

2024年2月17日

参加者各位

大田真次

## PBIJP Power Query 秘密特訓「虎の穴」炎の復活編

### #34

#### 内容

【自己紹介】 .....	2
【データの突合】 .....	2
【業務】 .....	2
【列項目】 .....	2
【Power Query の魅力】 .....	3
【元データ】 .....	3
【条件付き書式】 .....	4
【やりたいこと】 .....	5
【やること】 .....	5
【データの抽出】 → 【変換】 → 【読み込み】 (ELT) .....	6
【拡大図】 .....	7
【データ1とデータ2が抽出された】 .....	8
【拡大図】 .....	9
【データ1とデータ2が横につながった状態】 .....	10
【数値1からデータ2.数値1を減算することで相違を調べる】 .....	12
【数値2からデータ2.数値2を減算することで相違を調べる】 .....	12
【不要な列の削除】 .....	13
【項目名の変更】 .....	15
【違っているデータ】 .....	16
【読み込み】 .....	17
【Power BI のテーブルでビジュアル化】 .....	18

## 【自己紹介】

みなさん、こんにちは。私は、広島の病院で経理事務をしています。データの整理・加工には Excel をよく使います。私が業務で Power Query を使って実際に行っていることの一部をここで紹介します。

## 【データの突合】

データ1とデータ2

## 【業務】

業者が作成した請求書と病院が管理している台帳の突合

## 【列項目】

『日付』『ID』『数値 1』『数値 2』

参考までに【列項目】を業務用語に置き換えると↓のようになります。

『日付』	『検査日』
『ID』	『患者 ID』
『数値 1』	『レントゲン検査枚数』
『数値 2』	『金額』

4項目ですが、一月あたり 150 行前後あり、手作業でものさしを当てながらチェックするのはなかなか骨が折れます。順番もキレイには揃っていません。😅

## 【Power Query の魅力】

そこで、Power Query を使ってチェックすることを実施しています。

日付と金額の2項目なら、Excel の VLOOKUP 関数でできます。

もちろん、この課題も XLOOKUP を使えばできなくはないと思いますが、関数ではできない(ちょっと手間)ことが、Power Query ではできることが、魅力です。

それは、『複数の列でマージできること』と『完全外部(両方の行すべて)』です。

## 【元データ】

	A	B	C	D	E	F	G	H	I	J	K
1		データ 1						データ 2			
2											
3		日付	ID	数値 1	数値 2			日付	ID	数値 1	数値 2
4		2024/2/9	54	99	1,932			2024/2/9	54	99	1,933
5		2024/2/10	63	781	924			2024/2/10	63	781	924
6		2024/2/11	13	616	273			2024/2/11	13	616	273
7		2024/2/12	38	352	1,638			2024/2/12	37	693	1,071
8		2024/2/12	37	693	1,071			2024/2/12	38	352	1,638
9		2024/2/14	21	979	441			2024/2/14	21	980	441
10		2024/2/15	37	803	84						

説明用に、Excel ファイルの同じシートに配置していますが、実際はそれぞれが別ファイル (Excel と Csv) です。

## 【出来上がりのイメージ】

下図の赤枠の箇所をピックアップしたい

	A	B	C	D	E	F	G	H	I	J	K
1		データ 1						データ 2			
2											
3		日付	ID	数値 1	数値 2			日付	ID	数値 1	数値 2
4		2024/2/9	54	99	1,932			2024/2/9	54	99	1,933
5		2024/2/10	63	781	924			2024/2/10	63	781	924
6		2024/2/11	13	616	273			2024/2/11	13	616	273
7		2024/2/12	38	352	1,638			2024/2/12	37	693	1,071
8		2024/2/12	37	693	1,071			2024/2/12	38	352	1,638
9		2024/2/14	21	979	441			2024/2/14	21	980	441
10		2024/2/15	37	803	84						

## 【条件付き書式】

ちなみに Excel の条件付き書式で色を付けると(データ2に赤色)

	A	B	C	D	E	F	G	H	I	J	K	
1		データ 1						データ 2				
2												
3		日付	ID	数値 1	数値 2			日付	ID	数値 1	数値 2	
4		2024/2/9	54	9	1,932			2024/2/9	54	99	1,933	
5		2024/2/10	63	781	924			2024/2/10	63	781	924	
6		2024/2/11	13	616	273			2024/2/11	13	616	273	
7		2024/2/12	38	352	1,638			2024/2/12	37	693	1,071	
8		2024/2/12	37	693	1,071			2024/2/12	38	352	1,638	
9		2024/2/14	21	979	441			2024/2/14	21	980	441	
10		2024/2/15	37	803	84							

