

Data analysis tasks with python

Electricity consumption analysis

- Implementation: in python 3.9, environment according to https://github.com/DiLTAnalytics/default_model_environment/
- **steps**
 - write a python script for a standardized load profile analysis of electricity meter measurement data (kilowatt, 15-minute intervals)
 - Different grip operators provide load profiles of big buildings in different formats (3 examples attached) - prepare a standardized, modular analysis using python for the raw data including proper documentation (e.g. jupyter notebook, readme) for all the steps.
 - The results should be available as well as CSV and the plots as JPGs.
- **First step: think about meaningful plausibility checks, e.g. create scatter plot for outlier detection, etc.**
 - Make a suggestion of 2 plausibility checks and present it to me

Screenshots of raw data (3 csv files)

MEC	AT008000087000000000000000097058
UNIT	kWh
TGRID	15
Zeitbereich	01.01.2019 00:01 - 01.01.2020 00:00
Summe im Zeitraum	50.743,050 kWh
Max im Zeitraum	21,920 kW
01.01.2019 00:15	1,3
01.01.2019 00:30	0,77
01.01.2019 00:45	0,8
01.01.2019 01:00	0,78
01.01.2019 01:15	0,78
01.01.2019 01:30	0,76
01.01.2019 01:45	0,82
01.01.2019 02:00	0,8
01.01.2019 02:15	0,73
01.01.2019 02:30	0,83
01.01.2019 02:45	0,77
01.01.2019 03:00	0,79
01.01.2019 03:15	0,77
01.01.2019 03:30	0,75
01.01.2019 03:45	0,76
01.01.2019 04:00	0,89

Anlage					
Vertragsbeze					
Anlagenadre					
GP-Nummer					
Geschäftspar					
GP-Adresse					
VP-Nummer					
Vertriebsspa					
Zählpunkt					
Netzebene			6		
Netzgebiet			N08170		
Service			NE		
Branche					
Profil			10000308		
Profilbezeich			OP_L_AT00817008605000000000000000		
Profilrolle			1001		
Profiltyp/ME			01/kW		
Maximum			89,52		
Summe kW			453.685,20		
Summe kWh			113.421,30		
Profilwert	01.01.2020	00:00	6,72		
Profilwert	01.01.2020	00:15	6,8		
Profilwert	01.01.2020	00:30	6,64		
Profilwert	01.01.2020	00:45	6,72		
Profilwert	01.01.2020	01:00	6,72		
Profilwert	01.01.2020	01:15	6,96		
Profilwert	01.01.2020	01:30	7,6		
Profilwert	01.01.2020	01:45	7,92		
Profilwert	01.01.2020	02:00	7,28		
Profilwert	01.01.2020	02:15	6,56		
Profilwert	01.01.2020	02:30	6,96		
Profilwert	01.01.2020	02:45	6,8		
Profilwert	01.01.2020	03:00	7,6		

	A	B	C	D
1	Date	Time	1-1:1.5.0	
2	01.01.2019	00:15	21,76	
3	01.01.2019	00:30	20,64	
4	01.01.2019	00:45	21,84	
5	01.01.2019	01:00	20,96	
6	01.01.2019	01:15	21,12	
7	01.01.2019	01:30	20,32	
8	01.01.2019	01:45	21,84	
9	01.01.2019	02:00	22,24	
10	01.01.2019	02:15	21,6	
11	01.01.2019	02:30	19,04	
12	01.01.2019	02:45	20,32	
13	01.01.2019	03:00	22,16	
14	01.01.2019	03:15	21,68	
15	01.01.2019	03:30	21,84	

