

理系基礎演習 第2回

Excelでデータの傾向を把握する

データを視覚化する／特徴をつかむ

白松 俊
(情報工学教育類 准教授)

今回の目的

以下の2項目についてExcelで行う方法を習得

- データをグラフにして視覚化する
 - 棒グラフ, 折れ線グラフ, 円グラフ, 散布図, ...
- データの基本的な処理で特徴をつかむ
 - 合計, 比率, 分布曲線, 平均, 分散, ...

今日の演習で使うデータ

- Moodle からダウンロードできます
 - 教科書著者が提供している「datasets.xlsx」

第1回課題提出(ファイル名:ex-15.xlsx)
第1回課題提出(ファイル名:ex-16.xlsx)

Excelでデータの傾向を把握する

- 教科書2章: データを視覚化する
- 教科書3章: データの特徴をつかむ

データセットのダウンロード

- この演習で使うデータセット
- 演習問題3.8の平成20年度の人口のデータセット

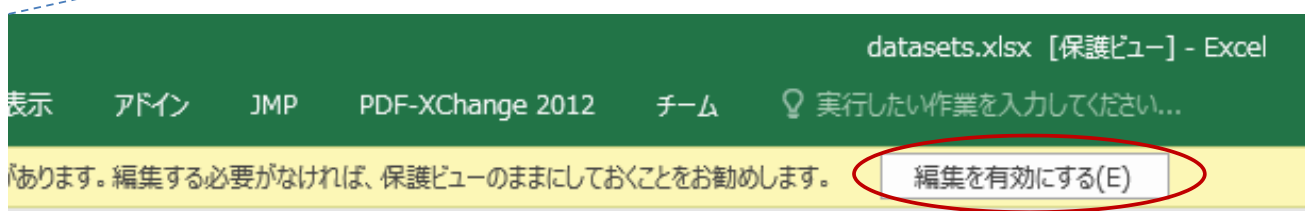
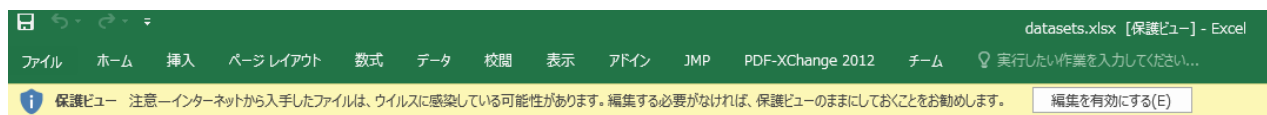
プレゼンテーション技法

- PowerPointでのプレゼン

- Moodleに入れない ⇒ 著者によるWebページから
 - 著者提供ページの短縮URL: <http://bit.ly/excel-stat>

編集を有効にする

- ダウンロードしたdatasets.xlsxをExcelで開くと,
「ウイルスに感染している可能性があります」
などと脅される ➡ 気にせず編集を有効にします



今日の内容

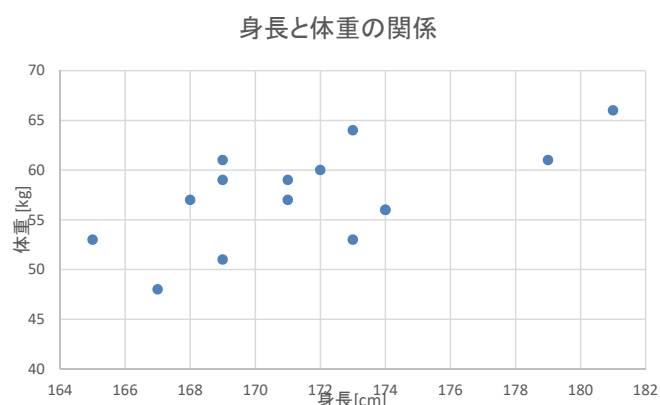
- 教科書 2 章: データを視覚化する
 - 棒グラフ, 折れ線グラフ, 円グラフ, ...
- 教科書 3 章: データの特徴をつかむ
 - 合計, 比率, 分布曲線, 平均, 分散, ...
- おまけ: 便利なTips
 - スクリーンショットを撮る
 - 結果をプレゼン資料や報告書に貼るとき便利
 - OneDriveでデータを共有する
 - 複数人で協力しながらデータを入力できる

なぜグラフで視覚化するのか

- 適切な種類のグラフを使えば,
一目でデータの傾向を理解してもらえる

身長	体重
179	61
172	60
165	53
171	57
169	51
169	59
173	53
174	56
169	61
171	59
167	48
174	56
168	57
173	64
181	66

