

# 490 Project - Python DataFrames

November 19, 2020

```
In [112]: import csv
          from pyspark.sql import SparkSession
          import pyspark.sql.functions as f
          # User Defined Functions
          from pyspark.sql.functions import udf
          from pyspark.sql.types import DoubleType
          # Stats
          from pyspark.ml.feature import VectorAssembler
          from pyspark.ml.stat import Correlation
          from math import sqrt
          import numpy as np
          import pandas as pd
```

```
In [2]: # Creates a new Spark session w/in Python
        spark = SparkSession.builder.appName("Final Project").getOrCreate()
```

## 1 Data Setup

Importing the data

```
In [165]: # Reads the local csv stored on my computer, it does have a header
          energy_df = spark.read.csv("C:\\Users\\Wes\\Desktop\\CS 490\\energy_dataset.csv", header=True)
          weather_df = spark.read.csv("C:\\Users\\Wes\\Desktop\\CS 490\\weather_features.csv", header=True)
          # Creates a temporary view from the dataframe we read from the file
          # This is how we can read SQL from it
          energy_df.createOrReplaceTempView("Energy")
          weather_df.createOrReplaceTempView("Weather")
```

Creating the Combined dataframe

```
In [166]: joined_df = energy_df.join(weather_df, energy_df.time == weather_df.dt_iso)
          joined_df.show(1, vertical=True)
```

```
-RECORD 0-----
time                | 2015-01-01 00:00:00
generation_biomass   | 447.0
generation_fossil_brown_coal_lignite | 329.0
```

generation_fossil_coal_derived_gas	0.0
generation_fossil_gas	4844.0
generation_fossil_hard_coal	4821.0
generation_fossil_oil	162.0
generation_fossil_oil_shale	0.0
generation_fossil_peat	0.0
generation_geothermal	0.0
generation_hydro_pumped_storage_aggregated	null
generation_hydro_pumped_storage_consumption	863.0
generation_hydro_run_of_river_and_poundage	1051.0
generation_hydro_water_reservoir	1899.0
generation_marine	0.0
generation_nuclear	7096.0
generation_other	43.0
generation_other_renewable	73.0
generation_solar	49.0
generation_waste	196.0
generation_wind_offshore	0.0
generation_wind_onshore	6378.0
forecast_solar_day_ahead	17.0
forecast_wind_offshore_eday_ahead	null
forecast_wind_onshore_day_ahead	6436.0
total_load_forecast	26118.0
total_load_actual	25385.0
price_day_ahead	50.1
price_actual	65.41
dt_iso	2015-01-01 00:00:00
city_name	Valencia
temp	270.475
temp_min	270.475
temp_max	270.475
pressure	1001
humidity	77
wind_speed	1
wind_deg	62
rain_1h	0.0
rain_3h	0.0
snow_3h	0.0
clouds_all	0
weather_id	800
weather_main	clear
weather_description	sky is clear
weather_icon	01n

only showing top 1 row

Creating columns for Farenheight temperatures and filling na values

```
In [167]: # Convert Kelvin temps to Farhenheight
k_to_f_udf = udf(lambda kelvin: (float(kelvin) - 273.15) * (9/5) + 32, DoubleType())
joined_df = joined_df.withColumn("temp_f", k_to_f_udf(joined_df.temp))
joined_df = joined_df.withColumn("temp_min_f", k_to_f_udf(joined_df.temp_min))
joined_df = joined_df.withColumn("temp_max_f", k_to_f_udf(joined_df.temp_max))
joined_df.na.fill(0)
```

```
Out[167]: DataFrame[time: string, generation_biomass: double, generation_fossil_brown_coal_lig
```

## 2 Data Aggregation

### 2.1 By City

```
In [168]: joined_df.groupBy("city_name", "weather_main")\
    .agg({"time" : "count", "price_actual" : "avg"})\
    .withColumnRenamed('count(time)', 'count')\
    .withColumnRenamed('avg(price_actual)', 'avg_price')\
    .filter("count > 100")\
    .orderBy("city_name", "weather_main")\
    .show(100, False)
```

city_name	weather_main	avg_price	count
Barcelona	clear	59.30181183960987	14764
Barcelona	clouds	57.059017502482575	16112
Barcelona	drizzle	56.17348837209302	301
Barcelona	mist	55.9887133182844	443
Barcelona	rain	55.62026551226556	3465
Barcelona	thunderstorm	62.42364820846908	307
Bilbao	clear	60.56761825922418	8456
Bilbao	clouds	57.37178233815964	16714
Bilbao	drizzle	60.67769072164944	485
Bilbao	fog	62.91078397212544	1148
Bilbao	mist	59.97895819508959	1507
Bilbao	rain	54.971499514495804	7209
Bilbao	snow	55.17807228915661	166
Bilbao	thunderstorm	58.42384615384614	208
Madrid	clear	58.671028877320325	20362
Madrid	clouds	56.2445749178017	10645
Madrid	drizzle	57.66627943485085	637
Madrid	fog	64.63083333333327	708
Madrid	mist	62.92892324093817	938
Madrid	rain	56.281083929243515	2657
Madrid	thunderstorm	63.46621621621619	222
Seville	clear	58.691902238426174	23588
Seville	clouds	56.31560800552104	7245
Seville	drizzle	49.820041152263364	243

Seville	dust	58.34052173913042	345	
Seville	fog	56.43466292134829	534	
Seville	haze	53.48614942528738	348	
Seville	mist	59.37509638554222	830	
Seville	rain	55.29623897707229	2268	
Seville	thunderstorm	57.56370370370372	135	
Valencia	clear	57.33447268515524	15541	
Valencia	clouds	58.23793175005739	17348	
Valencia	mist	56.55921052631578	190	
Valencia	rain	58.80600779510035	1796	
Valencia	thunderstorm	63.03674556213019	169	

```
+-----+-----+-----+-----+
```

## 2.2 By Wind Speed

```
In [38]: joined_df.groupby("wind_speed")\
        .agg({"time" : "count", "price_actual" : "avg"})\
        .withColumnRenamed('count(time)', 'count')\
        .withColumnRenamed('avg(price_actual)', 'avg_price')\
        .filter("count > 100")\
        .sort("wind_speed")\
        .show(100, False)
```

wind_speed	avg_price	count
------------	-----------	-------

```
+-----+-----+-----+
```

