

2016年10月

【ミニセミナー】

上司を説得できるデータ分析・資料作成3つのポイント

Tokyo
Business
Career Lab

東京ビジネスキャリアラボ

本日の内容

- 自己紹介(氏名、職業、今日の気分、将来の夢)
- ミニセミナー・ミニワーク(20～30分程度)
- フリーディスカッション(30分程度)
- 解散

本日のゴール

- データを活用した資料を
作成する際のポイントがわかる
- データ分析リテラシーを高めるきっかけを得る

なぜ今、データ分析なのか？

大量のデータを誰でも手軽に処理(解析)できるようになり、誰もがデータ分析のパワフルさを享受できるようになっている。

- だれでもPCを使えるような時代⇒過去には紙とペンによるデータ分析がPCで誰でもできる。
- ITによる業務のシステム化が進み、大量のデータを簡単に収集できるようになった。



第一刷発行: 2013年1月24日
著者: 西内啓(にしうち ひろむ)
東京大学医学部卒(生物統計学)
東京大学医学系研究科医療コミュニケーション学分野助教、ハーバードがん研究センター客員研究員などを経て現在は、現在はデータに基づいて社会にイノベーションを起こすための様々なプロジェクトでコンサルティングを行う。

データ分析を活用する目的

- データで自分の主張を示す

⇒科学やマーケティングにおいては、仮説を立証するとき。経営的な判断をする際に決断するための根拠とするなど。

- データから、事実や知見を読み取る

⇒ものづくりの現場では、異常を察知するために。仕事の改善に役立つ洞察を得るなど。

基本的なデータ分析の流れ

STEP1: どんなデータ分析をするか決める

ポイント① 誰のため？何のため？を考える

STEP2: データを収集、整理する

