



SASで実践！実務で使える販売・需要予測

SAS で実践！実務で使える 販売・需要予測

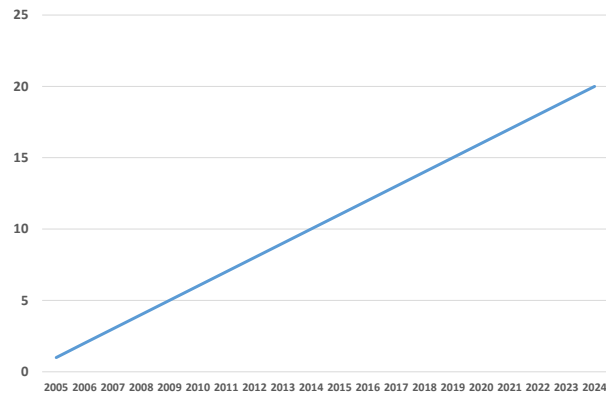
コースノート

実務で使える販売・需要予測

本日の内容

- ◇プロットして観察する。
- ◇パターンを抽出する。
 - ・時系列分析
トレンド、季節変動
- ◇因果関係を利用する。
 - ・重回帰分析
トレンド、季節指数、ダミー変数

売上高推移（設立～20年）



売上高は順調に伸びている!?

3

© Data Science Institute

年度	売上高	前年比
2005	1	-
2006	2	2.00
2007	3	1.50
2008	4	1.33
2009	5	1.25
2010	6	1.20
2011	7	1.17
2012	8	1.14
2013	9	1.13
2014	10	1.11
2015	11	1.10
2016	12	1.09
2017	13	1.08
2018	14	1.08
2019	15	1.07
2020	16	1.07
2021	17	1.06
2022	18	1.06
2023	19	1.06
2024	20	1.05

・年度間の差に着目

+1

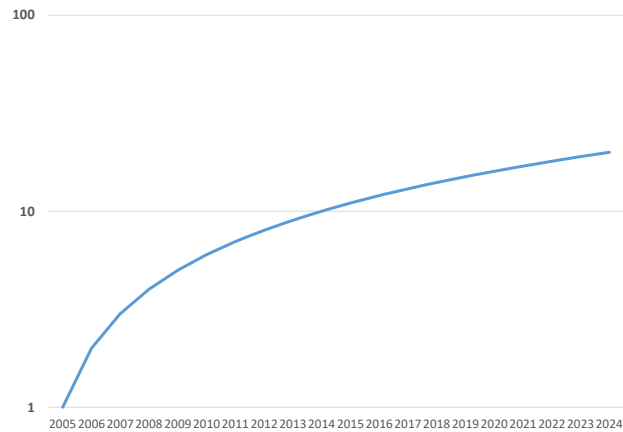
・年度間の比に着目



比率は年々減少！

4

© Data Science Institute

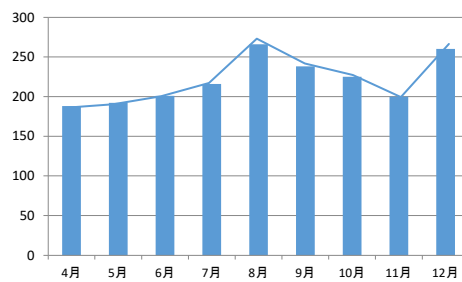


増加率は年々減少傾向！

5

© Data Science Institute

折れ線グラフと棒グラフの違いは何か？



目の動きは棒をイメージして上下に動かすべき



目の動きを斜めに動かすグラフは？

6

© Data Science Institute

支店（東北、関東、沖縄）別売上高

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
東北	188	192	201	216	266	238	225	200	260	213	198	218
関東	730	648	813	752	802	868	856	910	1183	1020	923	955
沖縄	41	40	55	63	113	42	53	50	65	48	45	47

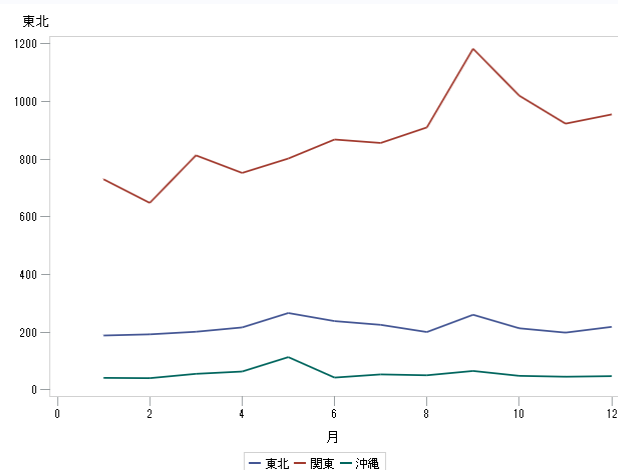
	4月	5月	差	前月比
東北	216	266	50	1.231
関東	752	802	50	1.066
沖縄	63	113	50	1.794

	8月	9月	差	前月比
東北	200	260	60	1.3
関東	910	1183	273	1.3
沖縄	50	65	15	1.3

7

© Data Science Institute

折れ線グラフ

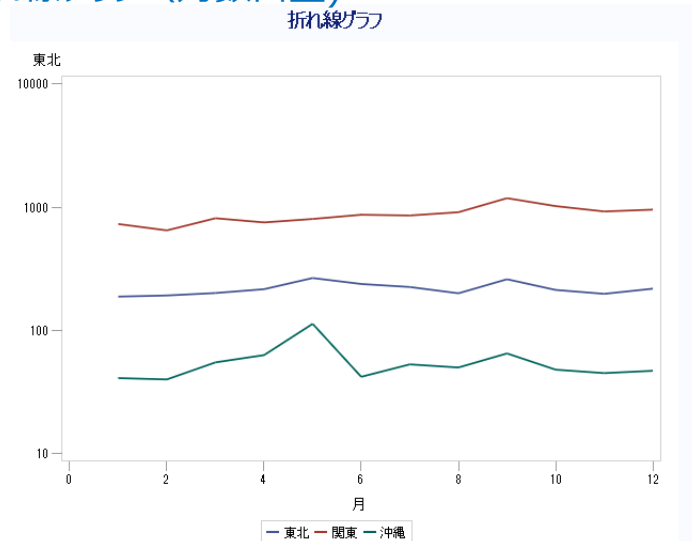


⇒ 普通目盛の折れ線グラフでは比較困難

8

© Data Science Institute

折れ線グラフ（対数目盛）



9

© Data Science Institute

新製品と主力製品の売上高

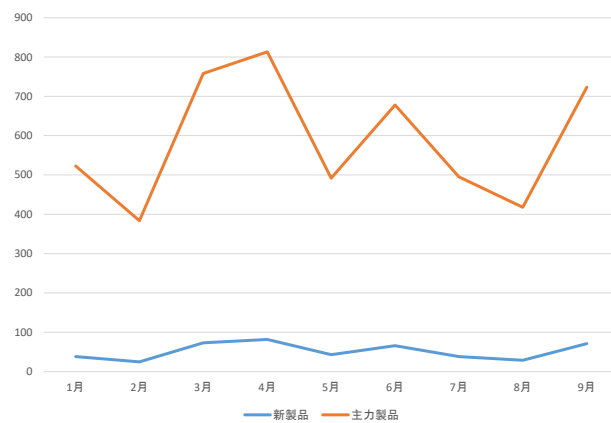
(百万円)