書式ルールの編集

ルールの種類を選択してください(S):

- セルの値に基づいてすべてのセルを書式設定
- 指定の値を含むセルだけを書式設定
- 上位または下位に入るセルだけを書式設定
- 平均より上または下のセルだけを書式設定
- 一意の値または重複するセルだけを書式設定
- 数式を使用して、書式設定するセルを決定**

ルールの内容を編集してください(E):

次の数式を満たす場合に値を書式設定(Q):

=B4<>H4

プレビュー: Aaあぁアァ亜宇

OK キャンセル

この方法だと同一行のみのチェックになります。

## 【やりたいこと】

改めて、やりたいことをまとめると  
データ1のデータ2を比べて、違うところを、調べたい。  
上から、

4 行目は、『数値2』が相違

7 行目 8 行目は、よくみると同じ行ではありませんが、順番が違うだけ  
(ここでは、合っていると考えます。)

9 行目は、『数値1』が相違

10 行目は、行ごと、データ 1 にはあるけど、データ 2 にはない

## これらを、Power Query を使って調べること

### 【やること】

データ1のデータ2を比べて、違うところを、調べること。  
Excel の条件付き書式では赤色をつけましたが、  
調べることが目的であり、違っているデータがわかればよしとします。

具体的には、

検索値:日付と ID として、データ1の数値1とデータ2の数値1が合っているか。

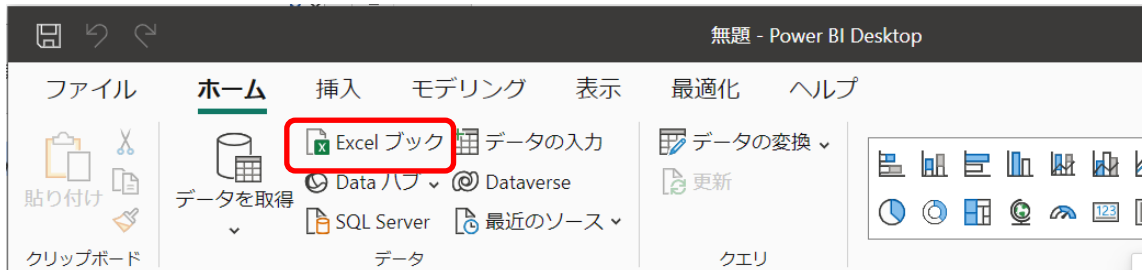
検索値:日付と ID として、データ1の数値2とデータ2の数値2が合っているか。

を調べること。

【データの抽出】 → 【変換】 → 【読み込み】 (ELT)

## データの抽出

[Excel ブック]をクリック



[データ 1 とデータ 2]をクリック

### ナビゲーター

表示オプション ▾

- 台帳照合.xlsxm [6]
  - データ1
  - データ1\_1
  - データ2
  - データ2\_2
  - 元データ
  - 条件付き書式

### データ2

日付	ID	数値1	数値2
2024/02/09	54	99	1933
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	37	693	1071
2024/02/12	38	352	1638
2024/02/14	21	980	441

## ナビゲーター

表示オプション

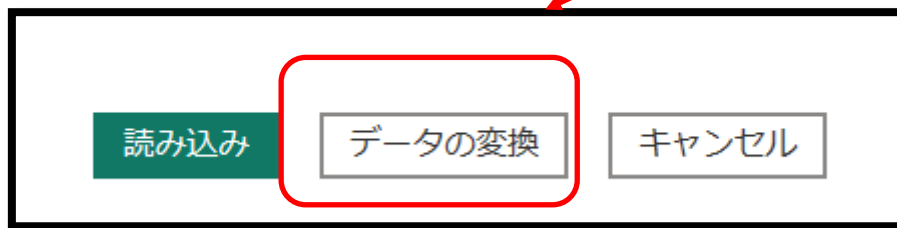
- 台帳照合.xlsx [6]
  - データ1
  - データ1\_1
  - データ2
  - データ2\_2
  - 元データ
  - 条件付き書式

## データ2

日付	ID	数値1	数値2
2024/02/09	54	99	1933
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	37	693	1071
2024/02/12	38	352	1638
2024/02/14	21	980	441