0	58.48113213667815	18496
1	59.43090143599589	55223
10	43.998986866791746	533
11	45.793608562691134	327
12	48.36364532019704	203
2	59.01943278830858	34555
3	58.0480008785592	25041
4	57.06761207885105	18313
5	55.98622442865706	11683
6	54.275993523697394	6794
7	51.87357766604919	3779
8	48.61905970850964	2127
9	46.72045927209707	1154

```
+-----+-----+-----+
```

## 2.3 By Month and Day

```
In [50]: joined_df.groupby(f.month('time'), f.dayofmonth('time'))\
        .agg({"time" : "count", "price_actual" : "avg"})\
```

```

.withColumnRenamed('dayofmonth(time)', 'day_of_month')\
.withColumnRenamed('month(time)', 'month')\
.withColumnRenamed('count(time)', 'count')\
.withColumnRenamed('avg(price_actual)', 'avg_price')\
.orderBy('month', 'day_of_month')\
.filter("count > 100")\
.show(100, False)

```

month	day_of_month	avg_price	count
1	1	46.19534161490682	483
1	2	56.790477178423245	482
1	3	54.05941414141414	495
1	4	55.018605577689236	502
1	5	60.3349007936508	504
1	6	55.52820512820514	507
1	7	59.728073217726376	519
1	8	65.08919087136927	482
1	9	64.61051792828685	502
1	10	56.60917647058824	510
1	11	58.63578313253013	498
1	12	66.3585655737705	488
1	13	63.68467065868262	501
1	14	62.59476861167003	497
1	15	60.28846790890269	483
1	16	61.2948065173116	491
1	17	61.06637113402061	485
1	18	64.53259336099586	482
1	19	70.2423217922607	491
1	20	71.31931392931394	481
1	21	65.19846311475408	488
1	22	64.41760162601625	492
1	23	67.78447154471543	492
1	24	64.34635627530368	494
1	25	66.61675456389453	493
1	26	69.434670781893	486
1	27	65.08360396039608	505
1	28	61.982448559670786	486
1	29	60.52966336633664	505
1	30	57.8948780487805	492
1	31	54.780824949698165	497
2	1	55.984112903225814	496
2	2	57.248588469184895	503
2	3	52.65928861788617	492
2	4	52.68297665369651	514
2	5	51.47737903225805	496
2	6	54.54255717255718	481

2	7	54.42419222903888	489	
2	8	56.40934693877553	490	
2	9	58.710910973084886	483	
2	10	59.3665731462926	499	
2	11	58.287439024390245	492	
2	12	55.75404	500	
2	13	50.59402414486921	497	
2	14	47.039126819126835	481	
2	15	49.80128099173554	484	
2	16	59.861366459627334	483	
2	17	55.82488888888886	495	
2	18	55.50045174537988	487	
2	19	58.44243243243243	481	
2	20	56.623929313929324	481	
2	21	49.219583333333334	480	
2	22	49.79297297297297	481	
2	23	53.33722222222222	486	
2	24	51.30213725490197	510	
2	25	51.41047422680412	485	
2	26	52.438012170385385	493	
2	27	48.79786138613861	505	
2	28	45.658138195777354	521	
2	29	30.163749999999997	120	
3	1	48.848900000000001	500	
3	2	51.46497975708502	494	
3	3	50.67409465020576	486	
3	4	43.0492531120332	482	
3	5	39.43883299798792	497	
3	6	47.92363265306124	490	
3	7	52.011443298969084	485	
3	8	50.85269938650307	489	
3	9	49.469940119760494	501	
3	10	47.69879345603272	489	
3	11	42.41525987525987	481	
3	12	43.564	485	
3	13	46.310897959183684	490	
3	14	47.92568047337277	507	
3	15	49.183134020618546	485	
3	16	54.21541152263375	486	
3	17	53.40634496919915	487	
3	18	51.95104166666665	480	
3	19	50.748416833667335	499	
3	20	48.857700205338816	487	
3	21	51.137582644628125	484	
3	22	50.82339506172839	486	
3	23	52.803757700205324	487	
3	24	46.30303643724696	494	
3	25	47.54305389221556	501	

3	26	47.61983298538623	479	
3	27	47.35381147540984	488	
3	28	46.87736514522822	482	
3	29	44.39	477	
3	30	40.541776859504125	484	
3	31	40.041351351351366	481	
4	1	40.07055785123967	484	
4	2	42.44409937888199	483	
4	3	42.82779625779625	481	
4	4	42.46078350515464	485	
4	5	42.83212678936605	489	
4	6	45.4385596707819	486	
4	7	47.995336048879835	491	
4	8	45.210269151138725	483	
4	9	49.79811983471073	484	

+-----+-----+-----+-----+

only showing top 100 rows

### 3 Correlation

```
In [188]: def correlation_matrix(df, corr_columns, method='pearson'):
    vector_col = "corr_features"
    assembler = VectorAssembler(inputCols=corr_columns, outputCol=vector_col)
    df_vector = assembler.setHandleInvalid("keep").transform(df).select(vector_col)
    np.nan_to_num(df_vector)
    matrix = Correlation.corr(df_vector, vector_col, method)

    result = matrix.collect()[0]["pearson({})".format(vector_col)].values
    return pd.DataFrame(result.reshape(-1, len(corr_columns)), columns=corr_columns,
```

