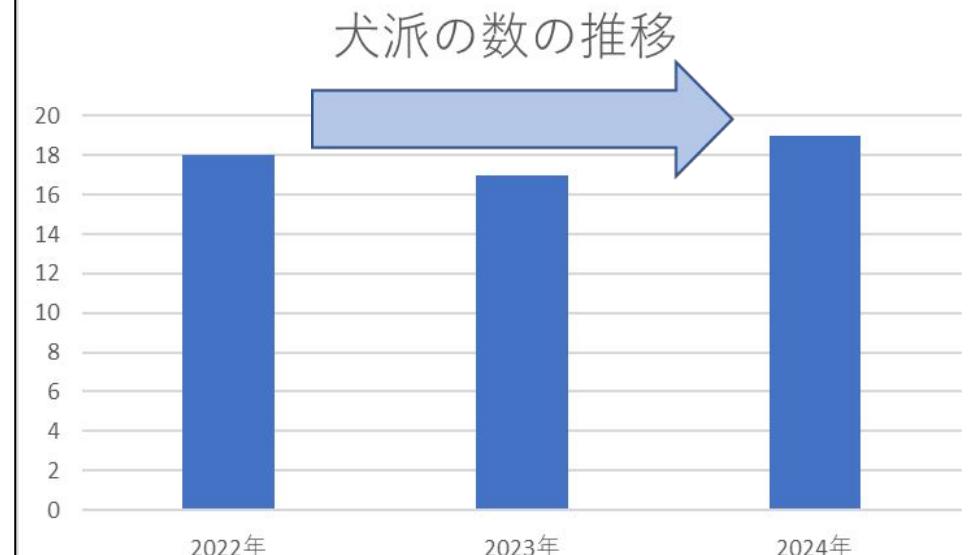
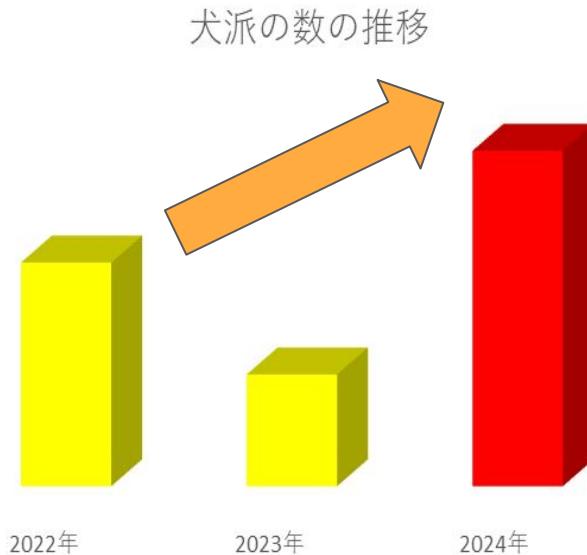


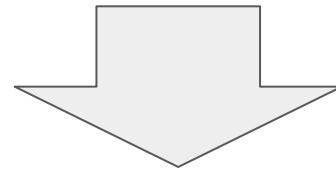
データサイエンス講座

データサイエンスを学ぶにあたって



データサイエンスを学ぶにあたって

データを正しく扱い相手に伝える力



自説に強い根拠や説得力が生まれる。



目的

具体的なデータを題材に数値データの取り扱いに関する一連の手法を学び、課題探究の充実を図る。

今後の日程

- 講座① 表計算アプリケーションを活用して統計量を求める。
- 講座② 表計算アプリケーションを活用してグラフを作成する。
- 講座③ 表計算アプリケーションを活用してデータを分析する。
- 講座④ 確認テスト

データサイエンス講座①

表計算アプリケーションを活用して統計量を求める。

《統計量の復習》

データの分析で学んだことを思い出し、教科書も参考にして①の問題に取り組もう！

平均値 $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$

分散 $s_x^2 = \frac{1}{n} \{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2\}$

標準偏差 $s_x = \sqrt{s_x^2}$

共分散 $s_{xy} = \frac{1}{n} \{(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})\}$

