

# aa-2. データサイエンス・AI の事例、技術 (人工知能社会の到来、最新の技術、産業 の変化、社会や生活の変化など)

(人工知能)

URL: <https://www.kkaneko.jp/ai/mi/index.html>

金子邦彦



- 人工知能や，データサイエンスは，データの活用に大いに役立つ。
- これらを実践的に学んで欲しい。
- 今回は Excel を利用して，データサイエンスを実践的に学ぶ

# アウトライン

1. データサイエンス
2. 表計算ソフトウェア Excel
3. 散布図 (Excel を使用)
4. 合計、平均 (Excel を使用)
5. 分布、密度 (Excel を使用)
6. 人工知能と機械学習

# 2-1 データサイエンスでできること

# データサイエンスでできること



- **データサイエンスは、データから正しい知見や結論を導き出すための学問であり、さまざまな分野で活用されている**
  - ビジネス分野：顧客の嗜好やニーズを分析し、マーケティング戦略の立案を行う
  - 医療分野：病気の早期発見や効果的な治療法の開発を行う
  - 工学分野：製品品質の改善や予測保全など、生産の最適化を行う
- **情報化社会において、多くのデータが生み出されている。データサイエンスは、将来の活躍につながる**
- **データサイエンスは、大量のデータを扱うもの。機械学習など人工知能や情報処理とも大きく関連する。さまざまな分野でデータを活用する実力につながる。**

# 研究レポートの書くべき6つの要素



データサイエンスのスキルは、**研究レポートの作成**でも有用

- ① **問題**： 研究の背景と目的を明確に説明し、どのような問題に対処するかを示す
- ② **仮説**： 研究で提起する仮説は、問題を解決するための自分のアイデアや仮説です。
- ③ **実験手順**： 研究の手順を明確に示す。
- ④ **結果**： **研究の結果を正確かつ明確に示すことが必要です。** グラフや表などを用いて視覚的に表現すると、結果が分かりやすくなります。また、研究結果に関連する数値や統計的な情報も提供しましょう。
- ⑤ **考察**： 研究結果に基づいた分析です。
- ⑥ **引用文献**： 関連する先行研究や参考にした文献を明示

## 2-2 表計算ソフトウェア Excel

# パソコンの威力



- **ワープロ**

文書の編集、清書、目次、表の作成など

- **表計算**

データの管理、計算、グラフ作成など

- **プレゼン**

ビジュアル資料作成

- **インターネット**

情報収集、コミュニケーション

データはすべて**デジタル**（ファイル）。

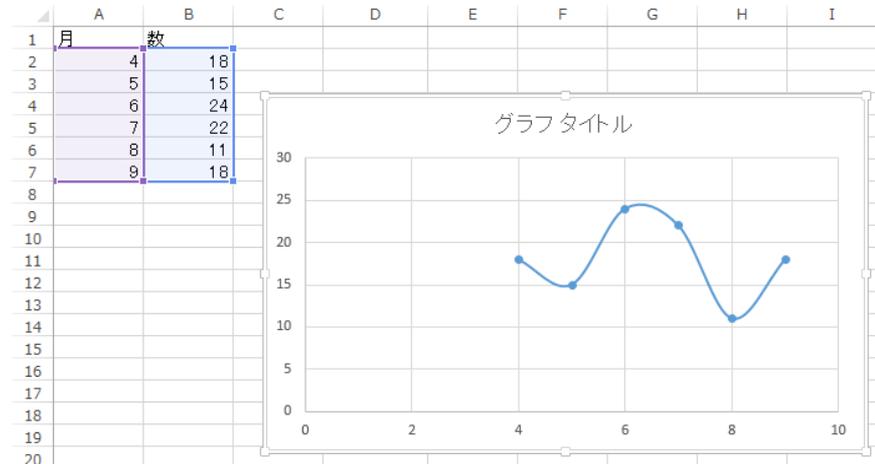
管理、共有、交換が簡単

# 表計算ソフトウェアは何の役に立つのか

- データの記録、保管、共有
- 表計算の機能 = 集計・集約、グラフ作成など

	A	B	C	D
1	品名	単価	数量	合計
2	りんご	100	10	1000
3	みかん	50	5	250
4				1250

	A	B	C
1	AA	算数	90
2	AA	国語	85
3	BB	算数	92
4	CC	国語	75
5	CC	理科	95
6			



# 例えば、こんなことが簡単にできます

## 表の作成

	A	B	C	D
1	品名	単価	数量	合計
2	りんご	100	10	1000
3	みかん	50	5	250
4				1250

単価を書き変えると

Excel の画面

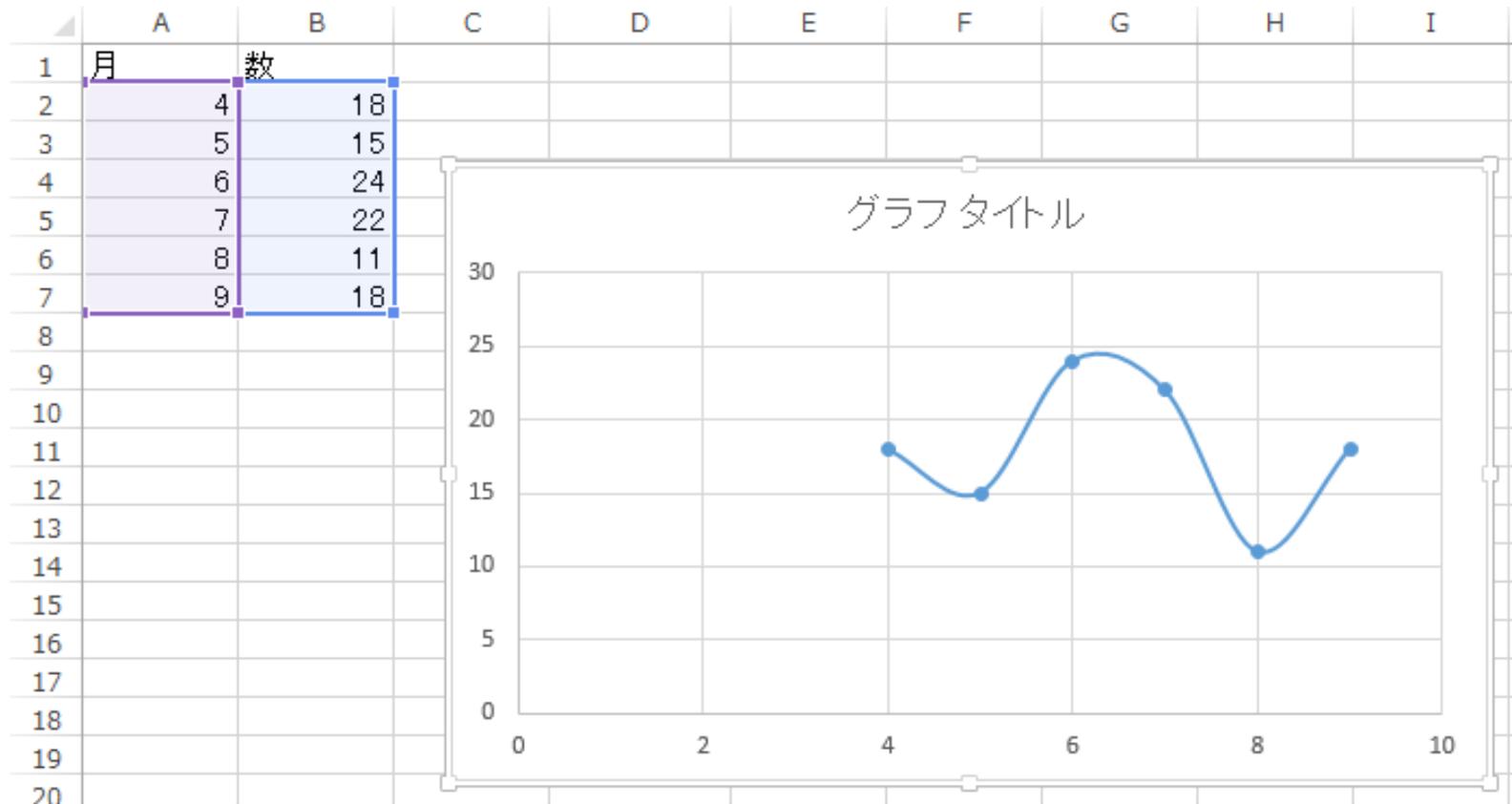
	A	B	C	D
1	品名	単価	数量	合計
2	りんご	80	10	800
3	みかん	50	5	250
4				1050

合計が**自動**で再計算される

Excel の画面

# 例えば、こんなことが簡単にできます

## グラフ



# 例えば、こんなことが簡単にできます

## 条件に合致するデータの 強調表示

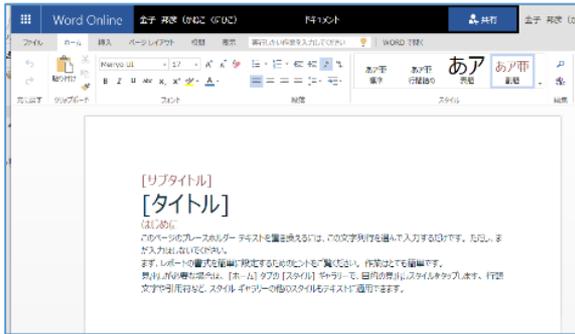
	A	B	C
1	AA	算数	90
2	AA	国語	85
3	BB	算数	92
4	CC	国語	75
5	CC	理科	95
6			

## 並べ替え

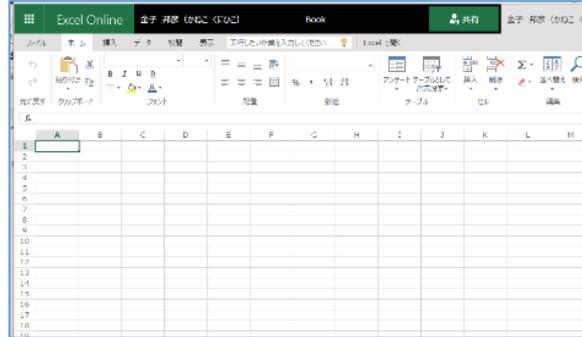
	A	B	C
1	CC	国語	75
2	AA	国語	85
3	AA	算数	90
4	BB	算数	92
5	CC	理科	95
6			

# Microsoft 365 と Excel

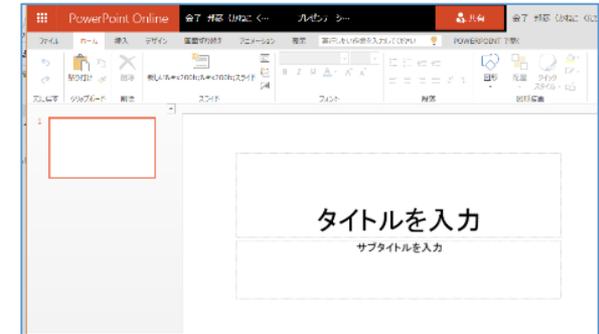
# Microsoft 365 の主な機能



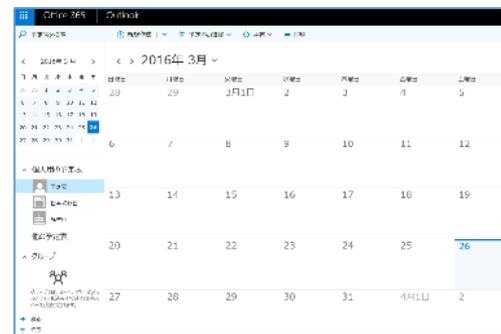
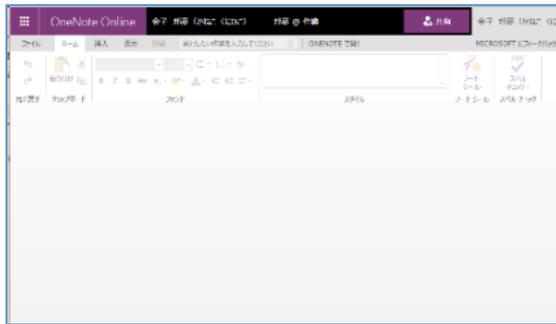
ワード (文書作成)



エクセル (表計算)



パワーポイント  
(プレゼン)



ワンノート (電子ノート)

アウトルック (電子メール)

- パソコンでレポートを作成したり，発表したり，データをまとめたりで便利

# Microsoft 365 の種類



- **Microsoft 365 のオンライン版**

**WEBブラウザ**で使う。

<https://portal.office.com>

各自の **ID** と **パスワード** でサインインが必要。

- **Microsoft 365 のアプリ版**

**前もってインストール**が必要。

インストールでは、大量の通信が行われる。

(時間がかかる。通信費用にも注意)