#### 3.1 Temperature, price, and actual load by year and month

```
In [198]: test_agg = joined_df.groupBy(f.year('time'), f.month('time'))\
    .agg({"price_actual" : "avg", \
        "total_load_actual" : "avg", \
        "temp_f": "avg", \
        "temp_min_f": "avg", \
        "temp_max_f": "avg"})\
    .withColumnRenamed('year(time)', 'year')\
    .withColumnRenamed('month(time)', 'month')\
    .withColumnRenamed('avg(price_actual)', 'avg_price')\
    .withColumnRenamed('avg(total_load_actual)', 'avg_total_load_actual')\
    .orderBy('year', 'month')
```

```
In [200]: correlation_matrix(test_agg, test_agg.columns)
```

```
Out [200]:
```

	year	month	avg(temp_f)	avg(temp_max_f)	\
year	1.000000	0.000000	-0.024369	-0.061868	
month	0.000000	1.000000	0.276436	0.286410	
avg(temp_f)	-0.024369	0.276436	1.000000	0.989586	
avg(temp_max_f)	-0.061868	0.286410	0.989586	1.000000	
avg(temp_min_f)	0.012814	0.260331	0.989929	0.959802	
avg_total_load_actual	0.304222	-0.210858	-0.104184	-0.124596	
avg_price	0.201002	0.389448	0.023399	-0.008685	

  

	avg(temp_min_f)	avg_total_load_actual	avg_price
year	0.012814	0.304222	0.201002
month	0.260331	-0.210858	0.389448
avg(temp_f)	0.989929	-0.104184	0.023399
avg(temp_max_f)	0.959802	-0.124596	-0.008685
avg(temp_min_f)	1.000000	-0.084801	0.047780
avg_total_load_actual	-0.084801	1.000000	0.283745
avg_price	0.047780	0.283745	1.000000

### 3.2 Generation\_ by Year

```
In [194]: test_agg = joined_df.groupBy(f.year('time'))\
    .agg({"price_actual" : "avg", \
        "total_load_actual" : "avg",\
        "generation_biomass" : "avg",\
        "generation_fossil_brown_coal_lignite" : "avg",\
        "generation_fossil_coal_derived_gas" : "avg",\
        "generation_fossil_gas" : "avg",\
        "generation_fossil_hard_coal" : "avg",\
        "generation_fossil_oil" : "avg",\
        "generation_fossil_oil_shale" : "avg",\
        "generation_fossil_peat" : "avg",\
        "generation_geothermal" : "avg",\
        "generation_hydro_pumped_storage_aggregated" : "avg",\
        "generation_hydro_pumped_storage_consumption" : "avg",\
        "generation_hydro_run_of_river_and_poundage" : "avg",\
        "generation_hydro_water_reservoir" : "avg",\
        "generation_marine" : "avg",\
        "generation_nuclear" : "avg",\
        "generation_other" : "avg",\
        "generation_other_renewable" : "avg",\
        "generation_solar" : "avg",\
        "generation_waste" : "avg",\
        "generation_wind_offshore" : "avg",\
        "generation_wind_onshore" : "avg",\
    })\
    .withColumnRenamed('year(time)', 'year')\
    .orderBy('year')
```



```
In [195]: correlation_matrix(test_agg, test_agg.columns)
```

```
Out[195]:
```

	year \
year	1.000000
avg(generation_fossil_gas)	0.707407
avg(generation_fossil_brown_coal_lignite)	-0.443560
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.112402
avg(total_load_actual)	0.988030
avg(generation_solar)	-0.376027
avg(generation_geothermal)	NaN
avg(generation_other)	-0.863683
avg(price_actual)	0.325889
avg(generation_other_renewable)	0.974178
avg(generation_nuclear)	-0.508627
avg(generation_hydro_pumped_storage_consumption)	-0.870629
avg(generation_biomass)	-0.863918
avg(generation_marine)	NaN
avg(generation_waste)	0.950379
avg(generation_fossil_hard_coal)	-0.701993
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.522804
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	0.514543
avg(generation_fossil_oil)	-0.796692

	avg(generation_fossil_gas) \
year	0.707407
avg(generation_fossil_gas)	1.000000
avg(generation_fossil_brown_coal_lignite)	0.296033
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.750713
avg(total_load_actual)	0.772630
avg(generation_solar)	0.386652
avg(generation_geothermal)	NaN
avg(generation_other)	-0.823875
avg(price_actual)	0.379506
avg(generation_other_renewable)	0.831806
avg(generation_nuclear)	-0.028780
avg(generation_hydro_pumped_storage_consumption)	-0.819886
avg(generation_biomass)	-0.678732
avg(generation_marine)	NaN
avg(generation_waste)	0.855207
avg(generation_fossil_hard_coal)	-0.241981

avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.230797
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.067438
avg(generation_fossil_oil)	-0.424401
	avg(generation_fossil_brown_coal_1.
year	-0
avg(generation_fossil_gas)	0
avg(generation_fossil_brown_coal_lignite)	1
avg(generation_hydro_pumped_storage_aggregated)	
avg(generation_fossil_peat)	
avg(generation_wind_offshore)	
avg(generation_hydro_water_reservoir)	-0
avg(total_load_actual)	-0
avg(generation_solar)	0
avg(generation_geothermal)	
avg(generation_other)	0
avg(price_actual)	0
avg(generation_other_renewable)	-0
avg(generation_nuclear)	0
avg(generation_hydro_pumped_storage_consumption)	0
avg(generation_biomass)	0
avg(generation_marine)	
avg(generation_waste)	-0
avg(generation_fossil_hard_coal)	0
avg(generation_fossil_coal_derived_gas)	
avg(generation_hydro_run_of_river_and_poundage)	-0
avg(generation_fossil_oil_shale)	
avg(generation_wind_onshore)	-0
avg(generation_fossil_oil)	0
	avg(generation_hydro_pumped_storage
year	
avg(generation_fossil_gas)	
avg(generation_fossil_brown_coal_lignite)	
avg(generation_hydro_pumped_storage_aggregated)	
avg(generation_fossil_peat)	
avg(generation_wind_offshore)	
avg(generation_hydro_water_reservoir)	
avg(total_load_actual)	
avg(generation_solar)	
avg(generation_geothermal)	
avg(generation_other)	
avg(price_actual)	
avg(generation_other_renewable)	
avg(generation_nuclear)	
avg(generation_hydro_pumped_storage_consumption)	

```

avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```

                                                    avg(generation_fossil_peat) \
year                                                    NaN
avg(generation_fossil_gas)                            NaN
avg(generation_fossil_brown_coal_lignite)             NaN
avg(generation_hydro_pumped_storage_aggregated)       NaN
avg(generation_fossil_peat)                           1.0
avg(generation_wind_offshore)                         NaN
avg(generation_hydro_water_reservoir)                 NaN
avg(total_load_actual)                                NaN
avg(generation_solar)                                 NaN
avg(generation_geothermal)                            NaN
avg(generation_other)                                 NaN
avg(price_actual)                                     NaN
avg(generation_other_renewable)                       NaN
avg(generation_nuclear)                               NaN
avg(generation_hydro_pumped_storage_consumption)      NaN
avg(generation_biomass)                               NaN
avg(generation_marine)                                NaN
avg(generation_waste)                                 NaN
avg(generation_fossil_hard_coal)                     NaN
avg(generation_fossil_coal_derived_gas)              NaN
avg(generation_hydro_run_of_river_and_poundage)       NaN
avg(generation_fossil_oil_shale)                     NaN
avg(generation_wind_onshore)                         NaN
avg(generation_fossil_oil)                           NaN

