

■ データの種類 … データには大きく分けて、質的データと量的データの区別がある。

データの種類によって整理・分析に用いる方法が違う。どのタイプのデータかを頭に置いておくことが大切である。

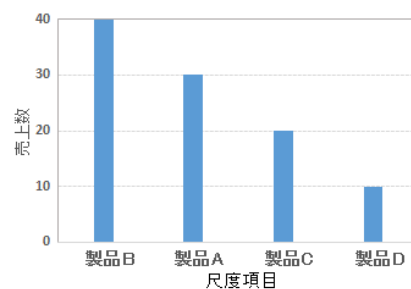
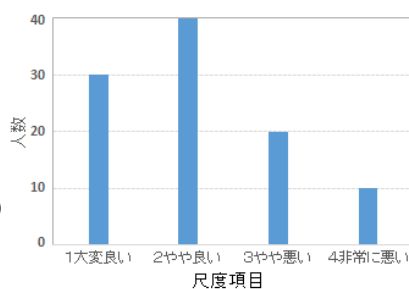
1. 質的データ … 分類項目であって、数量としての意味のないもの

- 1) 順序尺度：分類の順序には意味のあるもの 例：1 大変良い 2 やや良い 3 やや悪い 4 非常に悪い
- 2) 名義尺度：純粋な分類で順序にも意味のないもの 例：性別、職業、名前 など

※ 順序尺度をグラフ化する場合、尺度項目の並び替えは一般的に行わない。

(左図)

一方、名義尺度の場合は、大小関係で並び替えると見やすくなる。(右図)



尺度 (Scale)：物事を測る物差しのこと

2. 量的データ … 数量としての意味のあるもの。

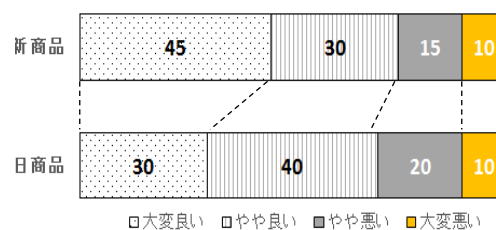
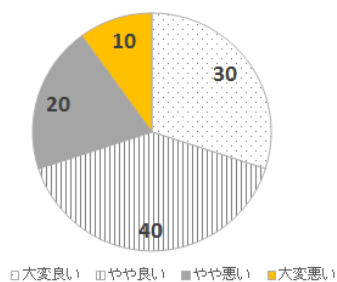
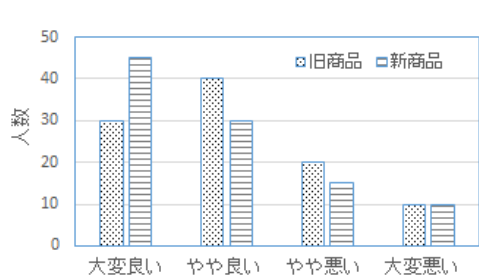
- 1) 間隔尺度：摂氏温度のようにゼロに意味のないもの。
- 2) 比例尺度：長さのようにゼロに意味のあるもの。

※ 上記のほかに、データの分類には離散データと連続データの区別がある。

離散データ：人数のようにとびとびの値をとるもの、 連続データ：身長のように連続的な値をとれるもの

■ 統計グラフの種類と特徴 その1 (質的データの度数や相対度数を表現するためのグラフ)

- 棒グラフ：項目間の度数の大小関係を比較するのに用いられる。集合棒グラフでは、複数の群で大小関係を比較できる。
- 円グラフ：項目の構成割合を省察するのに用いられる。
- 帯グラフ：複数の群での項目の構成割合を省察するのに用いられる。



