

既存資料の活用

75

演習 1 データの要約

76

既存資料の活用 演習 1

【目的1】

- 1) データの分布をみて要約方法の判断ができる
- 2) 端的なことばで結果を表現できる

母親の喫煙状況を知る目的で、乳児検診に来た母親に対し、1年間、自己申告による1日の喫煙本数を調べた。

結果をまとめたら、次表のようになった。

77

演習 1 データの要約

本数	0	2	3	4	6	7	8	9	10	12	
人数	14	2	1	1	1	1	2	2	1	1	2
本数	13	14	15	16	17	18	19	20	21	22	
人数	2	1	5	2	1	4	2	11	2	1	
本数	23	24	25	26	28	29	30	32	35	40	
人数	1	1	2	1	1	1	4	1	1	1	

上の表の内容を要約してみましょう

数量データ(離散データ)です。

- 1) 対象数(合計人数)を求めましょう
- 2) 喫煙本数の分布を図示しましょう
- 3) 喫煙状況の記述統計を求めてみましょう

78

演習1 データの要約

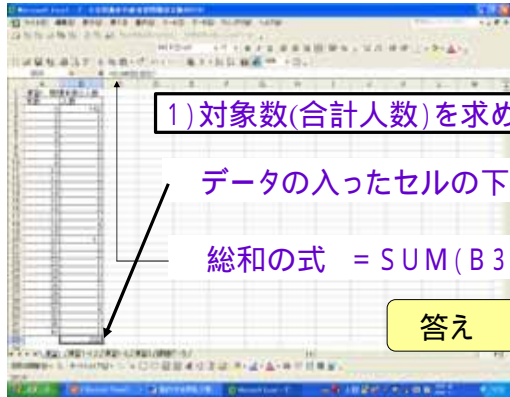
エクセルを立ち上げて、
演習1(1)(2)のシートを開いてください。

1) 対象数(合計人数)を求めましょう

データの入ったセルの下のセルをクリック

総和の式 =SUM(B3:B32) を入れる

答え 人

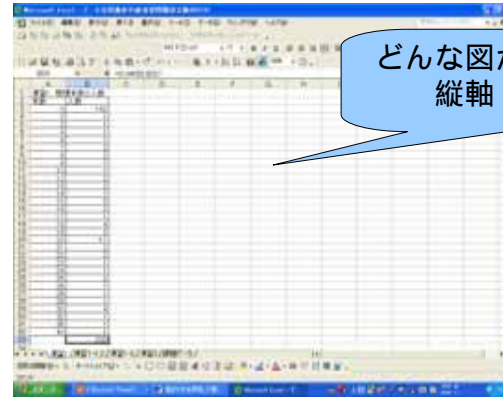


79

演習1 データの要約

2) 喫煙本数の分布を図示しましょう

どんな図が良いですか？
縦軸・横軸は？



80

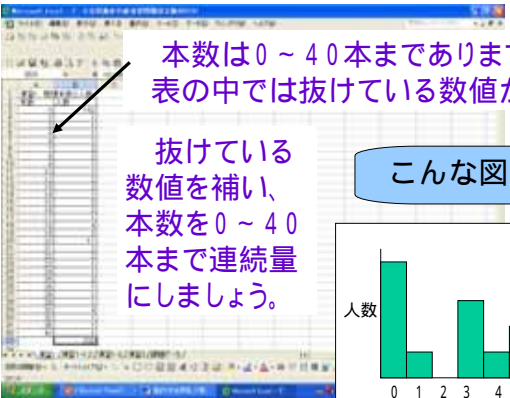
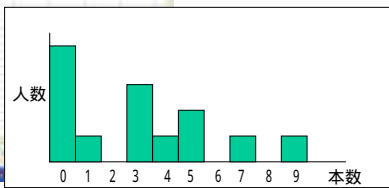
演習1 データの要約

2) 喫煙本数の分布を図示しましょう

本数は0~40本までありますが、
表の中では抜けている数値があります

抜けている
数値を補い、
本数を0~40
本まで連続量
にしましょう。

こんな図が書きたい！



81

演習1 データの要約

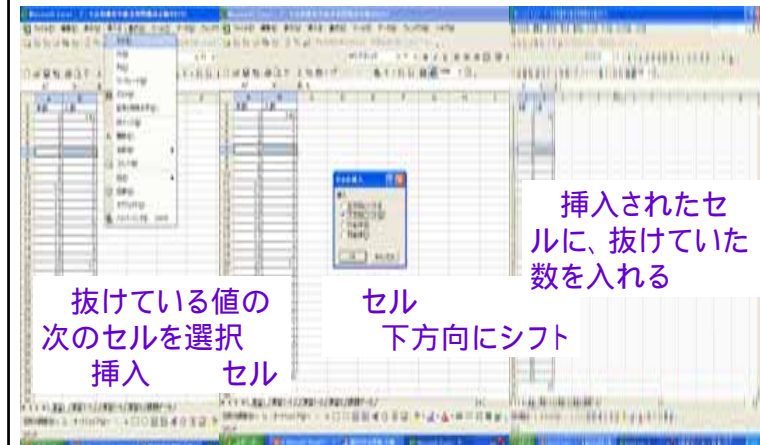
2) 喫煙本数の分布の図示

抜けた数値分のセルを挿
入し、抜けた値を入れる

挿入されたセ
ルに、抜けていた
数を入れる

抜けている値の
次のセルを選択
挿入 セル

セル
下方向にシフト

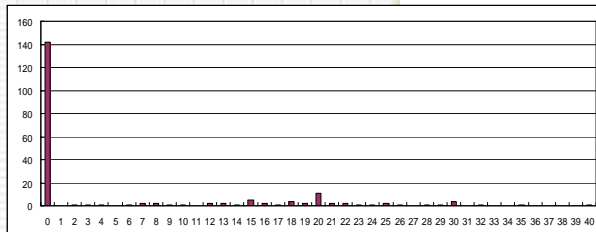
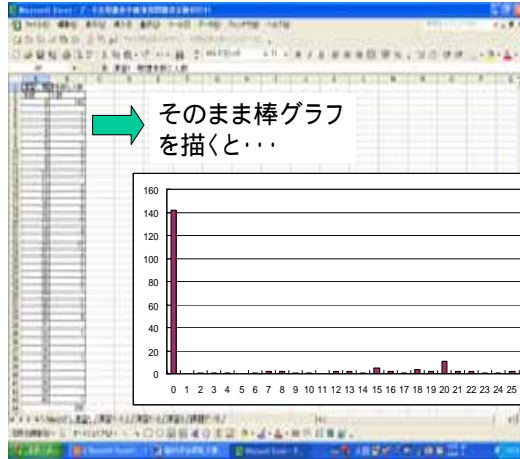


