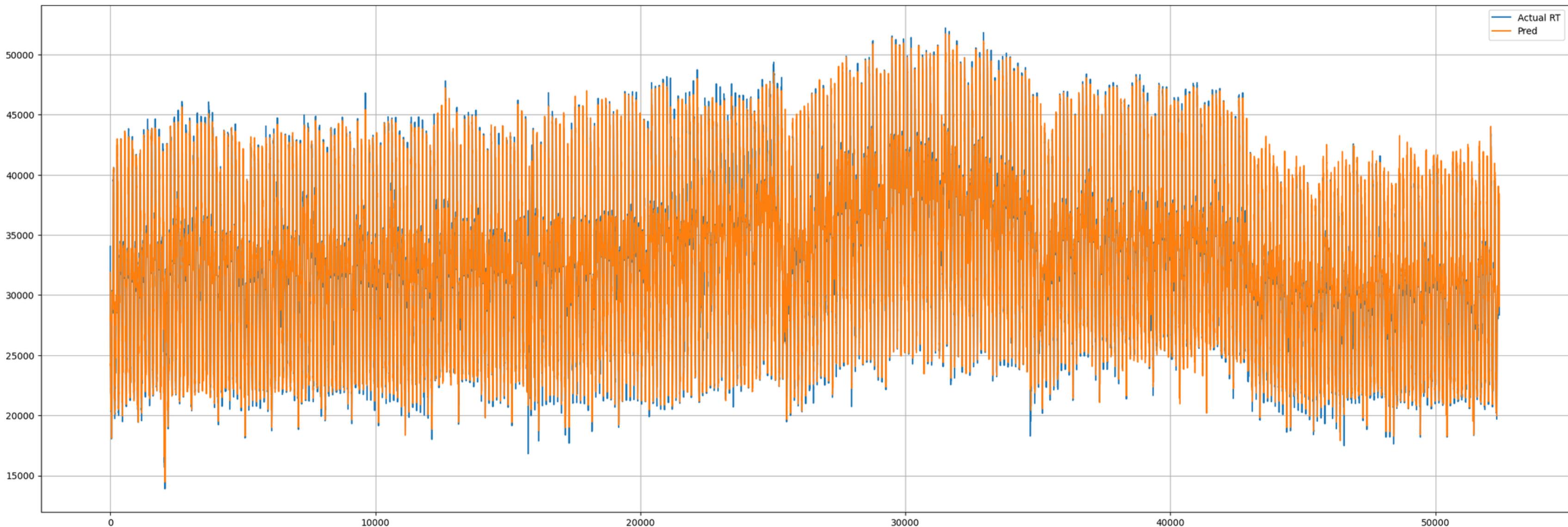


RF MODEL2

(Whether + Time)

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows
- Diffuse Flows
- Hour
- Day of Week