【データ】

	旧商品	新商品
大変良い	30	45
やや良い	40	30
やや悪い	20	15
大変悪い	10	10
計	100	100

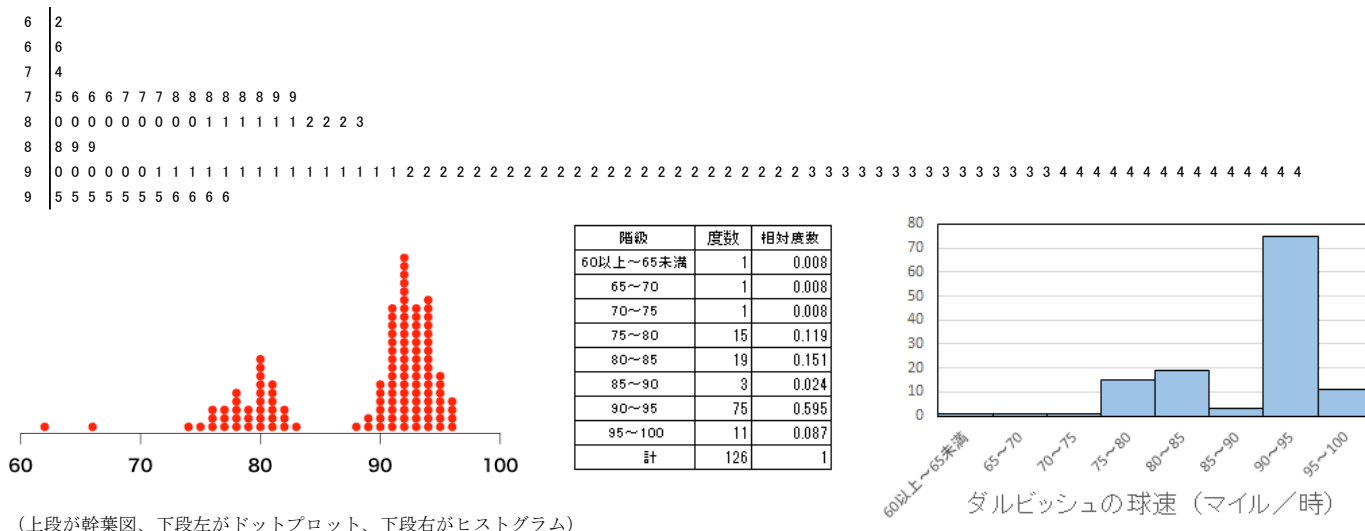
上の3つのグラフは、すべて左の度数分布表のデータに基づいてつくられたものである。3種類のグラフを比較し、グラフの種類とその特徴を確認しよう。

Q 1 種類の商品に関する消費者評価を最もわかり易く伝えるのには、どのグラフがよいか？

Q 新商品に対する消費者評価が旧商品に比べてどの程度変わったのかを、最もわかり易く伝えるのには、どのグラフがよいか？

■ 統計グラフの種類と特徴 その2 (量的データの分布を表す方法)

- ドットプロット：各値の数をドットの積み重ねで表現する。データの数が多いとグラフが大きくなり、大変である。
 - 幹葉図（ミキハズ と読む）：上位の桁を左（幹）に、最下位の桁の数字（葉）を幹の横に並べる。
 - ヒストグラム：量的データの分布を表すのに一般的に用いられるグラフ。度数分布表（データを階級に分けて、階級ごとの度数を表にしたもの。度数／全数で計算される相対度数を示すこともある）をグラフ化したものである。
- 棒グラフに似ているが、棒グラフはその高さがその値を示すのに対して、ヒストグラムはその面積がその値（度数）を示すことに注意。また、ヒストグラムでは、柱と柱の間をつめて描く点も、棒グラフと異なる。

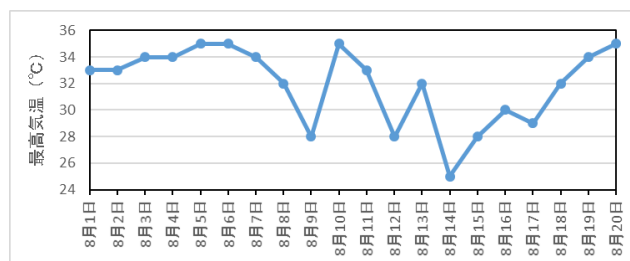


【データ】 ダルビッシュ有投手がメジャーリーグ(2014年5月9日)の試合で投げた全126球の球速 (マイル/時)

91 91 90 93 91 79 92 80 80 93 93 66 93 94 92 94 79 83 94 94 92 80 78 77
 92 76 62 96 81 96 96 93 91 92 91 91 93 80 94 94 94 92 92 78 78 81 92 94
 92 82 94 93 93 92 81 81 95 95 78 82 95 95 91 80 95 81 82 96 94 93 92 92
 94 92 89 95 92 92 78 78 92 92 91 95 92 92 90 93 94 92 94 93 81 80 93 93
 91 75 90 77 91 91 76 90 90 92 94 74 90 91 92 76 77 91 88 91 89 80 91 92
 92 80 94 93 80 93

