

データの視覚化（さまざまなグラフ）

与えられたデータ（標本）の性質を視覚的に把握する方法として、データのグラフ化があります。グラフにはさまざまな種類があり、そのグラフで何を表現したいかにより、適切なグラフは変わってきます。その為、それぞれのグラフの特性を理解し、適切なグラフを用いることが大切になります。ここでは、グラフの中でも特に基本的かつ重要で、よく用いられる**棒グラフ**、**折れ線グラフ**、**円グラフ**について、それぞれのグラフの特性、グラフからの情報の読み取り方、グラフ作成時の注意点について説明します。

棒グラフ

棒グラフは、長方形の棒の長さで何らかの値の大きさを表したグラフです。通常は、縦軸にデータの取り得る値をとります。棒グラフはデータの大小を比較するのに適しています。

例えば、50 人の学生のある教科に対する成績（A, B, C, D, F の 5 段階評価）に関するデータが表 1 のように与えられているとします。

表 1 学生 50 人の成績データ

成績	A	B	C	D	F
学生数	8	15	12	10	2

このデータを棒グラフで表すと、図.1 のようになります。このグラフからは、例えば、次のようなことを読み取ることができます。

- 最も多いのは成績が B の学生。
- 最も少ないのは成績が F の学生。
- 成績が B の学生は成績が A の学生の約 2 倍。

棒グラフを用いる場合に注意する点について説明します。まず、図 2 を見て下さい。2 つのグラフは同じデータをもとに作成した棒グラフですが、左側のグラフでは縦軸の一部が省略されており、原点での値が 1430 となっています。この例では、本来ほとんど差がないデータにも関わらず、A は他のものの 2 倍以上大きいというように、間違った印象を与えてしまいます。このように、縦軸の一部を省略すると、データの大小関係について間

