

情報の授業

xx高の生徒ってどうよ調査の分析



1.xx高の生徒ってどうよ調査

質問紙調査(Web)を行い、生徒のようすを明らかにします。

- ・質問紙を考えてみてください。2種類

タイプ1: 2種類の数値で回答するもの

例: スマホの一日の使用時間と勉強時間

身長と体重

タイプ2: 一つの属性(2区分)と一つの数値で

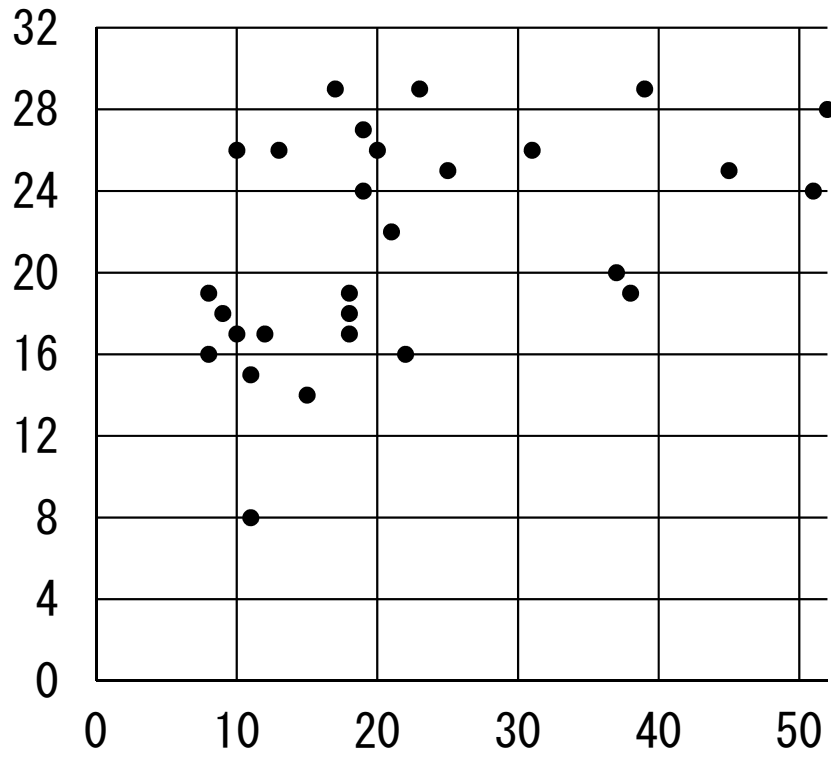
回答するもの

例: 男女と一日のLINEのメッセージ数

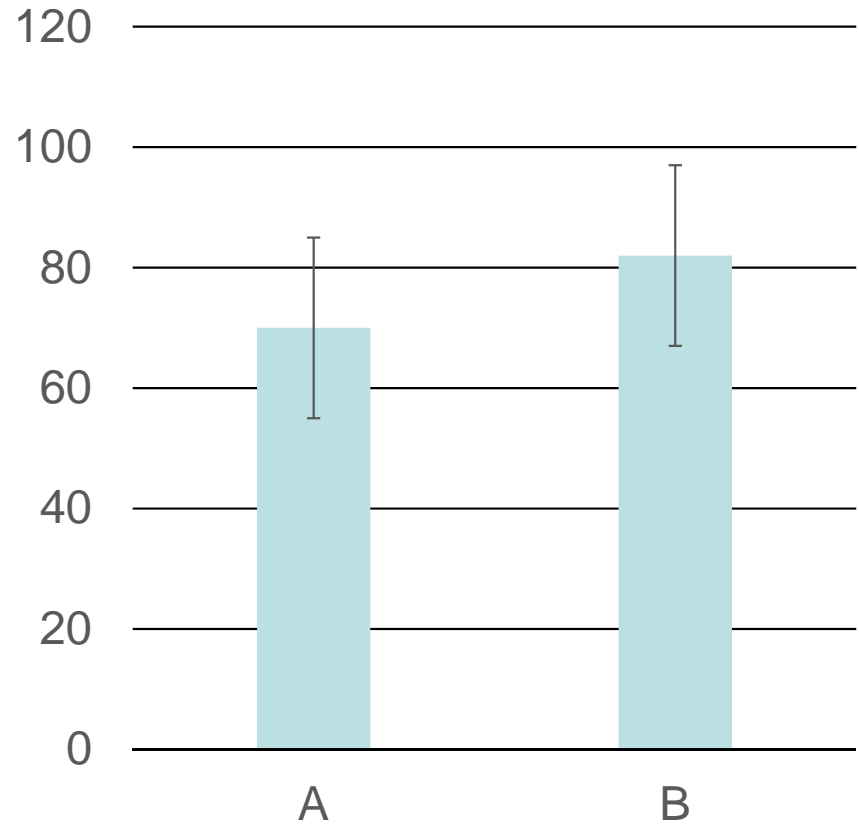