表だと一目ではわかりにくい



グラフをうまく使うと一目で傾向がわかる

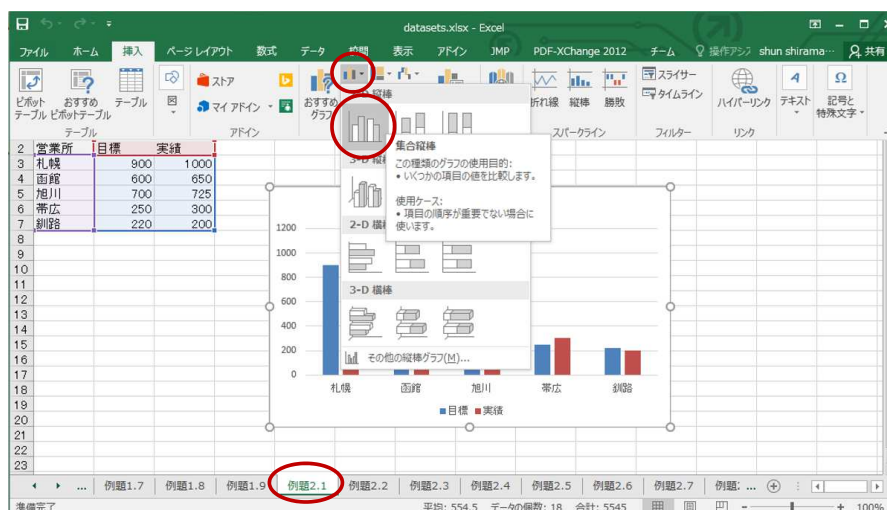
どんな時にどのグラフを使う？

- 棒グラフ: 幾つかのデータを比較したいとき
 - 例: 「従来手法」と「提案手法」の性能を比較
- 折れ線グラフ: 時系列的な変化を見たいとき
 - 例: 1800年から2015年までのCO2濃度の変化
- レーダーチャート: 複数項目のバランスを見たいとき
 - 例: 街ごとに「物価」「交通網」「福祉」などの項目で暮らしやすさを評価
- 円グラフ, ドーナツグラフ: 割合や構成比を見たいとき
- 散布図: 2つの変数の関係を見たいとき
- ヒストグラム: 頻度の分布(度数分布)を見たいとき
 - 例: 納税額の人数分布
- 2軸のグラフ: 単位の違う2種類のデータの変化傾向などを同時に見たいとき

縦棒グラフ (§2.2)

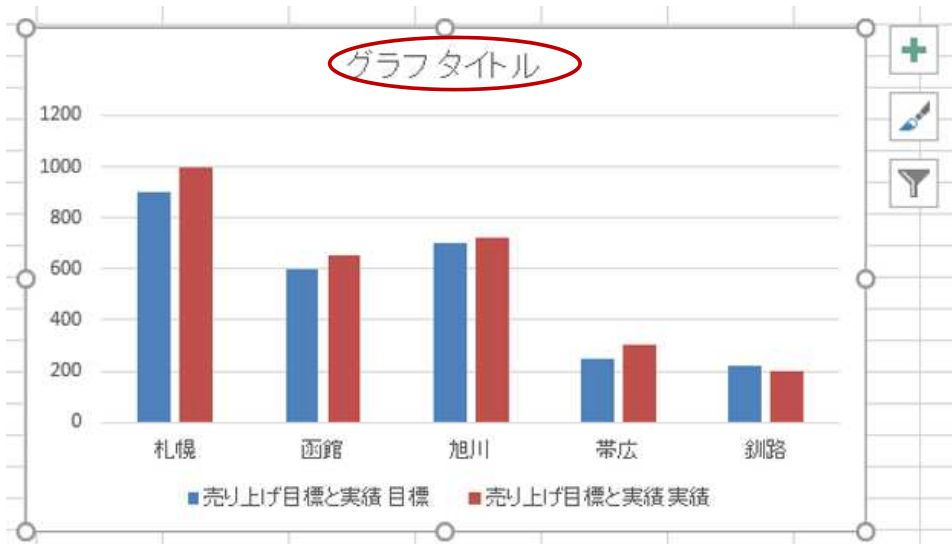
dataset.xlsxの例題2.1シートを開く

1. グラフにするセル範囲 (A2:C7) を選択
2. 「挿入」タブをクリック
3. 「グラフ」欄内の「縦棒」をクリック
4. 「2-D縦棒」⇒「集合縦棒」をクリック