ポイント② データを加工して、切り口を見出す

STEP3: 意味を解釈する

ポイント③ 数字に意味と解釈をつける

ポイント①

誰のため？何のため？を考える

ターゲットと目的を確認する

そもそも、相手の期待を理解しないと的外れの資料になる。

・ターゲットのプロファイル

人物像：どんなことに興味がある？どんな情報を聴きたい？

保有情報：用語はわかるか？理解のレベルは？

・目的の明確化

①どんな行動をとってもらいたいか？

・何かを決めてもらうのか、情報を共有するものなのか？

②何を理解してもらいたいか？



何をどのように伝えるかの方針が決まる

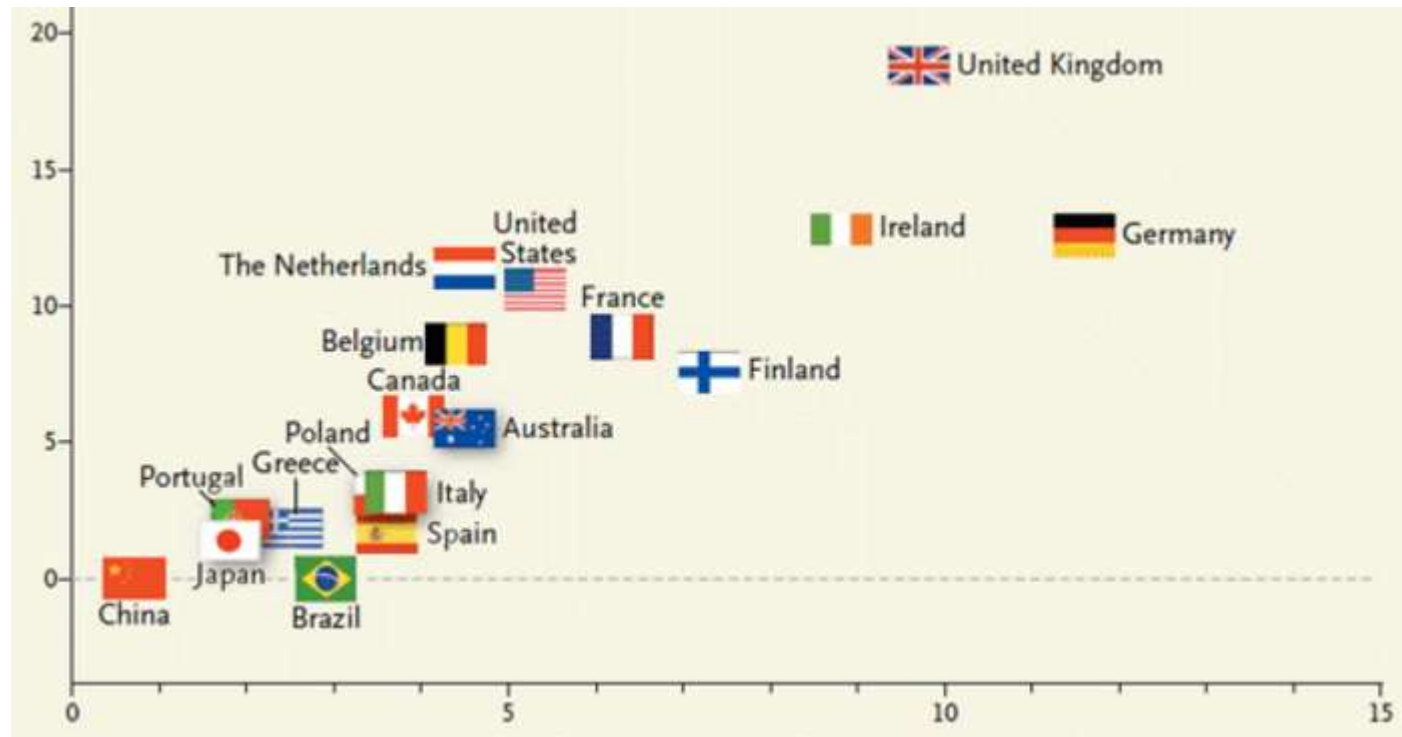
ポイント②

データを加工して、切り口を見出す

ここで問題

チョコレートをよく食べる国ほどたくさんのノーベル賞の受賞者を輩出しているとうデータは正しいか？

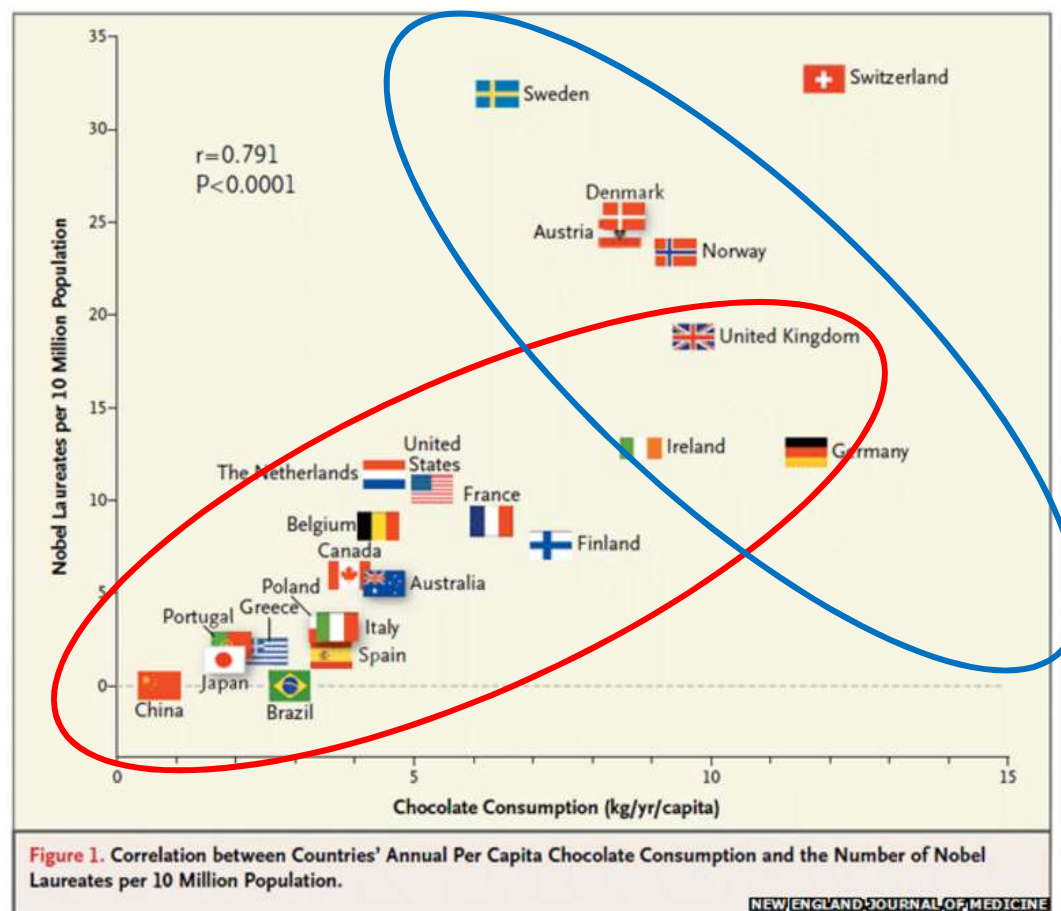
ノーベル賞の受賞者数



チョコレートの消費量 (kg/年)

データの見方で主張が変わる

ノーベル賞の受賞数とチョコレートの消費量



データを加工して、切り口を見出す

同じデータでも、切り口や見せ方を変えることで、違った意味を見出すことができる。

・集計基準を見直す

曜日別、月別、製品別、店舗別などの、集計の切り口を変える
⇒統計の用語では「層別」という

EXCELの「フィルター」機能や「ピポッド
テーブル」を使うと簡単に層別が可能

・比率に置き換えたり、単位を見直す

前年度との比較の比率を見たり、一人あたりに換算したりする。
⇒例えば、数字自体は大きくない売上の商品も、前月比のデータでみると伸び率が大きいことがわかれば有望だと判断できる

