

アメダス気象データ分析 チャレンジ！入門

5. データ分析の例（気象×電力）

主催：気象ビジネス推進コンソーシアム

共催：岐阜大学工学部附属応用気象研究センター

資料作成：吉野 純（岐阜大学）



本教材について

Copyright 2022 気象ビジネス推進コンソーシアム、岐阜大学 吉野純

(C) 2022 WXBC、岐阜大学 吉野純

<利用条件>

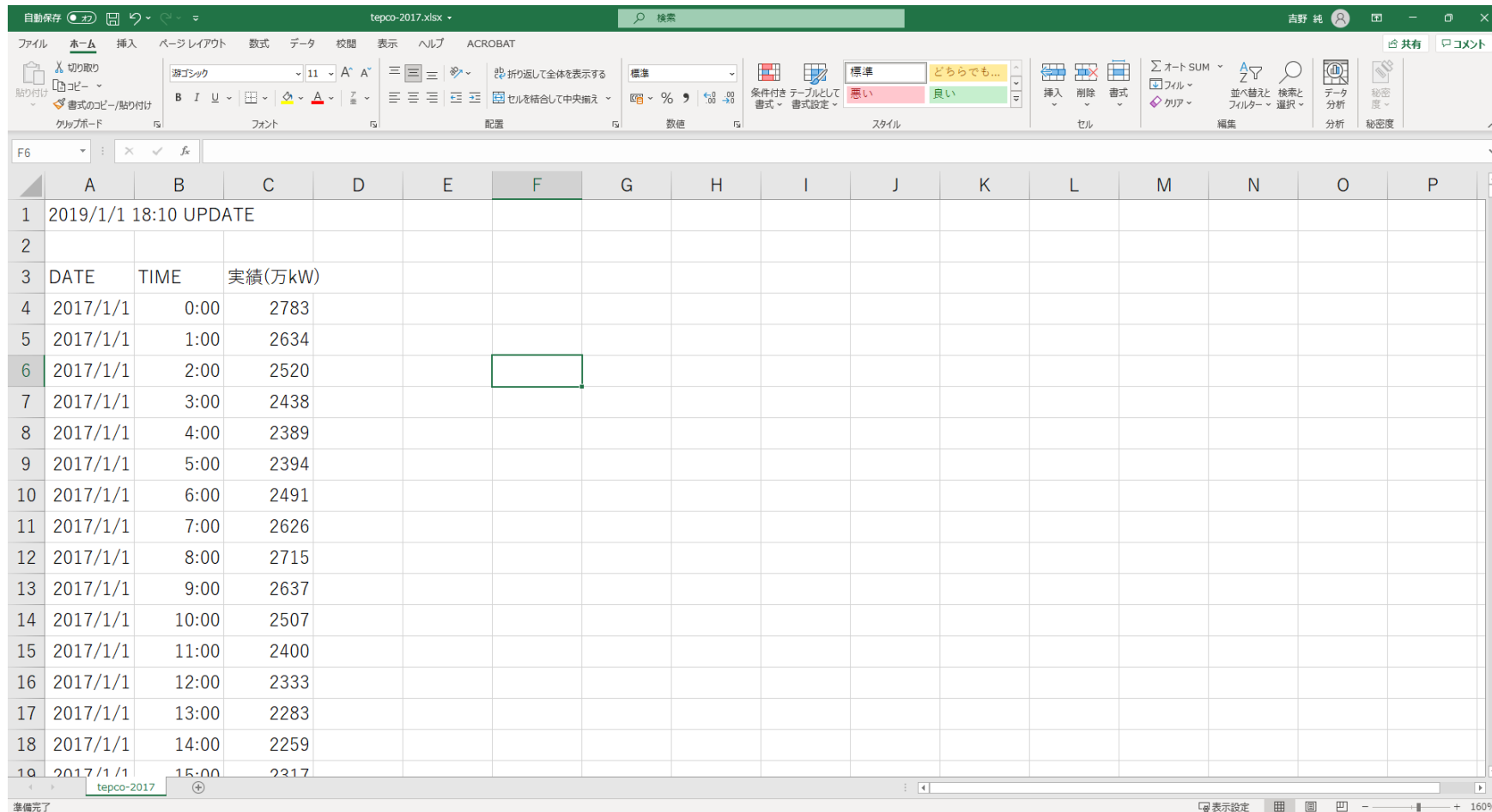
本書は、本書に記載した要件・技術・方式に関する内容が変更されないこと、および出典を明示いただくことを前提に、無償でその全部または一部を複製、翻案、翻訳、転記、引用、公衆送信等して利用できます。なお、全体を複製、翻案、翻訳された場合は、本書にある著作権表示および利用条件を明示してください。

<免責事項>

本書の著作権者は、本書の記載内容に関して、その正確性、商品性、利用目的への適合性等に関して保証するものではなく、特許権、著作権、その他の権利を侵害していないことを保証するものでもありません。本書の利用により生じた損害について、本書の著作権者は、法律上のいかなる責任も負いません。

仮説 1: 電力と気温の関係性

ここで、東京電力の電力消費量と東京の気温との間には関係性があるのではないかと仮説を立てて分析をしてみましょう。まず、電力データ「tepco-2017.xlsx」を開いてみましょう。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	2019/1/1 18:10 UPDATE															
2																
3	DATE	TIME	実績(万kW)													
4	2017/1/1	0:00	2783													
5	2017/1/1	1:00	2634													
6	2017/1/1	2:00	2520													
7	2017/1/1	3:00	2438													
8	2017/1/1	4:00	2389													
9	2017/1/1	5:00	2394													
10	2017/1/1	6:00	2491													
11	2017/1/1	7:00	2626													
12	2017/1/1	8:00	2715													
13	2017/1/1	9:00	2637													
14	2017/1/1	10:00	2507													
15	2017/1/1	11:00	2400													
16	2017/1/1	12:00	2333													
17	2017/1/1	13:00	2283													
18	2017/1/1	14:00	2259													
19	2017/1/1	15:00	2217													

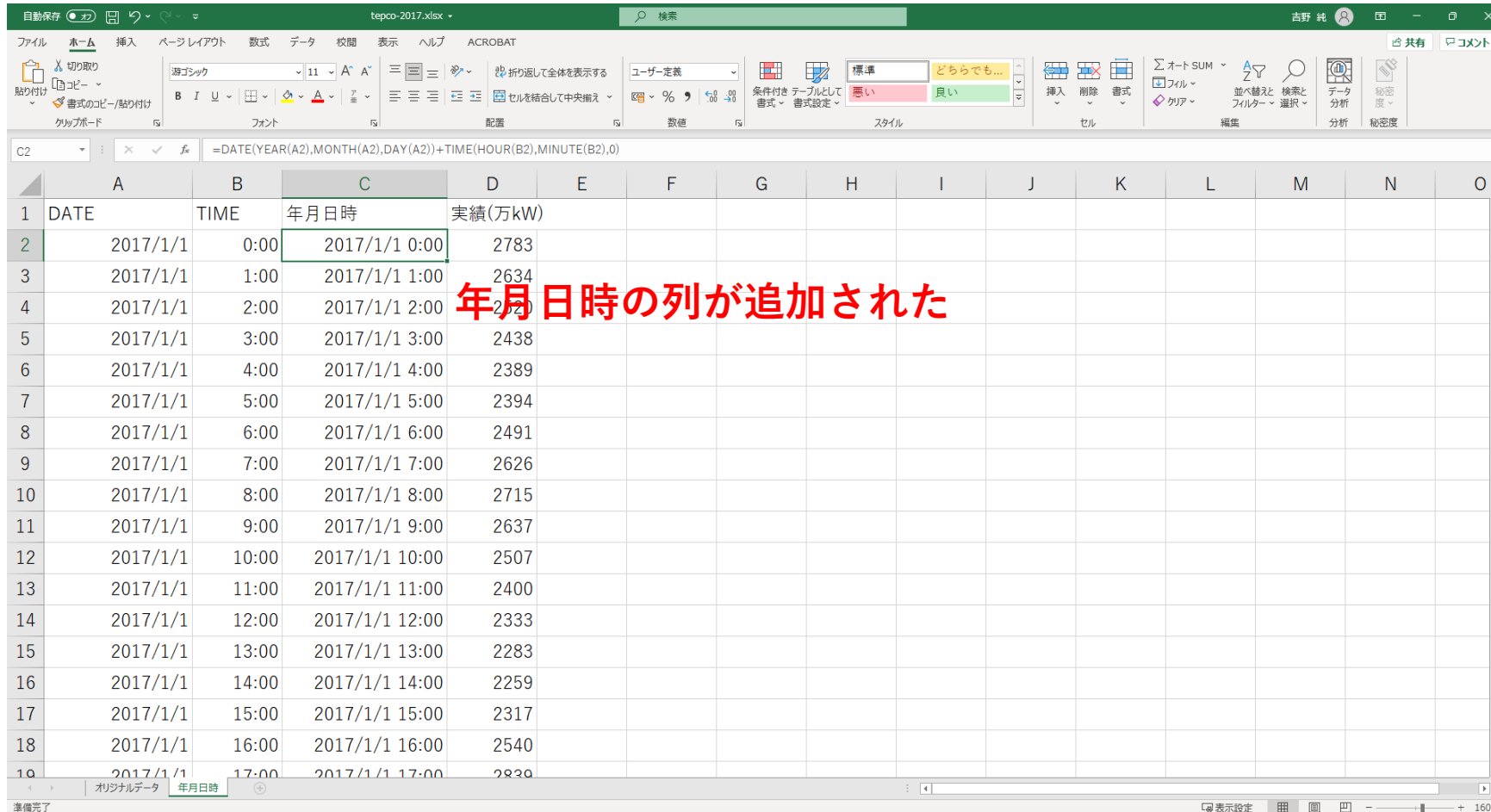
5_data/tepco-2017.xlsxに一連の処理が記録されています

仮説 1 : 電力と気温の関係性

DATE (A列) と TIME (B列) のデータから、DATE関数とTIME関数から年月日時のデータを作成しましょう。

=DATE(YEAR(A2),MONTH(A2),DAY(A2))+TIME(HOUR(B2),MINUTE(B2),0)

年 月 日 時間 分 秒

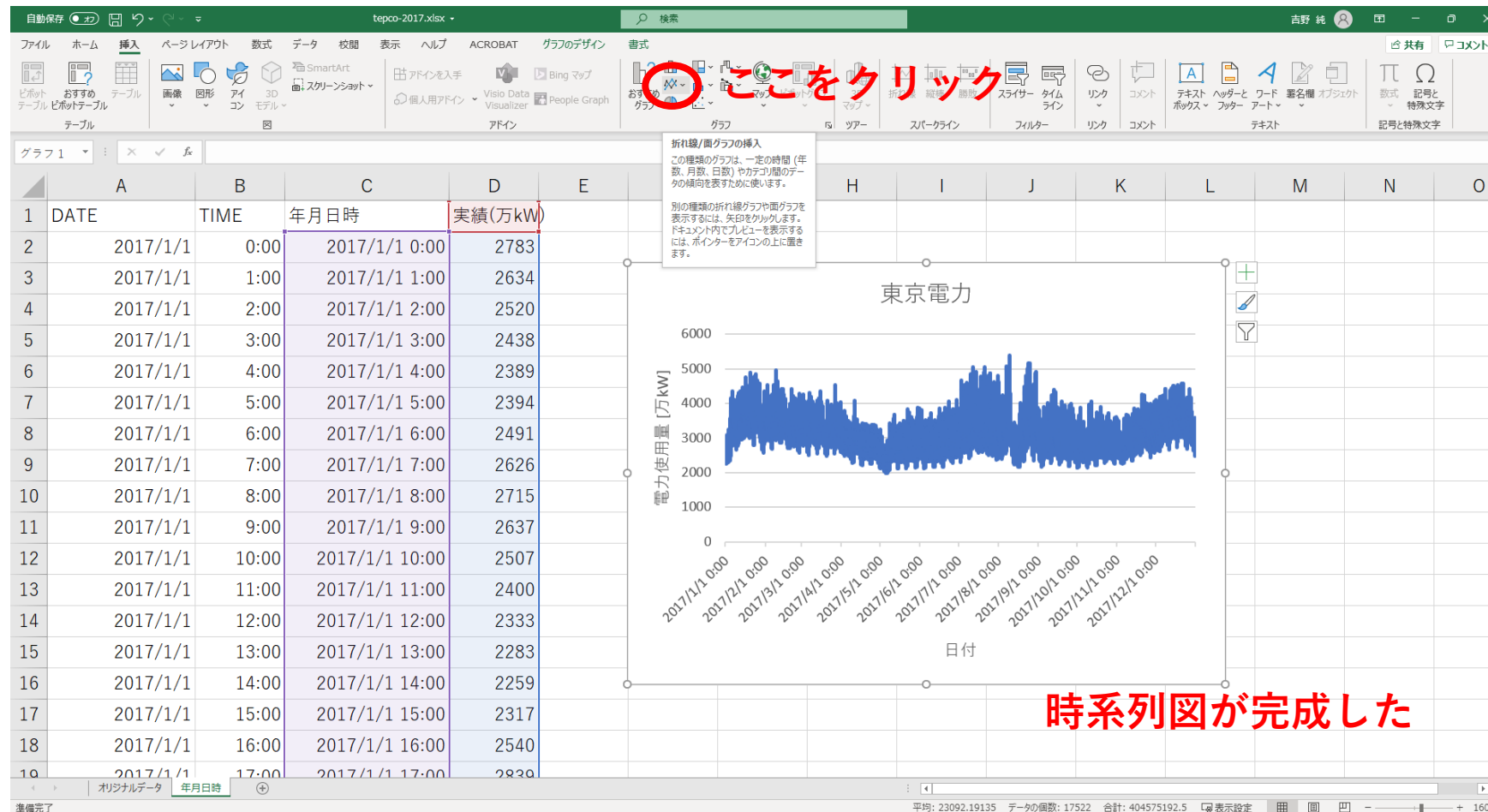


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	DATE	TIME	年月日時	実績(万kW)											
2	2017/1/1	0:00	2017/1/1 0:00	2783											
3	2017/1/1	1:00	2017/1/1 1:00	2634											
4	2017/1/1	2:00	2017/1/1 2:00	2429											
5	2017/1/1	3:00	2017/1/1 3:00	2438											
6	2017/1/1	4:00	2017/1/1 4:00	2389											
7	2017/1/1	5:00	2017/1/1 5:00	2394											
8	2017/1/1	6:00	2017/1/1 6:00	2491											
9	2017/1/1	7:00	2017/1/1 7:00	2626											
10	2017/1/1	8:00	2017/1/1 8:00	2715											
11	2017/1/1	9:00	2017/1/1 9:00	2637											
12	2017/1/1	10:00	2017/1/1 10:00	2507											
13	2017/1/1	11:00	2017/1/1 11:00	2400											
14	2017/1/1	12:00	2017/1/1 12:00	2333											
15	2017/1/1	13:00	2017/1/1 13:00	2283											
16	2017/1/1	14:00	2017/1/1 14:00	2259											
17	2017/1/1	15:00	2017/1/1 15:00	2317											
18	2017/1/1	16:00	2017/1/1 16:00	2540											
19	2017/1/1	17:00	2017/1/1 17:00	2839											

5_data/tepcO-2017.xlsxに一連の処理が記録されています

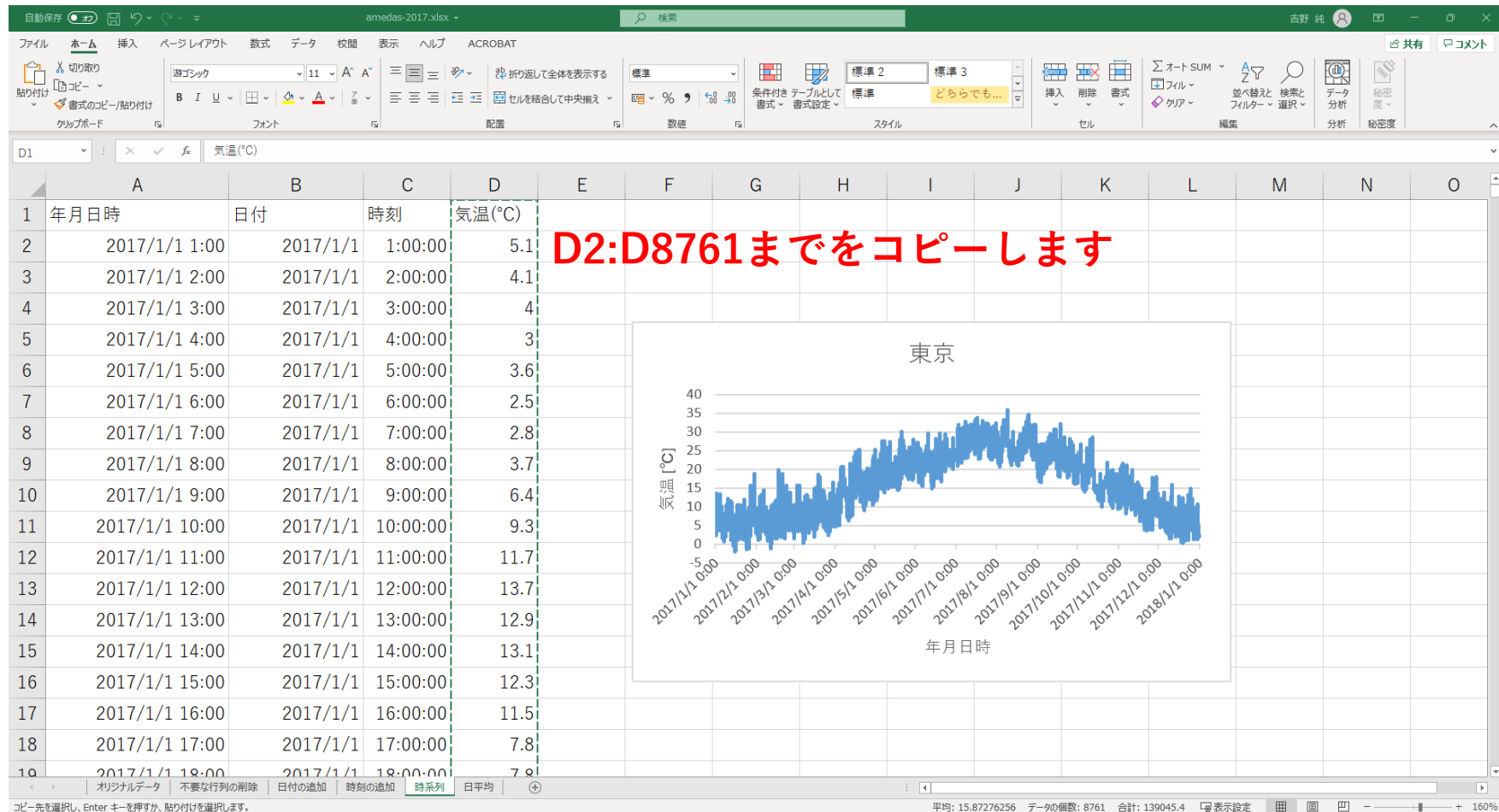
仮説 1 : 電力と気温の関係性

年月日時（C列）と電力データ（D列）を選択して、時系列図を作成しましょう。



仮説 1 : 電力と気温の関係性

再び、気象データ「amedas-2017.xlsx」を開いてみましょう。気象データは、2017年1月1日1時～2018年1月1日1時までのデータがあり、電力データは2017年1月1日0時～2018年1月1日0時までのデータがあり、データの範囲が1時間ずれていることがわかります。データをコピーします。