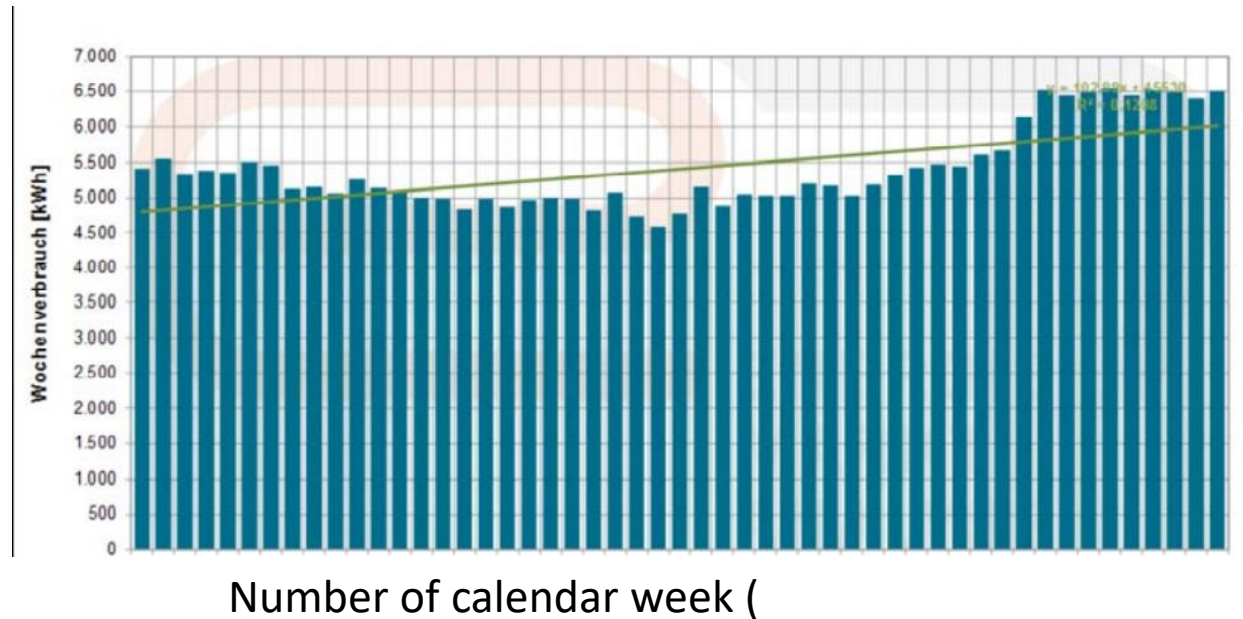
Create a report per dataset

Description of analysis	Result of analysis (write all results in a CSV)
Name of dataset	Name of analysed CSV
Time period of data set	dd.mm.yyyy-dd.mm.yyyy
Count amount of kWh values	1 integer (should be ~35000)
Total energy consumption (=sum of all values)	1 Decimal (kWh)
Maximum value in dataset	1 Decimal value (kWh) and corresponding timestamp
Minimum value in dataset	1 Decimal value (kWh) and corresponding timestamp
Average and median of daily sum in dataset	2 Decimal values (kWh) and standard deviation
Average and median of daily sum on weekdays	2 Decimal values (kWh) and standard deviation
Average and median of daily sum on Saturdays and Sundays	2 Decimal values (kWh) and standard deviation
Average and median of on Sundays at 03:00 in the morning	2 Decimal values (kWh) and standard deviation
Average daily sum between 08:00-17:00 on weekdays	1 Decimal value (kWh) and standard deviation
Average daily sum between 08:00-17:00 on Saturdays and Sundays	1 Decimal value (kWh) and standard deviation
Average daily sum between 17:00-08:00 on weekdays	1 Decimal value (kWh) and standard deviation
Average daily sum between 17:00-08:00 on Saturdays and Sundays	1 Decimal value (kWh) and standard deviation

Create 3 bar charts

- **Bar chart 1:** daily sums in kWh (y-axis), date on x-axis
- **Bar chart 2:** weekly sums in kWh (y-axis), number of calendar week on x-axis
- **Bar chart 3:** Monthly sums in kWh (y-axis), month (mm.yy) on x-axis