ポイント③

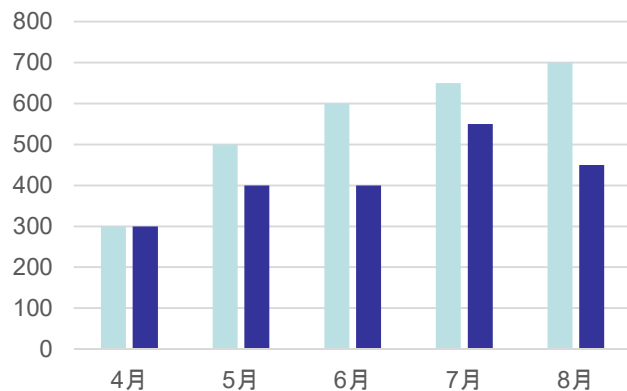
数字に意味と解釈を付ける

コメントで理解を早める

図表にコメントを入れると、主張を理解してもらいやすくなる。
主張についての賛成、反対の意見も出てきやすい。

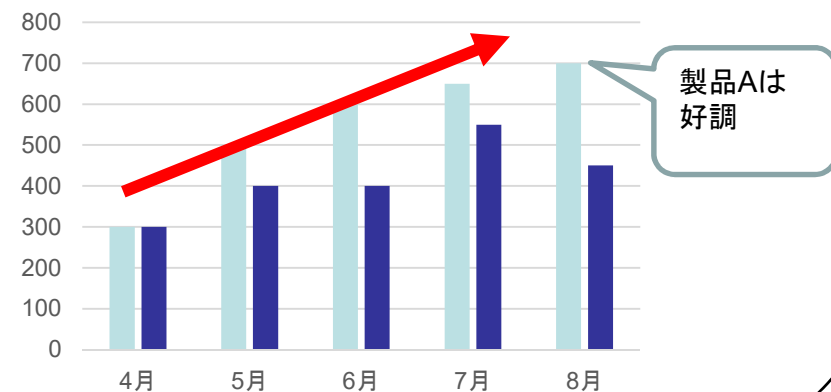
数字のみの表・グラフ

	4月	5月	6月	7月	8月	合計
製品A	300	500	600	650	700	2750
製品B	300	400	400	550	450	2200



傾向とコメントを入れた

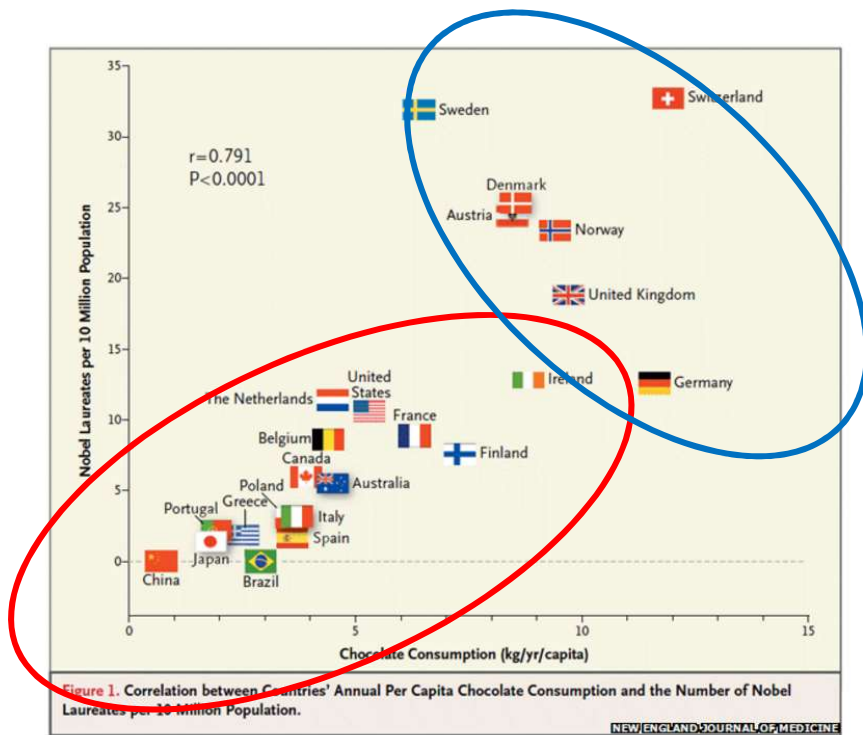
	4月	5月	6月	7月	8月	傾向
製品A	300	500	600	650	700	上昇 +
製品B	300	400	400	550	450	横ばい ・



