# データの分析と知識発見

Introduction to Data Analysis

# 今回の構成

Rにおける基本的な
グラフの書き方を学ぶ

関数のグラフの書き方を学ぶ

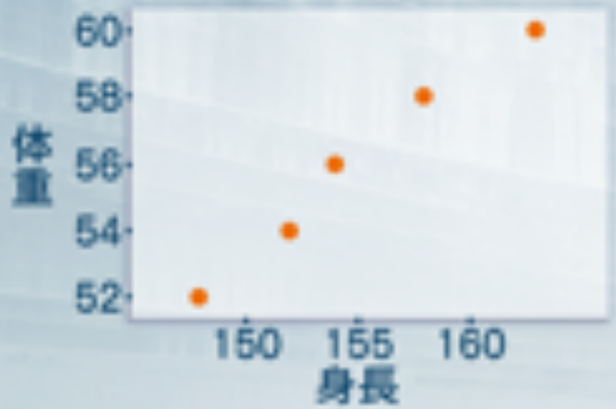
凡例や軸ラベルなどの
要素の書き方を学ぶ

# データの分析と知識発見

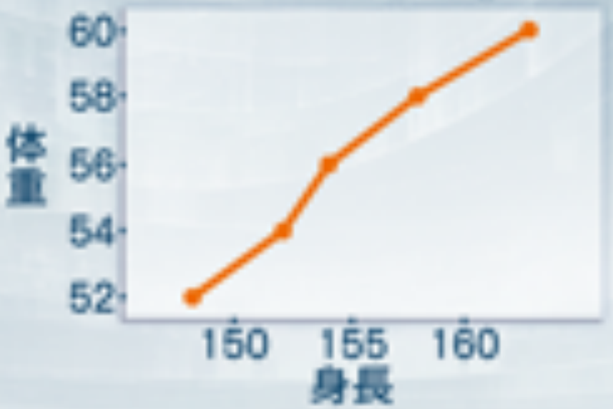
Introduction to Data Analysis

# グラフ

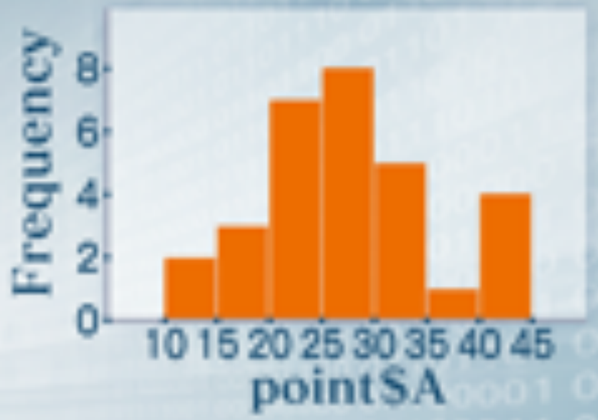
## 散布図(plot)