82

演習 1 データの要約

2) 喫煙本数の分布の図示

本数と人数のセルを選択し
棒グラフを描く

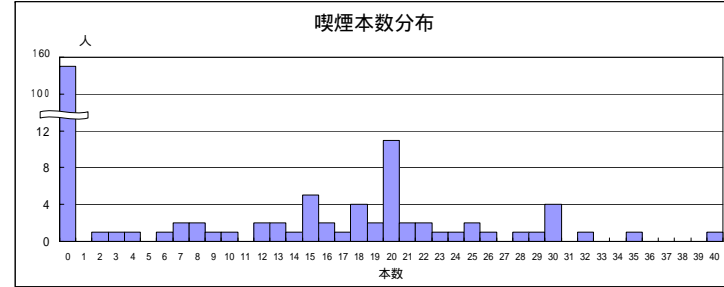


83

演習 1 データの要約

2) 喫煙本数の分布の図示

ちょっと工夫・・・



図を見てわかったことを書き出しましょう

84

演習 1 数値データの要約

3) 喫煙状況の記述統計を求めてみましょう



演習 1 (3) のシート
を開いてください

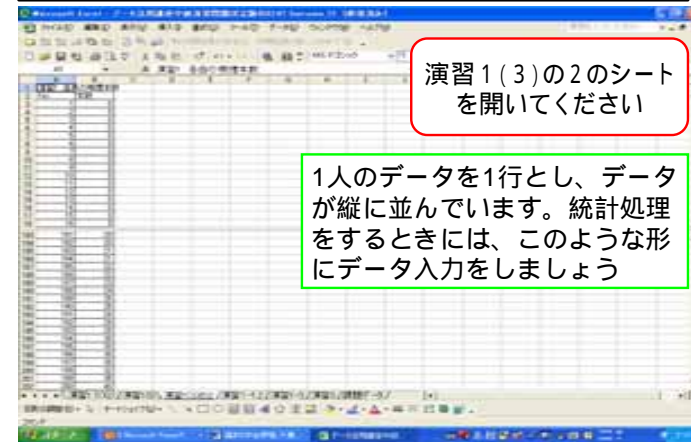
最初に値を入力した表です。
見やすいように、Noと本数が
25番毎に並べました。

しかし・・・ この形では、集
計はやりにくいので、データ
を縦に並べます。

85

演習 1 数値データの要約

3) 喫煙状況の記述統計を求めてみましょう



演習 1 (3) の 2 のシート
を開いてください

1人のデータを1行とし、データ
が縦に並んでいます。統計処理
をするときには、このような形
にデータ入力しましょう

86

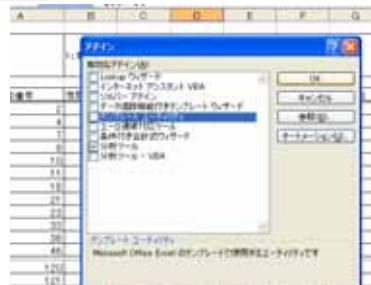
演習 1 数値データの要約

3) 喫煙状況の記述統計を求めてみましょう

分析ツールのインストール



(1) ツール アドイン
を選択

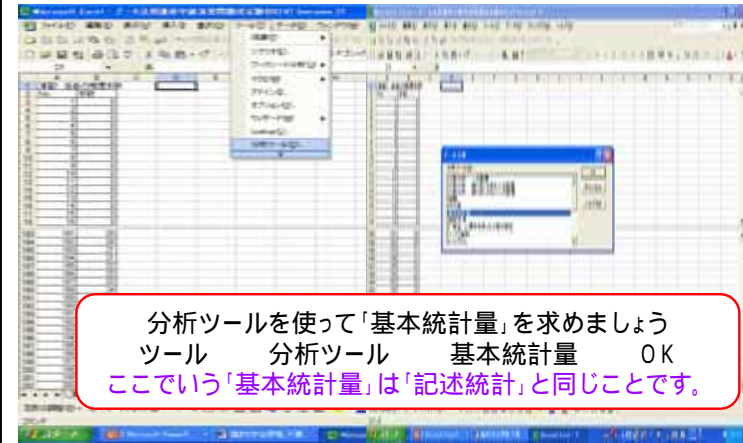


(2) アドインの分析ツール
をチェックし、OK

87

演習 1 数値データの要約

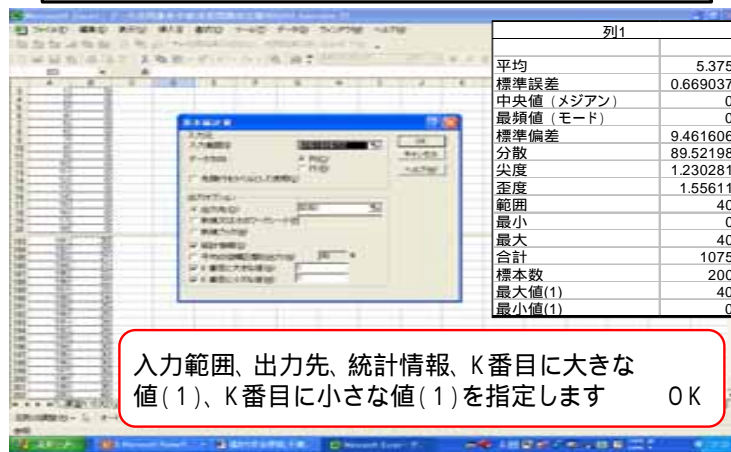
3) 喫煙状況の記述統計を求めてみましょう



分析ツールを使って「基本統計量」を求めましょう
ツール 分析ツール 基本統計量 OK
ここでいう「基本統計量」は「記述統計」と同じことです。

演習 1 数値データの要約

3) 喫煙状況の記述統計を求めてみましょう



入力範囲、出力先、統計情報、K番目に大きな値(1)、K番目に小さな値(1)を指定します OK

演習 1 数値データの要約

3) 喫煙状況の記述統計を求めてみましょう

代表値

Let's Try!