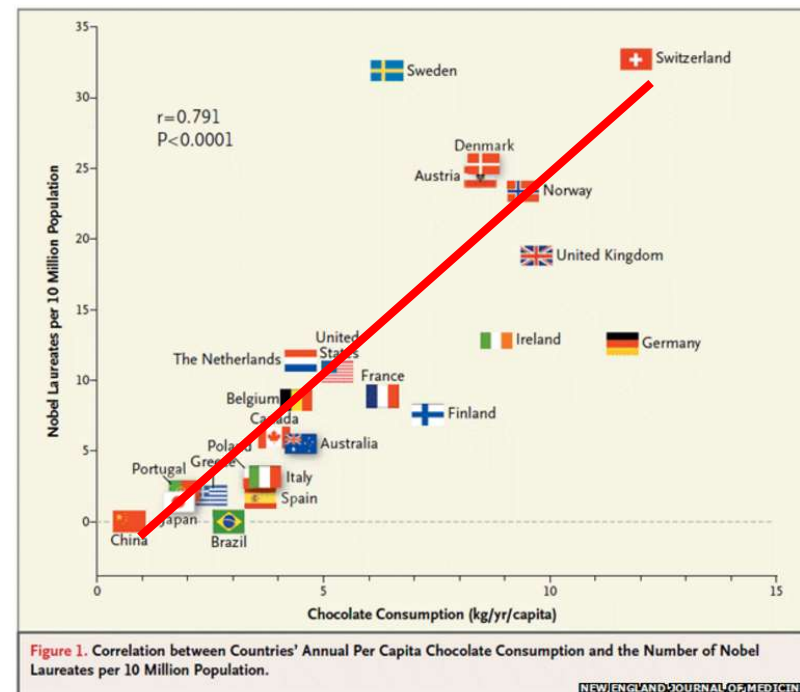
グラフに意味を込める

グラフに補助的な線や図形を入れることで、主張したいことを印象付けることができる。

例) ノーベル賞の受賞数とチョコレートの消費量



2つのグループの印象を付ける



直線的な関係がある印象を付ける

注意が必要なデータの見せ方

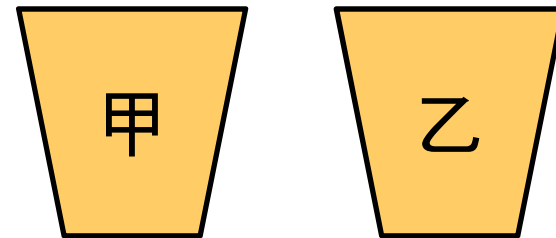
この例は、あまりにも恣意的になるため注意が必要。
基本的には使わない方がベター。

- グラフの目盛りを変える
0からではなく、都合のよい数字からのグラフとする。
- グラフを立体的にする
3Dのグラフ図を利用すると平面よりも印象が変わる
- グラフの縦横比を変える
横長だと、縦の変化が緩やかに見える

ミニワーク&参考 確率と相関について

ミニワーク1(確率)

