

# 演習問題の解答

## 第0章

### 問題 0.1

#### 問題 0.1.1

セル E1 に「20%」、セル F1 に「30%」、セル G1 に「50%」とそれぞれ入力する。

入力モードを「半角英数字」にし、セル E2 に「=」を入力する。番号 1 の小テスト 1 の点数が入力されているセル (B2) をクリックする。続けて、「\*」を入力し、「20%」が入力されているセル (E1) をクリックし、そのまま F4 キーを 2 回押す。すると、「=B2\*E\$1」というように、「1」の前に「\$」記号が付き、「1 (行)」が固定される。Enter キーを押し、それ (E2) を下にオートフィルし、そのまま (セル範囲 E2:E21 が選択されている状態で) セル範囲の右下あたりにマウスポインタを合わせると、マウスポインタが「+」の形になる。この状態のまま右にドラッグし G 列までオートフィルすれば、正しく求められる。

#### 問題 0.1.2

セル H1 に「合計」と入力する。

入力モードを「半角英数字」にし、セル H2 に「=su」などと入力すると、予測変換で関数の候補の一覧が出てくる。そこから「SUM」をダブルクリックし選択する。すると、「=SUM(」と入力されるので、合計をとるデータの範囲 (E2:G2) をドラッグして選択する。Enter キーを押すと合計 49.1 が計算される。セルには「=SUM(E2:G2)」と入力される。このセル (H2) が選択されている状態で、そのセルの右下あたりにマウスポインタを合わせると、マウスポインタが「+」の形になる。この状態のまま下に (セル H21 まで) ドラッグする (またはダブルクリックする) と、オートフィルすることができる。

	A	B	C	D	E	F	G	H	I
1	番号	小テスト 1	小テスト 2	定期試験	20%	30%	50%	合計	
2	1	84	41	40	16.8	12.3	20	49.1	
3	2	69	72	40	13.8	21.6	20	55.4	
4	3	66	60	44	13.2	18	22	53.2	
5	4	99	23	52	19.8	6.9	26	52.7	
6	5	63	58	52	12.6	17.4	26	56	
7	6	45	3	12	9	0.9	6	15.9	
8	7	76	25	72	15.2	7.5	36	58.7	
9	8	99	50	92	19.8	15	46	80.8	
10	9	81	58	48	16.2	17.4	24	57.6	
11	10	81	41	44	16.2	12.3	22	50.5	
12	11	45	26	52	9	7.8	26	42.8	
13	12	99	53	68	19.8	15.9	34	69.7	
14	13	96	9	48	19.2	2.7	24	45.9	
15	14	72	56	20	14.4	16.8	10	41.2	
16	15	75	41	68	15	12.3	34	61.3	
17	16	66	58	92	13.2	17.4	46	76.6	
18	17	69	9	56	13.8	2.7	28	44.5	
19	18	84	91	100	16.8	27.3	50	94.1	
20	19	87	29	24	17.4	8.7	12	38.1	
21	20	75	52	84	15	15.6	42	72.6	
22									
23									

### 問題 0.1.3

合計が入力されている H 列のどこかのセルを選択した状態で、ホームタブの（編集グループにある）「並べ替えとフィルター」をクリックし、「降順」を選択する。

	A	B	C	D	E	F	G	H	I
1	番号	小テスト 1	小テスト 2	定期試験	20%	30%	50%	合計	
2	18	84	91	100	16.8	27.3	50	94.1	
3	8	99	50	92	19.8	15	46	80.8	
4	16	66	58	92	13.2	17.4	46	76.6	
5	20	75	52	84	15	15.6	42	72.6	
6	12	99	53	68	19.8	15.9	34	69.7	
7	15	75	41	68	15	12.3	34	61.3	
8	7	76	25	72	15.2	7.5	36	58.7	
9	9	81	58	48	16.2	17.4	24	57.6	
10	5	63	58	52	12.6	17.4	26	56	
11	2	69	72	40	13.8	21.6	20	55.4	
12	3	66	60	44	13.2	18	22	53.2	
13	4	99	23	52	19.8	6.9	26	52.7	
14	10	81	41	44	16.2	12.3	22	50.5	
15	1	84	41	40	16.8	12.3	20	49.1	
16	13	96	9	48	19.2	2.7	24	45.9	
17	17	69	9	56	13.8	2.7	28	44.5	
18	11	45	26	52	9	7.8	26	42.8	
19	14	72	56	20	14.4	16.8	10	41.2	
20	19	87	29	24	17.4	8.7	12	38.1	
21	6	45	3	12	9	0.9	6	15.9	
22									
23									

### 問題 0.2

入力モードを「半角英数字」にし、セル B2 に「=」を入力する。すぐ左にある「1」が入力されているセル（A2）をクリックする。右にオートフィルする際にこれを固定したいので、「A（列）」を固定しないといけない。そのために F4 キーを 3 回押す。すると、「A」の前に「\$」記号が付き、「A（列）」が固定される。続けて、「\*」を入力し、すぐ上にある「1」が入力されているセル（B1）をクリックする。下にオートフィルする際にこれを固定したいので、「1（行）」を固定しないといけない。そのために F4 キーを 2 回押す。すると、「1」の前に「\$」記号が付き、「1（行）」が固定される。セル B2 には「=\$A2\*B\$1」が入力される。セル B2 を下にオートフィルし、そのまま（セル範囲 B2:B13 が選択されている状態で）セル範囲の右下あたりにマウスポインタを合わせると、マウスポインタが「+」の形になる。この状態のまま右にドラッグしオートフィルすれば、正しく求められる。

ために、たとえばセル C3 をクリックして、何が入力されているか確認すると、「=\$A3\*C\$1」となっていることがわかる（なお、スピルを使うなら、セル B2 に「=A2:A13\*B1:M1」と入力すればいい）。

## 第 1 章

### 問題 1.1

$$\text{個数} = \frac{\text{合計}}{\text{平均値}}$$

であるので，求めるクラスの人数は，

$$\frac{3933}{69} = 57$$

と計算され，57（人）となる．

この計算をエクセルで行う．まず，平均点の 69 と合計点の 3933 を入力する．

	A	B	C
1	平均点	69	
2	合計点	3933	
3			
4	人数		
5			
6			

次に，人数を求めるセル（B4）に「=」を入力し，合計点（3933）が入力されているセル（B2）をクリックする．

	A	B	C
1	平均点	69	
2	合計点	3933	
3			
4	人数	=B2	
5			
6			

続けて，「/」を入力し，平均点（69）が入力されているセル（B1）をクリックする．

	A	B	C
1	平均点	69	
2	合計点	3933	
3			
4	人数	=B2/B1	
5			
6			

Enter キーを押す．

	A	B	C
1	平均点	69	
2	合計点	3933	
3			
4	人数	57	
5			
6			

これで，人数 57（人）が求められた．

## 問題 1.2

### 問題 1.2.1

データを入力し、商品 A の売上個数の平均値を AVERAGE 関数で求める。

	A	B	C	D	E	F
1		商品A	商品B	商品C	商品D	
2	名古屋	1522	1443	1209	1570	
3	東京	1132	1321	1205	1061	
4	大阪	990	879	1561	1290	
5	商品ごとの平均値	=AVERAGE(B2:B4				
6		AVERAGE(数値1, [数値2], ...)				
7						
8						

オートフィルを使って、商品 B、商品 C、商品 D のそれぞれの売上個数の平均値も求める。

	A	B	C	D	E	F
1		商品A	商品B	商品C	商品D	
2	名古屋	1522	1443	1209	1570	
3	東京	1132	1321	1205	1061	
4	大阪	990	879	1561	1290	
5	商品ごとの平均値	1214.667	1214.333	1325	1307	
6						
7						

ホームタブの（数値グループにある）[小数点以下の表示桁数を減らす] ボタンを、値が整数表示になるまでクリックする。

	A	B	C	D	E	F
1		商品A	商品B	商品C	商品D	
2	名古屋	1522	1443	1209	1570	
3	東京	1132	1321	1205	1061	
4	大阪	990	879	1561	1290	
5	商品ごとの平均値	1215	1214	1325	1307	
6						
7						

売上個数の平均値が一番小さいのは商品 B であり、その平均値は約 1214（個）であることがわかる。

### 問題 1.2.2

都市ごとのそれぞれの売上個数の平均値を求め、整数の値で表示すると、東京の売上個数の平均値と大阪の売上個数の平均値が同じになってしまう。

	A	B	C	D	E	F	G
1		商品A	商品B	商品C	商品D	都市ごとの平均値	
2	名古屋	1522	1443	1209	1570	1436	
3	東京	1132	1321	1205	1061	1180	
4	大阪	990	879	1561	1290	1180	
5	商品ごとの平均値	1215	1214	1325	1307		
6							
7							

そこで、ホームタブの [小数点以下の表示桁数を増やす] ボタンを 1 回クリックすると、東京の売上個数の平均値のほうが小さいことがわかる。

	A	B	C	D	E	F	G
1		商品A	商品B	商品C	商品D	都市ごとの平均値	
2	名古屋	1522	1443	1209	1570	1436.0	
3	東京	1132	1321	1205	1061	1179.8	
4	大阪	990	879	1561	1290	1180.0	
5	商品ごとの平均値	1215	1214	1325	1307		
6							
7							

よって、売上個数の平均値が一番小さい都市は東京であり、その平均値は約 1180（個）ということになる。

### 問題 1.3

国語、英語、社会の平均点を AVERAGE 関数で求める。データを指定する際は、セル B2 をクリックして選択したあと、Ctrl キーを押しながら、セル D2 とセル F2 をそれぞれクリックして選択すればいい（または、セル B2 をクリックして選択したあと「,」を入力し、セル D2 をクリックして選択し、さらに「,」を入力し、セル F2 をクリックして選択してもいい）。A については「=AVERAGE(B2,D2,F2)」と入力される。数学、理科の平均点についても同様にすると、A については「=AVERAGE(C2,E2)」と入力される。これらの結果を下にオートフィルして全員分について求める。

	A	B	C	D	E	F	G	H	I
1		国語	数学	英語	理科	社会	国語、英語、社会の平均点	数学、理科の平均点	
2	A	56	87	94	57	81	77	72	
3	B	88	70	80	53	91	86.3	61.5	
4	C	79	70	74	70	70	74.3	70	
5	D	45	78	97	70	54	65.3	74	
6	E	89	78	60	80	79	76	79	
7	F	82	41	50	52	85	72.3	46.5	
8									
9									

これより、国語、英語、社会の平均点が最も高いのは B で、その平均点は約 86.3（点）であり、数学、理科の平均点が最も高いのは E で、その平均点は 79（点）であることがわかる。

### 問題 1.4

4 年目の売り上げは初年度の  $4 \times 1 \times 2$  倍に伸びている。つまり、初年度から 4 年目で 8 倍伸びていることになる。ということは、売り上げの伸びの倍率をならす（平均する）と、

初年度から 2 年目：2 倍

2 年目から 3 年目：さらに 2 倍

3 年目から 4 年目：さらに 2 倍

として、その結果  $2 \times 2 \times 2$  より 8 倍となったと考えるのが自然である。この 8 倍というのは、**4 と 1 と 2 の幾何平均**

$$\sqrt[3]{4 \times 1 \times 2} = \sqrt[3]{8} = 2$$

で求められている。よって、求める 1 年間の売り上げの伸びの平均倍率は 2（倍）となる。

これは、「ある数」を 3 回かけて、（4 と 1 と 2 の積）8 になるような「ある数」は何か、を求めているのである。

エクセルでは、セル B1, B2, B3 に 4, 1, 2 をそれぞれ入力した場合、他のセルに「=(B1\*B2\*B3)^(1/3)」と入力すると求めることができる。

### 問題 1.5

K くんの家から学校までの距離を  $x$  m とすると、今週において学校に行くのに移動した距離は

$$x \times 5 = 5x \text{ (m)}$$

であり、かかった時間の合計は

$$\left( \frac{x}{121} + \frac{x}{102} + \frac{x}{132} + \frac{x}{129} + \frac{x}{126} \right) \text{ 分} \quad \left( \text{時間} = \frac{\text{距離}}{\text{速さ}} \right)$$

である。よって、平均の速さは

$$\frac{\frac{x}{121} + \frac{x}{102} + \frac{x}{132} + \frac{x}{129} + \frac{x}{126}}{5} = \frac{5x}{\frac{1}{121} + \frac{1}{102} + \frac{1}{132} + \frac{1}{129} + \frac{1}{126}} \quad \left( \text{速さ} = \frac{\text{距離}}{\text{時間}} \right)$$

となる。つまり、**121, 102, 132, 129, 126 の調和平均**を使って求められるということである。

これは、121, 102, 132, 129, 126 のそれぞれの逆数の相加平均値

$$\frac{\frac{1}{121} + \frac{1}{102} + \frac{1}{132} + \frac{1}{129} + \frac{1}{126}}{5}$$

の逆数

$$\frac{5}{\frac{1}{121} + \frac{1}{102} + \frac{1}{132} + \frac{1}{129} + \frac{1}{126}}$$

である。

エクセルでは、セル B1, B2, B3, B4, B5 に 121, 102, 132, 129, 126 をそれぞれ入力した場合、他のセルに「=5/(1/B1+1/B2+1/B3+1/B4+1/B5)」と入力すると求めることができる。

今週での平均の速さは約 121.0 (m/分) であることがわかる。

## 第 2 章

---

### 問題 2.1

AVERAGE 関数と MEDIAN 関数をそれぞれ使うと、平均値が 55.25 (点)、中央値が 61 (点) であることがわかる。

### 問題 2.2

#### 問題 2.2.1

AVERAGE 関数を使う。5 科目の平均点が最も高いのは E で、その平均点は 77.2 (点) である。

#### 問題 2.2.2

AVERAGE 関数と MEDIAN 関数をそれぞれ使う。5 科目の中で、最も平均点が高い科目は社会で、その平均点は約 76.7 (点) であり、最も中央値が高い科目は国語で、その中央値は 80.5 (点) である。

	国語	数学	英語	理科	社会	平均値
A	56	87	94	57	81	75
B	88	70	80	53	91	76.4
C	79	70	74	70	70	72.6
D	45	78	97	70	54	68.8
E	89	78	60	80	79	77.2
F	82	41	50	52	85	62
平均値	73.2	70.7	75.8	63.7	76.7	
中央値	80.5	74	77	63.5	80	

### 問題 2.3

MODE.MULT 関数（または MODE.SNGL 関数）と AVERAGE 関数をそれぞれ使う。最頻値は 2（台）であり、平均値は約 1.62（台）である（関数を使うために、データ入力するのは数値だけにし「台」は入力してはいけない）。

### 問題 2.4

出身地の平均値、中央値、最頻値のうち、求められるものは最頻値である。

イタリア出身が 4 名、フランス出身が 2 名、香港出身が 3 名、イギリス出身が 1 名なので、一番多い出身地はイタリアである。よって、出身地の最頻値はイタリアである。

#### 注意

このように、最頻値は数値でないデータでも求められることがある。ただし、MODE.MULT 関数（または MODE.SNGL 関数）のみではエラーになってしまい、求められない。

### 問題 2.5

MODE.MULT 関数（または MODE.SNGL 関数）を使う。ペットの頭数の最頻値は 0（頭）である（データ入力の際、数値のみを入力し、文字列「頭」は入力してはいけない）。

### 問題 2.6

AVERAGE 関数、MEDIAN 関数、そして、MODE.MULT 関数（または MODE.SNGL 関数）をそれぞれ使う。平均値は約 19660（円）、中央値は 5000（円）、最頻値は 5000（円）である。