けやき坂/欒坂どっちが好きかと1日の

Youtubeの視聴時間

注意: 回答しても個人が特定されないもの



タイプ1



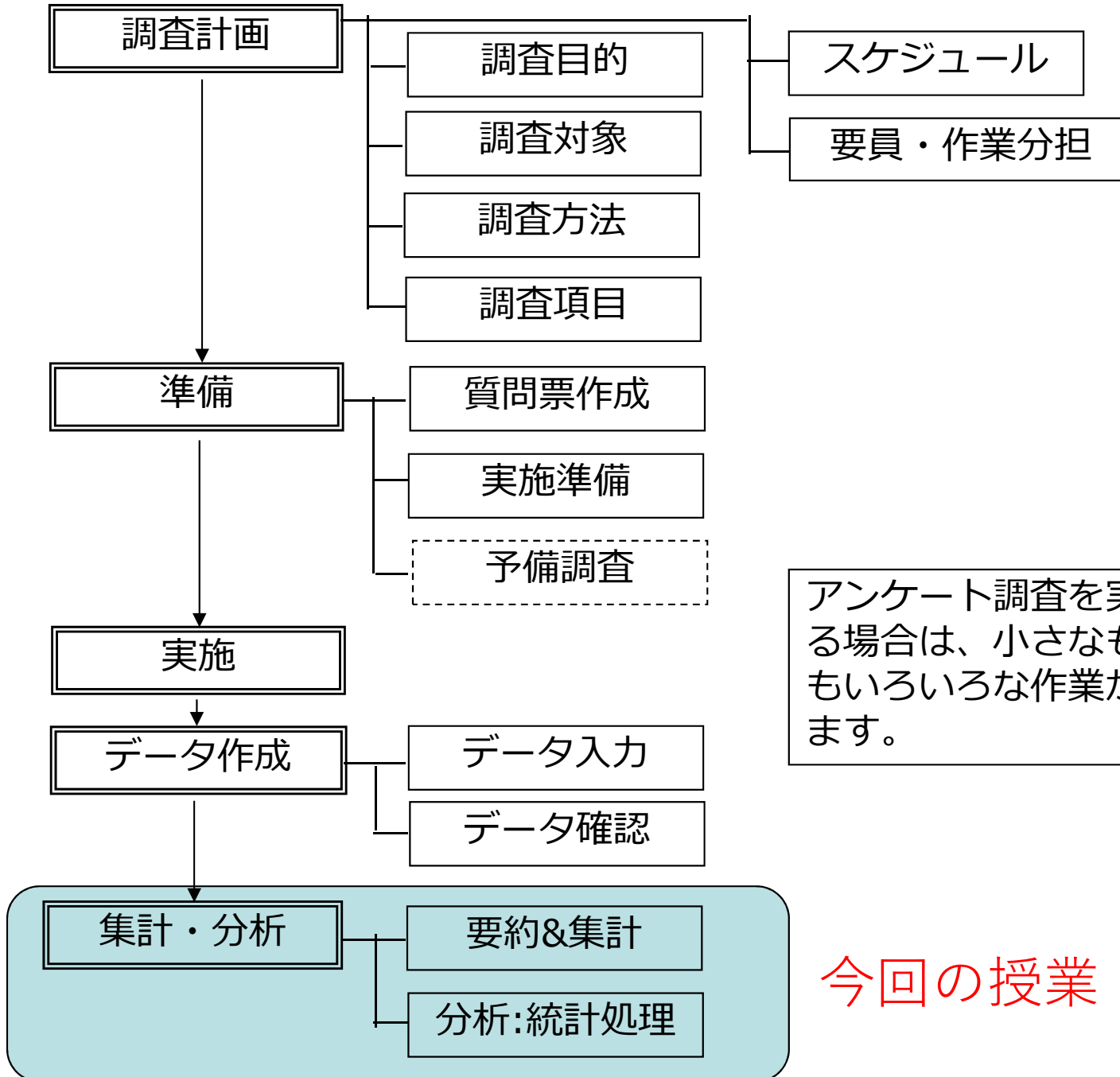
タイプB

今回の作成するパワポのサンプルです
xx東の生徒ってどうよ調査2018

調査パワポサンプル1.pdf
~調査パワポサンプル5.pdf

サンプル1	有意義に時間を使えている人はどんな人
サンプル2	部活動の参加より違い
サンプル3	男女のライフスタイルの違い
サンプル4	生活スタイルについて
サンプル5	平日に少しでも勉強する人としない人の違い

