

<u>社会人のためのデータサイエンス演習</u>

第3週:課題の補講 Excel分析演習

講師名:菅 由紀子

第3週の補講の内容紹介



※Microsoft[®], Microsoft[®] Office Excel[®]は、米国 Microsoft Corporation の, 米国およびその他の国における登録商標または商標です。 ※本資料は、Microsoft Corporation と提携しているものではなく、また、Microsoft Corporationが許諾、後援、その他の承認をするものでは ありません。

※本資料の本文では、©、®、™などの表記は割愛いたします。

※本資料ではデータ分析ツールとして Microsoft® Office Excel® 2013 を利用しています。

Microsoft® Office の他のバージョンや他の分析ツールを利用している場合は、ヘルプやインターネットなどで各自で調査し、該当機能に 置き換えて参照してください。

第3週のまとめ

比較と傾向の分析 クロス集計 / 散布図 / 相関 / 時系列

上段:n 下段:%	先月	今月	計			
ファミリー層	200,000円	100,000円	300,000円			
	67%	33%	100%			
単身層	180,000円	180,000円	360,000円			
	50%	50%	100%			
計	380,000円	280,000円	660,000円			
	58%	42%	100%			



✓ 変数の関係性把握✓ 外れ値の除外



✓ 解釈が容易に✓ 将来予測が可能







- 下表のような売上を示すデータがある
- クロス集計を行い、4月と5月で、ファミリーと単身者の売上に違いがあるか観察する

	Α	В	С
1	家族構成	売上月	売上
2	ファミリー	4月	¥15,000
3	ファミリー	4月	¥20,000
4	単身	4月	¥12,000
5	ファミリー	4月	¥18,000
6	単身	4月	¥9,000
7	単身	4月	¥5,000
8	単身	4月	¥10,000
9	ファミリー	4月	¥17,000
10	単身	4月	¥15,000
11	ファミリー	5月	¥12,000
12	単身	5月	¥15,000
13	ファミリー	5月	¥13,000
14	ファミリー	5月	¥12,000
15	単身	5月	¥16,000
10			

1)Excelのピボットテーブルという機能を使って、表頭に売上月、表側に家族構成を表示する集計表を作成



2) 集計軸を設定し、各セルの合計値を算出 Excelの右側に表示されるピボットテーブルのフィールドから、上部フィールドをドラッグして下部に追加し、 以下のように設定



⑥上部の「家族構成」を「行」に追加⇒ クロス表の表側になる

3)続いて、売上の値フィールドから計算の種類を選択して割合が表示されるように設定





4) クロス集計表が出力



- 全体の売上では5月に低下
- ファミリーと単身の総計では単身の方が売上が低い
- ファミリー、単身ともに、ほぼ同様の割合で5月の売上が低下

散布図の描画

● 水稲の作付面積と収穫量の関係を、散布図を描いて観察してみましょう。

1)Excelで、右のような表を作成。 X軸に示したい項目のデータをB列、Y軸に示した い項目のデータをC列に設定

	А	В	С	D
1	年次	作付面積	収穫量(千	-t)
2	1990年	2055	10463	
3	1991年	2033	9565	
4	1992年	2092	10546	
5	1993年	2127	7811	
6	1994年	2200	11961	
7	1995年	2106	10724	
8	1996年	1967	10328	
9	1997年	1944	10004	
10	1998年	1793	8939	
11	1999年	1780	9159	
12	2000年	1763	9472	
13	2001年	1700	9048	
14	2002年	1683	8876	
15	2003年	1660	7779	
16	2004年	1697	8721	

出典:農林水産省 作物統計より作成

http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&tclassID=000001024932&cycleCode=0&requestSender=dsearch