「Aさんは2種のビールの銘柄の味の違いがわかると言っている。そこで、Aさんに目隠しをして官能検査をしてもらうことにした。1回目があった。2回目もあつた。3回、4回・・・と何回あつたら彼の舌を信じるか？



回答: _____ 回

ミニワーク1(確率)

Aさんが味の違いがわからないとすると、当るのはまぐれと考えられるので、まぐれで当る確率は、

- 1回目 : 1/2 (50%)
- 2回連続 : 1/4 (25%)
- 3回連続 : 1/8 (12.5%)
- 4回連続 : 1/16 (6.25%)
- 5回連続 : 1/32 (3.125%)

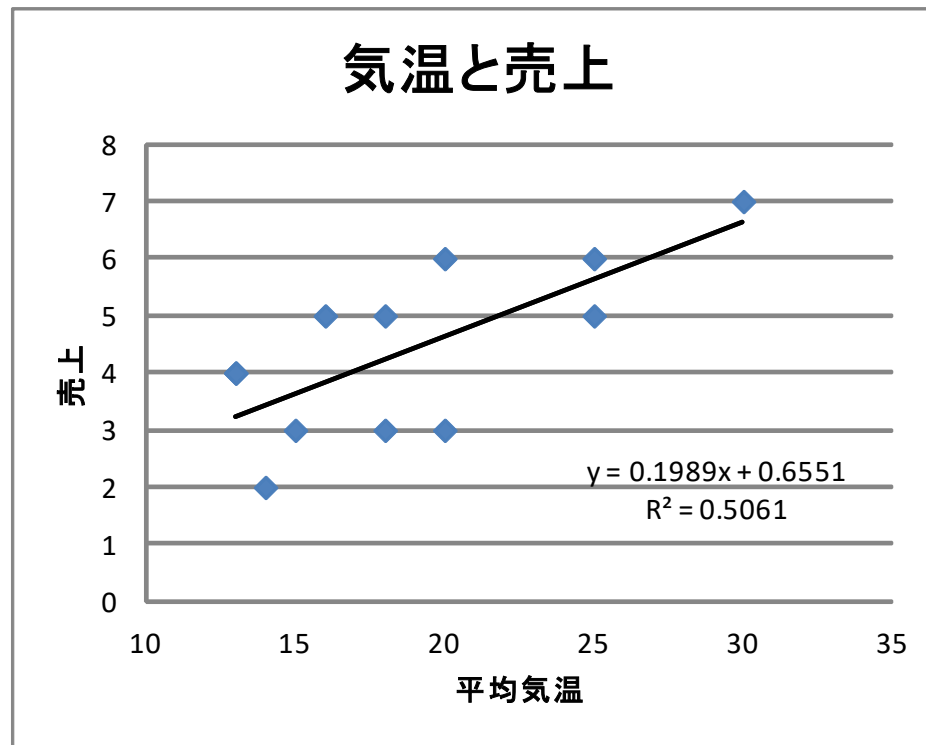
となる。そこで、判断基準を「まぐれ当りの可能性はゼロではないが、5%を切ったらまぐれ当りでないと判断する」($p < 0.05$)と慎重な人はより低い水準で判断すること。