平均(算術平均)

=AVERAGE(セルNo:セルNo)

中央値 =MEDIAN(セルNo:セルNo)

最頻値 =MODE(セルNo:セルNo)

余力のある方、時間が余っている方は、「挿入」「関数」を使って基本統計量を求めてみましょう

バラツキを表す指標

最小値 =MIN(セルNo:セルNo)

最大値 =MAX(セルNo:セルNo)

範囲 = 最大値-最小値

分散 =VAR(セルNo:セルNo)

標本標準偏差 =STDEV(セルNo:セルNo)

90

演習 1 数値データの要約

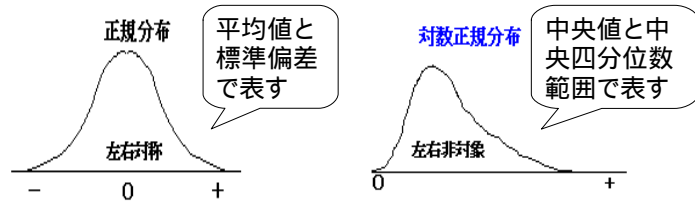
四分位数は基本統計量にありませんでした・・・

バラツキを表す指標

中心四分位数範囲(25 ~ 75パーセントイル)

第1四分位 =PERCENTILE(セルNo:セルNo, 0.25)

第3四分位 =PERCENTILE(セルNo:セルNo, 0.75)



平均値 = 中央値 = 最頻値

平均値 > 中央値 > 最頻値

91

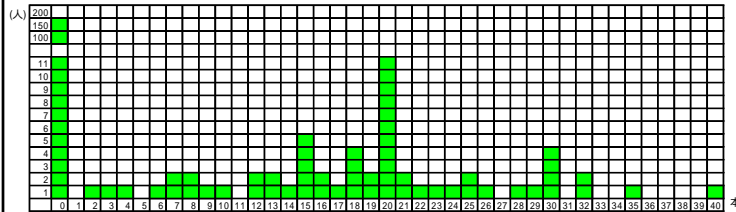
演習 1 数値データの要約

四分位数を求めましょう

値を入力するセルにカーソルをもっていく。
挿入 関数 統計 PERCENTILE
配列に本数のデータ範囲、率に0.25 OK
75パーセンタイルも同様に行う。

25パーセンタイル値	0
75パーセンタイル値	8.25

演習 1 全体についての基本統計量の結果です



全体の記述統計量

平均値 ± 標準偏差 5.4 ± 9.5

中央値 0 最頻値 0

最小値 0 最大値 40

中心四分位数範囲 8.3(0.0-8.3)

先に「気づいたこと」で書いた内容が反映されていますか？

分布が偏っている場合は、要約する時に配慮が必要です！

93

課題 1 層別の分析

全体の分析結果は「喫煙本数 0 本」の人が多いため、喫煙者の状況が分からなくなっていました。喫煙者と喫煙しない人は分けて分析する必要があります。

課題 1

喫煙者について、今までと同様の方法でデータを要約し、下記の()をうめてください。

喫煙者は全体の()%であり、その喫煙本数の特徴は、平均値()本、最頻値()本、中央値は()本である。本数の範囲は()から()本であり、()から()本の間に、中央値を中心とする50%が含まれている。

94

演習 2 カテゴリーデータの要約

95

既存資料の活用 演習 2

【目的2】

- (1) カテゴリーデータの扱いと作図ができる
- (2) 時系列にデータを見ることができる

居住市町村民の健康状態を、
健診判定区分から評価してみよう

基本健康診査の指導区分は毎年報告している。
これにより、地域の健康状態の特徴がわからないだろうか。(判定区分の比較・年次推移)

96

居住市町村住民の健康状態を 健診判定区分から評価する

知りたいこと

- (1) 基本健康診査の判定区分でみた健康状態の特徴
- (2) 経年的に見ると変化はあるのか

必要な情報

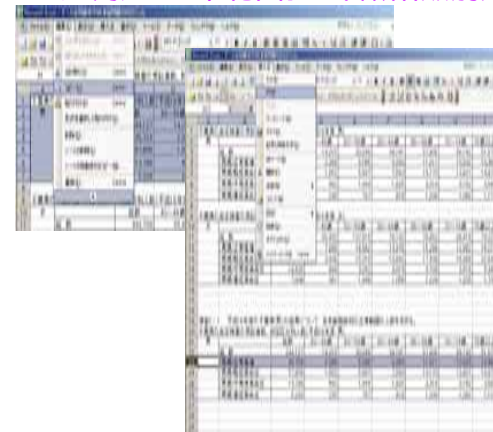
- (1) 基本健康診査の判定基準
- (2) 基本健康診査受診者の判定区分の結果

97

演習 2 カテゴリーデータの扱い

演習 2 - 1 ~ 2 - 3のシートを開いてください

血圧判定「正常範囲」の年齢階級別人数を求める



The screenshot shows a spreadsheet with multiple columns and rows. The columns include age groups (e.g., 10代, 20代, 30代, 40代, 50代, 60代, 70代, 80代) and blood pressure categories (e.g., 正常範囲, 正常高値, 高血圧). The rows represent different data points or years. The table is partially obscured by a text box on the right.

表の男性部分を
選択

編集 コピー

演習 2 - 1 の下
に貼り付け

正常高値の上に
1行挿入し、項目
に「正常範囲」と
入力

正常範囲の人数
を計算する

98

演習2 カテゴリーデータの扱い

血圧判定「正常範囲」の年齢階級別人数を求める

19	演習2-1	平成18年度の千葉県(男)の結果について、各年			
20		千葉県の血圧検査の受診者数、判定区分別人数(平成18年度			
21	男	総数	総数	40~49歳	
22		総数	244,127	14,018	
23		正常範囲	=C22-sum(
24		(再掲)正常高値	C24:C27)	2,361	
25		(再掲)軽症高血圧		57,809	1,851
26		(再掲)中等度高血圧		13,798	461
27		(再掲)重症高血圧		5,234	235
28					767

正常範囲の人数を計算する
 総数のセルを選択
 計算式を入力
 =C22-sum(C24:C27)
 式を他のセルにもコピーする

計算ができたなら、正常範囲のセルを全て選択し、
 コピー 編集 形式を選択して貼り付け 値 OK で値を確定する。



演習2 カテゴリーデータの扱い

貼り付けて表の枠を作る

血圧判定区分の年齢階級別割合の計算

19	演習2-1	平成18年度の千葉県(男)の結果について、各年齢階級別の正常範囲の人数を求める。						
20		千葉県の血圧検査の受診者数、判定区分別人数(平成18年度)						
21	男	総数	40~49歳	50~59歳	60~64歳	65~69歳	70~74歳	75歳以上
22		総数	244,127	14,018	30,348	34,199	57,805	36,192
23		正常範囲	24,560	9,300	17,373	17,668	27,805	27,260
24		(再掲)正常高値			5,005	6,290		
25		(再掲)軽症高血圧			5,897	7,590		
26		(再掲)中等度高血圧			1,444	1,851		
27		(再掲)重症高血圧			767	235		