読み込み データの変換 キャンセル

【拡大図】



【データ1とデータ2が抽出された】

スクリーンショット1: Power Query Editorのクエリ「データ1」の表示画面。メニューバーには「ファイル」「ホーム」「変換」「列の追加」「表示」「ツール」「ヘルプ」があります。クエリ名は「クエリ [2]」で、選択されているクエリは「データ1」です。公式バーには以下の式が記載されています:

```
= Table.TransformColumnTypes(データ1_Table,{{"日付", type date}, {"ID", Int64.Type},
```

データビューには4つの列が表示されています: 「日付」、「ID」、「数値1」、「数値2」。各列には有効率、エラー率、空の割合の統計情報と、個別の値ごとの頻度分布グラフが示されています。データ行は以下の通りです:

日付	ID	数値1	数値2
2024/02/09	54	99	1932
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	38	352	1638
2024/02/12	37	693	1071
2024/02/14	21	979	441
2024/02/15	37	803	84

スクリーンショット2: Power Query Editorのクエリ「データ2」の表示画面。クエリ名は「クエリ [2]」で、選択されているクエリは「データ2」です。公式バーには以下の式が記載されています:

```
= Table.TransformColumnTypes(データ2_Table,{{"日付", type date}, {"ID", Int64.Type},
```

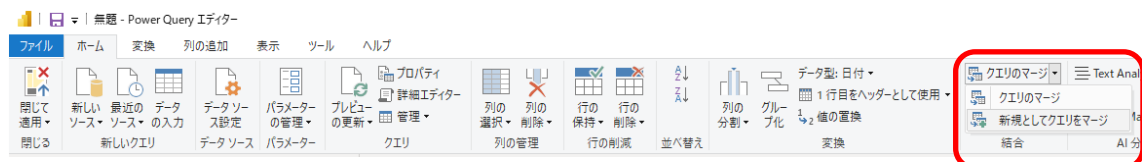
データビューには4つの列が表示されています: 「日付」、「ID」、「数値1」、「数値2」。各列には有効率、エラー率、空の割合の統計情報と、個別の値ごとの頻度分布グラフが示されています。データ行は以下の通りです:

日付	ID	数値1	数値2
2024/02/09	54	99	1933
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	37	693	1071
2024/02/12	38	352	1638
2024/02/14	21	980	441

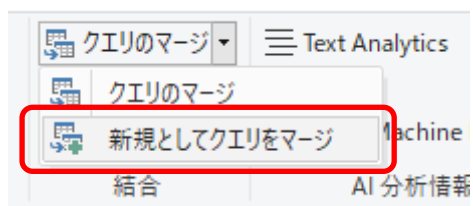


## 【変換】（マージ）

[クエリのマージ]をクリック



## 【拡大図】



[データ 1]を選択し、CTRL キーを押しながら[日付]と[ID]をクリック

[データ 2]を選択し、CTRL キーを押しながら[日付]と[ID]をクリック

## マージ

マージされたテーブルを作成するには、テーブルと照合列を選んでください。

データ1

日付	ID	数値1	数値2
2024/02/09	54	99	1932
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	38	352	1638
2024/02/12	37	693	1071

データ2

日付	ID	数値1	数値2
2024/02/09	54	99	1933
2024/02/10	63	781	924
2024/02/11	13	616	273
2024/02/12	37	693	1071
2024/02/12	38	352	1638

[完全外部 (両方のすべて)]をクリック

結合の種類

左外部 (最初の行すべて、および 2 番目の行のうち... ▼

左外部 (最初の行すべて、および 2 番目の行のうち一致するもの)

右外部 (2 番目の行すべて、および最初の行のうち一致するもの)

完全外部 (両方の行すべて)

内部 (一致する行のみ)

左反 (最初の行のみ)

右反 (2 番目の行のみ)