1(1) 解答

番号	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
①	36	27	-1	-6	1	36	6
②	42	39	5	6	25	36	30
③	35	35	-2	2	4	4	-4
④	33	25	-4	-8	16	64	32
⑤	38	41	1	8	1	64	8
⑥	32	23	-5	-10	25	100	50
⑦	39	43	2	10	4	100	20
⑧	40	31	3	-2	9	4	-6
⑨	34	29	-3	-4	9	16	12
⑩	41	37	4	4	16	16	16
計	370	330	0	0	110	440	164

x の最大値 42	x の最小値 32	y の最大値 43	y の最小値 23
x の分散 $S_x^2 = 11$	x の標準偏差 $S_x = 3.32$	y の分散 $S_y^2 = 44$	y の標準偏差 $S_y = 6.63$
x と y の共分散 $S_{xy} = 16.4$		x と y の相関係数 $r = 0.75$	

1(1) 解答

番号	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
①	36	27	-1	-6	1	36	6
②	42	39	5	6	25	36	30
③	35	35	-2	2	4	4	-4
④	33	25	-4	-8	16	64	32
⑤	38	41	1	8	1	64	8
⑥	32	23	-5	-10	25	100	50
⑦	39	43	2	10	4	100	20
⑧	40	31	3	-2	9	4	-6
⑨	34	29	-3	-4	9	16	12
⑩	41	37	4	4	16	16	16
計	370	330	0	0	110	440	164
							9

1(1) 解答

x の最大値	x の最小値	y の最大値	y の最小値
4.2	3.2	4.3	2.3
x の分散	x の標準偏差	y の分散	y の標準偏差
$S_x^2 = 11$	$S_x = 3.32$	$S_y^2 = 44$	$S_y = 6.63$
x と y の共分散		x と y の相関係数	
$S_{xy} = 16.4$		$r = 0.75$	

1(2) 解答

($r=0.75$ であることから)

右手の握力と左手の握力の大きさには正の相関があると考えられる。

《表計算アプリで統計量を計算しよう》

『学年クラスルーム/探究基礎/データサイエンス①』の色付きセルに数式等を入れて、統計量を求めよ。

=C2 - \$L\$4

(C2セルの値) - (L5セルの値)

=E2^2

E2セルの値の2乗

A	B	C	D	G	H	I	J	K	L	M
番号	地域	x : 年間収入 【千円】	y : 食料費 （全世帯） 【円】	xの 偏差	(xの偏差)^2	(yの偏差)^2	(xの偏差) × (yの偏差)		x : 年間収入 【千円】	y : 食料費 （全世帯） 【円】
1	北海道	4488	63638	=C2-\$L\$4	=E2^2				最大値	=max(C2:C48)
2	青森県	4952	65700						最小値	=min(C2:C48)
3	岩手県	5282	68745							
4	宮城県	5702	67836							
5	秋田県	5274	66097							
6	山形県	5856	68382							
7	福島県	5397	65103							
8	茨城県	5956	65872							
47	沖縄県	4215	50801						xとyの共分散	
	計	=SUM(C2:C48)								xとyの相関係数

=MAX(C2:C48)