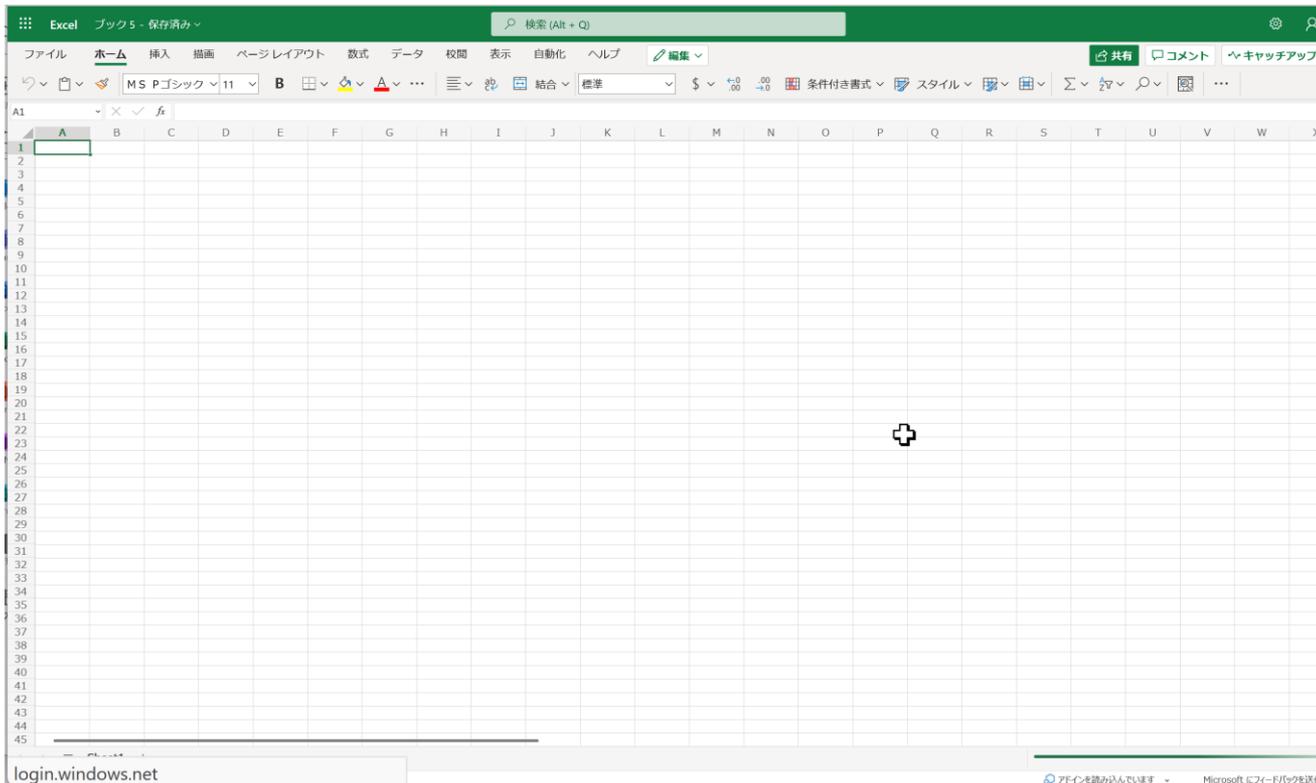
**2種類**ある。この授業では、**どちらを使用しても問題ない**

# Microsoft 365 オンライン版で Excel を起動



【要点】 Web ブラウザで、次のページを開き、各自の ID とパスワードでサインイン

<https://portal.office.com>



# Microsoft 365 オンライン版で Excel を起動



① Web ブラウザで, 次のページを開く

<https://portal.office.com>

② 電子メールアドレスを入れる. 「次へ」をクリック.

(例) [p1234567@fukuyama-u.ac.jp](mailto:p1234567@fukuyama-u.ac.jp)



## サインイン

メール、電話、Skype

---

アカウントをお持ちではない場合、[作成](#)できます。

[アカウントにアクセスできない場合](#)

戻る

次へ

# Microsoft 365 オンライン版で Excel を起動



## ③ パスワードを入れ, 「サインイン」をクリック

パスワードは, 各自が設定したもの

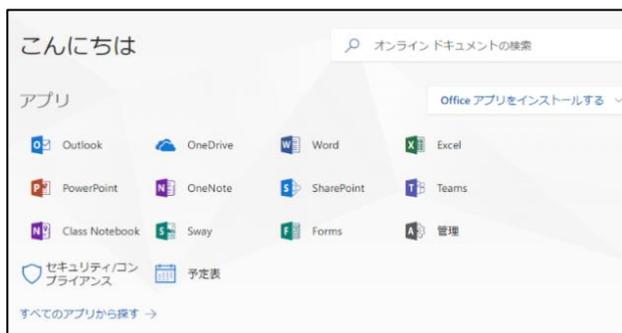
### パスワードの入力

パスワード

[パスワードを忘れた場合](#)

サインイン

## ④ Excel を使いたいときは, **メニュー**で Excel を選ぶ



さまざまなメニュー

# Microsoft 365 オンライン版で Excel を起動

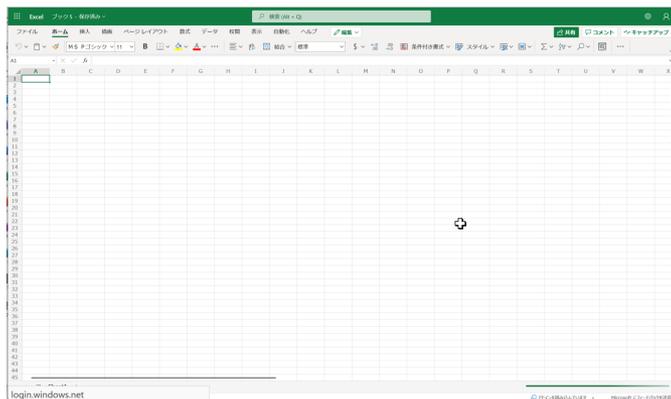


## ⑤ Excel のブックの種類を選ぶ



この授業では「新しい空白のブック」を使う

## ⑥ Excel の画面が開く



# Microsoft 365 アプリ版のインストールと Excel の起動

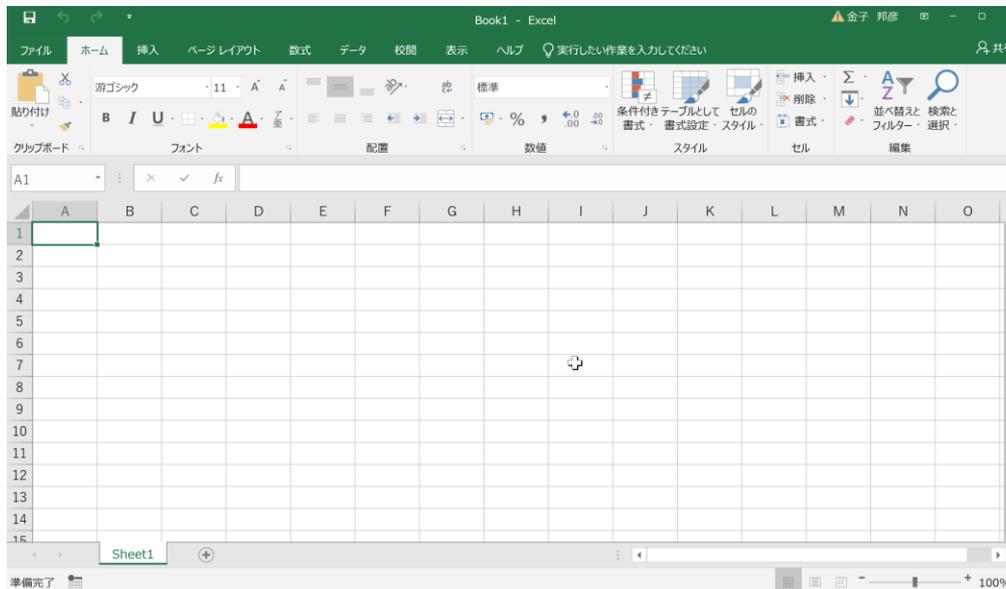


【要点】 **インストール**は、Microsoft 365 アプリ版を使えるようにするための作業（最初に行う）。

そのとき、次のページを開き、各自の **ID とパスワード** でサインイン

<https://portal.office.com>

**インストール**が終わったら、**スタートメニュー**等で Excel を起動



# Microsoft 365 アプリ版のインストールと Excel の起動



① **Web ブラウザ**で, 次のページを開く  
**<https://portal.office.com>**

② **電子メールアドレス**を入れる. 「**次へ**」をクリック.  
**(例) [p1234567@fukuyama-u.ac.jp](mailto:p1234567@fukuyama-u.ac.jp)**



## サインイン

メール、電話、Skype

---

アカウントをお持ちではない場合、[作成](#)できます。

[アカウントにアクセスできない場合](#)

戻る

次へ

# Microsoft 365 アプリ版のインストールと Excel の起動



- ③ パスワードを入れ, 「サインイン」をクリック  
パスワードは, 各自が設定したもの

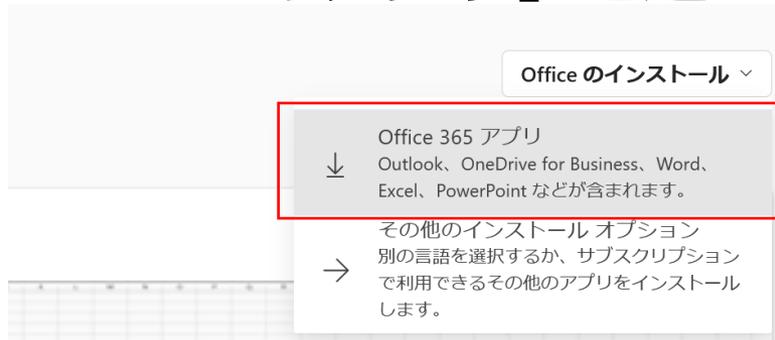
## パスワードの入力

パスワード

[パスワードを忘れた場合](#)

サインイン

- ④ 画面で「Office のインストール」をクリック. メニューで「Microsoft 365 のアプリ」を選ぶ



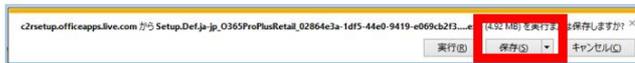
# Microsoft 365 アプリ版のインストールと Excel の起動



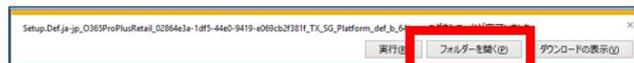
## ⑤ 画面の指示に従い、インストールを行う

インストールでは、大量の通信が行われる。  
(時間がかかる。通信費用にも注意)

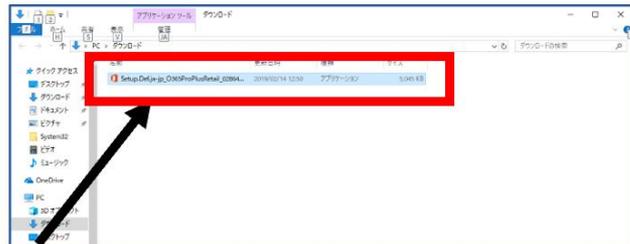
次のような指示がでる



1. 保存する



2. フォルダーを開く



3. 実行し、その後も、画面の指示に従う

Setup.Def.ja-jp\_O365ProPlusRetail\_02864...

# Microsoft 365 アプリ版のインストールと Excel の起動

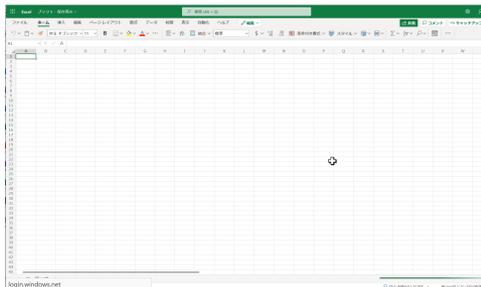
⑥ Excel を使うときは，スタートメニューなどで Excel を選ぶ

⑦ **Excel のブックの種類**を選ぶ



この授業では「新しい空白のブック」を使う

⑧ Excel の画面が開く



# Excel の基本

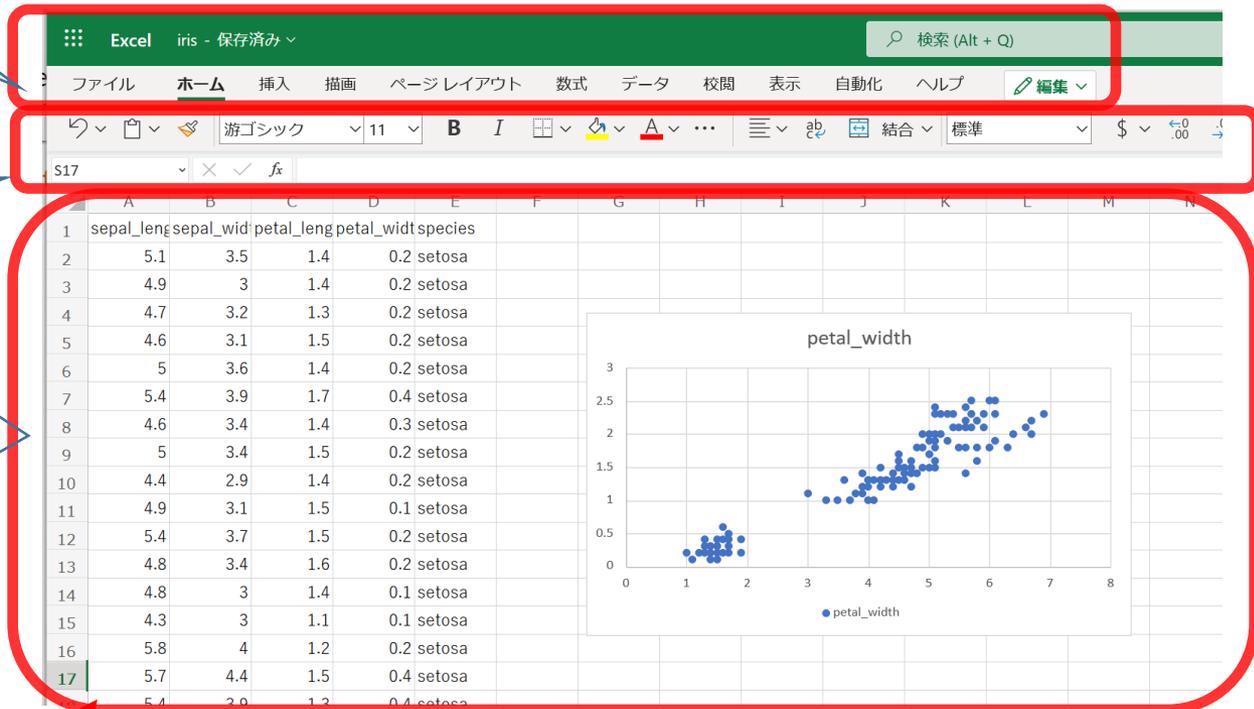
# オンライン版の Excel の画面 (メニュー、リボン、ワークシートなど)



