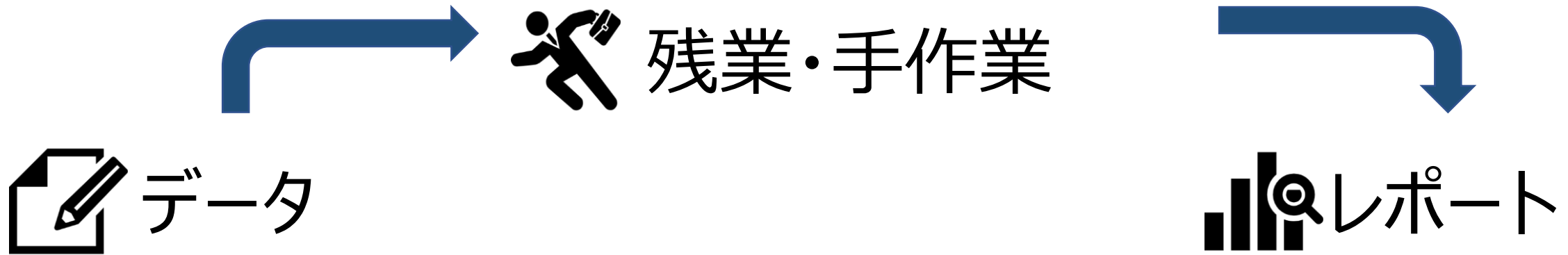


データ活用の実践

Excelのピボットテーブルを使いこなそう

日本電産株式会社 人事部
産業医 西田 典充
ma06063@gmail.com

データから何を語りたい？



データの集計業務は多くの産業保健職にとって負担
データの加工の効率化は
研究だけでなく、実務にも大きなメリット

効率的なデータ加工を見てみましょう



この動画では、データを効率的に加工する実例を提示して最初の1歩のイメージを固めます

状況設定

- あなたは産業看護職です
- 健康診断の受診結果のデータが労働衛生機関より納入されました
- 会社から傾向をまとめて報告してほしいと依頼がありました
- この動画では、「BMI25以上の人の割合」の集計を自動化してみます



まずはデータの持ち方を見てみます



データの記録方法

社員番号	受診年月日	受診日年齢	性別	BMI
1	2022/11/10	21	男性	21
2	2022/11/10	22	女性	19
3	2022/11/10	23	男性	26
...

このような健康診断データが納品されてきたとします
毎年、同じ形で新しいデータを手に入れる場合に、
あなたならどのような形で保存しますか？
動画を止めて考えてみてください

✕これは避けよう

社員 番号	受診 年月日	受診日 年齢	性別	BMI
1	2022/11/10	21	男性	21
2	2022/11/10	22	女性	19
3	2022/11/10	23	男性	26
...

ファイル ホーム 挿入 描画 ページレイアウト 数式 データ						
B3		×	✓	fx	2022/11/10	
	A	B	C	D	E	
1	社員番号	受診年月日	受診日年齢	性別	BMI	
2	1	2022/11/10	21	男性	21	
3	2	2022/11/10	22	女性	19	
4	3	2022/11/10	23	男性	26	
5						
		2019	2020	2021	2022	+
準備完了 アクセシビリティ: 問題ありません						

シート名に「データ」を含ませるのは非推奨です
(データをまとめて分析する際に、大きな手間が発生します)

✕これは避けよう

社員 番号	受診 年月日	受診日 年齢	性別	BMI
1	2022/11/10	21	男性	21
2	2022/11/10	22	女性	19
3	2022/11/10	23	男性	26
...

	A	B	C	D	E	F
1	社員番号	性別	年齢_2021	年齢_2022	BMI_2021	BMI_2022
2		1 男性	20	21	22	21
3		2 女性	21	22	18	19
4		3 男性	22	23	28	26

変数名に「データ」を含ませることは非推奨です
(よく見られる形ですが、記録と利用に手間が発生します)

○今回はこんな感じで！

社員番号	受診年月日	受診日年齢	性別	BMI
1	2022/11/10	21	男性	21
2	2022/11/10	22	女性	19
3	2022/11/10	23	男性	26
...

A	B	C	D	E	F
年度	社員番号	受診年月日	受診日年齢	性別	BMI
2021	①	2021/11/1	20	男性	22
2021	2	2021/11/1	21	女性	18
2021	3	2021/11/1	22	男性	28
2022	①	2022/11/10	21	男性	21
2022	2	2022/11/10	22	女性	19
2022	3	2022/11/10	23	男性	26

データを**縦**方向に伸ばしていきましょう

ワンポイント：データの保存の方法

データは特別な理由がない限り、
1枚のシートに保存しましょう

データを追加したときに**横に伸びる**場合、
そのデータは**手間がかかる**データとなります
縦に伸びる形で保存しましょう

次は自動化（ピボットテーブル）です

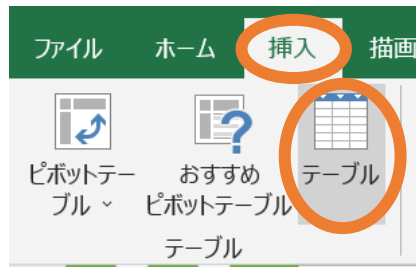


「BMI25以上の人の割合」
の集計を自動化

ここからはsampledata.xlsxを利用して解説します
ファイルは次の場所からダウンロード可能です



https://github.com/ironwest/2022_oh_data_analysis_video_samples



テーブル形式にすると作業が便利なので
挿入→テーブル
でデータをテーブルに

	A	B	C	D	E	F	G
1	yr	id	gender	age	bmi		
2	2017	1	男性	24	25.85247217		
3	2017	2	男性	24	25.97016442		
4	2017	3	男性	24	26.45460019		
5	2017	4	男性	24	17.49774458		
6	2017	5	男性	24	23.21597609		
7	2017	6	男性	24	22.11467568		
8	2017	7	男性	24	20.56100313		

