W11_PairplotConsumption

January 12, 2021

1 Pairplot of the consumption

Using all the features the team thinks are necessary

Niels van Drunen 18062814 [Navigator]

Jefry el Bhwash 16095065 [Driver]

[9]:
```
import pandas as pd
import seaborn as sns
import numpy as np
```

2 Reading features from the house data

[10]:
```
df = pd.read_pickle('consumption_df_28')
df['smartMeter_6'] = df['smartMeter_6'].diff()
df['solar_3'] = df['solar_3'].diff()
df['solar_2'] = df['solar_2'].diff()
df['smartMeter_7'] = df['smartMeter_7'].diff()
df['Consumption'] = df.apply(lambda x: x['smartMeter_6'] +␣
 ↪(x['solar_3']-x['solar_2'])-x['smartMeter_7'], axis=1)
df = df.dropna()
df = df.drop(['smartMeter_6', 'smartMeter_7', 'solar_2', 'solar_3'], axis=1)
df.columns = ['holiday', 'op_mode', 'outdoor_temp', 'target_temp', 'room_temp',␣
 ↪'co2', 'humidity', 'consumption']
df['dif_temp'] = df['target_temp'] - df['room_temp']
df = df.drop(['target_temp', 'room_temp'], axis=1)
df = df.resample('60min').agg({'holiday': np.median, 'op_mode':np.median,␣
 ↪'outdoor_temp': np.mean, 'co2': np.mean, 'humidity':np.mean, 'consumption':
 ↪np.sum, 'dif_temp':np.mean})
df.head()
```

[10]:

	holiday	op_mode	outdoor_temp	co2	humidity	\
2018-12-31 23:00:00	0	0.0	13.000000	781.672727	68.118972	
2019-01-01 00:00:00	0	0.0	12.750000	768.975000	66.555833	

```
2019-01-01 01:00:00        0      0.0     12.166667  688.800000  66.451667
2019-01-01 02:00:00        0      0.0     12.583333  602.076515  65.753409
2019-01-01 03:00:00        0      1.0      8.750000  486.162963  64.956204

                     consumption  dif_temp
2018-12-31 23:00:00        0.622 -1.590909
2019-01-01 00:00:00        0.402 -2.875000
2019-01-01 01:00:00        0.264 -2.500000
2019-01-01 02:00:00        0.163 -2.500000
2019-01-01 03:00:00        1.526 -2.375000
```

3 Reading Weather Data from KNMI

[11]:
```
weer = pd.read_pickle('KNMI_DF_W11_full')
weer.head()
```

[11]:
```
                     zonneschijn  bewolking  regen
DateTime
2018-12-31 00:00:00          0.0          8      0
2018-12-31 01:00:00          0.0          8      0
2018-12-31 02:00:00          0.0          8      0
2018-12-31 03:00:00          0.0          8      0
2018-12-31 04:00:00          0.0          8      0
```

4 Merging weather data with house data

[12]:
```
df = pd.merge(df, weer, left_index=True, right_on=weer.index)
df = df.drop('key_0', axis=1)
df['op_mode'] = df['op_mode'].astype(int)
df['hour'] = df.index.hour
df['weekday'] = df.index.dayofweek
df.head()
```