C2セルからC48セルまでの合計値

データサイエンス講座②

表計算アプリケーションを活用してグラフを作成する

《復習》zとx, wとxの相関係数を調べる。

0.100

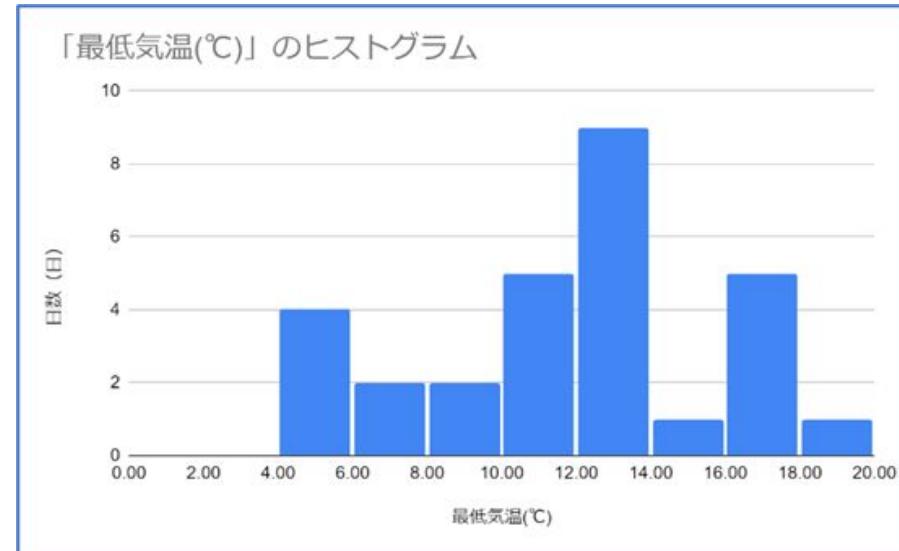
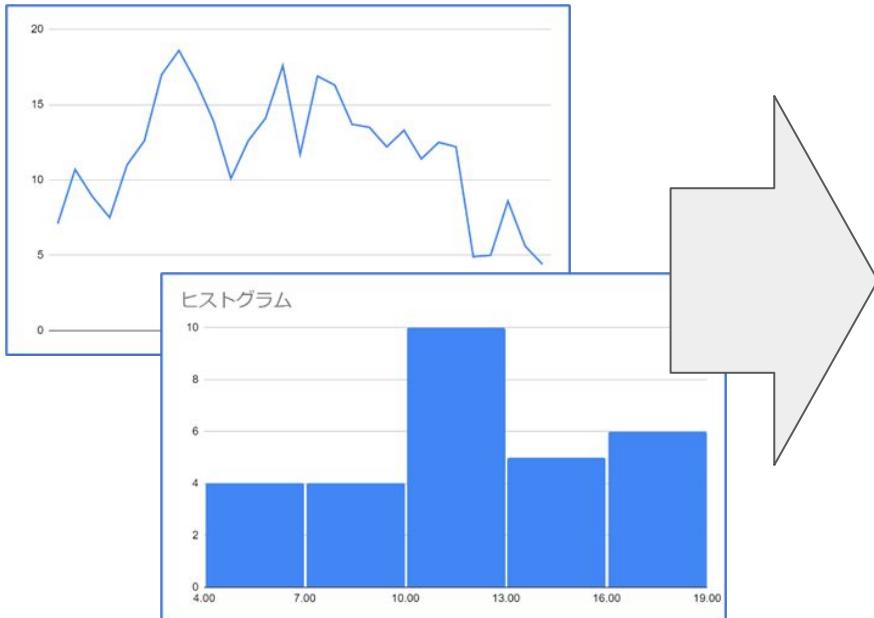
x - z の相関係数 =correl(D2:D48,C2:C48)

x - w の相関係数 =correl(E2:E48,C2:C48)

0.079

【ワーク①】ヒストグラムの作成

クラスルーム/探求基礎/データサイエンス講座②を開いて、グラフ機能を活用してデータを可視化しよう。



【ワーク①】ヒストグラムの作成

The screenshot shows the R6 Data Science Platform interface. A table titled "ある月の最低気温 (°C)" is displayed in the A:A range. The "Insert" menu (挿入) is open, with the "Graph" option (グラフ) highlighted by a blue oval. Another blue oval highlights the first column of the table, which contains the minimum temperatures for each day of the month.

A	B
1 ある月の最低気温 (°C)	
2 7.1	
3 10.1	
4 8.9	
5 7.5	
6 11	
7 12.6	
8 17	
9 18.6	
10 16.5	
11 13.9	
12 10.1	
13 12.6	
14 14.1	
15 17.6	
16 11.1	
17 16.9	
18 10.3	
19 3.7	

列Aを選択しながら,
挿入 → グラフ

【ワーク①】ヒストグラムの作成

R6データサイエンス（生徒用） ☆ 回 共有 幹太

ファイル 編集 表示 挿入 表示形式 データ ツール 拡張機能 ヘルプ

100% | ¥ % .0 .00 123 | デフォ... | - 10 + | B I ÷ A | 田 25 | : | ^

A:A ある月の最低気温 (°C)

A	ある月の最低気温 (°C)
1	7.1
2	10.7
3	8.9
4	7.5
5	11
6	12.6
7	17
8	18.6
9	16.5
10	13.9
11	10.1
12	12.6
13	14.1
14	17.6
15	11.7
16	16.9
17	16.3
18	13.7
19	

「ある月の最低気温 (°C)」のヒストグラム

ある月の最低気温 (°C)