	新製品	主力製品
1月	38	523
2月	25	384
3月	73	758
4月	82	813
5月	43	492
6月	66	678
7月	38	495
8月	29	418
9月	71	723

10

© Data Science Institute

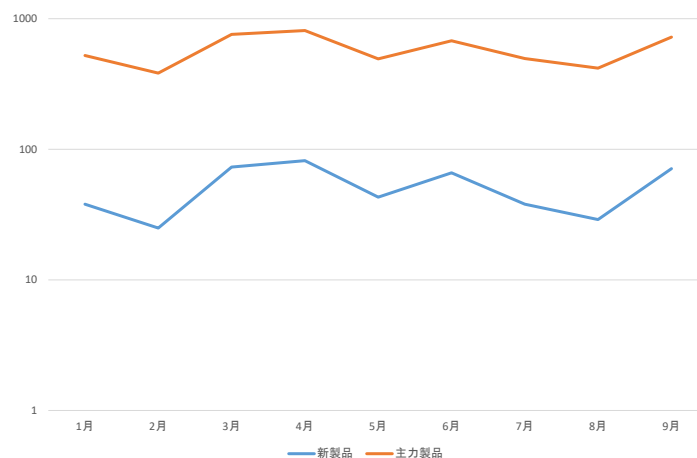
折れ線グラフ（普通目盛）



11

© Data Science Institute

折れ線グラフ（対数目盛）



12

© Data Science Institute

人間の五感は対数に変換されている

◇ウェーバー・フェヒナーの法則

弁別閾（気づくことができる最小の刺激差）は刺激の値に比例

手に重りを100gのせ、少しずつ重りを加え、重さの違いを感じたのが110gのとき、手に重りを200gのせ、1gずつ重りを加えると、重さの違いを感じるのは220gのときである。

デシベル、P H、マグニチュード、等星・・・

指数平滑法

- ・短期的な将来を予測する手法
- ・現在に近い値により多くのウェイトをおく。
過去よりも現在の出来事の影響が大きいときに有効

$$t+1\text{期の予測値} = \alpha \times t\text{期の実測値} + (1-\alpha) \times t\text{期の予測値}$$

指数平滑法による予測

金曜日の売上高	売上高
10月 4日	2105
10月11日	1816
10月18日	1842
10月25日	1907
11月 1日	1776
11月 8日	1794
11月15日	2067
11月22日	1778
11月29日	1811
12月 6日	1921
12月13日	1686
12月20日	1751
12月27日	?

15

© Data Science Institute

時系列分析

多くの場合、時系列の時間的变化は変動要素を持つ。

【 変動要素 】

(1) 傾向変動 (T : trend)

データが時系列に進む方向を示す要素

(2) 循環変動 (C : cycle movement)

景気変動などの長周期の変動要素

(3) 季節的変動 (S : seasonal movement)

季節的要因、12ヶ月の確定周期で繰り返す変動要素

(4) 不規則変動 (I : irregular component)

上記3つ以外のすべての変動要因

循環変動 (C) は傾向変動 (T) に含まれることが多い

$$Y = (TCI) \times S$$

16

© Data Science Institute

移動平均法

変動に一定の規則性があるとき、その変動をある程度排除

◇3項移動平均

年	1	2	3	4	5	6	7	8	9	10	11
売上高	32	58	63	25	72	52	31	43	52	39	61

求める項の値に前後の項の値を加えて3で除する。

1年：計算できない

2年： $(32 + 58 + 63) / 3 = 51.0$

3年： $(58 + 63 + 25) / 3 = 48.7$

.....

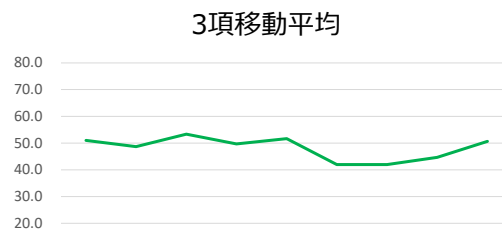
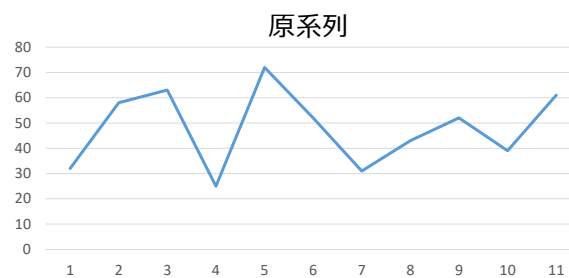
10年： $(52 + 39 + 61) / 3 = 50.7$

11年：計算できない

年	1	2	3	4	5	6	7	8	9	10	11
売上高	-	51.0	48.7	53.3	49.7	51.7	42.0	42.0	44.7	50.7	-

17

© Data Science Institute



18

© Data Science Institute

◇4項移動平均

年	1	2	3	4	5	6	7	8	9	10	11	12
売上高	32	58	63	25	72	52	31	43	52	39	61	55

偶数項の移動平均は、中央項が存在しない。

1年：計算できない

2年：計算できない

3年： $(32/2 + 58 + 63 + 25 + 72/2) / 4 = 49.5$

4年： $(58/2 + 63 + 25 + 72 + 52/2) / 4 = 53.8$

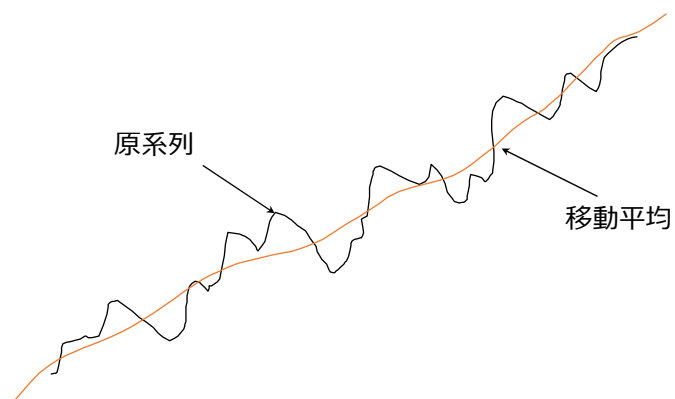
.....