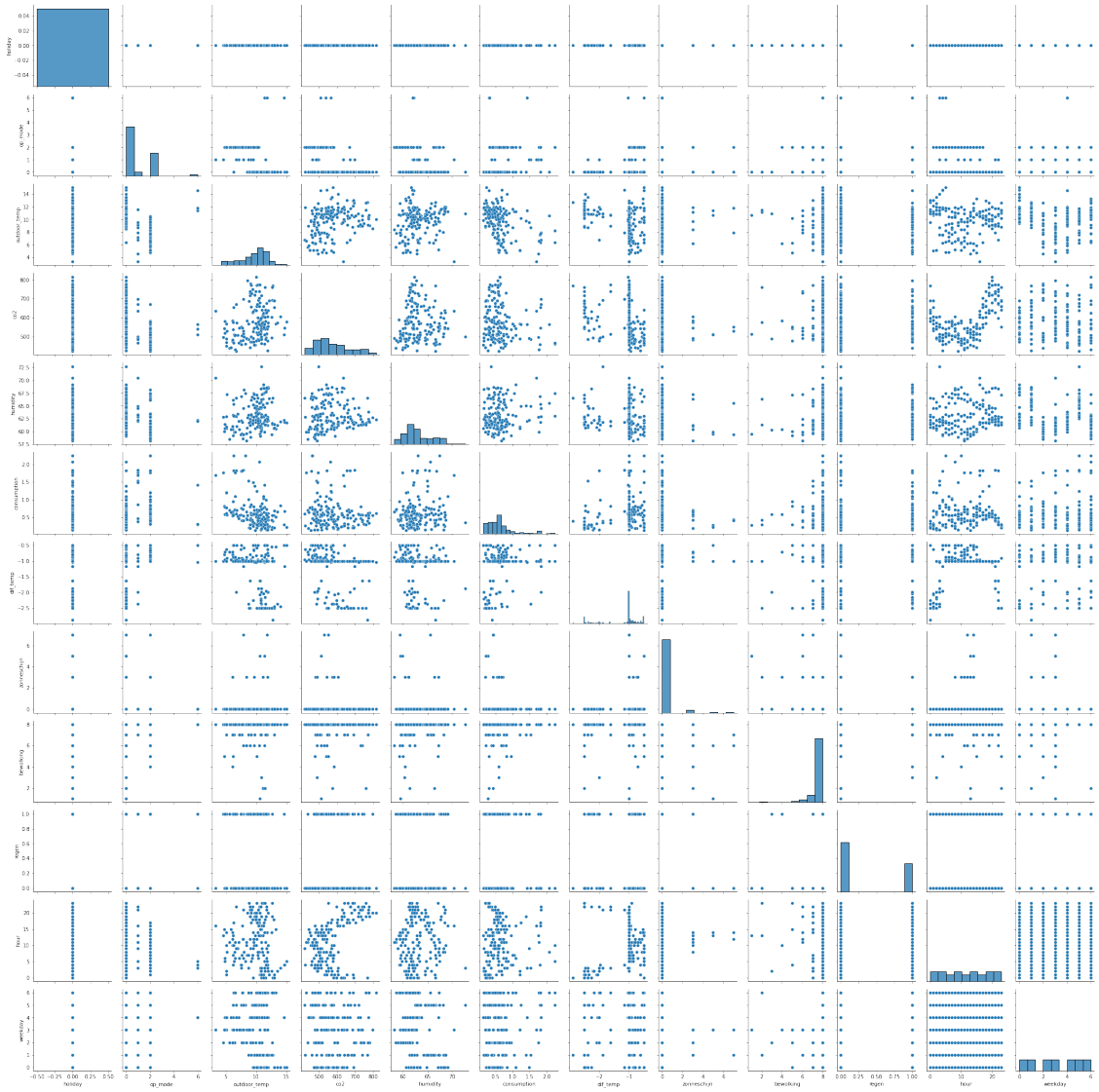
[12]:
```
                     holiday  op_mode  outdoor_temp         co2   humidity  \
DateTime
2018-12-31 23:00:00        0        0     13.000000  781.672727  68.118972
2019-01-01 00:00:00        0        0     12.750000  768.975000  66.555833
2019-01-01 01:00:00        0        0     12.166667  688.800000  66.451667
2019-01-01 02:00:00        0        0     12.583333  602.076515  65.753409
2019-01-01 03:00:00        0        1      8.750000  486.162963  64.956204

                     consumption  dif_temp  zonneschijn  bewolking  regen  \
DateTime
```

```
2018-12-31 23:00:00        0.622 -1.590909          0.0          8      0
2019-01-01 00:00:00        0.402 -2.875000          0.0          8      0
2019-01-01 01:00:00        0.264 -2.500000          0.0          8      0
2019-01-01 02:00:00        0.163 -2.500000          0.0          8      0
2019-01-01 03:00:00        1.526 -2.375000          0.0          7      0

                     hour  weekday
DateTime
2018-12-31 23:00:00    23        0
2019-01-01 00:00:00     0        1
2019-01-01 01:00:00     1        1
2019-01-01 02:00:00     2        1
2019-01-01 03:00:00     3        1
```

[15]:
```
week = df['2019-01-01':'2019-01-07']
pair_plot = sns.PairGrid(week)
pair_plot.map_diag(sns.histplot)
pair_plot.map_offdiag(sns.scatterplot)
#pair_plot.savefig('Pairplot_Consumption_firstweek.png')
```


holiday
op_mode
outdoor_temp
co2
humidity
consumption
dif_temp
zonneschijn
bewolking
regen
hour
weekday