

# データの確認①

「SSDSE-C」を使った分析例の、「データ加工」について解説しています。

データ分析を始める前に、扱うデータがどんなデータであるかを確認します。  
例えば、調査の対象、調査の規模（全体の何%くらいを調査したのか）、調査の時期、集計方法（実数か、平均かなど）などがどうなっているか、確認します。

① SSDSE-家計消費の解説を開いて内容を確認します。

独立行政法人 統計センター

文字サイズ 中 日本語 English Googleサイト内検索 検索

統計センターについて 統計をつくる 統計を活かす 統計を支える 統計技術 採用案内

正確な統計の作成のために 統計の利用者・研究者の皆様へ 各府省・自治体の統計作成を支援

## SSDSE-家計消費(SSDSE-C)

都道府県庁所在市別の家計消費データを集めたデータセットです。  
(全国・47都道府県庁所在市×家計消費226項目)  
(出典) 総務省統計局「家計調査」2022年(令和4年)～2024年(令和6年)

	SSDSE (EXCEL)	SSDSE (CSV)	SSDSEの解説
2025年4月25日	<a href="#">SSDSE-C-2025</a>	<a href="#">SSDSE-C-2025</a>	<a href="#">SSDSE-家計消費の解説</a> (PDF:1,697KB)

1

2025年4月25日 SSDSE-C-2025 SSDSE-C-2025 SSDSE-家計消費の解説 (PDF:1,697KB)

開いたSSDSE-Cは[こちらから](#)

### 収録しているデータ

収録地域 全国 及び 47の都道府県庁所在市

#### 地域コード

- 政府統計の総合窓口(e-Stat)では、数字5桁の市区町村コードが使われていますが、このデータセットでは、冒頭にRを付けた6桁のコードを使用しています。  
(「全国」の場合は、R00000)

5桁の市区町村コード

R 0 1 1 0 0

2桁の都道府県コード 1: 政令指定都市、2: 他の県庁所在市

収録項目 都道府県庁所在市別、二人以上の世帯の1世帯当たり、  
品目別(食料の全品目) 年間支出金額

- 総務省統計局「家計調査」から収録
- 収録しているデータ項目の一覧は、別表をご覧ください。
- 東京都は「東京都都区部」のデータです。
- 家計調査の「消費支出」は、10の大分類(10大費目)に分かれ、「食料」は、その中の一つです。SSDSE-Cには、「食料」の全品目(212)のほか、12の中分類(小計)も収録しています。

収録年次 2022年～2024年の平均値

$$= (2022\text{年暦年データ} + 2023\text{年暦年データ} + 2024\text{年暦年データ}) \div 3$$

※ 四捨五入して整数値としている。「世帯人員」のみ四捨五入して小数点以下2桁までの数値。

### 解説から分かったこと

- 出典: 総務省統計局が実施している「家計調査」を加工して作成したデータ
- 項目: 「二人以上の世帯」の1世帯当たり品目別年間支出金額
- 地域: 全国、47の都道府県庁所在市別のデータ
- 時点: 2022年から2024年の平均値

統計で見る京都府のすがた

# データの確認②

できれば、出典元の「家計調査」についてもどんな調査か確認できるとデータの理解が深まります。

## ・ [統計局ホームページ/家計調査](#)

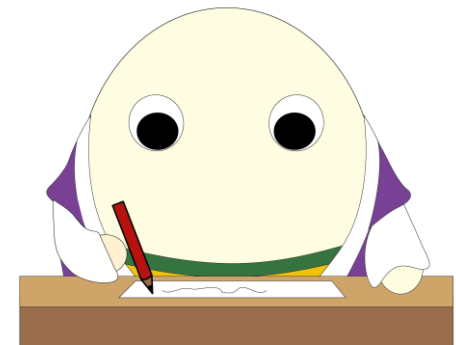
調査の概要から分かったこと

- 👉 調査の対象：全国の「世帯」（特定の世帯は対象外）
- 👉 調査世帯：標本調査であり、全国の168市町村から調査世帯約8700世帯を選定  
→全世帯を調べている訳でなく、調査対象でない市町村が多い（京都府内だと、京都市、長岡京市、木津川市から調査世帯を選定）  
→二人以上の世帯は6か月、単身世帯は3か月継続して調査される
- 👉 時期：毎月実施 など

「家計調査」は全国の一部の世帯にお願いして、毎月の買い物などでつかったお金（支出）や収入などを調べる調査です。

ただ、調べる世帯の数が多くはないので、1ヶ月の平均、1年の平均だと「たまたま」の影響が出ることがあります。（たまたま調査世帯で大きな出費があった など）

SSDSEでは、3年分の平均を使うことで「たまたま」をならして全体の傾向をつかんだデータに加工しているんですね。



# データの加工＞不要なデータの削除①

The image displays two side-by-side screenshots of the SSDSE-C-2025 spreadsheet application. The left screenshot shows the spreadsheet with column D selected, indicated by a yellow circle with the number 1 and a red box. The right screenshot shows the spreadsheet with columns D through DQ selected, indicated by a yellow circle with the number 2 and a red box. A context menu is open over the selected columns, with the 'Delete (D)' option highlighted, indicated by a yellow circle with the number 3 and a red box. The spreadsheet data includes columns for City, DK, DL, DP, DQ, DR, DS, DT, and others, with rows for various regions and products.