The screenshot shows an Excel spreadsheet with the following data in columns A through D:

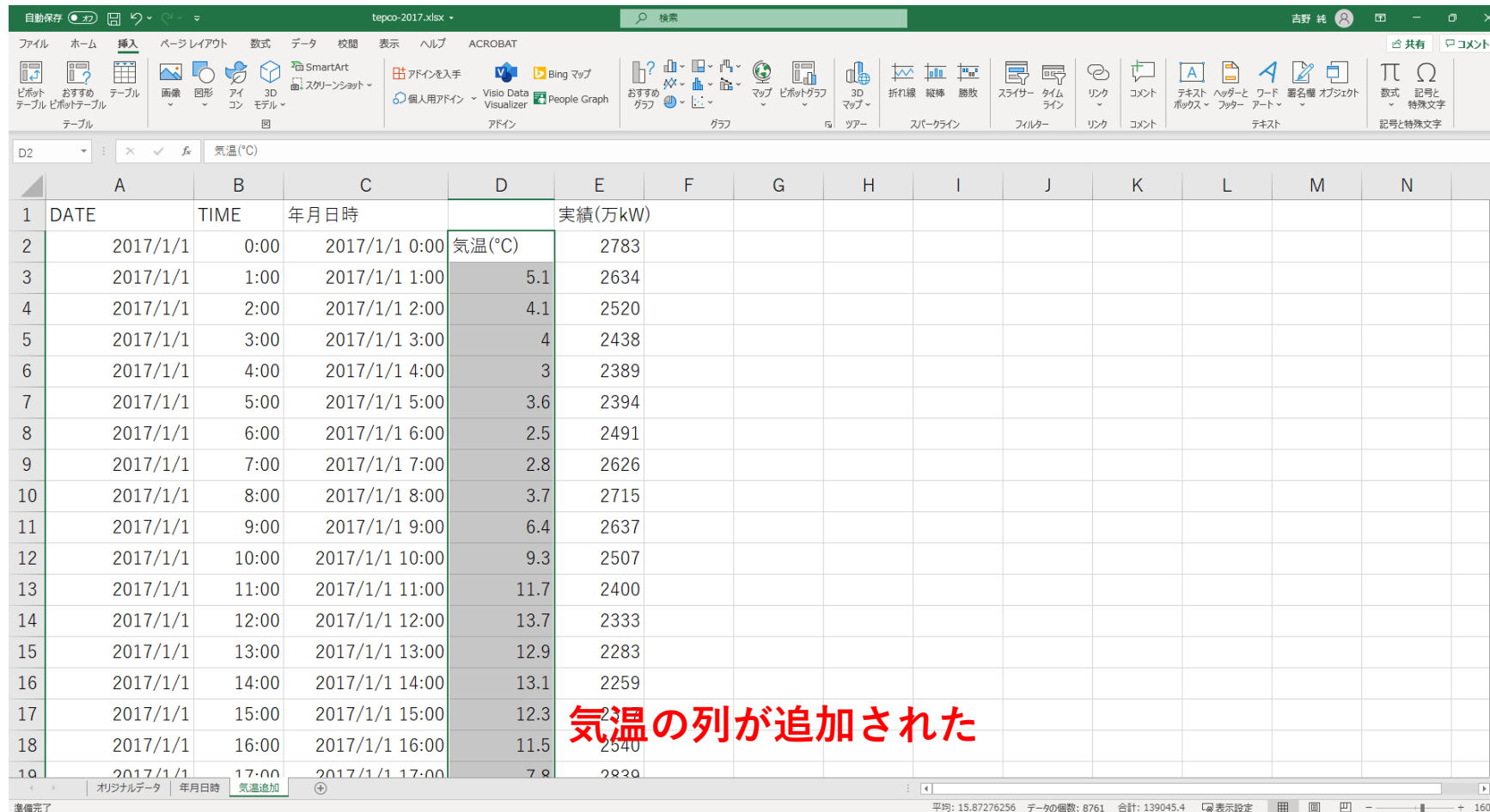
1	年月日時	日付	時刻	気温(°C)
2	2017/1/1 1:00	2017/1/1	1:00:00	5.1
3	2017/1/1 2:00	2017/1/1	2:00:00	4.1
4	2017/1/1 3:00	2017/1/1	3:00:00	4
5	2017/1/1 4:00	2017/1/1	4:00:00	3
6	2017/1/1 5:00	2017/1/1	5:00:00	3.6
7	2017/1/1 6:00	2017/1/1	6:00:00	2.5
8	2017/1/1 7:00	2017/1/1	7:00:00	2.8
9	2017/1/1 8:00	2017/1/1	8:00:00	3.7
10	2017/1/1 9:00	2017/1/1	9:00:00	6.4
11	2017/1/1 10:00	2017/1/1	10:00:00	9.3
12	2017/1/1 11:00	2017/1/1	11:00:00	11.7
13	2017/1/1 12:00	2017/1/1	12:00:00	13.7
14	2017/1/1 13:00	2017/1/1	13:00:00	12.9
15	2017/1/1 14:00	2017/1/1	14:00:00	13.1
16	2017/1/1 15:00	2017/1/1	15:00:00	12.3
17	2017/1/1 16:00	2017/1/1	16:00:00	11.5
18	2017/1/1 17:00	2017/1/1	17:00:00	7.8
19	2017/1/1 18:00	2017/1/1	18:00:00	7.8

An embedded line chart titled "東京" (Tokyo) shows the temperature trend from 2017/1/1 0:00 to 2018/1/1 0:00. The y-axis is labeled "気温(°C)" and ranges from -5 to 40. The x-axis is labeled "年月日時" and shows hourly intervals. A red text overlay on the spreadsheet reads "D2:D8761までをコピーします" (Copy from D2 to D8761).

5_data/amedas-2017.xlsxに一連の処理が記録されています

仮説 1 : 電力と気温の関係性

データが1時間ズレていることに気をつけながら、コピーした気温データを「tepco-2017.xlsx」のD列に貼り付けます。



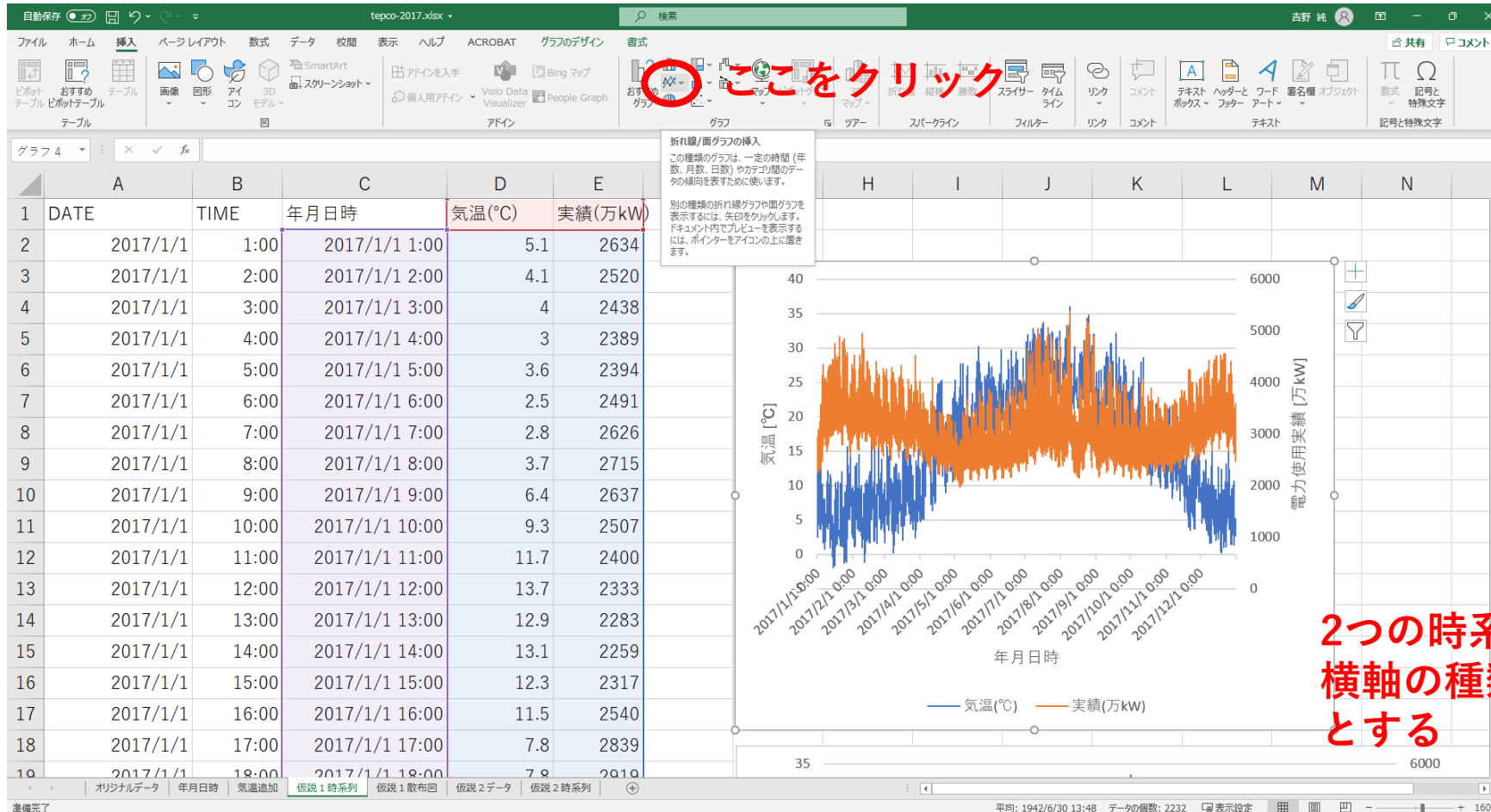
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	DATE	TIME	年月日時		実績(万kW)									
2	2017/1/1	0:00	2017/1/1 0:00	気温(°C)	2783									
3	2017/1/1	1:00	2017/1/1 1:00	5.1	2634									
4	2017/1/1	2:00	2017/1/1 2:00	4.1	2520									
5	2017/1/1	3:00	2017/1/1 3:00	4	2438									
6	2017/1/1	4:00	2017/1/1 4:00	3	2389									
7	2017/1/1	5:00	2017/1/1 5:00	3.6	2394									
8	2017/1/1	6:00	2017/1/1 6:00	2.5	2491									
9	2017/1/1	7:00	2017/1/1 7:00	2.8	2626									
10	2017/1/1	8:00	2017/1/1 8:00	3.7	2715									
11	2017/1/1	9:00	2017/1/1 9:00	6.4	2637									
12	2017/1/1	10:00	2017/1/1 10:00	9.3	2507									
13	2017/1/1	11:00	2017/1/1 11:00	11.7	2400									
14	2017/1/1	12:00	2017/1/1 12:00	13.7	2333									
15	2017/1/1	13:00	2017/1/1 13:00	12.9	2283									
16	2017/1/1	14:00	2017/1/1 14:00	13.1	2259									
17	2017/1/1	15:00	2017/1/1 15:00	12.3	2323									
18	2017/1/1	16:00	2017/1/1 16:00	11.5	2540									
19	2017/1/1	17:00	2017/1/1 17:00	7.8	2830									