メニュー

リボン

ワークシート  
表形式で値などが  
入る。  
グラフの挿入な  
ども可能



表形式で、値や数式を並べる

オンライン版の Excel の画面

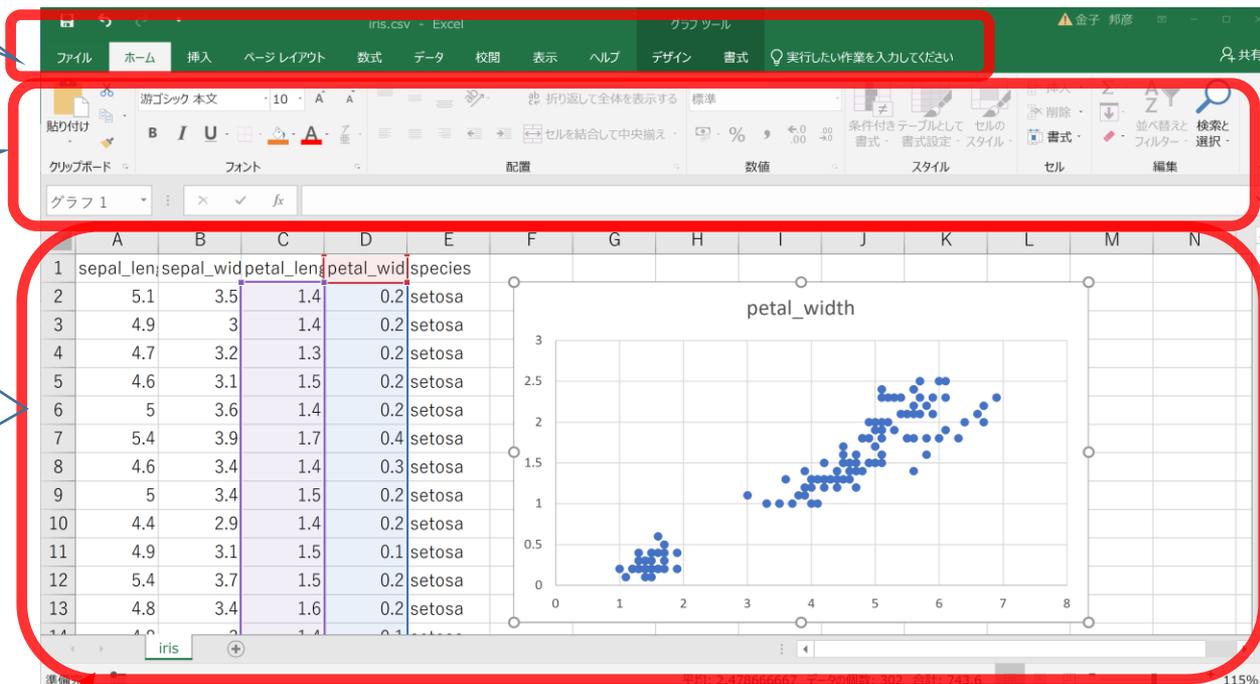
# アプリ版の Excel の画面（メニュー、リボン、ワークシートなど）



メニュー

リボン

ワークシート  
表形式で値などが  
入る。  
グラフの挿入な  
ども可能



表形式で、値や数式を並べる

アプリ版の Excel の画面  
(Excel 2019 の画面を示している) 27

# Excel のワークシート

Excel の**ワークシート**（シートともいう）には、表形式で、値や数式を並べる。

グラフの挿入なども可能

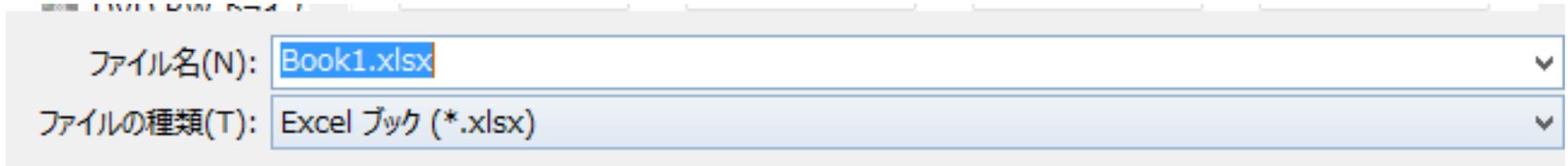
	A	B	C	D
1	品名	単価	数量	合計
2	りんご	100	10	1000
3	みかん	50	5	250
4				1250
5				

Excel のワークシートの例

# Excel のブック



- Excel の**ブック**は、Excel の**ファイル**のこと
- 1つあるいは複数の**ワークシート**を、1つの**ブック**に保存することができる



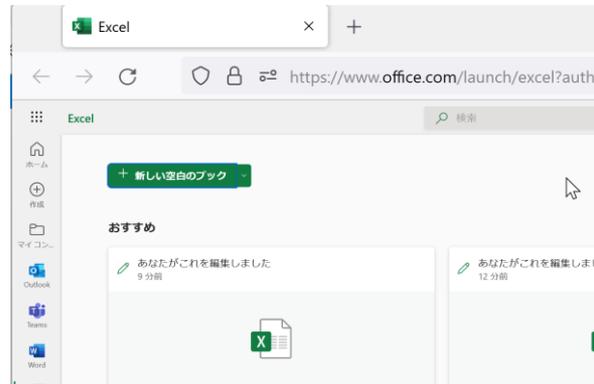
Excel で保存するときに、ファイル名などを設定できる。

# Excel のスタート画面

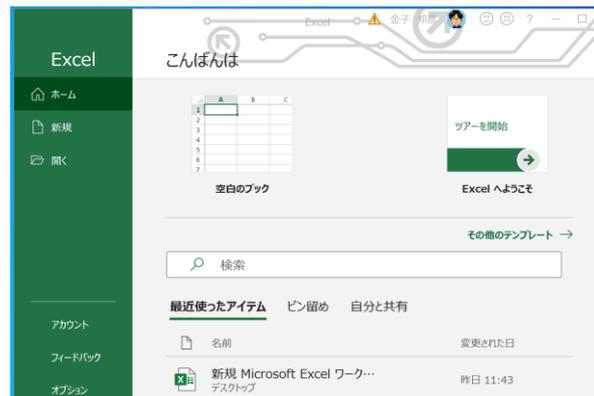
Excel を**起動**すると、**最初**にスタート画面が表示される。  
作成したいブックの種類を選ぶことができる。  
過去の履歴の確認もできる

この授業では「新しい空白のブック」を使う

オンライン版



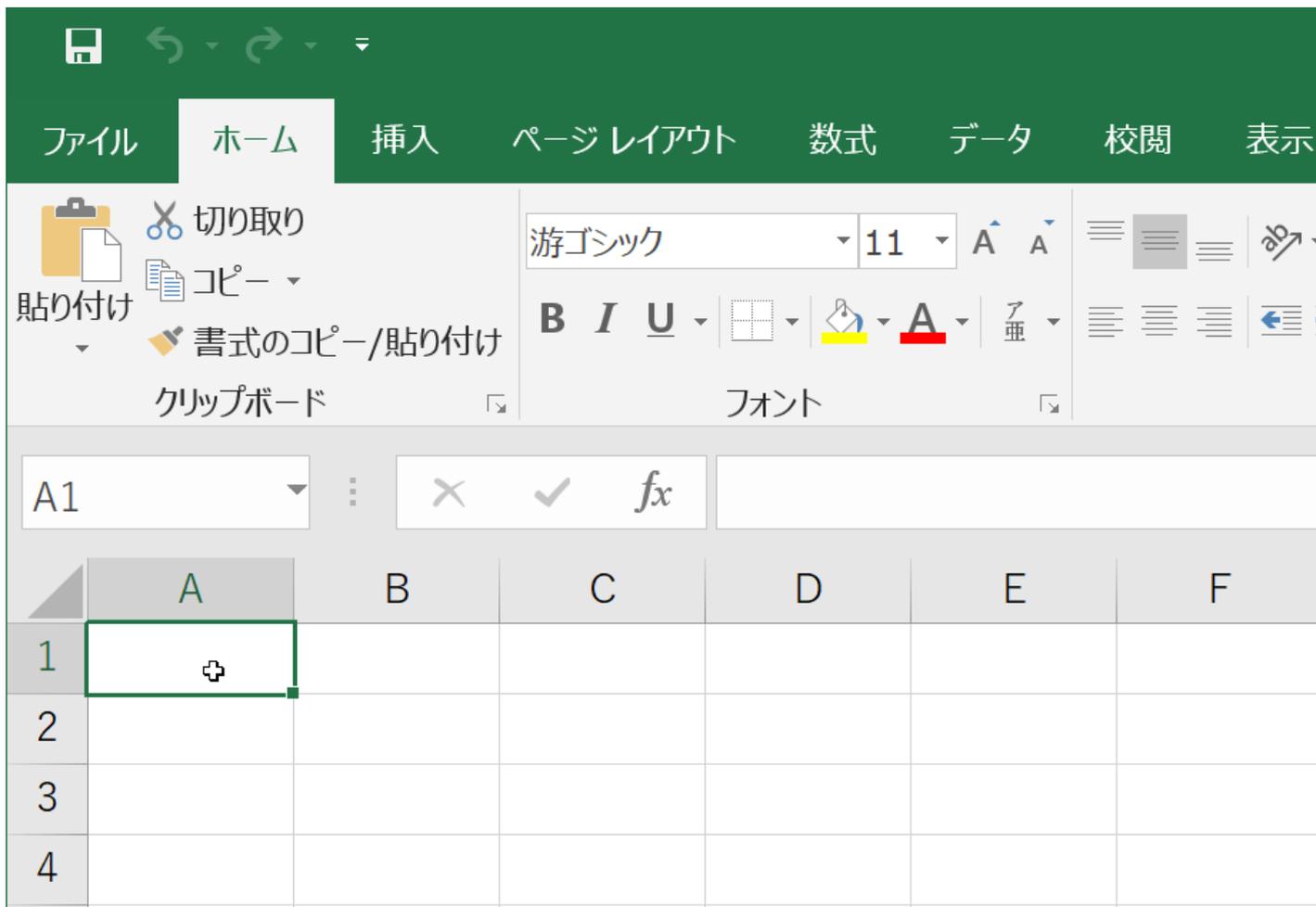
アプリ版



# アクティブセル



- Excel での編集集中のセル



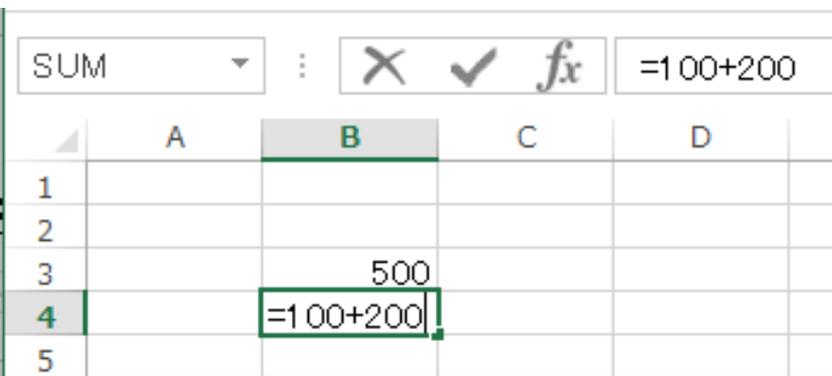
The screenshot displays the Microsoft Excel interface. The ribbon is set to the 'ホーム' (Home) tab. The 'クリップボード' (Clipboard) group is visible, showing '切り取り' (Cut), 'コピー' (Copy), and '書式のコピー/貼り付け' (Paste with Styles) options. The 'フォント' (Font) group shows the font name '游ゴシック' (Yu Gothic), size '11', and various formatting options like bold, italic, underline, and text color. The active cell is A1, which is highlighted with a green border and contains a small crosshair cursor, indicating it is in edit mode. The formula bar above the grid shows 'fx' and a confirmation button.