### 問題 2.7

$$\text{個数} = \frac{\text{合計}}{\text{平均値}}$$

であるので、求めるデータの個数は、

$$\frac{450}{56.25} = 8$$

と計算され、8（個）となる。

この計算をエクセルで行う。まず、平均値の 56.25 と合計の 450 をそれぞれ入力する。他のセルに「=」を入力し、合計（450）が入力されているセルをクリックする。続けて、「/」を入力し、平均値（56.25）が入力されているセルをクリックする。Enter キーを押すと、データの個数 8（個）が求められる。

## 問題 2.8

$$\text{合計} = \text{平均値} \times \text{個数}$$

であるので、求めるデータの合計は、

$$4.1 \times 20 = 82$$

と計算され、82 となる。

この計算をエクセルで行う。まず、平均値の 4.1 とデータの個数の 20 を入力する。合計を求めるセルに「=」を入力し平均 (4.1) が入力されているセルをクリックする。続けて、「\*」を入力し、データの合計 (82) が入力されているセルをクリックする。Enter キーを押すと、データの合計 82 が求められる。

## 問題 2.9

- (1) 平均値が中央値より大きい例 : 1, 1, 1, 3, 4      (平均値が 2, 中央値が 1)
- (2) 平均値が中央値より小さい例 : 0, 0, 3, 3, 4      (平均値が 2, 中央値が 3)
- (3) 平均値が最頻値より大きい例 : 1, 1, 1, 3, 4      (平均値が 2, 最頻値が 1)
- (4) 平均値が最頻値より小さい例 : 0, 1, 3, 3, 3      (平均値が 2, 最頻値が 3)
- (5) 中央値が最頻値より大きい例 : 1, 1, 2, 3, 4      (中央値が 2, 最頻値が 1)
- (6) 中央値が最頻値より小さい例 : 0, 1, 2, 3, 3      (中央値が 2, 最頻値が 3)

## 第 3 章

---

### 問題 3.1

TRIMMEAN 関数を使う。12 個のデータ中の 2 個を取り除いた平均値なので、TRIMMEAN 関数の「割合」には「2/12」と入力する。トリム平均は 55.8 (台) となる。

### 問題 3.2

TRIMMEAN 関数を使う。15 個のデータ中の 2 個を取り除いた平均値なので、TRIMMEAN 関数の「割合」には「2/15」と入力する。トリム平均は、小数第 3 位を四捨五入すると、約 64.62 (点) となる。

### 問題 3.3

最低点 (最小値) は 27 で、レンジは 75 である。

$$\text{レンジ} = \text{最大値} - \text{最小値}$$

なので、求める最高点 (最大値) は、

$$\text{最小値} + \text{レンジ} = 21 + 75 = 96$$

と計算され、96 となる。



### 問題 3.4

データ入力し，MAX 関数で最大値，MIN 関数で最小値を求める．

次に，レンジを求めるセルに「=」を入力し，最大値が計算されているセルをクリックする．続けて，「-」を入力し，最小値が計算されているセルをクリックする．国語のレンジは 44，数学のレンジは 46，英語のレンジは 47 と計算される．よって，レンジが最も大きい科目は英語であり，そのレンジは 47 であることがわかる．

	国語	数学	英語
A	76	32	94
B	88	70	80
C	78	70	74
D	45	78	47
E	89	78	60
F	80	41	50
最大値	89	78	94
最小値	45	32	47
レンジ	44	46	47

### 問題 3.5

	国語	数学	英語
A	75	=H9	65
B	=G8	70	80
C	66	70	74
D	70	77	47
E	89	78	96
F	47	56	50
最大値	=G9+G10	=MAX(H3:H7)	=MAX(I2:I7)
最小値	=MIN(G2,G4:G7)	=H8-H10	=MIN(I2:I7)
レンジ	43	43	=I8-I9

国語の最小値 b が 47，レンジが 43 なので，最大値 a は  $47 + 43 = 90$  である．

数学の最大値 d が 78，レンジは 43 なので，最小値 c は  $78 - 43 = 35$  である．

英語の最大値 e は 96，最小値 f は 47 なので，レンジ g は  $96 - 47 = 49$  である．

	国語	数学	英語
A	75	35	65
B	90	70	80
C	66	70	74
D	70	77	47
E	89	78	96
F	47	56	50
最大値	90	78	96
最小値	47	35	47
レンジ	43	43	49

## 第4章

### 問題 4.1

VAR.P 関数で分散を求めると、約 186.7 (台<sup>2</sup>) であることがわかる。

### 問題 4.2

STDEV.P 関数で標準偏差を求める。ホームタブの「小数点以下の表示桁数を減らす」ボタンを、値が小数第 1 位までの表示になるまでクリックする。

	国語	数学	英語	理科	社会
標準偏差	15.3	11.1	17.4	11.1	17.3

標準偏差が一番大きい科目は英語であり、その標準偏差は約 17.4 (点) であることがわかる。

数学の標準偏差と理科の標準偏差の表示が同じになってしまったので、ホームタブの「小数点以下の表示桁数を増やす」ボタンを 1 回クリックすると、理科の標準偏差のほうが小さいことがわかる。

	国語	数学	英語	理科	社会
標準偏差	15.35	11.07	17.38	11.06	17.34

よって、標準偏差が一番小さい科目は理科であり、その標準偏差は約 11.1 (点) である。

### 問題 4.3

分散が 4 のとき標準偏差は  $\sqrt{4}$ 、つまり、2 である。また、標準偏差が 9 のとき分散は  $9^2$ 、つまり、81 である。

### 問題 4.4

データ入力し、データタブの (分析グループにある) 「データ分析」を選択して、分析ツールの「基本統計量」を選んで求める。「統計情報」にチェックを入れ忘れないようにする。

列1	
平均	448.29
標準誤差	44.665
中央値 (メジアン)	431
最頻値 (モード)	#N/A
標準偏差	118.17
分散	13965
尖度	-1.7311
歪度	0.1359
範囲	302
最小	309
最大	611
合計	3138
データの個数	7

Excel アドインのデータ分析ツールの「基本統計量」で求められる「標準偏差」は約 118.2 (個) であることがわかる。

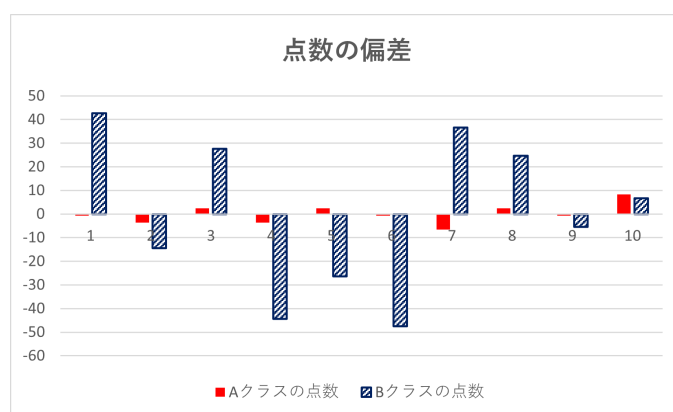
## 問題 4.5

平均点は AVERAGE 関数、中央値は MEDIAN 関数、標準偏差は STDEV.P 関数で求める。レンジは、MAX 関数で求めた最大値から MIN 関数で求めた最小値をひいて求める。

A クラスの点数、B クラスの点数それぞれについて、偏差を計算すると下記ようになる（（下記のように入力し、A クラスの点数の平均値をセル M2 に計算した場合）A クラスの最初の点数の偏差を「=B2-\$M2」と入力し、下にオートフィルして、さらに右にもオートフィルすればいい）。

B6    ↓    :    ×    ✓    fx    =B2-\$M2																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1													平均点	中央値	標準偏差	最高点	最低点	レンジ	
2	Aクラスの点数	51	48	54	48	54	51	45	54	51	60		51.6	51	3.98	60	45	15	
3	Bクラスの点数	93	36	78	6	24	3	87	75	45	57		50.4	51	31.17	93	3	90	
4																			
5																			
6																			
6	Aクラスの点数	-0.6	-3.6	2.4	-3.6	2.4	-0.6	-6.6	2.4	-0.6	8.4								
7	Bクラスの点数	42.6	-14.4	27.6	-44.4	-26.4	-47.4	36.6	24.6	-5.4	6.6								
8																			
9																			

この偏差の表の 2 行分を選択して、挿入タブの（グラフグループにある）「縦棒/横棒グラフの挿入」の「2-D 縦棒」の「集合縦棒」を選ぶ。



## 問題 4.6

平均点は AVERAGE 関数、トリム平均は TRIMMEAN 関数、最頻値は MODE.MULT 関数（または MODE.SNGL 関数）、分散は VAR.P 関数、標準偏差は STDEV.P 関数で求める。ここで、トリム平均は 20 個のデータ中の 2 個を取り除いた平均値なので、TRIMMEAN 関数の「割合」には「2/20」と入力する。

平均値	トリム平均	最頻値	分散	標準偏差
175.0	175.6	176	35.8	6.0

## 第5章

### 問題 5.1

クラスごとに AVERAGE 関数で平均値, STDEV.P 関数で標準偏差をそれぞれ求める (A クラスの平均値をセル B13 に, 標準偏差をセル B14 にそれぞれ求めて, 右にオートフィルする)。

次に, (セル G2 に)「=STANDARDIZE(」と入力し, A クラスの 1 番の点数 (B2) をクリックして指定する。続けて, 「,」を入力し, A クラスの平均点が計算されているセル (B13) をクリックし F4 キーを 2 回押す。さらに, 「,」を入力し, A クラスの標準偏差が計算されているセル (B14) をクリックし F4 キーを 2 回押す。Enter キーを押すと A クラスの 1 番のもともとの点数を標準化した値が計算される (セル G2 には「=STANDARDIZE(B2,B\$13,B\$14)」と入力される。これは, 下にオートフィルするときに「平均」と「標準偏差」の行番号を固定するために, その前に「\$」を付けているのである。一方, あとで右にオートフィルするために, 列番号 B の前には「\$」を付けてはいけない)。

下にオートフィルして, さらに右にもオートフィルしてデータ全部を標準化する (標準化したデータたちの最大値と最小値をそれぞれ MAX 関数と MIN 関数を用いて求めてもいい)。

G2											
	A	B	C	D	E	F	G	H	I	J	K
1	番号	Aクラス	Bクラス	Cクラス	Dクラス		Aクラス	Bクラス	Cクラス	Dクラス	
2	1	95	100	57	45		0.807	1.084	-0.247	-0.125	
3	2	98	80	70	70	⇒	1.082	-0.597	0.646	1.524	
4	3	70	98	45	55	標準化	-1.485	0.916	-1.072	0.534	
5	4	80	63	60	32		-0.568	-2.026	-0.041	-0.983	
6	5	97	88	87	43		0.990	0.076	1.814	-0.257	
7	6	90	90	60	46		0.348	0.244	-0.041	-0.059	
8	7	78	98	35	70		-0.752	0.916	-1.759	1.524	
9	8	88	100	57	55		0.165	1.084	-0.247	0.534	
10	9	98	77	80	23		1.082	-0.849	1.333	-1.577	
11	10	68	77	55	30		-1.669	-0.849	-0.385	-1.115	
12											
13	平均値	86.2	87.1	60.6	46.9			最大値	1.814		
14	標準偏差	10.907	11.895	14.555	15.156			最小値	-2.026		
15											
16											

これより, 標準化したときに最も大きくなる値は約 1.814 で, C クラスの 5 番の点数である。また, 標準化したときに最も小さくなる値は約 -2.026 で, B クラスの 4 番の点数である (なお, 標準化する際にスピルを使うなら, セル G2 に「=STANDARDIZE(B2:E11,B13:E13,B14:E14)」と入力すればいい)。

### 問題 5.2

K さんの点数のみ科目ごとに標準化すれば十分である。

科目ごとに AVERAGE 関数で平均値, STDEV.P 関数で標準偏差を求める (国語の平均値をセル B14 に, 標準偏差をセル B15 にそれぞれ求めて, 右にオートフィルする)。

そして, (セル H12 に)「=STANDARDIZE(B12,B14,B15)」と入力して, 右にオートフィルすればいい。この場合, 「\$」は不要。

H12														: × ✓ fx		=STANDARDIZE(B12,B14,B15)	
	A	B	C	D	E	F	G	H	I	J	K	L	M				
1		国語	数学	英語	理科	社会											
2	A	54	89	79	55	50											
3	B	76	78	54	30	80											
4	C	90	65	67	75	55											
5	D	76	98	80	80	63											
6	E	78	76	50	50	80											
7	F	58	76	78	45	87											
8	G	89	43	65	33	57											
9	H	55	56	72	70	54											
10	I	65	80	65	80	39											
11	J	77	78	49	65	77		国語	数学	英語	理科	社会					
12	K	100	100	91	98	98	⇒	1.778	1.471	1.779	1.778	1.774					
13							標準化										
14	平均値	74.36	76.27	68.18	61.91	67.27											
15	標準偏差	14.42	16.13	12.83	20.3	17.32											
16																	
17																	

これより、Kさんの点数で標準化した値が最も大きくなる科目は英語で、その標準化した値は約 1.779 である。また、Kさんの点数で標準化した値が 2 番目に大きくなる科目は、小数第 3 位までの表示にすると、国語と理科が同じ値で表示されるのでどちらかわからない。そこで、小数第 4 位まで表示させると、

国語	数学	英語	理科	社会
1.7781	1.4709	1.7790	1.7779	1.7740

となり、国語の値のほうが大きいことがわかる。よって、Kさんの点数で標準化した値が 2 番目に大きくなる科目は国語で、その標準化した値は約 1.778 である。

## 補足

全員の点数を科目ごとに標準化すると、下記のような結果になる

(セル H2 に「=STANDARDIZE(B2,B\$14,B\$15)」と入力して、下にオートフィルして、さらに右にもオートフィルし、データ全部を標準化している)(なお、標準化する際にスピルを使うなら、「=STANDARDIZE(B2:F12,B14:F14,B15:F15)」のように入力すればいい)。

H2													
	A	B	C	D	E	F	G	H	I	J	K	L	M
1		国語	数学	英語	理科	社会							
2	A	54	89	79	55	50		-1.412	0.789	0.843	-0.340	-0.997	
3	B	76	78	54	30	80	⇒	0.113	0.107	-1.106	-1.572	0.735	
4	C	90	65	67	75	55	標準化	1.085	-0.699	-0.092	0.645	-0.709	
5	D	76	98	80	80	63		0.113	1.347	0.921	0.891	-0.247	
6	E	78	76	50	50	80		0.252	-0.017	-1.418	-0.587	0.735	
7	F	58	76	78	45	87		-1.135	-0.017	0.765	-0.833	1.139	
8	G	89	43	65	33	57		1.015	-2.063	-0.248	-1.424	-0.593	
9	H	55	56	72	70	54		-1.343	-1.257	0.298	0.399	-0.766	
10	I	65	80	65	80	39		-0.649	0.231	-0.248	0.891	-1.632	
11	J	77	78	49	65	77		0.183	0.107	-1.496	0.152	0.562	
12	K	100	100	91	98	98		1.778	1.471	1.779	1.778	1.774	
13													
14	平均値	74.36	76.27	68.18	61.91	67.27							
15	標準偏差	14.42	16.13	12.83	20.3	17.32							
16													
17													

まず、女性の体重と体脂肪率についてそれぞれ、AVERAGE 関数で平均値、STDEV.P 関数で標準偏差を求める（女性の体重の平均値をセル B15 に、標準偏差をセル B16 にそれぞれ求めて、右にオートフィルする）。

そして、(セル H13 に)「=STANDARDIZE(B13,B15,B16)」と入力して、番号 11 の女性の標準化した体重の値を求める。これを右にオートフィルして、番号 11 の女性の標準化した体脂肪率の値も求める。この場合、「\$」は不要。

これより、番号 11 の女性の身長を標準化した値は約  $-1.05$  であり、体重を標準化した値は約  $-1.30$  であることがわかる（なお、標準化する際にスピルを使うなら「=STANDARDIZE(B13:C13,B15:C15,B16:C16)」と入力すればいい）。

作成した問題 5. 3 のファイルを開く.