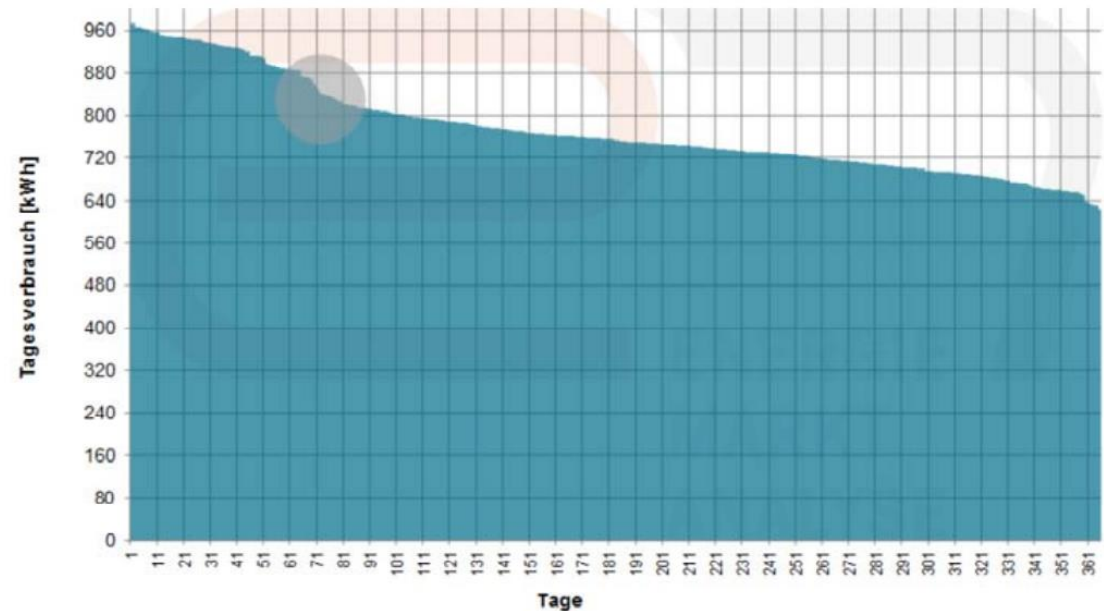
Example how bar chart 2 could look like



Create 2 bar charts

- **Sort values of bar chart 1 (slide before) by size (descending):** daily sums in kWh (y-axis), number of days on x-axis
- **Sort values of bar chart 2 (slide before) by size (descending):** weekly sums in kWh (y-axis), first day of related week on x-axis