10年： $(43/2 + 52 + 39 + 61 + 55/2) / 4 = 50.3$

11年：計算できない

12年：計算できない

年	1	2	3	4	5	6	7	8	9	10	11	12
売上高	-	-	49.5	53.8	49.0	47.3	47.0	42.9	45.0	50.3	-	-



◇エアコンの売上高

	2023	2024
1月	318	315
2月	298	305
3月	185	192
4月	173	205
5月	199	231
6月	328	334
7月	415	482
8月	298	353
9月	203	281
10月	192	187
11月	195	192
12月	381	380

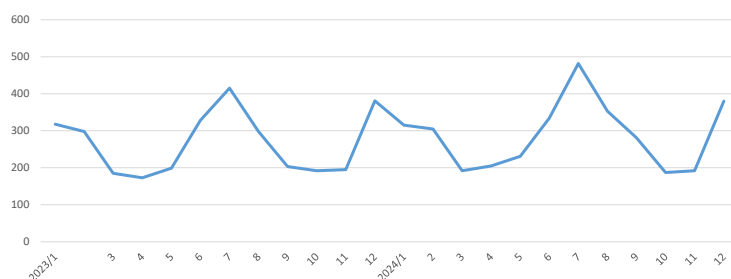
➡ 売上高の傾向変動（TC）を求める

21

© Data Science Institute

2年間（24ヶ月）推移

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
2023	318	298	185	173	199	328	415	298	203	192	195	381
2024	315	305	192	205	231	334	482	353	281	187	192	380

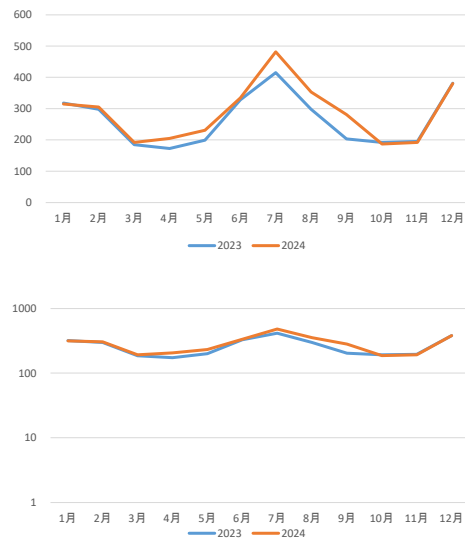


22

© Data Science Institute

季節変動の確認

	2023	2024
1月	318	315
2月	298	305
3月	185	192
4月	173	205
5月	199	231
6月	328	334
7月	415	482
8月	298	353
9月	203	281
10月	192	187
11月	195	192
12月	381	380



傾向変動（TC）を求める



季節変動を除去する

季節変動 : 12ヶ月の周期



12項（12ヶ月）移動平均による季節変動の除去

12ヶ月移動平均

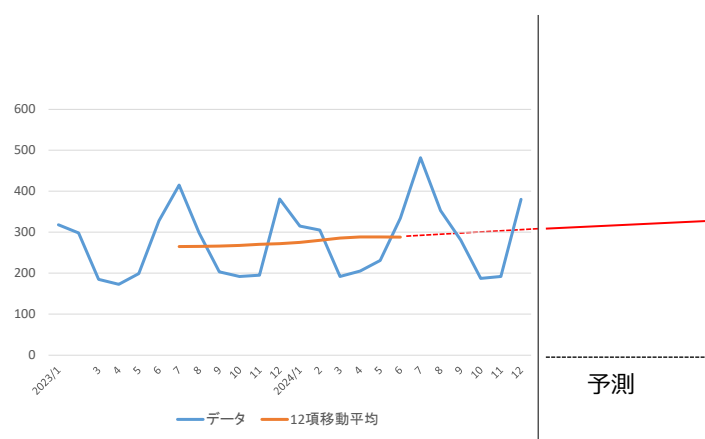
C8		fx		=(B2/2+SUM(B3:B13)+B14/2)/12			
1	A	B	C	D	E	F	G
2	年月	データ	12項移動平均				
3	2023/1	318					
4		298					
5	3	185					
6	4	173					
7	5	199					
8	6	328					
9	7	415	265.3				
10	8	298	265.5				
11	9	203	266.0				
12	10	192	267.7				
13	11	195	270.3				
14	12	381	271.9				
15	2024/1	315	275.0				
16	2	305	280.0				
17	3	192	285.6				
18	4	205	288.6				
19	5	231	288.3				
20	6	334	288.1				
21	7	482					
22	8	353					
23	9	281					
24	10	187					
25	11	192					
26	12	380					

傾向変動
(TCI)

25

© Data Science Institute

12項移動平均（TCI）の予測値を求める



26

© Data Science Institute

季節変動（季節指数）を求める

1) 売上高を傾向変動（TC）で割る。

$$Y = TC \cdot S \cdot I \quad S \cdot I = Y \div TC$$

A	B	C	D
年月	データ	12項移動平均	
2023/1	318		
	298		
3	185		
4	173		
5	199		
6	328		
7	415	265.3	156.4
8	298	265.5	112.3
9	203	266.0	76.3
10	192	267.7	71.7
11	195	270.3	72.1
12	381	271.9	140.1
2024/1	315	275.0	114.6
2	305	280.0	108.9
3	192	285.6	67.2
4	205	288.6	71.0
5	231	288.3	80.1
6	334	288.1	115.9
7	482		
8	353		
9	281		
10	187		
11	192		
12	380		

$$156.4 = 415 \div 265.3 \times 100$$

傾向変動
(SI)

27

© Data Science Institute

2) 各月の季節変動を求め、総和を1200に調整する。

$$156.4 + 112.3 + \dots + 115.9 = 1186.8$$

A	B	C	D	E
年月	データ	12項移動平均	季節変動	季節指数
2023/1	318			
	298			
3	185			
4	173			
5	199			
6	328			
7	415	265.3	156.4	158.2
8	298	265.5	112.3	113.5
9	203	266.0	76.3	77.2
10	192	267.7	71.7	72.5
11	195	270.3	72.1	72.9
12	381	271.9	140.1	141.7
2024/1	315	275.0	114.6	115.8
2	305	280.0	108.9	110.1
3	192	285.6	67.2	68.0
4	205	288.6	71.0	71.8
5	231	288.3	80.1	81.0
6	334	288.1	115.9	117.2
7	482			
8	353			
9	281			
10	187			
11	192			
12	380			

$$158.2 = 156.4 \times 1200 \div 1186.8$$

28

© Data Science Institute

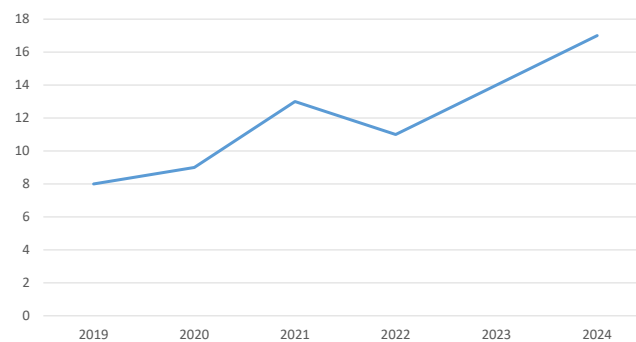
トレンドによる予測（傾向線の当てはめ）

年度（No.）	売上高
2019（1）	8
2020（2）	9
2021（3）	13
2022（4）	11
2023（5）	14
2024（6）	17
2025（7）	？