縦棒グラフ (§2.2)

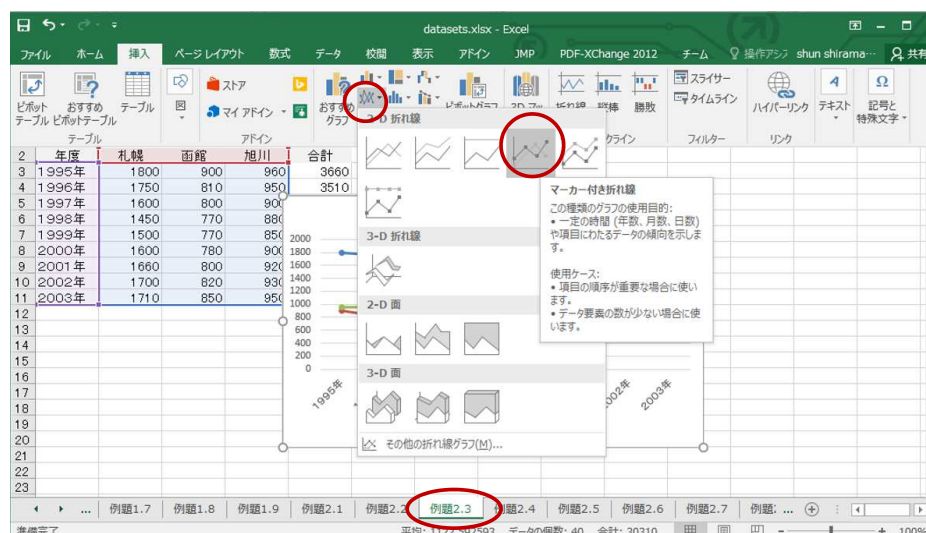
- グラフタイトルはダブルクリックで編集可能に



折れ線グラフ (§2.4)

datasets.xlsxの例題2.3シートを開く

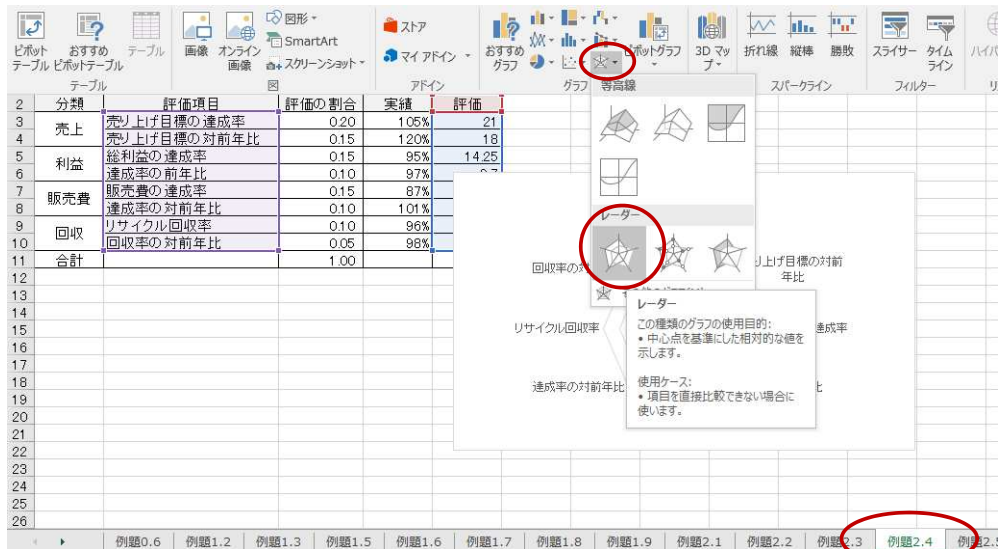
1. グラフにするセル範囲 (A2:D11) を選択
2. 「挿入」タブをクリック
3. 「グラフ」欄内の「折れ線」をクリック
4. 「2-D折れ線」⇒「マーカー付き折れ線」をクリック



