Building_ID	Temperature_C	Occupancy	Energy_Consumption_kWh	Humidity_%	CO2_Level_ppm	Motion_Sensor
7	25.7	180	126.26	63.3	510.3	1
4	20.8	41	98.96	70	379.8	1
8	20.7	185	119.65	57.2	512.7	1
5	23.1	422	170.02	65.7	638.5	1
7	27.8	397	179.08	38.2	612	1
10	22.2	222	130.71	43.2	480.8	1
3	20.8	121	102.19	33.2	475.8	1
7	24.6	132	120.08	43.9	465.2	1
8	24.2	162	115	57.6	443.1	1
5	19.1	214	118.2	52.1	526.5	1
4	20.4	220	120.36	58.7	531.7	0
8	25.3	234	130.07	40.6	488.4	0
8	23.3	330	152.82	41.2	587.6	1
3	14.6	145	99.61	52	493.3	1
6	23.4	238	115.31	44	537.5	1
5	24.8	75	100.86	63	450	1
2	19.5	8	78.26	43.7	364.3	1
8	24	73	105.06	45	447.4	1
6	27.6	441	176.78	54.2	617.2	1
2	23.2	491	195.7	46.1	660.4	1
5	20.1	400	167.13	37.7	603.8	1
1	20.4	252	129.05	39	525.7	1
10	22	229	134.55	55.4	524.1	1
6	23.1	6	88.48	36.4	400.4	0
9	22.2	173	119.07	64	486.1	0
1	23.5	140	117.26	43.1	461.7	0
10	21.6	167	116.07	56.2	459.4	1
3	22.5	169	118.75	43.7	493	1
7	24.9	392	169.87	40.8	583.9	1
4	18.8	433	175.06	52.5	613.5	1
9	22.3	282	141.11	60.5	545.3	0
6	22.3	121	105.38	47.5	429.2	1
5	24.8	193	115.65	49.3	468.6	0
3	21.3	4	92.56	58.1	397.5	1
7	24	28	90.28	58.4	420.8	0
5	25.3	164	123.42	49.6	493.2	1
9	16.5	421	168.48	45.6	586.6	0
7	25.5	338	156.99	50.8	578.5	1
2	19	135	104.53	48.1	463.8	1
4	18.7	495	184.15	53.1	634.3	1
9	13.4	364	145.2	38.1	572.3	1
2	19.6	320	146.33	52.8	572	1
10	22	341	143.58	42.8	574.6	1
7	20.9	499	187.18	59.1	650.3	1
10	20.6	144	111.58	43.3	475.7	1
5	18.7	326	136.39	43.4	576.9	1
2	24.6	216	132.53	54.5	506.5	1
4	24.9	300	137.04	50.1	550.8	0
7	23.3	131	122.86	72.8	482.6	1
8	16.3	291	127.14	33.7	540.5	1
3	28.8	69	115.06	51.1	443.8	1
1	22.4	251	129.74	61.3	517.9	1
4	23.6	414	171	44.2	600.8	1
2	17.5	274	126.49	40.9	535.4	1
8	22	444	176.79	41.3	629.7	0
4	23.2	363	156.22	33.9	599.4	1
2	21.6	181	120.94	63.5	491.9	1
5	24.6	166	120.18	46.7	460.5	1
6	25.9	90	100.02	45.5	454.5	1
10	18	201	119.72	28.8	537.3	0
4	24.6	345	153.38	35.7	574.1	1
6	23.2	18	93.63	64.7	383.4	1
2	22	38	89.42	53	421.5	1
10	17.3	125	100.57	61.7	462.5	1
2	26.4	450	188.01	60.4	623.7	1
10	23.2	172	131.11	40	500.2	1
3	16.7	140	106.98	74.3	490.7	1
8	21.3	241	128.71	69.5	530.8	0
7	23	219	126.43	44.4	507.5	1
9	23.5	125	106.16	14.3	479.4	0
8	21.6	57	90.67	46.4	406.1	1
5	23	147	112.53	52.4	462.8	1
5	28.4	475	197.42	44.3	628.4	1
5	26.6	455	179.91	47.8	627	1
8	19.2	316	135.87	41.7	565.6	0
10	20.2	382	156.91	59.2	588	1
9	21.7	360	155.44	50.6	566.4	1
9	16.8	422	159.78	59.5	625.6	1
1	18	0	81	62.2	396.5	1
9	17.1	386	153.2	28.3	594.5	0
7	18.6	460	173.74	42.3	638.3	1
9	26.2	347	154.8	43.8	574.1	1
9	19.4	189	125.8	75.7	495.1	1
1	18.9	190	117.61	51.5	510.5	0
8	24.6	368	162.5	58.1	586.3	0
8	20.7	408	158.73	31.2	616.9	1
8	23.3	311	149.51	51.4	563.7	1
1	19.8	416	166.16	59.8	638.3	1
8	16.6	421	153.25	58.6	616.6	1
3	23.4	116	100	56.2	467.6	1
4	17.3	133	104.02	50.1	483	0
1	20.5	57	89.71	50.6	429.5	1
5	19.8	43	89.6	33.1	411.9	1
10	22.6	172	111.84	41	478.9	1
7	20.3	159	104.45	52.3	472.5	0
10	21.6	172	117.83	42.8	471.7	1
9	17.3	316	137.4	56.9	562.8	1
7	28	302	141.85	54.7	541.6	1
9	18.4	148	113.92	63.7	476.9	1