同じ要領で、T列（油脂・調味料）以降の列も削除します。

今回は果物のデータを使いたいので、それ以外の項目、品目のデータを列ごと削除します。

- ① D列の表頭にカーソルを合わせ、クリックします。D列全体が選択されます。
- ② D列を選択したまま、スクロールバーでDQ列まで移動し、「shift」キーを押しながらDQ列の表頭をクリックします。D～DQ列が選択されます。
- ③ D～DQ列が選択された状態で、右クリックし、削除を選択します。

# データの加工＞不要なデータの削除②

1 2 3

メニューの検索

- 切り取り(I)
- コピー(C)
- 貼り付けのオプション:
- 形式を選択して貼り付け(S)...
- 挿入(I)
- 削除(D)
- 数式と値のクリア(N)
- クイック分析(Q)
- フィルター(E)
- 並べ替え(O)
- テーブルまたは範囲からデータを取得(G)...
- 新しいコメント(M)
- 新しいメモ(N)

地域コード	都道府県	市	06 果物	りんご
R00000	全国	全国	41349	4676
R01000	北海道	札幌市	38466	4616
R02000	青森県	青森市	40284	8800
R03000	岩手県	盛岡市	44132	8218
R04100	宮城県	仙台市	42336	4543
R05201	秋田県	秋田市	45700	6732
R06201	山形県	山形市	54312	6835
R07201	福島県	福島市	45746	6331

A～S (19) 列、48行の  
データになりました。

不要な行も削除します。

- ① 1行目の表側にカーソルを合わせ、クリックします。1行目全体が選択されます。
- ② 1行目を選択したまま、「ctrl」キーを押しながら3行目の表側をクリックします。1行目、3行目が選択されます。
- ③ 1行目、3行目が選択された状態で、右クリックし、削除を選択します。

# 代表値、四分位数を求める＞代表値、四分位数の計算①

The image illustrates the process of calculating statistics across multiple sheets in Excel. It consists of three main screenshots showing the steps:

- Step 1:** A new sheet named "SSDSE-C-2025" is added to the workbook. The sheet contains a table with columns for "地域コード" (Region Code), "都道府県" (Prefecture), "市" (City), and "06 果物" (Fruit). The data includes various regions like Hokkaido, Aomori, Iwate, Miyagi, Akita, Yamagata, Fukushima, Aichi, Nagano, Tokyo, and Chiba.
- Step 2:** The formula bar shows the input of the formula `=max(` in cell B2. The formula bar also shows the range `D2:D48` being selected.
- Step 3:** The formula bar shows the completion of the formula `=max(D2:D48)` in cell B2. The result of the calculation, 54,312, is displayed in cell B2.

別のシートに、代表値・四分位数を計算します。

- ① シートの「+」をクリックして新しいタブを開き、各項目をコピー、入力します。
- ② 各項目を計算します。最大値の場合、入力するセルに「=MAX(」と入力し、元のタブに戻ります。
- ③ 「D2:D48」と入力するか、セルを選択して「Enter」キーを押します。

最大値が入力できました。

## 代表値、四分位数を求める＞代表値、四分位数の計算②

B2		=MAX('SSDSE-C-2025'!D			
	A	B	C	D	E
1		06 果物	りんご	みかん	オレンジ
2	最大値	54,312			
3	最小値	=MIN(D2:D48) 1			
4	範囲				
5	平均値				
6	第1四分位数				
7	第2四分位数	=QUARTILE.EXC(D2:D48,2)			
8	第3四分位数	=QUARTILE.EXC(D2:D48,3)			
9					
10					

同様に、他の代表値と四分位数を計算します。

- ① エクセルで計算式を入力するか、関数を入力して計算します。  
各値は左の式で計算できます。

### Excelでの四分位数の計算結果

- QUARTILE.EXC関数は、データ数が奇数・偶数に関わらず、中央値を除いて四分位数を計算します。
- データ数が奇数の場合は教科書の方法で求めた値と一致しますが、データ数が偶数の場合は教科書の方法で求めた値と一致しないことがあります。
- ただしデータの数が多ければ、求め方による値の違いは小さくなるので、傾向を分析する上では問題にはなりません。

## 代表値、四分位数を求める＞代表値、四分位数の計算③

B2		fx		=MAX('SSDSE-C-2025'!D\$2:D\$48)													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		06 果物	りんご	みかん	オレンジ	他の柑橘	梨	ぶどう	柿	桃	すいか	メロン	いちご	パナナ	キウイ	他の果物	果物加工品
2	最大値	54312															
3	最小値	32217															
4	範囲	22095															
5	平均値	41719.81															
6	第1四分位数	38513															
7	第2四分位数	41458															
8	第3四分位数	45066															
9		1															
10																	

他の項目の代表値と四分位数も計算します。

① 関数を入力して計算したときは、コピーすると簡単です。

コピーする範囲（今回は、B2からB8のセル）を選択し、右下の角にカーソルを合わせると、エクセルの十字カーソルが変わります。（ → ）

「+」のカーソルになったら左クリックし、右方向にQ列までドラッグするとコピーできます。

コピーしたときは、関数の参照先が正しくコピーされているか、確認しましょう。