29

© Data Science Institute

売上高（年度別推移）



30

© Data Science Institute

回帰分析

売上高 \leftarrow 広告費

従属変数 \leftarrow 説明変数（独立変数）

説明変数が1つ：単回帰

説明変数が2つ以上：重回帰

目的：回帰式を求め予測する。

売上高 \leftarrow 年度

31

© Data Science Institute

EG

回帰分析（SAS EG）

1. データを入力する。

	年度	売上高
1	1	8
2	2	9
3	3	13
4	4	11
5	5	14
6	6	17

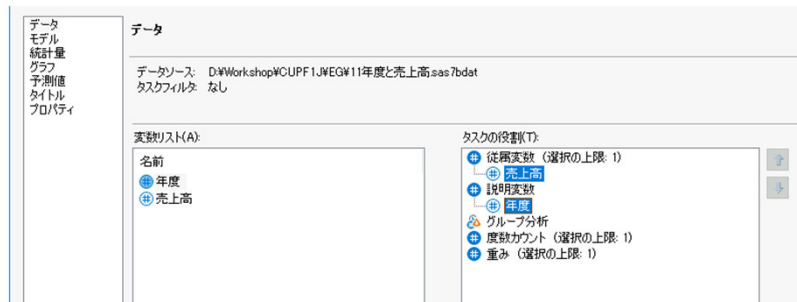
2. 「回帰分析」－「線形回帰分析」を選択する。

- > パレット
- > 回帰分析
 - HP ロジスティック回帰分析
 - HP 線形回帰分析
 - ロジスティック回帰分析
 - 一般化線形モデル
 - 線形回帰分析 ☆
 - 非線形回帰分析
- > 管理図
- > 記述統計
- > 工程能力分析

32

© Data Science Institute

3.「データ」をクリックし、売上高を「従属変数」、年度（No.）を「説明変数」に設定する。



売上高 \Leftarrow 年度 (No.)

Root MSE	1.40915	R2 値	0.8582
従属変数の平均	12.00000	調整済み R2 値	0.8227
変動係数	11.74295		

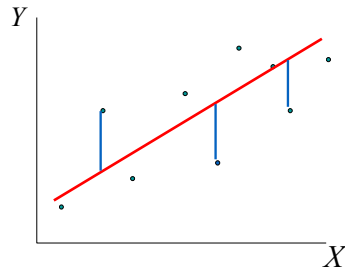
パラメータの推定					
変数	自由度	パラメータ推定値	標準誤差	t 値	Pr > t
Intercept	1	6.20000	1.31185	4.73	0.0091
年度	1	1.65714	0.33685	4.92	0.0079

$$\text{売上高} = 6.2 + 1.657 \times \text{年度 (No.)}$$

自由度調整済み決定係数 0.8227

⇒ 回帰式により約82.3%説明できる

傾向線のあてはめ（最小2乗法）

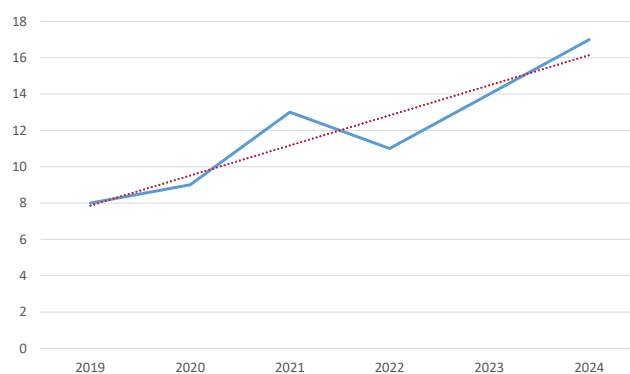


各データと傾向線との垂直距離の2乗和を最小

外れ値の影響を大きく受ける

35

© Data Science Institute



売上高 = $6.2 + 1.657 \times \text{年}$ (No.)

2025年予測値 = $6.2 + 1.6571 \times 7 = 17.80$

36

© Data Science Institute

トレンド（傾向線の当てはめ）による予測値の算出

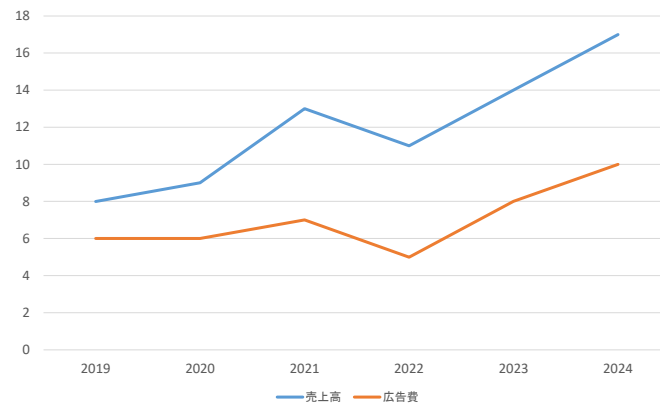
年度 (No.)	売上高	予測値
2019 (1)	8	7.86
2020 (2)	9	9.51
2021 (3)	13	11.17
2022 (4)	11	12.83
2023 (5)	14	14.49
2024 (6)	17	16.14
2025 (7)	?	17.80

$$2025\text{年予測値} = 6.2 + 1.657 \times 7 = 17.80$$

年度別広告費と売上高

年度	売上高	広告費
2019	8	6
2020	9	6
2021	13	7
2022	11	5
2023	14	8
2024	17	10
2025	?	11

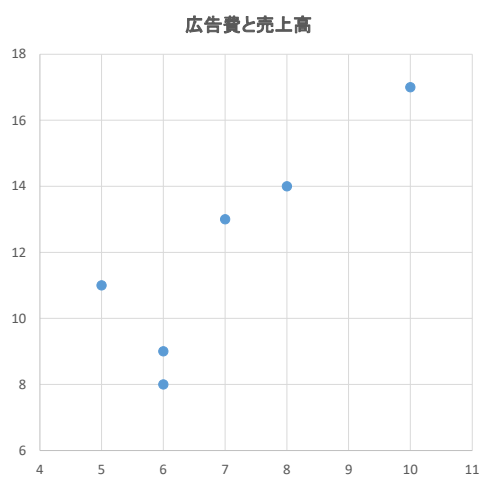
年度別推移



39

© Data Science Institute

散布図



40

© Data Science Institute

回帰分析 (SAS EG)

1. データを入力する。

	売上高	広告費	C	D	E
1	8	6			
2	9	6			
3	13	7			
4	11	5			
5	14	8			
6	17	10			

2. 「回帰分析」－「線形回帰分析」を選択する。

- > パレット
- ▼ 回帰分析
 - HP ロジスティック回帰分析
 - HP 線形回帰分析
 - ロジスティック回帰分析
 - 一般化線形モデル
 - 線形回帰分析 ☆
 - 非線形回帰分析
- > 管理図
- > 記述統計
- > 工程能力分析

3. 「データ」をクリックし、売上高を「従属変数」、広告費を「説明変数」に設定する。

データ
モデル
統計量
グラフ
予測値
タイトル
プロパティ

データ

データソース: LocalWORK.12広告費と売上高
タスクフィルタ: なし

変数リスト(A):