F1セルをクリックしてagekubunと入力
F2セルに
=FLOOR([@age],10)
と入力

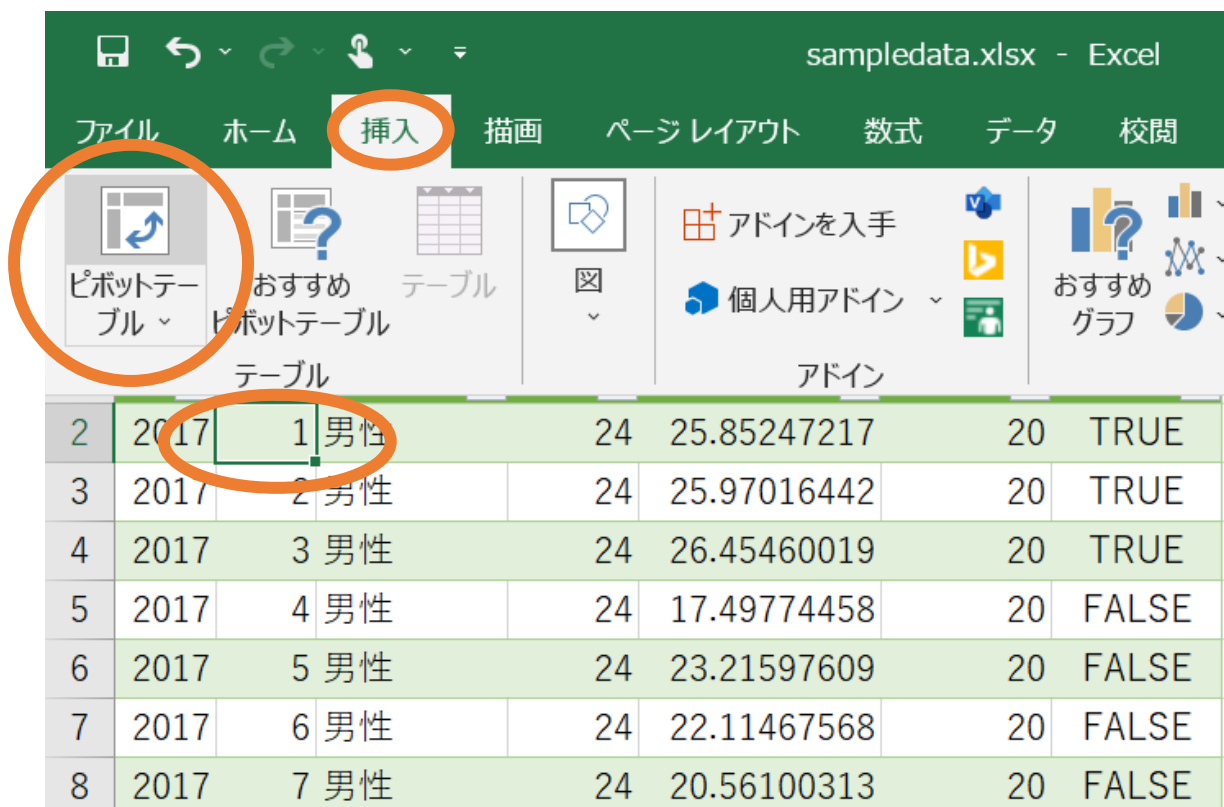
G1セルをクリックしてbmi25と入力
G2セルに
=[@bmi]>=25
と入力

	A	B	C	D	E	F	G
1	yr	id	gender	age	bmi		
2	2017	1	男性	24	25.85247217		
3	2017	2	男性	24	25.97016442		
4	2017	3	男性	24	26.45460019		
5	2017	4	男性	24	17.49774458		
6	2017	5	男性	24	23.21597609		



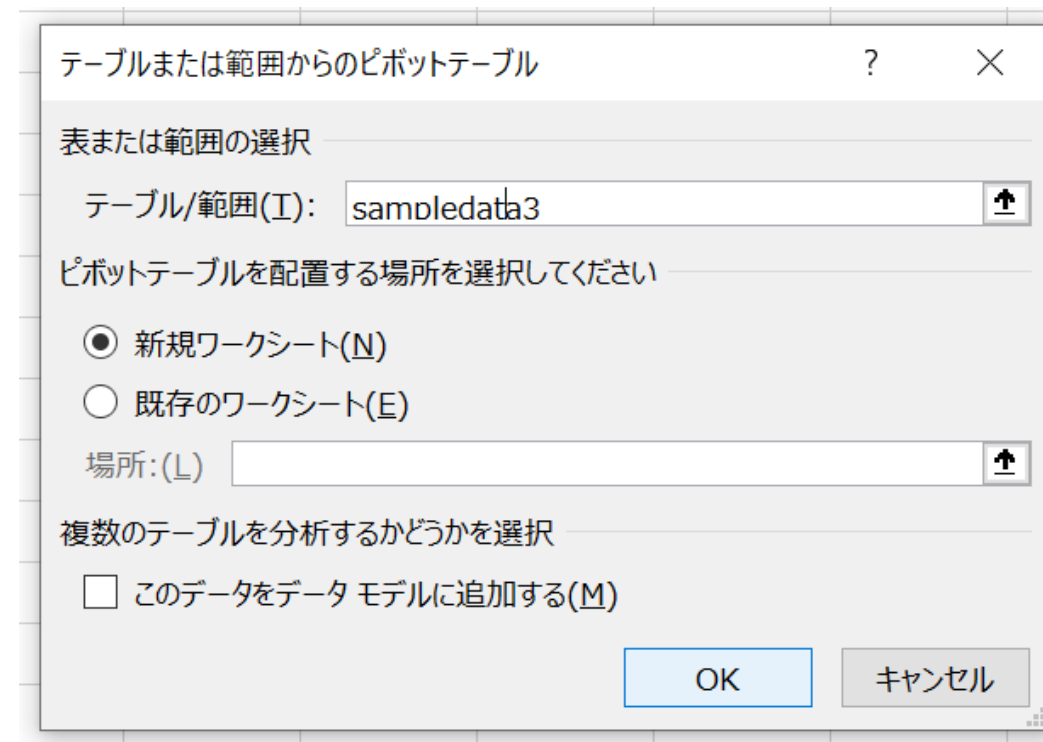
	A	B	C	D	E	F	G
1	yr	id	gender	age	bmi	agekut	bmi25
2	2017	1	男性	24	25.85247217	20	TRUE
3	2017	2	男性	24	25.97016442	20	TRUE
4	2017	3	男性	24	26.45460019	20	TRUE
5	2017	4	男性	24	17.49774458	20	FALSE
6	2017	5	男性	24	23.21597609	20	FALSE
7	2017	6	男性	24	22.11467568	20	FALSE
8	2017	7	男性	24	20.56100313	20	FALSE

テーブル形式なため、
自動的に列ごと関数が
適応されます



テーブル内のどこかを選択した状態で、
挿入→ピボットテーブルをクリックします

出てきたメニューでOKをクリックしましょう



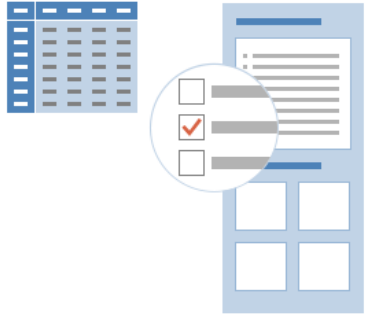
A3

✕ ✓ fx

	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						

ピボットテーブル8

レポートを作成するには、[ピボットテーブルのフィールド リスト] からフィールドを選択してください



ピボットテーブルのフィールド

レポートに追加するフィールドを選択してください:

検索

- ☒ yr
- ☐ id
- ☐ gender
- ☐ age
- ☐ bmi
- ☐ agekubun
- ☒ bmi25

その他のテーブル...