補足：調査の流れ



アンケート調査を実施する場合は、小さなものでもいろいろな作業があります。



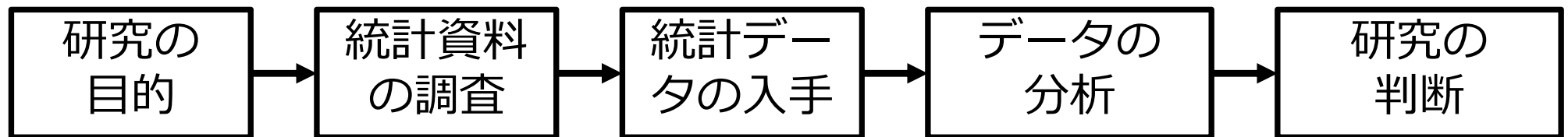
今回の授業

統計・調査もとにした研究

○研究目的の調査を実施する場合



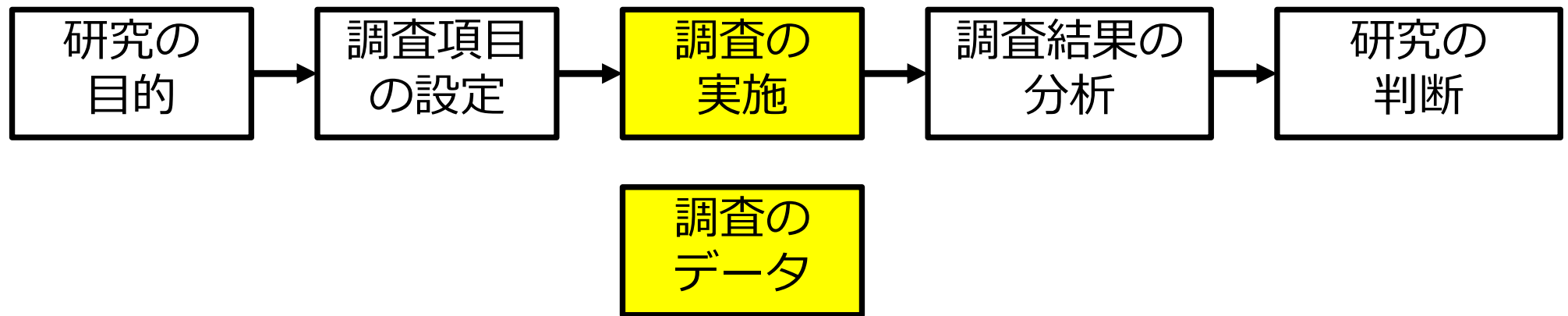
○研究目的の統計資料を収集する場合



今回の市東の生徒ってどうよ調査は、練習で上記の二つのことをする前に:

- ・調査でのデータとはどのようなものか
 - ・データをどのように分析するか
- を理解する。

今回の授業のアウトプット



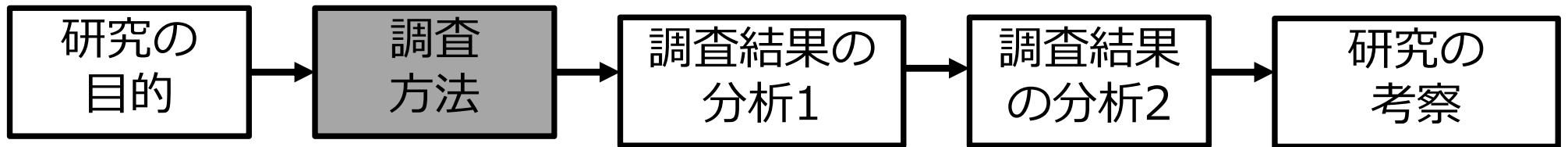
今あるデータを使ってできる研究の目的を考えて、その報告書としてパワポを作成する。

研究の目的例:

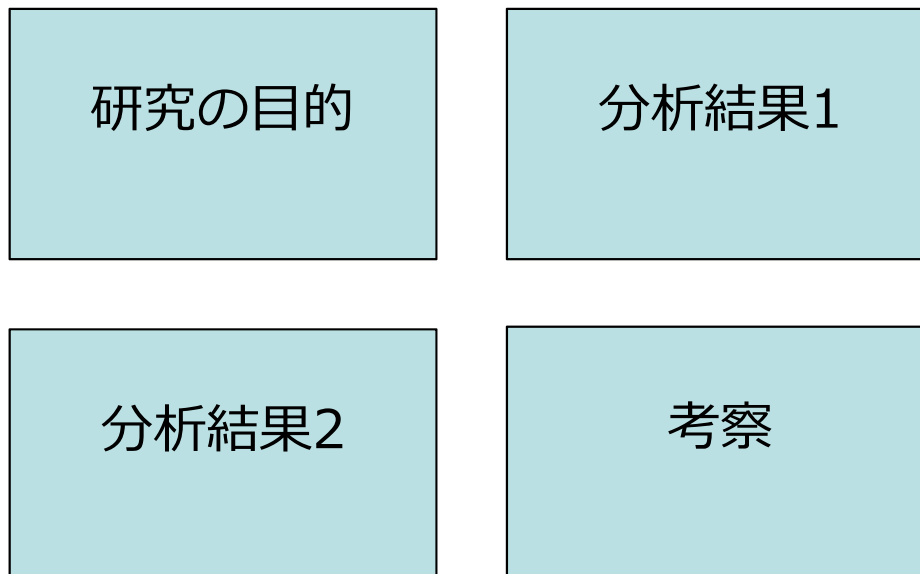
- ・市東生のライフスタイル/価値観
- ・市東生の部活参加による違い
- ・市東生のネットの利用について

.....

今回の授業のパワポ構成



今あるデータをいろいろいじって、どんな研究目的が、でっち上げられるか考えてみよう。



全体で3~4個のグラフを入れる。

データをいろいろ眺めてみよう

袖高調査元データ.xlsx - Excel

ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 開発 チーム 実行したい作業を入力してください...

貼り付け クリップボード フォント 配置 数値 スタイル

条件付き書式 テーブルとして書式設定

F3 : × ✓ fx パン

	A	B	C	D	E	F	G	H	I	J	K	L
1	タイムスタンプ	性別は	学年は	現在、部活動に参加していますか(3年は過去に2年以上参加した場合は参加)	どんなライフスタイルですか。	どちらが好き1	どちらが好き2	どちらが好き3	どちらが好き4	どちらが大切	スマホの種類は(両方持っている人はよく利用する方)	ドラクエで手に入るとしたらどの職業
2	#####	男子	2年	参加	インドア派	パン	ドラマ	たけのこ里	ディズニーシー	お金	iPhone	その他(興味なし)
3	#####	男子	2年	参加	インドア派	パン	アメ	きのこの山	ディズニーシー	お金	Android携帯	賢者

市東調査2018元データ.xlsxをコピーして使用する。

データ確認: 異常値だけは確認しよう

異常値例	回答	誤りの判談	対応
質問: あなたの性別は? 1. 男子 2. 女子	3	・記入又は入力ミス	・アンケート用紙の確認 ・回答者/記入者に問い合わせ ・異常値として印をつける又は未回答に変更
質問: あなたは平日に何時間ゲームをしますか(分で回答してください)?	2000分	・記入又は入力ミス 又は ・実は正しい値	・上記と同様の対応 ・正しい値の場合でも、あまりにも他のデータと違った場合は集計・統計の処理からは外す
	400分		



アンケート調査は質問票の作成から実施、データ入力までいろいろ手間がかかっています。それでもデータの中におかしな値が入っていると、それまでの苦労が無駄になるかもしれません。

そんな理由で最低でもデータの中に異常値が無いか確認しましょう。異常値には、記入や入力ミスによる、明らかな間違ったデータがあります。また、ミスなのか、例えば正しい値だけで他と比べて非常に大きな値などがあります。この時、データの平均などとりすると、正しい区判断できないことなどあるので、正しいデータでも集計や統計処理するときには外すことがあります。似たような考えかたは、オリンピックなどのフィギアや体操等の採点で、一番良い点と悪い点を除くというのも、極端な点数をはずすという意味があると考えられます。

実習：異常値を見つけよう

	A	B	C	D	E	F
1	No	q01	q02	q03a	q03b	q03c
2	1	1	2	7	0	0
3	2	1	4	2	0	0
38	37	2	4	16	0	43
39	38	2	1	10	0	0
40	39	2	4	0	10	87
41	40	2	3	10	0	9
42	Count	40	40	40	40	40
43	CountA	40	40	40	40	40
44	Ave	1.425	2.975	15.425	2	10.25
45	Min	1	1	0	0	0
46	Max	2	4	46	20	87