セル範囲 B15:B16（または C15:C16）を右にオートフィルし、男性の体重の平均値と標準偏差、および、男性の体脂肪率の平均値と標準偏差をそれぞれ求める（または、サポートページからダウンロードできる元データ「第5章 ファイル6」を開き、女性の体重の平均値をセル B15 に、標準偏差をセル B16 にそれぞれ求めて、右にオートフィルする）。

そして、(セル H3 に)「=STANDARDIZE(B3,B\$15,B\$16)」と入力して、番号 1 の女性の標準化した体重の値を求める(「平均」と「標準偏差」の行番号を固定するため、F4 キーを 2 回押すことにより行番号の前に「\$」を付ける)。下にオートフィルして、さらに右にもオートフィルしてデータ全部を標準化する。

[illegible]

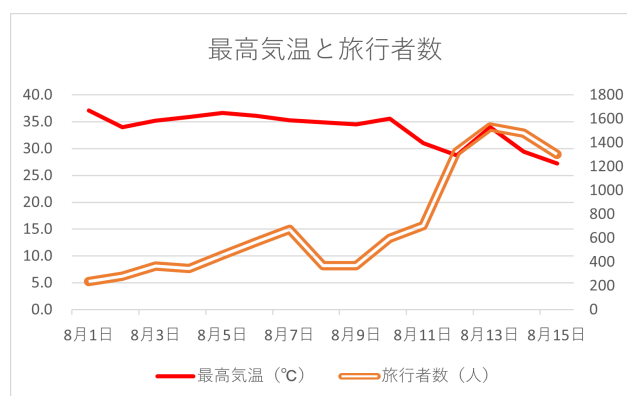
これより、(男女別に)標準化した体脂肪率の値のほうが(男女別に)標準化した体重の値より小さい女性の番号は2, 3, 5, 7, 9, 11であり、男性の番号は2, 3, 9であることがわかる(なお、標準化する際にSPILSを使うなら「=STANDARDIZE(B3:E13,B15:E15,B16:E16)」と入力すればいい)。

## 第6章

### 問題 6.1

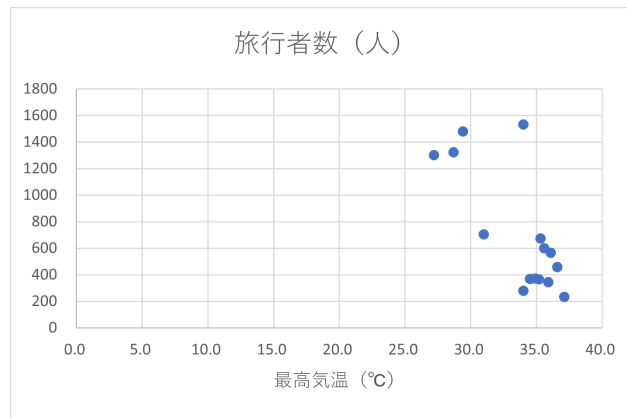
### 問題 6.1.1

「最高気温」と「旅行者数」の折れ線グラフ（日付を横軸にする）（第2軸を採用）:



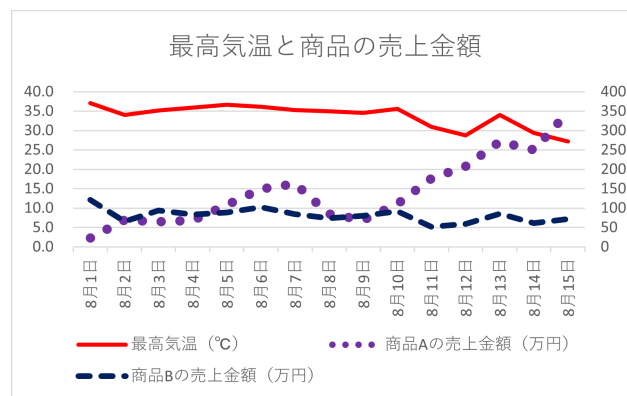
### 問題 6.1.2

「最高気温」（横軸）と「旅行者数」（縦軸）の散布図：



### 問題 6.1.3

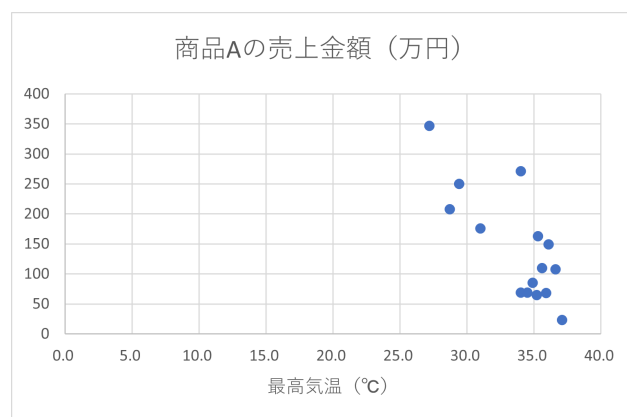
〔最高気温〕と〔商品 A の売上金額〕と〔商品 B の売上金額〕の折れ線グラフ（日付を横軸にする）（第 2 軸を採用）：



これより、〔最高気温〕、〔商品 A の売上金額〕、〔商品 B の売上金額〕のうち、増加傾向にあるように見えるのは〔商品 A の売上金額〕であることがわかる。

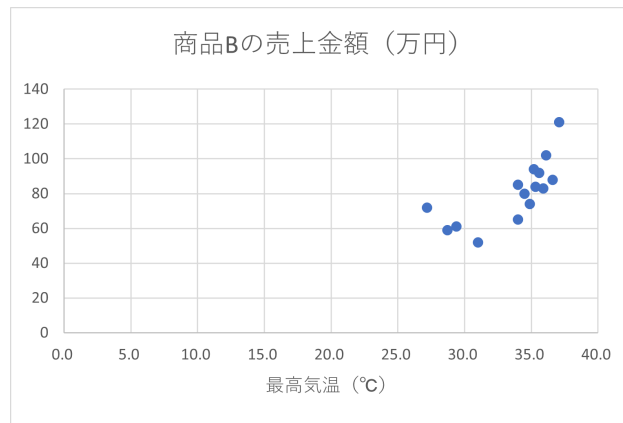
### 問題 6.1.4

〔最高気温〕（横軸）と〔商品 A の売上金額〕（縦軸）の散布図：



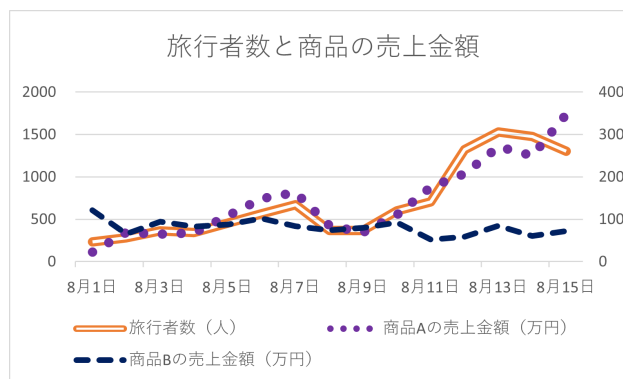
〔最高気温〕（横軸）と〔商品 B の売上金額〕（縦軸）の散布図：





### 問題 6.1.5

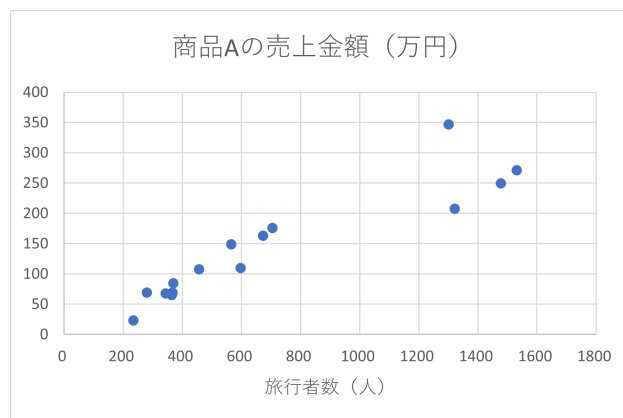
「旅行者数」と「商品 A の売上金額」と「商品 B の売上金額」の折れ線グラフ（日付を横軸にする）（第 2 軸を採用）：



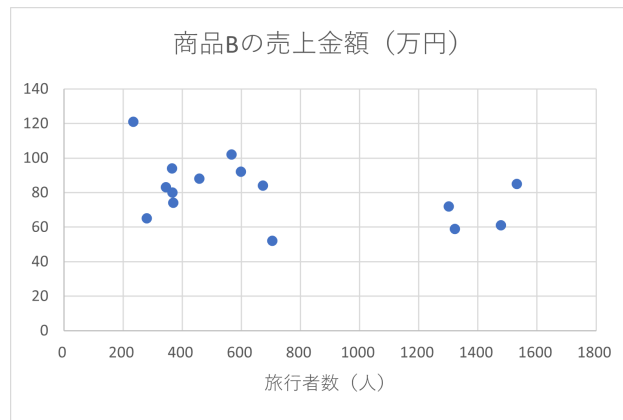
これより、「旅行者数」、「商品 A の売上金額」、「商品 B の売上金額」のうち、同じように増減しているように見え、深い関係がありそうに見えるのは、「旅行者数」と「商品 A の売上金額」の組み合わせであることがわかる。

### 問題 6.1.6

「旅行者数」（横軸）と「商品 A の売上金額」（縦軸）の散布図：



「旅行者数」（横軸）と「商品 B の売上金額」（縦軸）の散布図：



これより、直線的で右上がりの傾向があるものは「旅行者数」と「商品 A の売上金額」の散布図であることがわかる。

## 問題 6.2

### 問題 6.2.1

表中のどこかのセルを選択した状態で、挿入タブの（テーブルグループにある）[ピボットテーブル]（の「テーブルまたは範囲から」）を選択する。「テーブル/範囲」に表全体が選択されていることを確認する。OK ボタンを押すと「ピボットテーブルのフィールド」が出てくる。

「性別」、「車のボディタイプ」の順にチェックを入れる。すると、「行」ボックスに「性別」が上で「車のボディタイプ」が下に入る。

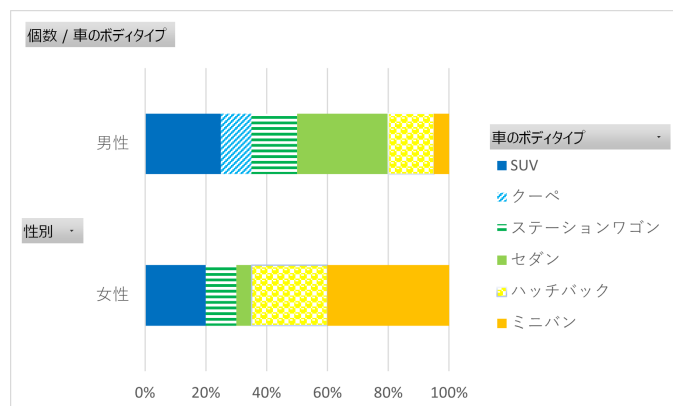
そして、上部にある、ボックスにチェックの入った「車のボディタイプ」（という文字列）をドラッグして「値」ボックスに移動させる。すると、「値」ボックスに「個数/車のボディタイプ」が入る（または、ボックスにチェックの入った「性別」（という文字列）をドラッグして「値」ボックスに移動させてもいい。この場合は、「値」ボックスに「個数/性別」が入る）。

行ラベル	個数 / 車のボディタイプ
<input checked="" type="checkbox"/> 女性	20
SUV	4
ステーションワゴン	2
セダン	1
ハッチバック	5
ミニバン	8
<input checked="" type="checkbox"/> 男性	20
SUV	5
クーペ	2
ステーションワゴン	3
セダン	6
ハッチバック	3
ミニバン	1
総計	40

次に、「行」ボックス内の「車のボディタイプ」を「列」ボックスに移動させる。

個数 / 車のボディタイプ	列ラベル						
行ラベル	SUV	クーペ	ステーションワゴン	セダン	ハッチバック	ミニバン	総計
女性	4			2	1	5	8
男性	5	2		3	6	3	1
総計	9	2		5	7	8	9

このピボットテーブルが選択されている状態において、(ピボットテーブル) 分析タブの (ツールグループにある) [ピボットグラフ] を選ぶ。[横棒] の [100%積み上げ横棒] を選択すると、下のようなグラフが得られる。



### 問題 6.2.2

表中のどこかのセルを選択した状態で、挿入タブの [ピボットテーブル] (の「テーブルまたは範囲から」) を選択する。「テーブル/範囲」に表全体が選択されていることを確認する。OK ボタンを押すと「ピボットテーブルのフィールド」が出てくる。

「車のボディタイプ」、「性別」の順にチェックを入れチェックを入れる。すると、「行」ボックスに「車のボディタイプ」が上で「性別」が下に入る。

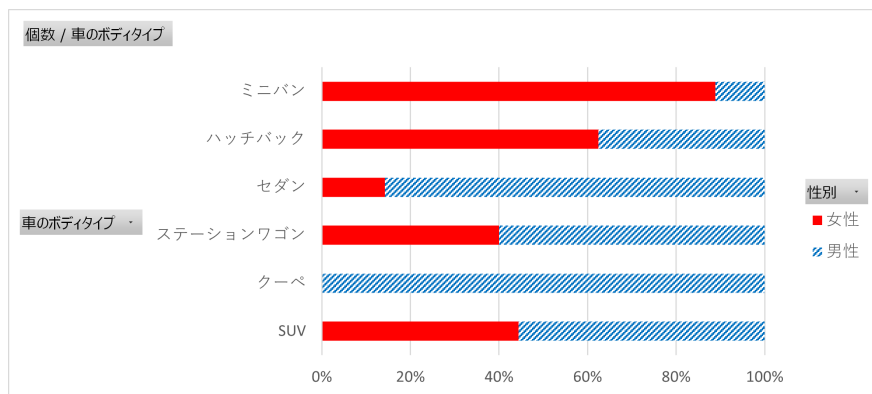
そして、上部にある、ボックスにチェックの入った「車のボディタイプ」(という文字列) をドラッグして「値」ボックスに移動させる。すると、「値」ボックスに「個数/車のボディタイプ」が入る (または、ボックスにチェックの入った「性別」(という文字列) をドラッグして「値」ボックスに移動させてもいい。この場合は、「値」ボックスに「個数/性別」が入る)。

行ラベル		個数 / 車のボディタイプ
SUV		9
女性		4
男性		5
クーペ		2
男性		2
ステーションワゴン		5
女性		2
男性		3
セダン		7
女性		1
男性		6
ハッチバック		8
女性		5
男性		3
ミニバン		9
女性		8
男性		1
総計		40

次に、「行」ボックス内の「性別」を「列」ボックスに移動させる。

個数 / 車のボディタイプ	列ラベル			
行ラベル	女性	男性	総計	
SUV	4	5	9	
クーペ		2	2	
ステーションワゴン	2	3	5	
セダン	1	6	7	
ハッチバック	5	3	8	
ミニバン	8	1	9	
総計	20	20	40	

このピボットテーブルが選択されている状態において、(ピボットテーブル) 分析タブの (ツールグループにある) [ピボットグラフ] を選ぶ。「横棒」の「100%積み上げ横棒」を選択すると、下のようなグラフが得られる。



## 第7章

### 問題 7.1

#### 問題 7.1.1

データタブの (分析グループにある) [データ分析] を選択して、分析ツールを使って相関係数を求める。

	最高気温 (°C)	旅行者数 (人)	商品Aの売上金額 (万円)	商品Bの売上金額 (万円)
最高気温 (°C)	1			
旅行者数 (人)	-0.751541901	1		
商品Aの売上金額 (万円)	-0.784218321	0.921868769	1	
商品Bの売上金額 (万円)	0.739698129	-0.443430024	-0.457807478	1