次のボックス間でフィールドをドラッグしてください:

フィルター	列
行	Σ 値

出てきたメニューでそれぞれ、
bmi25を行と値に、
yrを列にそれぞれドラッグします

▼ フィルター

≡ 行

bmi25 ▼

≡≡≡ 列

yr ▼

Σ 値

個数 / bmi... ▼

3	個数 / bmi25	列ラベル ▼						
4	行ラベル ▼		2017	2018	2019	2020	2021	総計
5	FALSE		169	170	159	156	154	808
6	TRUE		51	50	61	64	66	292
7	総計		220	220	220	220	220	1100
8								

そうすると、このような表がでてきます。
 これがピボットテーブルとよばれるものです。
 どのようにデータが入るかを考えてみましょう。

	A	B	C	D	E	F	G
1	yr ▾	id ▾	gender ▾	age ▾	bmi ▾	agekut ▾	bmi25 ▾
2	2017	1	男性	24	25.85247217	20	TRUE
3	2017	2	男性	24	25.97016442	20	TRUE
4	2017	3	男性	24	26.45460019	20	TRUE
5	2017	4	男性	24	17.49774458	20	FALSE

←Q1このデータはどこに行くでしょうか？

3	個数 / bmi25	列ラベル ▾					
4	行ラベル ▾	2017	2018	2019	2020	2021	総計
5	FALSE						
6	TRUE						
7	総計	220	220	220	220	220	1100
8							

	A	B	C	D	E	F	G
1	yr	id	gender	age	bmi	agekut	bmi25
2	2017	1	男性	24	25.85247217	20	TRUE
3	2017	2	男性	24	25.97016442	20	TRUE
4	2017	3	男性	24	26.45460019	20	TRUE
5	2017	4	男性	24	17.49774458	20	FALSE

←Q1こんな感じで1つとカウントされます

3	個数 / bmi25	列ラベル					
4	行ラベル		2017	2018	2019	2020	2021 総計
5	FALSE						
6	TRUE		1				
7	総計		220	220	220	220	1100
8							

	A	B	C	D	E	F	G
1	yr ▾	id ▾	gender ▾	age ▾	bmi ▾	agekut ▾	bmi25 ▾
2	2017	1	男性	24	25.85247217	20	TRUE
3	2017	2	男性	24	25.97016442	20	TRUE
4	2017	3	男性	24	26.45460019	20	TRUE
5	2017	4	男性	24	17.49774458	20	FALSE

←Q2ではこれは？

3	個数 / bmi25	列ラベル ▾					
4	行ラベル ▾		2017	2018	2019	2020	2021 総計
5	FALSE						
6	TRUE						
7	総計		220	220	220	220	1100
8							

	A	B	C	D	E	F	G
1	yr ▾	id ▾	gender ▾	age ▾	bmi ▾	agekut ▾	bmi25 ▾
2	2017	1	男性	24	25.85247217	20	TRUE
3	2017	2	男性	24	25.97016442	20	TRUE
4	2017	3	男性	24	26.45460019	20	TRUE
5	2017	4	男性	24	17.49774458	20	FALSE

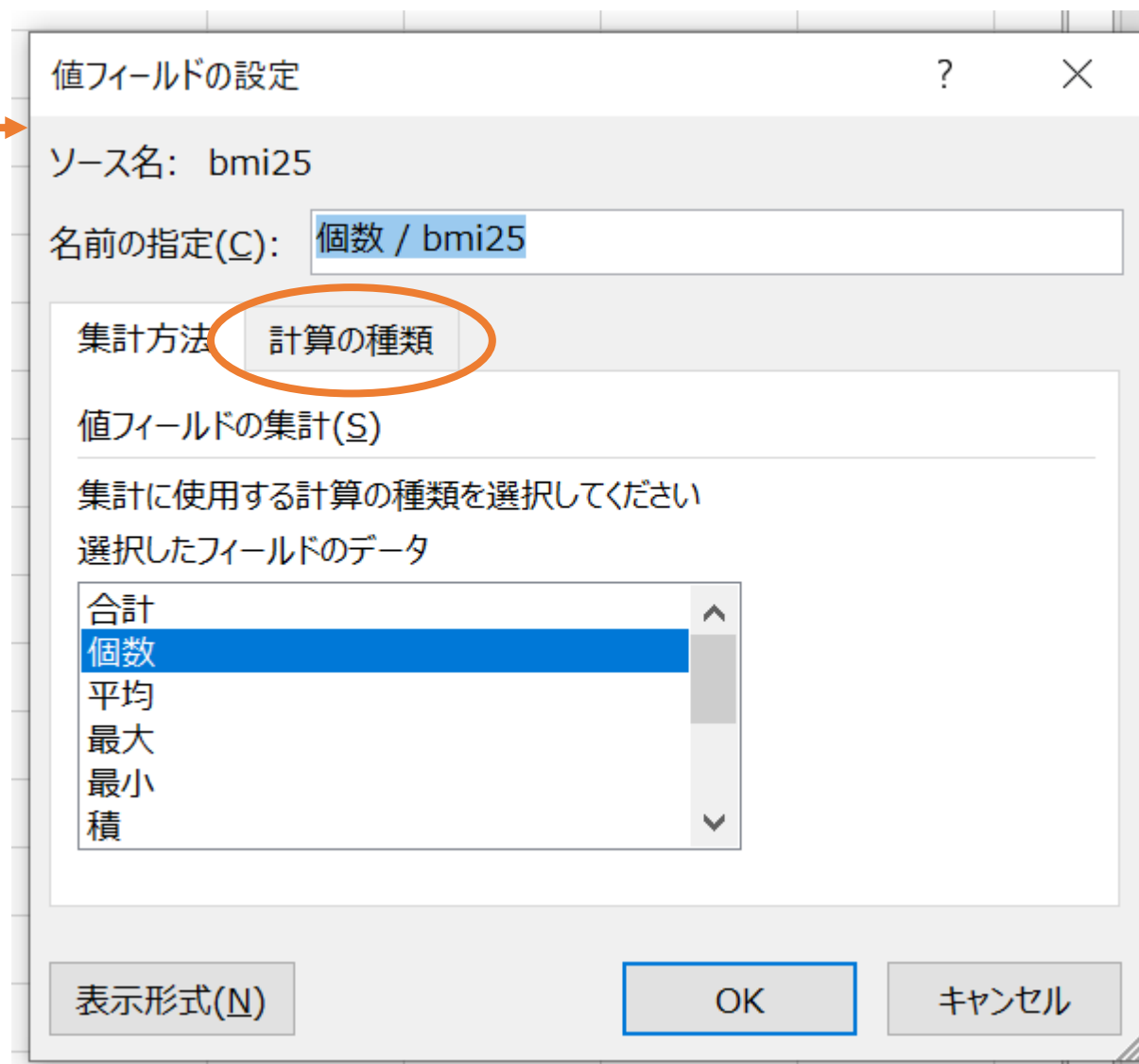
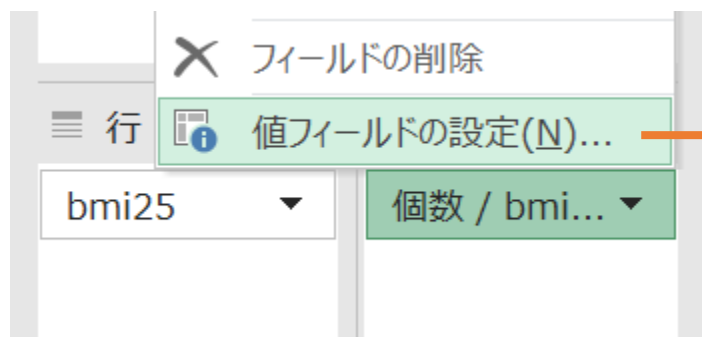
←Q2ではこれは？

3	個数 / bmi25	列ラベル ▾					
4	行ラベル ▾		2017	2018	2019	2020	2021 総計
5	FALSE		1				
6	TRUE						
7	総計		220	220	220	220	1100
8							