COUNT: 数値入力されているセルの数、未入力の確認

COUNTA: 入力されているセルの数、文字で入力されているセルの確認

AVERAGE=平均値: 異常値がないかの確認

MIN=最小値: 小さな異常値の確認

MAX=最大値: 大きな異常値の確認



=COUNT(範囲)

=COUNTA(範囲)

=AVERAGE(範囲)

=MIN(範囲)

=MAX(範囲)

異常値を見つけるためには、そんなに難しいことはしないで、入力されているデータの数、最大値、最小値を調べていきます。通常より異常に大きな数があれば異常値として扱うことになります。

あと、あまり未入力の多い質問は分析するかどうか検討することもあります。同様にも未入力の多い回答者も集計の対象からはずすこともあります。



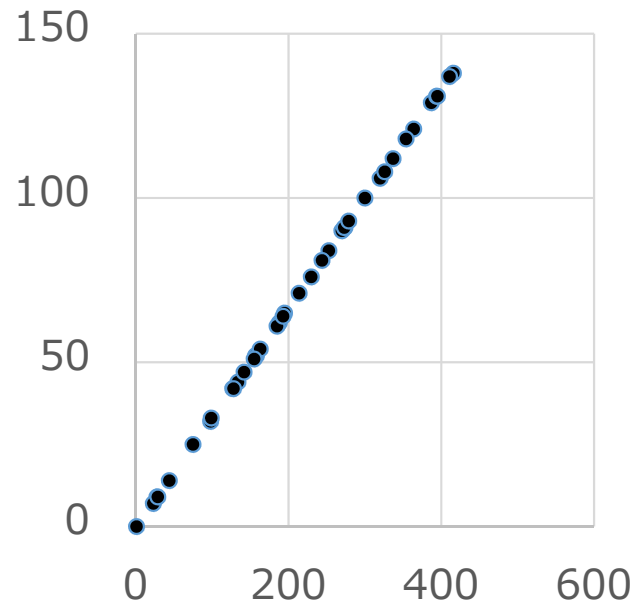
データの分析

	A	B	C	D	E	F	G
1	現在、部活動に参加していますか(3年は過去に2年以上参加した場合は参加)	登校時刻(8:40)の平均で何分前に来ていますか	授業後(15:30)平均で何分ぐらい学校にいますか	平日スマホは何分ぐらい使いますか	平日何分ぐらい勉強をしますか	休日何分ぐらい勉強しますか	一日平均どのぐらい睡眠時間をとっていますか(7時間30分の場合は7.5と回答してください)
2	参加	5	5	300	15	15	6
3	不参加	5	35	0	0	0	5
4	不参加	20	10	120	0	30	6
5	参加	60	120	120	0	10	6.5
6	参加	40	240	60		0	6
7	参加	60	10	120	60	180	9
8	参加	60	240	60	0	30	5.5
9	参加	20	10	120	0	30	6

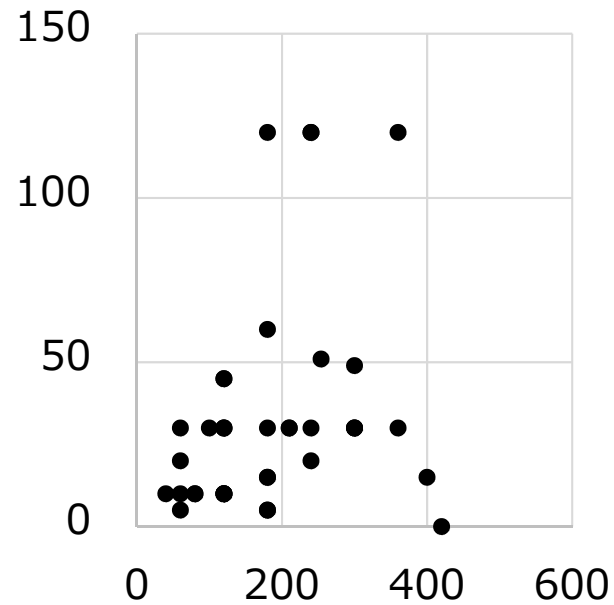
Sheet1 Sheet2 (+)