設定 カスタマイズ

- > グラフの種類
- > ヒストグラム
- > グラフと軸のタイトル
- > 系列
- > 凡例
- > 横軸
- > 縦軸

グラフエディタ

17

【ワーク②】度数分布表の作成

	A	B	C	D	E	F	G	H
1	地域	ガソリン1Lの販売価格（円）						
2	札幌市	180						
3	函館市	175			~	165		
4	旭川市	180			~			
5	青森市	172			~			
6	八戸市	169			~			
7	盛岡市	166			~			
8	仙台市	176			~			
9	秋田市	172			~			
10	山形市	189			~			
11	福島市	182			計			
12	郡山市	177						

G3セルに
=frequency(B:B,F:F)
を入力

【ワーク②】ヒストグラムの作成

The screenshot shows a spreadsheet application window titled "R6データサイエンス (生徒用)". The table has columns A and B, with data from row 1 to 20. Column A contains city names and row numbers, while column B contains values ranging from 172 to 189. A context menu is open over column B, with the "Insert" tab selected. Other options visible in the menu include "Timeline", "Chart", "PivotTable", "Image", "Diagram", "Function", "Link", "Checkboxes", "Dropdown", "Text", and "SmartTip".

A	B
1 地域	ガソリン1Lの値
2 札幌市	180
3 函館市	175
4 旭川市	180
5 青森市	172
6 八戸市	169
7 盛岡市	166
8 仙台市	176
9 秋田市	172
10 山形市	189
11 福島市	182
12 郡山市	177
13 水戸市	172
14 日立市	175
15 宇都宮市	176
16 小山市	172
17 前橋市	176
18 さいたま市	175
19 熊谷市	174
20 川口市	178
21 新潟市	177