```

```

                                                    avg(generation_wind_offshore) \
year                                                    NaN
avg(generation_fossil_gas)                            NaN
avg(generation_fossil_brown_coal_lignite)             NaN
avg(generation_hydro_pumped_storage_aggregated)       NaN
avg(generation_fossil_peat)                           NaN
avg(generation_wind_offshore)                         1.0
avg(generation_hydro_water_reservoir)                 NaN
avg(total_load_actual)                                NaN
avg(generation_solar)                                 NaN
avg(generation_geothermal)                            NaN
avg(generation_other)                                 NaN

```

avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	avg(generation_hydro_water_reservoir)	
year		-0.1124
avg(generation_fossil_gas)		-0.7507
avg(generation_fossil_brown_coal_lignite)		-0.8404
avg(generation_hydro_pumped_storage_aggregated)		1.0000
avg(generation_fossil_peat)		1.0000
avg(generation_wind_offshore)		1.0000
avg(generation_hydro_water_reservoir)		1.0000
avg(total_load_actual)		-0.2433
avg(generation_solar)		-0.8577
avg(generation_geothermal)		1.0000
avg(generation_other)		0.2644
avg(price_actual)		-0.5063
avg(generation_other_renewable)		-0.2711
avg(generation_nuclear)		-0.1763
avg(generation_hydro_pumped_storage_consumption)		0.4888
avg(generation_biomass)		0.0433
avg(generation_marine)		1.0000
avg(generation_waste)		-0.3020
avg(generation_fossil_hard_coal)		-0.4423
avg(generation_fossil_coal_derived_gas)		1.0000
avg(generation_hydro_run_of_river_and_poundage)		0.7288
avg(generation_fossil_oil_shale)		1.0000
avg(generation_wind_onshore)		0.3455
avg(generation_fossil_oil)		-0.2644

	avg(total_load_actual)	\
year	0.988030	
avg(generation_fossil_gas)	0.772630	
avg(generation_fossil_brown_coal_lignite)	-0.320710	
avg(generation_hydro_pumped_storage_aggregated)	NaN	
avg(generation_fossil_peat)	NaN	
avg(generation_wind_offshore)	NaN	
avg(generation_hydro_water_reservoir)	-0.243325	

avg(total_load_actual)	1.000000
avg(generation_solar)	-0.266476
avg(generation_geothermal)	NaN
avg(generation_other)	-0.838801
avg(price_actual)	0.448746
avg(generation_other_renewable)	0.972060
avg(generation_nuclear)	-0.535851
avg(generation_hydro_pumped_storage_consumption)	-0.935592
avg(generation_biomass)	-0.807328
avg(generation_marine)	NaN
avg(generation_waste)	0.946351
avg(generation_fossil_hard_coal)	-0.586183
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.435614
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	0.513215
avg(generation_fossil_oil)	-0.701659

	avg(generation_solar) \
year	-0.376027
avg(generation_fossil_gas)	0.386652
avg(generation_fossil_brown_coal_lignite)	0.982195
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.857189
avg(total_load_actual)	-0.266476
avg(generation_solar)	1.000000
avg(generation_geothermal)	NaN
avg(generation_other)	0.077832
avg(price_actual)	0.143228
avg(generation_other_renewable)	-0.187746
avg(generation_nuclear)	0.567939
avg(generation_hydro_pumped_storage_consumption)	0.030115
avg(generation_biomass)	0.274732
avg(generation_marine)	NaN
avg(generation_waste)	-0.131752
avg(generation_fossil_hard_coal)	0.644033
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.976062
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.710615
avg(generation_fossil_oil)	0.525526

	avg(generation_geothermal) \
year	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN

avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	NaN
avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	1.0
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	...	\
year	...	
avg(generation_fossil_gas)	...	
avg(generation_fossil_brown_coal_lignite)	...	
avg(generation_hydro_pumped_storage_aggregated)	...	
avg(generation_fossil_peat)	...	
avg(generation_wind_offshore)	...	
avg(generation_hydro_water_reservoir)	...	
avg(total_load_actual)	...	
avg(generation_solar)	...	
avg(generation_geothermal)	...	
avg(generation_other)	...	
avg(price_actual)	...	
avg(generation_other_renewable)	...	
avg(generation_nuclear)	...	
avg(generation_hydro_pumped_storage_consumption)	...	
avg(generation_biomass)	...	
avg(generation_marine)	...	
avg(generation_waste)	...	
avg(generation_fossil_hard_coal)	...	
avg(generation_fossil_coal_derived_gas)	...	
avg(generation_hydro_run_of_river_and_poundage)	...	
avg(generation_fossil_oil_shale)	...	
avg(generation_wind_onshore)	...	
avg(generation_fossil_oil)	...	

```

year
avg(generation_fossil_gas)
avg(generation_fossil_brown_coal_lignite)
avg(generation_hydro_pumped_storage_aggregated)
avg(generation_fossil_peat)
avg(generation_wind_offshore)
avg(generation_hydro_water_reservoir)
avg(total_load_actual)
avg(generation_solar)
avg(generation_geothermal)
avg(generation_other)
avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```
avg(generation_hydro_pumped_storage)
```

```

year
avg(generation_fossil_gas)
avg(generation_fossil_brown_coal_lignite)
avg(generation_hydro_pumped_storage_aggregated)
avg(generation_fossil_peat)
avg(generation_wind_offshore)
avg(generation_hydro_water_reservoir)
avg(total_load_actual)
avg(generation_solar)
avg(generation_geothermal)
avg(generation_other)
avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)