名前

- 売上高
- 広告費
- C
- D
- E
- F

タスクの役割(T):

- 従属変数 (選択の上限: 1)
- 説明変数
- グループ分析
- 度数カウント (選択の上限: 1)
- 重み (選択の上限: 1)

売上高 \Leftarrow 広告費

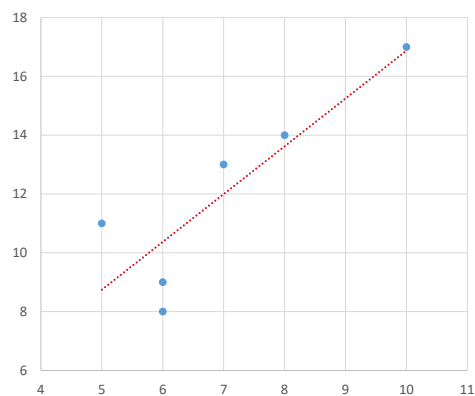
Root MSE	1.85406	R2 乗	0.7545
従属変数の平均	12.00000	調整済み R2 乗	0.6931
変動係数	15.45041		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	0.62500	3.33171	0.19	0.8603
広告費	1	1.62500	0.46351	3.51	0.0248

売上高 = $0.625 + 1.625 \times \text{広告費}$

自由度調整済み決定係数 : 0.6931

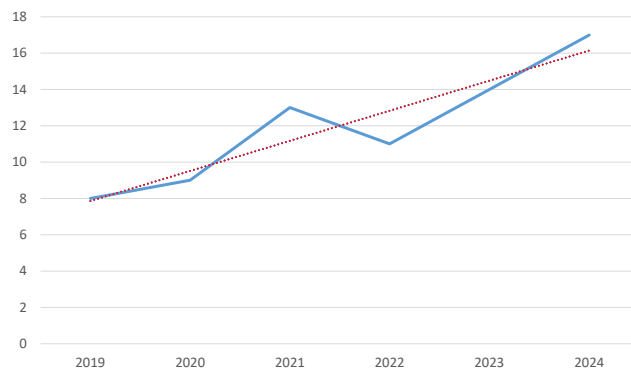
広告費と売上高



売上高 = $0.625 + 1.625 \times \text{広告費}$

2025年予測値 = $0.625 + 1.625 \times 11 = 18.5$

トレンドの検討



$$\text{売上高} = 6.2 + 1.657 \times \text{年 (No.)}$$

トレンドを説明変数に追加する

45

© Data Science Institute

トレンド（予測値）を説明変数に追加

年度 (No.)	売上高	広告費	トレンド (予測値)
2019 (1)	8	6	7.86
2020 (2)	9	6	9.51
2021 (3)	13	7	11.17
2022 (4)	11	5	12.83
2023 (5)	14	8	14.49
2024 (6)	17	10	16.14

$$\text{トレンド (予測値)} = 6.2 + 1.657 \times \text{年}$$

$$\text{例) 2019} \quad 7.86 = 6.2 + 1.657 \times 1$$

46

© Data Science Institute

回帰分析 (SAS EG)

1. データを入力する。

	売上高	広告費	トレンド
1	8	6	7.86
2	9	6	9.51
3	13	7	11.17
4	11	5	12.83
5	14	8	14.49
6	17	10	16.14

2. 「回帰分析」－「線形回帰分析」を選択する。

- > パレット
- ▼ 回帰分析
 - HP ロジスティック回帰分析
 - HP 線形回帰分析
 - ロジスティック回帰分析
 - 一般化線形モデル
 - 線形回帰分析 ☆
 - 非線形回帰分析
- > 管理図
- > 記述統計
- > 工程能力分析

3. 「データ」をクリックし、売上高を「従属変数」、広告費、トレンドを「説明変数」に設定する。

データ
モデル
統計量
グラフ
予測値
タイトル
プロパティ

データ

データソース: D:\Workshop\WCUPF1J\WEGW13トレンドを追加.sas7bdat
タスクフィルタ: なし

変数リスト(A):

名前

- 売上高
- 広告費
- トレンド

タスクの役割(T):

- 従属変数 (選択の上限: 1)
 - 売上高
- 説明変数
 - 広告費
 - トレンド
- グループ分析
 - 度数カウント (選択の上限: 1)
 - 重み (選択の上限: 1)

売上高 ← 広告費、トレンド（予測値）

Root MSE	1.02182	R2 乗	0.9441
従属変数の平均	12.00000	調整済み R2 乗	0.9068
変動係数	8.51517		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	-1.60102	1.96441	-0.82	0.4748
広告費	1	0.78676	0.36654	2.15	0.1211
トレンド	1	0.67450	0.21152	3.19	0.0498

売上高 = $-1.601 + 0.787 \times \text{広告費} + 0.675 \times \text{トレンド}$

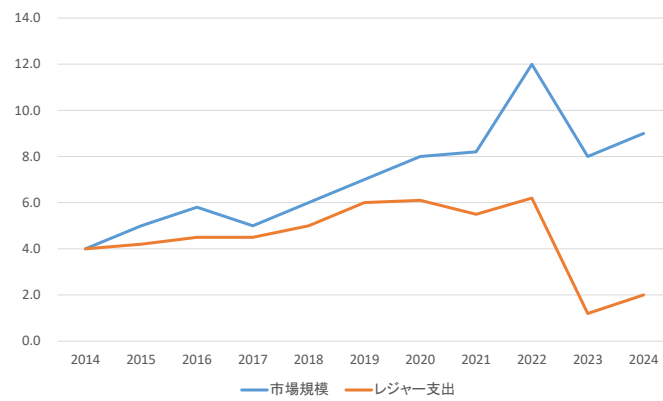
自由度調整済み決定係数 : 0.9068

0.6931 \Rightarrow 0.9068

事例1

年	市場規模	レジャー支出
2014	4.0	4.0
2015	5.0	4.2
2016	5.8	4.5
2017	5.0	4.5
2018	6.0	5.0
2019	7.0	6.0
2020	8.0	6.1
2021	8.2	5.5
2022	12.0	6.2
2023	8.0	1.2
2024	9.0	2.0
2025	?	3.0

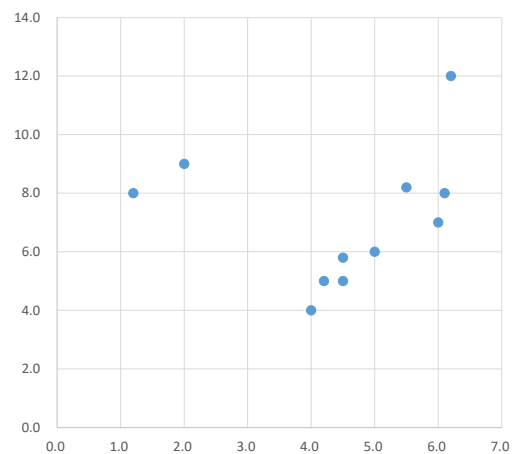
年度別推移



51

© Data Science Institute

散布図



相関係数 (r) = 0.125

市場規模とレジャー支出は無関係 !?