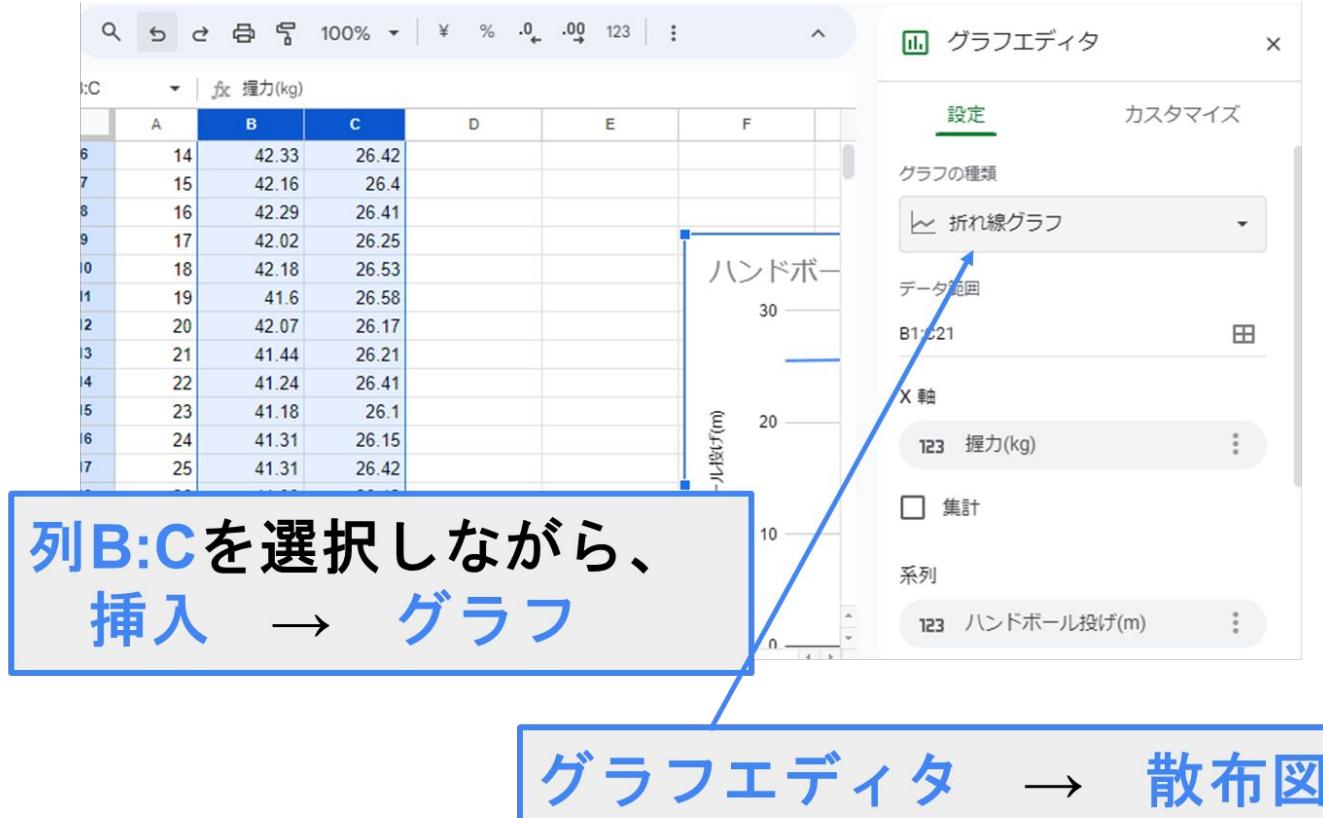
列Bを選択しながら、
挿入 → グラフ

【ワーク①】を参考にヒスト
グラムを作成する。

データサイエンス講座③

表計算アプリケーションを活用してデータを分析する

【ワーク①】散布図の作成

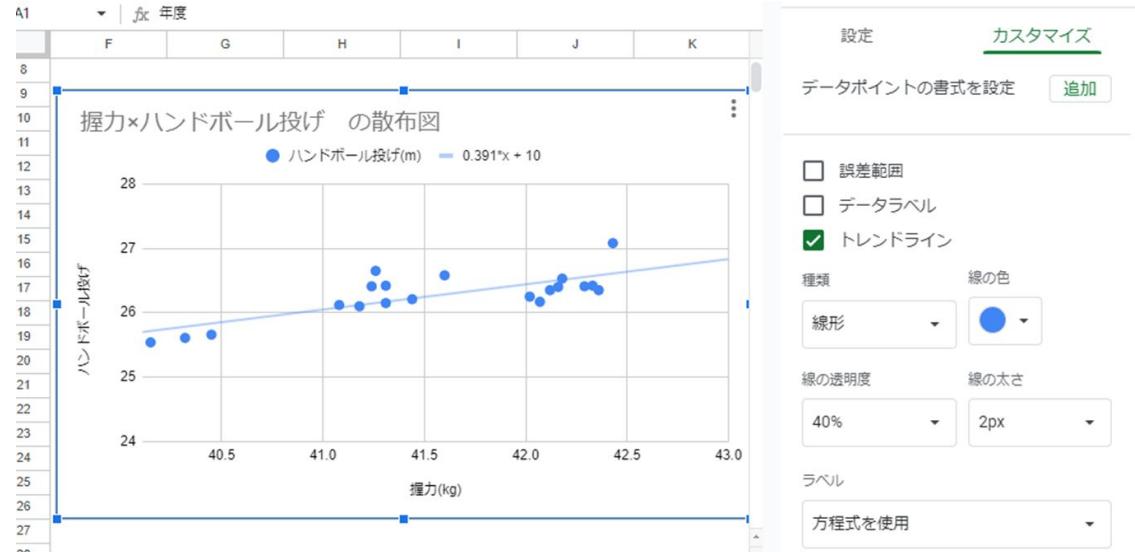


【ワーク①】散布図の作成(相関係数を求める)

	A	B	C
20	28	40.32	25.61
21	29	40.15	25.54
22			
23	相関係数(r)		
24	決定係数		
25	回帰直線方程式		
26			

B23セルに「=CORREL(B2:B21,C2:C21)」
と入力する。

【ワーク①】散布図の作成(トレンドライン描画)



回帰直線を引くには...
系列 → トレンドライン

①【ワーク①】散布図の作成(回帰直線の方程式と決定係数を求める)



ラベル → 方程式を使用
で表示される。
決定係数は するか, 相関
係数を2乗することで求めら
れる。

【ワーク①】散布図の作成(《おまけ》の解答例)