作成した表を
 選択

編集 コピー

演習2-2の
 下に貼り付け

貼り付けた表の
 数値を削除

28	演習2-2	平成18年度(男)の結果について、年齢階級別判定区分結果の割合を求める。						
29	男	総数	40~49歳	50~59歳	60~64歳			
30		総数						
31		正常範囲						
32		(再掲)正常高値						
33		(再掲)軽症高血圧						
34		(再掲)中等度高血圧						
35		(再掲)重症高血圧						

演習2 カテゴリーデータの扱い

血圧判定区分の年齢階級別割合の計算

31	演習2-2	平成18年度(男)の結果について、年齢階級			
32	男	総数	40~49歳	50~59歳	60~64歳
33		総数	100.0		
34		正常範囲	=C23/C\$22*		
35		(再掲)正常高値	1.00		
36		(再掲)軽症高血圧	23.7		
37		(再掲)中等度高血圧	5.7		
38		(再掲)重症高血圧	2.1		
39					

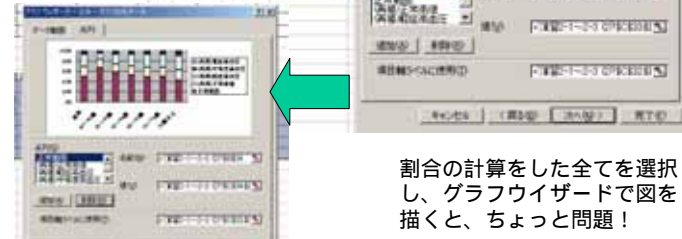
割合を求める式を最初のセルに書く
 =C23/C\$22*100
同じ列では分母は共通なので列固定
 これを全てのセルにコピー
 セルの書式設定をして、小数の桁数を統一
 数値の入ったセルを選択 書式 セル 表
 示形式 数値
 小数点以下の桁数を1に OK
 *小数1桁で表す方法は、他にもたくさんあります。

[テクニック] \$を使ったセルの固定
 (セルの番号にカーソルを置き、F4キーを押してもできます)
 \$C\$4 (全固定)・・・セルをコピーしても変更なし
 C\$4 (列固定)・・・列方向のコピーでは、セルの変更なし
 \$C4 (行固定)・・・行方向のコピーでは、セルの変更なし
 C4 (固定なし)・・・列方向、行方向ともコピーによりセル変更

演習2 カテゴリーデータの扱い

作図

35		(再掲)正常高値	45,724	2,360	5,005	6,290	11,213
36		(再掲)軽症高血圧	27,826				
37		(再掲)中等度高血圧	13,798				
38		(再掲)重症高血圧	5,234				



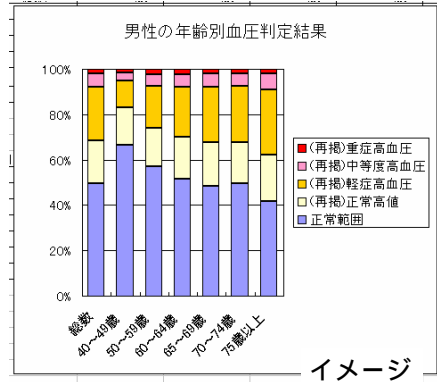
割合の計算をした全てを選択し、
 グラフウィザードで図を描くと、ちょっと問題!

完成イメージ

演習2 カテゴリーデータの作図

血圧判定区分の年齢階級別割合の作図

血圧判定の年齢階級別の割合が比較できるように、結果を作図をし、気づいたことを記載しましょう



図を見て気づいたことを記載しよう

余裕のある人は女性についてもやってみよう

103

演習2 糖尿病判定の年次推移

年齢階級別の糖尿病検査結果について、平成16年から18年の推移がわかるように、作図をしてみましょう

演習2 - 4のシートを開いてください

16年度
17年度
18年度

40~49歳
50~59歳
60~64歳
65~69歳
70~74歳
75歳以上

104

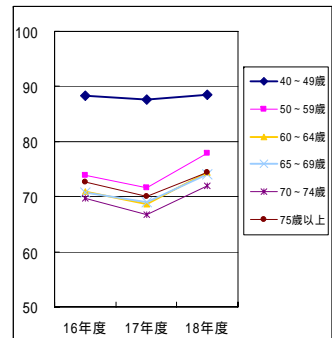
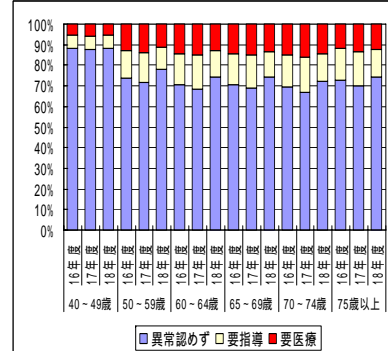
演習2 健康状態の年次推移

作りたい図をイメージします。

こちらを作ります

年齢階級別に判定割合を年次で比較したい場合

異常認めずの判定割合の年次推移を、年齢階級別に比較したい場合



105

演習2 健康状態の年次推移

16年度 17年度 18年度

40~49歳
50~59歳
60~64歳
65~69歳
70~74歳
75歳以上

年齢別に、各年度のデータが並ぶように、データを移動したり、並べ替えたりします

106

演習 2 健康状態の年次推移

	A	B	C	D	E	F	G	H	I
2			男						
3			総数	40~49歳	50~59歳	60~64歳	65~69歳	70~74歳	75歳以上
4	16年度	総数	202,517	13,282	26,861	33,851	47,822	41,703	38,788
5		異常認めず	146,834	11,739	18,937	24,049	33,811	29,092	28,206
6		要指導	28,892	799	3,544	5,084	7,140	6,262	5,854
7		要医療	26,801	745	3,480	4,818	6,871	6,249	4,638
8	17年度	総数	204,346	12,775	26,416	31,630	48,071	44,009	41,445
9		異常認めず	143,392	11,200	18,893	21,734	33,134	29,375	29,056
10		要指導	31,829	855	3,845	5,086	7,235	7,594	6,804
11		要医療	29,025	720	3,678	4,800	7,202	7,040	5,585
12	18年度	総数	235,829	13,361	28,153	32,802	56,219	53,010	51,184
13		異常認めず	176,811	11,816	22,728	24,470	41,667	38,154	38,076
14		要指導	28,184	845	3,213	4,158	7,039	7,199	6,739
15		要医療	29,734	700	3,212	4,274	7,513	7,657	6,378