気温の列が追加された

5_data/tepco-2017.xlsxに一連の処理が記録されています

仮説 1 : 電力と気温の関係性

日付 (C列) 気温 (D列) 電力 (E列) を選択して, 2つの時系列を重ねて表示しましょう。
電力の時系列図には第2軸を表示させましょう。



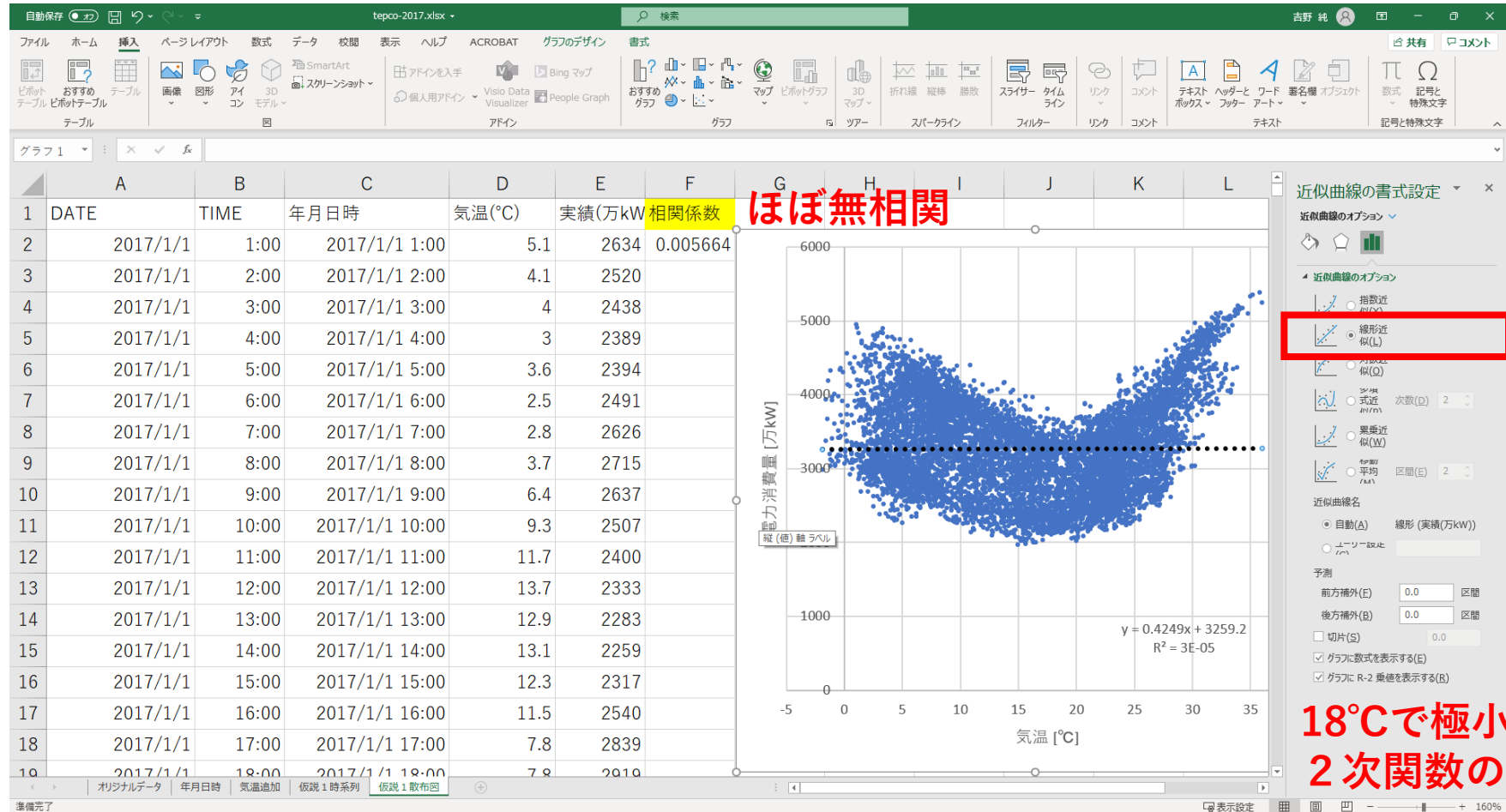
5_data/tepcO-2017.xlsxに一連の処理が記録されています

仮説 1 : 電力と気温の関係性

気温 (D列) 電力 (E列) を選択して、相関係数を計算し、散布図を作成しましょう。

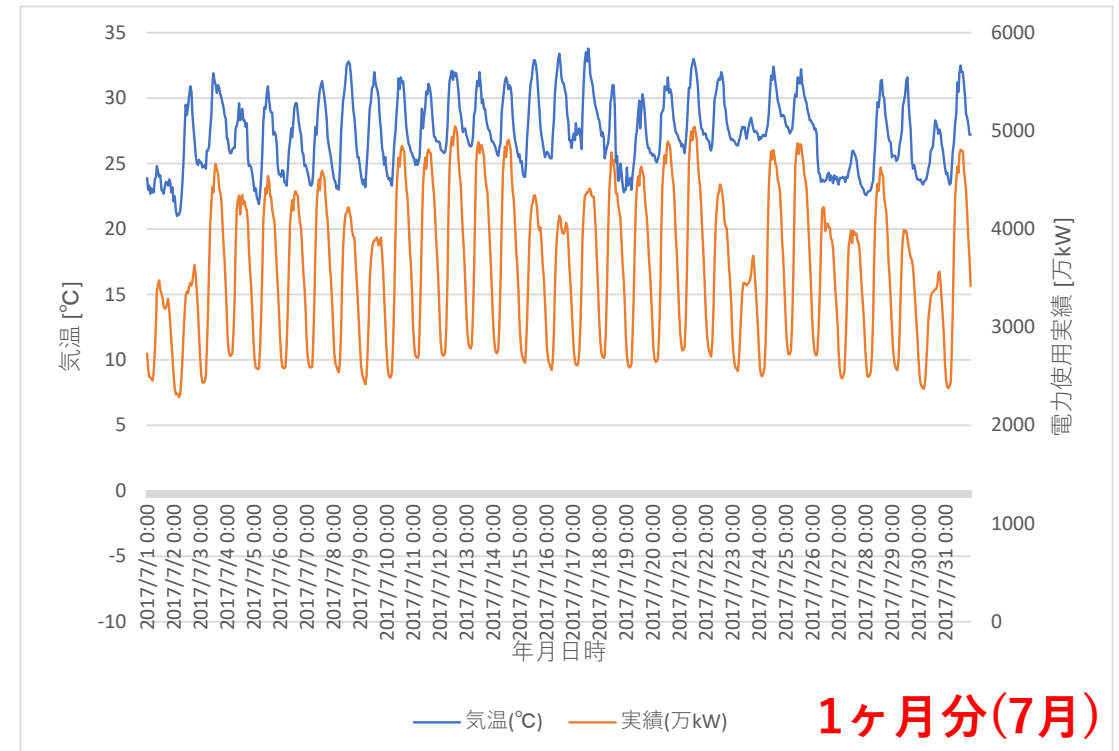
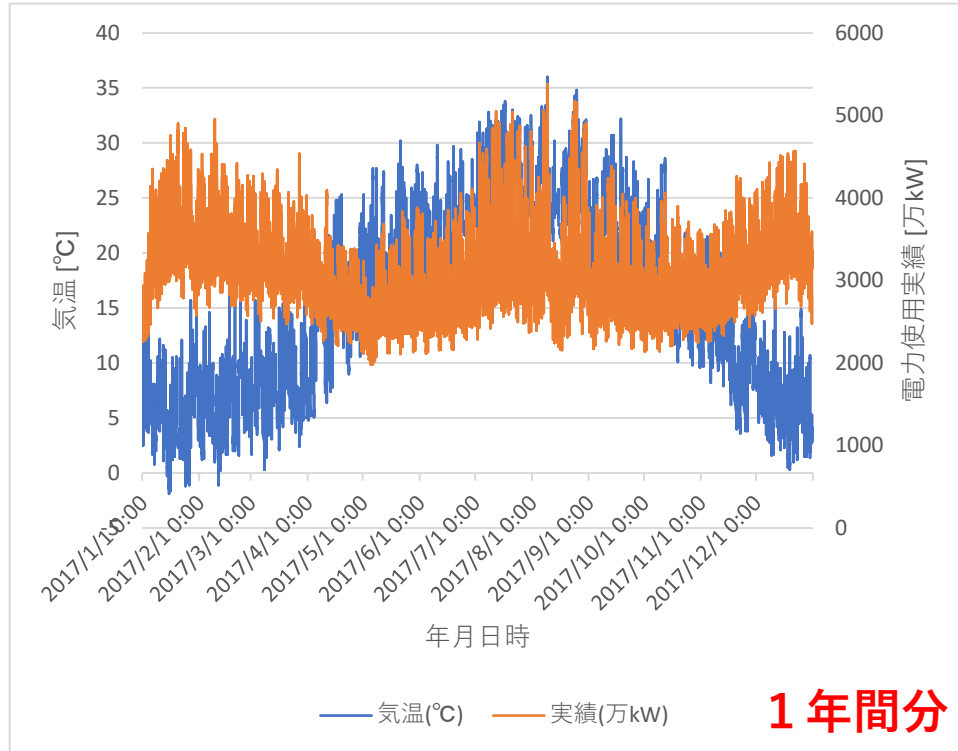
=CORREL(D2:D8760,E2:E8760)

気温 電力



仮説 1: 電力と気温の関係性

時系列図から読み取れることをぼんやりと考えてみましょう。

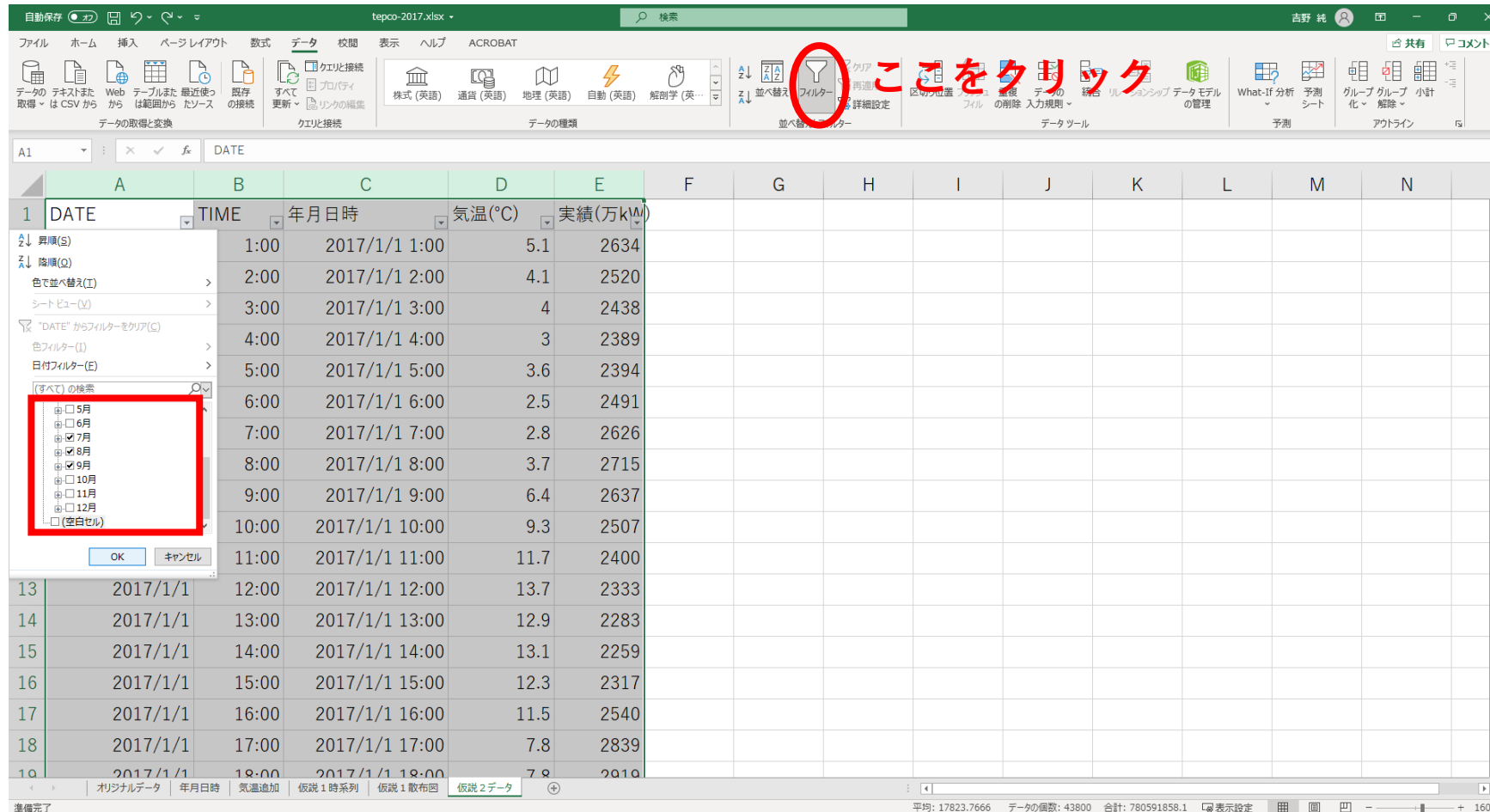


- 気温の高い夏季に電力消費量も上がる傾向にある（季節変化）
- 気温の低い冬季になると電力消費量が再び上昇する傾向にある（季節変化）
- 気温の高い昼間と気温の低い夜間の電力消費量の差が大きい（特に夏季に）（日変化）
- 気温では見られないが、電力消費量は1週間スケールでも周期的に変化している（週変化）

ここでの気づきが、以降のデータ分析の仮説を立てるうえで重要です。

仮説 2 : 夏季に限定する

次に、先の考察を踏まえて、「夏季と冬季に区別して分析すればより明瞭な関係性が得られる」と新たな仮説を立てます。フィルターを使ってデータを7月～9月の3ヶ月間に限定します。



7月, 8月, 9月の3ヶ月間にチェック

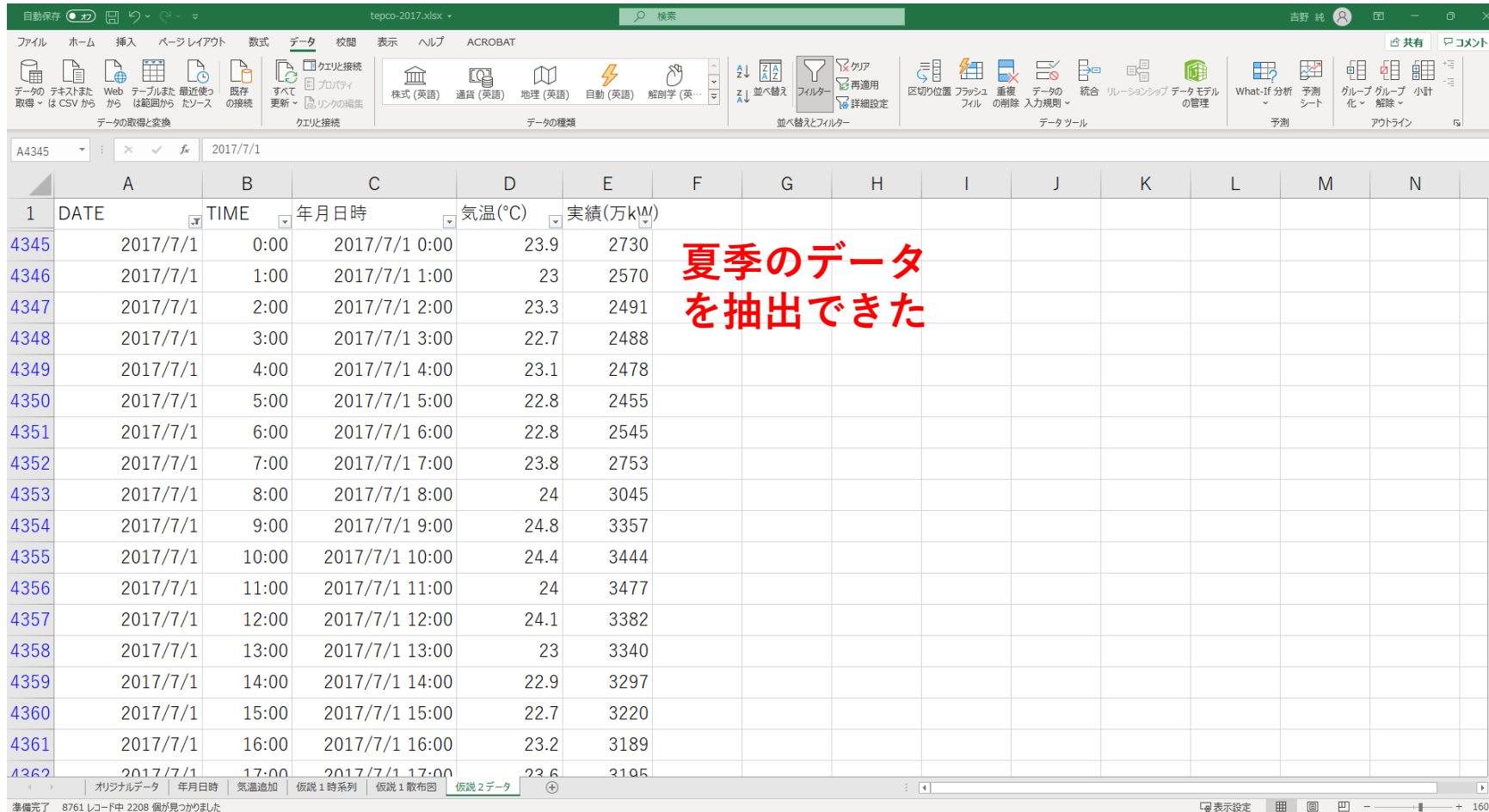
DATE	TIME	年月日時	気温(°C)	実績(万kW)	
	1:00	2017/1/1 1:00	5.1	2634	
	2:00	2017/1/1 2:00	4.1	2520	
	3:00	2017/1/1 3:00	4	2438	
	4:00	2017/1/1 4:00	3	2389	
	5:00	2017/1/1 5:00	3.6	2394	
	6:00	2017/1/1 6:00	2.5	2491	
	7:00	2017/1/1 7:00	2.8	2626	
	8:00	2017/1/1 8:00	3.7	2715	
	9:00	2017/1/1 9:00	6.4	2637	
	10:00	2017/1/1 10:00	9.3	2507	
	11:00	2017/1/1 11:00	11.7	2400	
13	2017/1/1	12:00	2017/1/1 12:00	13.7	2333
14	2017/1/1	13:00	2017/1/1 13:00	12.9	2283
15	2017/1/1	14:00	2017/1/1 14:00	13.1	2259
16	2017/1/1	15:00	2017/1/1 15:00	12.3	2317
17	2017/1/1	16:00	2017/1/1 16:00	11.5	2540
18	2017/1/1	17:00	2017/1/1 17:00	7.8	2839
19	2017/1/1	18:00	2017/1/1 18:00	7.8	2910

5_data/tepc0-2017.xlsxに一連の処理が記録されています

仮説 2 : 夏季に限定する

7~9月のデータのみを抽出することができた。

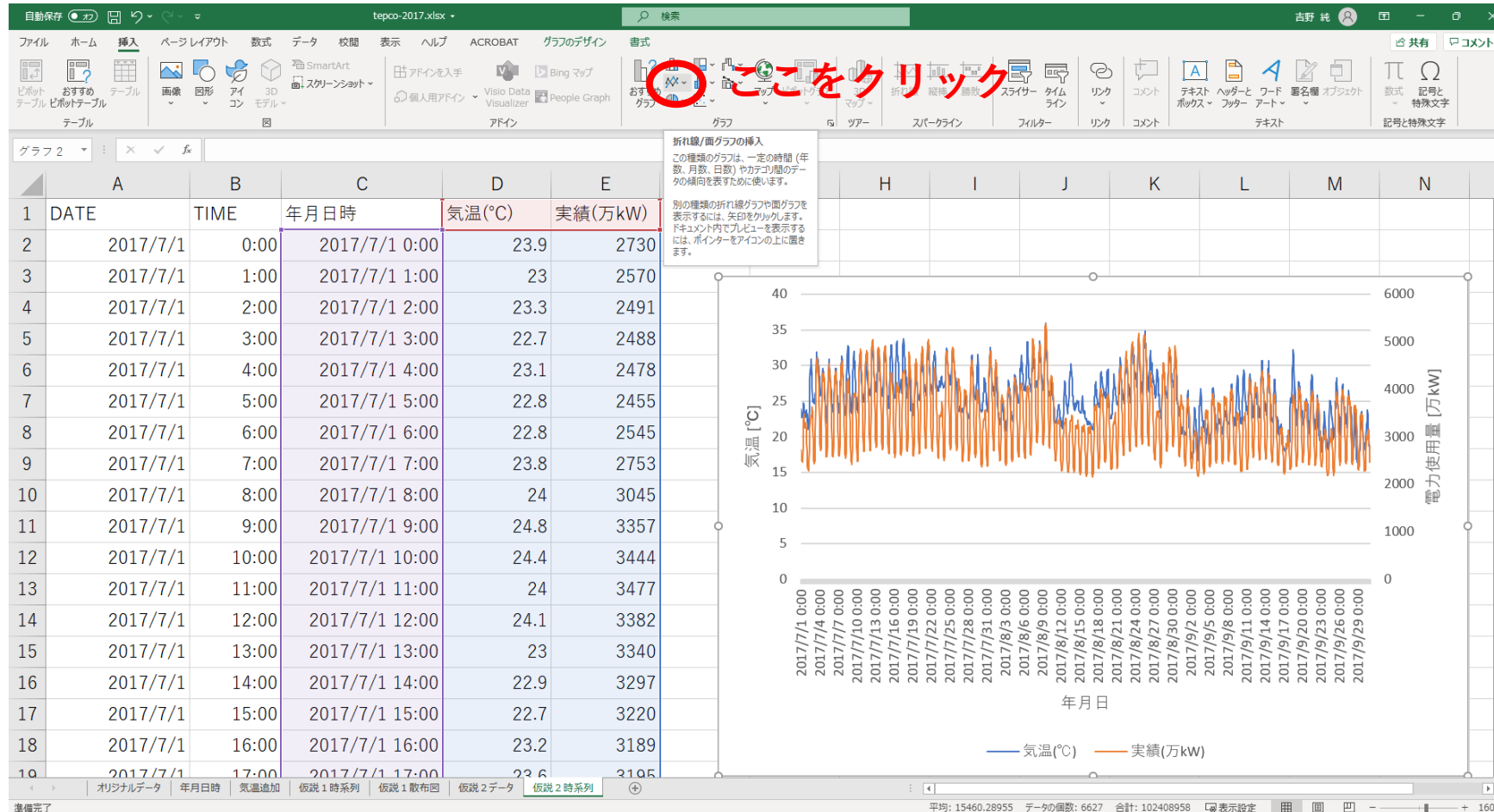
その他データは非表示になっただけで消えたわけではないことに注意。
別のワークシートにコピー&ペーストしましょう。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	DATE	TIME	年月日時	気温(°C)	実績(万kWh)									
4345	2017/7/1	0:00	2017/7/1 0:00	23.9	2730									
4346	2017/7/1	1:00	2017/7/1 1:00	23	2570									
4347	2017/7/1	2:00	2017/7/1 2:00	23.3	2491									
4348	2017/7/1	3:00	2017/7/1 3:00	22.7	2488									
4349	2017/7/1	4:00	2017/7/1 4:00	23.1	2478									
4350	2017/7/1	5:00	2017/7/1 5:00	22.8	2455									
4351	2017/7/1	6:00	2017/7/1 6:00	22.8	2545									
4352	2017/7/1	7:00	2017/7/1 7:00	23.8	2753									
4353	2017/7/1	8:00	2017/7/1 8:00	24	3045									
4354	2017/7/1	9:00	2017/7/1 9:00	24.8	3357									
4355	2017/7/1	10:00	2017/7/1 10:00	24.4	3444									
4356	2017/7/1	11:00	2017/7/1 11:00	24	3477									
4357	2017/7/1	12:00	2017/7/1 12:00	24.1	3382									
4358	2017/7/1	13:00	2017/7/1 13:00	23	3340									
4359	2017/7/1	14:00	2017/7/1 14:00	22.9	3297									
4360	2017/7/1	15:00	2017/7/1 15:00	22.7	3220									
4361	2017/7/1	16:00	2017/7/1 16:00	23.2	3189									
4362	2017/7/1	17:00	2017/7/1 17:00	23.6	3105									