3	個数 / bmi25	列ラベル					
4	行ラベル		2017	2018	2019	2020	2021 総計
5	FALSE		169	170	159	156	154 808
6	TRUE		51	50	61	64	66 292
7	総計		220	220	220	220	220 1100
8							

ということで、**1100行**のデータを
ピボットテーブルが自動で集計してくれました

ただ、これだと人数はわかりますが、割合はわかりません
「BMIが25以上の従業員の割合」を調べたい
場合は次のように設定します



値フィールドの設定

ソース名: bmi25

名前の指定(C): 個数 / bmi25

集計方法 計算の種類

計算の種類(A)

計算なし

基準フィールド(F):

id
gender
age
bmi
agekubun
bmi25

基準アイテム(I):

表示形式(N)

OK

キャンセル

値フィールドの設定

ソース名: bmi25

名前の指定(C): 個数 / bmi25

集計方法 計算の種類

計算の種類(A)

計算なし

計算なし

総計に対する比率

列集計に対する比率

行集計に対する比率

基準値に対する比率

親行集計に対する比率

agekubun

bmi25

表示形式(N)

OK

キャンセル

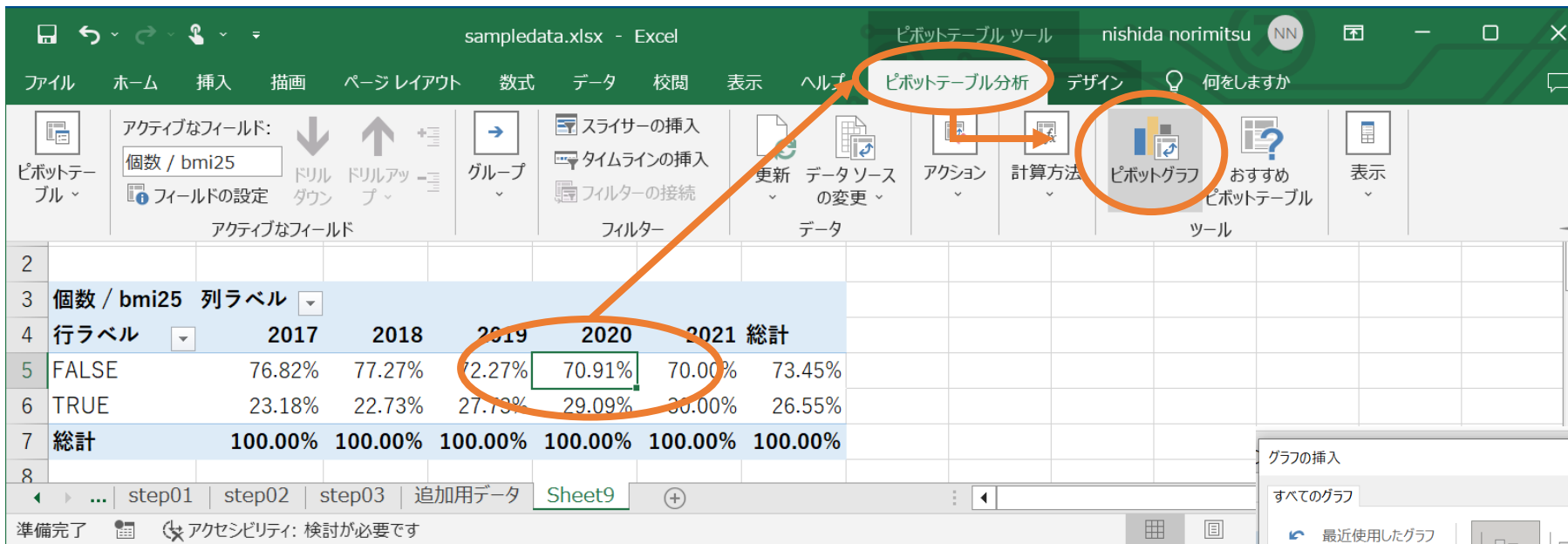
この手順で、列ごとの割合を求めることができました

グラフにしてみましょう

個数 / bmi25	列ラベル ▼					
行ラベル ▼	2017	2018	2019	2020	2021	総計
FALSE	169	170	159	156	154	808
TRUE	51	50	61	64	66	292
総計	220	220	220	220	220	1100

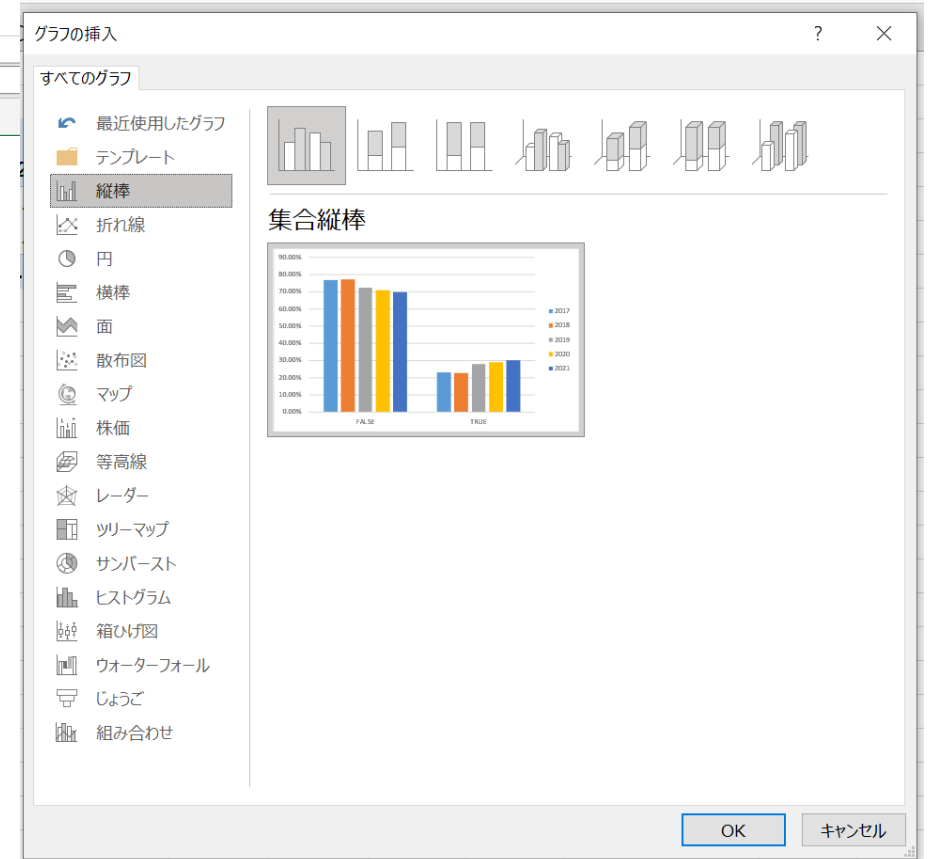


個数 / bmi25	列ラベル ▼					
行ラベル ▼	2017	2018	2019	2020	2021	総計
FALSE	76.82%	77.27%	72.27%	70.91%	70.00%	73.45%
TRUE	23.18%	22.73%	27.73%	29.09%	30.00%	26.55%
総計	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