52

© Data Science Institute

トレンド（予測値）を求める

年	市場規模
1	4.0
2	5.0
3	5.8
4	5.0
5	6.0
6	7.0
7	8.0
8	8.2
9	12.0
10	8.0
11	9.0

53

© Data Science Institute

EG

データを入力する。

	年	市場規模	レジャー支出
1	1	4.0	4.0
2	2	5.0	4.2
3	3	5.8	4.5
4	4	5.0	4.5
5	5	6.0	5.0
6	6	7.0	6.0
7	7	8.0	6.1
8	8	8.2	5.5
9	9	12.0	6.2
10	10	8.0	1.2
11	11	9.0	2.0

54

© Data Science Institute

市場規模を「従属変数」、年を「説明変数」に設定する。

The screenshot shows the SAS Data Editor window. On the left, a list of variables is shown: 名前, 年, 市場規模, and レジャー支出. The variable 年 is selected with a blue circle, and 市場規模 is selected with a blue square. On the right, the 'Tasks' pane shows the following settings: 従属変数 (選択の上限: 1) is set to 市場規模, 説明変数 is set to 年, グループ分析 is set to 度数カウント (選択の上限: 1), and 重み (選択の上限: 1) is set to 1.

Root MSE	1.27160	R2 乗	0.7190
従属変数の平均	7.09091	調整済み R2 乗	0.6878
変動係数	17.93284		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	3.60000	0.82231	4.38	0.0018
年	1	0.58182	0.12124	4.80	0.0010

$$\text{市場規模} = 3.6 + 0.582 \times \text{年}$$

トレンド（予測値）を算出し、説明変数に追加

年	市場規模	レジャー支出	トレンド
2014	4	4.0	4.18
2015	5	4.2	4.76
2016	5.8	4.5	5.35
2017	5	4.5	5.93
2018	6	5.0	6.51
2019	7	6.0	7.09
2020	8	6.1	7.67
2021	8.2	5.5	8.25
2022	12	6.2	8.84
2023	8	1.2	9.42
2024	9	2.0	10.00
2025	?	3.0	10.58

57

© Data Science Institute

EG

市場規模を「従属変数」、レジャー支出、トレンドを「説明変数」に設定する。

データ
モデル
統計量
グラフ
予測値
タイトル
プロパティ

データ

データソース: D:\Workshop\WCUPF1\JWEG\41事例1.sas7bdat
タスクフィルタ: なし

変数リスト(A):

名前

- 年
- 市場規模
- レジャー支出
- トレンド

タスクの役割(B):

- 従属変数 (選択の上限: 1)
- 説明変数
- レジャー支出
- トレンド
- グループ分析
- 度数カウント (選択の上限: 1)
- 重み (選択の上限: 1)

58

© Data Science Institute

Root MSE	1.00266	R2 乗	0.8447
従属変数の平均	7.09091	調整済み R2 乗	0.8053
変動係数	14.14004		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	-3.07941	1.70756	-1.80	0.1090
レジャー支出	1	0.51297	0.20180	2.54	0.0346
トレンド	1	1.11071	0.16996	6.54	0.0002

◇自由度調整済み決定係数 0.8053

$$\begin{aligned}\text{予測値} &= -3.076 + 0.513 \times 3.0 + 1.111 \times 10.58 \\ &= 10.22\end{aligned}$$

事例2

地域別男性用化粧品市場規模と男性人口

地域	市場規模	男性人口
1	130	93
2	290	234
3	235	250
4	260	260
5	140	119
6	173	180
7	135	151
8	190	192
9	220	273
10	181	185

Root MSE	28.91642	R2 乗	0.7525
従属変数の平均	195.40000	調整済み R2 乗	0.7216
変動係数	14.79858		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	43.99116	32.03230	1.37	0.2069
男性人口	1	0.78167	0.15849	4.93	0.0011

市場規模 \leftarrow 男性人口

$$\text{市場規模} = 43.99 + 0.782 \times \text{男性人口}$$

◇自由度調整済み決定係数 0.7216

◇偏回帰係数の有意確率

男性人口 0.0011 < 0.05

説明変数に「ホワイトカラー人口」を追加

地域	市場規模	男性人口	ホワイトカラー人口
1	130	93	150
2	290	234	311
3	235	250	182
4	260	260	245
5	140	119	149
6	173	180	160
7	135	151	98
8	190	192	180
9	220	273	113
10	181	185	105

*「ホワイトカラー」
(専門的・技術的職業、管理的職業、事務、販売従事者)

EG

Root MSE	10.04926	R2 乗	0.9738
従属変数の平均	195.40000	調整済み R2 乗	0.9664
変動係数	5.14292		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	7.15385	12.11739	0.59	0.5735
男性人口	1	0.60153	0.06985	10.05	<.0001
ホワイトカラー	1	0.42368	0.06505	7.70	0.0001

市場規模 \Leftarrow 男性人口、ホワイトカラー人口

◇自由度調整済み決定係数 : 0.9663

0.7216 \Rightarrow 0.9663

◇偏回帰係数の有意確率

男性人口 0.0000 $<$ 0.05

ホワイトカラー人口 0.0001 $<$ 0.05

65

© Data Science Institute

説明変数に地域所得をさらに追加

地域	市場規模	男性人口	ホワイトカラー人口	地域所得
1	130	93	150	143
2	290	234	311	284
3	235	250	182	320
4	260	260	245	302
5	140	119	149	182
6	173	180	160	225
7	135	151	98	190
8	190	192	180	242
9	220	273	113	320
10	181	185	105	235

66

© Data Science Institute

Root MSE	10.65938	R2 乗	0.9748
従属変数の平均	195.40000	調整済み R2 乗	0.9622
変動係数	5.45516		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	-1.82664	23.00272	-0.08	0.9393
男性人口	1	0.42311	0.38429	1.10	0.3131
ホワイトカラー	1	0.42144	0.06858	7.19	0.0004
地域所得	1	0.17978	0.38190	0.47	0.6544

市場規模 \Leftarrow 男性人口、ホワイトカラー人口、地域所得

◇自由度調整済み決定係数 : 0.9622

◇偏回帰係数の有意確率

男性人口 0.3131 > 0.05

ホワイトカラー人口 0.0004 < 0.05

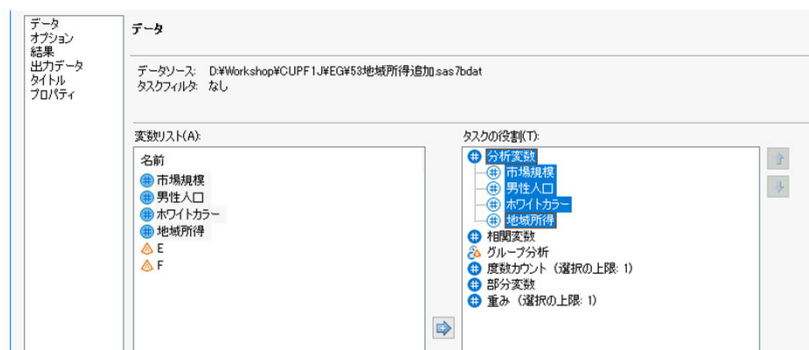
地域所得 0.6544 > 0.05

3つの説明変数で説明するモデルは、不成立

「多変量解析」－「相関分析」を選択する。

- 線形回帰分析
- 非線形回帰分析
- > 管理図
- > 記述統計
- > 工程能力分析
- > 時系列分析
- > 生存時間分析
- ▼ 多変量解析
 - クラスター分析
 - 因子分析
 - 主成分分析
 - 正準相関分析
 - 相関分析
 - 判別分析
- > 分散分析