ミニワーク3(相関)

問題

飲料の売り上げは季節(気温)と関係がありそうか？

月	平均気温	売上(百万)
1月	13	4
2月	15	3
3月	16	5
4月	18	3
5月	20	6
6月	20	3
7月	25	6
8月	30	7
9月	25	5
10月	18	5
11月	14	2
12月	13	4

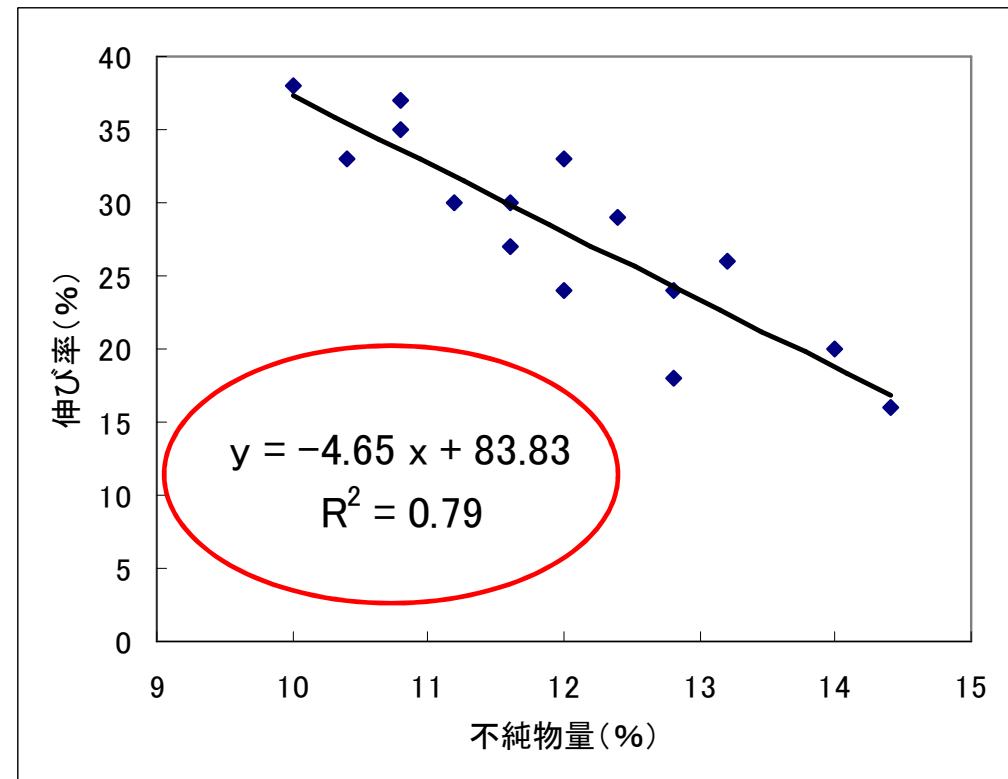


回帰分析＝相関の強さを調べる

作成手順

1. データを集める(一般的にn=30以上が望ましい)
2. グラフの座標軸を作る
3. プロットする
4. EXCELの場合、回帰直線と相関係数が平易に算出できる