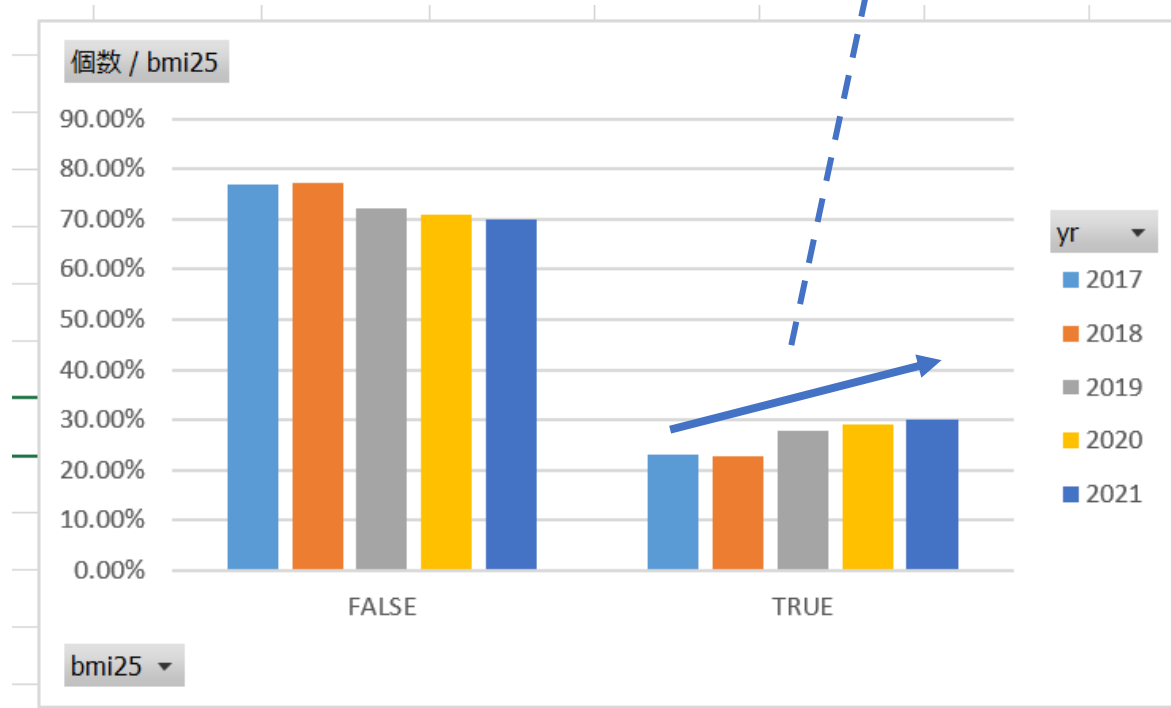


ピボットテーブルのエリアを選択します→
メニューのピボットテーブル分析を選択します→
ピボットグラフを選択します

でできたメニューから縦棒をえらび、OKをクリックします



できあがりました
どうやら、このデータからはBMI25以上の人の割合は
年々増加しているようです



それでは、
・性別別
・年代別
に分けて分析して、

**どの性別のどの年代でBMI25以上の
人の割合が増加しているか**

を調べることはできるでしょうか？
動画をとめて試みてください。

回答例：

▼ フィルター

凡例 (系列)

bmi25 ▼

≡ 軸 (分類項目)

gender ▼

agekubun ▼

yr ▼

Σ 値

個数 / bmi... ▼



表としては見づらいが、ピボットグラフを作成すると、

集計方法

計算の種類

計算の種類(A)