レーダーチャート (§2.5)

datasets.xlsxの例題2.4シートを開く

- 教科書の〔解答〕の手順を参照

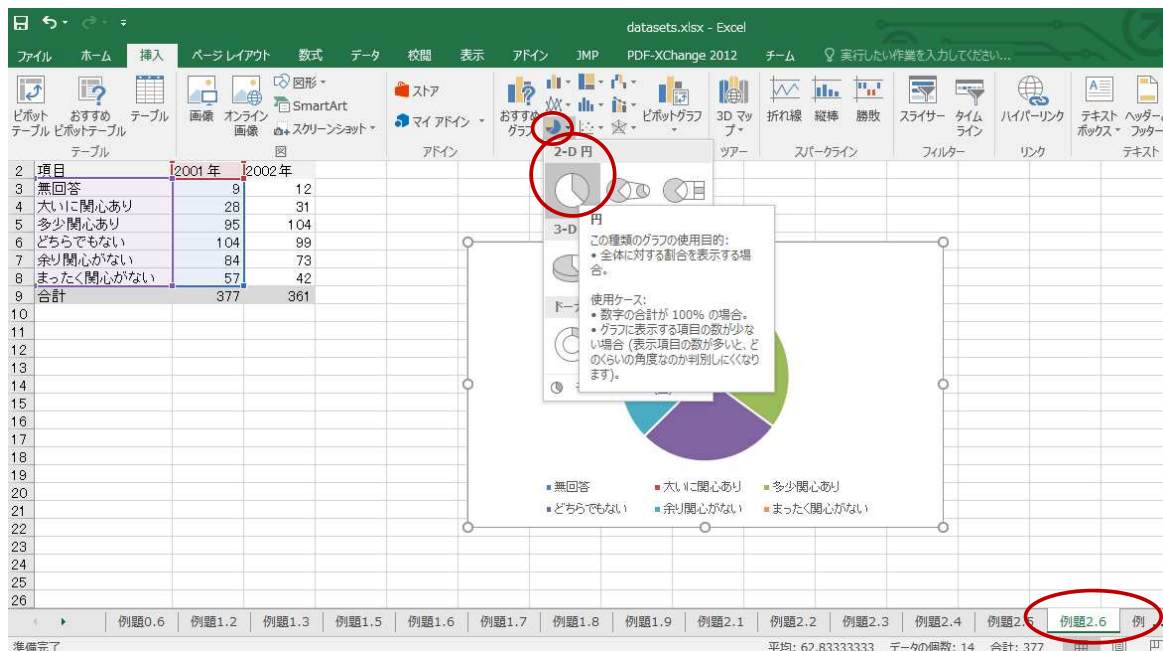


例題2.5についても教科書の手順でやってみる

円グラフ (§2.6)

datasets.xlsxの例題2.6シートを開く

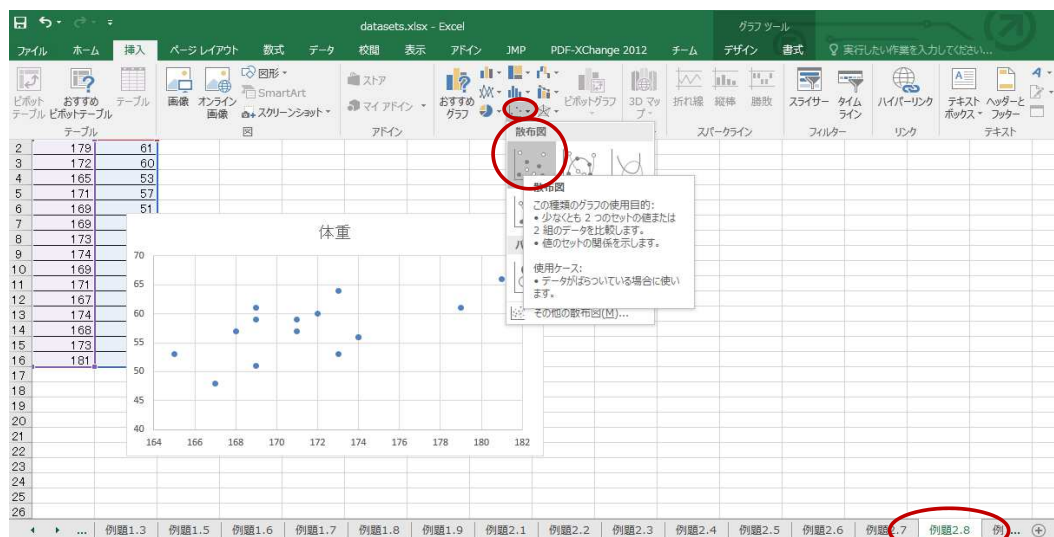
- 教科書の〔解答〕の手順を参照



散布図 (§2.8)

datasets.xlsxの例題2.8シートを開く

- 教科書の「解答」の手順を参照
 - 右ページ「注意」を参考にして縦軸の範囲を40以上に変える
 - 縦軸ラベルを「体重[kg]」に、横軸ラベルを「身長[cm]」に設定すると、軸の意味や単位を明示できる