希望の形のグラフを作成するのに適切な形でデータ入力をするための枠を作ります

演習 2 健康状態の年次推移

40~49歳の16年度の判定区分(異常認めず~要医療)を選択
コピー

カーソルをコピー先のセルに移し右クリック

形式を選択して貼り付け

行列を入れ替える OK

演習 2 健康状態の年次推移

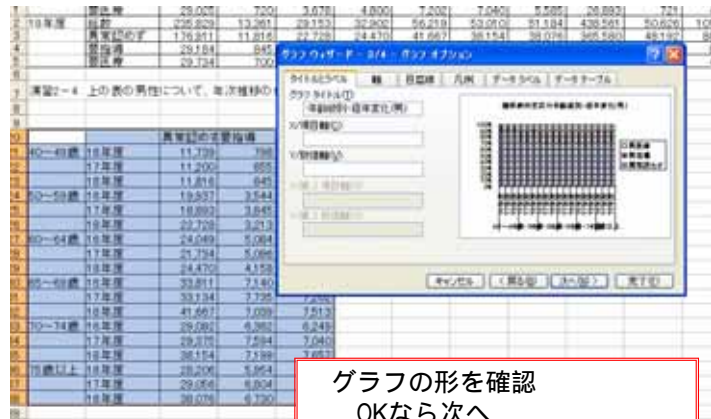
	A	B	C	D	E	F	G	H	I
2			男						
3			総数	40~49歳	50~59歳	60~64歳	65~69歳	70~74歳	75歳以上
4	16年度	総数	202,517	13,282	26,861	33,851	47,822	41,703	38,788
5		異常認めず	146,834	11,739	18,937	24,049	33,811	29,092	28,206
6		要指導	28,892	799	3,544	5,084	7,140	6,262	5,854
7		要医療	26,801	745	3,480	4,818	6,871	6,249	4,638
8	17年度	総数	204,346	12,775	26,416	31,630	48,071	44,009	41,445
9		異常認めず	143,392	11,200	18,893	21,734	33,134	29,375	29,056
10		要指導	31,829	855	3,845	5,086	7,235	7,594	6,804
11		要医療	29,025	720	3,678	4,800	7,202	7,040	5,585
12	18年度	総数	235,829	13,361	28,153	32,802	56,219	53,010	51,184
13		異常認めず	176,811	11,816	22,728	24,470	41,667	38,154	38,076
14		要指導	28,184	845	3,213	4,158	7,039	7,199	6,739
15		要医療	29,734	700	3,212	4,274	7,513	7,657	6,378

行と列が入れ替わりました
同様に、グラフ作成用の表を完成します

演習 2 健康状態の年次推移

作成した表を選択
グラフウィザードをクリック
100%積み上げグラフを選択

演習2 健康状態の年次推移

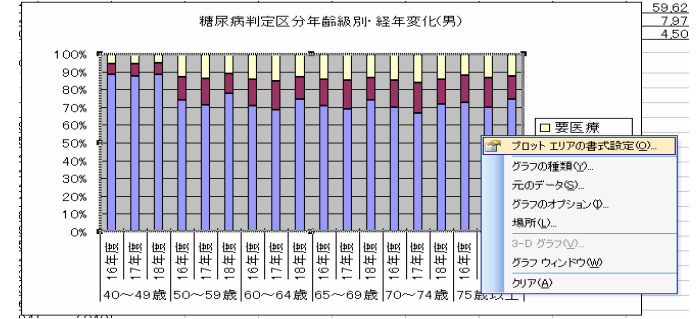


グラフの形を確認
OKなら次へ
グラフタイトル等の記入
完了

111

演習2 健康状態の年次推移

時間に余裕があれば、図のお化粧をしましょう



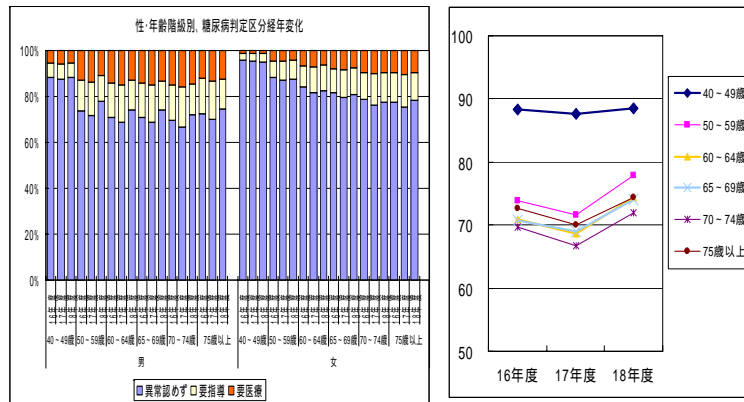
棒グラフエリアの背景を白くする = プロットエリアの書式設定
棒グラフの棒の色や太さを変えるには・
棒の1つをクリック 右クリック データ系列の書式設定
棒の色を変える・パターン画面
棒の太さを変える・オプション画面

112

演習2 健康状態の年次推移

更に時間に余裕があれば、こちらも作ってみましょう

男女別に、1つのグラフにする



113

演習3

数量データの扱い方

114

既存データの活用 演習 3

【目的3】 データ集計の基本作業の習得

- 1) 連続量を含むデータの要約・作図ができる
- 2) カテゴリー分けを適切にできる
- 3) 複数の条件を持つ人を抽出できる
(オートフィルターの活用)
- 4) クロス集計ができる(ピボットテーブルの活用)

性・年齢階級別に、肥満や生活習慣病のハイリスク者の割合を求め、対策を考える資料を作成する。

115

演習 3 連続量を含むデータの分析

演習3のシートを開いてください

データは、A市で行った基本健康診査の測定結果の一部です

* このデータを用いて、生活習慣病対策の資料を作成したい。

対象者の性別の特性を表す

1. 男性、女性の人数を数える
2. 年度末年齢を求め、性別に記述統計で示す
3. 年度末年齢を10歳階級に分け、度数分布を示す
4. BMIを小数点1桁で求める(計算式によるデータの追加)
$$BMI = \text{体重(kg)} / \text{身長(m)}^2$$
5. BMIの判定区分(3区分)をデータに追加(カテゴリーの作成)
6. 50歳代で、BMI判定区分が「肥満」または「やせ」を抽出する
(オートフィルターの利用による対象の抽出)
7. 50歳代女性において、BMI判定区分が「肥満」と「やせ」の人の収縮期血圧平均値と標準偏差をそれぞれ求め、作図する。