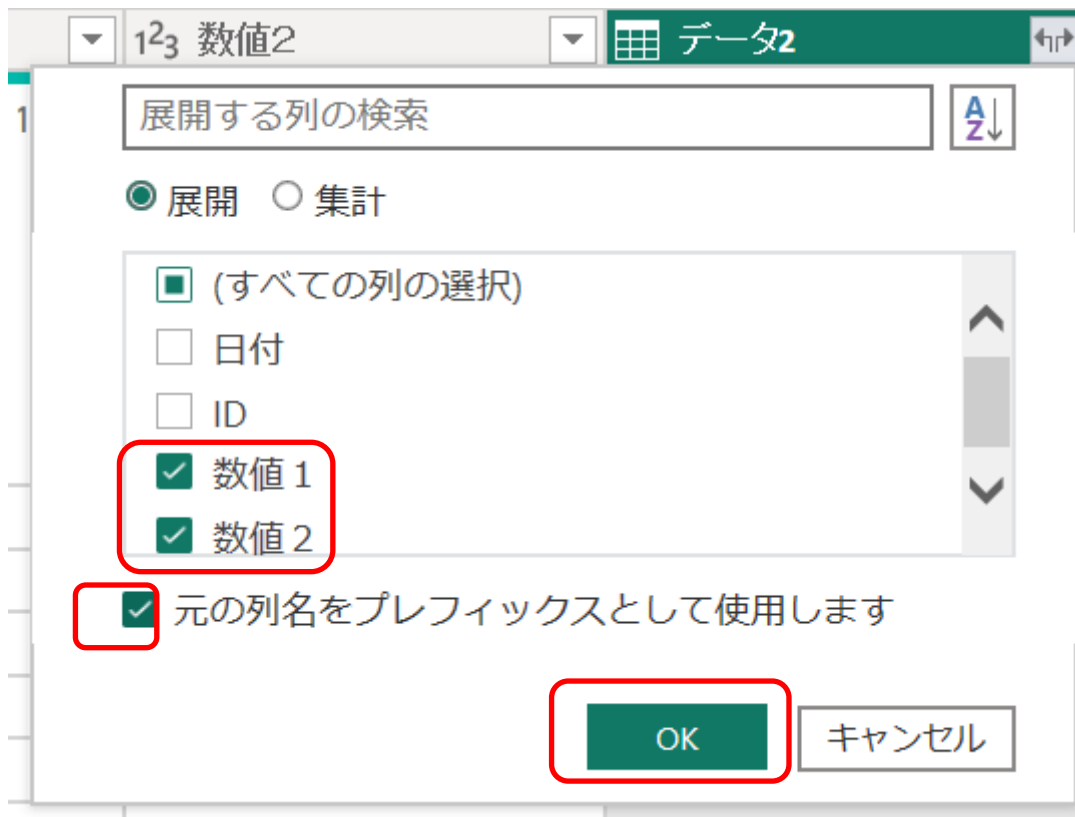
OK

キャンセル

【データ 1 とデータ 2 が横につながった状態】

日付	123 ID	123 数値1	123 数値2	データ2
● 有効 100% ● エラー 0% ● 空 0%	● 有効 100% ● エラー 0% ● 空 0%	● 有効 100% ● エラー 0% ● 空 0%	● 有効 100% ● エラー 0% ● 空 0%	● 有効 100% ● エラー 0% ● 空 0%
個別 6 個、一意 5 個	個別 6 個、一意 5 個	個別 7 個、一意 7 個	個別 7 個、一意 7 個	
1	2024/02/09	54	99	1932 Table
2	2024/02/10	63	781	924 Table
3	2024/02/11	13	616	273 Table
4	2024/02/12	38	352	1638 Table
5	2024/02/12	37	693	1071 Table
6	2024/02/14	21	979	441 Table
7	2024/02/15	37	803	84 Table

