

定性的なデータの分析

情報 I 第21回授業

04情報通信ネットワークとデータベース

対応データ 22exp21.csv

今日のテーマ

- 言葉を「数値化」する
- 見える化する
- わかりやすく示す

(復習) データの種類 (p.202)

ここに
注目!

データの種類	尺度	意味	単位	例
質的データ (定性的)	名義尺度	区別しかできない	ない	職業区分、電話番号
	順序尺度	大小比較ができる	ない	優良可の区分、震度
量的データ (定量的)	間隔尺度	差が意味を持つ	ある	気温、偏差値
	比率尺度	比が意味を持つ	ある	長さ、重さ

情報分析 (p.194)

☆データに対し、適切な分析方法を理解する
「定量(量的)」と「定性(質的)」

- 数値化されたもの (定量的なデータ)
 - 集計してグラフ化
 - 統計処理
- 数値化されていないもの (定性的なデータ)
 - 関係性や因果関係、順序などを図解
 - 同じような内容ごとや程度にまとめて数値化
 - テキストマイニングなどで数値化、分析

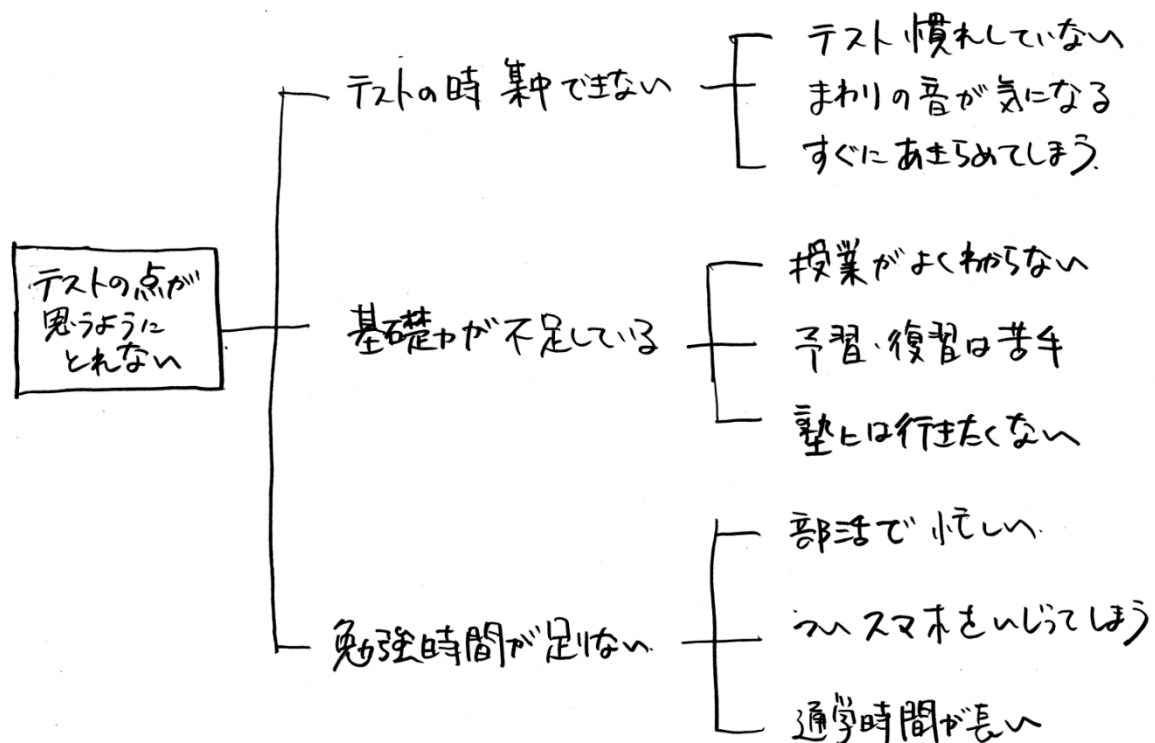
今回の
テーマ

定性的なデータを「見える化」する(教. 資料3、4)

- 関係をあらわす(授業で扱ったり紹介したりしましたね)
 - イメージマップ(キーワードマッピング)、マインドマップ
 - ロジックツリー、特性要因図(フィッシュボーン)、IE図
- 関係性や流れを示すための図
 - ベン図、プロセスチャート、循環図、ピラミッド図、ツリー図 など
- プロセスを表す
 - ガントチャート、PERT図
- 相対的な特性を表す
 - 座標軸、マトリックス図