RF MODEL2

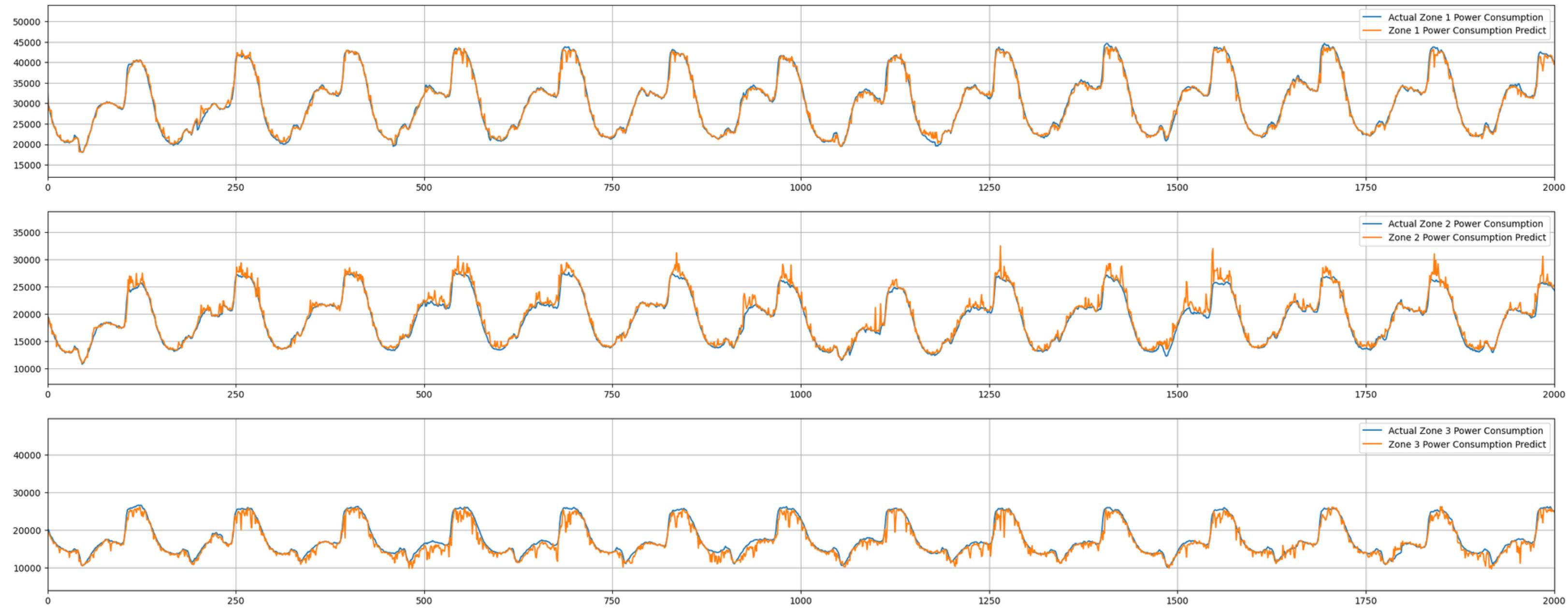
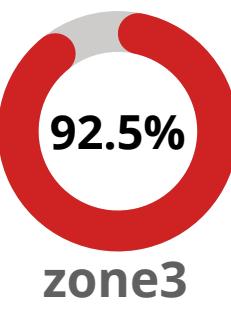
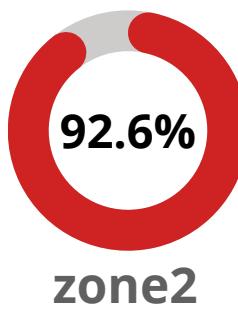
(Whether + Time)