【データ2の展開】

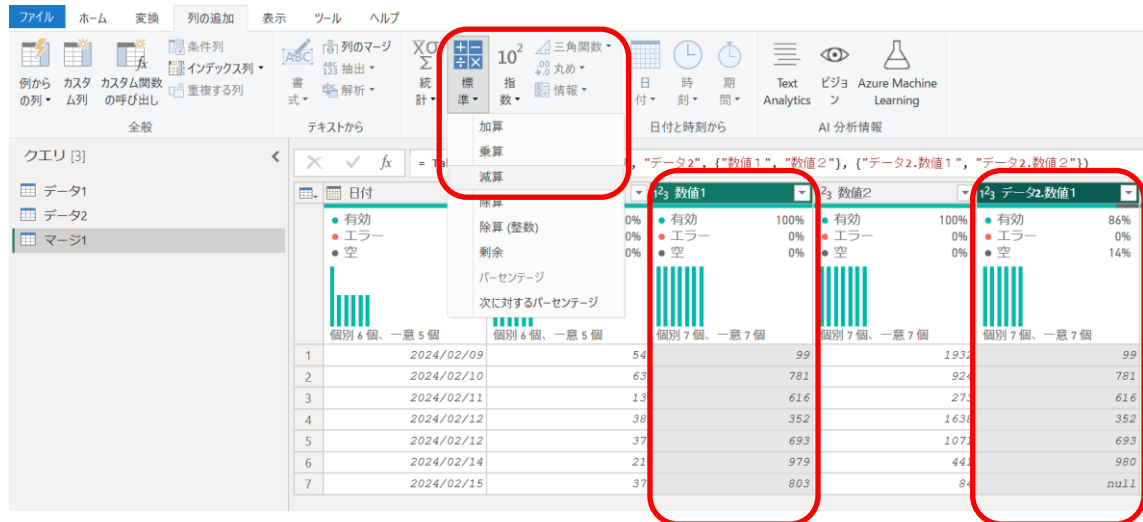


【データ2の数値1と数値2を表示】

	日付	i23 ID	i23 数値1	i23 数値2	i23 データ2.数値1	i23 データ2.数値2
1	2024/02/09		54	99	1932	99
2	2024/02/10		63	781	924	781
3	2024/02/11		13	616	273	616
4	2024/02/12		38	352	1638	352
5	2024/02/12		37	693	1071	693
6	2024/02/14		21	979	441	980
7	2024/02/15		37	803	84	null

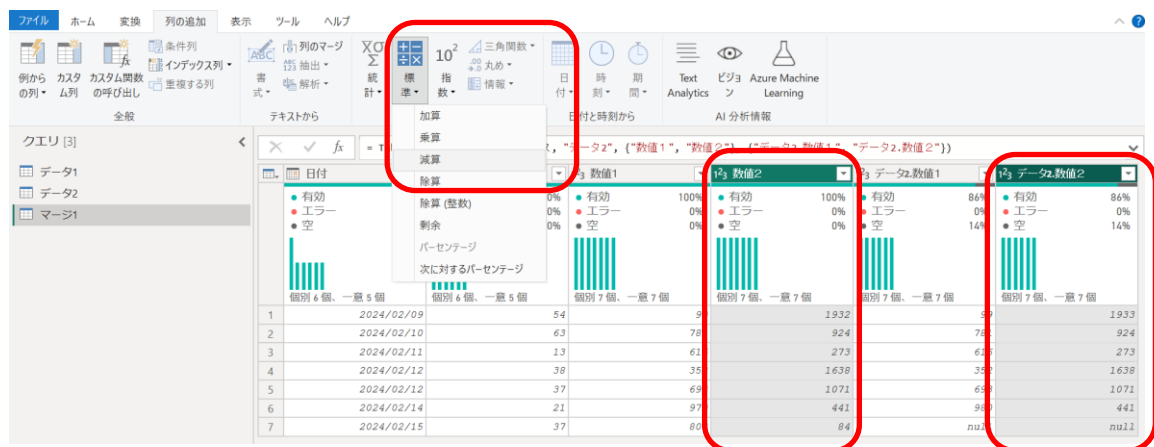
### 【数値1からデータ2.数値1を減算することで相違を調べる】

[数値1→データ2.数値1の順に選択し、[標準]→[減算]]をクリック



### 【数値2からデータ2.数値2を減算することで相違を調べる】

[数値2→データ2.数値2の順に選択し、[標準]→[減算]]をクリック



## 【不要な列の削除】

[数値 1、数値 2 を選択し、列の削除]をクリック

The screenshot shows a spreadsheet with two columns, '123 数値1' and '123 数値2', highlighted in green. A context menu is open over the '123 数値2' column, with the '列の削除' (Delete Column) option highlighted. The menu also includes options like 'コピー' (Copy), '他の列の削除' (Delete Other Columns), '例から列を追加する...' (Add Columns from Examples...), '重複の削除' (Delete Duplicates), 'エラーの削除' (Delete Errors), '値の置換...' (Replace Values...), 'フィル' (Filter), '型の変更' (Change Type), '変換' (Convert), '列のマージ' (Merge Columns), '合計' (Sum), '積' (Product), 'グループ化...' (Group...), '列のピボット解除' (Unpivot Columns), 'その他の列のピボット解除' (Unpivot Other Columns), '選択した列のみをピボット解除' (Unpivot Selected Columns), and '移動' (Move).