仮説 2: 夏季に限定する

日付 (C列) 気温 (D列) 電力 (E列) を選択して、2つの時系列を重ねて表示しましょう。
電力の時系列図には第2軸を表示させましょう。



2つの時系列を表示できた。

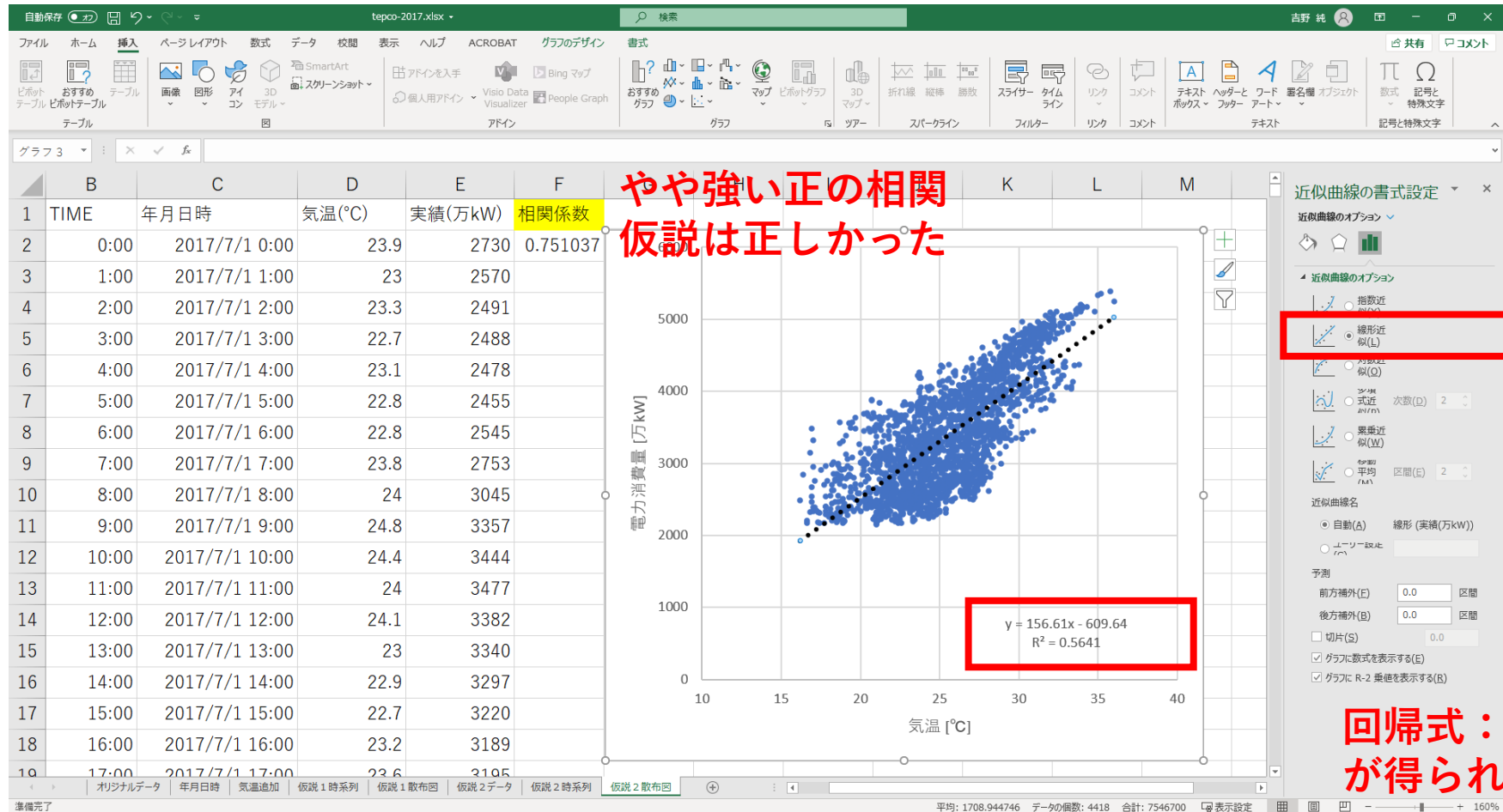
仮説 2: 夏季に限定する

気温 (D列) 電力 (E列) を選択して、相関係数を計算し、散布図を作成しましょう。

=CORREL(D2:D2209,E2:E2209)

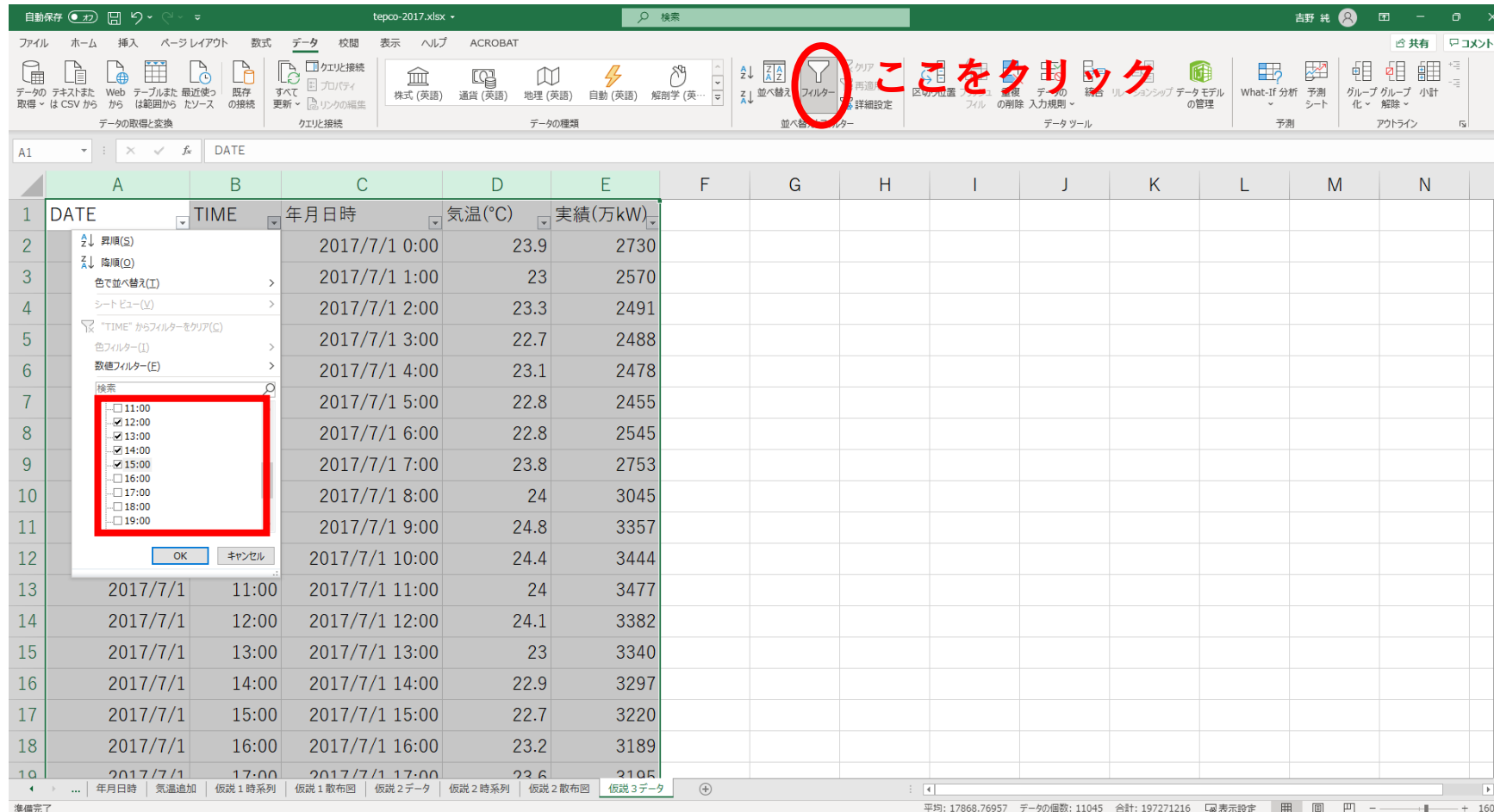
気温

電力



仮説 3 : 12~15時に限定する

次に、先の考察を踏まえて、「気温が最も高くなる12~15時のみを抽出して分析すればより明瞭な関係性が得られる」と新たな仮説を立てます。フィルターを使ってデータを12~15時の4時間に限定します。



自動保存 ●tepco-2017.xlsx 検索 吉野 純

ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 ヘルプ ACROBAT

データの種類 並べ替え フィルター 詳細設定

ここをクリック

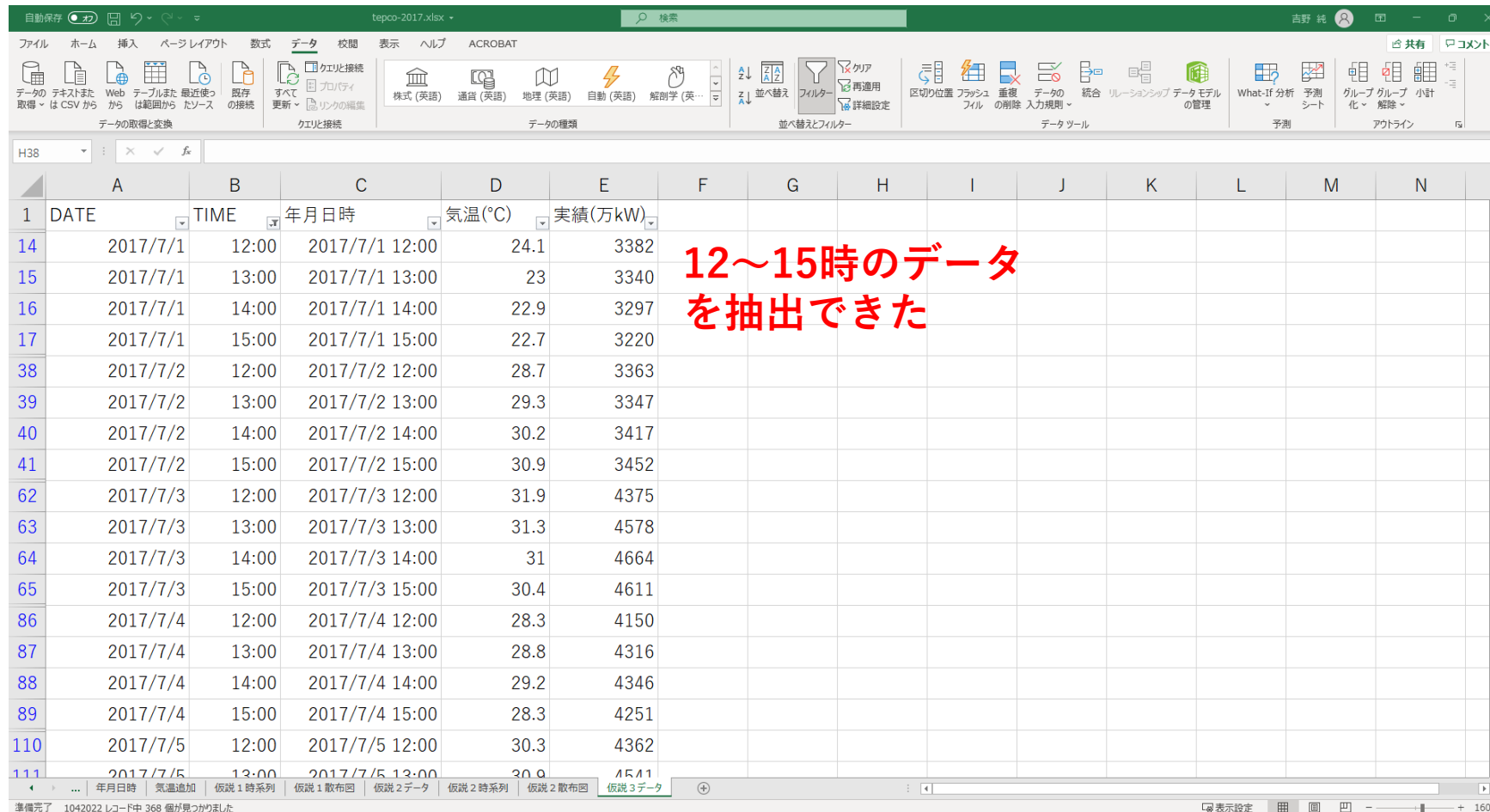
DATE	TIME	年月日時	気温(°C)	実績(万kW)
		2017/7/1 0:00	23.9	2730
		2017/7/1 1:00	23	2570
		2017/7/1 2:00	23.3	2491
		2017/7/1 3:00	22.7	2488
		2017/7/1 4:00	23.1	2478
		2017/7/1 5:00	22.8	2455
		2017/7/1 6:00	22.8	2545
		2017/7/1 7:00	23.8	2753
		2017/7/1 8:00	24	3045
		2017/7/1 9:00	24.8	3357
		2017/7/1 10:00	24.4	3444
	2017/7/1 11:00	2017/7/1 11:00	24	3477
	2017/7/1 12:00	2017/7/1 12:00	24.1	3382
	2017/7/1 13:00	2017/7/1 13:00	23	3340
	2017/7/1 14:00	2017/7/1 14:00	22.9	3297
	2017/7/1 15:00	2017/7/1 15:00	22.7	3220
	2017/7/1 16:00	2017/7/1 16:00	23.2	3189
	2017/7/1 17:00	2017/7/1 17:00	23.6	3105

準備完了 平均: 17868.76957 データの個数: 11045 合計: 197271216 表示設定 160%

12~15時
の4時間分に
チェック

仮説 3 : 12~15時に限定する

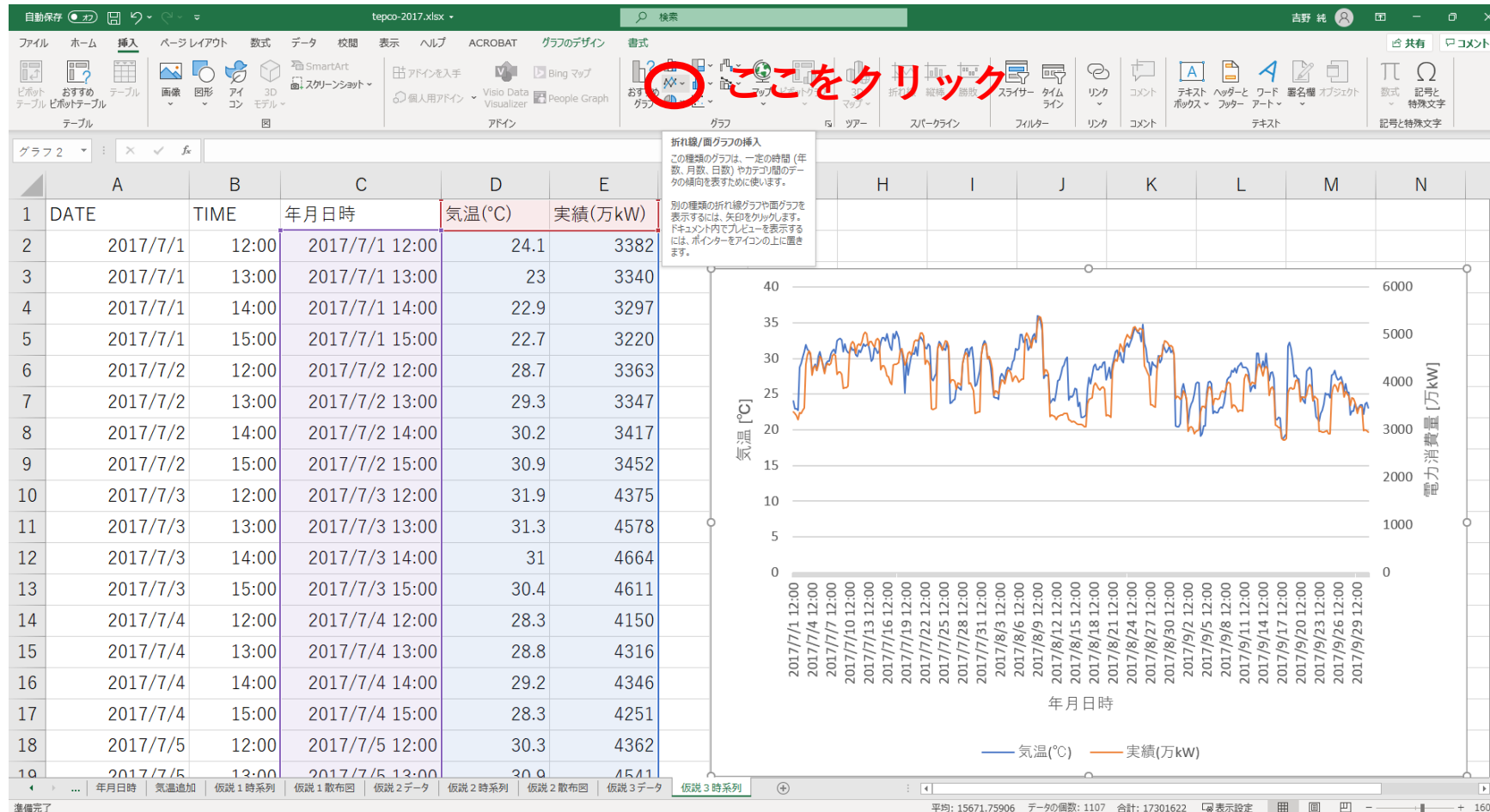
「7~9月」「12~15時」のデータのみを抽出することができた。
その他データは非表示になっただけで消えたわけではないことに注意。
別のワークシートにコピー&ペーストしましょう。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	DATE	TIME	年月日時	気温(°C)	実績(万kW)									
14	2017/7/1	12:00	2017/7/1 12:00	24.1	3382									
15	2017/7/1	13:00	2017/7/1 13:00	23	3340									
16	2017/7/1	14:00	2017/7/1 14:00	22.9	3297									
17	2017/7/1	15:00	2017/7/1 15:00	22.7	3220									
38	2017/7/2	12:00	2017/7/2 12:00	28.7	3363									
39	2017/7/2	13:00	2017/7/2 13:00	29.3	3347									
40	2017/7/2	14:00	2017/7/2 14:00	30.2	3417									
41	2017/7/2	15:00	2017/7/2 15:00	30.9	3452									
62	2017/7/3	12:00	2017/7/3 12:00	31.9	4375									
63	2017/7/3	13:00	2017/7/3 13:00	31.3	4578									
64	2017/7/3	14:00	2017/7/3 14:00	31	4664									
65	2017/7/3	15:00	2017/7/3 15:00	30.4	4611									
86	2017/7/4	12:00	2017/7/4 12:00	28.3	4150									
87	2017/7/4	13:00	2017/7/4 13:00	28.8	4316									
88	2017/7/4	14:00	2017/7/4 14:00	29.2	4346									
89	2017/7/4	15:00	2017/7/4 15:00	28.3	4251									
110	2017/7/5	12:00	2017/7/5 12:00	30.3	4362									
111	2017/7/5	13:00	2017/7/5 13:00	30.9	4541									