```

	avg(generation_biomass) \
year	-0.863918
avg(generation_fossil_gas)	-0.678732
avg(generation_fossil_brown_coal_lignite)	0.420684
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	0.043260
avg(total_load_actual)	-0.807328
avg(generation_solar)	0.274732
avg(generation_geothermal)	NaN
avg(generation_other)	0.974787
avg(price_actual)	0.154014
avg(generation_other_renewable)	-0.910937
avg(generation_nuclear)	0.037389
avg(generation_hydro_pumped_storage_consumption)	0.582766
avg(generation_biomass)	1.000000
avg(generation_marine)	NaN
avg(generation_waste)	-0.931177
avg(generation_fossil_hard_coal)	0.875800
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.337062

avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.088375
avg(generation_fossil_oil)	0.951989
	avg(generation_marine) \
year	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	NaN
avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	1.0
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN
	avg(generation_waste) \
year	0.950379
avg(generation_fossil_gas)	0.855207
avg(generation_fossil_brown_coal_lignite)	-0.240190
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.302046
avg(total_load_actual)	0.946351
avg(generation_solar)	-0.131752
avg(generation_geothermal)	NaN
avg(generation_other)	-0.968555
avg(price_actual)	0.205121
avg(generation_other_renewable)	0.995540
avg(generation_nuclear)	-0.237445
avg(generation_hydro_pumped_storage_consumption)	-0.833325
avg(generation_biomass)	-0.931177
avg(generation_marine)	NaN



avg(generation_waste)	1.000000
avg(generation_fossil_hard_coal)	-0.680025
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.265321
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	0.227072
avg(generation_fossil_oil)	-0.804868

	avg(generation_fossil_hard_coal)	-0.701993
year		-0.241981
avg(generation_fossil_gas)		0.771253
avg(generation_fossil_brown_coal_lignite)		NaN
avg(generation_hydro_pumped_storage_aggregated)		NaN
avg(generation_fossil_peat)		NaN
avg(generation_wind_offshore)		NaN
avg(generation_hydro_water_reservoir)		-0.442907
avg(total_load_actual)		-0.586183
avg(generation_solar)		0.644033
avg(generation_geothermal)		NaN
avg(generation_other)		0.746076
avg(price_actual)		0.413356
avg(generation_other_renewable)		-0.673309
avg(generation_nuclear)		0.081910
avg(generation_hydro_pumped_storage_consumption)		0.262458
avg(generation_biomass)		0.875800
avg(generation_marine)		NaN
avg(generation_waste)		-0.680025
avg(generation_fossil_hard_coal)		1.000000
avg(generation_fossil_coal_derived_gas)		NaN
avg(generation_hydro_run_of_river_and_poundage)		-0.631705
avg(generation_fossil_oil_shale)		NaN
avg(generation_wind_onshore)		-0.211109
avg(generation_fossil_oil)		0.981258

	avg(generation_fossil_coal_derived_gas)	-0.211109
year		0.981258
avg(generation_fossil_gas)		
avg(generation_fossil_brown_coal_lignite)		
avg(generation_hydro_pumped_storage_aggregated)		
avg(generation_fossil_peat)		
avg(generation_wind_offshore)		
avg(generation_hydro_water_reservoir)		
avg(total_load_actual)		
avg(generation_solar)		
avg(generation_geothermal)		
avg(generation_other)		
avg(price_actual)		
avg(generation_other_renewable)		

```

avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```
avg(generation_hydro_run_of_river_and_poundage)
```

```

year
avg(generation_fossil_gas)
avg(generation_fossil_brown_coal_lignite)
avg(generation_hydro_pumped_storage_aggregated)
avg(generation_fossil_peat)
avg(generation_wind_offshore)
avg(generation_hydro_water_reservoir)
avg(total_load_actual)
avg(generation_solar)
avg(generation_geothermal)
avg(generation_other)
avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```
avg(generation_fossil_oil_shale)
```

```

year
avg(generation_fossil_gas)
avg(generation_fossil_brown_coal_lignite)
avg(generation_hydro_pumped_storage_aggregated)
avg(generation_fossil_peat)
avg(generation_wind_offshore)
avg(generation_hydro_water_reservoir)
avg(total_load_actual)
avg(generation_solar)

```

```

NaN
NaN
NaN
NaN
NaN
NaN
NaN
NaN
NaN

```

avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	1.0
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	avg(generation_wind_onshore) \
year	0.514543
avg(generation_fossil_gas)	-0.067438
avg(generation_fossil_brown_coal_lignite)	-0.608724
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	0.345629
avg(total_load_actual)	0.513215
avg(generation_solar)	-0.710615
avg(generation_geothermal)	NaN
avg(generation_other)	-0.017169
avg(price_actual)	0.594049
avg(generation_other_renewable)	0.317127
avg(generation_nuclear)	-0.982664
avg(generation_hydro_pumped_storage_consumption)	-0.511716
avg(generation_biomass)	-0.088375
avg(generation_marine)	NaN
avg(generation_waste)	0.227072
avg(generation_fossil_hard_coal)	-0.211109
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.828306
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	1.000000
avg(generation_fossil_oil)	-0.186019

	avg(generation_fossil_oil)
year	-0.796692
avg(generation_fossil_gas)	-0.424401
avg(generation_fossil_brown_coal_lignite)	0.661934
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN

avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.264612
avg(total_load_actual)	-0.701659
avg(generation_solar)	0.525526
avg(generation_geothermal)	NaN
avg(generation_other)	0.859697
avg(price_actual)	0.308487
avg(generation_other_renewable)	-0.794180
avg(generation_nuclear)	0.084624
avg(generation_hydro_pumped_storage_consumption)	0.409152
avg(generation_biomass)	0.951989
avg(generation_marine)	NaN
avg(generation_waste)	-0.804868
avg(generation_fossil_hard_coal)	0.981258
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.545330
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.186019
avg(generation_fossil_oil)	1.000000