8	25.6	79	97.66	59.6	437.1	1
2	22.4	373	171.07	52.3	585.9	1
1	22.4	212	122.88	57.5	514.3	1
7	23.1	202	127.06	60.1	492.9	1
7	16.4	251	133.03	44.9	535.1	1
8	24.9	228	140.35	46.7	537.5	1
5	20.5	163	123.28	49	486.6	1
3	25.6	226	125.98	45.2	501.3	0
8	21.5	146	105.77	60.2	463.8	1
6	20.9	19	86.28	43.7	405.8	0
3	18.1	440	169.35	62.5	611.1	0
1	19.7	401	159.69	46.3	585.6	0
1	23.3	46	85.96	54.7	425.3	1
5	18.1	232	126.03	51.7	512.7	0
3	16	304	133.83	40.2	554.8	0
1	19.4	13	80.67	49.8	424.3	0
5	24.5	142	103.76	53.3	467.5	0
10	27.3	414	177.74	49.4	589.2	1
7	20	0	80.73	36.6	394	1
7	17.4	372	159.49	67	597.6	1
9	21	53	91.73	50.9	436.7	1
10	22.6	373	161.33	42.1	594.3	1
10	23.7	258	141.31	66.2	519.6	0
3	28.4	143	112.47	56.2	482.8	1
7	22.7	470	191.13	38	649.7	1
1	27	440	176.35	39.6	609.6	1
4	20.6	458	174.97	47.8	628.4	0
4	22.3	11	88.85	53	435.8	1
5	20.2	329	142.03	56.2	553.4	1
7	20	223	133.83	59.5	485.9	1
7	21.2	271	137.33	55.4	556.2	0
1	21.2	455	171.73	63.5	644.5	0
7	20.2	357	152.39	51.3	571.6	0
3	20.4	459	176.22	42.4	627.8	1
6	25.3	407	169.4	50	585.1	1
2	28.3	155	127.86	49.3	478.8	0
10	20.5	372	146.73	24.2	603.5	0
9	21	7	79.16	44.4	395.4	1
5	23.6	121	98.91	53.1	473.7	0
6	31.6	347	176.79	45.8	580.5	0
4	22.3	163	125.97	43.5	490.2	1
10	24.9	89	108.28	59.1	437.2	0
7	26.6	135	111.84	40.8	462.2	1
9	23	185	133.45	57.1	464.8	1
7	24.8	315	147.18	36.7	549	1
1	23.8	177	120.03	77.6	506	0
1	25.1	27	93.85	47.8	399.6	0
9	18.9	219	124.41	46.8	497.8	1
9	22.6	356	158.32	61.4	587.6	1
8	23.4	40	96.52	45.8	426	1
9	23.4	227	129.78	64.5	523.8	0
3	25.4	191	129.64	53.7	478.1	1
7	23.3	410	166	48.8	643.5	1
8	18.6	446	170	47.7	604.8	1
8	23.8	144	116.99	62.1	474.3	1
9	19.4	200	122.91	52.3	497.6	1
5	19.9	416	166.24	54.3	590.2	1
1	20.7	211	121.04	55.4	499.9	1
3	21.7	460	176.84	51.3	630.3	1
10	17.4	219	123.17	55.2	499.8	0
8	25.4	239	129.44	48.8	533.6	1
8	18.5	412	158.78	49.5	603.6	1
8	26.7	396	172.29	38.8	577.9	1
9	18.6	45	95.45	53	401.1	1
4	25.8	34	99.62	32.8	426.1	1
1	21.9	252	142.11	52.8	503.7	0
1	16.9	389	151.16	47.2	599.2	0
10	18.9	81	95.8	53.7	447.2	0
4	26	370	167.03	39.5	570.1	1
7	23.3	255	124.32	28.8	540.9	1
2	21.8	196	129.45	46.5	472.2	1
3	15.3	302	124.05	39.5	555.4	1
1	26.2	408	175.88	32.2	603.4	1
3	27.6	449	184.63	57.7	622.7	1
1	27.1	9	80.75	66	386.8	1
8	20.2	311	131.05	42	541.4	1
1	23.1	285	148.81	49.2	566.8	1
1	23.5	241	139.52	46.9	538.7	0
2	17.4	250	121.7	43.4	536.1	1
6	22.5	364	152.14	56.6	596.1	1
6	25.2	4	91.16	60.1	414.7	1
7	24.3	118	109.99	41.3	454.1	1
5	18.7	288	140.92	50.4	544.7	1
1	19.4	373	156.1	62.8	564.4	1
1	24.4	64	107.32	56.9	430.7	1
3	23.9	145	119.7	58.4	490.3	1
2	15.2	223	121	63.7	501.6	1
5	14.6	238	117.74	63.5	534.3	1
10	21.8	176	115.77	60.3	489.4	0
6	20.2	266	133.24	51.4	512.2	1
7	21.1	340	158.85	46.2	553.4	0
4	20.8	281	135.37	51	554.3	1
7	25	62	101.2	44.7	439.3	1
8	22.3	216	115.43	30.5	509.5	1
1	22.8	341	155.78	47.2	566.4	1
6	18	314	151.09	53.6	560.2	1
8	11.9	282	119.58	44.1	554.2	1
5	25.4	176	139.17	48.1	510.1	1
4	28.4	460	185.51	48.5	652.9	1
2	21	416	151.84	63.6	613.6	0
6	22.7	97	101.52	46.4	445.9	0