116

演習 3 連続量を含むデータの分析

8. 性・年齢階級別のBMI判定区分の分布を作表する
(ピボットテーブルを使ったクロス集計表の作成)
9. 作成したクロス表を使い、“「40~59歳」「60~79歳」の2群について、BMI判定区分が「肥満」または「やせ・標準」の性別の数”を表示させる
10. 性・年齢・肥満の有無別に、血圧、中性脂肪の記述統計(人数、平均、標準偏差、最大値、最小値)を求める。
(ピボットテーブルを使い、層別に記述統計を作成する)
11. 総コレステロールとHDLコレステロール値の関係をみるため、性別に散布図を作成し、相関関係があるかを考える
(散布図の作成)

117

結果のまとめ・発表

118

作図について

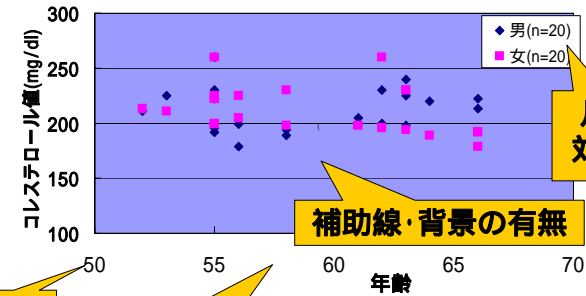
119

作図の注意

軸と単位

表題

性別の年齢とコレステロール値



凡例・対象数

目盛り

軸と単位

120

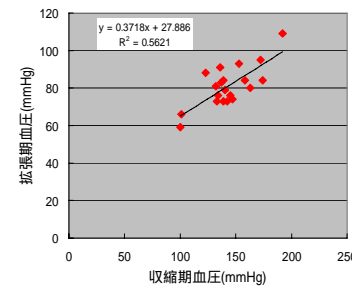
目的別、図の種類

1. 2 値の分布と関連をみる
散布図
2. カテゴリー別の分布状況を表す
棒グラフ・ヒストグラム
3. 比率を表す
円グラフ・積上げグラフ
4. 立体グラフの注意

121

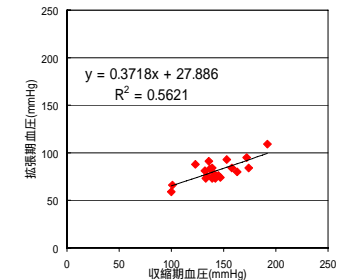
1. 各値の分布状況を表す：散布図

収縮期血圧と拡張期血圧の関連



軸の目盛により回帰式の印象が大きく異なる。目盛の間隔が縦軸と横軸で異なると、判断を誤る原因になることもある。

収縮期血圧と拡張期血圧の関連

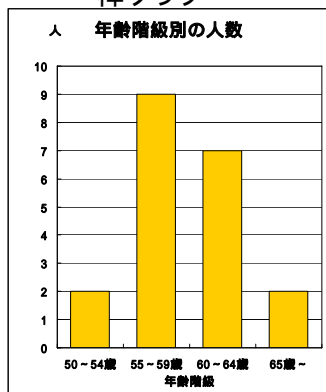


デフォルト(設定様式)では、横軸と縦軸の幅が同じではないので、注意が必要。

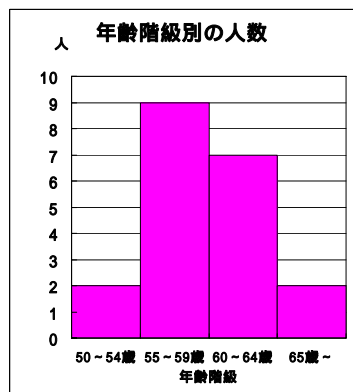
122

2. カテゴリー別の分布状況を表す ：棒グラフ・ヒストグラム

棒グラフ

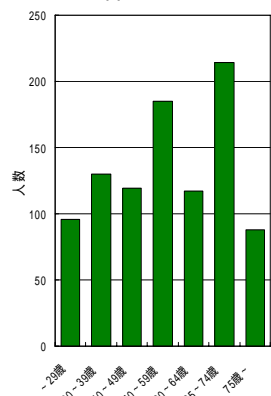


ヒストグラム

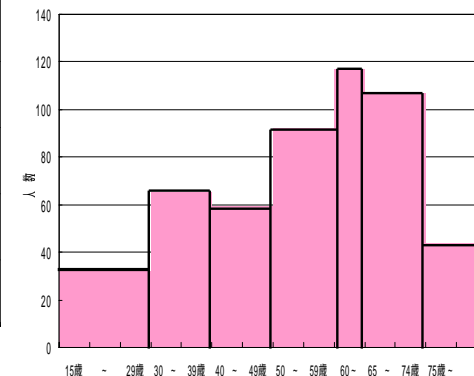


2. カテゴリー別の分布状況を表す 棒グラフ・ヒストグラム

棒グラフ

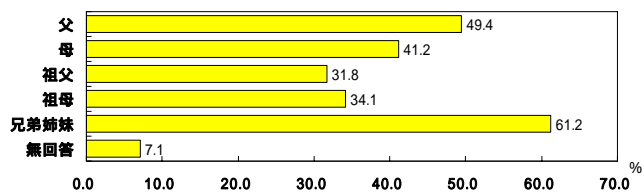


ヒストグラム

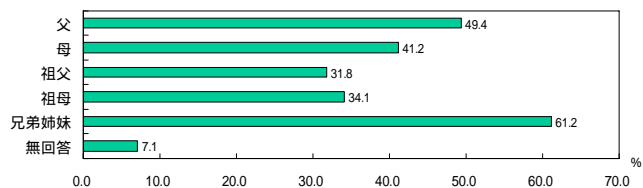


棒グラフ作成の注意

家族の高血圧保有者(n=85)



家族の高血圧保有者(n=85)



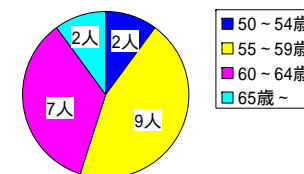
棒の幅とカテゴリーとの間隔に注意

3. 割合を表す：円グラフ・積上げ棒グラフ

円グラフ

構成比をみるのは、
円グラフが最適

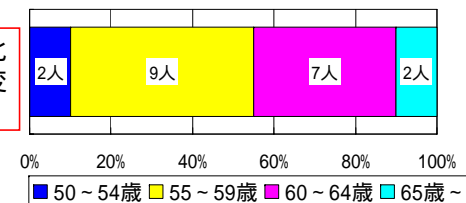
参加者の年齢構成



積上げ棒グラフ

積み上げ棒グラフは構成比
と全体および項目別の変
化を見るのに適している

参加者の年齢構成

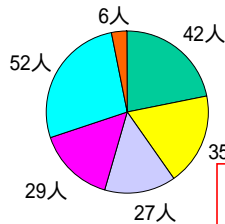


構成比(割合)を表すグラフの作成の注意点

複数回答の表記です、どこか変ではありませんか？

円グラフ

家族の高血圧保有者(n=85)