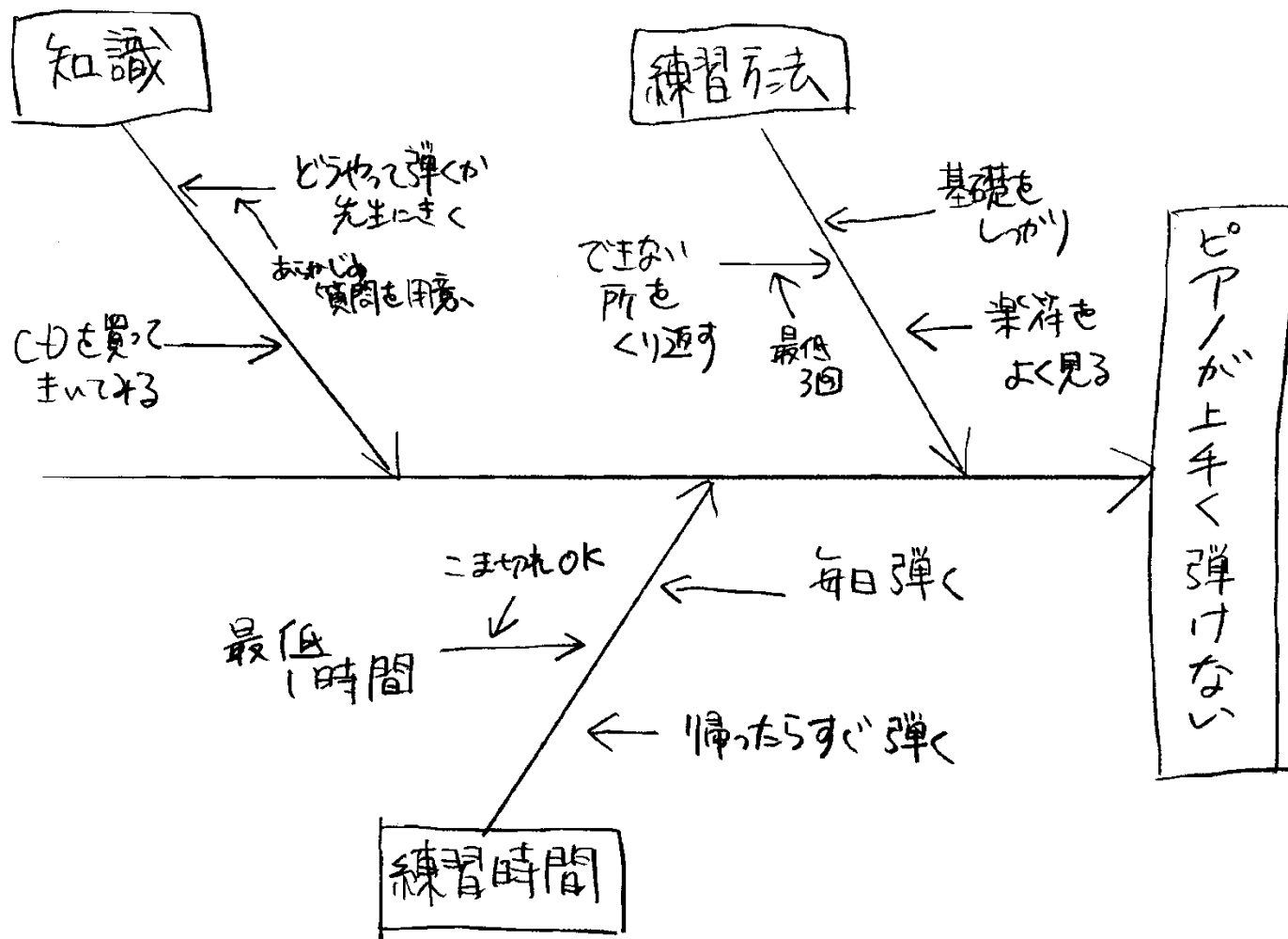
図解の例(1)

• ロジックツリー



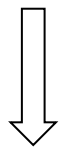
図解の例(2)