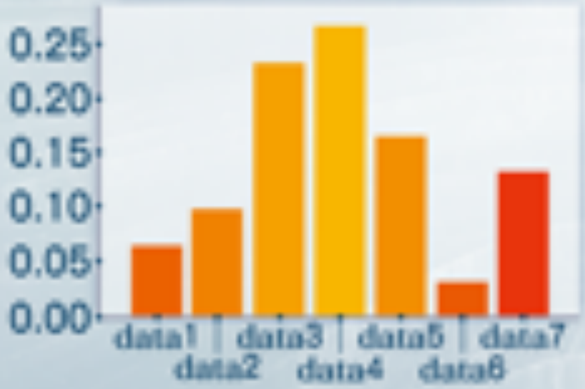

## 折れ線グラフ(plot)




## ヒストグラム(hist)



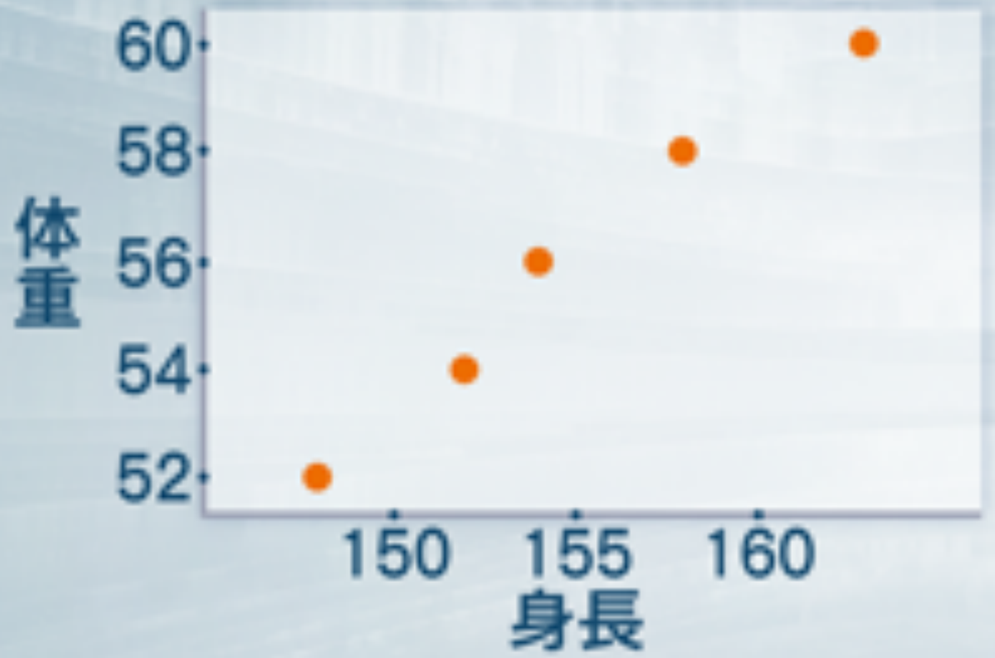

## 棒グラフ(barplot)




## 円グラフ(pie)

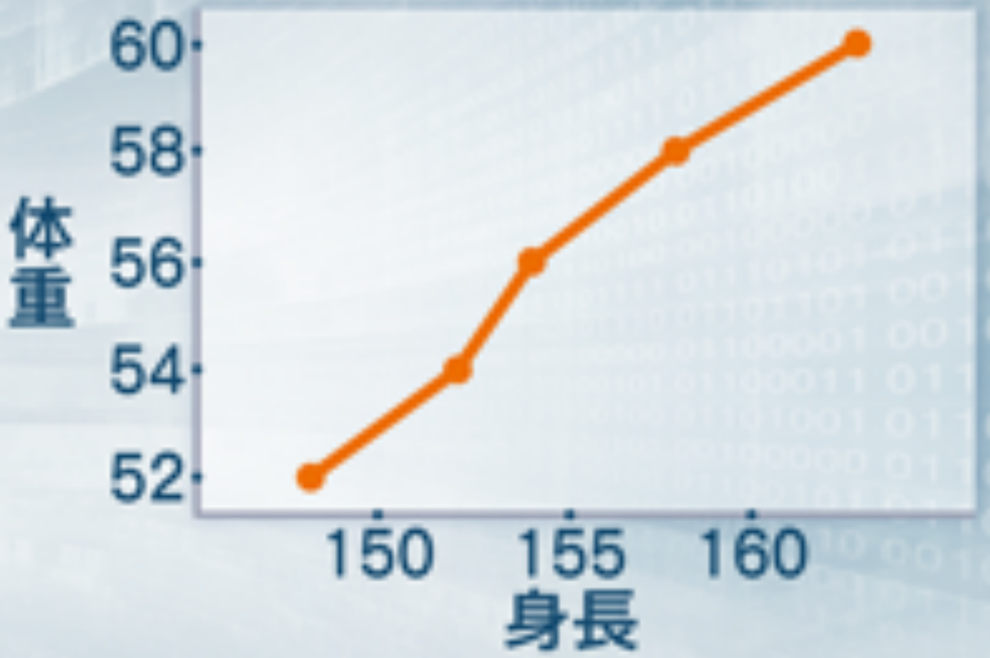
# 散布図と折れ線グラフ

散布図（関係）




plot（x座標のベクトル, y座標のベクトル）

折れ線グラフ（経過）