別シートを作り、必要なデータだけコピーしておく
と分析しやすい

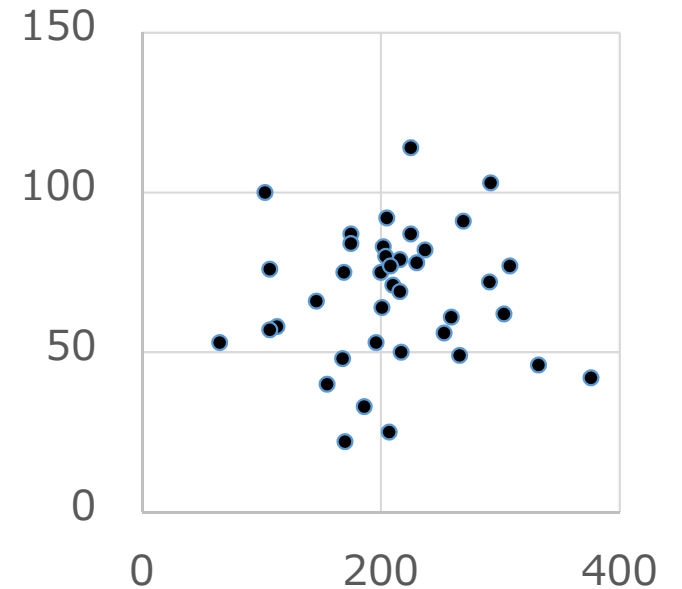
相関



相関が完全にある
 $y = ax$



相関の強さ



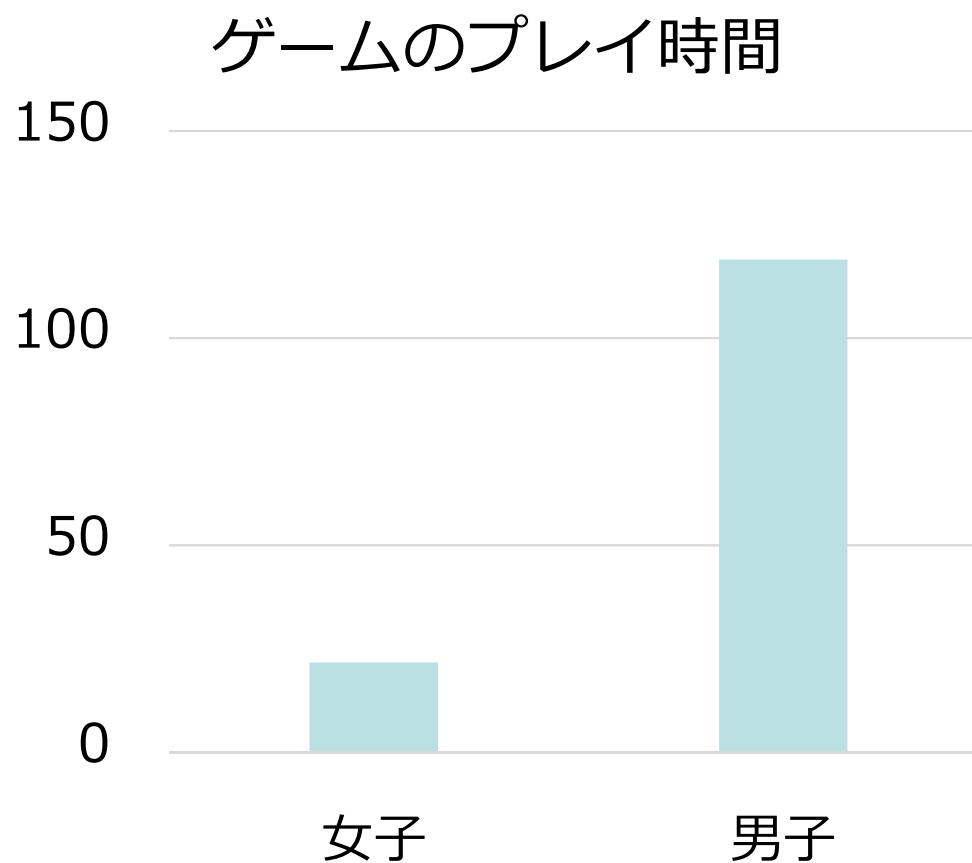
相関がない
バラバラ

グループ間の数値の違い: データの分析の考え方2

グループ間の数値の違いを比較する: **平均等を比較する**

表: 一日のゲームのプレイ時間の男女の比較

	平均
女子	22分
男子	119分



グループごとの平均を求める(1)

The screenshot shows the Microsoft Excel interface with the 'Data' ribbon selected. The 'Sort & Filter' group is highlighted, and the 'Sort' icon is circled in red. A 'Sort' dialog box is open, showing the 'Sort by' dropdown set to '現在、部活動に参加し...' and the 'Order' dropdown set to '昇順'. The spreadsheet data is visible in the background, with columns A through M and rows 1 through 4.