	A	B	C	D	E	F
1						
2						
3						
4						

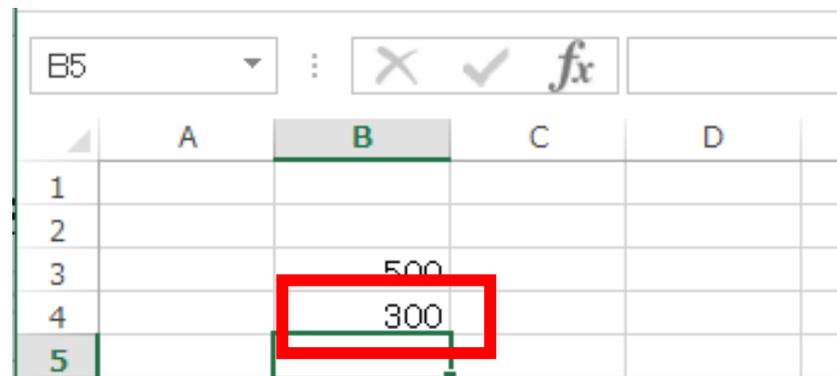
# アクティブセルでの数式の入力



数式を入力したいときは、頭に半角の「=」を付ける



	A	B	C	D
1				
2				
3		500		
4		=100+200		
5				



	A	B	C	D
1				
2				
3		500		
4		300		
5				

キーボードで「**=100+200**」と打つと、**アクティブセル**に数式が入る

入力を終わりたいので **Enter キー**を押す。

すると、数式が**自動計算**される

# 数式バーで数式の確認①

数式バーに  
数式が表示される  
(ここで修正もできる)

	A	B	C	D
1				
2				
3		500		
4		300		
5				



	A	B	C	D
1				
2				
3		500		
4		300		
5				

「300」のところを  
クリック

アクティブセルが動く

## 数式バーで数式の確認②

	A	B	C	D
1				
2				
3		500		
4		300		
5				



	A	B	C	D
1				
2				
3		500		
4		=100+200		
5				

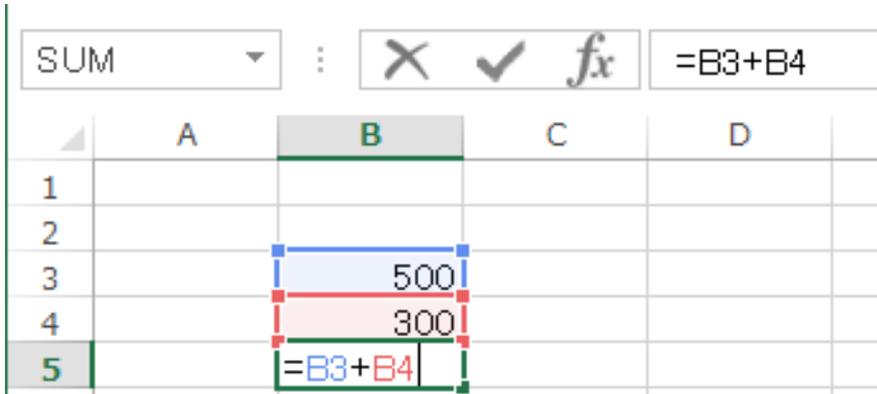
「300」のところを  
ダブルクリック

数式バーに  
数式が表示される  
(ここで修正もできる)

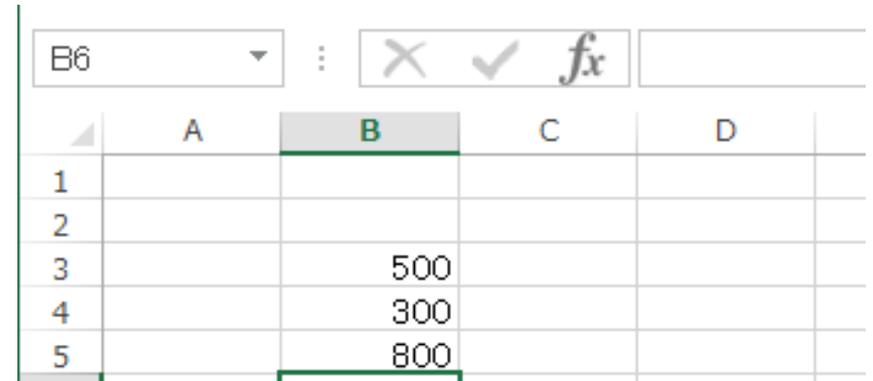
アクティブセルのところに  
数式が表示される  
(ここでも修正できる)

# アクティブセルでの数式の入力

「=B3+B4」のような数式もある



	A	B	C	D
1				
2				
3		500		
4		300		
5		=B3+B4		



	A	B	C	D
1				
2				
3		500		
4		300		
5		800		

キーボードで「=B3+B4」と  
打つと、**アクティブセル**に数式が入る

入力を終わりたいので  
**Enter** キーを押す。  
すると、数式が**自動計算**される

# Excel の数式

- Excel では、数式の頭に、**半角の「=」** を付ける
- 数式は、**半角文字**である
- 数式の中には、**番地**（「B3」や「B4」など）を書くことができる

# セルの数式と値のクリア

C4 : X ✓ fx 3.2

	A	B	C	D	E	F
1						
2		Sepal Length (cm)	Sepal Width (cm)			
3		7	3.2	4.7	1.4	Iris-versicolor
4		6.4	3.2			
5		6.9	3.1			
6		5.5	2.3			
7		6.5	2.8			
8		5.7	2.8			
9		6.3	3.3			
10		4.9	2.4			
11		6.6	2.9			
12		5.2	2.7			
13		5	2			
14		5.9	3			
15		6	2.2			
16		6.1	2.9			
17		5.6	2.9			

Context menu options: 切り取り(I), コピー(C), 貼り付けのオプション: 形式を選択して貼り付け(S)... 挿入(I)... 削除(D)... **数式と値のクリア(N)**, クリック分析(Q)



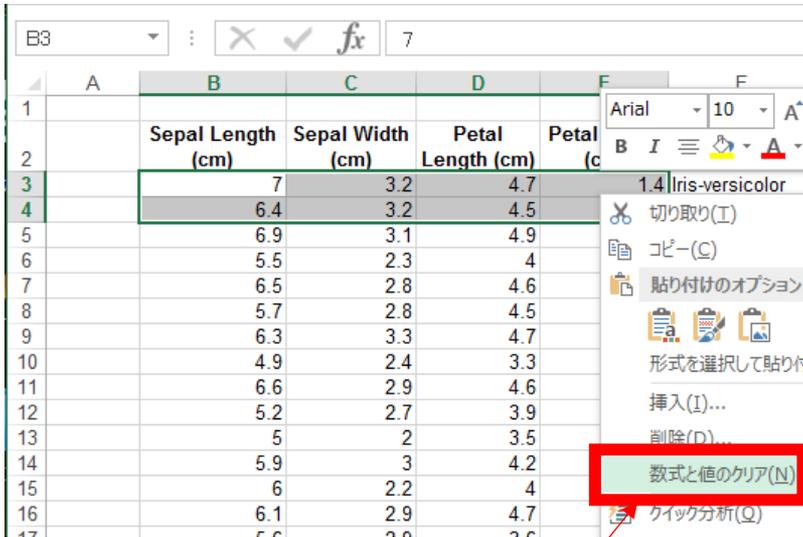
C4 : X ✓ fx

	A	B	C	D	E	F
1						
2		Sepal Length (cm)	Sepal Width (cm)	Petal Length (cm)	Petal Width (cm)	Class
3		7	3.2	4.7	1.4	Iris-versicolor
4		6.4		4.5	1.5	Iris-versicolor
5		6.9	3.1	4.9	1.5	Iris-versicolor
6		5.5	2.3	4	1.3	Iris-versicolor
7		6.5	2.8	4.6	1.5	Iris-versicolor
8		5.7	2.8	4.5	1.3	Iris-versicolor
9		6.3	3.3	4.7	1.6	Iris-versicolor
10		4.9	2.4	3.3	1	Iris-versicolor
11		6.6	2.9	4.6	1.3	Iris-versicolor
12		5.2	2.7	3.9	1.4	Iris-versicolor
13		5	2	3.5	1	Iris-versicolor
14		5.9	3	4.2	1.5	Iris-versicolor
15		6	2.2	4	1	Iris-versicolor
16		6.1	2.9	4.7	1.4	Iris-versicolor
17		5.6	2.9	3.6	1.3	Iris-versicolor

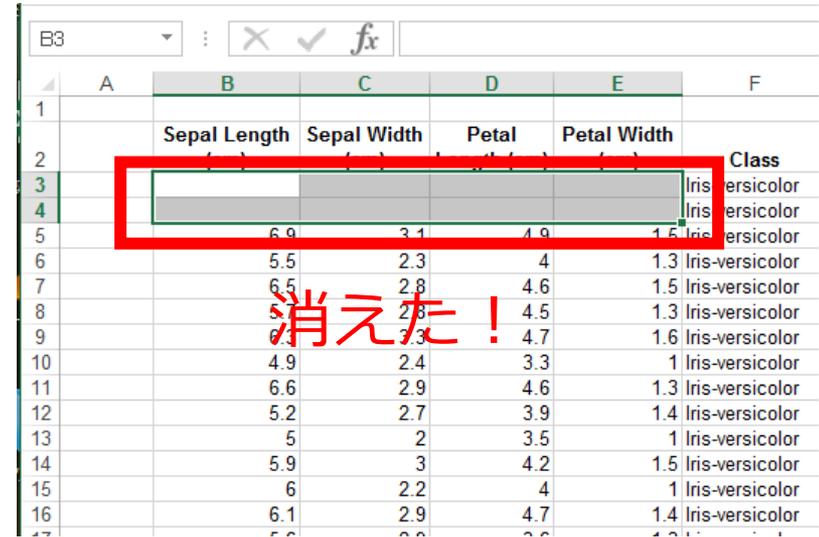
Red box highlights cell C4. Text: 消えた!

クリアしたいセルを右クリックして、  
「数式と値のクリア」を選ぶ

# セルの数値と値のクリア



	A	B	C	D	E	F
1						
2		Sepal Length (cm)	Sepal Width (cm)	Petal Length (cm)	Petal Width (cm)	
3		7	3.2	4.7	1.4	Iris-versicolor
4		6.4	3.2	4.5		
5		6.9	3.1	4.9		
6		5.5	2.3	4		
7		6.5	2.8	4.6		
8		5.7	2.8	4.5		
9		6.3	3.3	4.7		
10		4.9	2.4	3.3		
11		6.6	2.9	4.6		
12		5.2	2.7	3.9		
13		5	2	3.5		
14		5.9	3	4.2		
15		6	2.2	4		
16		6.1	2.9	4.7		
17						



	A	B	C	D	E	F
1						
2		Sepal Length	Sepal Width	Petal	Petal Width	Class
3						Iris-versicolor
4						Iris-versicolor
5		6.9	3.1	4.9	1.5	Iris-versicolor
6		5.5	2.3	4	1.3	Iris-versicolor
7		6.5	2.8	4.6	1.5	Iris-versicolor
8		5.7	2.8	4.5	1.3	Iris-versicolor
9		6.3	3.3	4.7	1.6	Iris-versicolor
10		4.9	2.4	3.3	1	Iris-versicolor
11		6.6	2.9	4.6	1.3	Iris-versicolor
12		5.2	2.7	3.9	1.4	Iris-versicolor
13		5	2	3.5	1	Iris-versicolor
14		5.9	3	4.2	1.5	Iris-versicolor
15		6	2.2	4	1	Iris-versicolor
16		6.1	2.9	4.7	1.4	Iris-versicolor
17						

消えた!

クリアしたいセルを範囲選択（マウスでドラッグ）したあと、右クリックして、「数式と値のクリア」を選ぶ

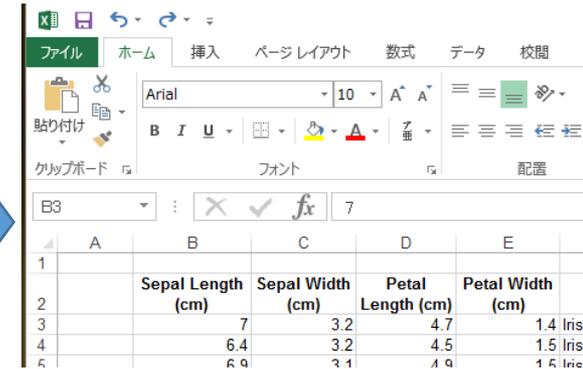
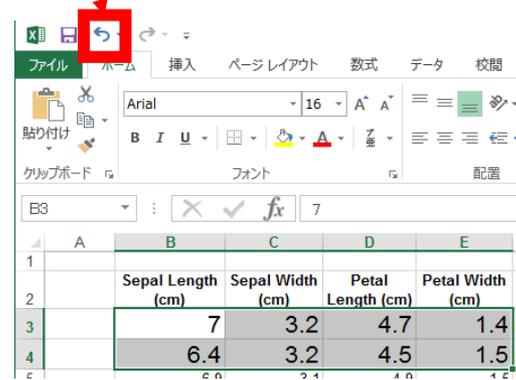
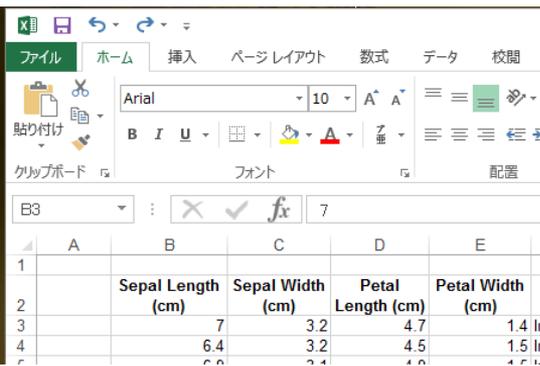
# 元に戻す操作