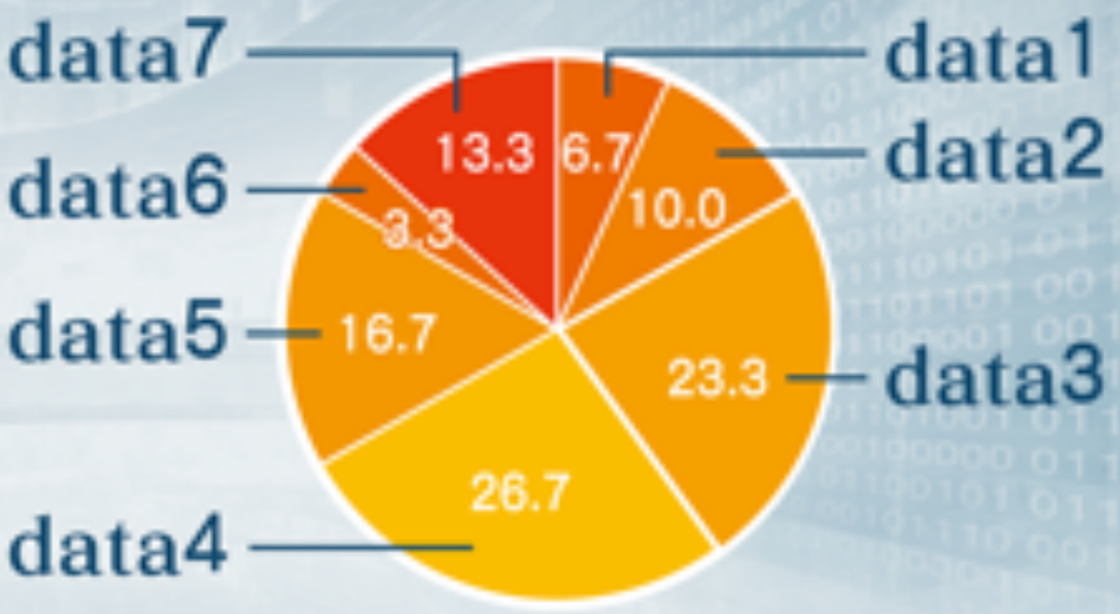



type＝o　b：点と線

# 円グラフ

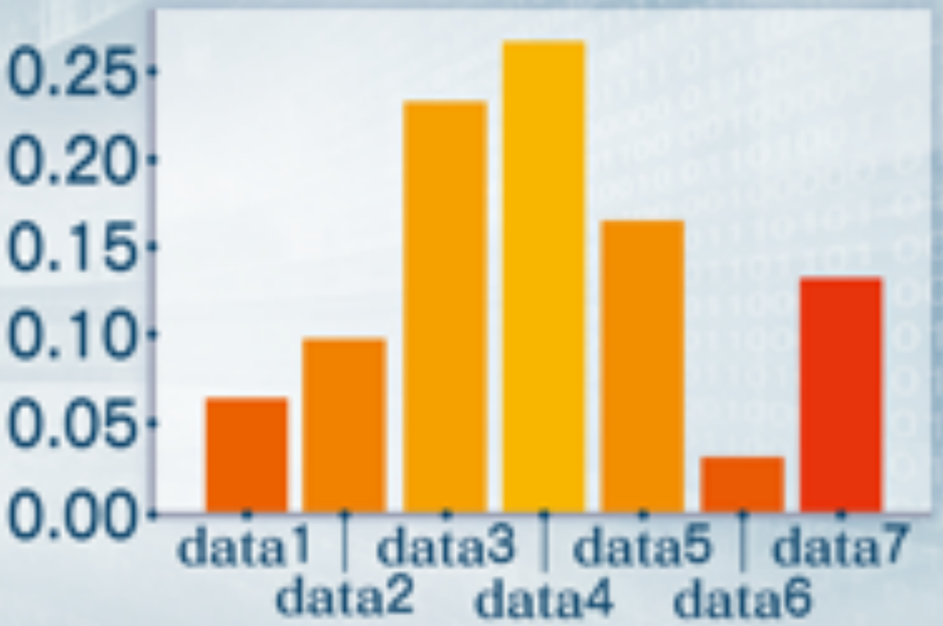
| 円グラフ(内訳) | |
|---|---|
| data1 | 6.7% |
| data2 | 10.0% |
| data3 | 23.3% |
| data4 | 26.7% |
| data5 | 16.7% |
| data6 | 3.3% |
| data7 | 13.3% |

pie(ベクトル)