	A	B	C	D	E	F	G	H	I	J	K	L	M
	現在、部活動に参加していますか(3年は過去に2年以上参加した場合は参加)	登校時刻 (8:40)	授業後	平日スマ	平日何分	休日何分	一日平均						
1													
2	参加												
3	不参加												
4	不参加												

グループが固まるように並べ方する

グループごとの平均を求める(2)

16	女子		0		
17	女子		10		
18	女子		0		
19	女子		0		
20	女子		0		
21	女子		2	21.7	
22	男子		180		
23	男子		30	=average(
24	男子		40	B22:B41)	
25	男子		30		
26	男子		60		
27	男子		60		
28	男子		300		

各グループごとの平均を求める

Excelの基本操作

よく使う関数(こけだけは覚えよう)

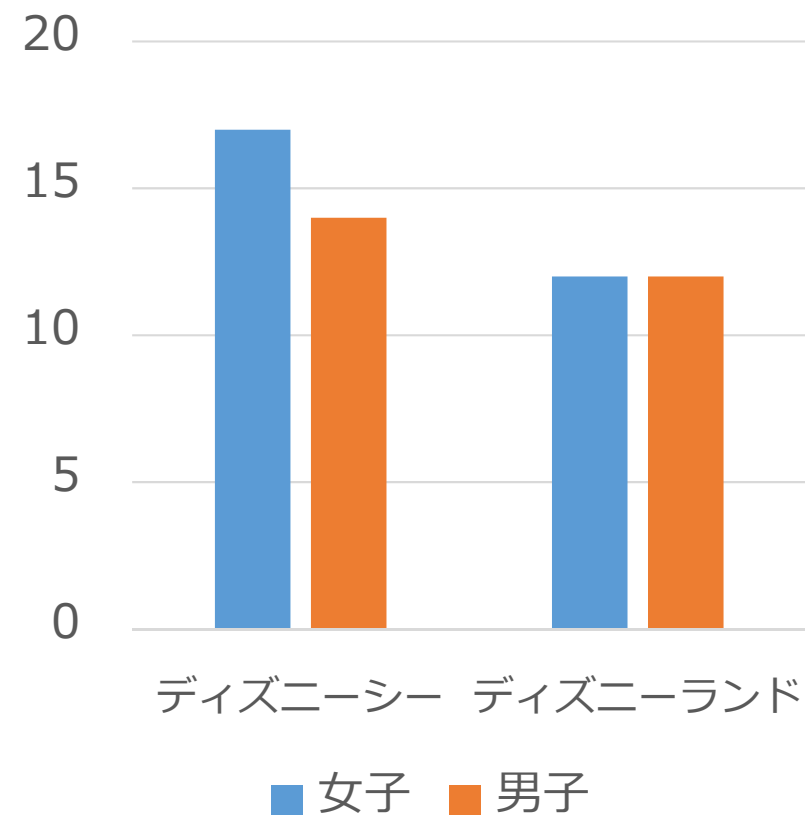
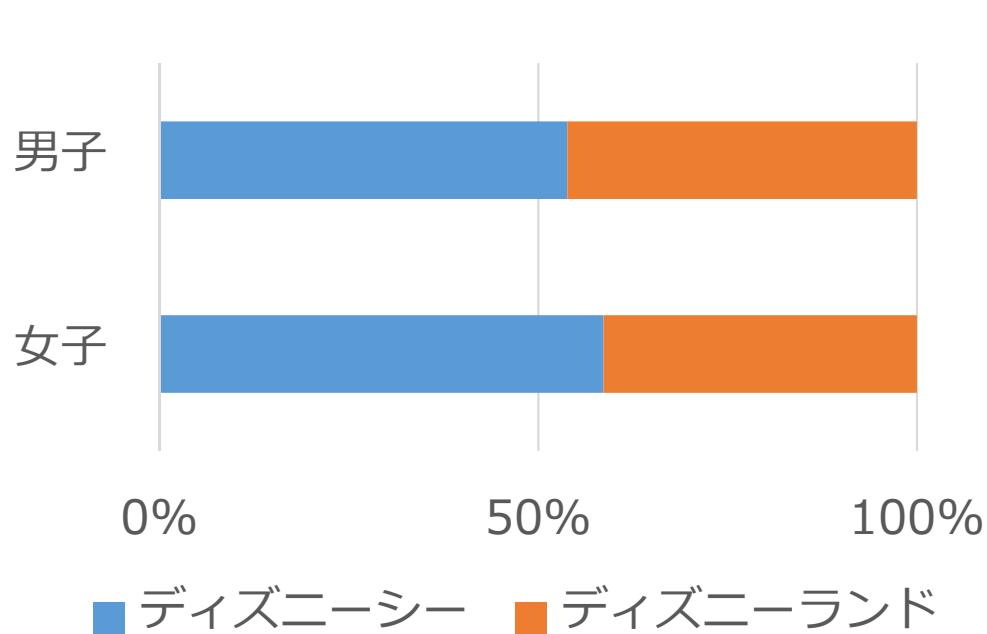
= 関数()で使う