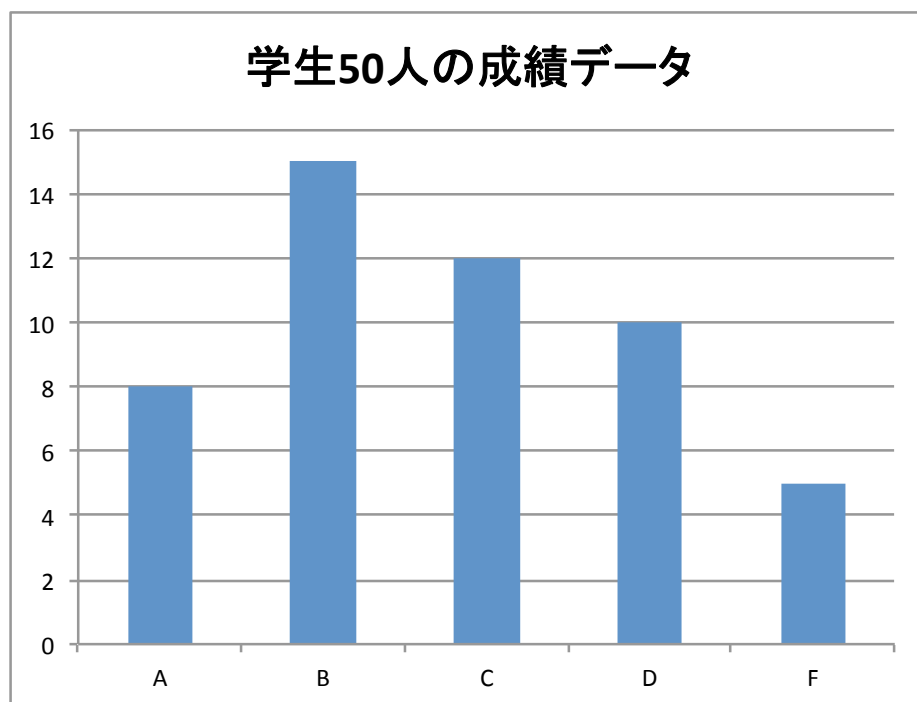


図1 棒グラフの例

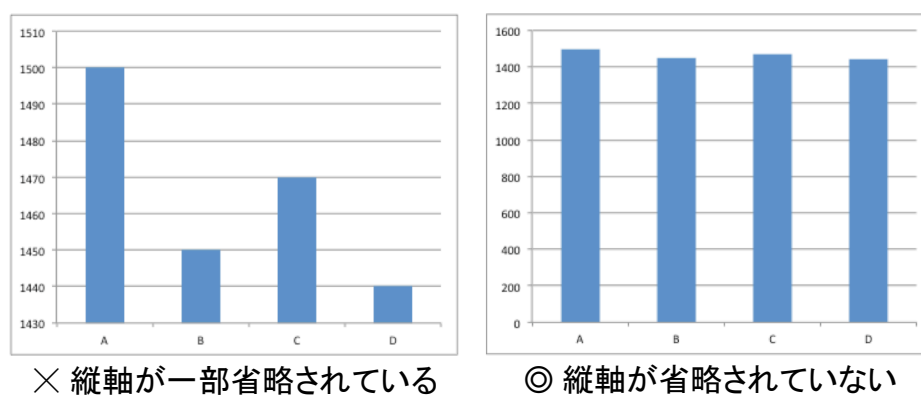


図2 縦軸の一部が省略されたグラフと省略されていないグラフ

違った印象を与えてしまいます。縦軸の一部をむやみに省略しないようにしましょう。

また、棒グラフを作成する際に、横軸の変数は質的データでなくてはいけません。横軸の変数をどのような順番で並べるかについては特に決まりはありません。状況に応じて、分かりやすい形で並べるのが良いでしょう。例えば、データの大きい順・名前の50音順・

アンケート調査での質問票の順・時間の順などが考えられます。

円グラフ

円グラフは、円を全体として、各項目が全体に対して占める構成比を扇形で表したグラフです。扇型の面積により構成比の大小を比較することができます。

例えば、200 人の人に好きな動物に関するアンケートをとった所、表 2 のようなデータが得られたとします。

表 2 好きな動物に関するアンケート結果

動物	犬	猫	鳥	その他
回答数	94	62	26	18

このデータを円グラフで表すと、図 3 のようになります。このグラフからは、例えば、次のようなことを読み取ることができます。

- 犬が全体の中のほぼ半分を占めている。
- 猫が全体の約 3 分の 1 を占めている。
- 犬，猫，鳥を足すと全体の 9 割以上を占めている。

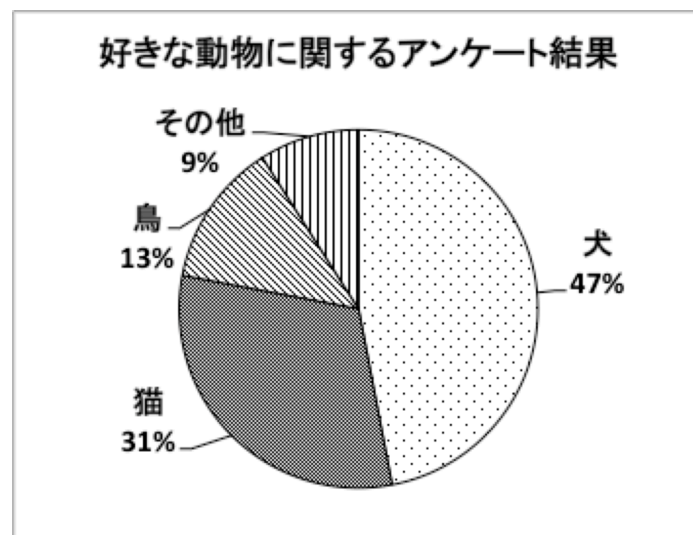


図 3 円グラフの例

円グラフを用いる場合に注意する点について説明します。円グラフは全体に対する構成比を表したグラフですので、全体で 100 % になるようにしなければいけません。その為、幾つかの項目を除いたり、複数回答を認めるようなアンケートの結果表示には用いることができません（図 4）。