Features

- Temperature
- Humidity
- Wind Speed
- General Diffuse Flows

- Diffuse Flows
- Hour
- Day of Week

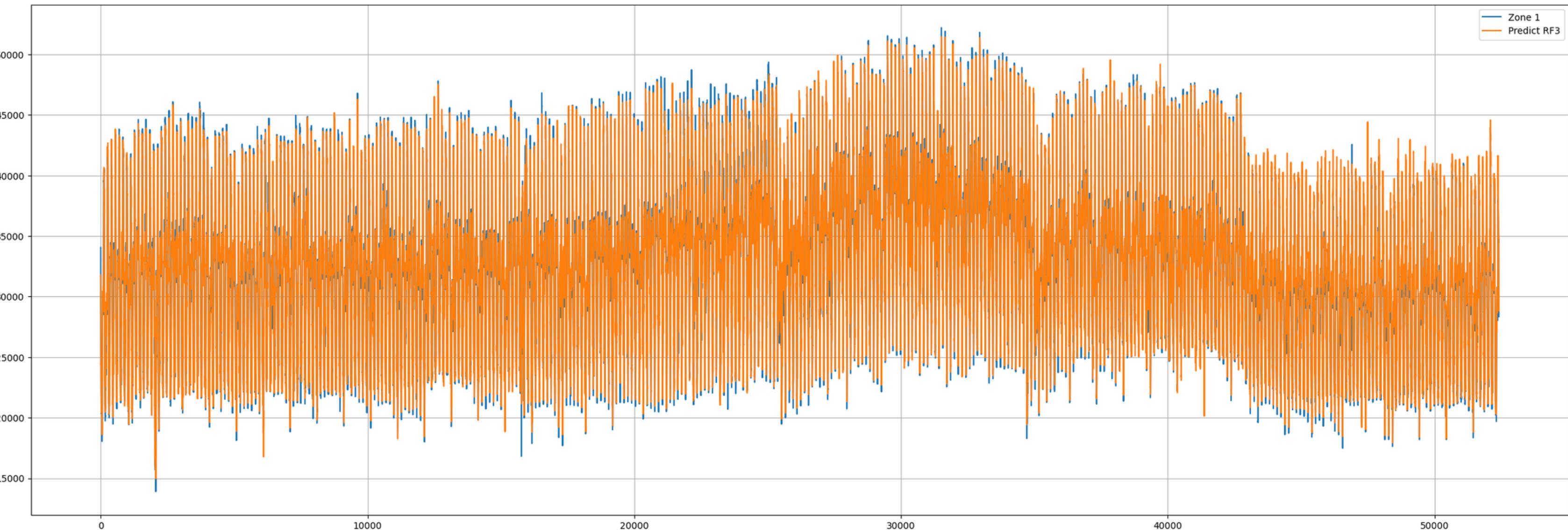
Evaluate: R Square



RF MODEL3 (PCA)

Features	Variance
1. General Diffuse Flows	69907.8674
2. Diffuse Flows	15428.3599
3. Humidity	241.8391
4. Hour	47.9175
5. Temperature	33.8198
6. Wind Speed	5.5171
7. Day of Week	4.0001

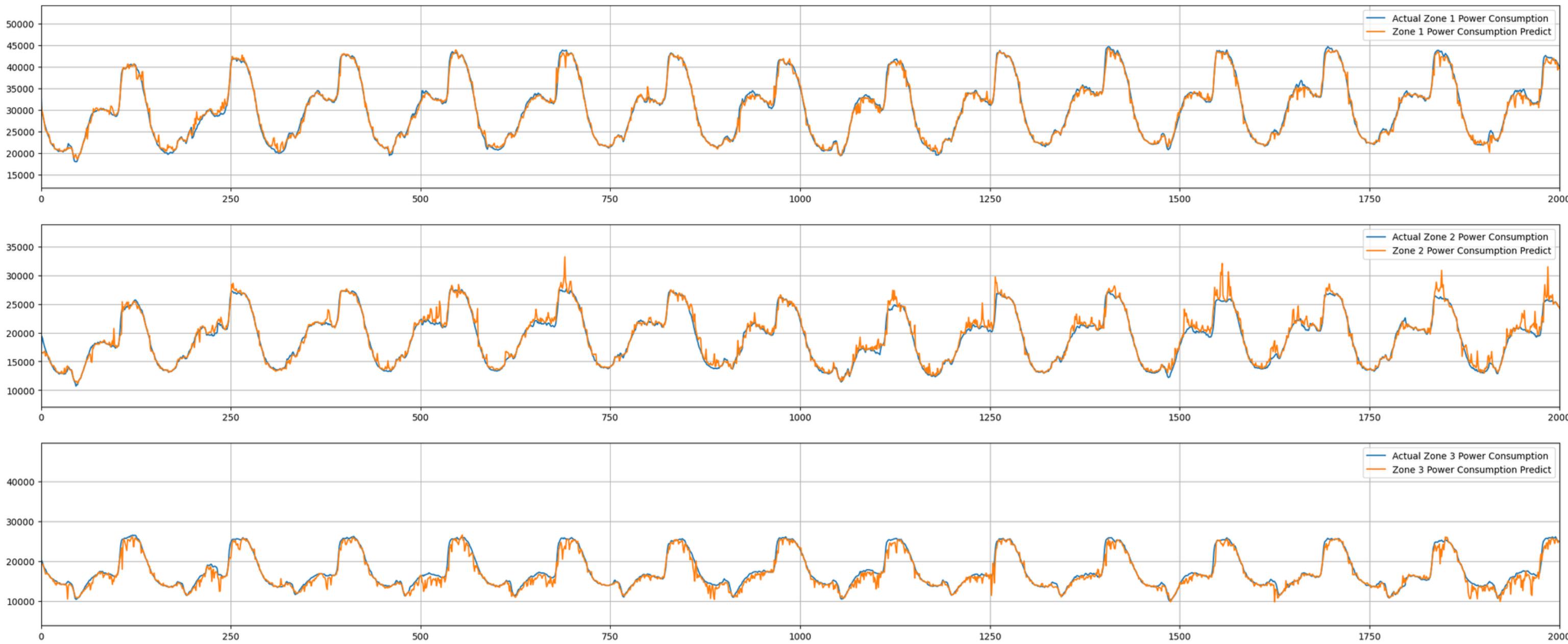
Opt for PCA1-PCA5 (Cumulative Explained Variance = 89%)
That is, it can represent 89% of actual data (7 Features)