親行集計に対する比率

計算なし

総計に対する比率

列集計に対する比率

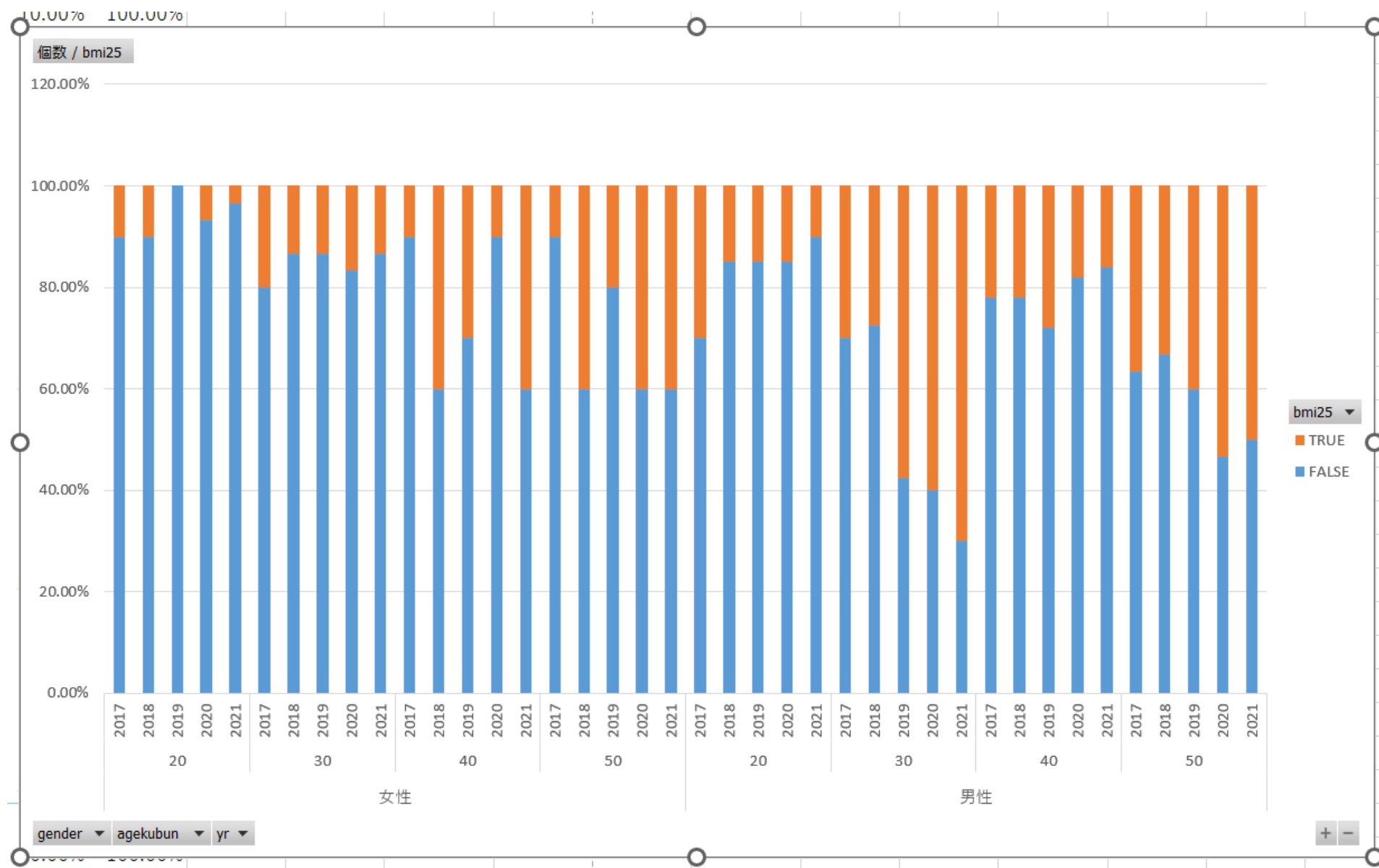
行集計に対する比率

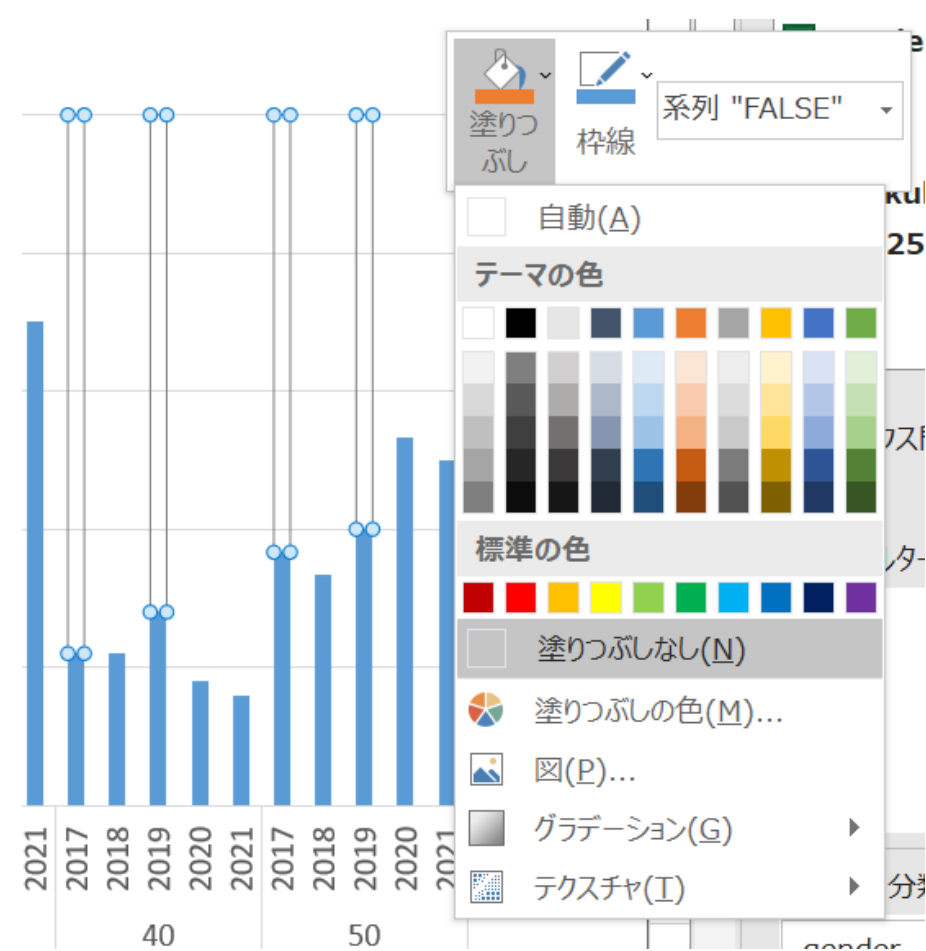
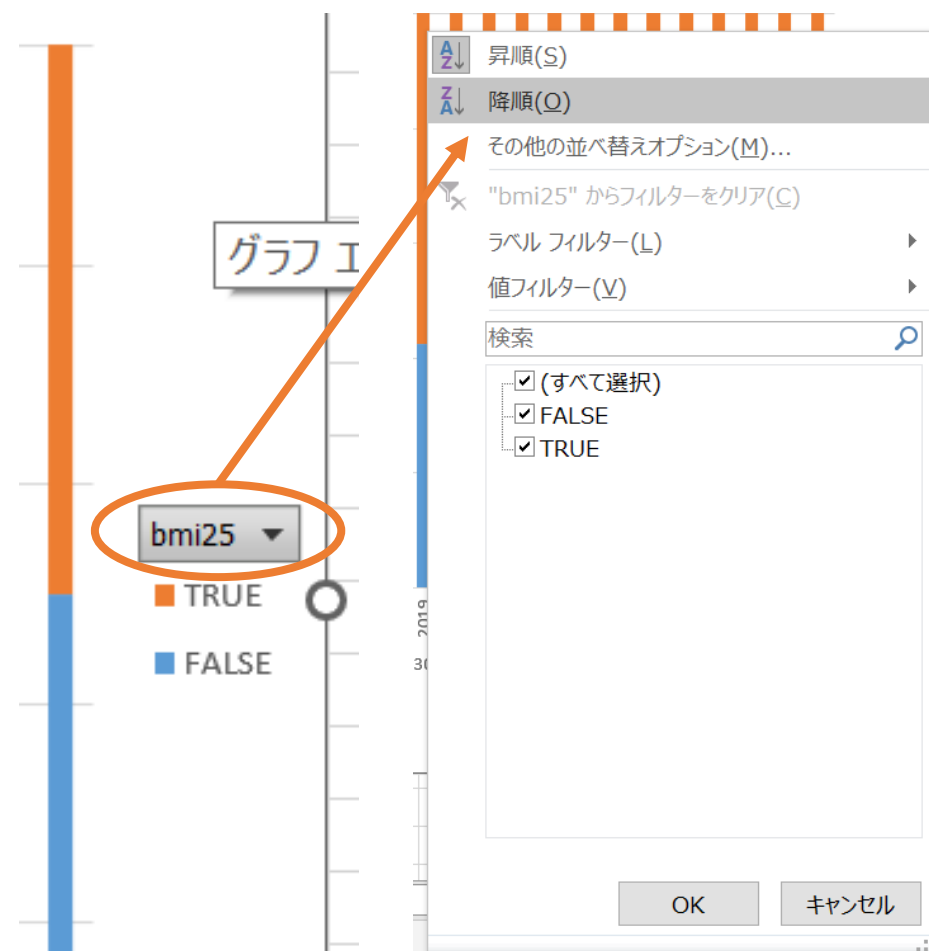
基準値に対する比率