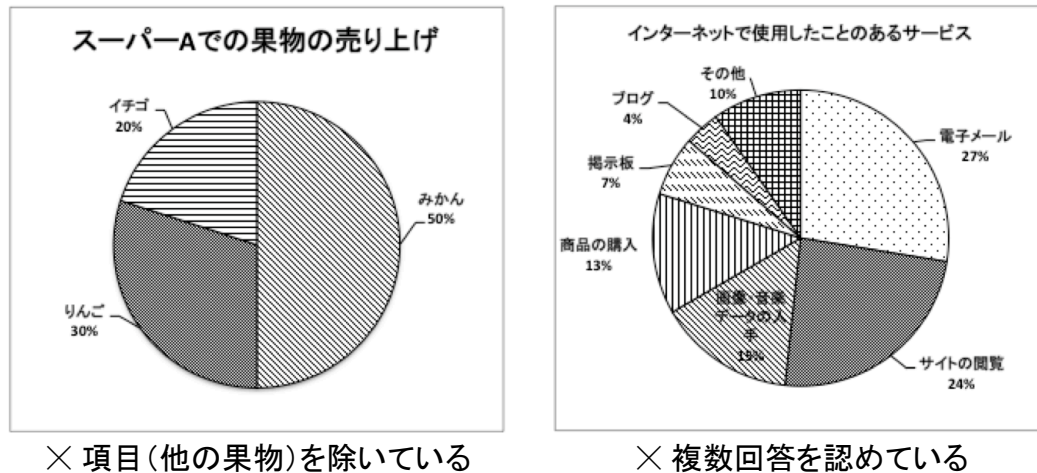


図 4 誤った円グラフの例

また、項目を並べる順番については、時計の 12 時の位置から、構成比の大きい順に並べるのが一般的です。このとき、「その他」に当たる項目については、どんなに大きくても一番最後に示すようにします。ただし図 5 のような例外もあります。これは、ある設問に対して、賛成、どちらかと言えば賛成、どちらかと言えば反対、反対の 4 項目で回答してもらった結果です。このような場合では、全体として、賛成のグループと反対のグループのどちらの方が多いかに興味があります。その為、値の大きい順に並べるのではなく、賛成のグループ、反対のグループをそれぞれまとめて表示したほうがわかりやすいグラフになります。

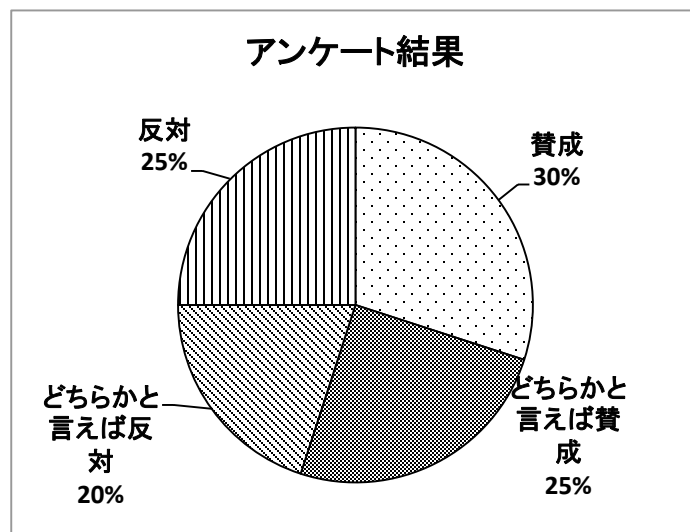


図 5 項目に順番がある場合の円グラフ

折れ線グラフ

折れ線グラフは、横軸に時間（日にち、時間、分、・・・）、縦軸にデータ量を取り、それぞれのデータを折れ線で結んだグラフです。折れ線グラフは、時間によるデータの変化を見るのに適したグラフです。