- Fish Bone図(特性要因図)



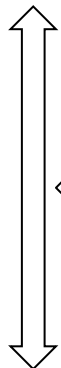
理想

いい点取りたい



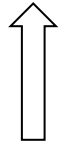
具体的な理想

数学 I で80点



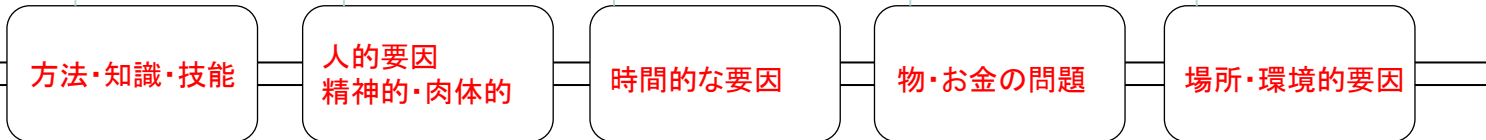
具体的な現実

数学 I が40点



現実

テストの点が悪い



基礎学力を磨く

中学校の内容から復習 — 中学校の問題集を購入

教科書の例題を確実にする — ノートに写してわからない所は質問する

勉強方法の見直し

友達に方法を聞いてみる

10分経ってわからないものは飛ばす

同じ問題を何度も繰り返す

携帯をいじらない — 勉強中は親に預ける

精神力を磨く — 勉強前に1分瞑想

成功したら自分へのご褒美を

勉強する時間帯を工夫 — 帰ったらすぐやる

勉強時間を増やす — 1日3時間

家用の基本問題集を買う

帰りに自習室を利用

方法・知識・技能

人的要因
精神的・肉体的

時間的な要因

物・お金の問題

場所・環境的要因

兄弟がテレビを見ていて集中できない

問題集が家がない

勉強する時間帯が悪い — すぐ眠くなる

勉強時間が少ない — 動画サイトを見てしまう

部活で帰ると8時

つい携帯をいじる — 目の前に携帯がある

すぐあきらめる — あまり集中力がないかも

ま、いいか、と思ってしまう

基礎学力がない

中学校の内容が不十分 — わからなくてもそのまま

教科書の例題がわからない

勉強方法が悪い

勉強方法がわからない

解けないといつまでも考えてしまう

定性的なデータを数値化する(1)

- 評価軸(賛成・中立・反対など)を設定し, 自由記述の内容をそれに当てはめる

例) アンケート調査の結果から, 立場を見極める

- ・とても興味があり, 楽しみです。
→ 賛成
- ・できればやりたくないと思っています。
→ 反対
- ・アイデアがないので、委員さんにお任せします。
→ 中立

定性的なデータを数値化する(2)