親行集計に対する比率

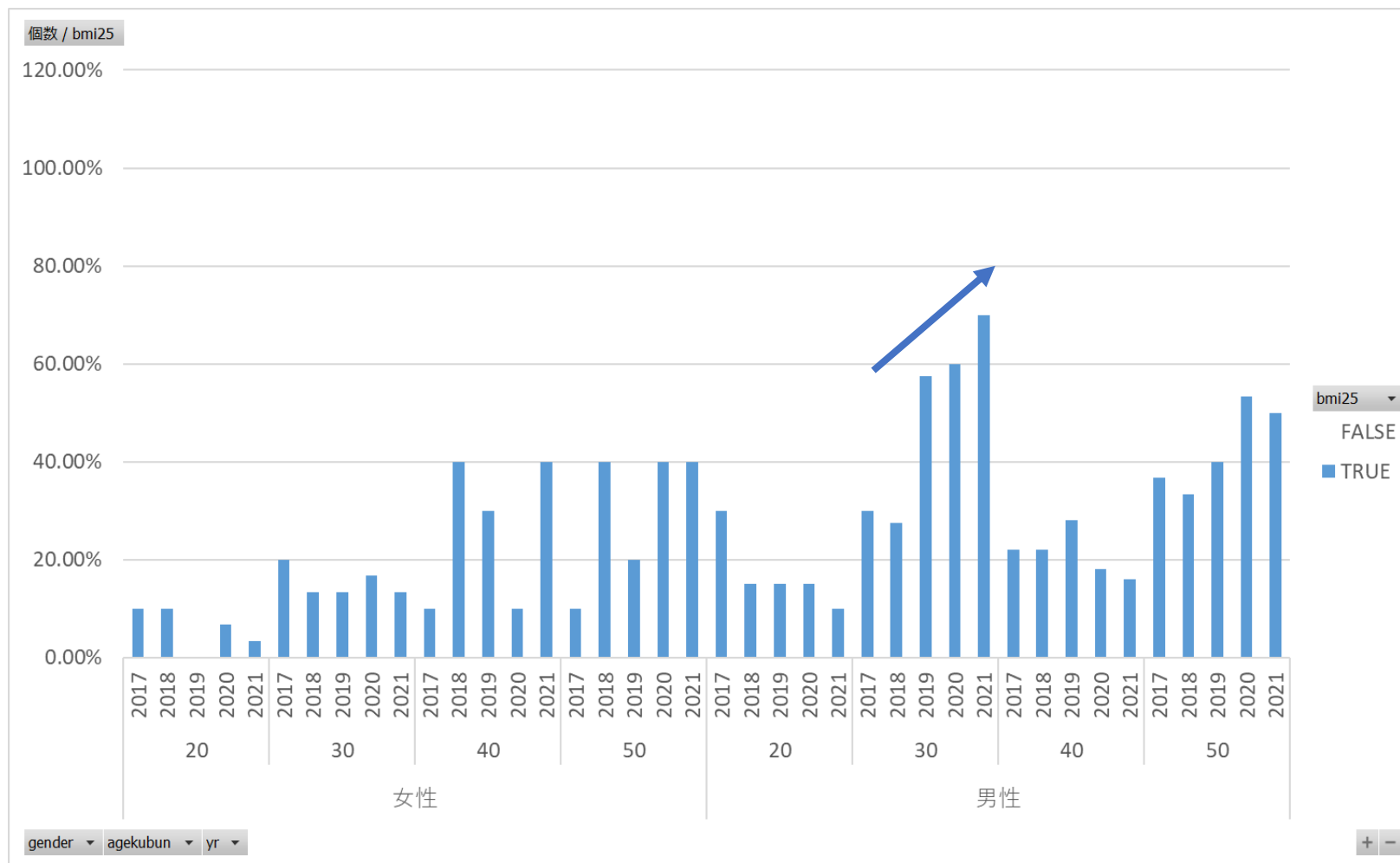
	A	B	C	D
1	個数 / bmi25	列ラベル		
2	行ラベル	FALSE	TRUE	総計
3	女性	85.00%	15.00%	100.00%
4	20	94.00%	6.00%	100.00%
5	2017	90.00%	10.00%	100.00%
6	2018	90.00%	10.00%	100.00%
7	2019	100.00%	0.00%	100.00%
8	2020	93.33%	6.67%	100.00%
9	2021	96.67%	3.33%	100.00%
10	30	84.67%	15.33%	100.00%
11	2017	80.00%	20.00%	100.00%
12	2018	86.67%	13.33%	100.00%
13	2019	86.67%	13.33%	100.00%
14	2020	83.33%	16.67%	100.00%
15	2021	86.67%	13.33%	100.00%
16	40	74.00%	26.00%	100.00%
17	2017	90.00%	10.00%	100.00%
18	2018	60.00%	40.00%	100.00%
19	2019	70.00%	30.00%	100.00%
20	2020	90.00%	10.00%	100.00%
21	2021	60.00%	40.00%	100.00%
22	50	70.00%	30.00%	100.00%
23	2017	90.00%	10.00%	100.00%
24	2018	60.00%	40.00%	100.00%
25	2019	80.00%	20.00%	100.00%
26	2020	60.00%	40.00%	100.00%
27	2021	60.00%	40.00%	100.00%
28	男性	66.86%	33.14%	100.00%
29	20	83.00%	17.00%	100.00%
30	2017	70.00%	30.00%	100.00%
31	2018	85.00%	15.00%	100.00%
32	2019	85.00%	15.00%	100.00%
33	2020	85.00%	15.00%	100.00%
34	2021	90.00%	10.00%	100.00%
35	30	51.00%	49.00%	100.00%
36	2017	70.00%	30.00%	100.00%
37	2018	72.50%	27.50%	100.00%
38	2019	42.50%	57.50%	100.00%
39	2020	40.00%	60.00%	100.00%
40	2021	30.00%	70.00%	100.00%
41	40	78.80%	21.20%	100.00%
42	2017	78.00%	22.00%	100.00%
43	2018	78.00%	22.00%	100.00%
44	2019	72.00%	28.00%	100.00%
45	2020	82.00%	18.00%	100.00%
46	2021	84.00%	16.00%	100.00%
47	50	57.33%	42.67%	100.00%
48	2017	63.33%	36.67%	100.00%
49	2018	66.67%	33.33%	100.00%
50	2019	60.00%	40.00%	100.00%
51	2020	46.67%	53.33%	100.00%
52	2021	50.00%	50.00%	100.00%
53	総計	73.45%	26.55%	100.00%

性別別、年齢別、年度別の割合の推移が一目でわかる
グラフの作成ができました。あとは、TRUEを下にしたいので、





「bmi25」の並び順を降順にしてあげて、FALSEの色を透明にすると、



このようなグラフが出来上がりました。ここからは、男性の30代が特にBMI25以上の方の割合が高く、かつ増加していると見て取れます。

自動化されていることを確認しましょう



「BMI25以上の人の割合」
の集計を自動化

ここまで作成したピボットテーブルとグラフはこれらのシートに含まれています
現時点では、2021年までのデータしか、「step01シート」に含まれていないため、
2021年までの結果がでています



「追加用データ」シートのデータをstep01の末尾に足します

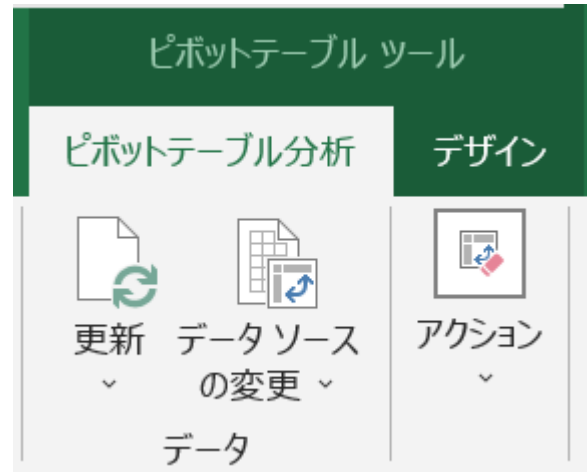
	A	B	C	D	E	F	G	
1	2022	1	男性	29	25.97395			
2	2022	2	男性	29	27.26526			
3	2022	3	男性	29	24.03712			
4	2022	4	男性	29	20.11853			
5	2022	5	男性	29	19.15855			
6	2022	6	男性	29	21.43208			
7	2022	7	男性	29	22.85778			
8	2022	8	男性	29	20.77133			
9	2022	9	男性	29	21.00907			
10	2022	10	男性	29	19.23974			
	◀ ▶	sampledata	step01	step02	step03	step03v2	追加用データ	