仮説 3 : 12~15時に限定する

日付 (C列) 気温 (D列) 電力 (E列) を選択して, 2つの時系列を重ねて表示しましょう。
電力の時系列図には第2軸を表示させましょう。



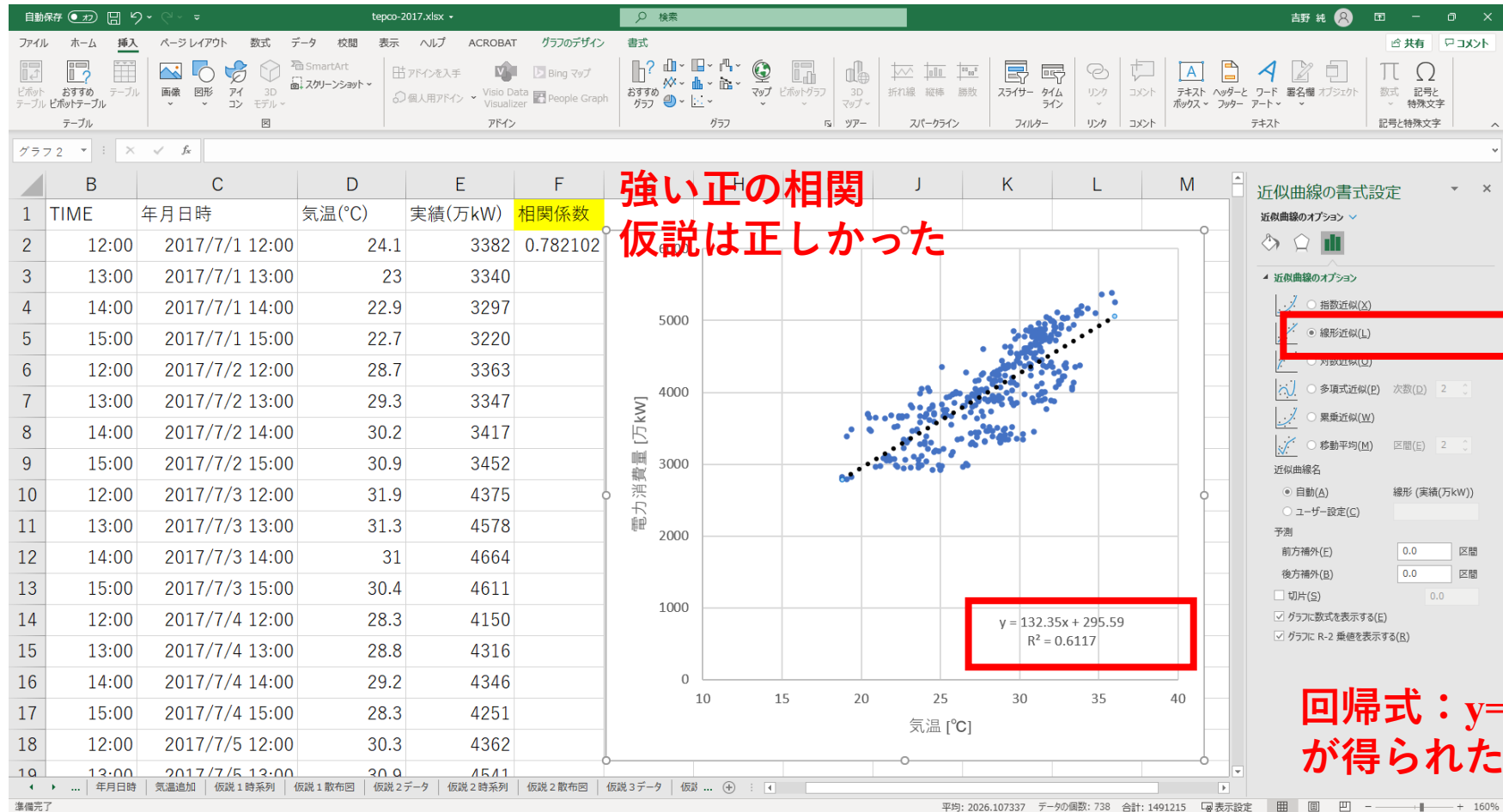
2つの時系列を表示できた。

仮説 3 : 12~15時に限定する

気温 (D列) 電力 (E列) を選択して, 相関係数を計算し, 散布図を作成しましょう。

=CORREL(D2:D369,E2:E369)

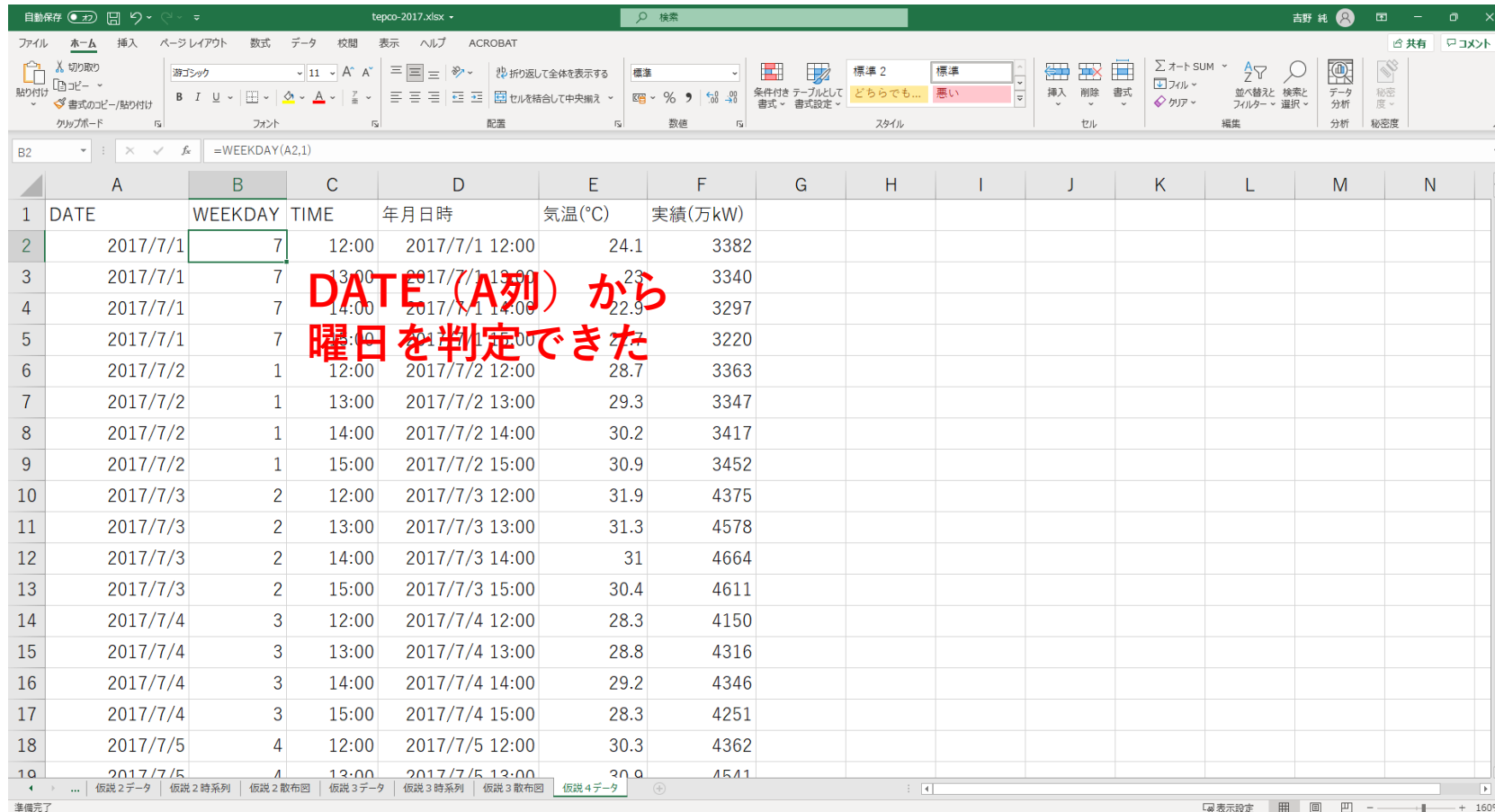
気温 電力



仮説4：平日に限定する

次に、先の考察を踏まえて、「土曜日と日曜日のデータを取り除いて平日のみを分析すればより明瞭な関係性が得られる」と新たな仮説を立てます。まず、WEEKDAY関数で行ごとに曜日を判定します。

=WEEKDAY(A2, 1)
DATE 1:日曜日~7:土曜日

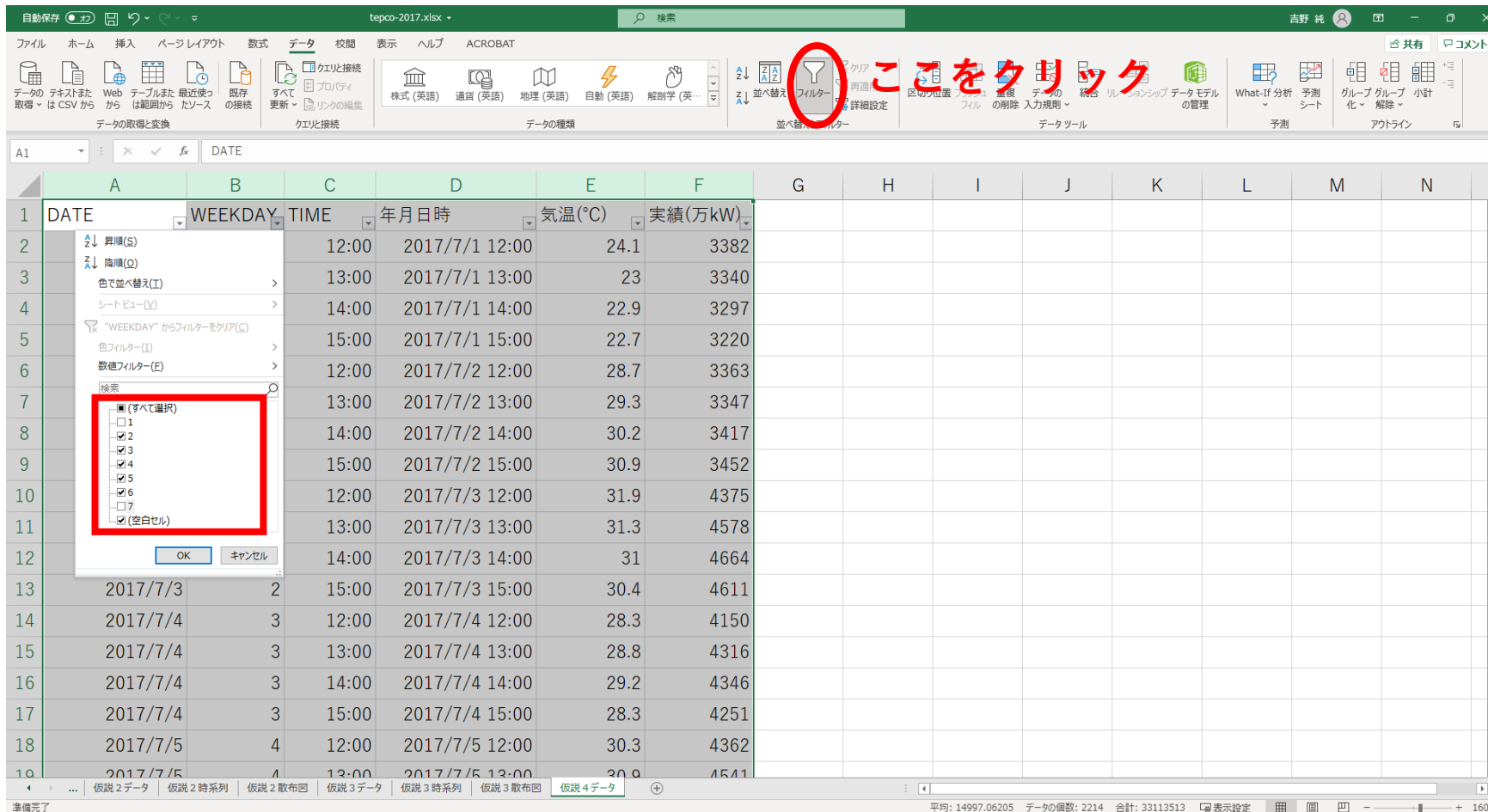


	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	DATE	WEEKDAY	TIME	年月日時	気温(°C)	実績(万kW)								
2	2017/7/1	7	12:00	2017/7/1 12:00	24.1	3382								
3	2017/7/1	7	13:00	2017/7/1 13:00	23.9	3340								
4	2017/7/1	7	14:00	2017/7/1 14:00	22.9	3297								
5	2017/7/1	7	15:00	2017/7/1 15:00	22.7	3220								
6	2017/7/2	1	12:00	2017/7/2 12:00	28.7	3363								
7	2017/7/2	1	13:00	2017/7/2 13:00	29.3	3347								
8	2017/7/2	1	14:00	2017/7/2 14:00	30.2	3417								
9	2017/7/2	1	15:00	2017/7/2 15:00	30.9	3452								
10	2017/7/3	2	12:00	2017/7/3 12:00	31.9	4375								
11	2017/7/3	2	13:00	2017/7/3 13:00	31.3	4578								
12	2017/7/3	2	14:00	2017/7/3 14:00	31	4664								
13	2017/7/3	2	15:00	2017/7/3 15:00	30.4	4611								
14	2017/7/4	3	12:00	2017/7/4 12:00	28.3	4150								
15	2017/7/4	3	13:00	2017/7/4 13:00	28.8	4316								
16	2017/7/4	3	14:00	2017/7/4 14:00	29.2	4346								
17	2017/7/4	3	15:00	2017/7/4 15:00	28.3	4251								
18	2017/7/5	4	12:00	2017/7/5 12:00	30.3	4362								
19	2017/7/5	4	13:00	2017/7/5 13:00	30.0	4541								