- テキストマイニング

- 大量の文字(=テキスト)情報を定量(数値)化, 分析するための手法

- アンケート調査

- TwitterなどのSNS

- 企業や大学などで広く活用されている

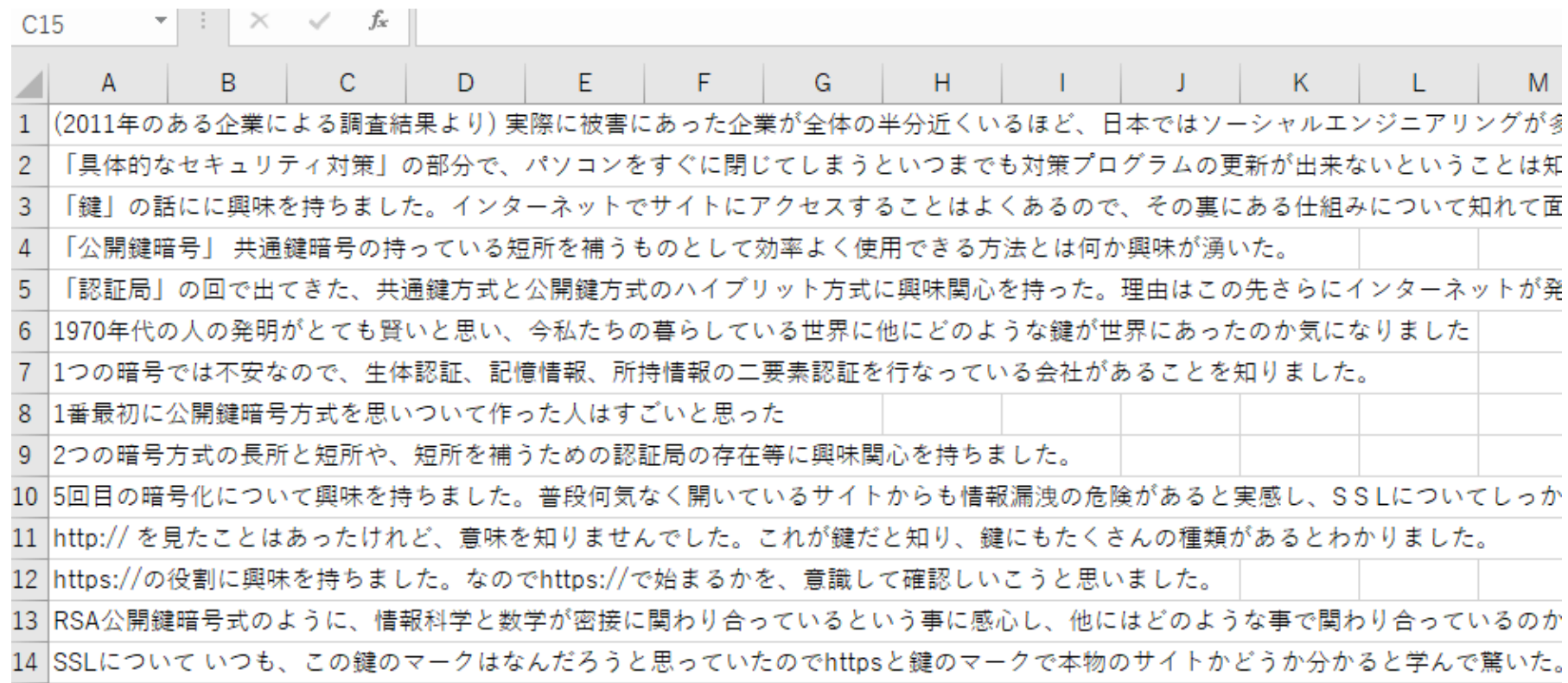
- マーケティング(売れ筋の商品の分析など)

- アンケート結果からの心理的な分析など

- アンケート結果を一目で見るための手法として

(例) 利用するデータ

- 表計算ソフトで扱えるデータ(.xlsxや.csv)を利用することが多い



The image shows a screenshot of a spreadsheet application. The active cell is C15. The spreadsheet contains 14 rows of text, which is a transcription of a document. The text discusses the author's interest in 'key' (鍵) related security concepts, such as public key cryptography, and their observations on the internet's security landscape in the 1970s and 1980s.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	(2011年のある企業による調査結果より) 実際に被害にあった企業が全体の半分近くいるほど、日本ではソーシャルエンジニアリングが多												
2	「具体的なセキュリティ対策」の部分で、パソコンをすぐに関じてしまうといつまでも対策プログラムの更新が出来ないということは知												
3	「鍵」の話にに興味を持ちました。インターネットでサイトにアクセスすることはよくあるので、その裏にある仕組みについて知れて直												
4	「公開鍵暗号」 共通鍵暗号の持っている短所を補うものとして効率よく使用できる方法とは何か興味湧いた。												
5	「認証局」の回で出てきた、共通鍵方式と公開鍵方式のハイブリット方式に興味関心を持った。理由はこの先さらにインターネットが発												
6	1970年代の人の発明がとても賢いと思い、今私たちの暮らしている世界に他にどのような鍵が世界にあったのか気になりました												
7	1つの暗号では不安なので、生体認証、記憶情報、所持情報の二要素認証を行なっている会社があることを知りました。												
8	1番最初に公開鍵暗号方式を思いついて作った人はすごいと思った												
9	2つの暗号方式の長所と短所や、短所を補うための認証局の存在等に興味関心を持ちました。												
10	5回目の暗号化について興味を持ちました。普段何気なく開いているサイトからも情報漏洩の危険があると実感し、SSLについてしっか												
11	http://を見たことはあったけれど、意味を知らませんでした。これが鍵だと知り、鍵にもたくさんの種類があるとわかりました。												
12	https://の役割に興味を持ちました。なのでhttps://で始まるかを、意識して確認しようと思いました。												
13	RSA公開鍵暗号式のように、情報科学と数学が密接に関わり合っているという事に感心し、他にはどのような事に関わり合っているのか												
14	SSLについていつも、この鍵のマークはなんだろうと思っていたのでhttpsと鍵のマークで本物のサイトかどうか分かると学んで驚いた。												