すべての変数を「分析変数」に設定する。



相関行列が表示される。

Pearson の相関係数 N=10 H0: Rho=0 に対する Prob > r				
	市場規模	男性人口	ホワイトカラー	地域所得
市場規模	1.00000	0.86747 0.0011	0.77224 0.0089	0.86784 0.0011
男性人口	0.86747 0.0011	1.00000	0.39108 0.2638	0.98837 <.0001
ホワイトカラー	0.77224 0.0089	0.39108 0.2638	1.00000	0.39792 0.2548
地域所得	0.86784 0.0011	0.98837 <.0001	0.39792 0.2548	1.00000

各変数間の相関係数の検討

◇従属変数⇔説明変数

市場規模	男性人口	0.867
	ホワイトカラー人口	0.722
	地域所得	0.868

いずれも高い値⇒売上高を説明する説明変数として妥当

◇説明変数間の相関係数

男性人口	：	ホワイトカラー人口	0.391
男性人口	：	地域所得	0.988
ホワイトカラー人口	：	地域所得	0.398

男性人口と地域所得の値 0.988は高い!

説明変数相互の相関係数は高くない方が良い!

説明変数 ⇒ 独立変数

事例3

広告費と売上高

	2022		2023		2024	
	売上高	広告費	売上高	広告費	売上高	広告費
1月	78	9	84	20	120	18
2月	138	23	162	24	168	24
3月	144	25	150	20	138	24
4月	156	24	162	26	174	25
5月	84	12	90	13	96	13
6月	102	18	96	19	96	19
7月	84	8	90	14	132	19
8月	132	19	114	22	120	20
9月	132	22	138	29	138	29
10月	108	19	108	17	108	16
11月	102	18	114	19	108	24
12月	96	17	84	15	102	15

73

© Data Science Institute

EG

売上高 \Leftarrow 広告費

Root MSE	16.89396	R2 乗	0.6235
従属変数の平均	118.00000	調整済み R2 乗	0.6125
変動係数	14.31692		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	35.90780	11.29601	3.18	0.0031
広告費	1	4.23398	0.56421	7.50	<.0001

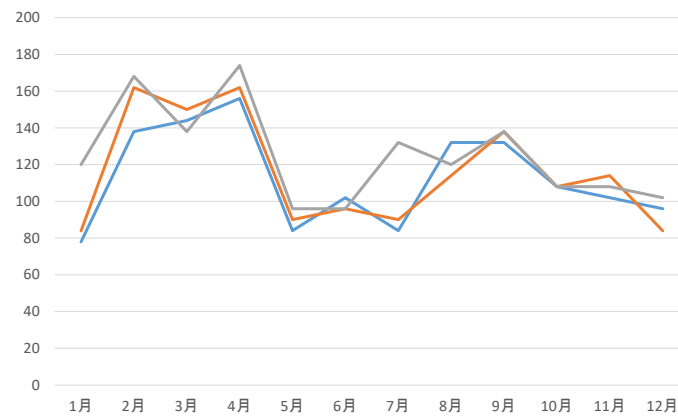
$$\text{売上高} = 35.9 + 4.234 \times \text{広告費}$$

自由度調整済み決定係数 0.6125

74

© Data Science Institute

季節変動の検討



季節変動（季節指数）を説明変数に追加

75

© Data Science Institute

季節指数（月別平均法）の算出

- 1) 月別に平均を求める。
- 2) 月別平均の全平均を求める。

月	2022	2023	2024	平均
1月	78	84	120	94
2月	138	162	168	156
3月	144	150	138	144
4月	156	162	174	164
5月	84	90	96	90
6月	102	96	96	98
7月	84	90	132	102
8月	132	114	120	122
9月	132	138	138	136
10月	108	108	108	108
11月	102	114	108	108
12月	96	84	102	94

平均 118

76

© Data Science Institute

3) 各月ごとに月別平均を全平均で割る。

(例 1月 : $94/118=0.80$)

月	2022	2023	2024	平均	季節指数
1月	78	84	120	94	0.80
2月	138	162	168	156	1.32
3月	144	150	138	144	1.22
4月	156	162	174	164	1.39
5月	84	90	96	90	0.76
6月	102	96	96	98	0.83
7月	84	90	132	102	0.86
8月	132	114	120	122	1.03
9月	132	138	138	136	1.15
10月	108	108	108	108	0.92
11月	102	114	108	108	0.92
12月	96	84	102	94	0.80

売上高 \Leftarrow 広告費、季節指数

回帰分析 (SAS EG)

1. データを入力する。

季節指数は3年間
同じ値を入力する。

	④ 売上高	⑤ 広告費	⑥ 季節指数
1	78	9	0.80
2	138	23	1.32
3	144	25	1.22
4	156	24	1.39
5	84	12	0.76
6	102	18	0.83
7	84	8	0.86
8	132	19	1.03
9	132	22	1.15
10	108	19	0.92
11	102	18	0.92
12	96	17	0.80
13	84	20	0.80
14	162	24	1.32
15	150	20	1.22
16	162	26	1.39
17	90	13	0.76
18	96	19	0.83
19	90	14	0.86
20	114	22	1.03
21	138	29	1.15
22	108	17	0.92
23	114	19	0.92
24	84	15	0.80
25	120	18	0.80

EG

季節指数を説明変数に追加

Root MSE	9.92735	R2 乗	0.8738
従属変数の平均	118.00000	調整済み R2 乗	0.8662
変動係数	8.41301		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	-1.31267	8.07608	-0.16	0.8719
広告費	1	1.06003	0.51362	2.06	0.0470
季節指数	1	98.75981	12.20621	8.09	<.0001

売上高 = $-1.313 + 1.060 \times \text{広告費} + 98.760 \times \text{季節指数}$

◇自由度調整済み決定係数 0.8662

0.6125 \Rightarrow 0.8662

事例4

年度別広告費と売上高

年度	売上高	広告費
2016	5,123	623
2017	4,986	612
2018	4,867	543
2019	4,421	463
2020	2,985	427
2021	3,123	441
2022	4,422	513
2023	4,123	523
2024	4,256	532