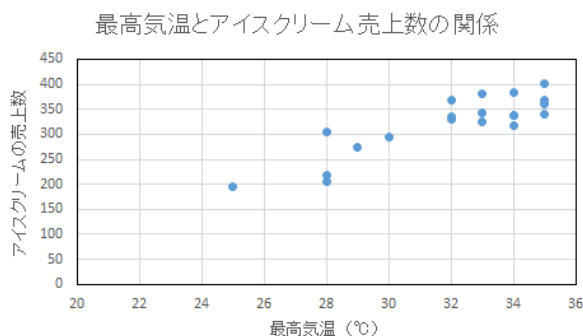
■ 統計グラフの種類と特徴 その3 (時系列データの変化や2次元データの関係を表す方法)

- 折れ線グラフ：経時変化を表すときに多く用いられる。横軸に時間（時刻）、縦軸に注目している値をとる。
- 散布図：横軸と縦軸にそれぞれ異なる量を取り、2種類のデータの関係を表すのに用いられる。



【データ】 最高気温とある店のアイスクリームの売上数

日付	最高気温	売上数	日付	最高気温	売上数
8月1日	33	382	8月11日	33	342
8月2日	33	324	8月12日	28	205
8月3日	34	338	8月13日	32	368
8月4日	34	317	8月14日	25	196
8月5日	35	341	8月15日	28	304
8月6日	35	360	8月16日	30	294
8月7日	34	339	8月17日	29	275
8月8日	32	329	8月18日	32	336
8月9日	28	218	8月19日	34	384
8月10日	35	402	8月20日	35	368



<参考>

統計グラフには、様々なものがあります。
 Web ページ「なるほど統計学園高等部（総務省統計局による統計学習サイト）」には、各グラフの特徴やエクセルによるグラフ作成法の説明が載っています。

2020 課題研究基礎 統計学① 事前アンケート&確認テスト（宿題）

今回は、研究を進めるための強力な武器となる統計学について学びます。少しだけ予習が必要です。必ず予習を行い、アンケートと問題に答えて下さい。

□ アンケート Q1~Q4 の質問に、下の①~④で回答してください。

- ① 前からよく理解している ② 資料を読んで理解することができた ③ 資料を読んだが、まだ理解できない
④ 資料を読んでおらず、理解できていない

Q1 データの種類には、質的データと量的データがあることを理解していますか？

Q2 ドットプロットや幹葉図とはどのようなものかを理解していますか？

Q3 度数分布表とヒストグラムがどのようなものかを理解していますか？

Q4 散布図とはどのようなものかを理解していますか？

□ 問題 Q5~Q7 の間に答えてください。

Q5 投球速度は、どの種類のデータか？

- ① 間隔尺度で離散データである ② 間隔尺度で連続データである
③ 比例尺度で離散データである ④ 比例尺度で連続データである

Q6 新商品に対する消費者評価が旧商品に比べてどの程度変わったのかを、最もわかり易く伝えるグラフは次のどれか？

- ① 横軸に消費者評価の4項目を、縦軸に各項目の人数をとって、新商品と旧商品について比較をする集合棒グラフ
② 新商品に対する消費者評価の割合（人数）を示す円グラフと旧商品に対するそれを示す円グラフを並べて示す
③ 新商品に対する消費者評価の割合（人数）を示す帯グラフと旧商品に対するそれを示す帯グラフを並べて示す

Q7 資料の「アイスクリームの売上数のデータ」を使って立てた予想として、最も適切でないものを1つ選べ。

- ① 翌日の最高気温の予報が26℃であれば、売上数は150~250と予想できる
② 翌日の最高気温の予報が30℃であれば、売上数は250~350と予想できる
③ 翌日の最高気温の予報が40℃であれば、売上数は400~500と予想できる

2020 課題研究基礎 統計学 事前プリント（宿題） 解答用紙

1年1組（ ）番（ ）

アンケート	Q1	Q2	Q3	Q4
問題	Q5	Q6	Q7	

2020 課題研究基礎 統計学 事前プリント（宿題） の答え

Q5 ④ （球速 0 は止まっている状態で、ゼロ本来の意味「無」を示すから「比例尺度」である。また、速度は連続した値であるから「連続データ」である。与えられたデータが整数値なのは、測定器具の精度のためなどの理由であり、これを理由に離散データと考えてはいけない。）

Q6 ③ （複数の群での項目の構成割合を比較するには、帯グラフが適切である。）

Q7 ③ （データから得られた傾向を使って予測する場合、予測可能な範囲を考えておく必要がある。今回のデータは気温 25℃～35℃での売上高を調べたものであり、この範囲外で

の予測（外挿）には慎重でなければならない。グラフは範囲内で右上がりであるが、32℃～35℃では売上数の増加率は鈍っており、40℃でも右上がりのグラフと予測することはできない。余りに暑いと、シャーベットやかき氷の方が売上げを伸ばすかも…？）