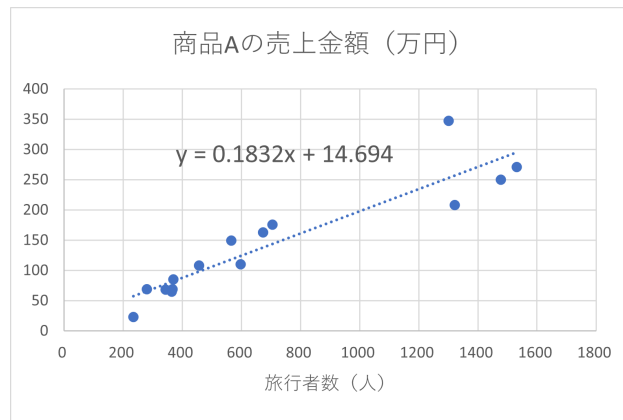
これより、[最高気温] と [旅行者数] の相関係数は約  $-0.752$ 、[最高気温] と [商品 A の売上金額] の相関係数は約  $-0.784$ 、[最高気温] と [商品 B の売上金額] の相関係数は約  $0.740$ 、[旅行者数] と [商品 A の売上金額] の相関係数は約  $0.922$ 、[旅行者数] と [商品 B の売上金額] の相関係数は約  $-0.443$  となる。それぞれの散布図は問題 6.1 の解答参照。

#### 問題 7.1.2

線形近似の近似曲線は下記のようになり、回帰式は

$$y = 0.1832x + 14.694$$

となることがわかる。



### 問題 7.1.3

問題 7.1.2 より，「旅行者数」( $x$ ) と「商品 A の売上金額」( $y$ ) についての回帰式は

$$y = 0.1832x + 14.694$$

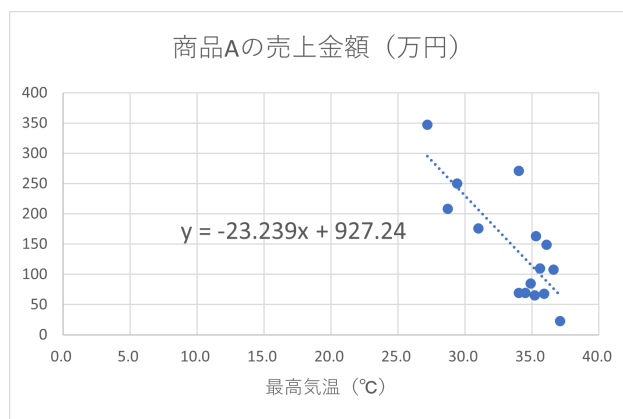
である．よって，セルに計算式「=0.1832\*1000+14.694」を入れると，「旅行者数が 1000 人のときの商品 A の売上金額の予測値」約 197.894（万円）が求められる．また，セルに計算式「=0.1832\*0+14.694」を入れると，「旅行者が誰もいないときの商品 A の売上金額の予測値」約 14.694（万円）が求められる．後者はたしかに回帰式の切片と一致する．

### 問題 7.1.4

線形近似の近似曲線は下記のようになり，回帰式は

$$y = -23.239x + 927.24$$

となることがわかる．



### 問題 7.1.5

問題 7.1.4 より，「最高気温」( $x$ ) と「商品 A の売上金額」( $y$ ) についての回帰式は

$$y = -23.239x + 927.24$$

である．よって，セルに計算式「=-23.239\*30+927.24」を入れると，「最高気温が 30.0℃のときの商品 A の売上金額の予測値」約 230.07（万円）が求められる．

また、セルに計算式「=-23.239\*31+927.24」を入れると、[最高気温が 31.0℃のときの商品 A の売上金額の予測値] 約 206.831（万円）が求められる。

#### 問題 7.1.6

[8 月 11 日における商品 A の売上金額の実測値] は 176（万円）である。また、問題 7.1.5 で求めた [最高気温が 31.0℃のときの商品 A の売上金額の予測値] は約 206.831（万円）である。よって、 $176 - 206.831$  をエクセルで計算すると、差（残差）約  $-30.831$ （万円）が求められる。

## 第 8 章

### 問題 8.1

#### 問題 8.1.1

分析ツールの「回帰分析」を選び、「入力 Y 範囲」は結果系のデータ（この場合は「商品 A の売上金額」、D1:D16）、「入力 X 範囲」は原因系のデータ（この場合は「最高気温」、B1:B16）、「ラベル」にチェックする。「残差」にもチェックを入れておく。

回帰式の傾きは約  $-23.24$ 、切片は約 927.24 であることがわかる。

	係数
切片	927.2378
最高気温 (°C)	-23.2395

つまり、[最高気温] ( $x$ ) と [商品 A の売上個数] ( $y$ ) についての回帰式は

$$y = -23.24x + 927.24 \quad (\text{傾きと切片は概数})$$

となる。一般に、直線の方程式「 $y = ax + b$  ( $a, b$  は定数)」において、 $x$  が 1 増えると  $y$  は傾き  $a$  の分だけ増える。つまり、気温が 1℃上がると、商品 A の売上金額は約  $-23.24$ （万円）だけ増えると予測できることになる。いいかえると、気温が 1℃上がると、商品 A の売上金額は約 23.24（万円）だけ減ると予測できる。

#### 問題 8.1.2

回帰分析の結果、残差が最も大きくなるのは観測値 13、つまり、8 月 13 日であり、その日の商品 A の売上金額の予測値は約 137.09（万円）であることがわかる。

#### 問題 8.1.3

問題 8.1.1 より、回帰式の傾きは約  $-23.24$ 、切片は約 927.24 である。

	係数
切片	927.2378
最高気温 (°C)	-23.2395

よって、[最高気温が 33.3℃のときの商品 A の売上金額] は、空いているセルに「=[傾きが入力されたセル番地] \*33.3+ [切片が入力されたセル番地]」と入力すると求められる。約 153.36（万円）と計算される。

#### 問題 8.1.4

回帰分析の結果、[最高気温]と[商品 A の売上金額]についての R-2 乗値は約 0.62 であり、[最高気温]と[商品 B の売上金額]についての R-2 乗値は約 0.55 である。

よって、[最高気温]と[商品 A の売上金額]についての回帰式による予測精度と、[最高気温]と[商品 B の売上金額]についての回帰式による予測精度のうち、大きいほうは[最高気温]と[商品 A の売上金額]についての回帰式による予測精度である。

### 問題 8.2

#### 問題 8.2.1

分析ツールの「回帰分析」を選び、「入力 Y 範囲」は結果系のデータ（この場合は「点数」、D1:C16）、「入力 X 範囲」は原因系のデータ（この場合は「勉強時間」、C1:C16）、「ラベル」にチェックする。「残差」にもチェックを入れておく。

観測値	予測値: 点数 (点)	残差
1	43.59086041	-11.5909
2	81.88083655	16.11916
3	57.33598005	1.66402
4	81.88083655	-47.8808
5	29.35484364	-11.3548
6	55.37239153	-2.37239
7	48.99072884	5.009271
8	72.55379108	21.44621
9	42.60906615	7.390934
10	76.97186525	13.02813
11	74.5173796	9.48262
12	49.9725231	3.027477
13	91.20788201	-3.20788
14	101.5167217	-11.5167
15	89.24429349	10.75571

これより、一番予測が当たっている、つまり、残差の絶対値が最も小さくなる番号は 3 であり、その残差は約 1.664 (点)、予測値は約 57.336 (点)、実測値は 59 (点) であることがわかる。

また、この実測値から予測値をひくと、約 1.664 (点) になり、その残差と一致することが確認できる。

#### 問題 8.2.2

回帰分析の結果、回帰式の傾きは約 0.491、切片は約 27.882 である。

	係数
切片	27.882152
勉強時間 (時間)	0.4908971

つまり、勉強時間 ( $x$ ) と点数 ( $y$ ) についての回帰式は

$$y = 0.491x + 27.882 \quad (\text{傾きと切片は概数})$$

となることがわかる。よって、[勉強時間が 120 時間のときの点数の予測値] は、空いているセルに「= [傾きが入力されたセル番地] \*120+ [切片が入力されたセル番地]」と入力すると求められる。約 86.790 (点) と計算される。

同様にして、[勉強時間が 130 時間のときの点数の予測値] は約 91.699 (点) と計算される。

### 問題 8.2.3

問題 8.2.2 と同様にし、[勉強時間が 121 時間のときの点数の予測値] から順番に計算すると、

勉強時間	点数の予測値
121	87.280705
122	87.7716021
123	88.2624992
124	88.7533964
125	89.2442935
126	89.7351906
127	90.2260878

となる。はじめて点数の予測値が 90 点以上になるのは勉強時間が（整数値では）127 時間のときである。よって、90 点以上の点数を取るには、127 時間勉強すれば十分であると予測できる（もちろん 128 時間（またはそれ以上）勉強すれば十分であるとも予測できるが、求める「最小の」整数値は 127（時間）ということになる）。

### 問題 8.2.4

勉強時間と点数についての回帰分析における R-2 乗値は約 0.618 であることがわかる。

回帰統計	
重相関 R	0.7862593
重決定 R2	<b>0.6182036</b>
補正 R2	0.5888347
標準誤差	17.300882
観測数	15

## 問題 8.3

### 問題 8.3.1

データタブの（分析グループにある）「データ分析」を選択して、分析ツールを使って相関係数を求める。

	価格（円）	売上個数（個）
価格（円）	1	
売上個数（個）	-0.8611401	1

これより、価格と売上個数の相関係数は約  $-0.86$  となることがわかる（CORREL 関数で求めてもいい）。

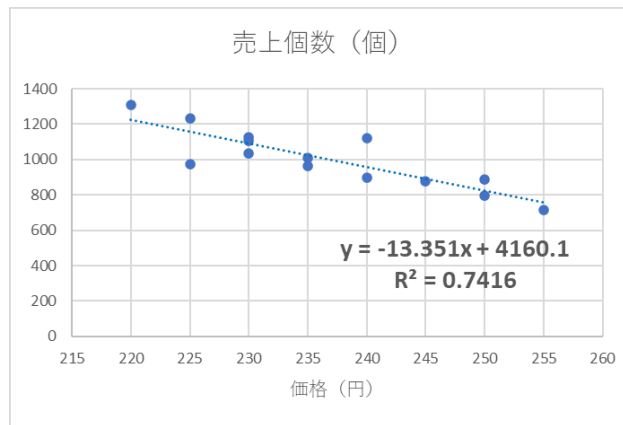
### 問題 8.3.2

散布図と線形近似の近似曲線は下記のようになり、価格（ $x$ ）と売上個数（ $y$ ）についての回帰式は

$$y = -13.351x + 4160.1$$

となることがわかる。





### 問題 8.3.3

問題 8.3.2 より，価格（ $x$ ）と売上個数（ $y$ ）についての回帰式は

$$y = -13.351x + 4160.1$$

である．よって，セルに計算式「=-13.351\*260+4160.1」を入れると，「価格を 260 円としたときの売上個数の予測値」が計算され，約 689（個）であることがわかる．

### 問題 8.3.4

問題 8.3.2 より，価格（ $x$ ）と売上個数（ $y$ ）についての回帰式は

$$y = -13.351x + 4160.1$$

である．一般に，直線の方程式「 $y = ax + b$ （ $a, b$  は定数）」において， $x$  が 1 増えると  $y$  は傾き  $a$  の分だけ増える．つまり，価格を 1 円上げると売上個数は約  $-13.35$ （個）だけ増えると予測できることになる．いいかえると，価格を 1 円上げると売上個数は約  $13.35$ （個）だけ減ると予測できる．

### 問題 8.3.5

問題 8.3.4 の結果より，価格を 1 円下げると売上個数は約  $13.35$ （個）だけ増えると予測できる．

## 第 9 章

### 問題 9.1

セル A1 に「価格」，A2 に「仕入れ値」，A3 に「売上個数」，A4 に「粗利」，A5 に「利益」と入力する．セル B1（価格）に「100」，セル B2（仕入れ値）に「100」を入力する．セル B3（売上個数）には

$$\text{「} -1.69 \times [\text{価格のセル}] + 369.08 \text{」, つまり, 「} =-1.69*B1+369.08 \text{」}$$

を入力する．セル B4（粗利）は

$$\text{「} [\text{価格のセル}] - [\text{仕入れ値のセル}] \text{」, つまり, 「} =B1-B2 \text{」}$$

を入力する．セル B5（利益）は

$$\text{「} [\text{売上個数のセル}] \times [\text{粗利のセル}] \text{」, つまり, 「} =B3*B4 \text{」}$$

を入力する。

価格	100
仕入れ値	100
売上個数	200.08
粗利	0
利益	0

データタブの（分析グループにある）ソルバーを選択し、「ソルバーのパラメーター」ダイアログボックスを出す。

「目的セルの設定」には利益のセル B5 を指定する。

「目標値」は「最大値」を選択する。

「変数セルの変更」には価格のセル B1 を指定する。

「制約条件の対象」には制約条件を入れる。ここでは、原因となる価格を整数のみで考えたいので、「価格のセル B1 は整数しかとらない」という制約条件を入れる。「制約条件の対象」の右にある「追加」をクリックし、「セル参照」に B1 を指定し、真ん中を「int」として OK ボタンを押す。

そして、「ソルバーのパラメーター」ダイアログボックスの「解決」を押すと、計算され、「ソルバーの結果」ダイアログボックスが出てくる。OK ボタンを押すと、最適化の結果の値、つまり、最大になった予測利益「約 5922（円）」が求められていることが確認できる。

価格	159
仕入れ値	100
売上個数	100.37
粗利	59
利益	5922

## 問題 9.2

### 問 9.2.1

分析ツールの「回帰分析」を選び、「入力 Y 範囲」は結果系のデータ（この場合は「売上個数」、C1:C11）、「入力 X 範囲」は原因系のデータ（この場合は「価格」、B1:B11）、「ラベル」にチェックする。「残差」にもチェックを入れておく。

観測値	予測値: 売上個数 (個)	残差
1	126.5	-4.5
2	111	2.84E-14
3	103.25	4.75
4	95.5	-8.5
5	111	9
6	126.5	0.5
7	142	1
8	134.25	-6.25
9	118.75	7.25
10	103.25	-3.25

これより、残差の絶対値が最も大きくなるのは 5 日目であり、その日の売上個数の実測値は 120（個）、予測値は 111（個）であることがわかる。

### 問題 9.2.2

回帰分析の結果、回帰式の傾きは-1.55、切片は 266 であることがわかる。

	係数
切片	<b>266</b>
価格 (円)	<b>-1.55</b>

つまり、価格 ( $x$ ) と売上個数 ( $y$ ) についての回帰式は

$$y = -1.55x + 266$$

となることがわかる。よって、[価格が 75 円のときの売上個数の予測値] は、空いているセルに「= [傾きが入力されたセル番地] \*75+ [切片が入力されたセル番地]」と入力すると求められる。149.75 (個) と計算される。

同様にすると、[価格が 115 円のときの売上個数の予測値] は 87.75 (個) と計算される。

### 問題 9.2.3

価格と売上個数についての回帰分析における R-2 乗値は約 0.871 であることがわかる。

回帰統計	
重相関 R	0.933151375
重決定 R2	<b>0.870771488</b>
補正 R2	0.854617924
標準誤差	6.118619125
観測数	10

### 問題 9.2.4

回帰分析の結果、価格 ( $x$ ) と売上個数 ( $y$ ) についての回帰式は

$$y = -1.55x + 266$$

となる。一般に、直線の方程式「 $y = ax + b$  ( $a, b$  は定数)」において、 $x$  が 1 増えると  $y$  は傾き  $a$  の分だけ増える。つまり、価格を 1 円上げると売上個数は  $-1.55$  (個) だけ増えると予測できることになる。これより、価格を 1 円下げると売上個数は約 1.55 (個) だけ増えると予測できる。