ヒストグラム (§2.9)

準備: 「分析ツール」アドインを読み込む

- [ファイル] タブをクリックし、[オプション] をクリックします。
- [アドイン] をクリックし、[管理] ボックスの一覧の [Excel アドイン] が選ばれているのを確認して [設定] をクリックします。
- [有効なアドイン] ボックスの一覧で、[分析ツール] チェック ボックスをオンにし、[OK] をクリックします。
 - [有効なアドイン] ボックスの一覧に [分析ツール] が表示されない場合は、[参照] をクリックしてアドインファイルを見つけます。
 - 分析ツールが現在コンピューターにインストールされていないというメッセージが表示されたら、[はい] をクリックして分析ツールをインストールします。

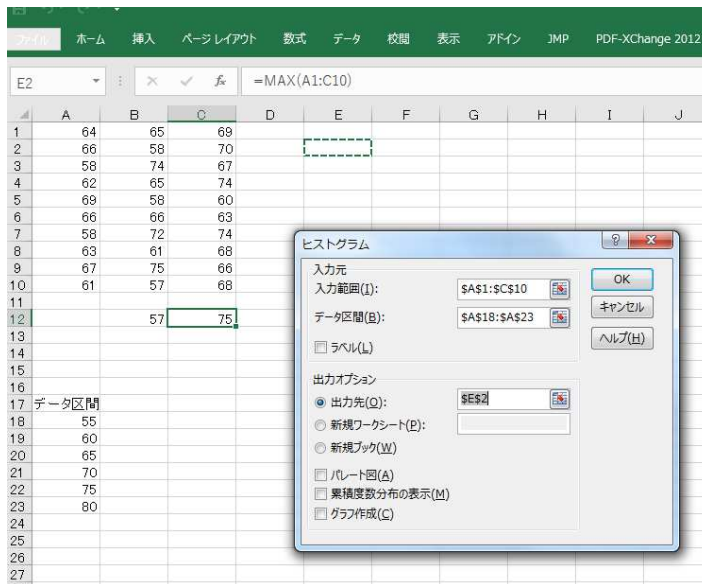
これで、[データ] タブの [分析] で [データ分析] を利用できます。



ヒストグラム (§2.9)

datasets.xlsxの例題2.9シートを開く

1. 「解答」の手順を参照し、度数分布表を作成

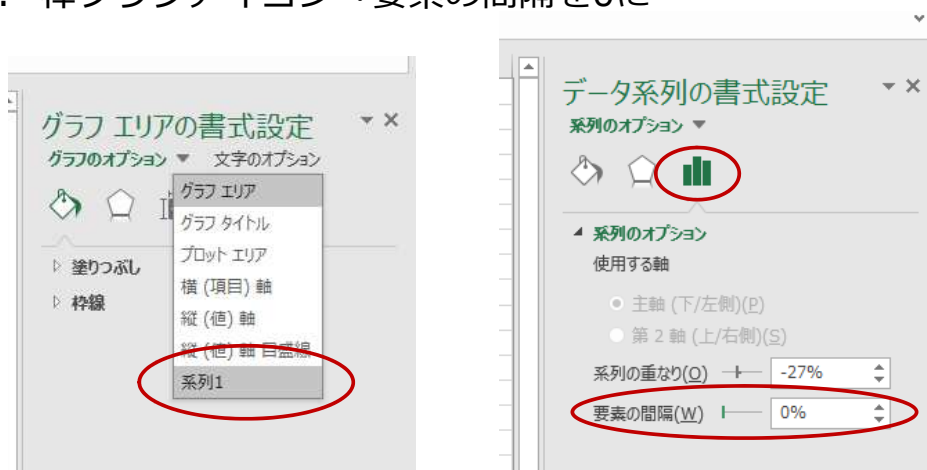


ヒストグラム (§2.9)