例えば、ある学生が数学の問題集を毎日解き、学習を開始した日から正しく解けた問題の数を記録した所、表 3 のようなデータが得られたとします。

表 3 解けた数学の問題の数

経過日数	1 日	2 日	3 日	4 日
問題の数	3	4	13	15

このデータを折れ線グラフで表すと、図 6 のようになります。このグラフからは、例えば、次のようなことを読み取ることができます。

- 日が進むにつれて数が増えている。
- 2 日から 3 日にかけての伸び率が大きい。

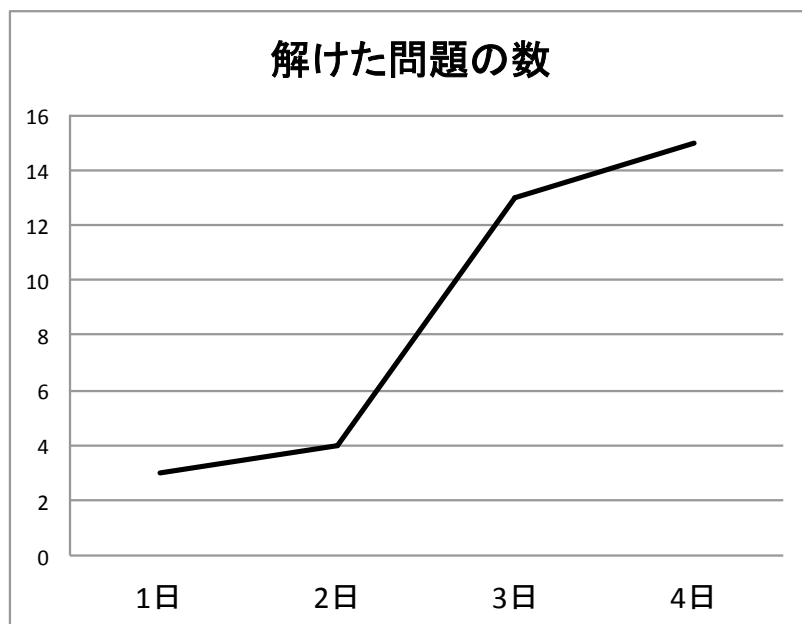


図 6 折れ線グラフの例

折れ線グラフを用いる場合に注意する点について説明します。折れ線グラフは、時間とともにデータがどのように変化するかを見るグラフですので、基本的に横軸は時間を取るようにします^{*1}。横軸の並び順に意味がないようなデータ（名義尺度）については、折れ線グラフよりも棒グラフを用いるべきです（図 7）。

^{*1} 横軸が時間以外の場合もあります。例えば、緯度が変わった時の平均気温の変化を調べるような場合には、横軸に緯度を取り、縦軸に気温を取ります。

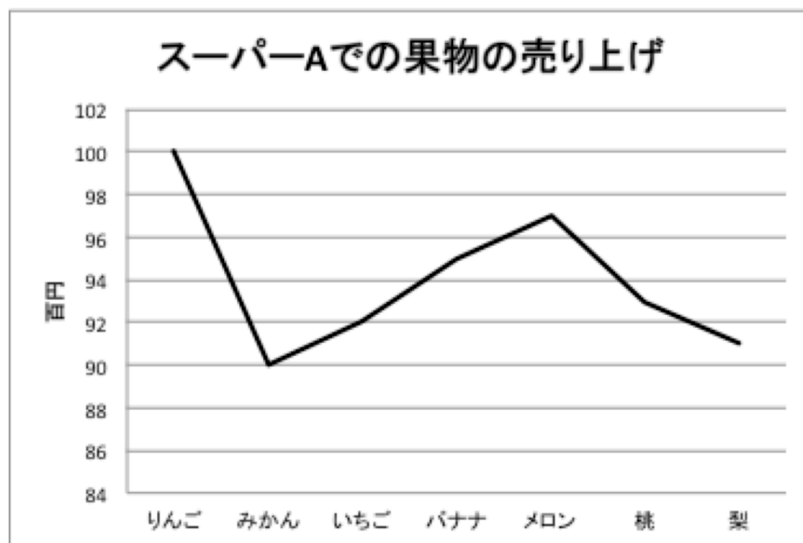
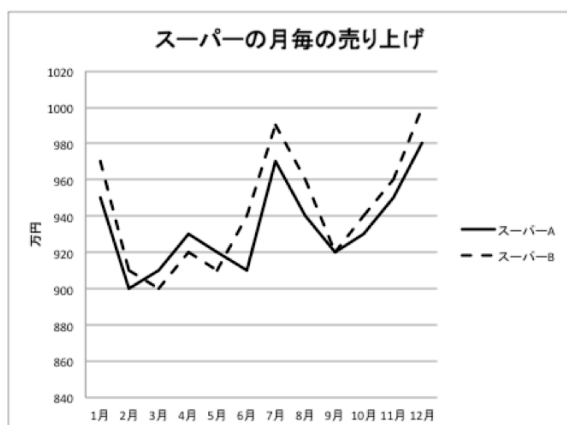
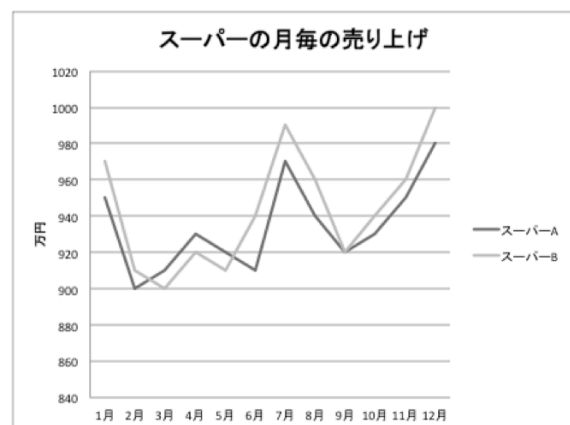