# 棒グラフ

| 棒グラフ(比較) | |
|---|---|
| data1 | 6.7% |
| data2 | 10.0% |
| data3 | 23.3% |
| data4 | 26.7% |
| data5 | 16.7% |
| data6 | 3.3% |
| data7 | 13.3% |

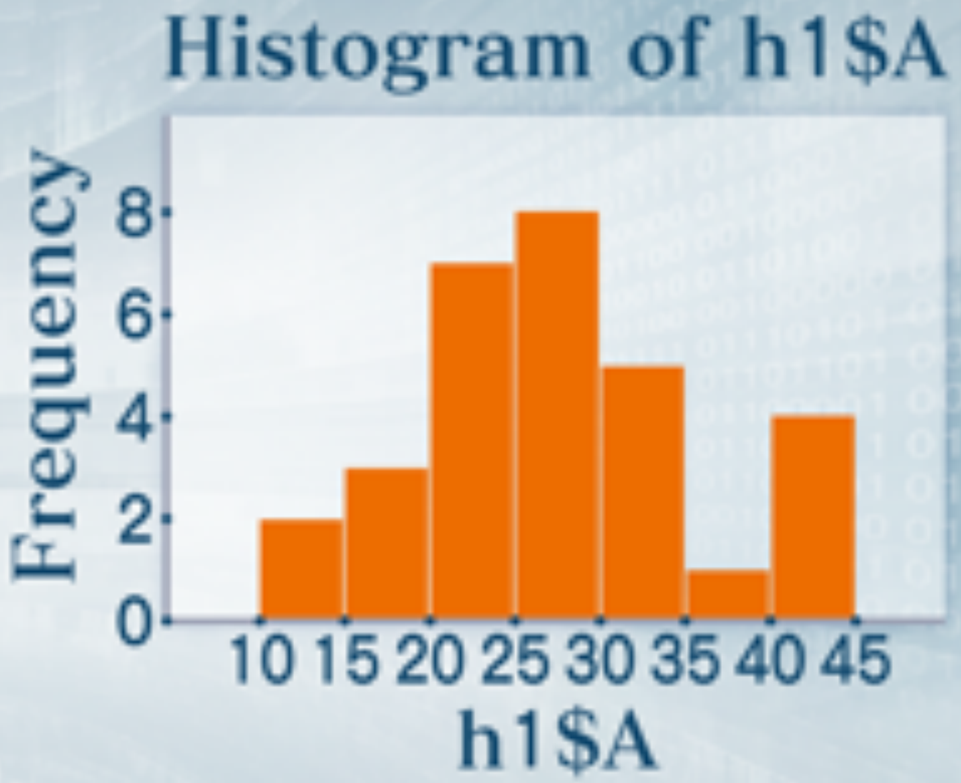



barplot (ベクトル or 行列)
行列で積み上げ棒グラフ

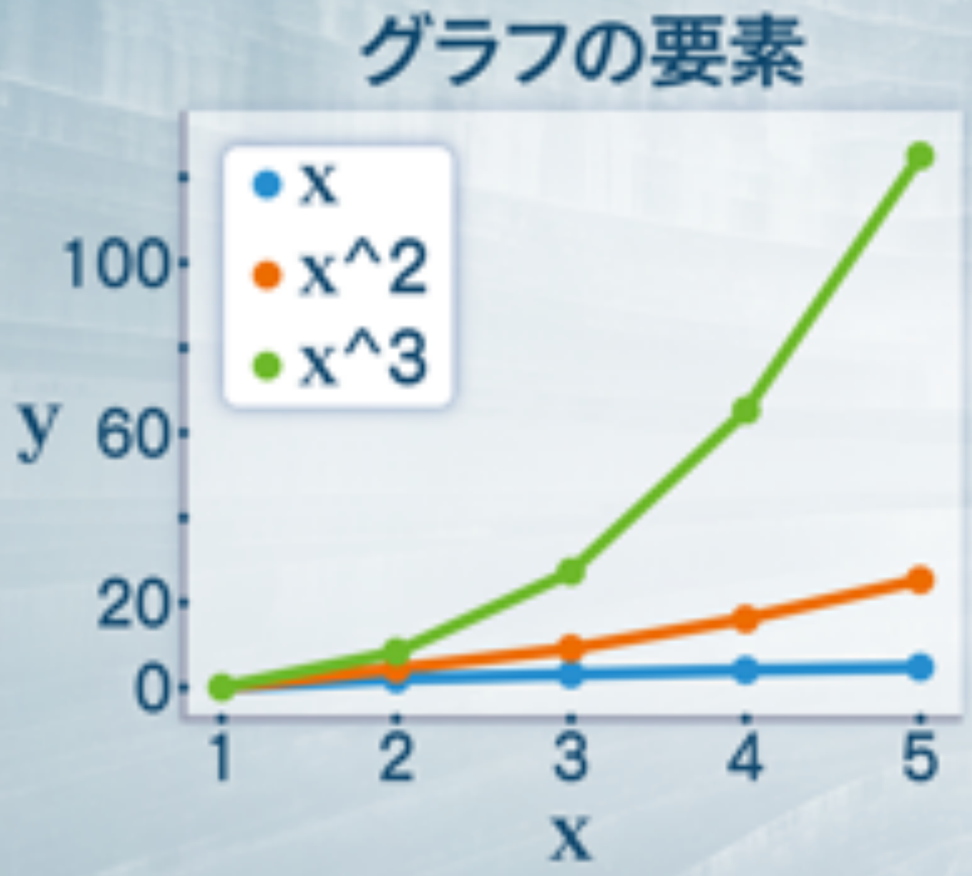
# ヒストグラム

| ヒストグラム | | | | |
|---|---|---|---|---|
| 24 | 31 | 26 | 24 | 43 |
| 34 | 32 | 27 | 28 | 38 |
| 23 | 28 | 42 | 27 | 13 |
| 31 | 15 | 41 | 17 | 29 |
| 29 | 24 | 35 | 23 | 19 |
| 21 | 28 | 44 | 16 | 22 |

hist（ベクトル）



# グラフの要素




**要素の指定**

- plot(オプションを指定する)
- main＝“図の見出し”
- xlab＝“x軸の名前”
- ylab＝“y軸の名前”

**凡例**

- legend(座標)

# グラフ

## グラフの種類

1. 散布図
2. 折れ線グラフ
3. 円グラフ
4. 棒グラフ
5. ヒストグラム
6. 箱ひげ図

## Rによるグラフの作成

### 高水準作図関数

- 単体で図を作成
- plot(), barplot()...

### 低水準作図関数

- 図に点などを追加
- points(), text()...

# Rによるグラフの作成(1)

散布図

```
> h1 <- read.table("ch051.dat",header=T, row.names=1)
> plot(h1, type="n")
> text(h1, rownames(h1))
```

n : none
p : point
l : line
b : both
o : overfit
s : step

凡例

```
> plot(h1, main="plot()", xlab="height",ylab="weight", pch=1)
> legend("topleft", legend="e01.dat", pch=1)
```

# Rによるグラフの作成(2)

## 折れ線グラフ

```
> h1 <- read.table("ch051.dat",header=T, row.names=1)
> plot(h1, type="o")
```

## 棒グラフと円グラフ

```
> c1 <- read.table("ch052.dat", header=T, row.names=1)
> pie(c1$freq, labels=rownames(c1), clockwise=T)
> barplot(c1$freq, names=rownames(c1))
```

# ファクター(因子)

```
> size1 <- c ("S","M","L","M","XL","S","L")
> size2 <- factor (size1)
> str (size2)
```

## カテゴリーデータの作成

$a < x \leqq b$

```
> sh <- c (0,3,15,15,10,8,12)
> cut (sh,breaks=c(min(sh)-1,10,max(sh)),label=c("N","Y"))
> cut (sh,breaks=c(min(sh),10,max(sh)+1),label=c("N","Y"),right=F)
```

$a \leqq x < b$