5_data/tepcO-2017.xlsxに一連の処理が記録されています

仮説4：平日に限定する

フィルターを使ってデータを日曜日(1)と土曜日(7)以外の平日のデータに限定します。



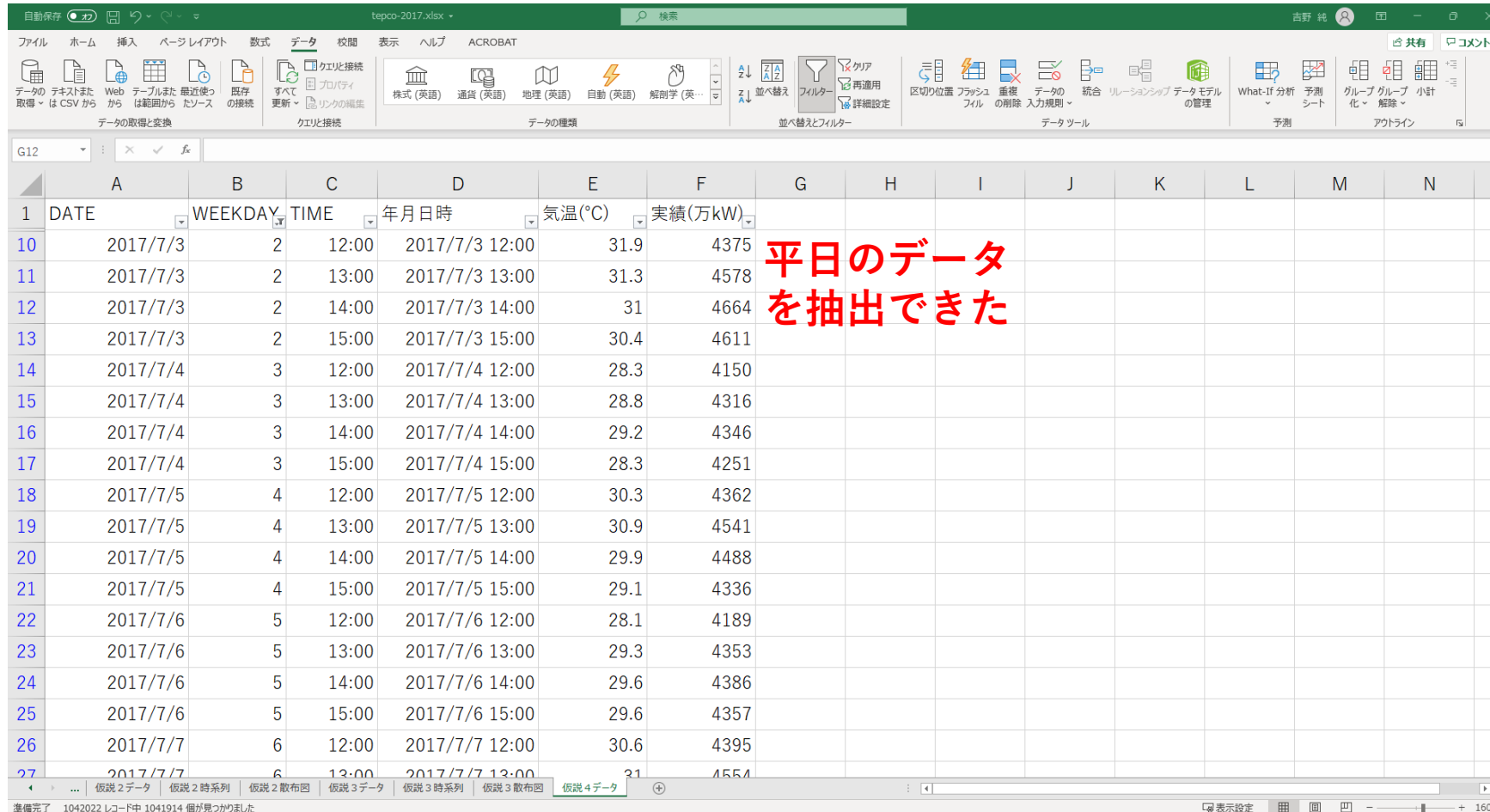
ここをクリック

DATE	WEEKDAY	TIME	年月日時	気温(°C)	実績(万kW)
		12:00	2017/7/1 12:00	24.1	3382
		13:00	2017/7/1 13:00	23	3340
		14:00	2017/7/1 14:00	22.9	3297
		15:00	2017/7/1 15:00	22.7	3220
		12:00	2017/7/2 12:00	28.7	3363
		13:00	2017/7/2 13:00	29.3	3347
		14:00	2017/7/2 14:00	30.2	3417
		15:00	2017/7/2 15:00	30.9	3452
		12:00	2017/7/3 12:00	31.9	4375
		13:00	2017/7/3 13:00	31.3	4578
		14:00	2017/7/3 14:00	31	4664
2017/7/3	2	15:00	2017/7/3 15:00	30.4	4611
2017/7/4	3	12:00	2017/7/4 12:00	28.3	4150
2017/7/4	3	13:00	2017/7/4 13:00	28.8	4316
2017/7/4	3	14:00	2017/7/4 14:00	29.2	4346
2017/7/4	3	15:00	2017/7/4 15:00	28.3	4251
2017/7/5	4	12:00	2017/7/5 12:00	30.3	4362
2017/7/5	4	13:00	2017/7/5 13:00	30.9	4541

火(2),水(3),
木(4),金(5),
土(6)
にチェック

仮説4：平日に限定する

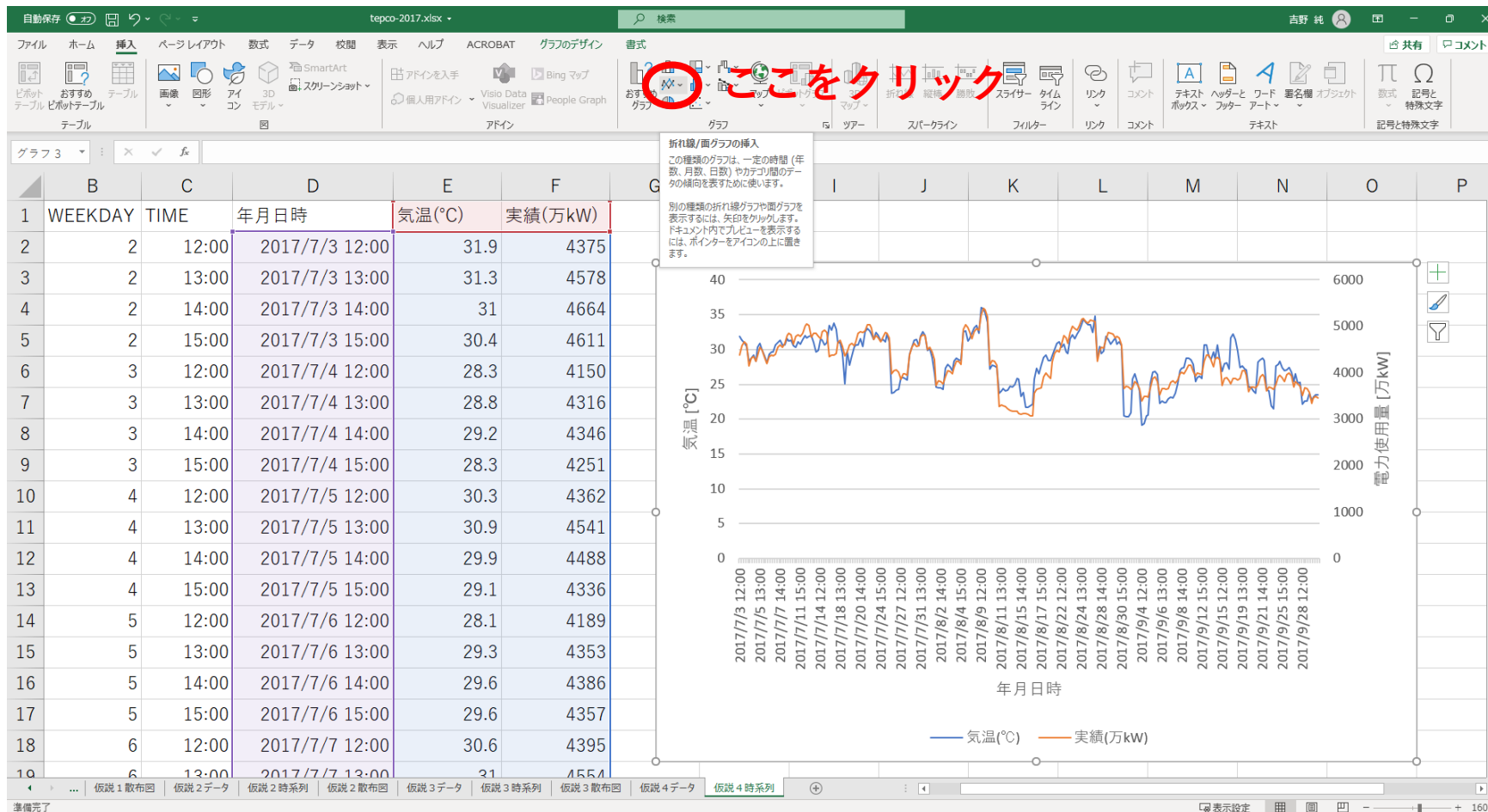
「7~9月」「12~15時」「平日」のデータのみを抽出することができた。
 その他データは非表示になっただけで消えたわけではないことに注意。
 別のワークシートにコピー&ペーストしましょう。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	DATE	WEEKDAY	TIME	年月日時	気温(°C)	実績(万kW)								
10	2017/7/3	2	12:00	2017/7/3 12:00	31.9	4375								
11	2017/7/3	2	13:00	2017/7/3 13:00	31.3	4578								
12	2017/7/3	2	14:00	2017/7/3 14:00	31	4664								
13	2017/7/3	2	15:00	2017/7/3 15:00	30.4	4611								
14	2017/7/4	3	12:00	2017/7/4 12:00	28.3	4150								
15	2017/7/4	3	13:00	2017/7/4 13:00	28.8	4316								
16	2017/7/4	3	14:00	2017/7/4 14:00	29.2	4346								
17	2017/7/4	3	15:00	2017/7/4 15:00	28.3	4251								
18	2017/7/5	4	12:00	2017/7/5 12:00	30.3	4362								
19	2017/7/5	4	13:00	2017/7/5 13:00	30.9	4541								
20	2017/7/5	4	14:00	2017/7/5 14:00	29.9	4488								
21	2017/7/5	4	15:00	2017/7/5 15:00	29.1	4336								
22	2017/7/6	5	12:00	2017/7/6 12:00	28.1	4189								
23	2017/7/6	5	13:00	2017/7/6 13:00	29.3	4353								
24	2017/7/6	5	14:00	2017/7/6 14:00	29.6	4386								
25	2017/7/6	5	15:00	2017/7/6 15:00	29.6	4357								
26	2017/7/7	6	12:00	2017/7/7 12:00	30.6	4395								
27	2017/7/7	6	13:00	2017/7/7 13:00	31	4554								

仮説 4 : 平日に限定する

日付 (D列) 気温 (E列) 電力 (F列) を選択して、2つの時系列を重ねて表示しましょう。
電力の時系列図には第2軸を表示させましょう。



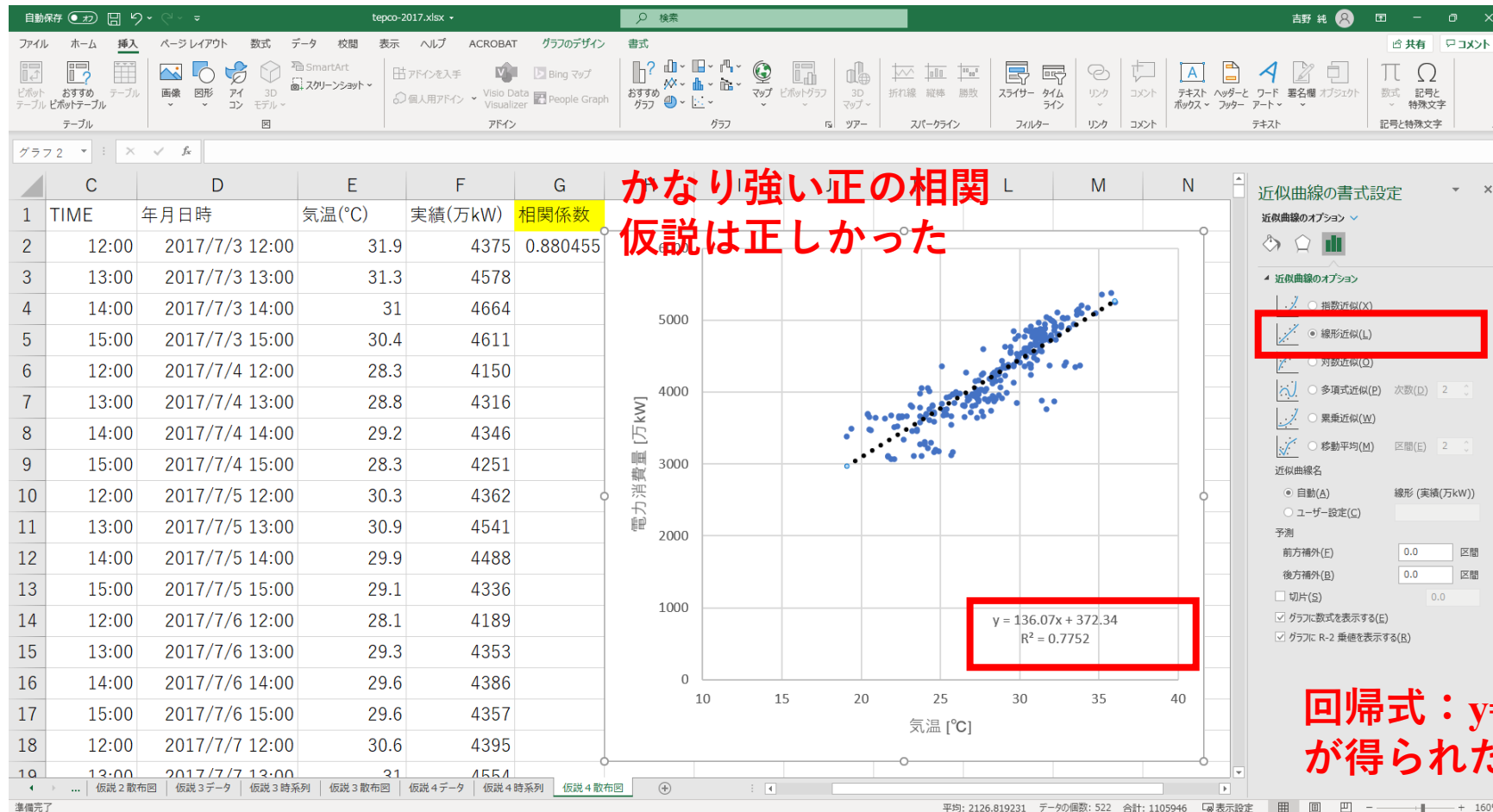
2つの時系列を表示できた。

仮説 4 : 平日に限定する

気温 (E列) 電力 (F列) を選択して, 相関係数を計算し, 散布図を作成しましょう.

=CORREL(E2:E261,F2:F261)

気温 電力



仮説5：さらに祝日や夏季休暇を取り除く



次に、先の考察を踏まえて、「祝日（7/17, 9/18）と夏季休暇（8/7-8/20）のデータを取り除いて分析すればより明瞭な関係性が得られる」と新たな仮説を立てます。手動で取り除きます。

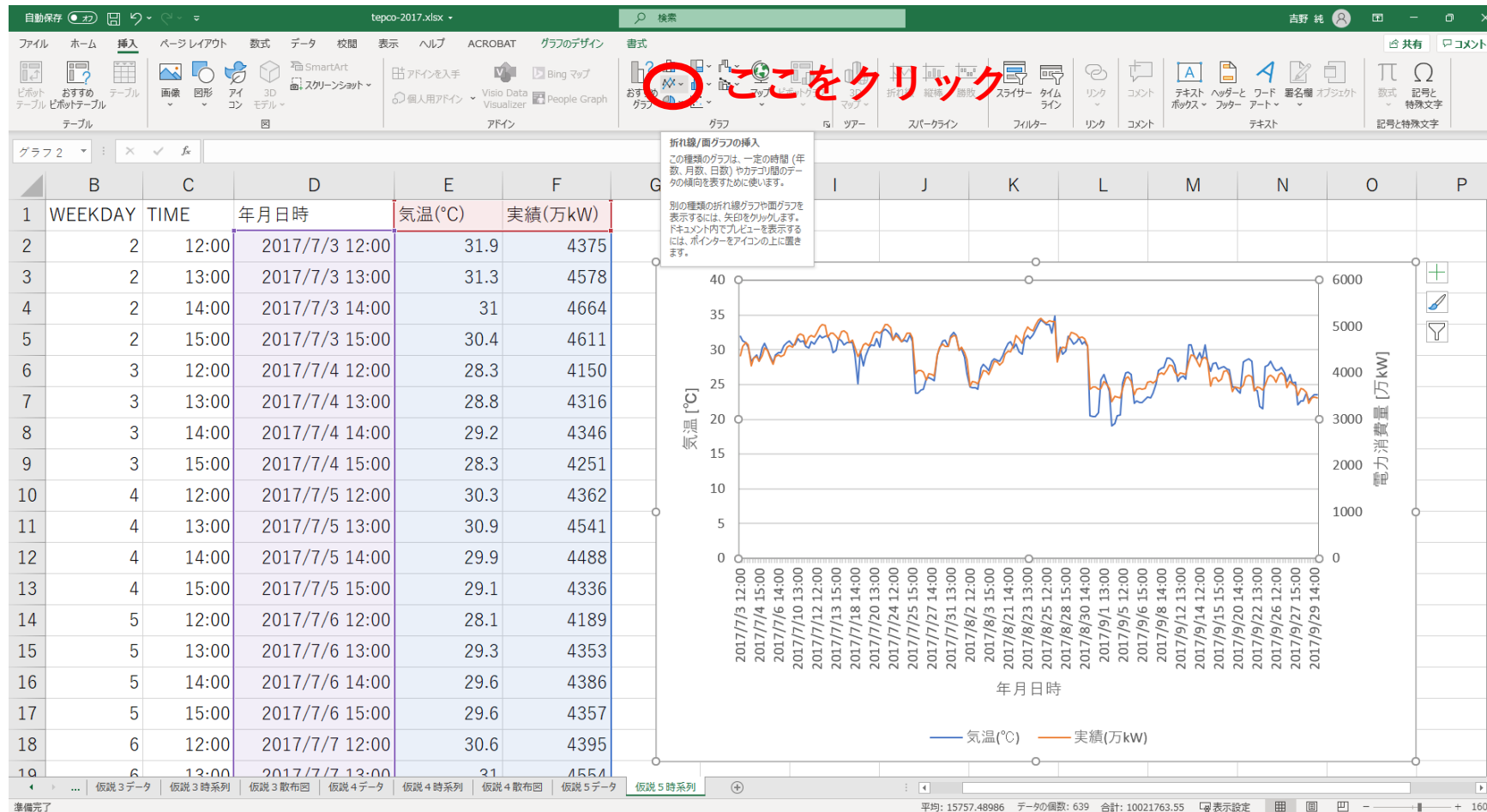
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
34	2017/7/13	5	12:00	2017/7/13 12:00	32	4772								
35	2017/7/13	5	13:00	2017/7/13 13:00	31	4856								
36	2017/7/13	5	14:00	2017/7/13 14:00	29.6	4850								
37	2017/7/13	5	15:00	2017/7/13 15:00	29.9	4786								
38	2017/7/14	6	12:00	2017/7/14 12:00	31.6	4731								
39	2017/7/14	6	13:00	2017/7/14 13:00	31.3	4886								
40	2017/7/14	6	14:00	2017/7/14 14:00	30.7	4912								
41	2017/7/14	6	15:00	2017/7/14 15:00	31	4864								
42	2017/7/18	3	12:00	2017/7/18 12:00	31	4656								
43	2017/7/18	3	13:00	2017/7/18 13:00	31	4706								
44	2017/7/18	3	14:00	2017/7/18 14:00	29.4	4562								
45	2017/7/18	3	15:00	2017/7/18 15:00	25.1	4358								
46	2017/7/19	4	12:00	2017/7/19 12:00	29.8	4440								
47	2017/7/19	4	13:00	2017/7/19 13:00	27.7	4605								
48	2017/7/19	4	14:00	2017/7/19 14:00	29.1	4636								
49	2017/7/19	4	15:00	2017/7/19 15:00	30.3	4582								
50	2017/7/20	5	12:00	2017/7/20 12:00	30.7	4668								
51	2017/7/20	5	13:00	2017/7/20 13:00	30.6	4854								
52	2017/7/20	5	14:00	2017/7/20 14:00	31.6	4891								