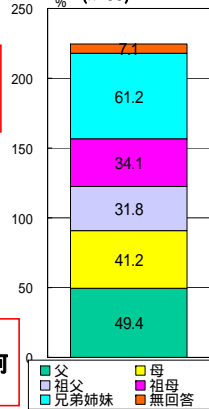


人数の合計は191人だが、回答者は85人

85人が100%積み重ねて何が見たい？

積上げグラフ

家族の高血圧保有者(n=85)

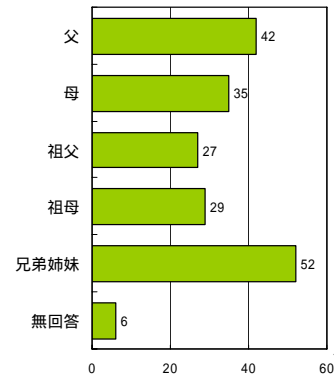


複数回答のグラフ表記

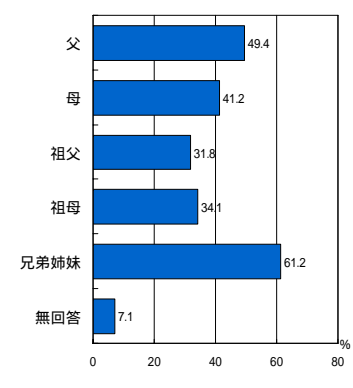
実数の棒グラフ

割合の棒グラフ

家族の高血圧保有者(n=85)

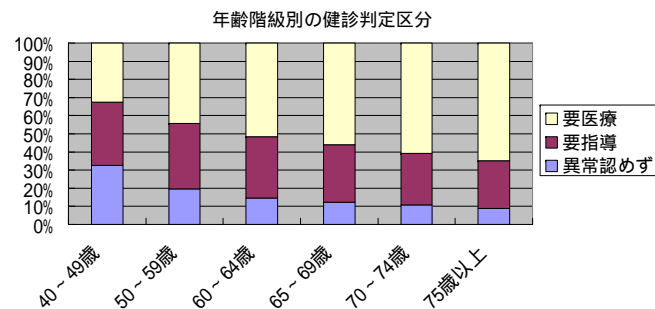


家族の高血圧保有者(n=85)



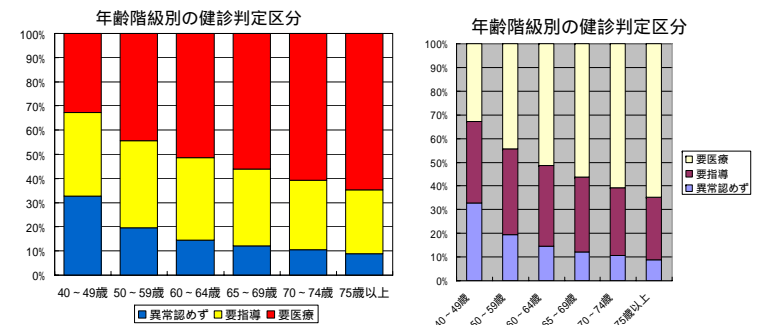
演習2で作成した図をみてみよう

- 演習2 - 2で作成した図を出してください。
- もう少し、見やすくなりませんか？

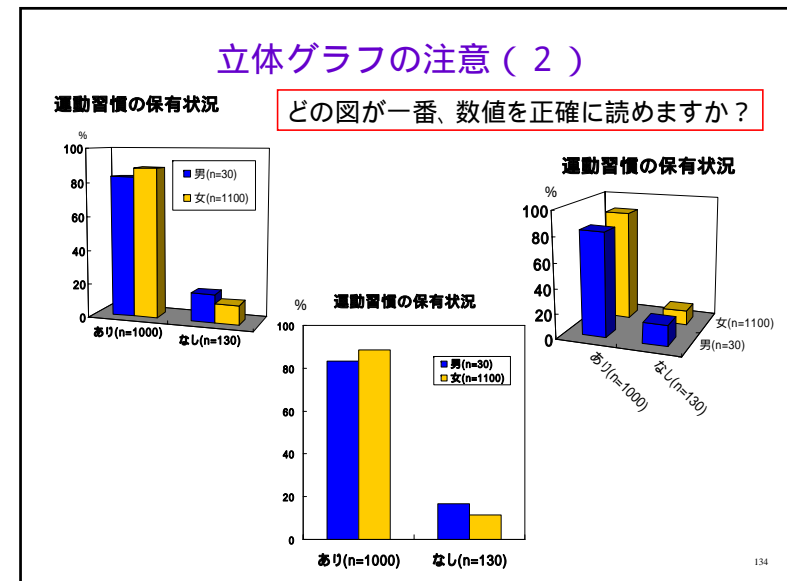
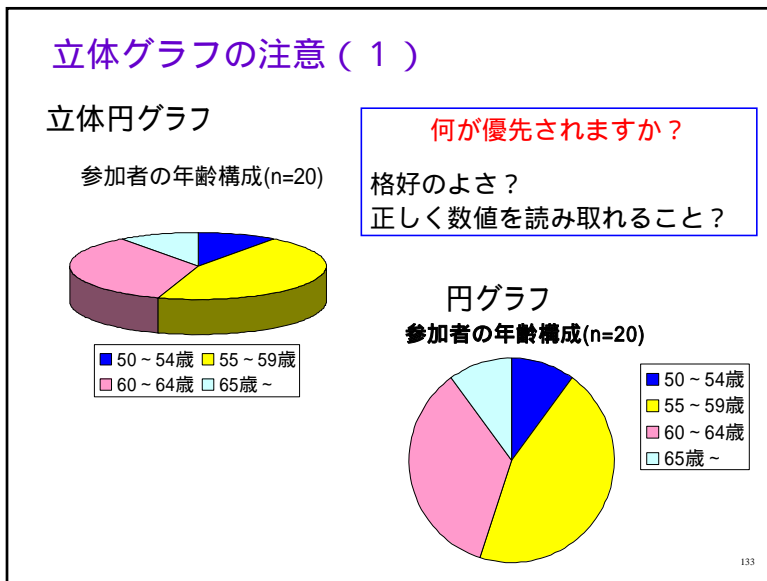
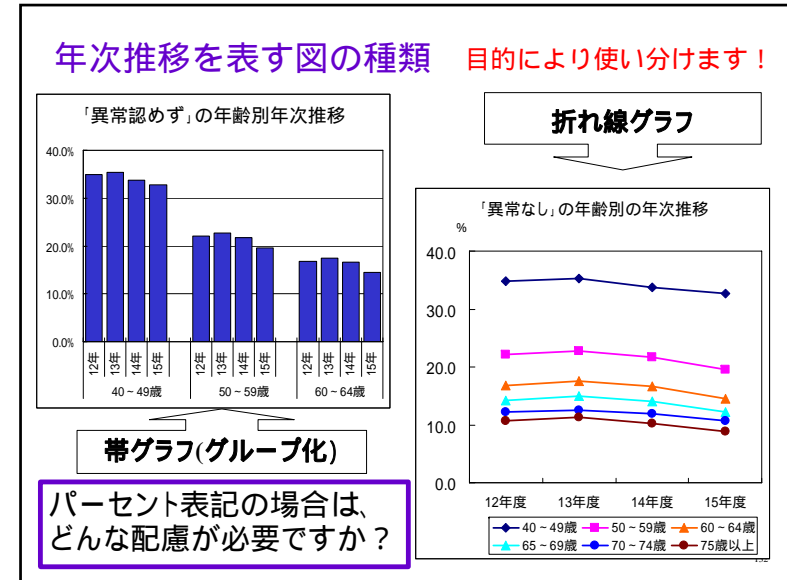
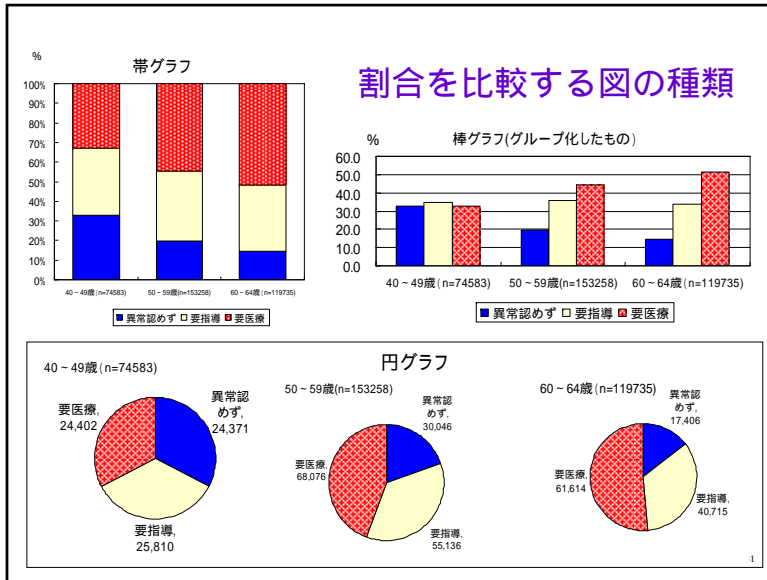