[24 rows x 24 columns]

### 3.3 Generation\_ By Month

```
In [196]: test_agg = joined_df.groupby(f.month('time'))\
        .agg({"price_actual" : "avg", \
            "total_load_actual" : "avg", \
            "generation_biomass" : "avg", \
            "generation_fossil_brown_coal_lignite" : "avg", \
            "generation_fossil_coal_derived_gas" : "avg", \
            "generation_fossil_gas" : "avg", \
            "generation_fossil_hard_coal" : "avg", \
            "generation_fossil_oil" : "avg", \
            "generation_fossil_oil_shale" : "avg", \
            "generation_fossil_peat" : "avg", \
            "generation_geothermal" : "avg", \
            "generation_hydro_pumped_storage_aggregated" : "avg", \
            "generation_hydro_pumped_storage_consumption" : "avg", \
            "generation_hydro_run_of_river_and_poundage" : "avg", \
            "generation_hydro_water_reservoir" : "avg", \
            "generation_marine" : "avg", \
            "generation_nuclear" : "avg", \
            "generation_other" : "avg", \
            "generation_other_renewable" : "avg", \
            "generation_solar" : "avg", \
            "generation_waste" : "avg", \
            "generation_wind_offshore" : "avg", \
            "generation_wind_onshore" : "avg", \
```

```

    })\
    .withColumnRenamed('month(time)', 'month')\
    .orderBy('month')

```

```
In [197]: correlation_matrix(test_agg, test_agg.columns)
```

```

Out[197]:
      month \
month      1.000000
avg(generation_fossil_gas)      0.721932
avg(generation_fossil_brown_coal_lignite)      0.627535
avg(generation_hydro_pumped_storage_aggregated)      NaN
avg(generation_fossil_peat)      NaN
avg(generation_wind_offshore)      NaN
avg(generation_hydro_water_reservoir)      -0.801728
avg(total_load_actual)      -0.250626
avg(generation_solar)      -0.152388
avg(generation_geothermal)      NaN
avg(generation_other)      0.036723
avg(price_actual)      0.688920
avg(generation_other_renewable)      0.502845
avg(generation_nuclear)      -0.250136
avg(generation_hydro_pumped_storage_consumption)      -0.496155
avg(generation_biomass)      -0.007149
avg(generation_marine)      NaN
avg(generation_waste)      0.534335
avg(generation_fossil_hard_coal)      0.566349
avg(generation_fossil_coal_derived_gas)      NaN
avg(generation_hydro_run_of_river_and_poundage)      -0.707855
avg(generation_fossil_oil_shale)      NaN
avg(generation_wind_onshore)      -0.610899
avg(generation_fossil_oil)      -0.009340

      avg(generation_fossil_gas) \
month      0.721932
avg(generation_fossil_gas)      1.000000
avg(generation_fossil_brown_coal_lignite)      0.893011
avg(generation_hydro_pumped_storage_aggregated)      NaN
avg(generation_fossil_peat)      NaN
avg(generation_wind_offshore)      NaN
avg(generation_hydro_water_reservoir)      -0.815705
avg(total_load_actual)      0.326877
avg(generation_solar)      -0.127170
avg(generation_geothermal)      NaN
avg(generation_other)      0.204300
avg(price_actual)      0.864872
avg(generation_other_renewable)      0.400812
avg(generation_nuclear)      -0.266553
avg(generation_hydro_pumped_storage_consumption)      -0.462019

```

avg(generation_biomass)	0.408114
avg(generation_marine)	NaN
avg(generation_waste)	0.713947
avg(generation_fossil_hard_coal)	0.915238
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.766604
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.501046
avg(generation_fossil_oil)	0.339675

	avg(generation_fossil_brown_coal_lignite)	1
month		0
avg(generation_fossil_gas)		0
avg(generation_fossil_brown_coal_lignite)		1
avg(generation_hydro_pumped_storage_aggregated)		
avg(generation_fossil_peat)		
avg(generation_wind_offshore)		
avg(generation_hydro_water_reservoir)		-0
avg(total_load_actual)		0
avg(generation_solar)		-0
avg(generation_geothermal)		
avg(generation_other)		0
avg(price_actual)		0
avg(generation_other_renewable)		0
avg(generation_nuclear)		-0
avg(generation_hydro_pumped_storage_consumption)		-0
avg(generation_biomass)		0
avg(generation_marine)		
avg(generation_waste)		0
avg(generation_fossil_hard_coal)		0
avg(generation_fossil_coal_derived_gas)		
avg(generation_hydro_run_of_river_and_poundage)		-0
avg(generation_fossil_oil_shale)		
avg(generation_wind_onshore)		-0
avg(generation_fossil_oil)		0