### 問題 9.2.5

回帰分析の結果、回帰式は下記のようなになる。

$$\text{売上個数} = -1.55 \times \text{価格} + 266$$

データタブの (分析グループにある) [ソルバー] を使って最適化の結果の値を求める。

	A	B	C	D	E	F	G
1		価格（円）	売上個数（個）		価格	60	
2	1日目	90	122		仕入れ値	60	
3	2日目	100	111		売上個数	=-1.55*F1+266	
4	3日目	105	108		粗利	=F1-F2	
5	4日目	110	87		利益	=F3*F4	
6	5日目	100	120				
7	6日目	90	127				
8	7日目	80	143				
9	8日目	85	128				
10	9日目	95	126				
11	10日目	105	100				
12							
13							

最大になった予測利益は「約 4827（円）」であり、それを与える販売価格は「116（円）」であることがわかる。

価格	116
仕入れ値	60
売上個数	86.2
粗利	56
利益	4827

## 問題 9.3

### 問題 9.3.1

分析ツールの「回帰分析」を選び、「入力 Y 範囲」は結果系のデータ（この場合は「売上個数」、C1:C15）、「入力 X 範囲」は原因系のデータ（この場合は「価格」、B1:B15）、「ラベル」にチェックする。「残差」にもチェックを入れておく。

観測値	予測値: 売上個数（個）	残差
1	143.1986385	-23.199
2	114.4419159	-3.4419
3	114.4419159	3.5581
4	75.22820326	-4.2282
5	143.1986385	7.8014
6	140.584391	-18.584
7	114.4419159	-13.442
8	101.3706783	-17.371
9	140.584391	27.416
10	101.3706783	1.6293
11	101.3706783	13.629
12	75.22820326	13.772
13	169.3411135	7.6589
14	143.1986385	4.8014

これより、[1 日目の売上個数の予測値] は約 143.2（個）であり、[2 日目の売上個数の予測値] は約 114.4（個）であることがわかる。

### 問題 9.3.2

6 日目（または 9 日目）の価格は 1000 円なので，「[価格が 1000 円のときの売上個数の予測値]」は「6 日目（または 9 日目）の売上個数の予測値」であり，約 140.6（個）であることがわかる．

### 問題 9.3.3

回帰分析の結果，回帰式の傾きは約-0.1307，切片は約 271.3 であることがわかる．

	係数
切片	<b>271.2967664</b>
価格（円）	<b>-0.130712375</b>

よって，「[価格が 1100 円のときの売上個数の予測値]」は，空いているセルに「= [傾きが入力されたセル番地] \*1100+ [切片が入力されたセル番地]」と入力すると求められる．約 127.5（個）と計算される．

### 問題 9.3.4

価格と売上個数についての回帰分析における R-2 乗値は約 0.789 であることがわかる．

回帰統計	
重相関 R	0.888399619
重決定 R2	<b>0.789253883</b>
補正 R2	0.771691707
標準誤差	14.94609869
観測数	14

### 問題 9.3.5

回帰分析の結果，価格（ $x$ ）と売上個数（ $y$ ）についての回帰式は

$$y = -0.1307x + 271.3$$

となる．

### 問題 9.3.6

一般に，直線の方程式「 $y = ax + b$ （ $a, b$  は定数）」において， $x$  が 1 増えると  $y$  は傾き  $a$  の分だけ増える．つまり，価格を 1 円上げると売上個数は約-0.1307（個）だけ増えると予測できることになる．これより，価格を 1 円下げると売上個数は約 0.1307（個）だけ増える，つまり，価格を 100 円下げると売上個数は約 13.07（個）だけ増えると予測できる．

### 問題 9.3.7

問題 9.3.5 より，回帰式は下記のようなになる．

$$\text{売上個数} = -0.1307 \times \text{価格} + 271.3$$

データタブの（分析グループにある）ソルバーを使って最適化の結果の値を求める．

	A	B	C	D	E	F	G
1		価格（円）	売上個数（個）		価格	800	
2	1日目	980	120		仕入れ値	800	
3	2日目	1200	111		売上個数	=-0.1307*F1+271.3	
4	3日目	1200	118		粗利	=F1-F2	
5	4日目	1500	71		利益	=F3*F4	
6	5日目	980	151				
7	6日目	1000	122				
8	7日目	1200	101				
9	8日目	1300	84				
10	9日目	1000	168				
11	10日目	1300	103				
12	11日目	1300	115				
13	12日目	1500	89				
14	13日目	780	177				
15	14日目	980	148				
16							
17							

最大になった予測利益は「約 53179（円）」であることがわかる（それを与える販売価格は「1438（円）」であることもわかる）。最大になった予測利益「約 53179（円）」の百の位を四捨五入して、千の位までの値にすると「約 53000（円）」となる。

価格	1438
仕入れ値	800
売上個数	83.3534
粗利	638
利益	53179

## 第 10 章

### 問題 10.1

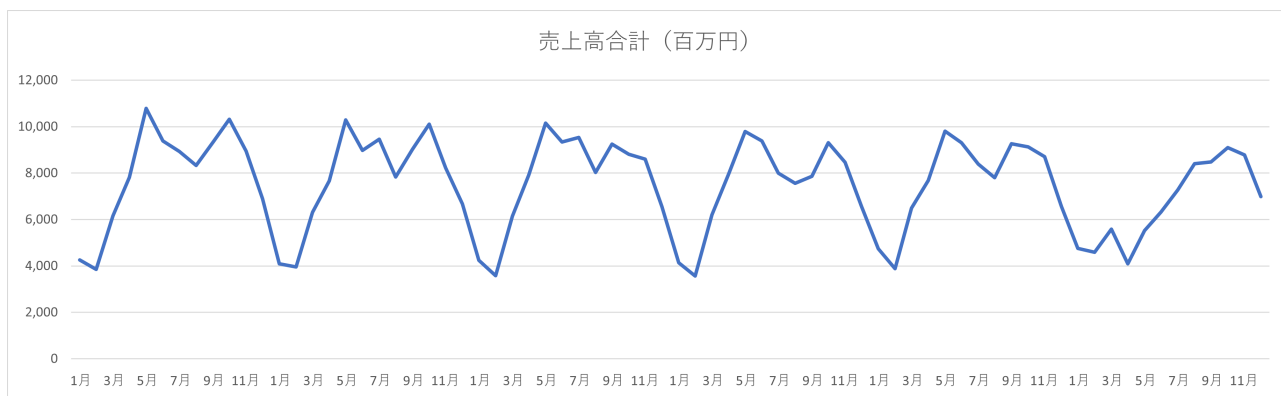
データタブの（分析グループにある）[データ分析] を選択して、分析ツールの「移動平均」を選ぶ。「入力範囲」は学習塾の売上高合計のデータ（C2:C25）、「区間」は「12」, 「出力先」は最初の学習塾の売上高合計データのとなり（D2）にする。

季節変動値は「売上高合計（原数値）÷ 移動平均」で求められる。2020 年 5 月の季節変動値を求めるセル（E18）に「=C18/D18」と計算式を入れると、約 0.590 であることがわかる。

### 問題 10.2

#### 問題 10.2.1

月と売上高合計の 2 列分（B14:C86）を選択して、挿入タブの [折れ線/面グラフの挿入] の「2-D 折れ線」の「折れ線」を選ぶ。

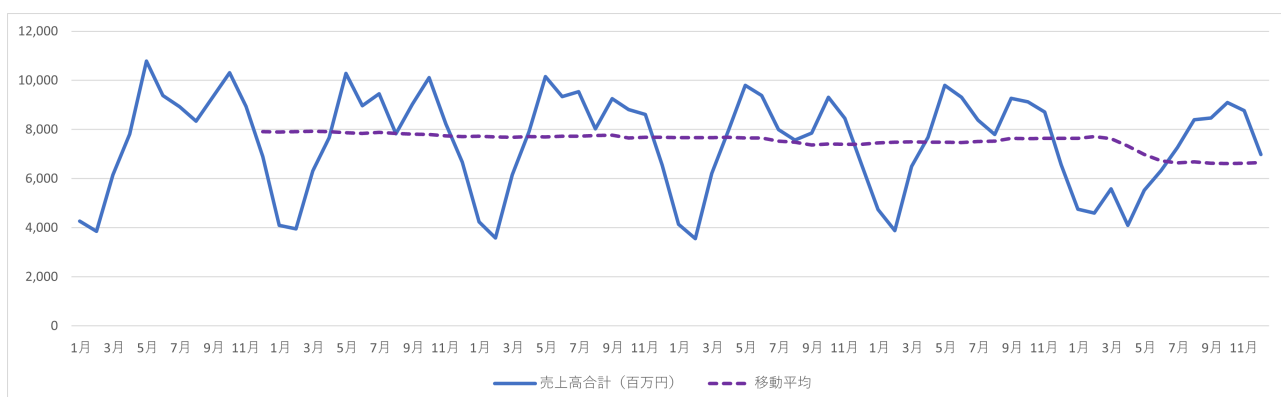


### 問題 10.2.2

データタブの（分析グループにある）[データ分析] を選択して、分析ツールの「移動平均」を選ぶ。「入力範囲」は売上高合計のデータ（C15:C86）、「区間」は「12」,「出力先」は最初の売上高合計データのとなり（D15）にする。2016 年 2 月の移動平均値は約 7912（百万円）であることがわかる。

### 問題 10.2.3

月と売上高合計と移動平均の 3 列分（B14:D86）を選択して、挿入タブの [折れ線/面グラフの挿入] の「2-D 折れ線」の「折れ線」を選ぶ。



### 問題 10.2.4

季節変動値は「売上高合計（原数値）÷ 移動平均」で求められる。セル E26 に「=C26/D26」と計算式を入れ、下にオートフィルする。2016 年 2 月の季節変動値は約 0.500 であることがわかる。

## 問題 10.3

### 問題 10.3.1

作成した問題 10.2 のファイルを開き、E 列の季節変動値のデータ（E26:E86）を選択する。右クリックをし、「コピー」を選択する。そのまま同じ場所で右クリックし、「貼り付けのオプション」から「値」を選ぶ。

### 問題 10.3.2

E 列の季節変動値のデータの 2016 年の分（E27:E38）を選択し、右クリックし、「コピー」を選択する。そして、セル J5 を右クリックし、「貼り付けのオプション」から「行/列の入れ替え」を選ぶ。同様のことを 2020 年の分まで繰り返す。

## 問題 10.4

### 問題 10.4.1

各月の平均値を AVERAGE 関数で求める。9 月の値は 1201464 (円) であり、百の位を四捨五入して、千の位までの値にすると、約 1201000 (円) となる。

### 問題 10.4.2

各月の中央値を MEDIAN 関数で求める。9 月の値は 1173420 (円) であり、百の位を四捨五入して、千の位までの値にすると、約 1173000 (円) となる。

### 問題 10.4.3

各月の最大値と最小値を取り除いたトリム平均を TRIMMEAN 関数で求める (1 月のトリム平均を「=TRIMMEAN(B2:B6,2/5)」で求め、右にオートフィルする)。9 月の値は 1159890 (円) であり、百の位を四捨五入して、千の位までの値にすると、約 1160000 (円) となる。

### 問題 10.4.4

各月のレンジを、[MAX 関数で求めた最大値] から [MIN 関数で求めた最小値] をひいて求める。9 月の値は 1147590 (円) であり、百の位を四捨五入して、千の位までの値にすると、約 1148000 (円) となる。

### 問題 10.4.5

各月の標準偏差を STDEV.P 関数で求める。9 月の値の百の位を四捨五入して、千の位までの値にすると、約 445000 (円) となる。

### 問題 10.4.6

支店ごとに AVERAGE 関数で平均値、STDEV.P 関数で標準偏差を求める (A 店の平均値をセル N2 に、標準偏差をセル O2 にそれぞれ求めて、下にオートフィルする)。

次に、(セル B17 に)「=STANDARDIZE(」と入力し、A 店の 1 月のももとの売上金額 (B2) をクリックして指定する。続けて、「,」を入力し、A 店の平均値が計算されているセル (N2) をクリックし F4 キーを 3 回押す。さらに、「,」を入力し、A 店の標準偏差が計算されているセル (O2) をクリックし F4 キーを 3 回押す。Enter キーを押すと A 店の 1 月の売上金額を標準化した値が計算される (セル B17 には「=STANDARDIZE(B2,\$N2,\$O2)」と入力される)。

下にオートフィルして、さらに右にもオートフィルしてデータ全部を標準化する (標準化したデータの最大値と最小値をそれぞれ MAX 関数と MIN 関数を用いて求めてもいい)。



B17 $\times \vee f_x$ =STANDARDIZE(B2,\$N2,\$O2)																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	平均値	標準偏差	
2	A店	1959390	1926180	1553490	1889280	1472310	738000	1512900	1771200	1837620	1852380	918810	1771200	1600230	379817	
3	B店	1114380	1184490	1036890	1066410	1143900	1033200	1143900	1118070	1173420	1221390	1287810	1228770	1146053	75072.2	
4	C店	594090	520290	450180	671580	741690	675270	675270	690030	752760	553500	675270	708480	642368	88911.1	
5	D店	1516590	1590390	1863450	1468620	1778580	1619910	1638360	1284120	1553490	1697400	1394820	1398510	1567020	160351	
6	E店	409590	453870	608850	682650	896670	988920	988920	1214010	690030	1483380	1767510	1767510	995993	453226	
7	平均値	1118808	1135044	1102572	1155708	1206630	1011060	1191870	1215486	1201464	1361610	1208844	1374894			
8	中央値	1114380	1184490	1036890	1066410	1143900	988920	1143900	1214010	1173420	1483380	1287810	1398510			
9	トリム平均	1075020	1098390	1066410	1072560	1170960	920040	1215240	1205400	1159890	1467390	1200480	1464930			
10	最大値	1959390	1926180	1863450	1889280	1778580	1619910	1638360	1771200	1837620	1852380	1767510	1771200			
11	最小値	409590	453870	450180	671580	741690	675270	675270	690030	690030	553500	675270	708480			
12	レンジ	1549800	1472310	1413270	1217700	1036890	944640	963090	1081170	1147590	1298880	1092240	1062720			
13	標準偏差	573262.13	579241.15	539620.84	469506.06	377860.32	334363.69	350079.7	346157.4	445489.22	456320.39	379934.56	394264.71			
14																
15																
16																
17		1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月			
17	A店	0.9456129	0.8581761	-0.123059	0.7610241	-0.336794	-2.270119	-0.229926	0.4501377	0.6250113	0.6638721	-1.794074	0.4501377			
18	B店	-0.421894	0.5120071	-1.4541	-1.060879	-0.028672	-1.503253	-0.028672	-0.372741	0.364549	1.0035338	1.8882821	1.1018392			
19	C店	-0.542986	-1.373029	-2.161569	0.3285585	1.1170988	0.3700606	0.3700606	0.5360691	1.2416052	-0.999509	0.3700606	0.7435797			
20	D店	-0.314497	0.1457426	1.8486303	-0.613653	1.3193544	0.3298386	0.4448986	-1.764253	-0.084377	0.8130905	-1.073893	-1.050881			
21	E店	-1.293842	-1.196142	-0.854194	-0.691361	-0.219146	-0.015605	-0.015605	0.4810352	-0.675078	1.0753748	1.702281	1.702281			
22																
23																
24																
25																
26																

これより、最も大きい標準化後の値は約 1.888 で、B 店の 11 月の売上金額を標準化したものである。また、最も小さい標準化後の値は約-2.270 で、A 店の 6 月の売上金額を標準化したものである（なお、標準化する際にスピルを使うなら、「=STANDARDIZE(B2:M6,N2:N6,O2:O6)」のように入力すればいい）。