⇒意味を抑えておくこと
とは重要



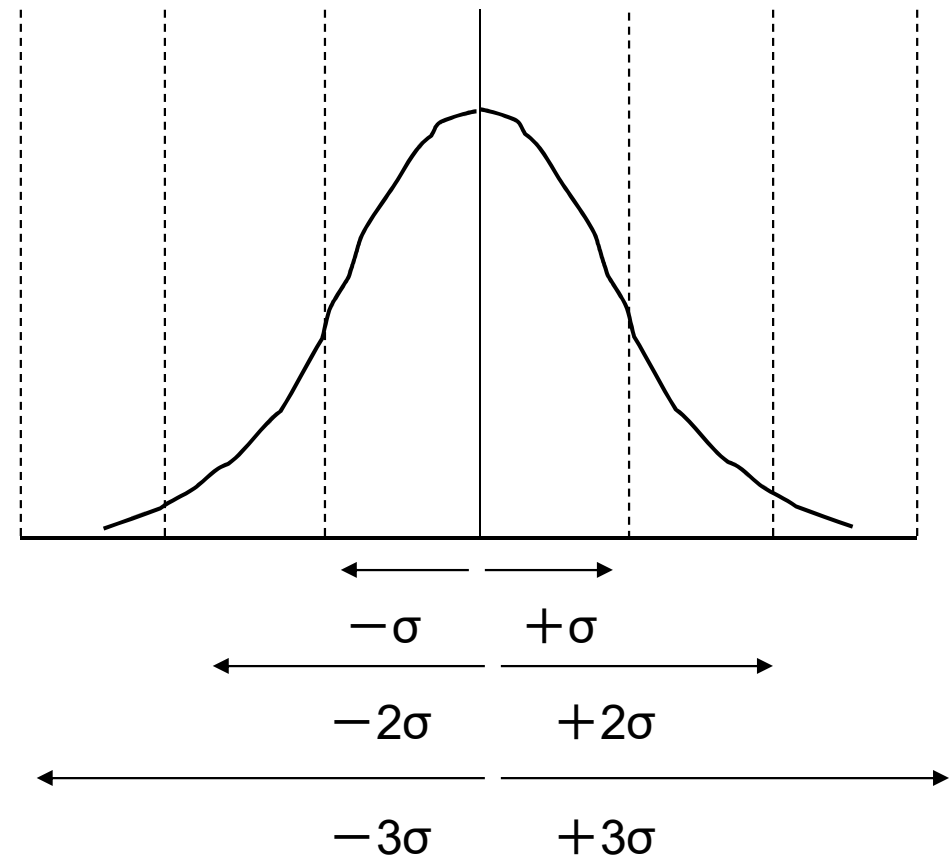
相関係数

- ・2つの変数xとyとの関係の強さを表す尺度として、
相関係数(R)がある： $R = \pm \sqrt{|AC|}$
- ・xとyとの相関の有無や強さの度合いは、相関係数を求め、検定を行うことで統計的に判断される
- ・Excelのグラフで自動計算されるR2は、寄与率(決定係数)に相当
- ・寄与率(R^2)
変数yの変動のうち、変数xが変化したために影響を受ける割合を示す尺度で、 $R = 0.5$ とはyの変動のうち、xの影響によるものが $R^2 = 0.25$ 、つまり25%であるということ(75%はx以外の要因)
- ・寄与率は0.5以上(相関係数0.7以上)でない通常使えない

<参考> 標準偏差の表すもの

正規分布の場合、平均値(\bar{X}) \pm 標準偏差(σ)の範囲内に全データの68.27%が含まれる

$n\sigma$	平均 $\pm n\sigma$ の範囲に含まれる率
σ	68.27%
2σ	95.45%
3σ	99.73%
4σ	99.9937%
5σ	99.999943%
6σ	99.99999980%



一般的に工程管理では $\pm 3\sigma$ を用いる
 $\Rightarrow 99.73\%$ は規格内と予想される