「元に戻す」  
ボタン



オンライン版の Excel でも  
「元に戻す」ボタンはある



何かの操作を  
したとする

「元に戻す」ボタン  
を押すと元に戻る

もとに戻す操作は CTRL + Z

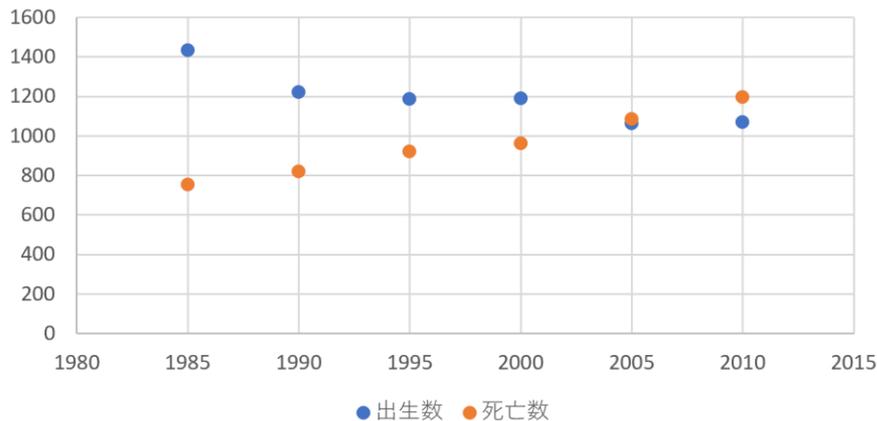
(コントロールキーと「Z」を同時押し)でも、できる

## 2-3 散布図 (Excel を使用)

# 散布図の用途

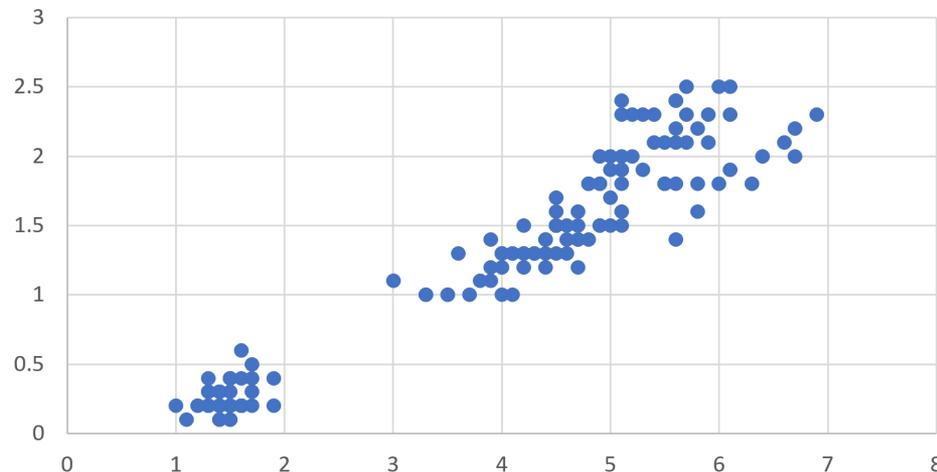


## • 時間変化



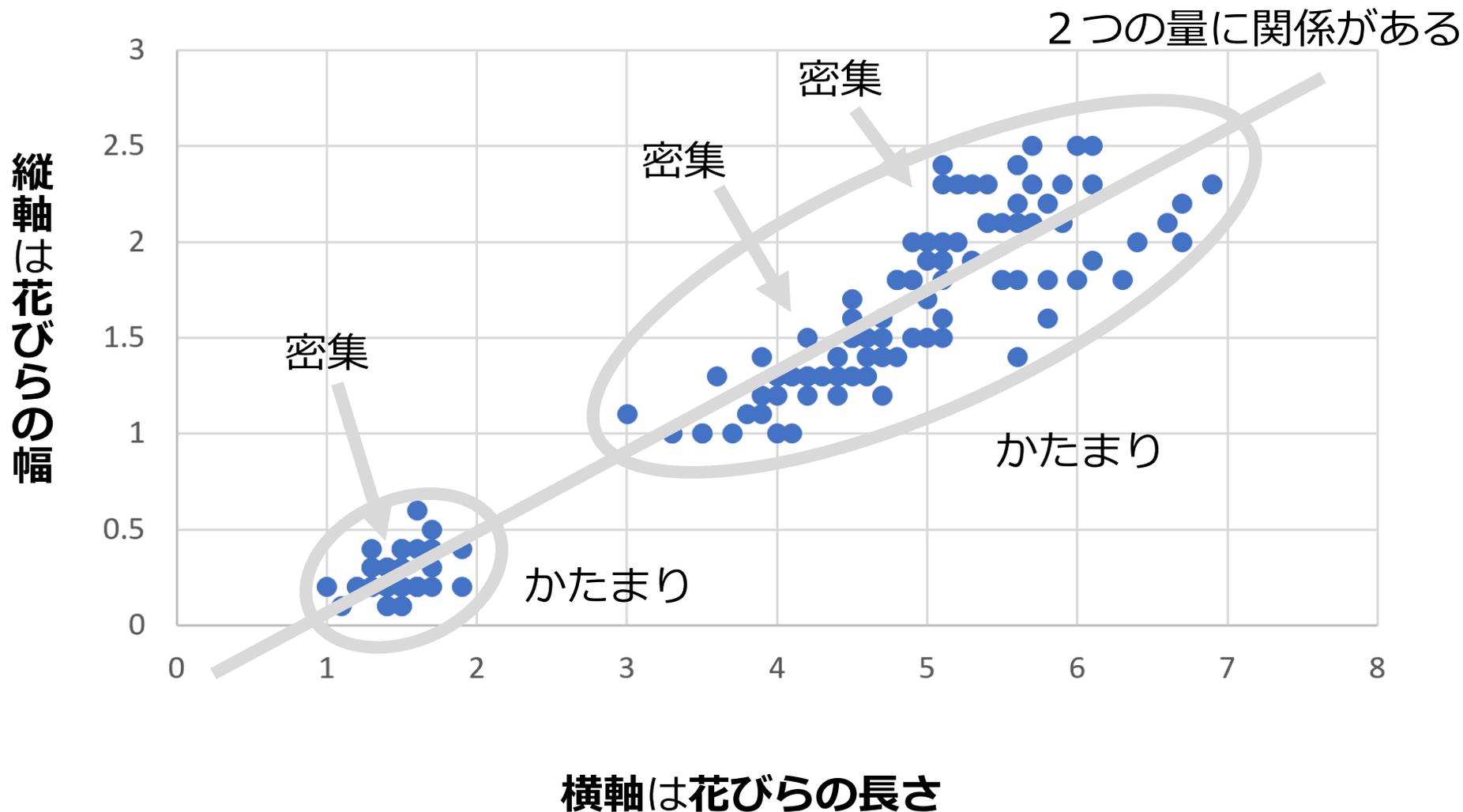
横軸は時間。  
散布図から、  
**時間変化**を読み取る

## • 分布



横軸と縦軸は、2つの量。  
散布図から、  
2つの量の間の**関係**を見る

# 分布から読み取れること



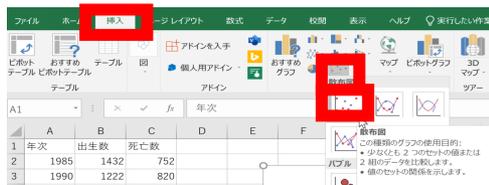
# Excel での散布図の作成手順

	A	B	C
1	年次	出生数	死亡数
2	1985	1432	752
3	1990	1222	820
4	1995	1187	922
5	2000	1191	962
6	2005	1063	1084
7	2010	1071	1197

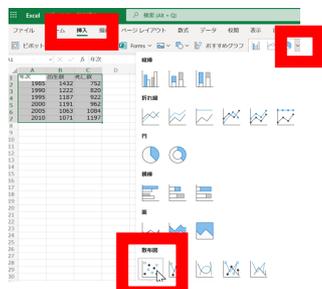
元データ

	A	B	C	D
1	年次	出生数	死亡数	
2	1985	1432	752	
3	1990	1222	820	
4	1995	1187	922	
5	2000	1191	962	
6	2005	1063	1084	
7	2010	1071	1197	
8				

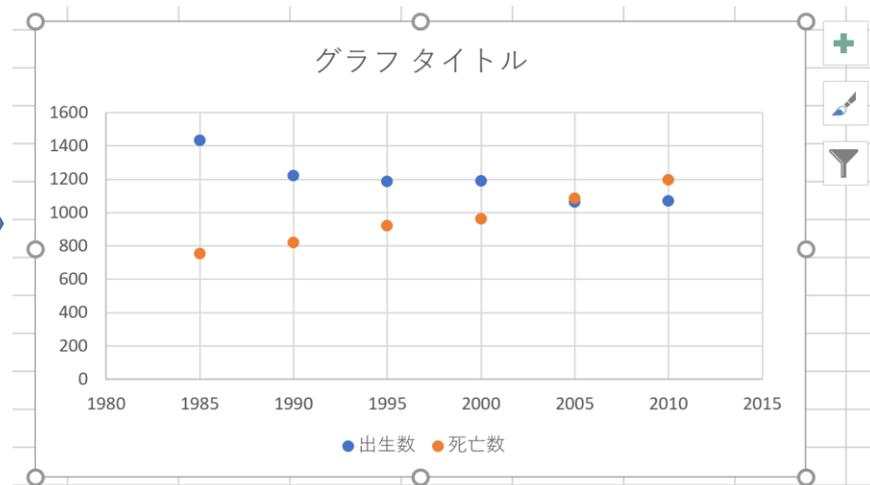
① グラフ化したい部分を範囲選択



アプリ版の Excel



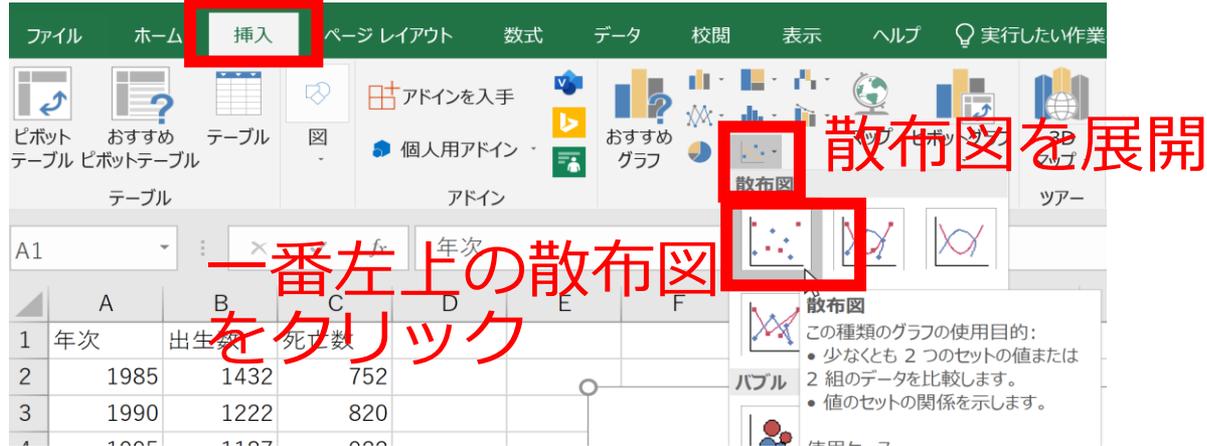
オンライン版の Excel



散布図が得られる

② リボンで「挿入」→散布図

# Excel での散布図の種類を選択 挿入



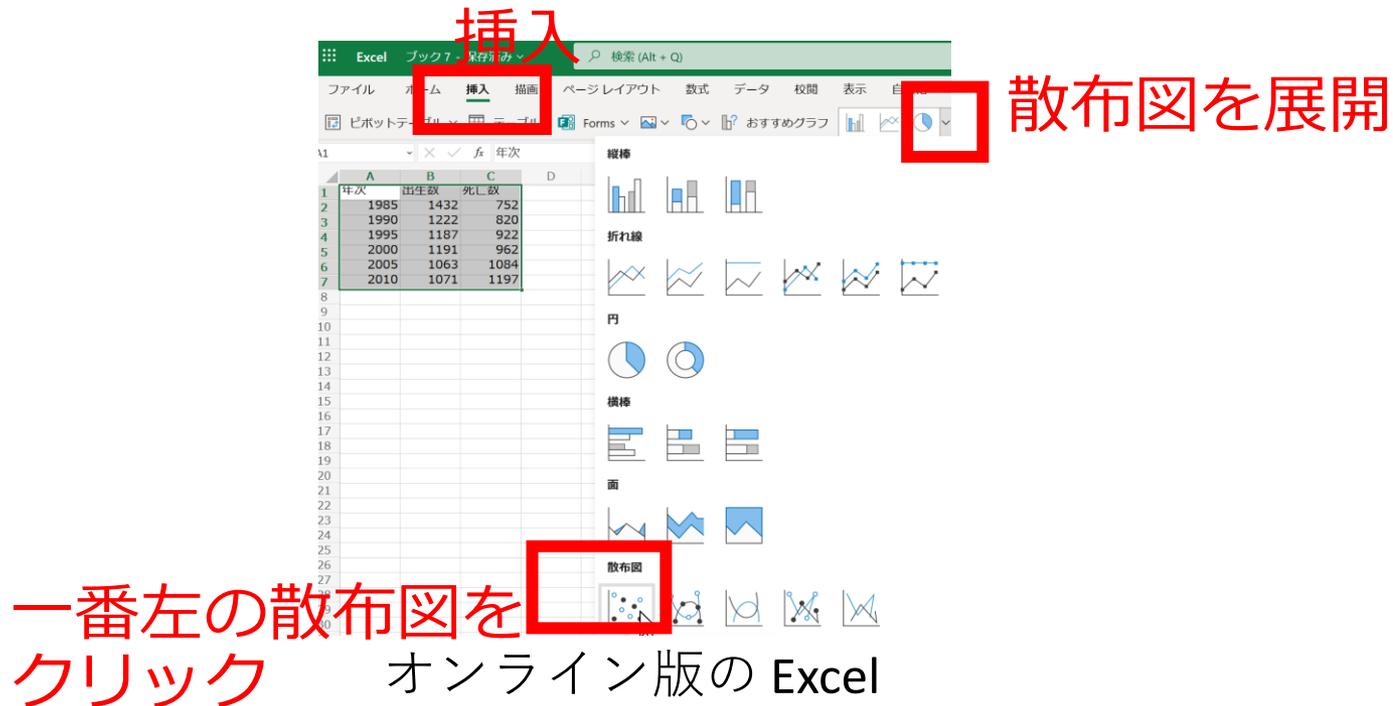
Excel desktop application interface. The '挿入' (Insert) tab is highlighted in the ribbon. The '散布図' (Scatter) chart type is selected and expanded, showing various scatter plot options. A tooltip for the selected scatter plot is visible, explaining its purpose: comparing two sets of data and showing their relationship.