RF MODEL3 (PCA)

- Features**
- General Diffuse Flows
 - Diffuse Flows
 - Humidity
 - Hour
 - Temperature

Evaluate: R Square



EVALUATE

	Model	Zone 1	Zone 2	Zone 3
1	MAE	3233.6401	2393.1856	2470.7624
	MSE	22116671.3085	11677545.3489	12667753.5063
	RSME	4702.8365	3417.2423	3559.1787
2	MAE	1005.6263	948.0474	1209.6280
	MSE	200011763.5549	1946933.1871	3356146.6749
	RSME	1529.2593	1395.3254	1831.9788
3	MAE	1160.5930	1080.9362	1370.0767
	MSE	2985146.6584	2544661.1557	4424015.1465
	RSME	1727.7576	1595.1994	2103.3342



EVALUATE



Model		Zone 1	Zone 2	Zone 3
1	MAE	3233.6401	2393.1856	2470.7624
	MSE	22116671.3085	11677545.3489	12667753.5063
	RSME	4702.8365	3417.2423	3559.1787
2	MAE	1005.6263	948.0474	1209.6280
	MSE	200011763.5549	1946933.1871	3356146.6749
	RSME	1529.2593	1395.3254	1831.9788
3	MAE	1160.5930	1080.9362	1370.0767
	MSE	2985146.6584	2544661.1557	4424015.1465
	RSME	1727.7576	1595.1994	2103.3342