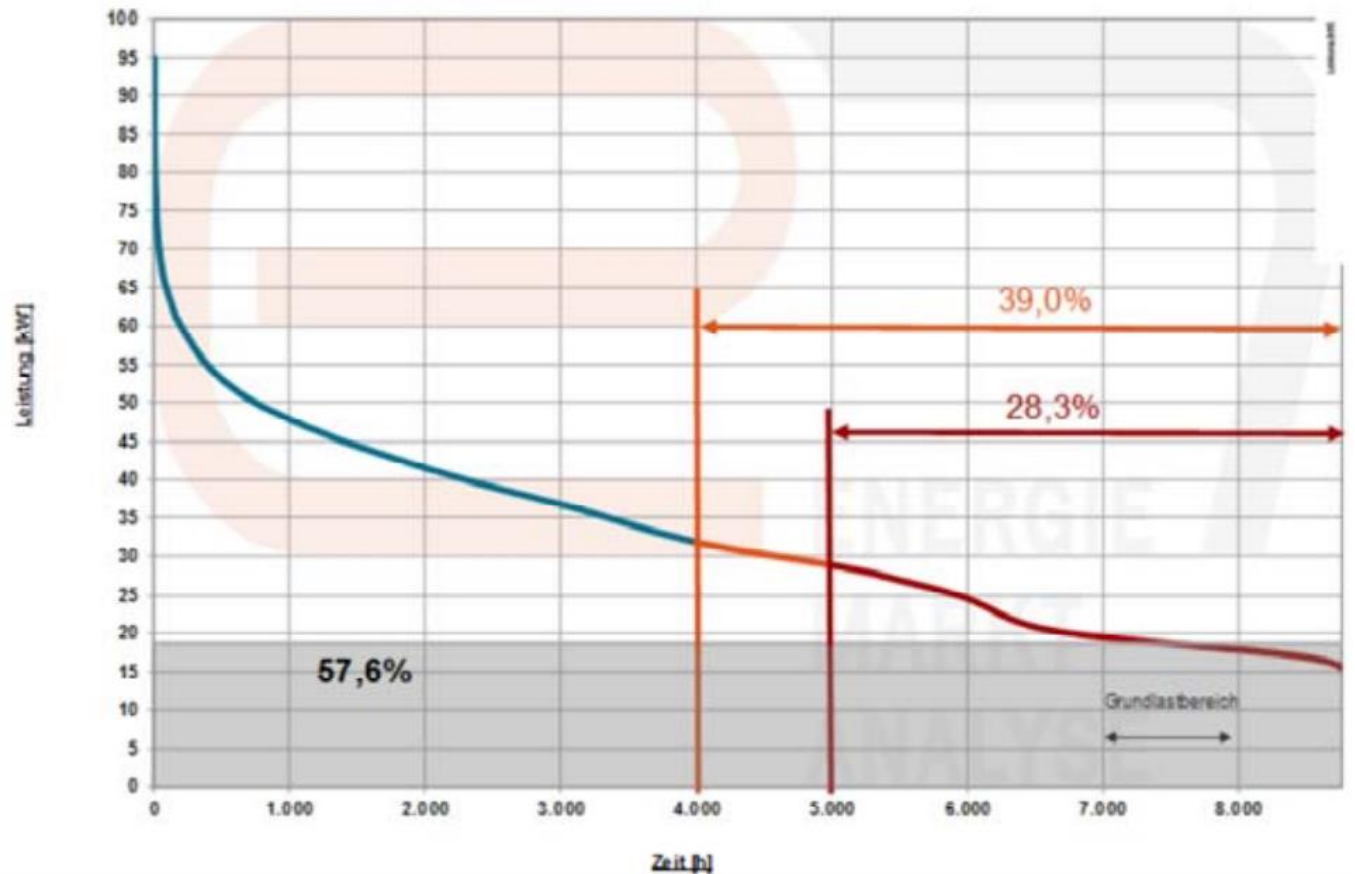
Example how bar chart 1 could look like



Create a line diagram

- Multiply each 15-min-value with 4 (converts kWh in kW)
- Create a plot which arranges the kW values by size (y-axis, descending) and time in hours on x-axis

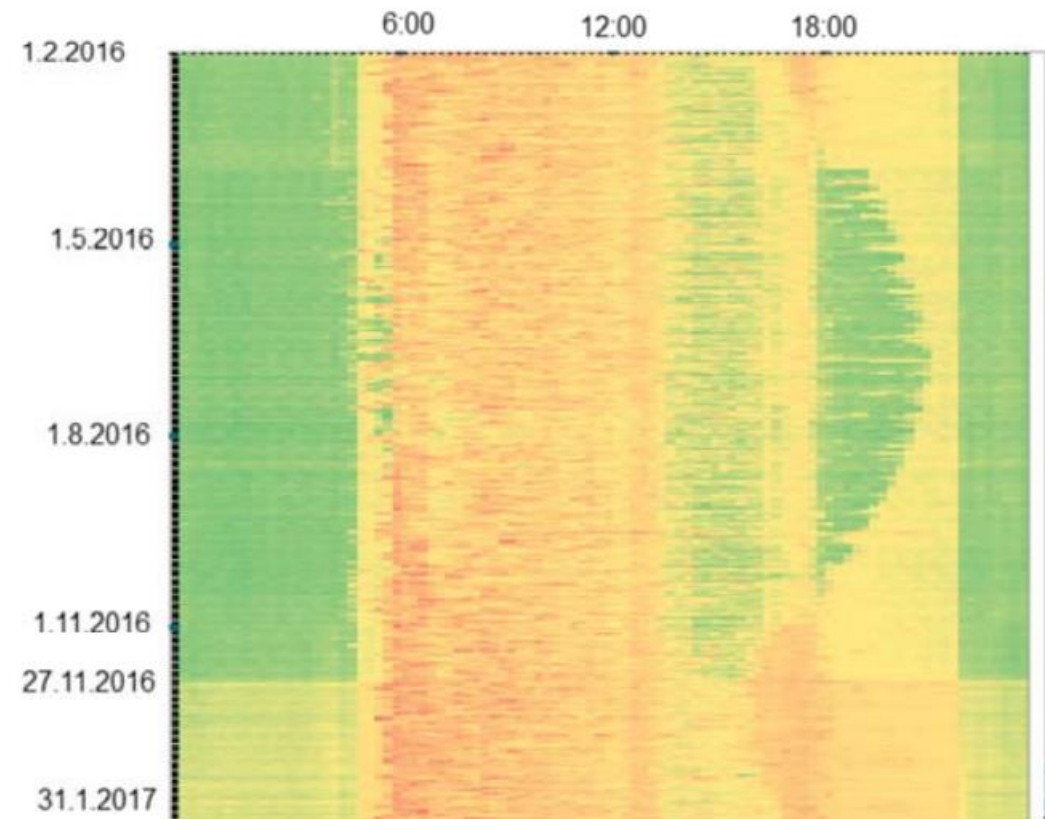
Example plot how it could look like (blue, orange & red line, don't take care about the labels in the plot)