## 第 11 章

### 問題 11.1

#### 問題 11.1.1

月ごとのトリム平均は TRIMMEAN 関数で求める。たとえば、セル J10 は「=TRIMMEAN(J5:J9,2/5)」で求める。

#### 問題 11.1.2

合計は SUM 関数で、つまり、「=SUM(J10:U10)」で求める。約 11.972 が計算される。

#### 問題 11.1.3

まず、補正值（セル J12）を「12÷ [トリム平均の合計値] (=12/V10)」とする。そして、各月の補正トリム平均（11 行目）を「トリム平均 × 補正值」（セル J11 は「=J10\*\$J\$12」または「=J10\*\$J12」）で求める。それらの合計値を SUM 関数でセル V11 に求めると 12 になるはずである。

	季節変動値												
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	
2016年	0.51764	0.49977	0.79623	0.96941	1.30704	1.14497	1.19954	0.99856	1.15489	1.29689	1.06302	0.86443	
2017年	0.54924	0.46532	0.8003	1.02727	1.31966	1.2086	1.23398	1.03592	1.19088	1.15064	1.11908	0.85259	
2018年	0.53955	0.46443	0.80856	1.03817	1.28031	1.22677	1.0636	1.01042	1.0663	1.25667	1.14332	0.88684	
2019年	0.63575	0.51865	0.86526	1.02718	1.31075	1.24628	1.11687	1.03589	1.21276	1.19667	1.13951	0.85883	
2020年	0.621	0.59633	0.73196	0.55898	0.79236	0.94071	1.0958	1.25679	1.27993	1.37534	1.32489	1.04857	合計値
トリム平均	0.56993	0.49458	0.8017	1.00795	1.29937	1.19344	1.1374	1.02741	1.18618	1.25008	1.13397	0.87003	11.972
補正トリム平均	0.57126	0.49574	0.80357	1.0103	1.3024	1.19623	1.14006	1.02981	1.18895	1.25299	1.13662	0.87207	12
補正值	1.00234												

#### 問題 11.1.4

補正トリム平均の値（J11:U11）を選択し、右クリックし、「コピー」を選択する。そして、セル J17 を右クリックし、「貼り付けのオプション」から「値」を選ぶ。

季節指数											
1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
0.57126	0.49574	0.80357	1.0103	1.3024	1.19623	1.14006	1.02981	1.18895	1.25299	1.13662	0.87207

#### 問題 11.1.5

季節指数の値（J17:U17）を選択し、右クリックし、「コピー」を選択する。そして、セル F15 を右クリックし、「貼り付けのオプション」から「行/列の入れ替え」を選ぶ。

同様のことを 2020 年の分まで繰り返す。

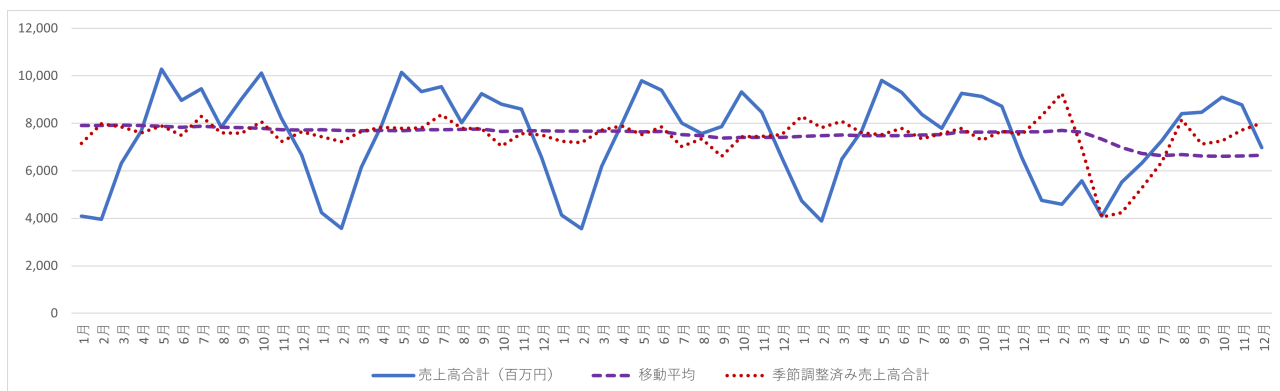
#### 問題 11.1.6

季節調整済み売上高合計は、「売上高合計（原数値）÷ 季節指数」で求められるので、セル G15 に「=C15/F15」と入力し、下にオートフィルする。2016 年 2 月の季節調整済み売上高合計は約 7976（百万円）であることがわかる。

年	月	売上高合計（百万円）	移動平均	季節変動値	季節指数	季節調整済み売上高合計
2016年	1月	4,091	7,903	0.517640608	0.57126	7161.375835
	2月	3,954	7,912	0.499768275	0.49574	7976.028823
	3月	6,310	7,925	0.796231256	0.80357	7852.463377
	4月	7,671	7,913	0.969407205	1.01030	7592.763914
	5月	10,288	7,871	1.307035096	1.30240	7899.254502
	6月	8,973	7,837	1.144965601	1.19623	7501.053178
	7月	9,453	7,881	1.199543176	1.14006	8291.668912
	8月	7,827	7,838	0.998564731	1.02981	7600.440734
	9月	9,023	7,813	1.154894724	1.18895	7589.051765
	10月	10,110	7,796	1.296888195	1.25299	8068.669429
	11月	8,223	7,736	1.063021136	1.13662	7234.613748
	12月	6,670	7,716	0.864428197	0.87207	7648.497987

#### 問題 11.1.7

月と売上高合計と移動平均と季節調整済み売上高合計の 4 列分（B14:D86 と G14:G86）を選択して、挿入タブの「折れ線/面グラフの挿入」の「2-D 折れ線」の「折れ線」を選ぶ。売上高合計、移動平均、季節調整済み売上高合計の折れ線グラフは下記ようになる。



### 問題 11.1.8

2月の季節指数が一番小さいことがわかる。

よって、翌年（2021年）の売上高合計を予測するとき、求めた季節指数のみの影響を受けると仮定すると、最も売上高合計が小さくなるのは2月であると予想できる。

## 問題 11.2

### 問題 11.2.1

月ごとのトリム平均は TRIMMEAN 関数で求める。1月についてのトリム平均（セル B11）は「=TRIMMEAN(B3:B8,2/6)」で求める。これをオートフィルすると、10月についてのトリム平均は約 1.208 であることがわかる。

### 問題 11.2.2

まず、問題 11.2.1 で求めた月ごとのトリム平均の合計をセル N11 に「=SUM(B11:M11)」と入力して求める。次に、合計値が 12 になるように各月のトリム平均を補正したもの（補正トリム平均）を 12 行目にそれぞれ求めたい。そのために、補正值（セル B13）を「 $12 \div [\text{トリム平均の合計値}] (=12/N11)$ 」とし、各月の補正トリム平均を「トリム平均  $\times$  補正值」（セル B12 は「=B11\*\$B\$13」または「=B11\*\$B13」）で求める（それらの合計値を SUM 関数でセル N12 に求めると 12 になるはずである）。10月についての補正トリム平均は約 1.179 であることがわかる。

なお、最大値と最小値については、問題に答えるためだけならば求める必要はない。

	季節変動値												
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	
2015年	0.71	0.63	1.09	0.99	1.18	0.69	1.01	1.89	1.11	1.22	0.98	1.21	
2016年	0.7	0.66	1.21	0.11	1.13	0.87	1.11	2.01	1.13	1.23	1.21	1.17	
2017年	0.68	0.72	1.12	0.12	1.14	0.76	1.2	1.98	1.21	1.23	1.32	1.1	
2018年	0.81	0.59	1.15	0.11	1.11	0.89	1.14	1.88	1.14	1.15	1.22	0.99	
2019年	0.77	0.62	1.22	0.13	1.22	0.65	1.03	1.82	1.21	1.09	1.32	1.08	
2020年	0.73	0.56	1.08	0.1	1.08	0.59	1.04	2.13	1.2	1.23	1.28	1.25	
最大値	0.81	0.72	1.22	0.99	1.22	0.89	1.2	2.13	1.21	1.23	1.32	1.25	
最小値	0.68	0.56	1.08	0.1	1.08	0.59	1.01	1.82	1.11	1.09	0.98	0.99	合計値
トリム平均	0.728	0.625	1.143	0.118	1.14	0.743	1.08	1.94	1.17	1.208	1.258	1.14	12.29
補正トリム平均	0.71	0.61	1.116	0.115	1.113	0.725	1.055	1.894	1.142	1.179	1.228	1.113	12
補正值	0.976												

## 問題 11.3

### 問題 11.3.1

表 11.9 の 8 月の季節指数は、各月の合計が 12 になるような値なので、12 から [8 月以外の合計] をひいて求める (SUM 関数の範囲に E4:K4 と M4:P4 を指定するには、E4:K4 をドラッグして選択したあと、Ctrl キーを押しながら M4:P4 をドラッグして選択すればいい。または、E4:K4 をドラッグして選択したあと、「,」を入力し、そのあと、M4:P4 をドラッグして選択してもいい)。1.34 となることがわかる。

季節指数 (補正トリム平均)											
1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
0.79	0.80	1.17	0.92	0.97	0.80	0.95	=12-SUM(E4:K4,M4:P4)	0.95	1.07	1.06	1.18

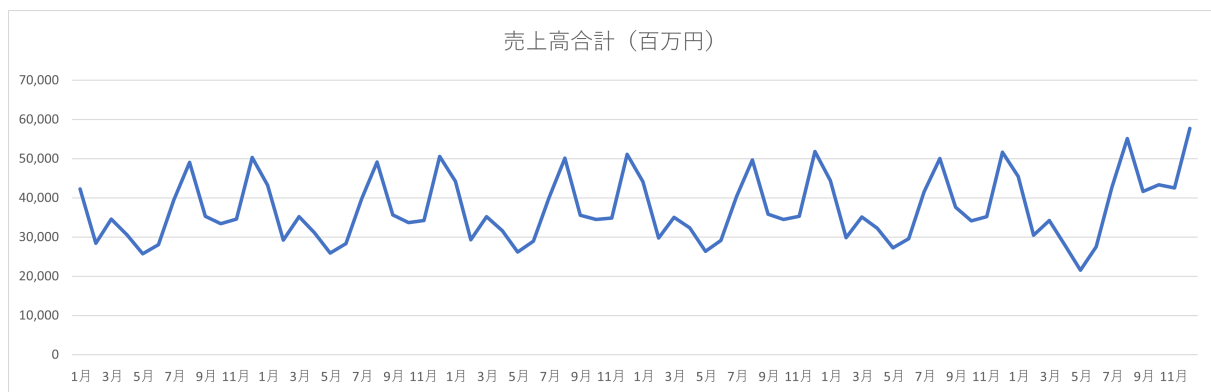
### 問題 11.3.2

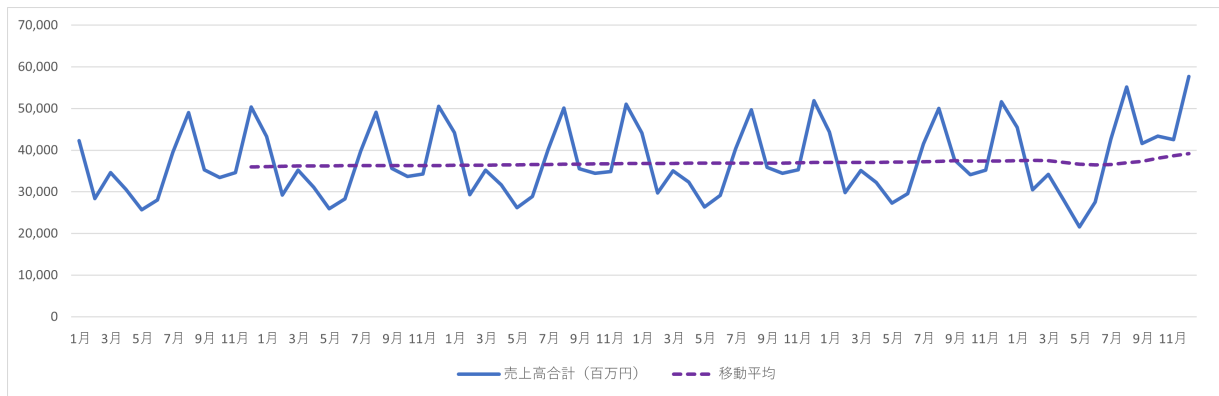
季節調整済み売上高合計は、「売上高合計 ÷ 季節指数」で求められる。表 11.8 の 2019 年 8 月の季節調整済み売上高合計は「2019 年 8 月の売上高合計 (原数値) ÷ 8 月の季節指数」(「=C21/L4」) で求められ、約 57911 (百万円) となる。

## 問題 11.4

### 問題 11.4.1

月と売上高合計の 2 列分 (B14:C86) を選択して、挿入タブの [折れ線/面グラフの挿入] の「2-D 折れ線」の「折れ線」を選ぶ。





### 問題 11.4.4

季節変動値は「売上高合計（原数値）÷ 移動平均」で求められる。セル E26 に「=C26/D26」と計算式を入れ、下にオートフィルする。2019 年 12 月の季節変動値は約 1.381 であることがわかる。

### 問題 11.5

### 問題 11.5.1

作成した問題 11.4 のファイルを開き、E 列の季節変動値のデータ (E26:E86) を選択する。右クリックをし、「コピー」を選択する。そのまま同じ場所で右クリックし、「貼り付けのオプション」から「値」を選ぶ。

### 問題 11.5.2

E 列の季節変動値のデータの 2016 年の分 (E27:E38) を選択し、右クリックし、「コピー」を選択する。  
そして、セル J5 を右クリックし、「貼り付けのオプション」から「行/列の入れ替え」を選ぶ。  
同様のことを 2020 年の分まで繰り返す。

### 問題 11.6

### 問題 11.6.1

月ごとのトリム平均は TRIMMEAN 関数で求める。たとえば、セル J10 は「=TRIMMEAN(J5:J9,2/5)」で求める。

### 問題 11.6.2

合計は SUM 関数で、つまり、「=SUM(J10:U10)」で求める。約 12.0544 が計算される。

### 問題 11.6.3

まず、補正值（セル J12）を「 $12 \div [\text{トリム平均の合計値}] (=12/V10)$ 」とする。そして、各月の補正トリム平均（11 行目）を「トリム平均  $\times$  補正值」（セル J11 は「 $=J10*\$J\$12$ 」または「 $=J10*\$J12$ 」）で求める。それらの合計値を SUM 関数でセル V11 に求めると 12 になるはずである。

	季節変動値												
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	
2016年	1.19975	0.80847	0.97304	0.85738	0.71616	0.78095	1.09485	1.35383	0.98069	0.92688	0.94319	1.39066	
2017年	1.21361	0.80544	0.96603	0.86756	0.71791	0.7914	1.09782	1.36723	0.96958	0.93858	0.94822	1.38766	
2018年	1.20046	0.80715	0.95193	0.87661	0.71503	0.79003	1.09276	1.34708	0.97077	0.93376	0.9548	1.40023	
2019年	1.19842	0.80556	0.94742	0.87074	0.73448	0.79639	1.11507	1.34148	1.00561	0.91144	0.94214	1.38139	
2020年	1.2123	0.81186	0.91342	0.75086	0.58952	0.75545	1.16956	1.49099	1.11651	1.13897	1.1005	1.47311	合計値
トリム平均	1.20417	0.80706	0.95513	0.86523	0.71637	0.78746	1.10258	1.35605	0.98569	0.93307	0.94874	1.39285	12.0544
補正トリム平均	1.19874	0.80342	0.95082	0.86132	0.71314	0.78391	1.09761	1.34993	0.98124	0.92886	0.94446	1.38657	12
補正值	0.99549												

#### 問題 11.6.4

補正トリム平均の値（J11:U11）を選択し、右クリックし、「コピー」を選択する。そして、セル J17 を右クリックし、「貼り付けのオプション」から「値」を選ぶ。