散布図の描画

2)ツールバーの「挿入」をクリックし、データの範囲となるセルB2からセルC16までを選択して、 グラフツールの散布図を選択

フ	アイル ジ	ホーム 挿 入		①挿	入」をク	リック	表	示開発	ヘルプ		
ビオ テー	ううしょう しょう しょう しょう しょう しょう しょう しょう しょう し	■ すすめ トテーブル	▲ ●像 ◆ ⑦	図形 ~ アイコン 3D モデル ~	Pa Smart る スクリーン	Art ンショット 〜	日アドインを の個人用ア	入手 🔮 ドイン Y 🛃	ting the second		2 プ ピボットグラフ ~
B2	τ-	-),, - : X	$\sqrt{f_x}$	2055	l		ፖቦ	12	1		③散布図をクリック
	A	В	С	D	Е	F	G	Н			
1	年次	作付面積	収穫量(千	t)							
2	1990年	2055	10463							I I I I	
3	1991年	2033	9565							バブル	
4	1992年	2092	10546								④左上の図形を選択
5	1993年	2127	7811								
6	1994年	2200	11961								-
7	1995年	2106	10724							き その他の散布	图(<u>M</u>)
8	1996年	1967	10328								
9	1997年	1944	10004								
10	1998年	1793	0150								
11	2000年	1763	9159		(2) ,	ータの賃	範囲				
13	2000年	1703	9048		(P ²		(c)				
14	2002年	1683	8876		(B2	נטפיתי	[0]				
15	2003年	1660	7779		を選打	尺					
16	2004年	1697	8721	J							



3)グラフ描画



※上図は結果が見やすいように軸の目盛りを変更している

外れ値が存在

右肩上がりの形状 ⇒正の相関

相関係数の算出

● 次に、相関係数を算出してみましょう。

1) 散布図を描いたデータと同じデータを用意

	А	В	С	D
1	年次	作付面積	収穫量(千	-t)
2	1990年	2055	10463	
3	1991年	2033	9565	
4	1992年	2092	10546	
5	1993年	2127	7811	
6	1994年	2200	11961	
7	1995年	2106	10724	
8	1996年	1967	10328	
9	1997年	1944	10004	
10	1998年	1793	8939	
11	1999年	1780	9159	
12	2000年	1763	9472	
13	2001年	1700	9048	
14	2002年	1683	8876	
15	2003年	1660	7779	
16	2004年	1697	8721	

出典:農林水産省 作物統計より作成

http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_toGL08020103_&tclassID=000001024932&cycleCode=0&requestSender=dsearch

相関係数の算出

1)データ分析ツールを起動



2)ダイアログが表示されたら[相関] を選択



3)[入力範囲]にデータ範囲であるセルB1~C16を選択



相関係数の算出 [結果]

4)相関係数が出力



【補足1】 分析ツールアドインではなく、関数で 算出することも可能

=CORREL(変量1、変量2)

【補足2】 分析ツールアドインでは、3つ以上の 変数を対象に、それぞれの相関係数 を一度に算出可能



移動平均の算出

 下の図表は自動車教習料の支出額を時系列に示したものです。12ヶ月 移動平均値を算出して季節変動を除いたグラフを作成してみましょう。

	年月	支出額	
	平成14年1月	1,405	
	平成14年2月	1,563	
	平成14年3月	1,198	
	平成14年4月	617	
	平成14年5月	720	
	平成14年6月	560	
	平成14年7月	1,091	
	平成14年8月	1,369	
	平成14年9月	655	
	<u>平成14年10月</u>	631	
/		\land	
· · .			\sim
1		1 1 1 4	\geq
	平成22年2月	1,114	
/	平成22年2月 平成22年3月	1,114 1,136	
	平成22年2月 平成22年3月 平成22年4月	1,114 1,136 438	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月	1,114 1,136 438 344	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月	1,114 1,136 438 344 410	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月 平成22年7月	1,114 1,136 438 344 410 724	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月 平成22年7月 平成22年8月	1,114 1,136 438 344 410 724 1,016	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月 平成22年7月 平成22年8月 平成22年8月 平成22年9月	1,114 1,136 438 344 410 724 1,016 543	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月 平成22年7月 平成22年8月 平成22年8月 平成22年9月 平成22年10月	1,114 1,136 438 344 410 724 1,016 543 395	
	平成22年2月 平成22年3月 平成22年4月 平成22年5月 平成22年6月 平成22年7月 平成22年8月 平成22年9月 平成22年10月 平成22年11月	1,114 1,136 438 344 410 724 1,016 543 395 450	