5_data/tepcO-2017.xlsxに一連の処理が記録されています

仮説5：さらに祝日や夏季休暇を取り除く



日付（D列） 気温（E列） 電力（F列） を選択して、2つの時系列を重ねて表示しましょう。
電力の時系列図には第2軸を表示させましょう。



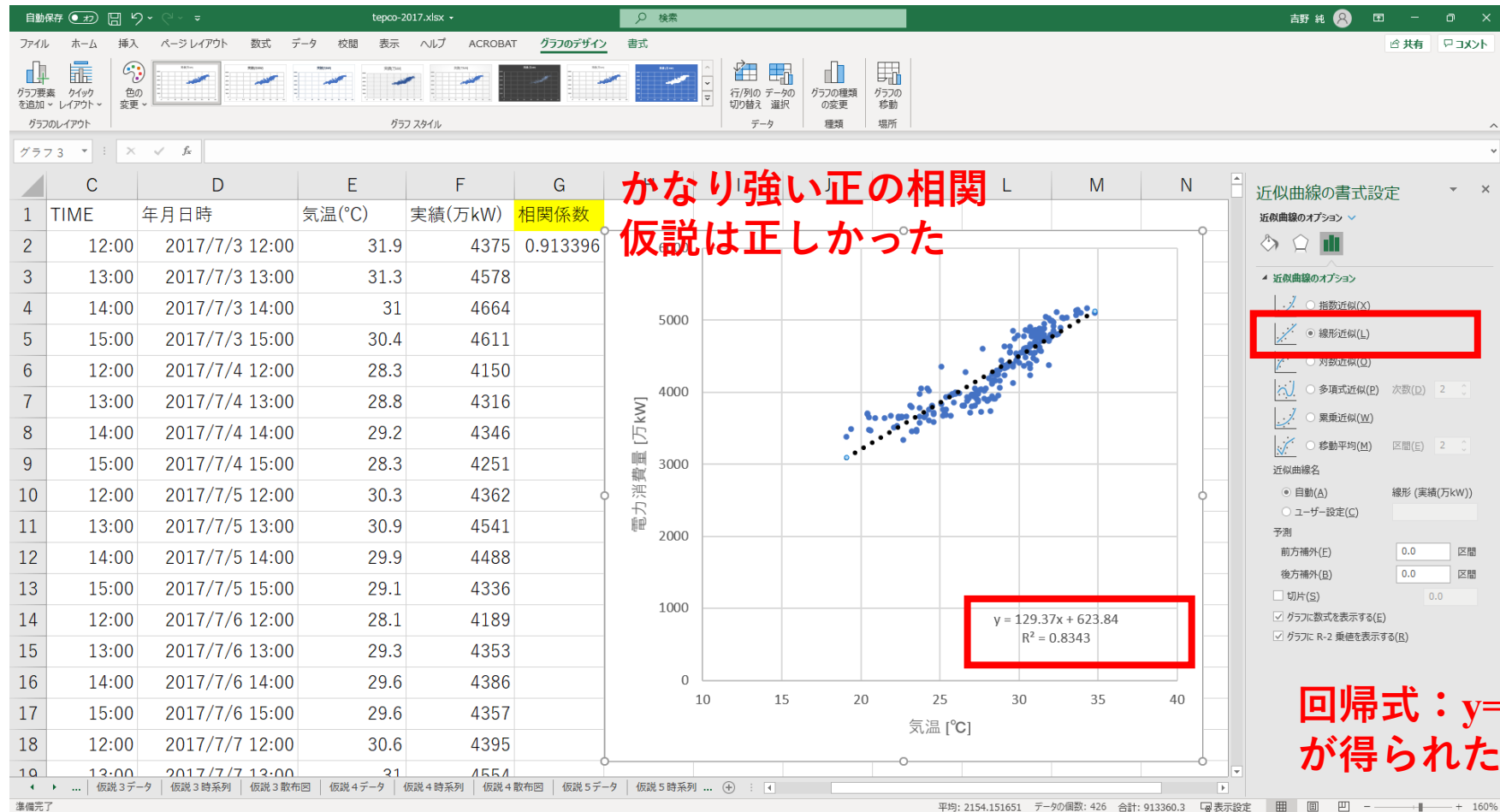
2つの時系列を表示できた。
2つの時系列はよく似た変化傾向を示している。

仮説5：さらに祝日や夏季休暇を取り除く

気温（E列）電力（F列）を選択して、相関係数を計算し、散布図を作成しましょう。

=CORREL(E2:E213,F2:F213)

気温 電力



予測モデルを作ってみよう

これまでの分析により以下のことが分かりました。

- ✓電力消費量は気温の2次関数で表せる。
- ✓電力消費量は平日か休日かの違いによる影響が大きい。
- ✓電力消費量は時間帯の違いによる影響が大きい。

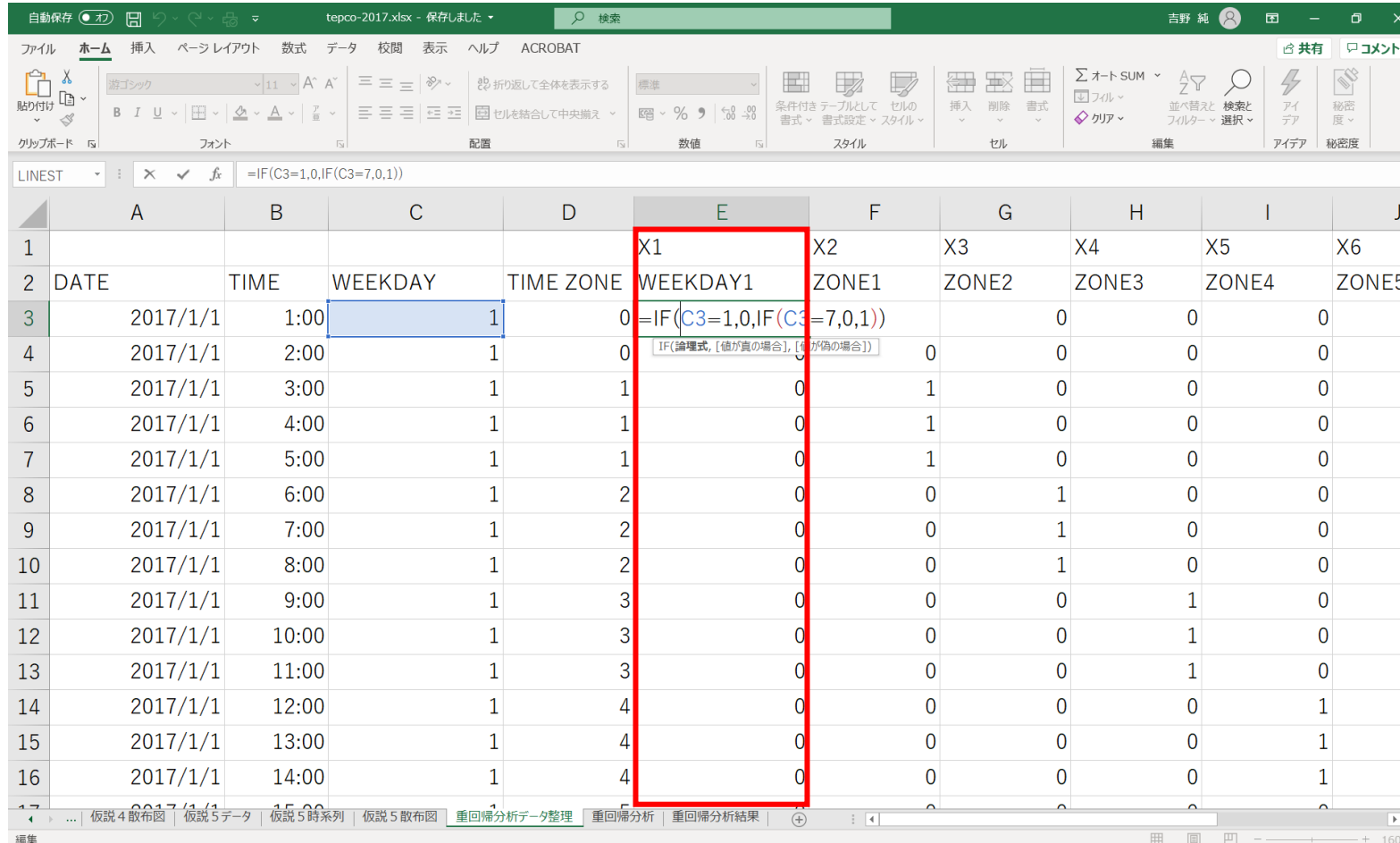
よって複数の説明変数の影響を考慮に入れて、**重回帰分析**により以下のような予測モデルを作る必要があります。

$$\begin{array}{|c|} \hline \text{電力} \\ \text{消費量} \\ \hline \end{array} \mathbf{y} = \mathbf{a} \begin{array}{|c|} \hline \text{気温} \\ \hline \end{array} \mathbf{x}_1 + \mathbf{b} \begin{array}{|c|} \hline \text{曜日} \\ \hline \end{array} \mathbf{x}_2 + \mathbf{c} \begin{array}{|c|} \hline \text{時} \\ \hline \end{array} \mathbf{x}_3 + \mathbf{d}$$

予測モデルを作ってみよう

ダミー変数（E列）を導入して、平日ならば1，休日ならば0と設定します。

=IF(C3=1,0,IF(C3=7,0,1)) C3列が1(日曜日)か7(土曜日)ならば0とし，それ以外(平日)は1とする。

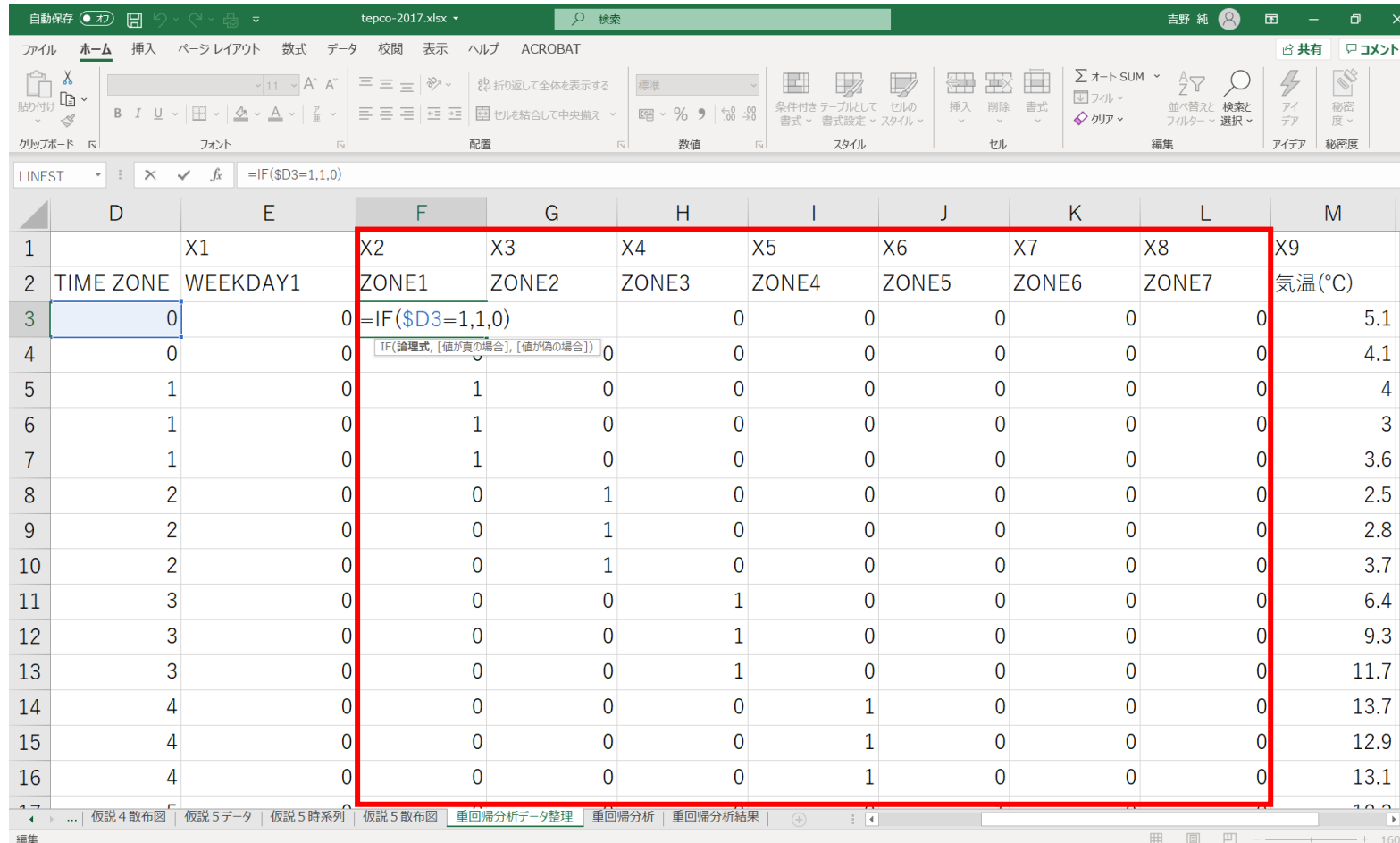


	A	B	C	D	E	F	G	H	I	J
1					X1	X2	X3	X4	X5	X6
2	DATE	TIME	WEEKDAY	TIME ZONE	WEEKDAY1	ZONE1	ZONE2	ZONE3	ZONE4	ZONE5
3	2017/1/1	1:00	1	0	=IF(C3=1,0,IF(C3=7,0,1))		0	0	0	0
4	2017/1/1	2:00	1	0	IF(論理式,[値が真の場合],[値が偽の場合])	0	0	0	0	0
5	2017/1/1	3:00	1	1		1	0	0	0	0
6	2017/1/1	4:00	1	1		0	1	0	0	0
7	2017/1/1	5:00	1	1		0	1	0	0	0
8	2017/1/1	6:00	1	2		0	0	1	0	0
9	2017/1/1	7:00	1	2		0	0	1	0	0
10	2017/1/1	8:00	1	2		0	0	1	0	0
11	2017/1/1	9:00	1	3		0	0	0	1	0
12	2017/1/1	10:00	1	3		0	0	0	1	0
13	2017/1/1	11:00	1	3		0	0	0	1	0
14	2017/1/1	12:00	1	4		0	0	0	0	1
15	2017/1/1	13:00	1	4		0	0	0	0	1
16	2017/1/1	14:00	1	4		0	0	0	0	1
17	2017/1/1	15:00	1	4		0	0	0	0	1

質的データを定量化することを「数量化I類」という。

予測モデルを作ってみよう

ダミー変数 (F~L列) を導入して、時間帯1 (3-5時), 時間帯2 (6-8時), 時間帯3 (9-11時), 時間帯4 (12-14時), 時間帯5 (15-17時), 時間帯6 (18-20時), 時間帯7 (21-23時)ごとに、該当は1, 非該当は0とする。
 $=IF(\$D3=1,1,0)$ D3列が1(3-5時)ならば1, それ以外は0とする。