一番左上の散布図をクリック

散布図を展開

年次	出生数	死亡数
1985	1432	752
1990	1222	820
2005	1063	1084
2010	1071	1197

アプリ版の Excel



Excel online application interface. The '挿入' (Insert) tab is highlighted in the ribbon. The '散布図' (Scatter) chart type is selected and expanded, showing various scatter plot options. A tooltip for the selected scatter plot is visible, explaining its purpose: comparing two sets of data and showing their relationship.

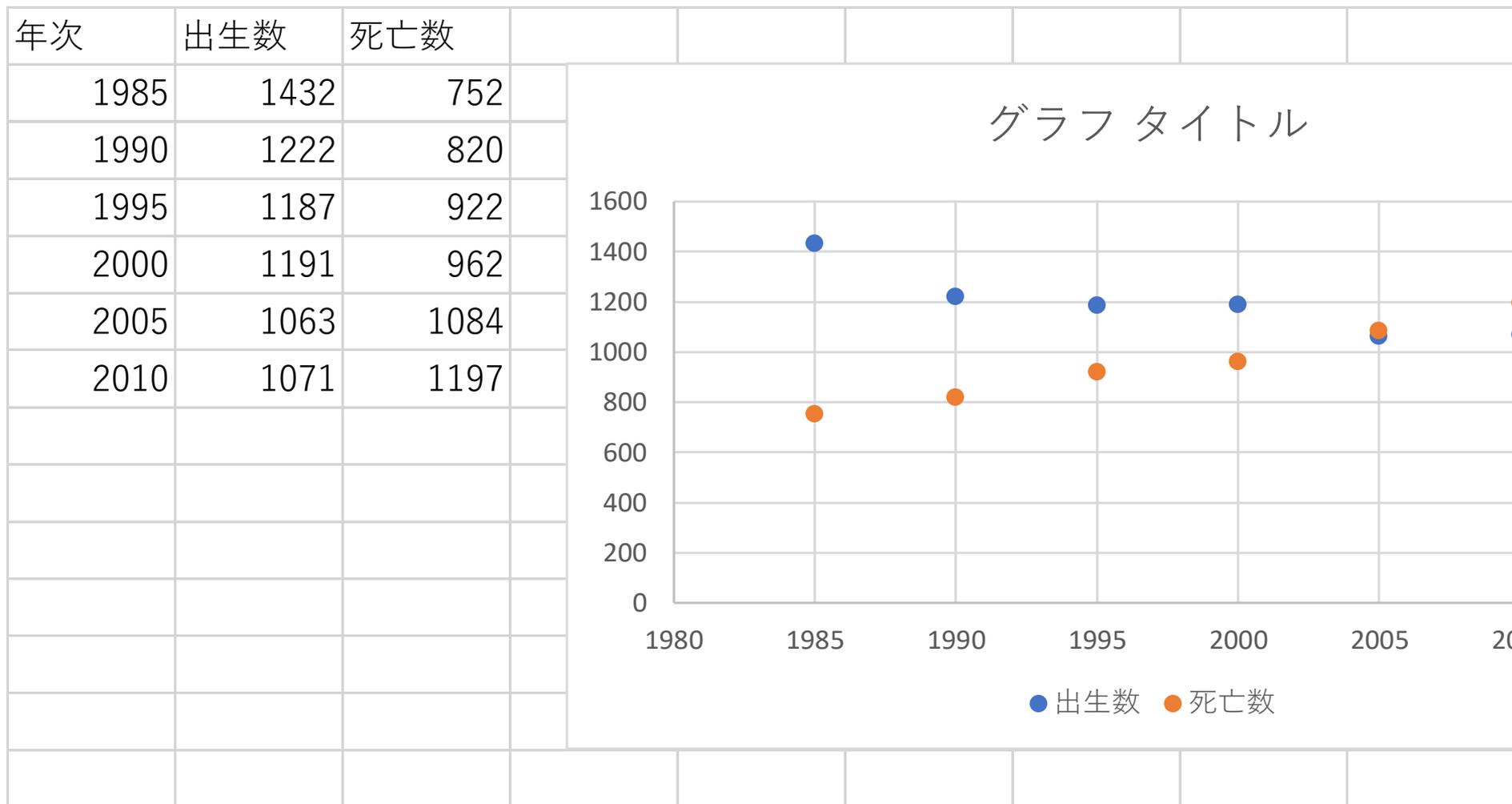
一番左の散布図をクリック

オンライン版の Excel

散布図を展開

年次	出生数	死亡数
1985	1432	752
1990	1222	820
1995	1187	922
2000	1191	962
2005	1063	1084
2010	1071	1197

# 元データ



## 散布図

## 2-4 データの合計、平均 (Excel を使用)

# Excel で合計を求める SUM



	A	B	C	D	E
1	年次	出生数	死亡数		
2	1985	1432	752		
3	1990	1222	820		
4	1995	1187	922		
5	2000	1191	962		
6	2005	1063	1084		
7	2010	1071	1197		
8			5737		

**=SUM(C2:C7)** は、  
範囲 **C2** から **C7** の  
**合計**を求める

# Excel で平均を求める AVERAGE



	A	B	C	D	E
1	年次	出生数	死亡数		
2	1985	1432	752		
3	1990	1222	820		
4	1995	1187	922		
5	2000	1191	962		
6	2005	1063	1084		
7	2010	1071	1197		
8		1194.333	5737		

**=AVERAGE(B2:B7)** は、  
範囲 **B2** から **B7** の  
**平均**を求める

# 平均



- **平均**は、データの**合計**を、**データの個数**で割ったもの

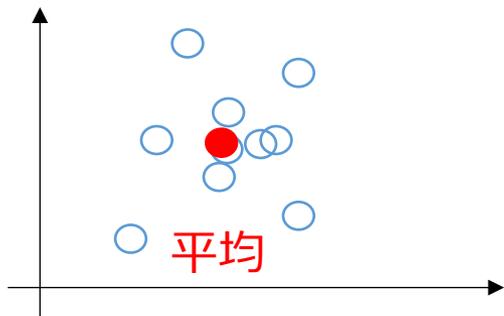
10, 40, 30, 40 の**平均**:  $120 \div 4$  で **30**

- **複数の値の組の平均**を考えることもある

(10, 5), (40, 10), (30, 5), (40, 20) の平均:

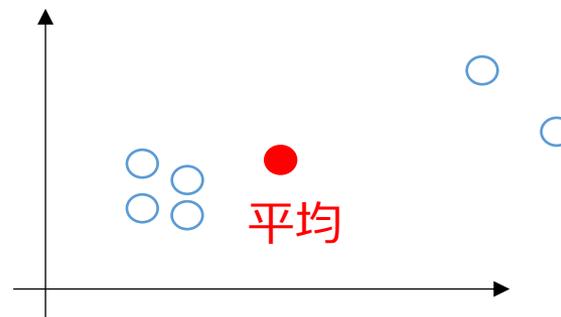
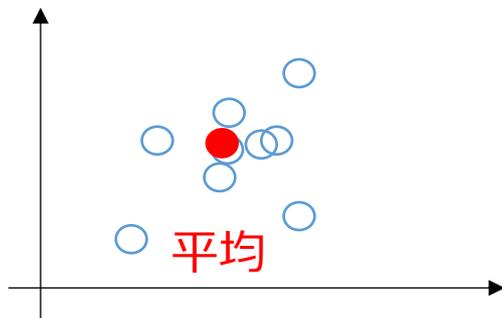
合計は 120 と 40. 4で割って (30, 10)

**平均**は、**データ集合の代表**とみることができるところがある



計測に**誤差**があるとき、  
複数の計測を繰り返して、**平均**をとることで、**誤差を軽減**できることも

# 平均を使うときの注意点



このような平均に、  
意味があるでしょうか？

**データの分布によっては、平均では役に立たないこともある。**  
(平均は万能ではない)

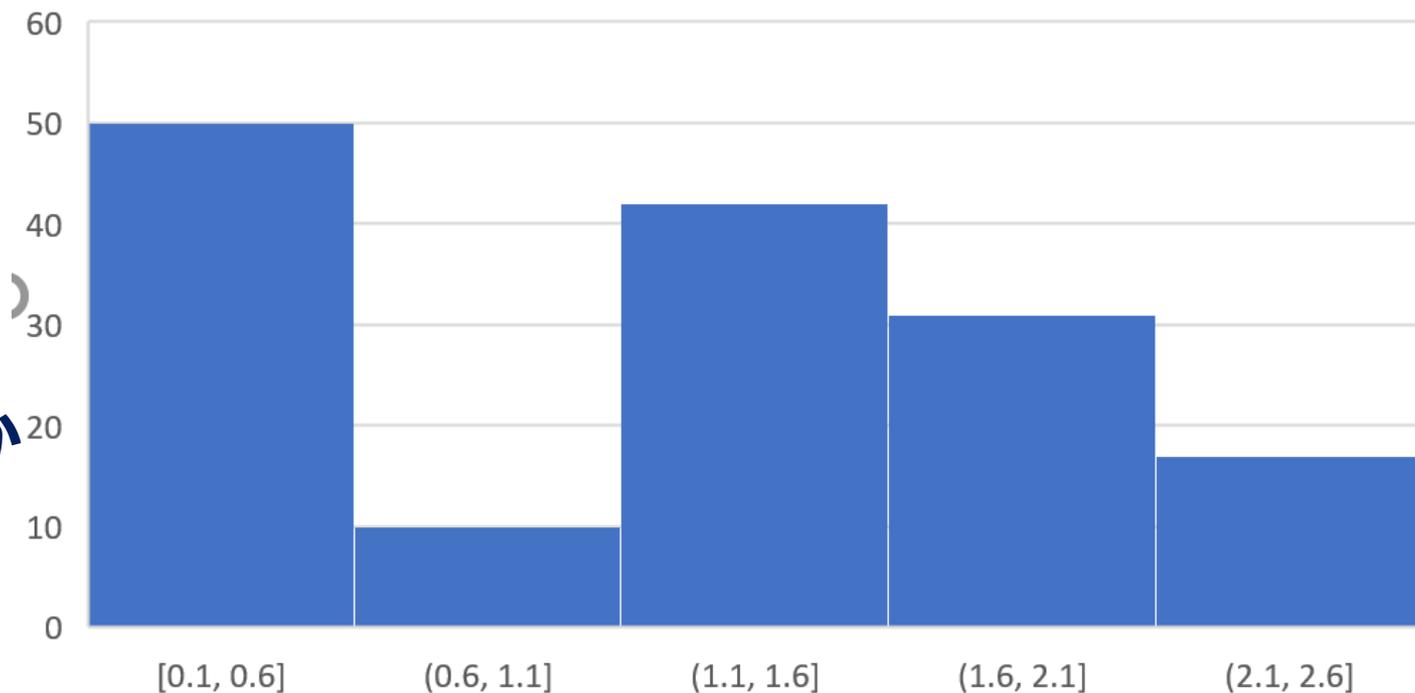
## 2-5 データの分布、密度 (Excel を使用)

# ヒストグラム



**ヒストグラム**は、区間ごとに、データを数え上げたもの

↑  
データが  
何個あるのか



↑  
区間 0.6 ~ 1.1 の  
データは 10個

# Excel でのヒストグラムの作成手順



	A	B	C	D	E
1	sepal_len	sepal_wid	petal_len	petal_wid	species
2	5.1	3.5	1.4	0.2	setosa
3	4.9	3	1.4	0.2	setosa
4	4.7	3.2	1.3	0.2	setosa
5	4.6	3.1	1.5	0.2	setosa
6	5	3.6	1.4	0.2	setosa
7	5.4	3.9	1.7	0.4	setosa
8	4.6	3.4	1.4	0.3	setosa
9	5	3.4	1.5	0.2	setosa

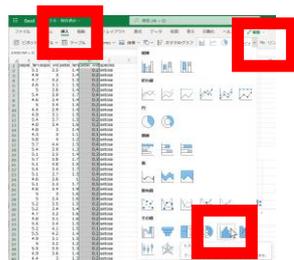
元データ

	A	B	C	D	E
1	sepal_len	sepal_wid	petal_len	petal_wid	species
2	5.1	3.5	1.4	0.2	setosa
3	4.9	3	1.4	0.2	setosa
4	4.7	3.2	1.3	0.2	setosa
5	4.6	3.1	1.5	0.2	setosa
6	5	3.6	1.4	0.2	setosa
7	5.4	3.9	1.7	0.4	setosa
8	4.6	3.4	1.4	0.3	setosa
9	5	3.4	1.5	0.2	setosa
10	4.4	2.9	1.4	0.2	setosa
11	4.9	3.1	1.5	0.1	setosa
12	5.4	3.7	1.5	0.2	setosa