売上高 \Leftarrow 広告費

Root MSE	391.81596	R2 乗	0.7679
従属変数の平均	4256.22222	調整済み R2 乗	0.7347
変動係数	9.20572		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	-773.87865	1063.44180	-0.73	0.4864
広告費	1	9.67948	2.01151	4.81	0.0019

$$\text{売上高} = -773.9 + 9.679 \times \text{広告費}$$

自由度調整済み決定係数 0.7347

説明変数にダミー変数を追加

年度	売上高	広告費	ダミー
2016	5,123	623	0
2017	4,986	612	0
2018	4,867	543	0
2019	4,421	463	0
2020	2,985	427	1
2021	3,123	441	1
2022	4,422	513	0
2023	4,123	523	0
2024	4,256	532	0

売上高 \Leftarrow 広告費、復興ダミー

Root MSE	245.90361	F2 乗	0.9216
従属変数の平均	4256.22222	調整済み F2 乗	0.8955
変動係数	5.77751		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	1677.21682	973.37820	1.72	0.1356
広告費	1	5.37083	1.78066	3.02	0.0235
ダミー	1	-954.15599	278.09754	-3.43	0.0140

$$\text{売上高} = 1677.2 + 5.371 \times \text{広告費} - 954.16 \times \text{ダミー}$$

自由度調整済み決定係数 0.8955

0.7347 \Rightarrow 0.8955

* 広告費550のときの予測値

$$1677.2 + 5.371 \times 550 - 954.16 \times 0 = 4631.3$$

ダミー変数の作成方法

(2区分)	ダミー-1
有	1
無	0

区分数 - 1 \Rightarrow ダミー変数

(3区分)	ダミー-1	ダミー-2
大	1	0
中	0	1
小	0	0

(4区分)	ダミー-1	ダミー-2	ダミー-3
20歳代	1	0	0
30歳代	0	1	0
40歳代	0	0	1
50歳代	0	0	0

売上個数 \Leftarrow 曜日、温度

売上個数	曜日	温度
356	月	28
245	火	21
128	水	15
189	木	15
215	金	28
412	土	24
388	日	22
312	月	19
301	火	22
355	水	25

曜日 : 7区分



ダミー変数 : 6個

85

© Data Science Institute

入力データ

売上個数	ダミー-1	ダミー-2	ダミー-3	ダミー-4	ダミー-5	ダミー-6	温度
356	1	0	0	0	0	0	28
245	0	1	0	0	0	0	21
128	0	0	1	0	0	0	15
189	0	0	0	1	0	0	15
215	0	0	0	0	1	0	28
412	0	0	0	0	0	1	24
388	0	0	0	0	0	0	22
312	1	0	0	0	0	0	19
301	0	1	0	0	0	0	22
355	0	0	1	0	0	0	25

86

© Data Science Institute

◇広告費と売上高（3区分）

年度	売上高	広告費	ダミー1	ダミー2
2016	5,123	623	0	0
2017	4,986	612	0	0
2018	4,867	543	0	0
2019	4,421	463	0	0
2020	2,985	427	1	0
2021	3,523	441	0	1
2022	4,422	513	0	0
2023	4,123	523	0	0
2024	4,256	532	0	0

87

© Data Science Institute

EG

売上高 \Leftarrow 広告費、ダミー1、ダミー2

Root MSE	268.63688	R2 乗	0.9221
従属変数の平均	4256.22222	調整済み R2 乗	0.8753
変動係数	6.31163		

パラメータの推定					
変数	自由度	パラメータ 推定値	標準誤差	t 値	Pr > t
Intercept	1	1689.82626	1066.08275	1.59	0.1738
広告費	1	5.34765	1.95029	2.74	0.0407
ダミー1	1	-988.27475	366.97455	-2.69	0.0431
ダミー2	1	-925.14192	350.62789	-2.64	0.0461

$$\text{売上高} = 1689.8 + 5.348 \times \text{広告費} - 988.3 \times \text{ダミー1} - 925.1 \times \text{ダミー2}$$

88

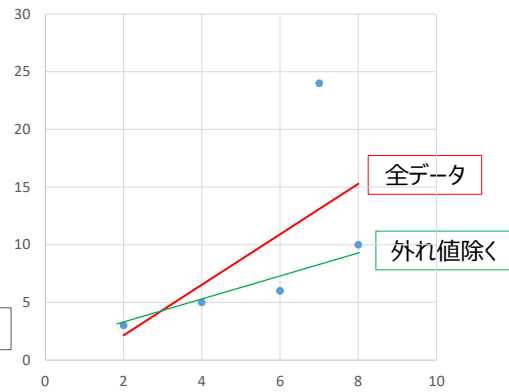
© Data Science Institute

外れ値を含む場合

売上高	広告費
3	2
5	4
6	6
10	8
24	7

$$\text{売上高} = -2.22 + 2.19 \times \text{広告費}$$

$$\text{売上高} = 0.5 + 1.1 \times \text{広告費}$$



回帰直線は外れ値の影響を受ける。

89

© Data Science Institute

ダミー変数の設定

売上高	広告費	ダミー
3	2	0
5	4	0
6	6	0
10	8	0
24	7	1

Root MSE	0.94868	R2 垂	0.9937
従属変数の平均	9.60000	調整済み R2 垂	0.9874
変動係数	9.88212		

パラメータの推定					
変数	自由度	パラメータ推定値	標準誤差	t 値	Pr > t
Intercept	1	0.50000	1.16190	0.43	0.7089
広告費	1	1.10000	0.21213	5.19	0.0052
ダミー変数	1	15.80000	1.14237	13.83	0.00062

$$\text{売上高} = 0.5 + 1.1 \times \text{広告費} + 15.8 \times \text{ダミー}$$

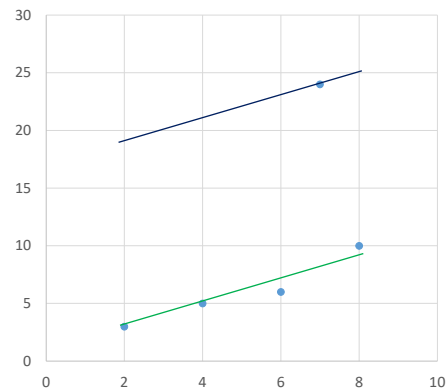
自由度調整済み決定係数 0.9874

EG

90

© Data Science Institute

ダミー : 0 売上高 = $0.5 + 1.1 \times \text{広告費}$
 ダミー : 1 売上高 = $16.3 + 1.1 \times \text{広告費}$



91

© Data Science Institute

まとめ

- ◇ グラフにおける対数目盛の活用
- ◇ 指数平滑法
- ◇ 時系列分析
 - ・ 移動平均法
 - ・ 季節指数
 - ・ 傾向線のあてはめ
- ◇ 回帰分析
 - ・ 重回帰
 - ・ トレンドを説明変数に追加
 - ・ 季節指数を説明変数に追加
 - ・ ダミー変数の活用

92

© Data Science Institute

SASで実践！実務で使える販売・需要予測

2025年2月26日 初版 第1刷

発行元 SAS Institute Japan株式会社

〒106-6661 東京都港区六本木6-10-1 六本木ヒルズ森タワー11F
TEL 03(6434)3690

本書内容の一部、全体を問わず、SAS Institute Japan株式会社の文書による承諾なく引用複製することを禁じます。