Problem 1: Understanding the Data

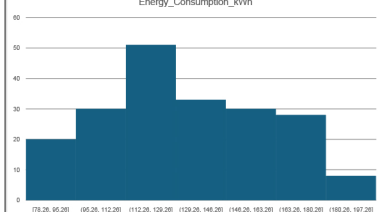
Numerical variables: Temperature_C [Continuous numerical values]
Occupancy [Integer counts]
Energy_Consumption_kWh [Continuous numerical values]
Humidity_% [Continuous numerical values]
CO2_Level_ppm [Continuous numerical values]

Categorical variables: Building_ID [Number- but serves as an identifier rather than a quantity]
Motion_Sensor [Binary categorial variable]

Summary Statistics:	Temperature_C	Occupancy	Energy_Consumption_kWh	Humidity_%	CO2_Level_ppm
Mean:	21.94	243.15	132.37	50.17	522.13
Median:	22.20	228.50	128.88	50.10	518.75
Standard Deviation:	3.29	136.58	26.67	10.02	70.64
Min:	11.90	0.00	78.26	14.30	364.30
Max:	31.60	499.00	197.42	77.60	660.40
Missing Values:	N/A	N/A	N/A	N/A	N/A
Extreme Values:	N/A	N/A	N/A	N/A	N/A

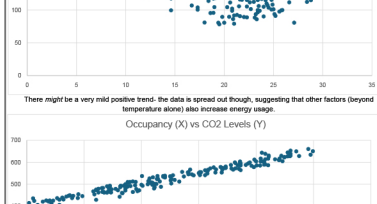
Not many observations to make here- I don't think there are any missing/extreme values. The only values that would seem to near extreme would be values on the higher end of the building data but the extreme nature of them would depend entirely on context that I don't have.

Problem 2: Analyzing Energy Consumption Trends

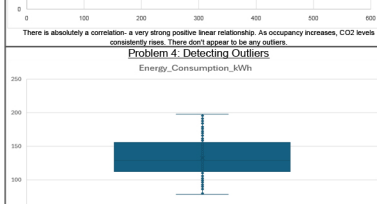


The distribution looks unimodal with a slight right skew. There is one main peak around values (112.26, 129.26] and there don't seem to be any significant peaks or unusual patterns outside of that.

Problem 3: Identifying Relationships Between Variables

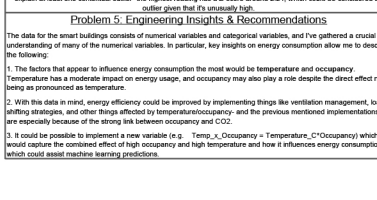


There *might* be a very mild positive trend- the data is spread out though, suggesting that other factors (beyond temperature alone) also increase energy usage.



There is absolutely a correlation- a very strong positive linear relationship. As occupancy increases, CO2 levels consistently rises. There don't appear to be any outliers.

Problem 4: Detecting Outliers



There aren't any points that lie beyond the whiskers, so there aren't any outliers. If there were, it could be a result of things like faulty meter or sensor readings, unusually high occupancy, equipment malfunction, etc. Using Z-score to explain at least one contextual outlier- the highest reading has a Z-score around 2.24, which *could* be considered an outlier given that it's unusually high.

Problem 5: Engineering Insights & Recommendations

The data for the smart buildings consists of numerical variables and categorical variables, and I've gathered a crucial understanding of many of the numerical variables. In particular, key insights on energy consumption allow me to describe the following:

1. The factors that appear to influence energy consumption the most would be **temperature** and **occupancy**. Temperature has a moderate impact on energy usage, and occupancy may also play a role despite the direct effect not being as pronounced as temperature.

2. With this data in mind, energy efficiency could be improved by implementing things like ventilation management, load-shifting strategies, and other things affected by temperature/occupancy- and the previous mentioned implementations are especially because of the strong link between occupancy and CO2.

3. It could be possible to implement a new variable (e.g. `Temp_x_Occupancy = Temperature_C*Occupancy`) which would capture the combined effect of high occupancy and high temperature and how it influences energy consumption- which could assist machine learning predictions.