

# 情報リテラシー（第5回 後期）

## 情報の加工2

# 今日のねらい

- オフィスソフトを使って数学的な処理を行うことができる
- データの取得・可視化・分析の基本的なスキルを習得できる
- 実践的な演習を通じてデータ処理の流れを体験できる

# 演習1：ローカル保存の設定

## OneDriveではなくローカルに保存する設定

### 設定手順

1. **Excelを起動**（まだファイルは開かない）
2. 「**ファイル**」→「**オプション**」をクリック
3. 左側メニューから「**保存**」を選択
4. 「**既定でコンピューターに保存する**」に**チェック**を入れる
5. 「**既定のローカルファイルの保存場所**」を確認
  - 推奨：ドキュメントフォルダやデスクトップ
6. 「**OK**」をクリック

### なぜローカル保存？

- **作業がスムーズ**：同期待ちがない
- **オフラインでも安心**：ネット環境に左右されない
- **ファイル管理が明確**：保存場所がわかりやすい

# データの取得

## データ取得の重要性

- 分析の第一歩は**正しいデータを集めること**
- データの質が分析結果の質を左右する

## データ取得の方法

- **手動入力**： アンケート結果、観測データ
- **ファイルからの読み込み**： CSV、Excel、テキストファイル
- **Web上のデータ**： オープンデータ、API経由のデータ
- **データベース**： 企業の業務データ、公的データベース

## 注意点

- データの信頼性を確認する
- 個人情報保護に配慮する
- データの出典を明記する

# オープンデータの活用

## オープンデータとは

- **誰でも自由に利用・再配布できる** 公開データ
- 政府、自治体、研究機関などが提供
- 商用利用も可能（利用規約を確認）

## 主なオープンデータサイト

- **e-Stat（政府統計の総合窓口）**： 人口、経済、労働統計など
- **データカタログサイト**： 各省庁・自治体のデータ
- **気象庁**： 気象データ、地震情報
- **国土地理院**： 地図データ、標高データ

## 利用時の注意点

- **利用規約を確認**する（著作権、再配布条件）
- **データの出典を明記**する
- データの**更新日**や**調査方法**を確認する

# Excelでのデータ取得

## データの入力とインポート

- **直接入力**：セルに手動でデータを入力
- **CSVファイルの読み込み**：「データ」→「テキストまたはCSVから」
- **Webからの取得**：「データ」→「Webから」

## データの整形

- **不要な行・列の削除**
- **データ型の変換**（文字列→数値、日付など）
- **欠損値の処理**（空白セルの確認と対処）
- **重複データの削除**

## データの確認

- データ範囲の確認、データ件数の把握
- 各列の意味を理解する

# データの並べ替え

## 並べ替え（ソート）

データを特定の順序に並べ替える

### 基本的な並べ替え

1. 並べ替えたい列のセルを選択
2. 「データ」タブ→「昇順」または「降順」
3. **昇順**：小さい順、A→Z、古い→新しい
4. **降順**：大きい順、Z→A、新しい→古い

### 複数列での並べ替え

- 「データ」→「並べ替え」で詳細設定
- 例：クラスごとに並べ替え、その中で成績順

# データの可視化

## 可視化の目的

- データの傾向を一目で把握する
- 数値だけでは見えないパターンを発見する
- 他者への説明を容易にする

## 代表的なグラフの種類

- 棒グラフ：カテゴリ別の比較
- 折れ線グラフ：時系列の変化
- 円グラフ：全体に占める割合
- 散布図：2つの変数の関係
- ヒストグラム：データの分布

→ 目的に応じて適切なグラフを選択することが重要



# Excelでのグラフ作成

## グラフ作成の基本手順

1. グラフにしたいデータ範囲を選択
2. 「挿入」タブから適切なグラフを選択
3. グラフタイトル、軸ラベルを追加
4. 凡例、データラベルを調整

## グラフ作成のポイント

- **わかりやすいタイトル**をつける
- **軸ラベル**で単位を明示する
- **色使い**に配慮する（色覚多様性への配慮）
- **データラベル**で具体的な数値を表示
- 過度な装飾は避け、**シンプルに**

# 尺度水準（データの種類）

## 尺度水準とは

データの性質による分類。適切な分析方法を選ぶために重要

## 4つの尺度水準

### 1. 名義尺度（質的データ）

- 分類・区別するためのラベル
- 例：性別、血液型、都道府県名
- 可能な分析：度数、最頻値、円グラフ

### 2. 順序尺度（質的データ）

- 順序・順位に意味があるが、間隔は等しくない
- 例：成績（優・良・可）、満足度（5段階評価）
- 可能な分析：中央値、順位相関

# 尺度水準（続き）

## 3. 間隔尺度（量的データ）

- **間隔が等しい**が、絶対的なゼロ点がない
- 例：気温（℃）、西暦年
- 可能な分析：平均値、標準偏差、相関係数
- 注意：「2倍」などの比率は意味を持たない（20℃は10℃の2倍暑いわけではない）