回帰直線による予想

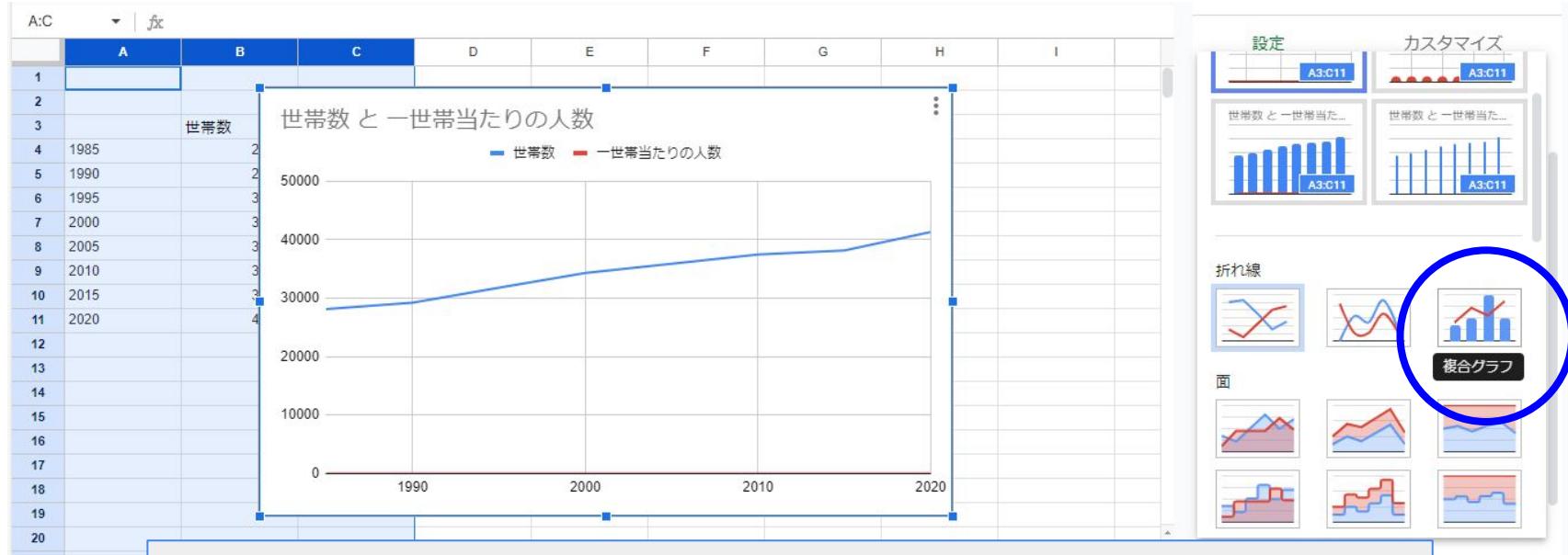
握力が

kgならば

ハンドボール投げは

=A29*0.319+10 mになるはず

【ワーク②】複合グラフの作成



A～C列を全て選択し、[挿入→グラフ](#)
グラフの種類→複合グラフを選択

【ワーク②】複合グラフの作成

The screenshot shows the Microsoft Excel Graph Editor interface. On the left, there's a sidebar with tabs for '設定' (Settings) and 'カスタマイズ' (Customize). Under 'カスタマイズ', a section titled '系列' (Series) is expanded, showing two options: 'すべての系列に適用' (Apply to all series) and two radio buttons: '世帯数' (Households) and '一世帯当たりの人数' (Number of people per household). Below this is a '種類' (Type) dropdown menu. On the right, there are two large text boxes with arrows pointing from specific UI elements.

系列 → 一世帯当たりの人数
に変更後

左軸 → 右軸に変更することで右
側に新しい縦軸が出現する。

【ワーク②】複合グラフの作成(訂正と応用)



調査年数までグラフに含まれてしまつた場合は、
①系列からA列を削除し、
②X軸にA3:A11を設定することで解決します。

【ワーク②】複合グラフの作成(訂正と応用)



「系列」からグラフの種類を変更したものを選択すると

階段面や、折れ線グラフなどに、グラフの種類を変更できる。

The screenshot shows the Microsoft Power BI Graph Editor interface. On the left, there's a sidebar with tabs for '設定' (Settings) and 'カスタマイズ' (Customize), with 'カスタマイズ' being active. Below that is a section titled '系列' (Series) with a dropdown menu open, showing three options: 'すべての系列に適用' (Apply to all series), '世帯数' (Households), and '一世帯当たりの人数' (Number of people per household). A callout arrow points from the text '「系列」からグラフの種類を変更したものを選択すると' to the 'すべての系列に適用' button. Another callout arrow points from the text '階段面や、折れ線グラフなどに、グラフの種類を変更できる。' to the '種類' (Type) section, which includes a '縦棒' (Bar) button and a '軸' (Axis) section with a '左軸' (Left axis) dropdown.

データポイントの書式を設定

追加