季節指数											
1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
1.19874	0.80342	0.95082	0.86132	0.71314	0.78391	1.09761	1.34993	0.98124	0.92886	0.94446	1.38657

#### 問題 11.6.5

季節指数の値（J17:U17）を選択し、右クリックし、「コピー」を選択する。そして、セル F15 を右クリックし、「貼り付けのオプション」から「行/列の入れ替え」を選ぶ。

同様のことを 2020 年の分まで繰り返す。

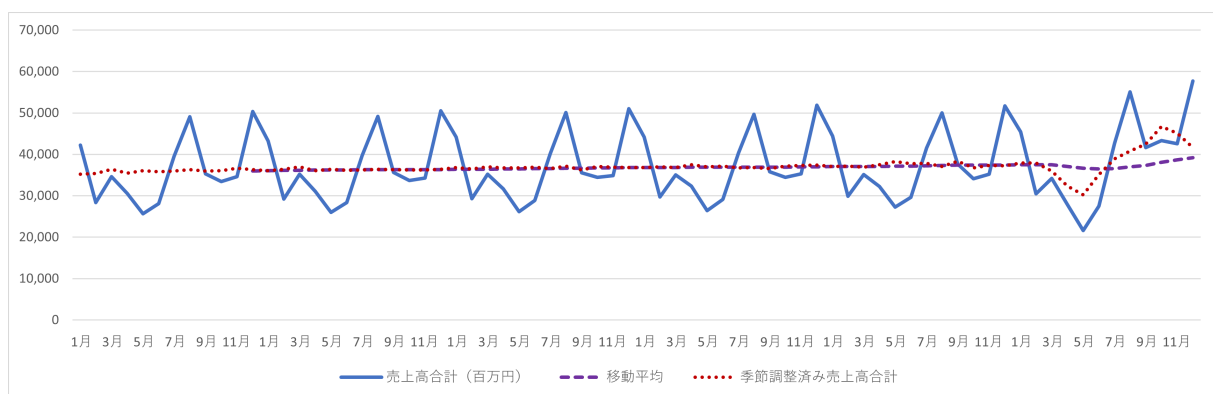
#### 問題 11.6.6

季節調整済み売上高合計は、「売上高合計（原数値）÷ 季節指数」で求められるので、セル G15 に「=C15/F15」と入力し、下にオートフィルする。2019 年 12 月の季節調整済み売上高合計は約 37257（百万円）であることがわかる。

年	月	売上高合計（百万円）	移動平均	季節変動値	季節指数	季節調整済み売上高合計
2019年	1月	44,398	37,047	1.198423624	1.198737	37037.31173
	2月	29,852	37,057	0.80556074	0.803418	37156.23562
	3月	35,114	37,063	0.947418123	0.950817	36930.35887
	4月	32,268	37,058	0.870743159	0.861324	37463.2407
	5月	27,273	37,133	0.734477883	0.713137	38243.70004
	6月	29,601	37,169	0.796387679	0.783907	37760.83593
	7月	41,558	37,269	1.11507476	1.097607	37862.37356
	8月	50,033	37,297	1.341484178	1.349927	37063.49151
	9月	37,658	37,448	1.005610014	0.981241	38377.91901
	10月	34,103	37,417	0.911442813	0.928862	36714.83137
	11月	35,248	37,413	0.942136611	0.944456	37320.95988
	12月	51,660	37,397	1.381388073	1.386567	37257.49582

#### 問題 11.6.7

月と売上高合計と移動平均と季節調整済み売上高合計の 4 列分（B14:D86 と G14:G86）を選択して、挿入タブの「折れ線/面グラフの挿入」の 2-D 折れ線を選ぶ。売上高合計、移動平均、季節調整済み売上高合計の折れ線グラフは下記ようになる。



### 問題 11.6.8

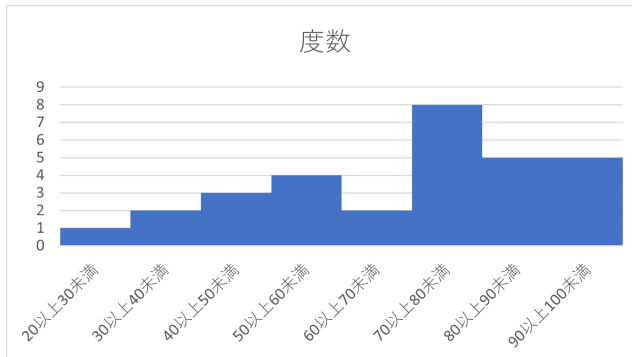
12月の季節指数が一番大きいことがわかる。よって、翌年（2021年）の売上高合計を予測するとき、求めた季節指数のみの影響を受けると仮定すると、最も売上高合計が大きくなるのは12月であると予想できる。

## 第12章

### 問題 12.1

度数分布表とヒストグラムの例は下記のようなになる。条件さえ満たせば階級はこれ以外の分け方でもいい。ここで、データを大きさの順に並べる際は、まずデータを1列（または1行）に配置してから並べ替えればいい。

階級	度数	相対度数	累積度数	累積相対度数
20以上30未満	1	0.033333	1	0.033333333
30以上40未満	2	0.066667	3	0.1
40以上50未満	3	0.1	6	0.2
50以上60未満	4	0.133333	10	0.333333333
60以上70未満	2	0.066667	12	0.4
70以上80未満	8	0.266667	20	0.666666667
80以上90未満	5	0.166667	25	0.833333333
90以上100未満	5	0.166667	30	1



### 問題 12.2

下記のように計算式を入れる。

階級	度数	相対度数	累積度数	累積相対度数
4.0～6.0	1	0.02	1	0.02
6.0～8.0	3	0.06	4	=E2+C3
8.0～10.0	7	=B4/D9	11	0.22
10.0～12.0	9	0.18	20	0.4
12.0～14.0	9	0.18	29	0.58
14.0～16.0	=D9-(SUM(B2:B6,B8:B9))	=B7/D9	41	0.82
16.0～18.0	8	0.16	=D7+B8	0.98
18.0～20.0	1	0.02	50	=E8+C9
			↑	
最後の階級の累積度数が50なので、全度数は50である。				



$$A = \text{全度数} - [\text{他の度数の合計}] = 12$$

$$(\text{または}, A = [\text{この階級の累積度数}] - [\text{前の階級の累積度数}] = 12)$$

$$B = \text{度数} / \text{合計} = 0.14$$

$$C = \text{度数} / \text{合計} = 0.24$$

$$D = [\text{前の階級の累積度数}] + [\text{この階級の度数}] = 49$$

$$E = [\text{前の階級の累積相対度数}] + [\text{この階級の相対度数}] = 0.08$$

$$F = 1 \quad (\text{最後の階級の累積相対度数は} 1 \text{ になる})$$

次のような結果が得られる。

階級	度数	相対度数	累積度数	累積相対度数
4.0～6.0	1	0.02	1	0.02
6.0～8.0	3	0.06	4	0.08
8.0～10.0	7	0.14	11	0.22
10.0～12.0	9	0.18	20	0.4
12.0～14.0	9	0.18	29	0.58
14.0～16.0	12	0.24	41	0.82
16.0～18.0	8	0.16	49	0.98
18.0～20.0	1	0.02	50	1

### 問題 12.3

4 が正解である。

実際に度数分布表やヒストグラムを完成させる必要はない。問題の選択肢の 4 つのヒストグラムはどれも階級が 20～30, 30～40, 40～50, 50～60, 60～70, 70～80, 80～90 であることがわかる。よって、データを大きさの順に並べ替えて、この階級で度数だけ調べ、それに一致するヒストグラムを選べば正解はわかる。ここで、データを大きさの順に並べる際は、まずデータを 1 列に配置してから並べ替えればいい。

この階級で度数を調べると下記のようなになる。

階級	度数
20～30	2
30～40	6
40～50	8
50～60	12
60～70	10
70～80	8
80～90	4

なお、問題に答えるためだけなら、すべての階級についての度数は調べる必要はなく、

階級	度数
20～30	2
30～40	6
40～50	8

だけ調べれば、4 が正解だということはわかる。



## 問題 12.4

選択肢 1 のヒストグラムを見て各階級における度数を入力し、相対度数、累積度数、累積相対度数について下記のように計算式を入力する。相対度数については最初のセル（C2）を下にオートフィルする。累積度数、累積相対度数については 2 番目のセル（D3, E3）を下にオートフィルする。

	A	B	C	D	E	F
1	階級	度数	相対度数	累積度数	累積相対度数	
2	20～30	3	=B2/\$B\$9	=B2	=C2	
3	30～40	5	=B3/\$B\$9	=D2+B3	=E2+C3	
4	40～50	8	=B4/\$B\$9	=D3+B4	=E3+C4	
5	50～60	9	=B5/\$B\$9	=D4+B5	=E4+C5	
6	60～70	13	=B6/\$B\$9	=D5+B6	=E5+C6	
7	70～80	7	=B7/\$B\$9	=D6+B7	=E6+C7	
8	80～90	5	=B8/\$B\$9	=D7+B8	=E7+C8	
9	合計	=SUM(B2:B8)	=SUM(C2:C8)			
10						
11						

これより、選択肢 1 のヒストグラムを与える度数分布表は

	A	B	C	D	E	F
1	階級	度数	相対度数	累積度数	累積相対度数	
2	20～30	3	0.06	3	0.06	
3	30～40	5	0.1	8	0.16	
4	40～50	8	0.16	16	0.32	
5	50～60	9	0.18	25	0.5	
6	60～70	13	0.26	38	0.76	
7	70～80	7	0.14	45	0.9	
8	80～90	5	0.1	50	1	
9	合計	50	1			
10						
11						

となる。50 以上 60 未満の階級で累積相対度数が 0.5 となっているので、この階級までのデータ（つまり 60 未満のデータ）が全体のちょうど半数であることがわかる。

## 問題 12.5

最後の階級の累積度数が 40 なので、全度数は 40 である。「相対度数 = 度数 ÷ 全度数」より、「度数 = 相対度数 × 全度数」であるので、度数を求めることができる。また、最初の階級の累積度数は度数そのものであり、最初の階級の累積相対度数は相対度数そのものである。

下記のように計算式を入れると、

階級	度数	相対度数	累積度数	累積相対度数
20～30	=D2	0.1	4	=C2
30～40	=C3*D7	0.175	=D2+B3	0.275
40～50	=C4*D7	0.3	=D3+B4	0.575
60～70	9	=B5/D7	32	=E4+C5
80～90	7	=B6/D7	39	0.975
90～100	=D7-D6	0.025	40	1

次のような結果が得られる。

階級	度数	相対度数	累積度数	累積相対度数
20～30	4	0.1	4	0.1
30～40	7	0.175	11	0.275
40～50	12	0.3	23	0.575
60～70	9	0.225	32	0.8
80～90	7	0.175	39	0.975
90～100	1	0.025	40	1

## 問題 12.6

一般に、最後の階級の累積相対度数は 1 である。これより、最後の階級の相対度数を求めることができる。また、「相対度数 = 度数 ÷ 全度数」より、「全度数 = 度数 ÷ 相対度数」であるので、全度数（つまり、最後の階級の累積度数）も求めることができる。

下記のように計算式を入れると、

階級	度数	相対度数	累積度数	累積相対度数
0～10	=C2*D9	=E2	=B2	0.0625
10～20	=C3*D9	0.125	=D2+B3	=E2+C3
20～30	15	=B4/D9	=D3+B4	=E3+C4
30～40	=C5*D9	=E5-E4	=D4+B5	0.475
40～50	=D6-D5	=B6/D9	43	=E5+C6
60～70	=D7-D6	=B7/D9	52	=E6+C7
80～90	=C8*D9	=E8-E7	=D7+B8	0.9125
90～100	7	=E9-E8	=B9/C9	1

次のような結果が得られる。

階級	度数	相対度数	累積度数	累積相対度数
0～10	5	0.0625	5	0.0625
10～20	10	0.125	15	0.1875
20～30	15	0.1875	30	0.375
30～40	8	0.1	38	0.475
40～50	5	0.0625	43	0.5375
60～70	9	0.1125	52	0.65
80～90	21	0.2625	73	0.9125
90～100	7	0.0875	80	1

## 第 13 章

### 問題 13.1

#### 問題 13.1.1

「性別」と「年代」は質的変数であり、「スーパーに行く回数」と「コンビニに行く回数」は量的変数である。

#### 問題 13.1.2

スーパーに行く回数とコンビニに行く回数の 2 列分 (D1:E32) を選択して、挿入タブの「散布図 (x, y) またはバブルチャートの挿入」の「散布図」を選ぶ。

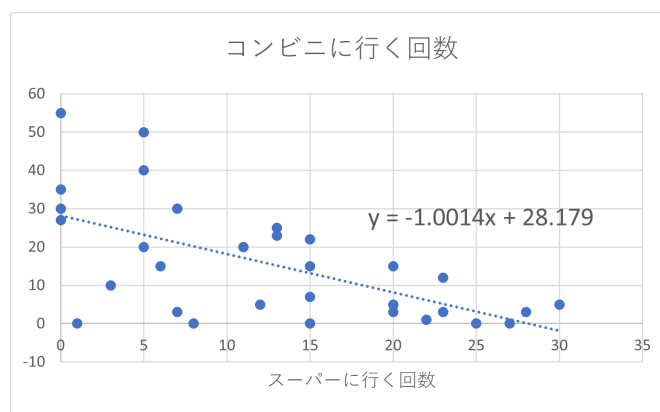
作成したグラフを選択した状態で、(グラフの) デザインタブの [グラフ要素の追加] をクリックし、「軸ラベル」の「第 1 横軸」を選択すると (横) 軸ラベルが出てくる。(横) 軸ラベルに「スーパーに行く回数」と入力する。

### 問題 13.1.3

セルに「=CORREL(」と入力し、相関係数を求める 2 変数のうちの片方の変数 (スーパーに行く回数) のデータが入力されているセル範囲 (D2:D32) をドラッグして選択したあと、Ctrl キーを押しながらもう片方の変数 (コンビニに行く回数) が入力されているセル範囲 (E2:E32) をドラッグして選択する (または、セル範囲 (D2:D32) をドラッグして選択したあと「,」を入力し、そのあと、セル範囲 (E2:E32) をドラッグして選択する)。Enter キーを押すと相関係数が計算される。小数第 3 位までの表示にすると約  $-0.606$  であることがわかる (セルには「=CORREL(D2:D32,E2:E32)」と入力される)。

### 問題 13.1.4

散布図のマーカー (点) の上で右クリックして「近似曲線の追加」を選択する。近似曲線の書式設定において、「線形近似」が選ばれていることを確認し、「グラフに数式を表示する」にチェックを入れる。



すると、右下がりの直線に近似されることがわかる。その直線の式 (回帰式) は

$$y = -1.0014x + 28.179$$

であることもわかる (この式をことばで書くと、

$$\text{コンビニに行く回数} = -1.0014 \times \text{スーパーに行く回数} + 28.179$$

ということである)。

### 問題 13.1.5

問題 13.1.4 より、スーパーに行く回数 ( $x$ ) とコンビニに行く回数 ( $y$ ) についての回帰式は

$$y = -1.0014x + 28.179$$

である。よって、空いているセルに「=-1.0014\*10+28.179」と入力すると、[スーパーに行く回数が 10 回のときのコンビニに行く回数の予測値] 約 18.165 (回) が計算される。

### 問題 13.1.6

問題 13.1.4 より、スーパーに行く回数 ( $x$ ) とコンビニに行く回数 ( $y$ ) についての回帰式は

$$y = -1.0014x + 28.179$$

である。ここで、一般に、直線の方程式「 $y = ax + b$  ( $a, b$  は定数)」において、 $x$  が 1 増えると  $y$  は傾き  $a$  の分だけ増える。つまり、スーパーに行く回数が 1 回増えると、コンビニに行く回数は傾き約  $-1.0014$  (回) だけ増える、つまり、約  $1.0014$  (回) だけ減ると予測できることになる。