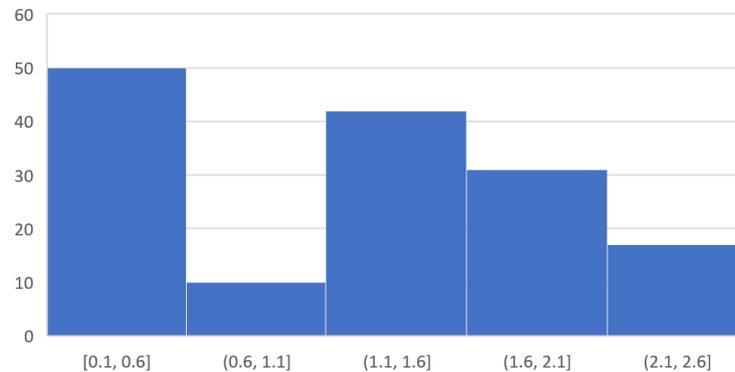
① ヒストグラム化したい列を選択



アプリ版の Excel



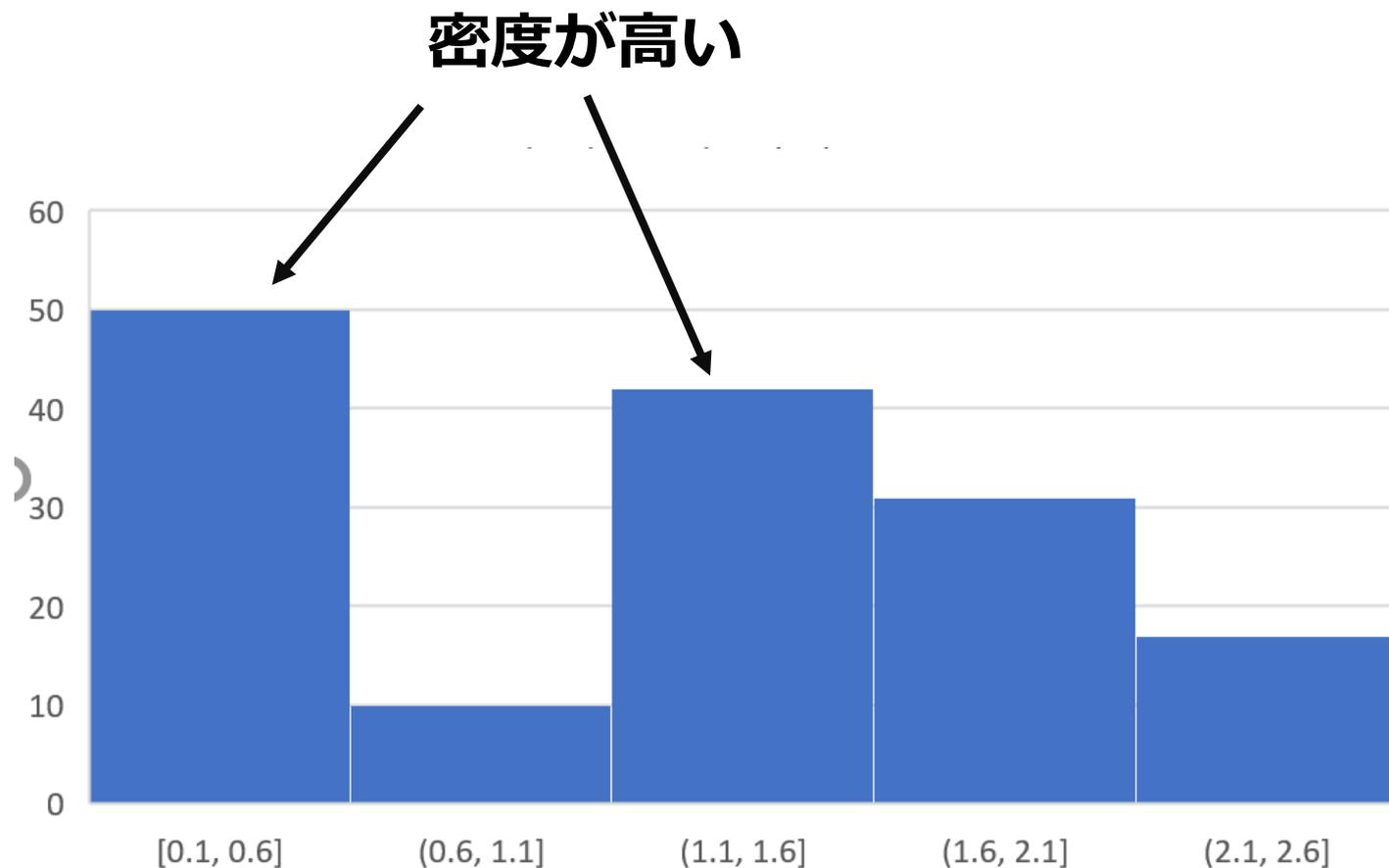
オンライン版の Excel



ヒストグラムが得られる

② リボンで「挿入」→ヒストグラムの選択

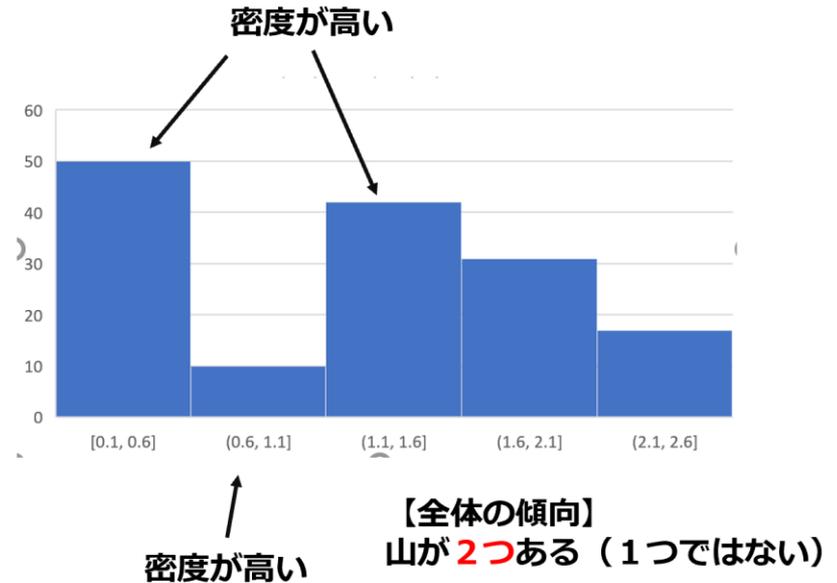
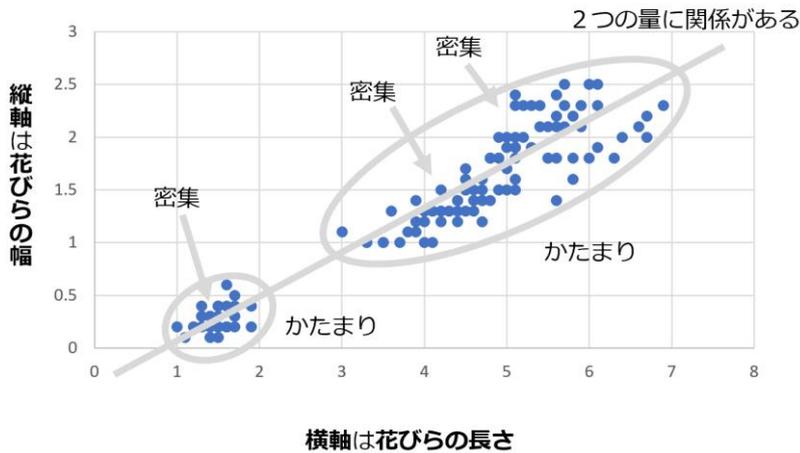
# ヒストグラムから読み取れること



密度が低い

【全体の傾向】  
山が2つある（1つではない）

- データサイエンスは、データから正しい知見や結論を導くための学問



- 数式を使うこともあるが、それがデータサイエンスの全てではない
- データから有益な情報を引き出すことが可能に。
- 大学生にとって、大切なスキル

## 2-6 人工知能と機械学習

# 機械学習の特徴



**機械学習**は、**コンピュータ**が**データ**を使用して**学習**することにより**知的能力を向上**させる技術

- **情報の抽出**：データの中からパターンや関係性を自動で見つけ出す能力
- **簡潔さ**：人間が設定しなければならなかったルールを、自動で生成できるようになる
- **限界の超越**：他の方法では難しかった課題でも、機械学習を用いることで解決策や視点を得られる可能性がある

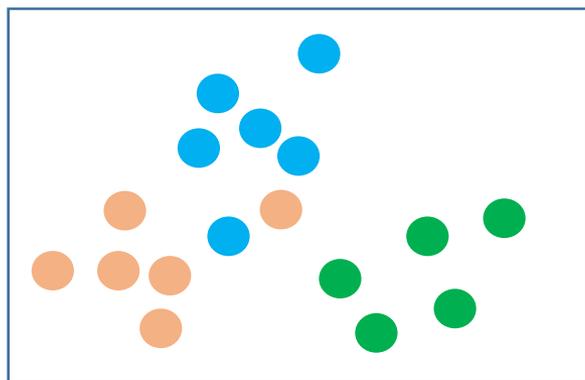


# 機械学習と訓練データ



機械学習は、コンピュータがデータを使用して学習することにより知的能力を向上させる技術

## 訓練データ



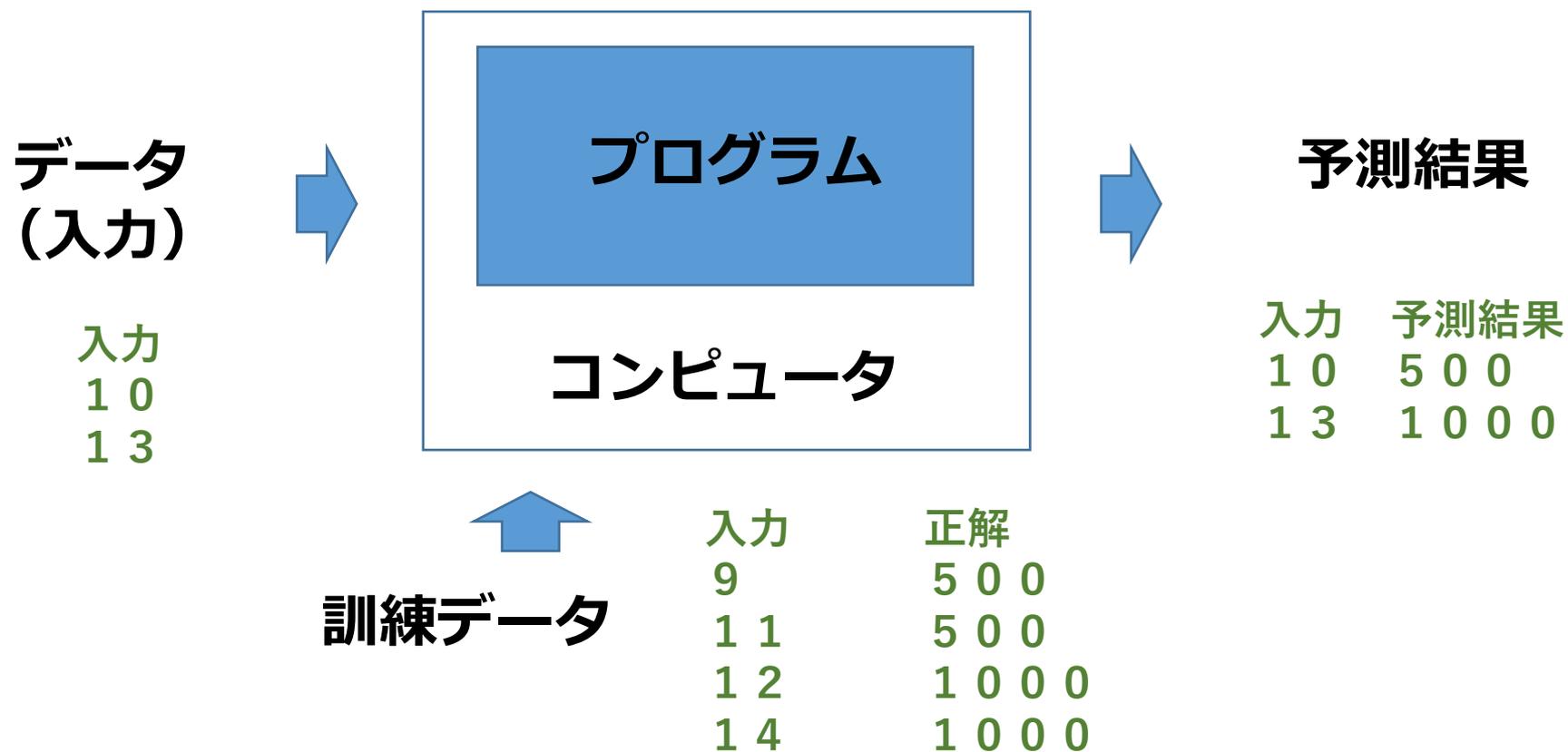
3種類に分類済み



学習者

大量の訓練データを用いて  
学習を行う

# 機械学習での予測



訓練データにより、学習が行われる。

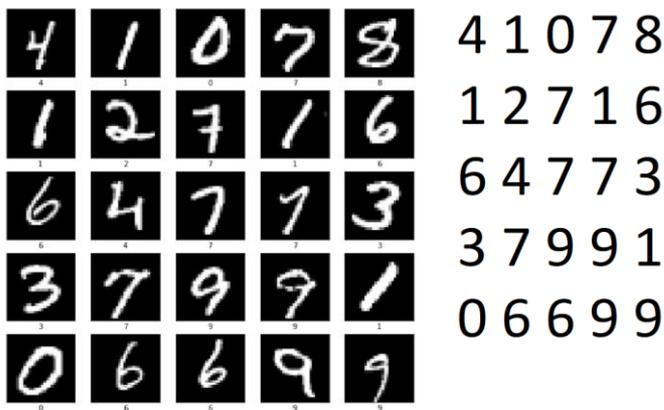
新しいデータに対しては、自動で処理が行われる。

# 機械学習と訓練データとプログラム



## 機械学習のプログラム

### 学習に使用する訓練データ (抜粋)



画像 60000枚  
(うち一部)

正解 60000個  
(うち一部)