1100	2021	219 女性	56	29.92238693	50	TRUE
1101	2021	220 女性	56	23.78356207	50	FALSE
1102	2022	1 男性	29	25.97395159	20	TRUE
1103	2022	2 男性	29	27.2652618	20	TRUE
1104	2022	3 男性	29	24.03712133	20	FALSE
1105	2022	4 男性	29	20.11853015	20	FALSE
1106	2022	5 男性	29	19.15854656	20	FALSE
1107	2022	6 男性	29	21.43208054	20	FALSE
1108	2022	7 男性	29	22.85778397	20	FALSE
1109	2022	8 男性	29	20.77133229	20	FALSE
1110	2022	9 男性	29	21.00907279	20	FALSE
1111	2022	10 男性	29	19.23973682	20	FALSE
1112	2022	11 男性	29	22.27174611	20	FALSE
1113	2022	12 男性	29	11.84760616	20	FALSE

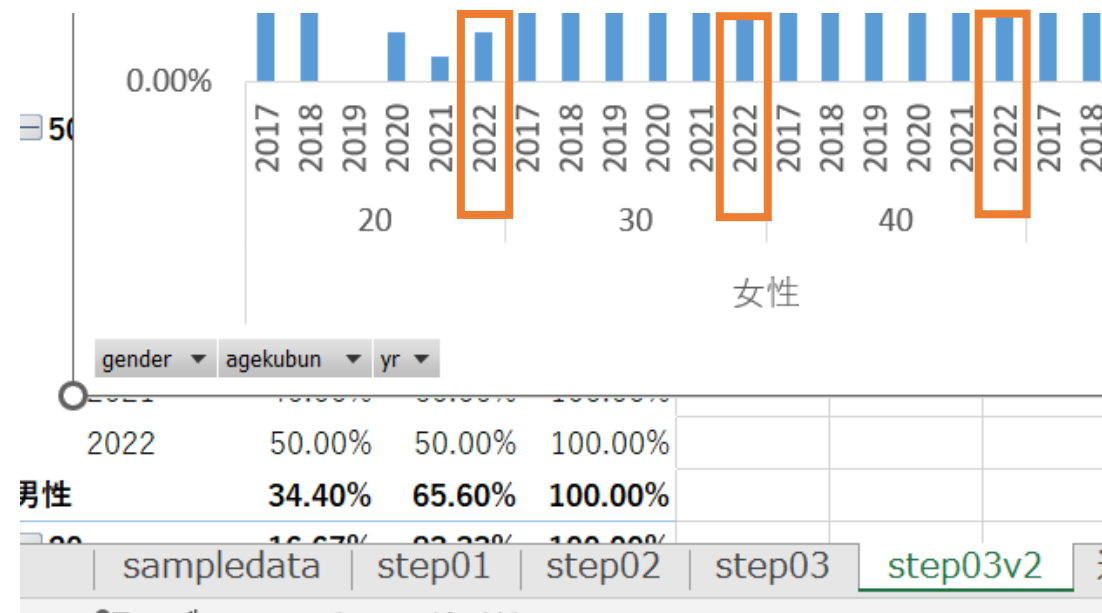
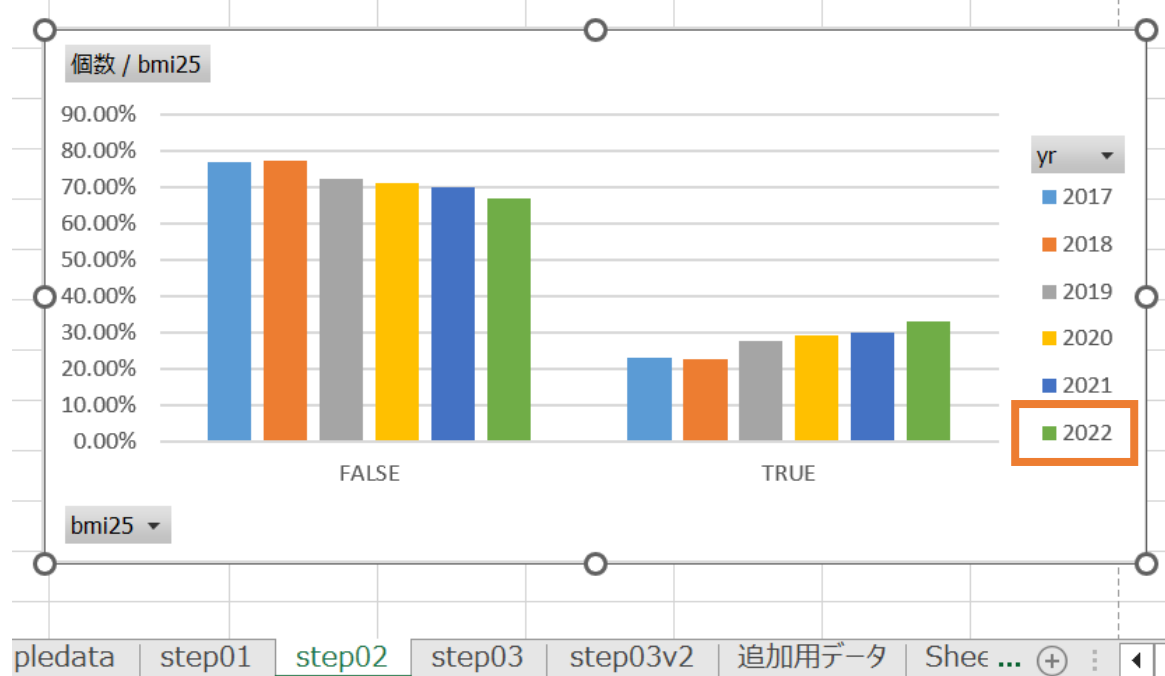
step02

step03

step03v2



データを追加した後、どれか一つのシートの
「ピボットテーブル分析」から、
「更新」をクリックすると・・・



表とグラフがデータに合わせて
自動更新されました！

レポートまで自動更新を行うと便利です



PowerPointと連携してみよう

パワーポイントファイルにエクセルのピボットグラフを貼り付ける場合に、
「貼り付けのオプション」から「～データをリンク」を選びましょう

貼り付けのオプション:



それだけで、エクセルファイルのグラフの内容が新しくなると、
パワーポイントファイルの内容も新しくなります

注) エクセルファイルの保存場所が変わるとうまくいきません

まとめ

- データの保存方法は後の工程に大きな影響を与えます。基本は縦に伸びるデータを意識しましょう
- エクセルであればピボットテーブルとグラフの機能を使い、一度求める形を作れば、元データを追加するだけで自動でグラフや表を作成することができます
- データをリンクしてパワーポイントに貼り付けることで、レポートの内容を自動的に更新するようなことも可能です

効率的なデータのワークフローを構築することで、
産業保健業務や研究をより効率的に行うことができます

次のステップのためのヒント

- データの保存方法について
 - 「Tidyデータ」でGoogle検索
- もっと効率的にデータを取り扱いたい
- 統計解析などを行ってみたい！
- きれいなレポートを自動で出力したい！

無料の統計ソフト
R言語をお勧めします