### 問題 13.1.7

性別ごとにデータの並べ替えをする。そのため、B 列の性別が入力されているどこかのセルを選択した状態で、ホームタブの(編集グループにある)「並べ替えとフィルター」をクリックし、「昇順」または「降順」を選択する。男女別に、それぞれの相関係数を CORREL 関数で求めればよい。女性については約  $-0.683$  となり、男性については約  $-0.535$  であることがわかる。

## 問題 13.2

### 問題 13.2.1

表中のどこかのセルを選択した状態で、挿入タブの(テーブルグループにある)「ピボットテーブル」の「テーブルまたは範囲から」を選択する。「テーブル/範囲」に表全体が選択されていることを確認する。OK ボタンを押すと「ピボットテーブルのフィールド」が出てくる。

「性別」、「コンビニに行く回数」にチェックを入れる。すると、「行」ボックスに「性別」、「値」ボックスに「合計/コンビニに行く回数」が入る。

ピボットテーブルの行(横)に性別の項目が出てきて、それぞれの項目(男女)ごとの「コンビニに行く回数」の合計が出てくる。求めたいのは平均値なので、「値」ボックス内の「合計/コンビニに行く回数」をクリックし、値フィールドの設定を選び、「選択したフィールドのデータ」を「平均」にする。

行ラベル	平均 / コンビニに行く回数
女性	11.125
男性	20.06666667
総計	15.4516129

女性のコンビニに行く回数の平均値は約 11.13 (回) であることがわかる。

### 問題 13.2.2

「ピボットテーブルのフィールド」の「年代」、「スーパーに行く回数」、「コンビニに行く回数」にチェックが入っている状態にする(これら以外にチェックが入っていたら外す)。

「値」ボックス内が「平均/コンビニに行く回数」、「平均/スーパーに行く回数」になるように、値フィールドの設定の「選択したフィールドのデータ」を「平均」にする。

行ラベル	平均 / コンビニに行く回数	平均 / スーパーに行く回数
10代	29.8	6.6
20代	27	7
30代	24.16666667	9.666666667
40代	7.5	14.25
50代	2.166666667	17.33333333
60代	1.4	21.4
総計	15.4516129	12.70967742

コンビニに行く回数の平均値がスーパーに行く回数の平均値より多い年代は10代、20代、30代であることがわかる。

### 問題 13.2.3

「ピボットテーブルのフィールド」の「性別」、「年代」、「スーパーに行く回数」にチェックが入っている状態にする（「コンビニに行く回数」はこの問題では不要だが、チェックがあっても差し支えない）。

「値」ボックス内が「平均／スーパーに行く回数」になるように、値フィールドの設定の「選択したフィールドのデータ」を「平均」にする（下記では、「行」ボックスに「年代」、「列」ボックスに「性別」が入っているが、これ以外の配置でも問題ない）。

平均 / スーパーに行く回数	列ラベル		
行ラベル	女性	男性	総計
10代	6.5	6.666666667	6.6
20代	10	2.5	7
30代	13.33333333	6	9.666666667
40代	17.5	11	14.25
50代	27.66666667	7	17.33333333
60代	20	23.5	21.4
総計	16.3125	8.866666667	12.70967742

これより、30代の女性のスーパーに行く回数の平均値は約13.33（回）、30代の男性のスーパーに行く回数の平均値は6（回）であることがわかる。

空いているセルに「=」を入力し、30代の女性のスーパーに行く回数の平均値が計算されているセルをクリックする。続けて、「-」を入力し、30代の男性のスーパーに行く回数の平均値が計算されているセルをクリックする。30代の女性と30代の男性のスーパーに行く回数の平均値の差は約7.33（回）であることがわかる。

### 問題 13.2.4

「ピボットテーブルのフィールド」の「性別」、「年代」、「コンビニに行く回数」にチェックが入っている状態にする（「スーパーに行く回数」はこの問題では不要だが、チェックがあっても差し支えない）。

「値」ボックス内が「平均／コンビニに行く回数」になるように、値フィールドの設定の「選択したフィールドのデータ」を「平均」にする（下記では、「行」ボックスに「年代」、「列」ボックスに「性別」が入っているが、これ以外の配置でも問題ない）。

平均 / コンビニに行く回数	列ラベル		
行ラベル	女性	男性	総計
10代	26	32.33333333	29.8
20代	18.33333333	40	27
30代	15.66666667	32.66666667	24.16666667
40代	6	9	7.5
50代	2.666666667	1.666666667	2.166666667
60代	1.333333333	1.5	1.4
総計	11.125	20.06666667	15.4516129

男性のうち、コンビニに行く回数の平均値が最も大きいのは20代（平均値40回）であることがわかる。

### 問題 13.2.5

「ピボットテーブルのフィールド」の「性別」,「コンビニに行く回数」にチェックが入っている状態にする（これら以外にチェックが入っていたら外す）。

「値」ボックス内の「合計/コンビニに行く回数」（または「平均/コンビニに行く回数」）を「列」ボックスに移動させる。そして,（「ピボットテーブルのフィールド」の上部にある）ボックスにチェックの入った「性別」（という文字列）をドラッグして「値」ボックスに移動させる。すると,「値」ボックスに「個数/性別」が入る（この結果より,女性のうち,コンビニに行く回数が10回以下の人が占める割合をわり算で求めてもいい）。

「値」ボックス内の「個数/性別」をクリックし,「値フィールドの設定」を選び,「計算の種類」タブを選択する。「計算の種類」を「行集計に対する比率」に変更する。ピボットテーブルのコンビニに行く回数の項目（最上行の0,1,3,5,7,⋯,55）が書かれたセルで右クリックし,「グループ化」を選択する。「グループ化」ダイアログボックスが出てくるので,単位を「11」にする（10回以下とは「0回から10回」なので,11である）。

個数 / 性別	列ラベル					
行ラベル	0-10	11-21	22-32	33-43	44-55	総計
女性	56.25%	25.00%	18.75%	0.00%	0.00%	100.00%
男性	40.00%	13.33%	20.00%	13.33%	13.33%	100.00%
総計	48.39%	19.35%	19.35%	6.45%	6.45%	100.00%

女性のうち,コンビニに行く回数が10回以下の人が占める割合は56.25%であることがわかる。

#### 問題 13.2.6

「ピボットテーブルのフィールド」の「年代」,「スーパーに行く回数」にチェックが入っている状態にする（これら以外にチェックが入っていたら外す）。

「値」ボックス内の「合計/スーパーに行く回数」を「列」ボックス内に移動させる。また,ボックスにチェックの入った「年代」（という文字列）をドラッグして「値」ボックスに移動させる。すると,「値」ボックスに「個数/年代」が入る（この結果より,10代の人のうち,スーパーに行く回数が0回の人割合をわり算で求めてもいい）。

「値」ボックス内の「個数/年代」をクリックし,「値フィールドの設定」を選び,「計算の種類」タブを選択する。「計算の種類」を「行集計に対する比率」に変更する。

個数 / 年代	列ラベル																		
行ラベル	0	1	3	5	6	7	8	11	12	13	15	20	22	23	25	27	28	30	総計
10代	40.00%	0.00%	0.00%	20.00%	0.00%	0.00%	0.00%	0.00%	0.00%	20.00%	20.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
20代	20.00%	0.00%	20.00%	20.00%	0.00%	20.00%	0.00%	0.00%	0.00%	0.00%	0.00%	20.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
30代	16.67%	0.00%	0.00%	16.67%	16.67%	0.00%	0.00%	16.67%	0.00%	16.67%	0.00%	0.00%	0.00%	16.67%	0.00%	0.00%	0.00%	0.00%	100.00%
40代	0.00%	0.00%	0.00%	0.00%	0.00%	25.00%	0.00%	0.00%	0.00%	0.00%	50.00%	25.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%
50代	0.00%	16.67%	0.00%	0.00%	0.00%	0.00%	16.67%	0.00%	16.67%	0.00%	0.00%	0.00%	0.00%	0.00%	16.67%	0.00%	16.67%	16.67%	100.00%
60代	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	20.00%	20.00%	20.00%	20.00%	0.00%	20.00%	0.00%	0.00%	100.00%
総計	12.90%	3.23%	3.23%	9.68%	3.23%	6.45%	3.23%	3.23%	3.23%	6.45%	12.90%	9.68%	3.23%	6.45%	3.23%	3.23%	3.23%	3.23%	100.00%

すると,10代の人のうち,スーパーに行く回数が0回の人割合は40%であることがわかる。

#### 問題 13.2.7

問題 13.2.6 の結果の状態において,「ピボットテーブルのフィールド」の「値」ボックス内の「個数/年代」をクリックし,「値フィールドの設定」を選び,「計算の種類」タブを選択する。「計算の種類」を「列集計に対する比率」に変更する。

個数 / 年代	列ラベル	0	1	3	5	6	7	8	11	12	13	15	20	22	23	25	27	28	30	総計
行ラベル		0																		
10代		50.00%	0.00%	0.00%	33.33%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%	25.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	16.13%
20代		25.00%	0.00%	100.00%	33.33%	0.00%	50.00%	0.00%	0.00%	0.00%	0.00%	0.00%	33.33%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	16.13%
30代		25.00%	0.00%	0.00%	33.33%	100.00%	0.00%	0.00%	100.00%	0.00%	50.00%	0.00%	0.00%	0.00%	50.00%	0.00%	0.00%	0.00%	0.00%	19.35%
40代		0.00%	0.00%	0.00%	0.00%	0.00%	50.00%	0.00%	0.00%	0.00%	0.00%	50.00%	33.33%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	12.90%
50代		0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	100.00%	100.00%	19.35%
60代		0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	25.00%	33.33%	100.00%	50.00%	0.00%	100.00%	0.00%	0.00%	0.00%	16.13%
総計		100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

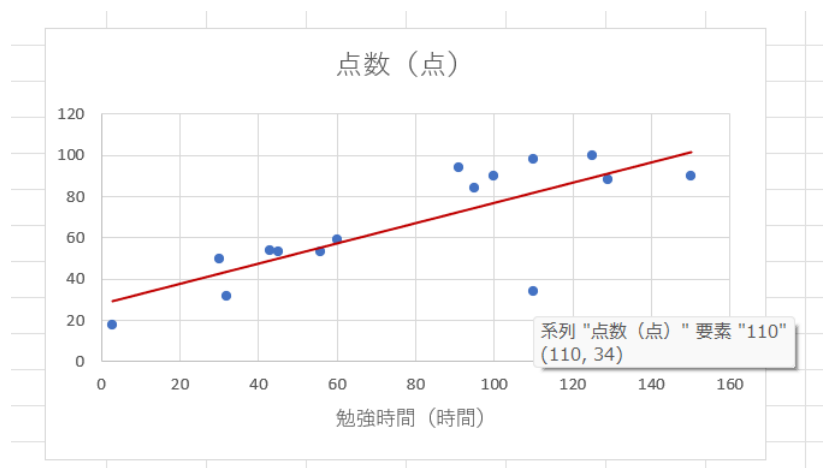
すると、スーパーに行く回数が0回の人のうち、10代の人の割合は50%であることがわかる（問題13.2.6において「計算の種類」を変更する前の状態より、スーパーに行く回数が0回の人のうち、10代の人の割合をわり算で求めてもいい）。

## 第14章

### 問題14.1

#### 問題14.1.1

勉強時間（横軸）と点数（縦軸）の散布図を作成し、近似曲線を追加すると、1点だけ下に飛び出しているのが見える。その点にマウスポインタを近づけ確認すると、勉強時間110（時間）、点数34（点）のデータであり、それは4番のデータであることがわかる。よって、外れ値ではないかと考えられるのは4番のデータである。



#### 問題14.1.2

データタブの（分析グループにある）[データ分析]を選択する。分析ツールの「回帰分析」を選び、「入力Y範囲」は結果系のデータ（この場合は「点数」、D1:D16）、「入力X範囲」は原因系のデータ（この場合は「勉強時間」、C1:C16）、「ラベル」にチェックする。「残差」にもチェックを入れる。

これより、勉強時間が原因で点数が結果という因果関係を想定して回帰分析を行ったとき、一番予測が外れている（残差の絶対値が一番大きい）のは4番のデータ（残差：約-47.88）であることがわかる。



観測値	予測値: 点数 (点)	残差
1	43.59086041	-11.59086
2	81.88083655	16.119163
3	57.33598005	1.6640199
4	81.88083655	-47.88084
5	29.35484364	-11.35484
6	55.37239153	-2.372392
7	48.99072884	5.0092712
8	72.55379108	21.446209
9	42.60906615	7.3909338
10	76.97186525	13.028135
11	74.5173796	9.4826204
12	49.9725231	3.0274769
13	91.20788201	-3.207882
14	101.5167217	-11.51672
15	89.24429349	10.755707

### 問題 14.1.3

データタブの（分析グループにある）「データ分析」を選択し、分析ツールの「相関」で相関係数を求める。

	勉強時間（時間）	点数（点）
勉強時間（時間）	1	
点数（点）	0.78625926	1

これより、勉強時間と点数の相関係数は約 0.786 であることがわかる。

次に、外れ値である 4 番のデータを削除し、分析ツールの「相関」で相関係数を求める。

	勉強時間（時間）	点数（点）
勉強時間（時間）	1	
点数（点）	0.923625013	1

これより、外れ値を取り除いた勉強時間と点数の相関係数は約 0.924 であることがわかる。

### 問題 14.2

科目ごとに AVERAGE 関数で平均値、STDEV.P 関数で標準偏差をそれぞれ求める（国語の平均値をセル B8 に、標準偏差をセル B9 に求めて、右にオートフィルする）。

次に、（セル I2 に）「=STANDARDIZE(」と入力し、国語の A のもともとの点数 (B2) をクリックして指定する。続けて、「,」を入力し、国語の平均点が計算されているセル (B8) をクリックし F4 キーを 2 回押す。さらに、「,」を入力し、国語の標準偏差が計算されているセル (B9) をクリックし F4 キーを 2 回押す。Enter キーを押すと国語の A の点数を標準化した値が計算される（セル I2 には「=STANDARDIZE(B2,B\$8,B\$9)」と入力される。これは、下にオートフィルするときに「平均」と「標準偏差」の行番号を固定するために、その前に「\$」を付けているのである。一方、あとで右にオートフィルするために、列番号 B の前には「\$」を付けてはいけない）。

下にオートフィルして、さらに右にもオートフィルしてデータ全部を標準化する（なお、標準化する際にスピルを使うなら、セル I2 に「=STANDARDIZE(B2:F7,B8:F8,B9:F9)」と入力すればいい）。



I2 : ✕ ✓ fx =STANDARDIZE(B2,B\$8,B\$9)														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		国語	数学	英語	理科	社会			国語	数学	英語	理科	社会	
2	A	56	87	94	57	81		A	-1.02871	1.1292	1.073957	-0.64483	0.362512	
3	B	88	70	80	53	91		B	0.88889	-0.04609	0.24632	-1.03172	1.199079	
4	C	79	70	74	70	70		C	0.349563	-0.04609	-0.10838	0.612585	-0.55771	
5	D	45	78	97	70	54		D	-1.68789	0.506988	1.251307	0.612585	-1.89622	
6	E	89	78	60	80	79		E	0.948815	0.506988	-0.93602	1.579823	0.195199	
7	F	82	41	50	52	85		F	0.529339	-2.051	-1.52719	-1.12845	0.697139	
8	平均値	73.17	70.67	75.83	63.67	76.67								
9	標準偏差	16.68749	14.46452	16.91564	10.33871	11.95361								
10														
11														

標準化した値の絶対値が2を超えるのはFの数学の点数であることがわかる。