### プログラム

データを用いて学習を行う  
学習ののち、画像分類を行う

```
[4] !pip install -U scikit-learn matplotlib

import torch
import torch.nn as nn
import torch.optim as optim
from sklearn import datasets
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt

# データの取得と前処理
iris = datasets.load_iris()
X = iris.data
y = iris.target

# データの標準化
scaler = StandardScaler()
X = scaler.fit_transform(X)

# 訓練データとテストデータの分割
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

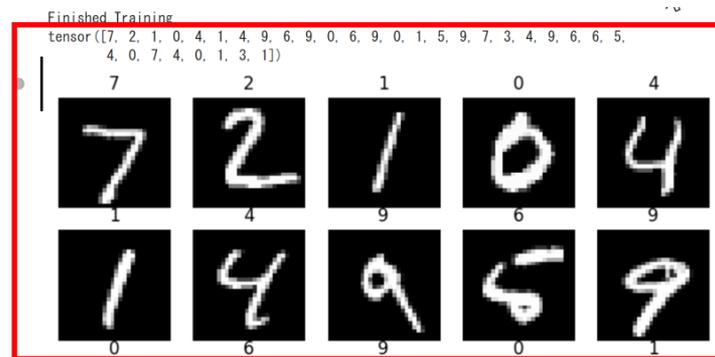
X_train = torch.tensor(X_train, dtype=torch.float32)
y_train = torch.tensor(y_train, dtype=torch.long)
X_test = torch.tensor(X_test, dtype=torch.float32)
y_test = torch.tensor(y_test, dtype=torch.long)

# ニューラルネットワークの定義
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.fc1 = nn.Linear(4, 10) # 入力4次元 (Irisの特徴量)
        self.fc2 = nn.Linear(10, 3) # 出力は3クラス

    def forward(self, x):
        x = torch.relu(self.fc1(x))
        x = self.fc2(x)
        return x

net = Net()
criterion = nn.CrossEntropyLoss()
optimizer = optim.SGD(net.parameters(), lr=0.01)
```

学習の結果、文字認識の能力を獲得



# 機械学習まとめ



## 機械学習の特徴

- データを用いて知的能力を向上
- 自動でデータのパターンを抽出
- さまざまなタスクを自動実行

## 応用事例

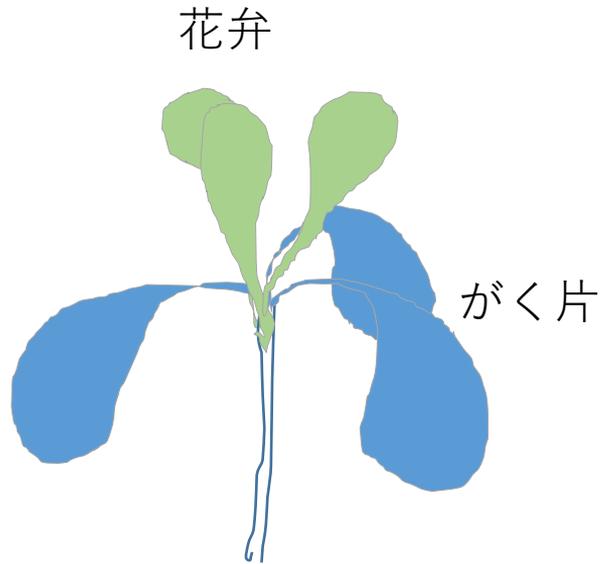
画像理解、自然言語処理、予測  
など多数

## 機械学習の定義：

- **訓練データ**を用いて**学習**し、その結果として知的能力が向上
- **訓練データの追加**により、さらに**知的能力が向上**する可能性



# アヤメ属 (Iris)



- 多年草
- 世界に 150種. 日本に 9種.
- **がく片** Sepal  
3個 (大型で下に垂れる)
- **花弁** Petal  
3個 (直立する)

# Iris データセット

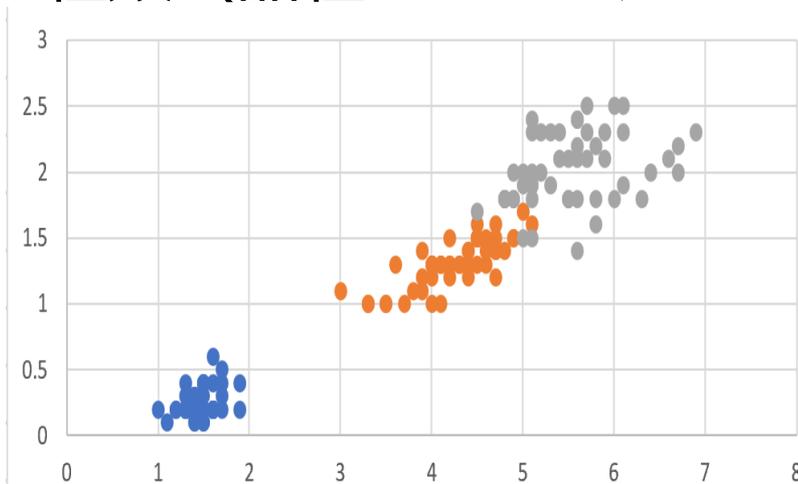
## 【概要】

- 有名なデータセット.
- 3種類のアヤメ (Iris) の花の特徴を記録したもの.

## 【データの内容】

- 150個のサンプル (各種類50個ずつ) .
- 4つの特徴 (**がく片**の長さ と 幅、**花弁**の長さ と 幅) .
- 3つの種類 (品種 : setosa、versicolor、virginica) .

縦軸  
..  
花弁の幅



横軸 : 花弁の長さ

次の **3種類** で色を付けた散布図  
setosa  
versicolor  
virginica

# Iris データセット

## Iris データセット

150 サンプルのうち先頭 10

sepal_length	sepal_width	petal_length	petal_width	species
5.1	3.5	1.4	0.2	setosa
4.9	3	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa

がく片(Sepal)  
の長さ と 幅

花弁(Petal)  
の長さ と 幅

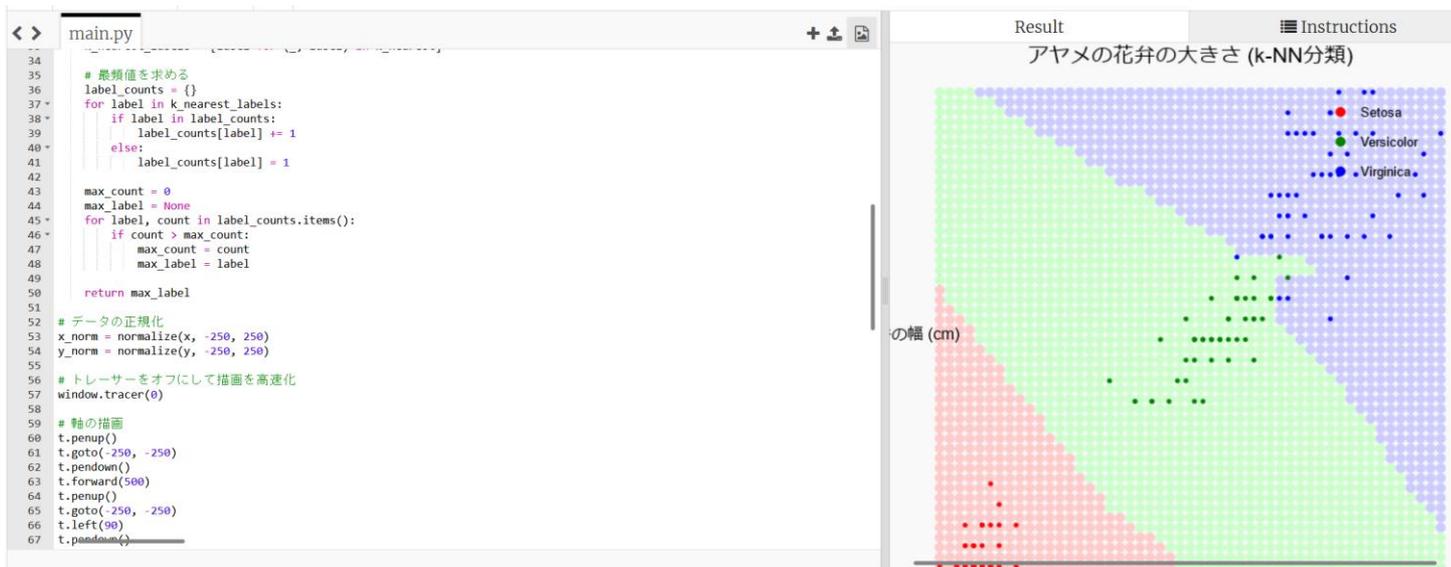
種類

# 分類のモデル

① trinket の次のページを開く

<https://trinket.io/python/918ff17a9ad9>

② 実行結果が、次のように表示されることを確認



- 実行が開始しないときは、「**実行ボタン**」で**実行**
- ソースコードを**書き替えて再度実行**することも可能

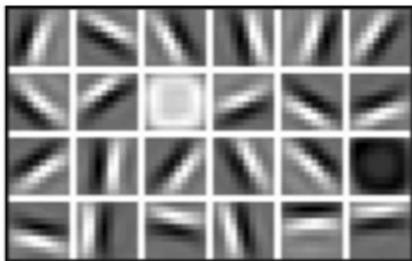
## さまざまなレベルのパターン

### Why Deep Learning?

Hand engineered features are time consuming, brittle, and not scalable in practice

Can we learn the **underlying features** directly from data?

Low Level Features



Lines & Edges

Mid Level Features



Eyes & Nose & Ears

High Level Features



Facial Structure



線や点のレベル 目, 鼻, 耳のレベル 顔の構造のレベル

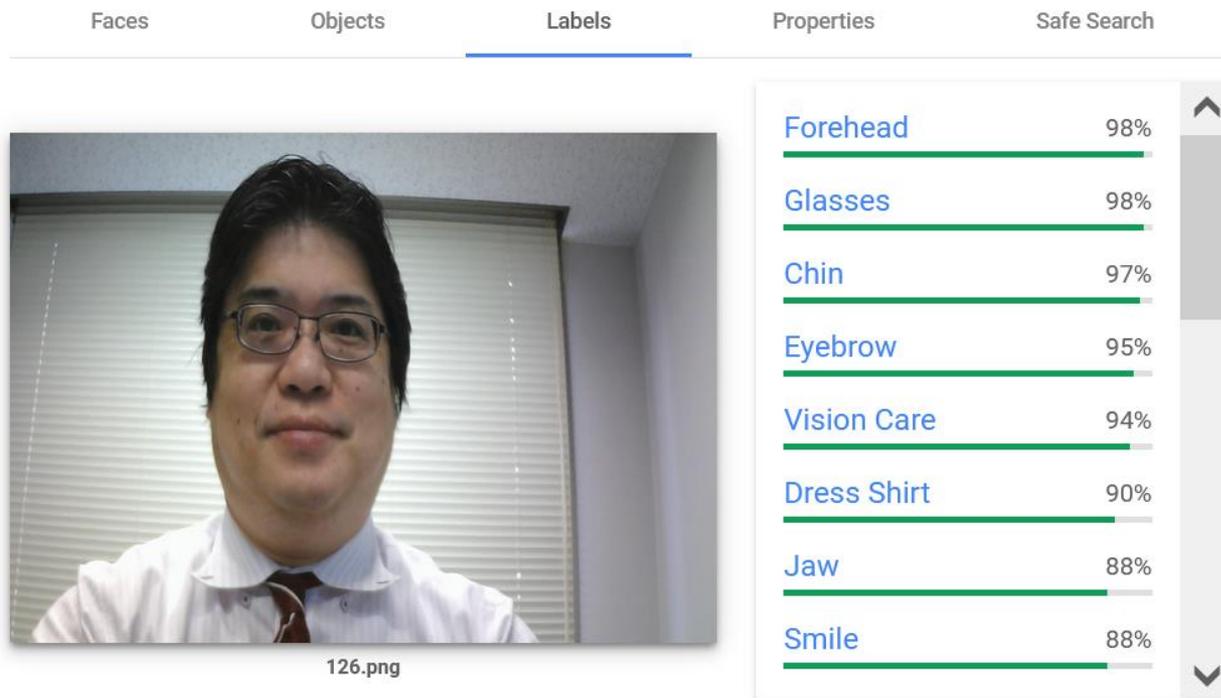
MIT Introduction to Deep Learning | 6.S191,  
[https://www.youtube.com/watch?v=5tvmMX8r\\_OM](https://www.youtube.com/watch?v=5tvmMX8r_OM)  
の「Why Deep Learning」のページ

# 画像分類を行うオンラインサービス



元画像

Faces      Objects      **Labels**      Properties      Safe Search



Label	Confidence
Forehead	98%
Glasses	98%
Chin	97%
Eyebrow	95%
Vision Care	94%
Dress Shirt	90%
Jaw	88%
Smile	88%

126.png

画像分類の結果

URL : <https://cloud.google.com/vision/docs/drag-and-drop>

# 機械学習の特徴まとめ



- **機械学習**は、**データによる学習**を行うための手法
- **大量の学習用データ**を用いることで、**パターンの抽出やデータ分類の精度が向上**
- **機械学習**の特徴の一つに、**学習によって未知のデータにも当てはまるパターンや規則を抽出する能力がある。**
- **機械学習**は、自然言語処理、画像認識、音声認識などの問題に応用され、**我々の生活に浸透しつつある**。大切な分野である。

- いずれも、**データを活用し、知見や結論を導くもの**
- 機械学習やデータサイエンスは、**結果に、必ずしも100%の正確性が保証されているものではなく、結果の根拠の確認が大切**
- 特に、使用するデータの選択、データの前処理で注意が必要。この段階で誤りなどがあると、結果も間違ったものになる
- データを使用するため、**データの漏洩や改竄などの危険があることにも注意**

以上の注意点を理解しながら、実践的な経験を積むことで、機械学習やデータサイエンスのスキルに上達でき、将来、活躍できる。

# まとめ



- **データサイエンス**は、データから有益な情報を導き出すための学問
- **機械学習**は、大量のデータからパターンを抽出し、精度の高いデータ分類などを可能にする手法
- いずれも、ビジネス、医療、工学分野などで**広く活用されている**
- 100%の正確性が保証されているわけではなく、データ選択や前処理で誤りがあると結果も誤ったものになる可能性があることに注意

これらの技術を**理解し、実践的な経験を積む**ことは、**将来の活躍に大いに役立つ。**