	123 数値1	123 数値2
	有効 100%	有効 100%
	エラー 0%	エラー 0%
	空 0%	空 0%
	個別 7 個、一意 7 個	個別 7 個、一意 7 個
4	99	
3	781	
3	616	
8	352	
7	693	
1	979	
7	803	

[データ 2.数値 1、データ 2.数値 2 を選択し、列の削除]をクリック

The screenshot shows a pivot table with three columns: 'データ2.数値1', 'データ2.数値2', and '減算'. The 'データ2.数値1' and 'データ2.数値2' columns are selected. A context menu is open over the 'データ2.数値2' column, and the '列の削除' option is highlighted with a red box. The menu also includes options like '他の列の削除', '例から列を追加する...', '重複の削除', 'エラーの削除', '値の置換...', 'フィル', '型の変更', '変換', '列のマージ', '合計', '積', 'グループ化...', '列のピボット解除', 'その他の列のピボット解除', '選択した列のみをピボット解除', and '移動'.

データ2.数値1	データ2.数値2	減算
有効 86%	有効 86%	
エラー 0%	エラー 0%	
空 14%	空 14%	
個別 7 個、一意 7 個	個別 7 個	
		99
		781
		616
		352
		693
		980
		null

【項目名の変更】

[減算を、減算\_数値1へ、減算.1を減算\_数値2へ]変更

日付	123 ID	123 減算	123 減算.1
<ul style="list-style-type: none"> <li>有効 100%</li> <li>エラー 0%</li> <li>空 0%</li> </ul> <p>個別6個、一意5個</p>	<ul style="list-style-type: none"> <li>有効 100%</li> <li>エラー 0%</li> <li>空 0%</li> </ul> <p>個別6個、一意5個</p>	<ul style="list-style-type: none"> <li>有効 86%</li> <li>エラー 0%</li> <li>空 14%</li> </ul> <p>個別3個、一意2個</p>	<ul style="list-style-type: none"> <li>有効 86%</li> <li>エラー 0%</li> <li>空 14%</li> </ul> <p>個別3個、一意2個</p>
1	2024/02/09	54	0
2	2024/02/10	63	0
3	2024/02/11	13	0
4	2024/02/12	38	0
5	2024/02/12	37	0
6	2024/02/14	21	-1
7	2024/02/15	37	null

日付	123 ID	123 減算_数値1	123 減算_数値2
<ul style="list-style-type: none"> <li>有効 100%</li> <li>エラー 0%</li> <li>空 0%</li> </ul> <p>個別6個、一意5個</p>	<ul style="list-style-type: none"> <li>有効 100%</li> <li>エラー 0%</li> <li>空 0%</li> </ul> <p>個別6個、一意5個</p>	<ul style="list-style-type: none"> <li>有効 86%</li> <li>エラー 0%</li> <li>空 14%</li> </ul> <p>個別3個、一意2個</p>	<ul style="list-style-type: none"> <li>有効 86%</li> <li>エラー 0%</li> <li>空 14%</li> </ul> <p>個別3個、一意2個</p>
1	2024/02/09	54	0
2	2024/02/10	63	0
3	2024/02/11	13	0
4	2024/02/12	38	0
5	2024/02/12	37	0
6	2024/02/14	21	-1
7	2024/02/15	37	null

【違うデータ】

減算\_数値1と減算\_数値2の0以外が違っていることになる

日付	1 <sup>2</sup> ID	1 <sup>2</sup> 減算_数値1	1 <sup>2</sup> 減算_数値2
1	2024/02/09	54	-1
2	2024/02/10	63	0
3	2024/02/11	13	0
4	2024/02/12	38	0
5	2024/02/12	37	0
6	2024/02/14	21	-1
7	2024/02/15	37	null