単語の出現頻度

■ 名詞	スコア	出現頻度	■ 動詞	スコア	出現頻度
鍵	298.90	139	思う	15.45	167
興味	133.10	115	持つ	48.12	136
暗号	372.59	72	知る	14.91	79
情報	45.18	69	使う	7.56	58
方式	176.10	51	できる	3.91	56
公開鍵暗号	605.71	46	驚く	23.57	33
仕組み	108.72	42	開ける	10.74	27
対策	47.48	42	守る	13.44	24
公開鍵暗号方式	461.30	36	感じる	2.57	22
暗号化	278.39	36	いく	0.85	21
証明	77.63	36	閉める	25.64	19
公開	32.58	36	考える	1.02	19
認証局	432.87	34	聞く	0.88	19
方法	23.81	34	わかる	0.70	19
関心	77.97	33	しまう	0.51	18

■ 形容詞	スコア	出現頻度	■ 感動詞	スコア	出現頻度
面白い	2.55	26	---	---	---
すごい	1.50	26	---	---	---
多い	0.49	13	---	---	---
詳しい	2.11	10	---	---	---
難しい	0.65	9	---	---	---
良い	0.11	9	---	---	---
うまい	0.49	7	---	---	---
興味深い	5.97	6	---	---	---
いい	0.03	6	---	---	---

単語どうしの係り受けなど

■ 名詞 - ■ 動詞

名詞 - 動詞	スコア	出現頻度
興味 - 持つ	76.69	102
関心 - 持つ	6.35	29
情報 - 守る	7.28	13
鍵 - 持つ (否: 10.00%)	0.80	10 (否: 1)
素因数分解 - 使う	1.22	8
鍵 - かける	7.00	7
暗号化 - 持つ	0.41	7
鍵 - 渡す (否: 16.67%)	6.00	6 (否: 1)
鍵 - 作る (否: 16.67%)	2.47	6 (否: 1)
証明 - できる	0.74	6
暗号 - 使う	0.71	6
部分 - 持つ	0.31	6
鍵 - 使う	0.51	5
方式 - 使う	0.51	5
被害 - 遭う (否: 25.00%)	3.33	4 (否: 1)

「コンピュータ」と「人間」の役割

コンピュータ

人間



データ処理

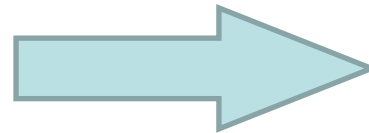


解釈

■ 名詞	スコア	出現頻度	■ 動詞	スコア	出現頻度
鍵	298.90	139	思う	15.45	167
興味	133.10	115	持つ	48.12	136
暗号	372.59	72	知る	14.91	79
情報	45.18	69	使う	7.56	58
方式	176.10	51	できる	3.91	56
公鍵暗号	605.71	46	驚く	23.57	33
仕組み	108.72	42	開ける	10.74	27
対策	47.48	42	守る	13.44	24
公鍵暗号方式	461.30	36	感じる	2.57	22
暗号化	278.39	36	いく	0.85	21
証明	77.63	36	認める	25.64	19
公鍵	32.58	36	考える	1.02	19
認証局	432.87	34	開く	0.88	19
方法	23.81	34	わかる	0.70	19
関心	77.97	33	しまう	0.51	18

■ 形容詞	スコア	出現頻度	■ 感動詞	スコア	出現頻度
面白い	2.55	36	---	---	---
すごい	1.50	36	---	---	---
多い	0.49	13	---	---	---
詳しい	2.11	10	---	---	---
難しい	0.65	9	---	---	---
良い	0.11	9	---	---	---
うまい	0.49	7	---	---	---
興味深い	5.97	6	---	---	---
いい	0.03	6	---	---	---

データ提供



興味関心を持ったり、面白い、と感じている人がとても多いようだ。
この内容は意味があるに違いない！

演習

- 実際にワードクラウドを作り、解釈を試みよう！