2. 「解答」の手順を参照し、ヒストグラムを描画

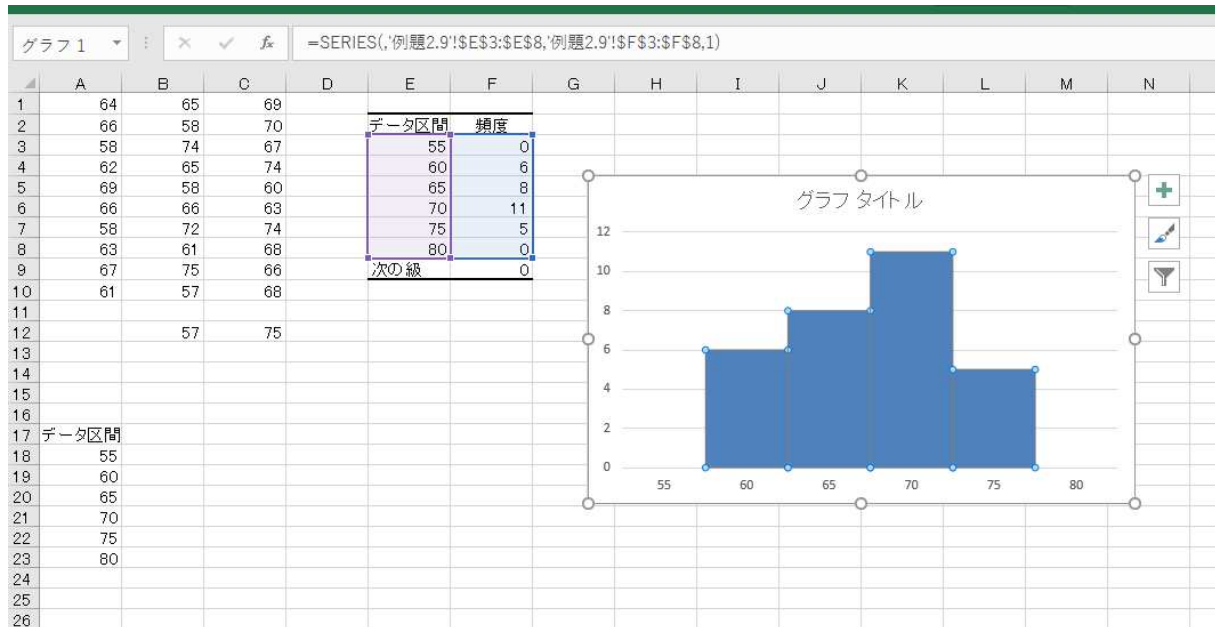
– 「棒の間の隙間をなくす」のやり方がExcel 2016では少し違う

1. グラフ上で右クリック→「グラフ エリアの書式設定」
2. グラフのオプション→「系列1」
3. 棒グラフアイコン→要素の間隔を0に



ヒストグラム (§2.9)

- こうなればOK
- グラフタイトルや軸ラベル(横軸: データ区間, 縦軸: 頻度)も設定してみよう



演習問題

- 教科書P.51～52の演習問題2.1, 2.4, 2.7, 2.8
 - 演習2.1の成果はファイルex-21.xlsxに保存
 - 2.4はex-24.xlsx, 2.7はex-27.xlsx, 2.8はex-28.xlsx
 - Moodle提出期限: 5/9(火) 24:00
- 周囲の人と話し合っても構いません

今日の内容

- 教科書 2 章: データを視覚化する
 - 棒グラフ, 折れ線グラフ, 円グラフ, ...
- 教科書 3 章: データの特徴をつかむ
 - 合計, 比率, 分布曲線, 平均, 分散, ...
- おまけ: 便利なTips
 - スクリーンショットを撮る
 - 結果をプレゼン資料や報告書に貼るとき便利
 - OneDriveでデータを共有する
 - 複数人で協力しながらデータを入力できる

量的データの集計

- 量的データと質的データ
 - 質的: 性別 (男性/女性), 血液型 (A/B/O/AB), ...
 - 量的: 長さ, 速さ, 点数, 成功率, ...
- 量的データの特徴をつかむための集計方法
 - 合計
 - 比率 (パーセンテージ)
 - 分布 (ヒストグラムで見る, 分布曲線で見る)
 - 基本統計量: 平均, 最頻値, 中央値, 分散, 標準偏差

合計 (§3.1)

datasets.xlsxの例題3.1-3.2シートを開く

1. セールスマンA～Jの売り上げを入力する
2. L2セルに「=SUM(B2:K2)」と入力してEnterキー
 - 「=SUM()」と入力しておき、括弧()の間にカーソルをあわせてB2からK2まで選択すると、B2:K2と自動入力される
 - 他のセルをクリックする前にEnterキーを忘れずに

Figure 1-10: Entering a formula into a cell. The formula bar shows the formula `=SUM(B2:K2)` being entered into cell L2. The spreadsheet shows data for sales volume, sales increase, and sales ratio across columns A to Q.