Create a carpet plot

- Create a carpet plot or heat map (kWh values on y-axis, hour of the day (e.g. 06:00, 12:00, 18:00) on x-axis)

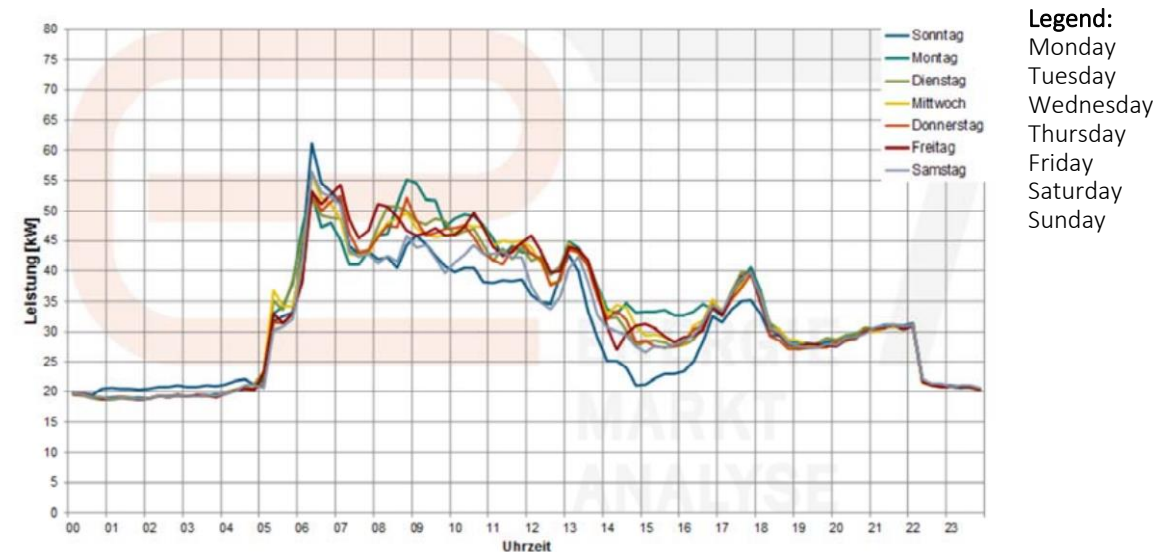
Example how the plot should look like



Create a line diagram of the average daily profile per day of the week

- Multiply each 15-min-value with 4 (converts kWh in kW)
- Calculate the average 15-minute value per weekday and time of the day (*e.g. on all Mondays at 08:00 the average value is 45 kW*)
- Create a line diagram out of the calculated values (kWh values on y-axis, hour of the day (*e.g. 06:00, 12:00, 18:00*) on x-axis which shows the a line diagram per day

Example how the plot could look like



Hour of the day (on y-axis)