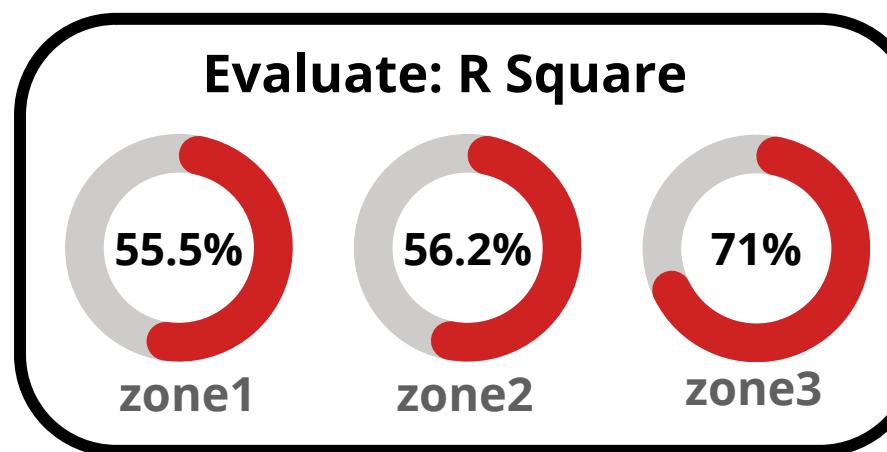
EVALUATE

Model		Zone 1	Zone 2	Zone 3
1	MAE	3233.6401	2393.1856	2470.7624
	MSE	22116671.3085	11677545.3489	12667753.5063
	RSME	4702.8365	3417.2423	3559.1787
2	MAE	1005.6263	948.0474	1209.6280
	MSE	200011763.5549	1946933.1871	3356146.6749
	RSME	1529.2593	1395.3254	1831.9788
3	MAE	1160.5930	1080.9362	1370.0767
	MSE	2985146.6584	2544661.1557	4424015.1465
	RSME	1727.7576	1595.1994	2103.3342



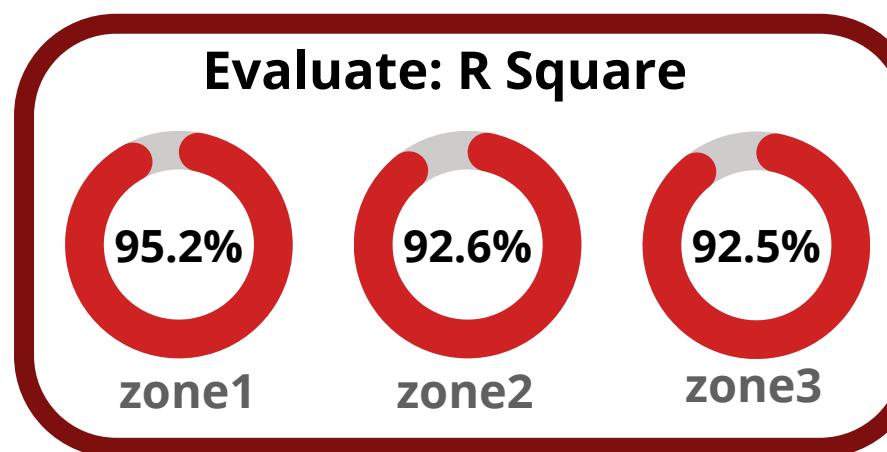
EVALUATE

RF MODEL1 (Whether)



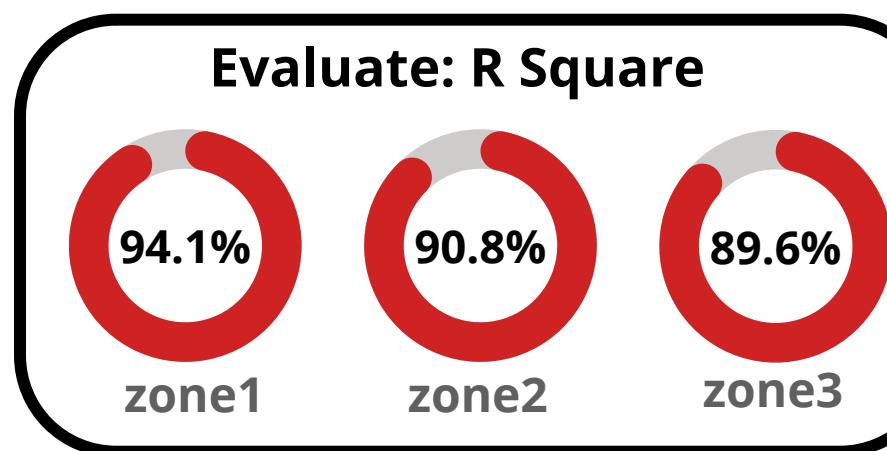
"Model obtained from a dataset that is not optimized"

RF MODEL2 (Whether+Time)



"Model that adds time features "
BEST MODEL

RF MODEL3 (PCA)



"Model that uses features from PCA "