比率 (§3.2)

同じシートのまま, p.56〔解答〕③以降の手順を参照

- 「\$L2」は複合参照 (p.27 解説1.5)
 - \$をつけておくと、コピーしたときにずれない

[illegible]

ヒストグラム (§3.3-4)

datasets.xlsxの例題3.3-3.5シートを開く

1. 例題3.3〔解答〕の手順でデータ区間を作る
2. 例題3.4〔解答〕の手順でデータを集計
3. 例題3.5〔解答〕の手順でヒストグラムを描く

分布曲線 (§3.5)

同じシートのまま, p.62〔解答〕の手順を参照して
正規分布の近似曲線を描く

- NORMDIST関数を使う

平均 (§3.6-7)

datasets.xlsxの例題3.7-3.9シートを開く

- 例題3.7: SUM関数で総和を出し, 個数で割る方法
- 例題3.8: AVERAGE関数を使う方法
- 例題3.9: オートカルクを使う方法
 - B2:B7を選択するだけ → 平均などがステータスバーに表示される

最頻値と中央値 (§3.8)

datasets.xlsxの例題3.10シートを開く

- 〔解答〕の手順を参照
 - 最頻値: MODE関数を使う
 - 中央値: MEDIAN関数を使う

データの存在する範囲 (§3.9)

datasets.xlsxの例題3.11シートを開く

- 「解答」の手順を参照し、範囲を求める
 - 教科書では、範囲 = 最大値 - 最小値 と定義
 - 最大値: MAX関数を使う
 - 最小値: MIN関数を使う

分散と標準偏差 (§3.10-11)

datasets.xlsxの例題3.12シートを開く

- 「解答」の手順を参照し、分散と標準偏差を求める
 - 分散: VAR関数を使う
 - 標準偏差: STDEV関数を使う
 - なお、標準偏差 = $\sqrt{\text{分散}}$

演習問題

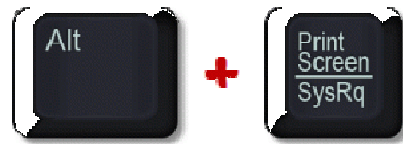
- 教科書P.86～87の演習問題3.1, 3.3, 3.4, 3.8
 - ex-31.xlsx, ex-33.xlsx, ex-34.xlsx, ex-38.xlsx
にそれぞれ保存
 - Moodle提出期限: 5/9(火) 24:00
- 周囲の人と話し合っても構いません

今日の内容

- 教科書 2 章: データを視覚化する
 - 棒グラフ, 折れ線グラフ, 円グラフ, ...
- 教科書 3 章: データの特徴をつかむ
 - 合計, 比率, 分布曲線, 平均, 分散, ...
- おまけ: 便利なTips
 - スクリーンショットを撮る
 - 結果をプレゼン資料や報告書に貼るとき便利
 - OneDriveでデータを共有する
 - 複数人で協力しながらデータを入力できる

スクリーンショットを撮る

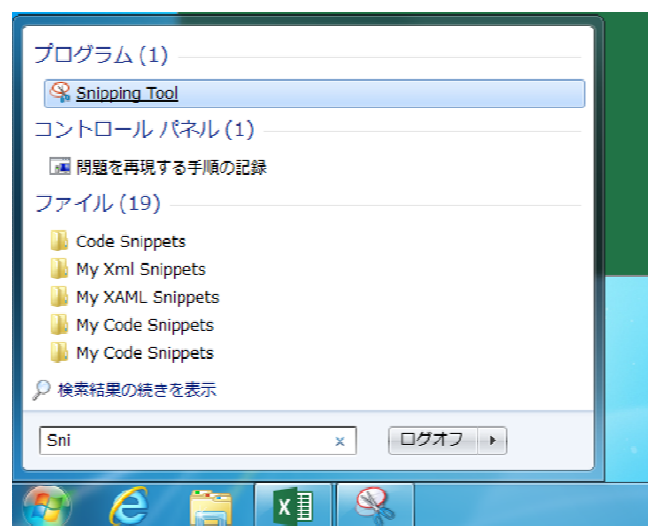
- キーボード右上の[Print Screen]ボタンでスクショがクリッピング領域にコピーされる
 - [Print Screen]だけだと全画面
 - [Alt]+[Print Screen]だとアクティブウィンドウだけ