=SUM()	合計	=SUM(領域)
=AVERAGE()	平均	=AVERAGE(領域)
=STDEV.P()	標準偏差	=STDEV.P (領域)
=MAX()	最大値	=MAX(領域)
=MIN()	最小値	=MIN(領域)
=COUNT()	数値の数	=COUNT(領域)
=COUNTA()	数値・文字の数	=COUNTA(領域)
=INT()	整数部のみ	=INT(セル)
=IF()	条件での内容変更	=IF(条件、正の場合の値、誤の場合の値)

クロス集計:データの分析の考え方3

複数の属性を組み合わせて人数を調べる(クロス集計)

	ディズニーシー	ディズニーランド
女子	17	12
男子	14	12



クロス集計の仕方

The screenshot shows the Microsoft Excel interface with the 'Data' ribbon selected. The 'Sort' dialog box is open, displaying the following configuration:

列	並べ替えのキー	順序
最優先されるキー	性別	昇順
次に優先されるキー	ディズニーランド	昇順

The background spreadsheet shows the following data:

	A	B
1	性別	ディズニー
2	女子	ディズニー
3	女子	ディズニーランド
4	女子	ディズニー
5	女子	ディズニーランド

二つのキーを使って並び替えすると、それぞれのグループが固まってカウントしやすい

Excelチャレンジしたい人(1)

= AVERAGEIF (範囲, 検索条件, 平均対象範囲)
を使うと、並び替えせずに各位グループの平均を計算することができます。

=COUNTIFS (範囲1, 検索条件1, 範囲2, 検索条件2, ...)
を使用すると、並び替えせずに、クロス集計の各人数を計算することができます。

インターネット等で調べてみよう。

Excelチャレンジしたい人(2)

かなり高度ですが、非常に便利です。

Work市東調査2018元データ.xlsx - Excel

ファイル ホーム **挿入** ページレイアウト 数式 データ 校閲 表示 開発 チーム

ピボットテーブル おすすめ テーブル

ピボットテーブル おすすめ テーブル

ピボットテーブル

複雑なデータをピボットテーブルに簡単に配置し、集計します。

参考: 値をダブルクリックすると、どの値が合計に含まれるかを確認することができます。

詳細情報

			C	D	E	F	G	H
			ライフスタイル	女子の髪	小説/アニメ	父親/母親	ディズニー	兄弟姉妹
			インドア派	ショート派	アニメ	母親/ママ	ディズニー	いない
3	2019/2/5 9:51	女子	アウトドア派	どうでもいい	小説	母親/ママ	ディズニー	兄弟(年上)
4	2019/2/5 9:51	女子	インドア派	ショート派	小説		ディズニー	弟妹(年下)

ピボットテーブルを使えば、元データのグループごとの平均やクロス集計が楽にできます。
インターネット等で調べてみよう。

プレゼンを意識したパワポ

チラシは人の興味を引くことを目的としているので、
見栄えや読んでわかることに注意して作ります。

プレゼン用パワポ

- ・ 発表する内容を的確に伝えることが目的です。
- ・ 話すことの補助的/補完的な位置づけです。
- ・ チラシのように、あまりごてごてしない。
- ・ 見やすく・理解しやすいこと。

プレゼン用のパワポの形式(評価ポイント)

□ 情報の凝縮/単純化

必要なことだけ書く

例えばタイトルがカッコよくでるアニメーションなどはムダ

サンプルで示したものは文書として長い

□ フォントは、メイリオ又は游ゴシックで 最低18pt以上 (できたら20~24p)

フォントサイズの意味の統一

□ そろえたレイアウト

□ 色は少なく/色の意味の統一

プレゼン用のパワポの内容(評価ポイント)

□ 結果と考察が明確に書かれているもの

結果: データおよびデータをもとにして作成してグラフから客観的に誰でも納得できる事実

考察: 結果の背景、結果を生じる理由などの自分の考えなどで、結果から無理なく推測できるもの。

おまけ：考察を考える時に因果関係の考え方。