## 4. 比例尺度（量的データ）

- **間隔が等しく、絶対的なゼロ点がある**
- 例：身長、体重、金額、時間
- 可能な分析：すべての統計分析が可能
- 比率にも意味がある（100万円は50万円の2倍）

## 重要ポイント

**データの尺度水準によって使える統計手法が変わる！**

# データの基本的な分析

## 記述統計

データの特徴を数値で要約する

### 代表値

- **平均値 (mean)** : データの合計÷個数
- **中央値 (median)** : データを並べたときの中央の値
- **最頻値 (mode)** : 最も頻繁に現れる値

### ばらつき

- **範囲 (range)** : 最大値－最小値
- **標準偏差 (standard deviation)** : データの散らばりの程度
- **分散 (variance)** : 標準偏差の2乗

# Excelでの基本的な統計関数

## よく使う関数

- **AVERAGE()** : 平均値を計算

```
=AVERAGE(A2:A10)
```

- **MEDIAN()** : 中央値を計算

```
=MEDIAN(A2:A10)
```

- **MODE.SNGL()** : 最頻値を計算

```
=MODE.SNGL(A2:A10)
```

- **STDEV.S()** : 標準偏差を計算

```
=STDEV.S(A2:A10)
```

- **MAX()**、**MIN()** : 最大値、最小値

```
=MAX(A2:A10)    =MIN(A2:A10)
```

## 演習2：テストの成績分析

以下の10人の数学テストの点数を分析してみよう。

**点数：** 85, 72, 90, 68, 88, 95, 78, 82, 70, 92

### 課題

1. Excelに点数データを入力
2. 平均値を計算（AVERAGE関数）
3. 中央値を計算（MEDIAN関数）
4. 最大値と最小値を計算（MAX、MIN関数）
5. 標準偏差を計算（STDEV.S関数）

**どんなことがわかるか考えてみよう！**

# 演習2の解答

## 計算結果

- **平均値**： 82点

```
=AVERAGE(A2:A11) → 82
```

- **中央値**： 83.5点

```
=MEDIAN(A2:A11) → 83.5
```

- **最大値**： 95点、**最小値**： 68点

```
=MAX(A2:A11) → 95
```

```
=MIN(A2:A11) → 68
```

- **標準偏差**： 約9.63点

```
=STDEV.S(A2:A11) → 9.63
```

## 分析結果の解釈

- 平均値と中央値がほぼ近い → データは比較的対称的
- **標準偏差が約9.6** → ある程度ばらつきがある

# 相関分析

## 相関とは

- **2つの変数の関係性**を調べる分析
- 一方が増えるともう一方も増える？減る？関係ない？

## 相関係数

- **-1から+1の値**をとる
- **+1に近い**：強い正の相関（片方が増えるともう片方も増える）
- **-1に近い**：強い負の相関（片方が増えるともう片方は減る）
- **0に近い**：相関なし（関係性が薄い）

## 例

- 勉強時間と成績 → **正の相関**
- 気温とコート販売数 → **負の相関**
- 身長と学力 → **相関なし**



# Excelでの相関分析

## CORREL関数

2つのデータ範囲の相関係数を計算

```
=CORREL(範囲1, 範囲2)
```

## 散布図の作成

1. 2つの変数のデータを準備
2. データ範囲を選択
3. 「挿入」→「散布図」を選択
4. 近似曲線を追加して傾向を確認

## 注意点

- **相関関係≠因果関係**
- 相関があっても原因と結果の関係とは限らない
- 第三の要因が関係している可能性も

## 演習3：確率統計と線形代数の成績の関係

以下のデータで確率統計と線形代数の成績の相関を分析しよう。

学生	確率統計（点）	線形代数（点）
A	55	75
B	60	85
C	88	80
D	75	80
E	72	85
F	83	90
G	70	70
H	85	100

### 課題

1. データをExcelに入力
2. 散布図を作成
3. 相関係数を計算（CORREL関数）

# 提出課題

## 課題内容

演習3の散布図をExcelファイルで提出してください。

## 提出物

- 演習3のデータと散布図を含むExcelファイル
- 相関係数の計算結果も記載すること

## 提出方法

指定されたフォームから提出

## 確認ポイント

- グラフタイトルと軸ラベルが適切についているか
- データが正しく入力されているか
- 相関係数が計算されているか

# まとめ

- **演習1：ローカル保存の設定**で作業環境を整える
- **Excel**を使えば基本的なデータ処理が可能
- **統計関数**で記述統計を計算できる
- **グラフ化**でデータの傾向を視覚的に理解
- **相関分析**で変数間の関係を調べられる

**今日学んだスキルを様々な場面で活用しよう！**

**感想をチャット欄に書こう！**