- コピーしたら, WordやPower Pointに貼り付け
 - [Ctrl]+[V]で貼り付け

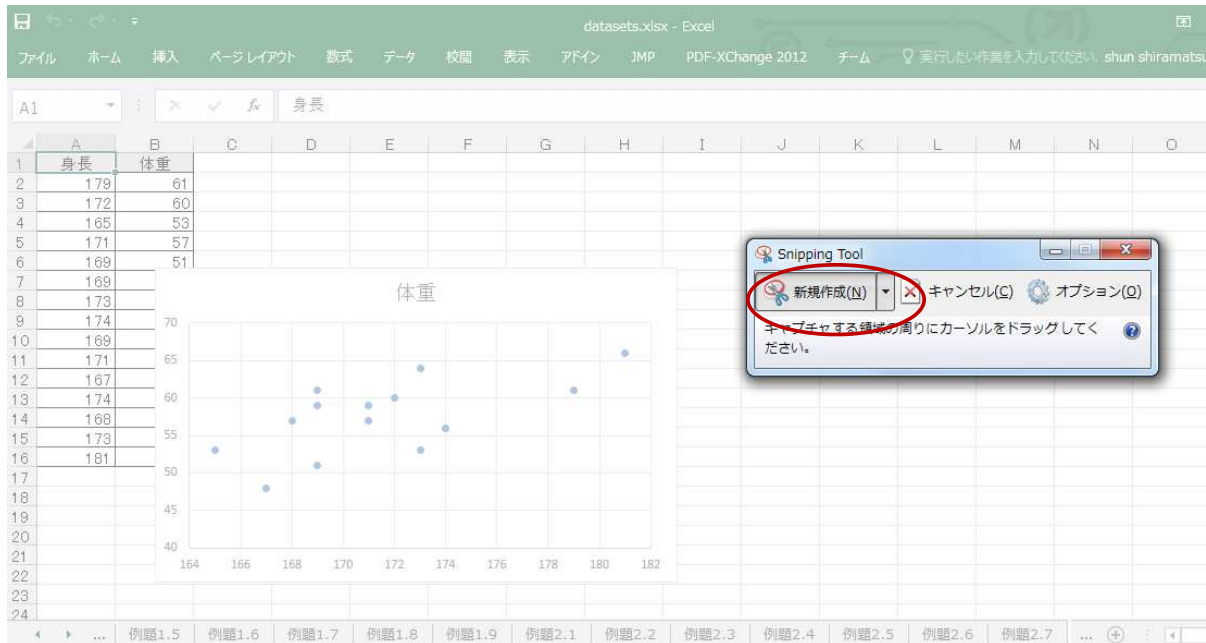
Snipping Tool

- 範囲を指定して撮りたいならSnipping Tool
- 起動: まず左下のスタートボタン
 - 「プログラムとファイルの検索」でSni と入力
 - 上にSnipping Toolが出てくるので選択して起動



Snipping Tool

- 「新規作成」を押し、キャプチャ範囲をマウスで指定
- 「編集」→「コピー」でクリッピング領域にコピー
 - コピーしたらWordやPower Pointに貼り付け ([Ctrl]+[V]で)



今日の内容

- 教科書 2 章: データを視覚化する
 - 棒グラフ, 折れ線グラフ, 円グラフ, ...
- 教科書 3 章: データの特徴をつかむ
 - 合計, 比率, 分布曲線, 平均, 分散, ...
- おまけ: 便利なTips
 - スクリーンショットを撮る
 - 結果をプレゼン資料や報告書に貼るとき便利
 - OneDriveでデータを共有する
 - 複数人で協力しながらデータを入力できる

OneDrive for Business

- <https://onedrive.live.com/about/ja-JP/business/>
- ExcelやWordなどのファイルをWebブラウザで共有
- 右上の「サインイン」から
 - メールアドレス: [基盤ID]@ict.nitech.ac.jp
 - 基盤IDに設定済みのパスワードでOK



Excelファイルをアップロード



Excelファイルを他の人と共有

1. 自分の基盤IDでOneDriveを開く
 - WebブラウザからOffice365のサインインすると、Webブラウザ上でフォルダが開けます。
2. OneDrive内に、データ共有用のフォルダを作成
3. 作成したフォルダを共有できるようアクセス権を設定する
 - フォルダー一覧画面で、共有したいフォルダの共有アイコンを表示
 - 「ユーザの招待」で共有する相手を追加
 - 招待した相手にメールが行くはず
4. 共有したいファイルをそのフォルダにアップロード