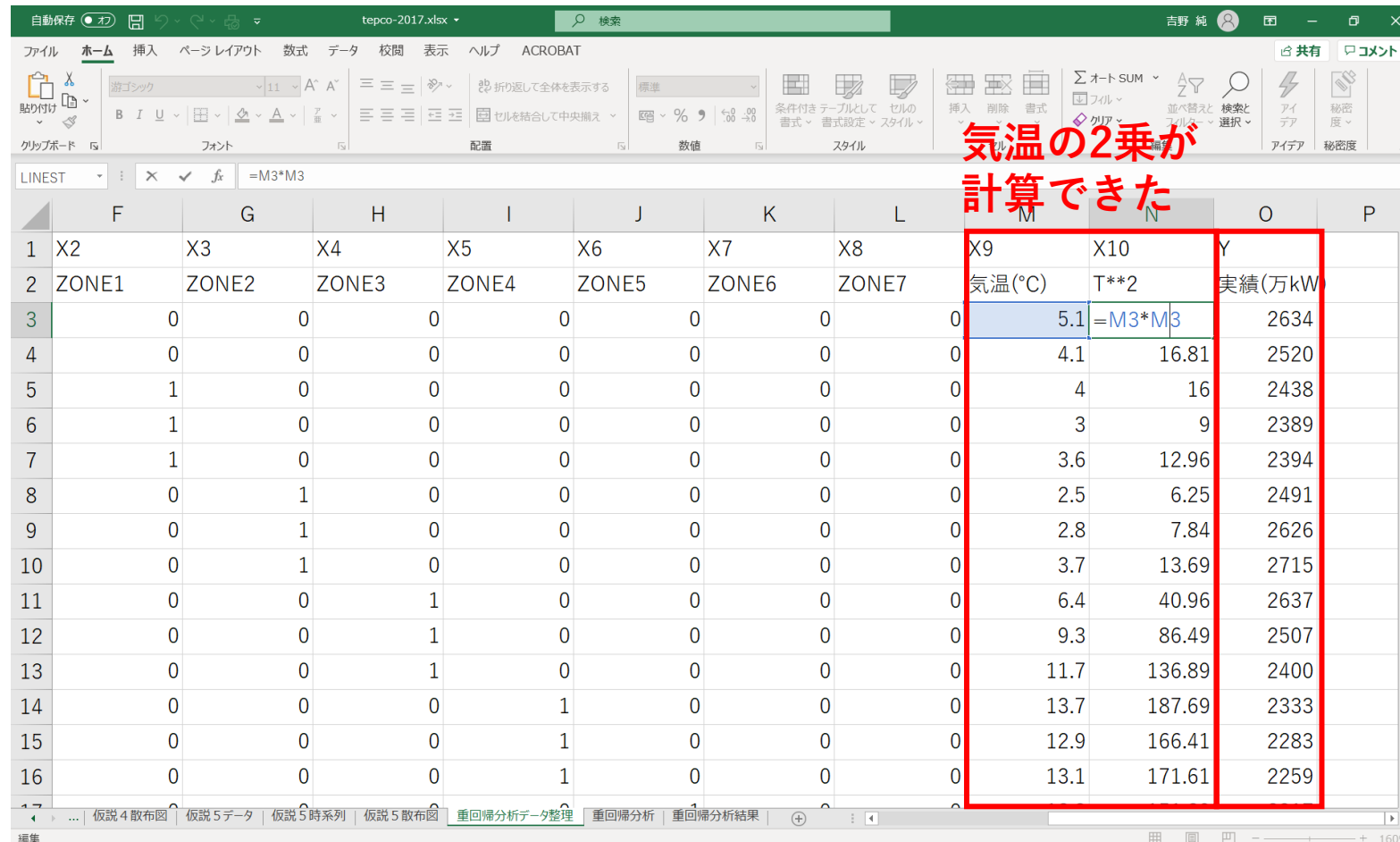
	D	E	F	G	H	I	J	K	L	M
1		X1	X2	X3	X4	X5	X6	X7	X8	X9
2	TIME_ZONE	WEEKDAY1	ZONE1	ZONE2	ZONE3	ZONE4	ZONE5	ZONE6	ZONE7	気温(°C)
3	0	0	=IF(\$D3=1,1,0)		0	0	0	0	0	5.1
4	0	0	IF(論理式, [値が真の場合], [値が偽の場合])	0	0	0	0	0	0	4.1
5	1	0		1	0	0	0	0	0	4
6	1	0		1	0	0	0	0	0	3
7	1	0		1	0	0	0	0	0	3.6
8	2	0		0	1	0	0	0	0	2.5
9	2	0		0	1	0	0	0	0	2.8
10	2	0		0	1	0	0	0	0	3.7
11	3	0		0	0	1	0	0	0	6.4
12	3	0		0	0	1	0	0	0	9.3
13	3	0		0	0	1	0	0	0	11.7
14	4	0		0	0	0	1	0	0	13.7
15	4	0		0	0	0	1	0	0	12.9
16	4	0		0	0	0	1	0	0	13.1

質的データを定量化することを「数量化I類」という。

時間帯0 (0-2時) は作る必要がない。7つのダミー変数すべてが0となる。

予測モデルを作ってみよう

電力消費量(y)は気温の2次関数で表現されるため、気温の1乗だけでなく気温の2乗も説明変数とする。
=M3*M3 気温の2乗を計算する。

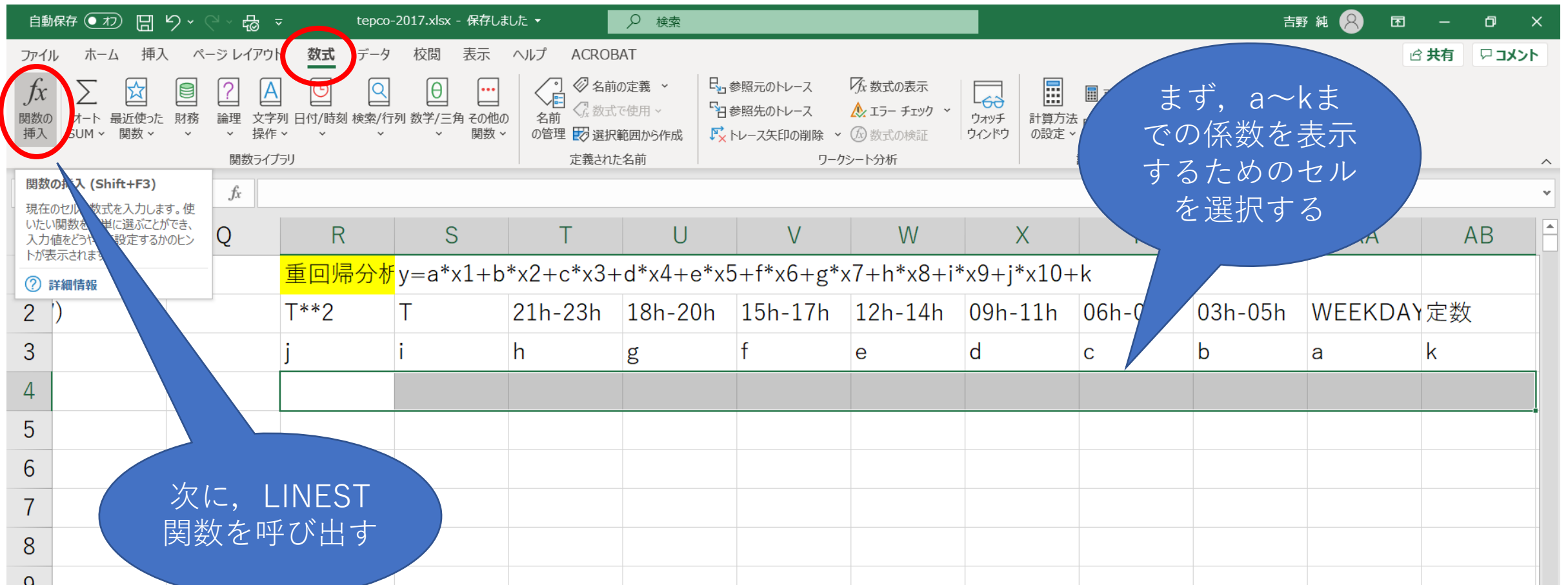


	F	G	H	I	J	K	L	M	N	O	P
1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y	
2	ZONE1	ZONE2	ZONE3	ZONE4	ZONE5	ZONE6	ZONE7	気温(°C)	T**2	実績(万kW)	
3		0	0	0	0	0	0	5.1	=M3*M3	2634	
4		0	0	0	0	0	0	4.1	16.81	2520	
5		1	0	0	0	0	0	4	16	2438	
6		1	0	0	0	0	0	3	9	2389	
7		1	0	0	0	0	0	3.6	12.96	2394	
8		0	1	0	0	0	0	2.5	6.25	2491	
9		0	1	0	0	0	0	2.8	7.84	2626	
10		0	1	0	0	0	0	3.7	13.69	2715	
11		0	0	1	0	0	0	6.4	40.96	2637	
12		0	0	1	0	0	0	9.3	86.49	2507	
13		0	0	1	0	0	0	11.7	136.89	2400	
14		0	0	0	1	0	0	13.7	187.69	2333	
15		0	0	0	1	0	0	12.9	166.41	2283	
16		0	0	0	1	0	0	13.1	171.61	2259	

O列は目的変数 (y) となる電力消費量である。

予測モデルを作ってみよう

- 10個の説明変数と定数項からなる重回帰式をLINEST関数で評価する。
重回帰式は、 $Y=a*X1+b*X2+c*X3+d*X4+e*X5+f*X6+g*X7+h*X8+i*X9+j*X10+k$ となる。



Excelの「数式」タブの「fx」アイコンが選択されています。また、ワークシートのセルR2に「重回帰分析」として、 $y=a*x1+b*x2+c*x3+d*x4+e*x5+f*x6+g*x7+h*x8+i*x9+j*x10+k$ という式が入力されています。

セルR3には、係数aからjまでの値が、セルR4には定数kの値がそれぞれ入力されています。

青い吹き出しの指示: 「まず、a~kまでの係数を表示するためのセルを選択する」

青い吹き出しの指示: 「次に、LINEST関数を呼び出す」

予測モデルを作ってみよう

関数の引数

LINEST

既知のy	O3:O8761	= {2634;2520;2458;2389;2394;2491;2626
既知のx	E3:N8761	= {0,0,0,0,0,0,0,0,5,1,2,6,01;0,0,0,0,0,0,0,
定数	TRUE	=
補正	FALSE	= FALSE

数式の結果 = 5.082315031

[この関数のヘルプ\(H\)](#)

OK キャンセル

既定値を含む

補正項を求めない

目的変数のデータの範囲

説明変数のデータの範囲

CTRL + SHIFT + ENTER

法を使って直線を当てはめることで、既知のデータポイントに対応する線形トレンドを表す補正項を計算します。

既知のy には $y = mx + b$ となる、既にわかっている y 値の系列を指定します。

=LINEST(A3:A102, B3:D102, TRUE, FALSE)

販売量
の範囲

説明変数
の範囲

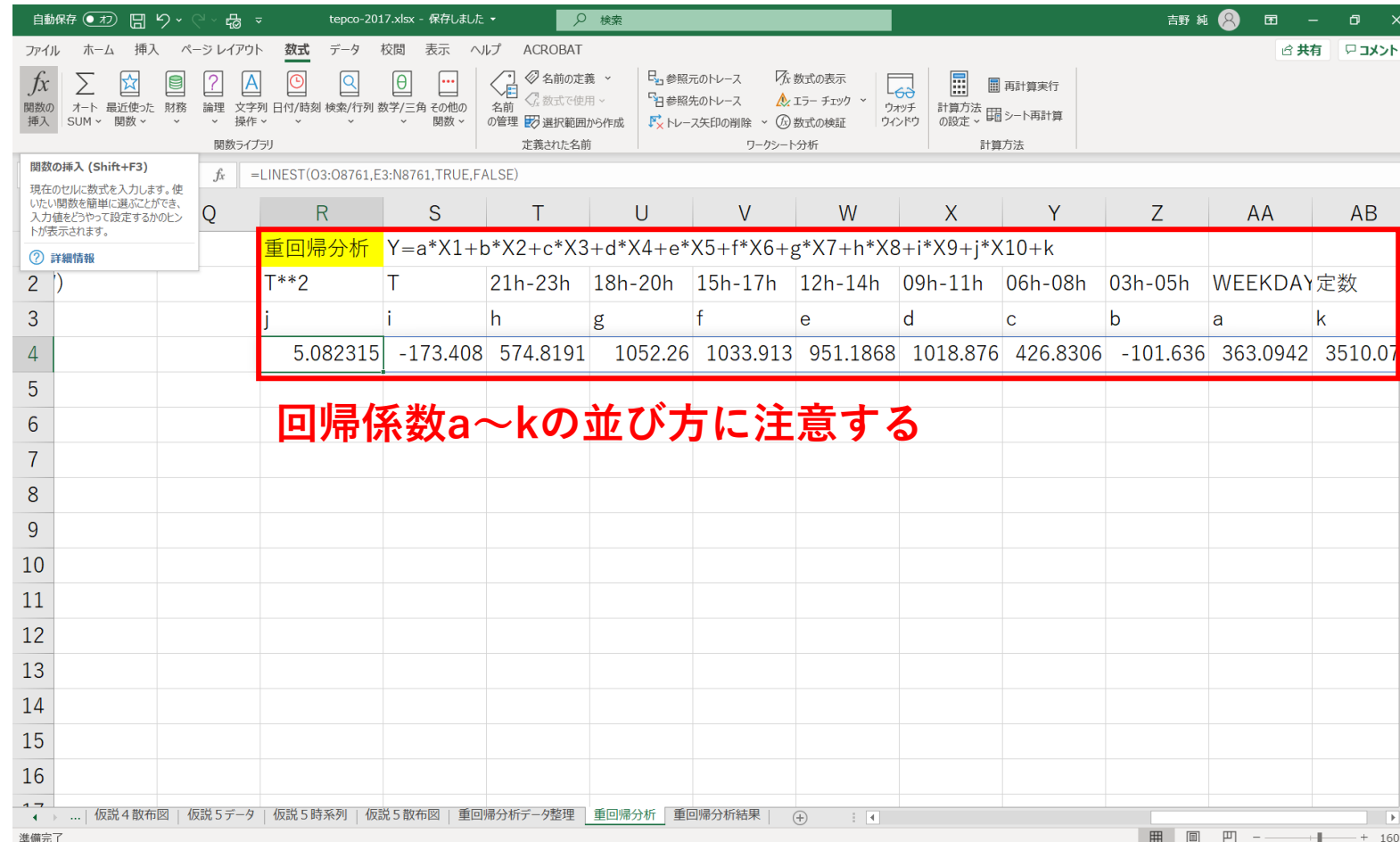
定数項
の有無

補正項
の有無

予測モデルを作ってみよう

重回帰式を求めることができた。

$$Y=363.09*X1-101.64*X2 + 426.83*X3 + 1018.88*X4 + 951.19*X5 + 1033.91*X6 + 1052.26*X7 + 574.82*X8 - 173.41*X9 + 5.08*X10 + 3510.07$$



重回帰分析

$$Y=a*X1+b*X2+c*X3+d*X4+e*X5+f*X6+g*X7+h*X8+i*X9+j*X10+k$$

重回帰分析	Y=a*X1+b*X2+c*X3+d*X4+e*X5+f*X6+g*X7+h*X8+i*X9+j*X10+k
T**2	T
j	i
5.082315	-173.408
	574.8191
	1052.26
	1033.913
	951.1868
	1018.876
	426.8306
	-101.636
	363.0942
	3510.07

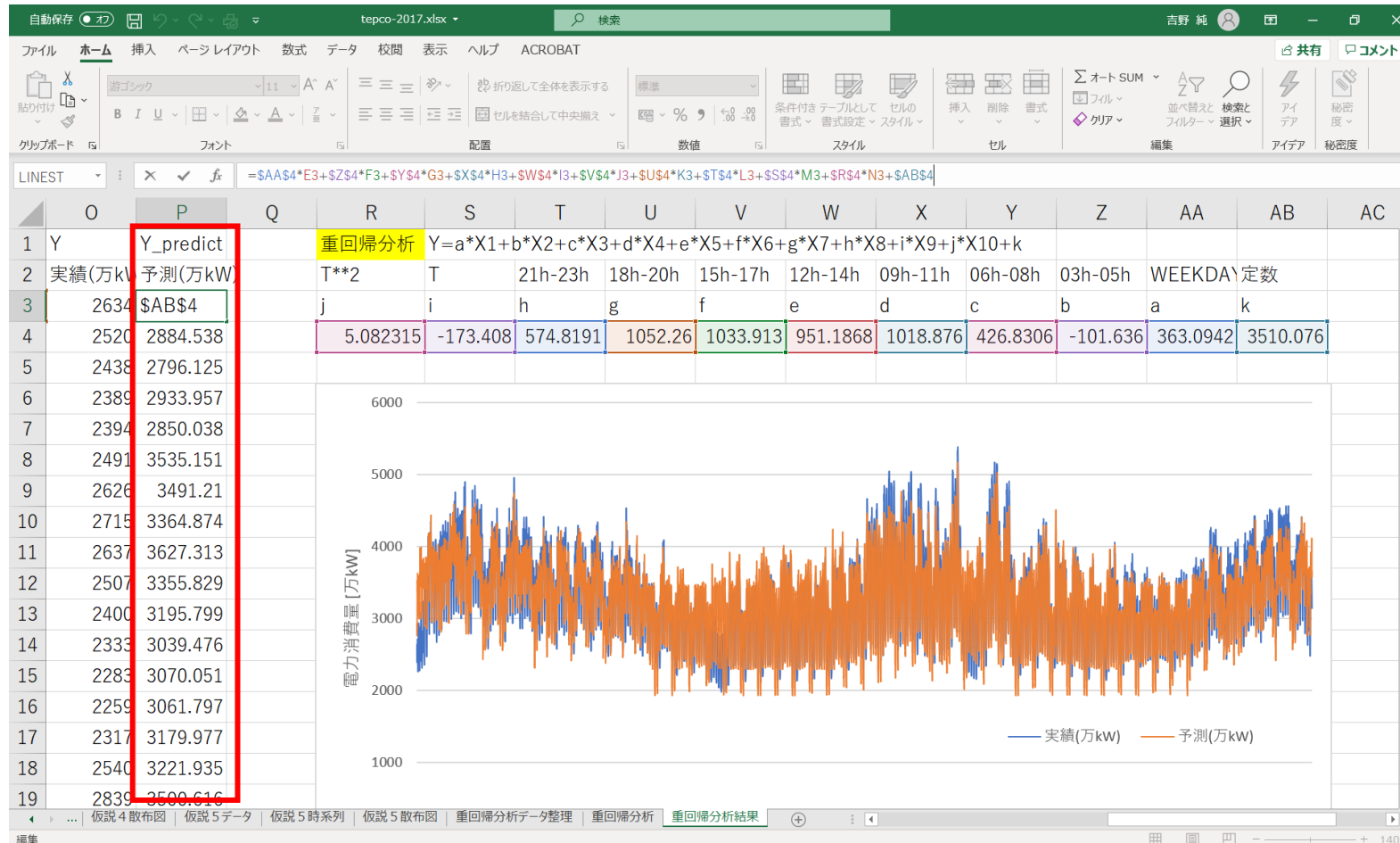
重回帰分析結果

重回帰分析により
回帰係数を計算で
きた

回帰係数a~kの並び方に注意する

予測モデルを作ってみよう

出来上がった重回帰式により電力消費量を予測（P列）し，実績値（O列）と比較する。
 時系列図と散布図を作成し，相関係数を計算する。



1年間を通じて，
 電力消費量の変化
 を適切に表現でき
 ている。

相関係数：0.90

やってみよう

- 気温と電力消費量との間には関係があり，正確に電力消費量を予測するためには，さらに，季節，時間，曜日等を考慮する必要があることがわかりました．
- 冬季のデータを使って気温と電力消費量との関係性を調べてみましょう．
- 2017年の重回帰式により，2018年の気象データを説明変数として電力消費量を精度良く予測できるか調べてみましょう．
- 2020年（コロナ禍）において，仮説2，仮説3，仮説4，仮説5は正しいでしょうか？また，2017年の重回帰式により2020年を予測できるでしょうか？調べてみましょう．