	avg(generation_hydro_pumped_storage_aggregated)	
month		
avg(generation_fossil_gas)		
avg(generation_fossil_brown_coal_lignite)		
avg(generation_hydro_pumped_storage_aggregated)		
avg(generation_fossil_peat)		
avg(generation_wind_offshore)		
avg(generation_hydro_water_reservoir)		
avg(total_load_actual)		
avg(generation_solar)		
avg(generation_geothermal)		
avg(generation_other)		

```

avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

	avg(generation_fossil_peat) \
month	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	1.0
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	NaN
avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	avg(generation_wind_offshore) \
month	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	1.0
avg(generation_hydro_water_reservoir)	NaN

avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN
	avg(generation_hydro_water_reservoir)
month	-0.8017
avg(generation_fossil_gas)	-0.8157
avg(generation_fossil_brown_coal_lignite)	-0.8594
avg(generation_hydro_pumped_storage_aggregated)	1.0000
avg(generation_fossil_peat)	1.0000
avg(generation_wind_offshore)	1.0000
avg(generation_hydro_water_reservoir)	1.0000
avg(total_load_actual)	-0.0444
avg(generation_solar)	-0.0478
avg(generation_geothermal)	1.0000
avg(generation_other)	-0.2053
avg(price_actual)	-0.8681
avg(generation_other_renewable)	-0.3888
avg(generation_nuclear)	-0.0000
avg(generation_hydro_pumped_storage_consumption)	0.6504
avg(generation_biomass)	-0.4309
avg(generation_marine)	1.0000
avg(generation_waste)	-0.8508
avg(generation_fossil_hard_coal)	-0.8358
avg(generation_fossil_coal_derived_gas)	1.0000
avg(generation_hydro_run_of_river_and_poundage)	0.9600
avg(generation_fossil_oil_shale)	1.0000
avg(generation_wind_onshore)	0.7511
avg(generation_fossil_oil)	-0.3491
	avg(total_load_actual) \
month	-0.250626
avg(generation_fossil_gas)	0.326877
avg(generation_fossil_brown_coal_lignite)	0.351370



avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.044412
avg(total_load_actual)	1.000000
avg(generation_solar)	0.014620
avg(generation_geothermal)	NaN
avg(generation_other)	0.388074
avg(price_actual)	0.295006
avg(generation_other_renewable)	-0.327580
avg(generation_nuclear)	0.281043
avg(generation_hydro_pumped_storage_consumption)	0.061434
avg(generation_biomass)	0.664148
avg(generation_marine)	NaN
avg(generation_waste)	0.289135
avg(generation_fossil_hard_coal)	0.464990
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.062771
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	0.218531
avg(generation_fossil_oil)	0.626397

	avg(generation_solar) \
month	-0.152388
avg(generation_fossil_gas)	-0.127170
avg(generation_fossil_brown_coal_lignite)	-0.162395
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.047851
avg(total_load_actual)	0.014620
avg(generation_solar)	1.000000
avg(generation_geothermal)	NaN
avg(generation_other)	-0.299850
avg(price_actual)	-0.304441
avg(generation_other_renewable)	-0.474513
avg(generation_nuclear)	-0.095104
avg(generation_hydro_pumped_storage_consumption)	-0.725989
avg(generation_biomass)	0.067916
avg(generation_marine)	NaN
avg(generation_waste)	0.155315
avg(generation_fossil_hard_coal)	-0.128250
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.047322
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.554381
avg(generation_fossil_oil)	0.294699

	avg(generation_geothermal)	\
month		NaN
avg(generation_fossil_gas)		NaN
avg(generation_fossil_brown_coal_lignite)		NaN
avg(generation_hydro_pumped_storage_aggregated)		NaN
avg(generation_fossil_peat)		NaN
avg(generation_wind_offshore)		NaN
avg(generation_hydro_water_reservoir)		NaN
avg(total_load_actual)		NaN
avg(generation_solar)		NaN
avg(generation_geothermal)		1.0
avg(generation_other)		NaN
avg(price_actual)		NaN
avg(generation_other_renewable)		NaN
avg(generation_nuclear)		NaN
avg(generation_hydro_pumped_storage_consumption)		NaN
avg(generation_biomass)		NaN
avg(generation_marine)		NaN
avg(generation_waste)		NaN
avg(generation_fossil_hard_coal)		NaN
avg(generation_fossil_coal_derived_gas)		NaN
avg(generation_hydro_run_of_river_and_poundage)		NaN
avg(generation_fossil_oil_shale)		NaN
avg(generation_wind_onshore)		NaN
avg(generation_fossil_oil)		NaN
	...	\
month		...
avg(generation_fossil_gas)		...
avg(generation_fossil_brown_coal_lignite)		...
avg(generation_hydro_pumped_storage_aggregated)		...
avg(generation_fossil_peat)		...
avg(generation_wind_offshore)		...
avg(generation_hydro_water_reservoir)		...
avg(total_load_actual)		...
avg(generation_solar)		...
avg(generation_geothermal)		...
avg(generation_other)		...
avg(price_actual)		...
avg(generation_other_renewable)		...
avg(generation_nuclear)		...
avg(generation_hydro_pumped_storage_consumption)		...
avg(generation_biomass)		...
avg(generation_marine)		...
avg(generation_waste)		...
avg(generation_fossil_hard_coal)		...
avg(generation_fossil_coal_derived_gas)		...
avg(generation_hydro_run_of_river_and_poundage)		...

avg(generation_fossil_oil_shale)	...
avg(generation_wind_onshore)	...
avg(generation_fossil_oil)	...
	avg(generation_hydro_pumped_storage)
month	
avg(generation_fossil_gas)	
avg(generation_fossil_brown_coal_lignite)	
avg(generation_hydro_pumped_storage_aggregated)	
avg(generation_fossil_peat)	
avg(generation_wind_offshore)	
avg(generation_hydro_water_reservoir)	
avg(total_load_actual)	
avg(generation_solar)	
avg(generation_geothermal)	
avg(generation_other)	
avg(price_actual)	
avg(generation_other_renewable)	
avg(generation_nuclear)	
avg(generation_hydro_pumped_storage_consumption)	
avg(generation_biomass)	
avg(generation_marine)	
avg(generation_waste)	
avg(generation_fossil_hard_coal)	
avg(generation_fossil_coal_derived_gas)	
avg(generation_hydro_run_of_river_and_poundage)	
avg(generation_fossil_oil_shale)	
avg(generation_wind_onshore)	
avg(generation_fossil_oil)	
	avg(generation_biomass) \
month	-0.007149
avg(generation_fossil_gas)	0.408114
avg(generation_fossil_brown_coal_lignite)	0.545007
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.430973
avg(total_load_actual)	0.664148
avg(generation_solar)	0.067916
avg(generation_geothermal)	NaN
avg(generation_other)	0.702590
avg(price_actual)	0.494994
avg(generation_other_renewable)	-0.104189
avg(generation_nuclear)	0.236782
avg(generation_hydro_pumped_storage_consumption)	-0.219000
avg(generation_biomass)	1.000000
avg(generation_marine)	NaN

avg(generation_waste)	0.396271
avg(generation_fossil_hard_coal)	0.619570
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.490738
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.071424
avg(generation_fossil_oil)	0.850527

	avg(generation_marine) \
month	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	NaN
avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	1.0
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	avg(generation_waste) \
month	0.534335
avg(generation_fossil_gas)	0.713947
avg(generation_fossil_brown_coal_lignite)	0.804039
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.850827
avg(total_load_actual)	0.289135
avg(generation_solar)	0.155315
avg(generation_geothermal)	NaN
avg(generation_other)	0.110818
avg(price_actual)	0.748662
avg(generation_other_renewable)	0.094508

avg(generation_nuclear)	0.303606
avg(generation_hydro_pumped_storage_consumption)	-0.611523
avg(generation_biomass)	0.396271
avg(generation_marine)	NaN
avg(generation_waste)	1.000000
avg(generation_fossil_hard_coal)	0.795894
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.869397
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.696604
avg(generation_fossil_oil)	0.377811

	avg(generation_fossil_hard_coal)	0.566349
month		0.915238
avg(generation_fossil_gas)		0.976464
avg(generation_fossil_brown_coal_lignite)		NaN
avg(generation_hydro_pumped_storage_aggregated)		NaN
avg(generation_fossil_peat)		NaN
avg(generation_wind_offshore)		-0.835560
avg(generation_hydro_water_reservoir)		0.464990
avg(total_load_actual)		-0.128250
avg(generation_solar)		NaN
avg(generation_geothermal)		0.301946
avg(generation_other)		0.928118
avg(price_actual)		0.403654
avg(generation_other_renewable)		-0.058392
avg(generation_nuclear)		-0.426933
avg(generation_hydro_pumped_storage_consumption)		0.619570
avg(generation_biomass)		NaN
avg(generation_marine)		0.795894
avg(generation_waste)		1.000000
avg(generation_fossil_hard_coal)		NaN
avg(generation_fossil_coal_derived_gas)		-0.827654
avg(generation_hydro_run_of_river_and_poundage)		NaN
avg(generation_fossil_oil_shale)		-0.486231
avg(generation_wind_onshore)		0.471489
avg(generation_fossil_oil)		

	avg(generation_fossil_coal_derived_gas)	
month		
avg(generation_fossil_gas)		
avg(generation_fossil_brown_coal_lignite)		
avg(generation_hydro_pumped_storage_aggregated)		
avg(generation_fossil_peat)		
avg(generation_wind_offshore)		
avg(generation_hydro_water_reservoir)		
avg(total_load_actual)		
avg(generation_solar)		