ここでいったん終了。

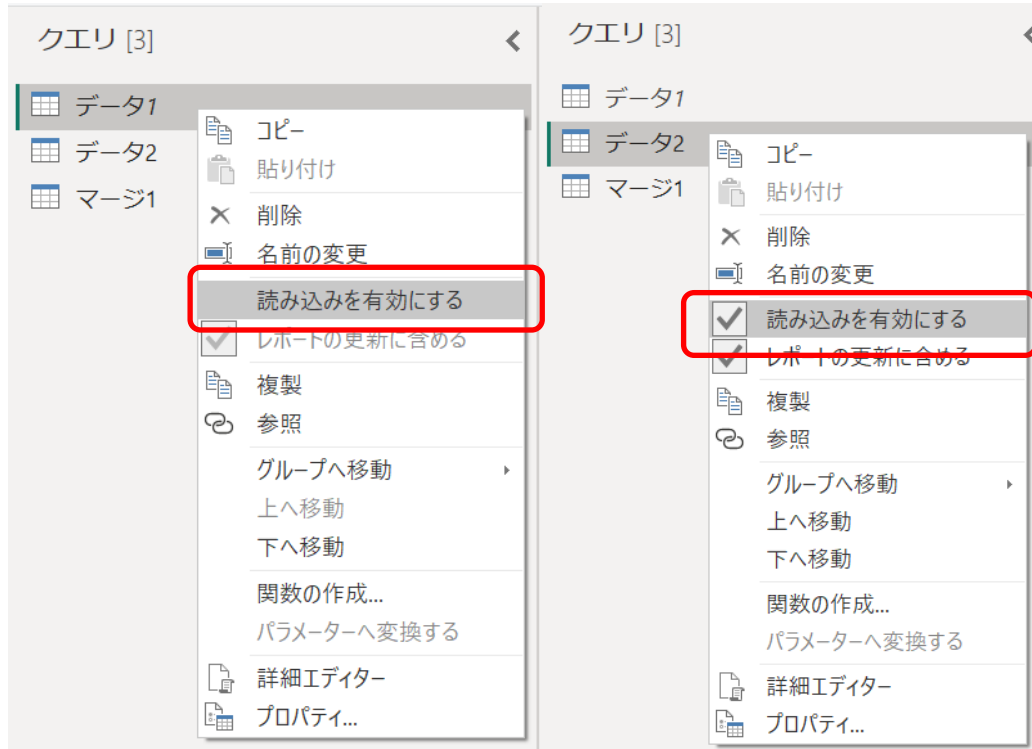


## 【読み込み】

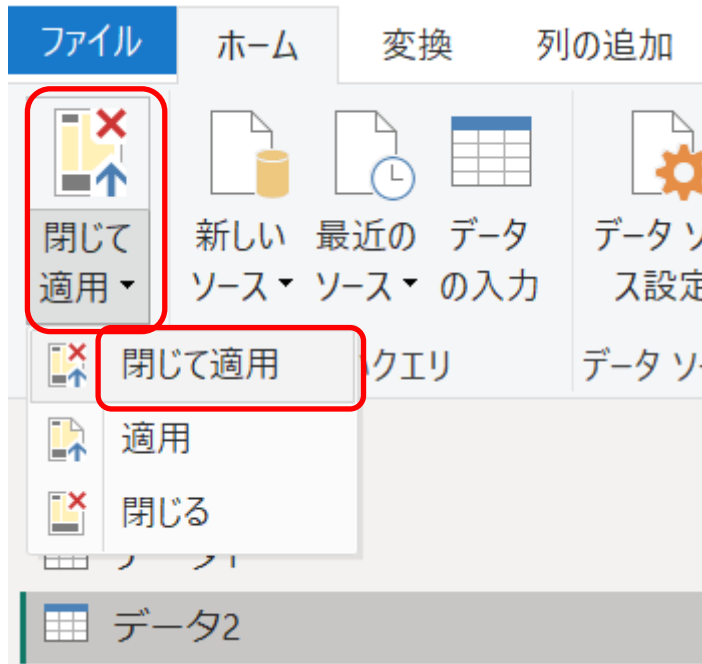
Power BI に読み込む場合は、マージ 1 のみとする

読み込みを有効にするの  を外す

理由：データ 1 とデータ 2 は読み込み不要のため



[閉じて適用]



【Power BI のテーブルでビジュアル化】

The image shows the Microsoft Power BI ribbon with the 'Home' tab selected. The ribbon includes 'File', 'Home', 'Insert', and 'Modeling' tabs. The 'Home' tab is active, showing icons for 'Paste' (貼り付け), 'Get Data' (データを取得), 'Excel Workbook' (Excelブック), 'Data Hub' (Dataハブ), 'Data Warehouse' (Dataverse), 'SQL Server', and 'Recent Sources' (最近のソース). Below the ribbon, a table visualization is shown with the following data:

月	日	IDのカウンタ	減算_数値1の合計	減算_数値2の合計
February	9	1	0	-1
February	10	1	0	0
February	11	1	0	0
February	12	2	0	0
February	14	1	-1	0
February	15	1		
<b>合計</b>		<b>7</b>	<b>-1</b>	<b>-1</b>