演習2 カテゴリーデータの作図

どちらが見やすい？



エクセルの設定しているフォーマットで完了とせず、背景や棒の太さ、凡例の位置なども調整して、見栄えも気にしましょう！



研究成果の公表

研究発表の目的

- 1) 事業結果の公表
- 2) 他の自治体や研究者からのアドバイスや第三者からの評価を得る

発表形式

- 1) 口頭発表
- 2) ポスター発表
- 3) 概要(抄録)
- 4) 投稿論文

↓
視聴者に研究内容を正しく理解してもらうことが重要！
正しい日本語・適切な表現のチェックを忘れずに！

135

プレゼンテーションの注意点

発表原稿の構成

- 1) 緒言……研究の背景・目的
- 2) 研究方法……研究・調査・実験・解析に関する手法の記述
資料・材料の集め方
- 3) 研究結果……研究などの結果・成績
- 4) 考察……結果の考察・評価
- 5) 結語……結論(省略も可)
- 6) 文献

研究方法は丁寧に詳しく

結果は事実を述べる

考察は結果に基づいて科学的・理論的に

結論は研究目的・結果と矛盾しない

136

プレゼンテーションの注意点

スライド(パワーポイント)の構成

- 1) 緒言
- 2) 研究方法
- 3) 研究結果
- 4) 考察
- 5) 結語

結果と考察

結果は図を活用
表は文字の量・大きさに注意

動画や早い切り替えは、見ている人がついていけない

スライド1枚 = 1分が目安
1枚のスライドは10行程度まで

137

プレゼンテーションの注意点

スライド(パワーポイント)の構成の目安

- 0) 表紙………1枚
- 1) 研究の背景・目的……1~2枚
- 2) 研究方法………1~2枚
- 3) 研究結果………1~3枚
- 4) 考察………1~2枚
- 5) 結語………1枚

図表はスライド1枚に1つが原則
発表時間が7分程度 スライドは10枚以内
1分で300字~350字を目安に

138

プレゼンテーションの注意点

抄録を書くとき

- 1) 緒言
- 2) 研究方法
- 3) 研究結果
- 4) 考察
- 5) 結語

「である」体が普通
「です・ます」の混用を
避ける

結果と考察

考察が「感想」になっていませんか？
「仮説」に対する「結果」の検討
先行研究との比較
研究の限界・問題点 などを記述

139

まとめ(1)

- データを収集するときには、
目的を明確にする。
結果が集団の代表値または集団特性を
客観的に判断できるよう、対象設定をする。
他と比較可能な調査票・測定法を採用する。
- 結果の整理・表現では、
パーセントの算出は、母数が何かを考える。
データの特徴が一目でわかるような作表・
作図をする。

140

まとめ(2)

- 結果を言葉で表すときには、
事実を正確に述べる。
主観的な判断や思い込みは禁物。
比較するときは、比較対照を明確にする。
必要に応じて、検定を行い統計的な差異を
検討をする。

141

理解度の評価のための課題

本日の講義はいかがでしたか？

- 演習3の資料の最後に記載されている「課題」は
各自が解いてみてください。
- 回答をお送りいただければ、コメントをつけてお返し
します。よく分からなかった点などは遠慮なく質問
してください。

142

お 願 い

調査にやり直しはききません。
無駄な調査や使えない・使わないデータの蓄積を避け、
千葉のエビデンスを、協同して作っていきましょう。

調査デザイン・質問紙の作成・データ解析・データの読み方などの疑問や質問、アドバイスが欲しい時は遠慮なく、衛生研究所・健康疫学研究室にご連絡ください。

また、私たちの実施する調査・研究にも是非、ご協力下さい。よろしくお願いいたします。

電話:043-266-7989 Fax:043 - 265 - 5544

柳堀: r.yngbr@mb.pref.chiba.lg.jp

143