```

avg(generation_geothermal)
avg(generation_other)
avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```
avg(generation_hydro_run_of_river_and_poundage)
```

```

month
avg(generation_fossil_gas)
avg(generation_fossil_brown_coal_lignite)
avg(generation_hydro_pumped_storage_aggregated)
avg(generation_fossil_peat)
avg(generation_wind_offshore)
avg(generation_hydro_water_reservoir)
avg(total_load_actual)
avg(generation_solar)
avg(generation_geothermal)
avg(generation_other)
avg(price_actual)
avg(generation_other_renewable)
avg(generation_nuclear)
avg(generation_hydro_pumped_storage_consumption)
avg(generation_biomass)
avg(generation_marine)
avg(generation_waste)
avg(generation_fossil_hard_coal)
avg(generation_fossil_coal_derived_gas)
avg(generation_hydro_run_of_river_and_poundage)
avg(generation_fossil_oil_shale)
avg(generation_wind_onshore)
avg(generation_fossil_oil)

```

```
avg(generation_fossil_oil_shale)
```

month	NaN
avg(generation_fossil_gas)	NaN
avg(generation_fossil_brown_coal_lignite)	NaN
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN

avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	NaN
avg(total_load_actual)	NaN
avg(generation_solar)	NaN
avg(generation_geothermal)	NaN
avg(generation_other)	NaN
avg(price_actual)	NaN
avg(generation_other_renewable)	NaN
avg(generation_nuclear)	NaN
avg(generation_hydro_pumped_storage_consumption)	NaN
avg(generation_biomass)	NaN
avg(generation_marine)	NaN
avg(generation_waste)	NaN
avg(generation_fossil_hard_coal)	NaN
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	NaN
avg(generation_fossil_oil_shale)	1.0
avg(generation_wind_onshore)	NaN
avg(generation_fossil_oil)	NaN

	avg(generation_wind_onshore) \
month	-0.610899
avg(generation_fossil_gas)	-0.501046
avg(generation_fossil_brown_coal_lignite)	-0.542033
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	0.751825
avg(total_load_actual)	0.218531
avg(generation_solar)	-0.554381
avg(generation_geothermal)	NaN
avg(generation_other)	0.314799
avg(price_actual)	-0.419002
avg(generation_other_renewable)	-0.196771
avg(generation_nuclear)	0.187460
avg(generation_hydro_pumped_storage_consumption)	0.925689
avg(generation_biomass)	-0.071424
avg(generation_marine)	NaN
avg(generation_waste)	-0.696604
avg(generation_fossil_hard_coal)	-0.486231
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	0.670355
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	1.000000
avg(generation_fossil_oil)	-0.159446

	avg(generation_fossil_oil)
month	-0.009340

avg(generation_fossil_gas)	0.339675
avg(generation_fossil_brown_coal_lignite)	0.429313
avg(generation_hydro_pumped_storage_aggregated)	NaN
avg(generation_fossil_peat)	NaN
avg(generation_wind_offshore)	NaN
avg(generation_hydro_water_reservoir)	-0.349102
avg(total_load_actual)	0.626397
avg(generation_solar)	0.294699
avg(generation_geothermal)	NaN
avg(generation_other)	0.591783
avg(price_actual)	0.324440
avg(generation_other_renewable)	-0.226934
avg(generation_nuclear)	0.268101
avg(generation_hydro_pumped_storage_consumption)	-0.376654
avg(generation_biomass)	0.850527
avg(generation_marine)	NaN
avg(generation_waste)	0.377811
avg(generation_fossil_hard_coal)	0.471489
avg(generation_fossil_coal_derived_gas)	NaN
avg(generation_hydro_run_of_river_and_poundage)	-0.409427
avg(generation_fossil_oil_shale)	NaN
avg(generation_wind_onshore)	-0.159446
avg(generation_fossil_oil)	1.000000

[24 rows x 24 columns]