出典:総務省 家計消費状況調査(二人以上の世帯)より作成 http://www.stat.go.jp/data/joukyou/12.htm

移動平均の算出



2)ダイアログが表示されたら[移動平均] を選択

移動平均の算出 [結果]

4)12か月移動平均が算出される。

	А	В		С	
1	年月	支出額		12ヶ月移動平均	
2	平成14年1月	1,40)5		
3	平成14年2月	1,56	33		
4	平成14年3月	1,19	98		
5	平成14年4月	61	7		
6	平成14年5月	72	20		
7	平成14年6月	56	60		
8	平成14年7月	1,09	91		
9	平成14年8月	1,36	39		
10	平成14年9月	65	55		
11	平成14年10月	63	31		
12	平成14年11月	8	14		
13	平成14年12月	8	-1	955	
14	平成15年1月	1,2	5	940	
15	平成15年2月	1,3	0	918	
16	平成15年3月	1,1	17	918	
17	平成15年4月	7:	9	928	
18	平成15年5月	7.	4	930	
19	平成15年6月	5	3	931	
20	平成15年7月	8	2	913	
21	平成15年8月	1,3	0	910	
/		\frown			
/			_		
107	平成22年10月	3	15	666	
108	平成22年11月	4.	0	671	
109	平成22年12月	6	9	673	

※分析ツールで求められるのは後方移動平均。 ※各月とそれ以前の月の平均値を使って算出する移動平均を 「後方移動平均」という。



移動平均の算出 [補足]

【補足】分析ツールアドインではなく、関数で算出することも可能

		A	В		C		D		D	
	1	年月	支出額	1:	12ヶ月移動 平均			3動 12ヶ月中/i 移動平 ⁱ		
	2	平成14年1月	1,405				ן			
	3	平成14年2月	1,563					ך		
	4	平成14年3月	1,198							
	5	平成14年4月	617							
	6	平成14年5月	720						\mathbf{O}	
	7	平成14年6月	560				l			
	8	平成14年7月	1,091			\downarrow	⊥		947	
	9	平成14年8月	1,369					Ĺ	929	
	10	平成14年9月	655						918	
	11	平成14年10月	631						923	
	12	平成14年11月	804						929	
	13	平成14年12月	841		95	5	3		930	
	14	平成15年1月	1,235		94	0		J	922	
	15	平成15年2月	1,300		91	8	911			
	16	平成15年3月	1,197		91	8	91-		914	
_	17			_	\smile	_	\subseteq	_		_
	100			_						_
	101	平成22年4月	438		680				668	
	102	平成22年5月	344		678	8			668	
	103	平成22年6月	410		679	9			672	
	104	半成22年7月	724		688					
	105	半成22年8月	1,016		680					
	106	半成22年9月	543		671	1				
	107	平成22年10月	395		666	6				
	108	平成22年11月	450		671	1				
	109	平成22年12月	659		673	3				

①12ヶ月移動平均(後方移動平均) 各月における過去12ヶ月間の平均値

C1312, = AVERAGE(B2:B13) C1412, = AVERAGE(B3:B14)

.

ο

C109に、=AVERAGE(B98:B109)

<u>②12ヶ月中心化移動平均</u>

各月を中心にした12ヶ月間の平均値。月数は偶数のため、 区間を1ヶ月ずらした平均値を2つ求め、両者の平均を取る

C8に、=(AVERAGE(B2:B13)+ AVERAGE(B3:B14))/2 C9に、=(AVERAGE(B3:B14)+ AVERAGE(B4:B15))/2

C103に、=(AVERAGE(B97:B108)+ AVERAGE(B98:B109))/2