図7 折れ線グラフよりも棒グラフが適している例

また、複数の線を1つのグラフに同時に描く場合は、線のスタイルを変えたり、線の色を変えるなど、それぞれの区別が付きやすいようにしましょう（図8）。



線のスタイルを変える



線の色を変える

図8 2つの折れ線による折れ線グラフの例

棒グラフのときには「縦軸をむやみに省略しない」ということを注意しましたが、折れ線グラフの場合は、必要であればこれを行っても構いません。実際、図8では縦軸の一部が省略され、原点の値が0にはなっていません。これは、折れ線グラフにより表現したい

のが、各項目の絶対的な大きさではなく、時間による変化具合にあるためです。図8と同じデータを、縦軸を省略せずに、原点の値を0にして折れ線グラフを描くと、図9のようになります（ここでは縦軸の目盛の幅も変えています）。グラフの下側に意味のない空白ができており、見難くなっています。折れ線グラフの場合、必要に応じて、縦軸の原点の値を変更して構いません。ただし、縦軸の目盛の幅の設定によっては、本来ほとんど変化していないデータを変化しているように見せる（またはその逆）ことも出来るため、やはり注意が必要です。

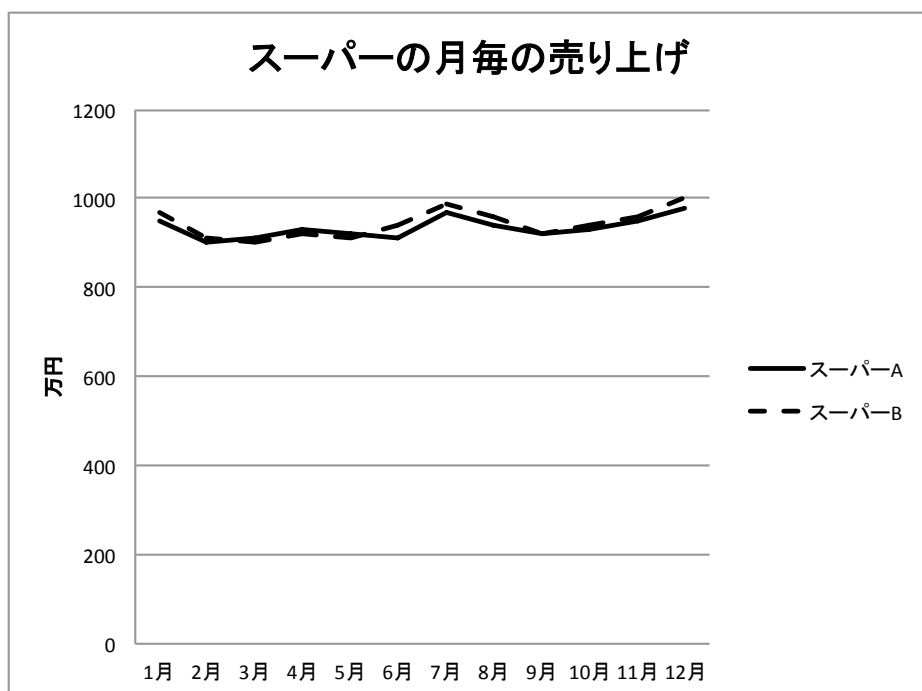


図9 図8と同じデータを縦軸の目盛を省略せずに描いた折れ線グラフ