

速習！情報 ～共通テスト対策講座～

データベース

**ビッグデータ、オープンデータ、テキストマイニング、
データベース管理システム、関係データベース、表、行、
列、
主キー、外部キー、複合キー、射影、選択、結合**

データベース

赤字：用語も意味も理解が必要、緑字：意味や概念の理解が必要

行政機関や企業の組織内、ネット上などで大量のデータが日々生まれている

- **ビッグデータ**：人やモノ、社会、自然などから収集される大量のデータのこと

速習!情報
@info-study · チャンネル登録者数 30人 · 36本の動画
高校情報!で共通テストに必要な内容のみにギュッと絞った動画解説をしています。 >

チャンネル登録

ホーム 動画 再生リスト コミュニティ 検索

作成した再生リスト

- 速習!情報 ~共通テスト対策講座~
ネットワークの仕組み
7本の動画
情報!「情報通信ネットワーク」共通テスト対策講座
更新: 2日前
再生リストの全体を見る
- 速習!情報 ~共通テスト対策講座~
モデル化とシミュレーション
2本の動画
情報!「モデル化とシミュレーション」共通テスト対策講座
再生リストの全体を見る
- 速習!情報 ~共通テスト対策講座~
アルゴリズム
7本の動画
情報!「プログラミングの基本」共通テスト対策講座
再生リストの全体を見る
- 速習!情報 ~共通テスト対策講座~
コンピュータの構成
4本の動画
情報!「コンピュータの仕組み」共通テスト対策講座
再生リストの全体を見る

出典：速習!情報 YouTubeチャンネル (<https://www.youtube.com/@info-study>)

行政機関や企業の組織内、ネット上などで大量のデータが日々生まれている

- ・ **オープンデータ**：行政機関などから提供される無償の公開データで、コンピュータが扱いやすく、二次利用可能なルールが設定されたもの

e-GOV データポータル

Language ログイン 新規登録

e-Govポータル

中央行政のオープンデータポータルです。
オープンデータをご自身のプロジェクトや業務にご活用ください。

データセット >

データセットとは、ファイルやURLなどの「オープンデータ」が登録された入れ物を指します。
データポータルでは、複数の切り口からデータセットを探すことができます。

データセットをキーワードで検索

出典：e-GOV (<https://data.e-gov.go.jp/info/ja>)

1. 学校数・学級数・児童生徒数(学校種別)

(1) 全国計

令和3年5月1日現在

区分	学校数	学級数	学級数	学級数	児童生徒数
		単式・複式	特別支援学級	合計	
	校	cl	cl	cl	人
小学校	19173	219422	51469	270891	6144915
中学校	9409	91246	21615	112861	2986670
小中計	28582	310668	73084	383752	9131585
高等学校	3555	58046	***	58046	2000069
特別支援学校	1100	36059	***	36059	142525
幼稚園	3965	12356	***	12356	201808
計	37202	417129	73084	490213	11475987

(注) 学校基本調査による。

出典：e-Stat(政府統計の総合窓口)「公立学校施設実態調査_令和3年度」

行政機関や企業の組織内、ネット上などで大量のデータが日々生まれている

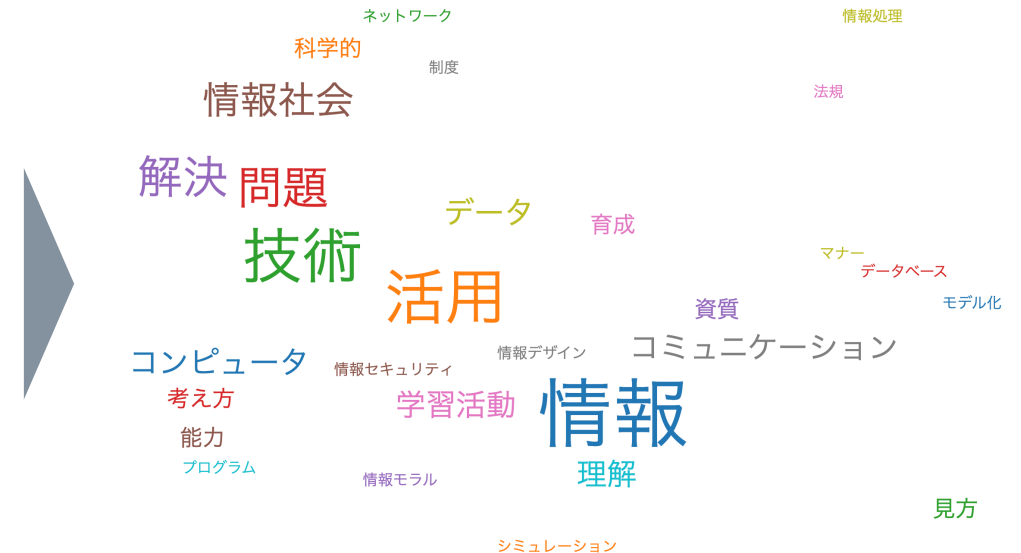
- ・ **テキストマイニング**：文章を単語や文節で区切り、それらの出現傾向や他との関係性などを分析する手法

「情報I」の目標は、次のように示されている。

情報に関する科学的な見方・考え方を働かせ、情報技術を活用して問題の発見・解決を行う学習活動を通して、問題の発見・解決に向けて情報と情報技術を適切かつ効果的に活用し、情報社会に主体的に参画するための資質・能力を次のとおり育成することを旨とする。

- (1) 効果的なコミュニケーションの実現、コンピュータやデータの活用について理解を深め技能を習得するとともに、情報社会と人との関わりについて理解を深めるようにする。
- (2) 様々な事象を情報とその結び付きとして捉え、問題の発見・解決に向けて情報と情報技術を適切かつ効果的に活用する力を養う。
- (3) 情報と情報技術を適切に活用するとともに、情報社会に主体的に参画する態度を養う。

出典：高等学校学習指導要領(平成30年告示)解説 情報編 平成30年7月



行政機関や企業の組織内、ネット上などで大量のデータが日々生まれている

- **構造化データ**：集計や分析がしやすいようにデータ間の関係性を使って構造化されたデータ
 例) 表にまとめられたデータ
- **非構造化データ**：文書や画像、音声、動画などの構造化されていないデータ
 例) SNSでのやり取り、写真、動画
- **情報技術の発達・普及により、非構造化データが急激に増加している**

1. 学校数・学級数・児童生徒数(学校種別)

(1) 全国計 令和3年5月1日現在

区分	学校数	学級数			児童生徒数	
		単式・複式	特別支援学級	合計		
	校	cl	cl	cl	人	
小学校	19173	219422	51469	270891	6144915	
中学校	9409	91746	21615	112861	2986670	
小中計	28582	構造化データ			383752	9131585
高等学校	3555	58046	***	58046	2000069	
特別支援学校	1100	36059	***	36059	142525	
幼稚園	3965	12356	***	12356	201808	
計	37202	417129	73084	490213	11475987	

(注) 学校基本調査による。

出典：e-Stat(政府統計の総合窓口)「公立学校施設実態調査_令和3年度」



出典：速習!情報 YouTubeチャンネル (<https://www.youtube.com/@info-study>)
 無断転載禁止 Copyright © 速習!情報 <https://info-study.jp>

データを検索・集計しやすくするためにデータベースが利用される

- **データベース**
 - データを使用しやすいように整理して、検索や集計などの活用ができるようにしたもの
- **データベース管理システム=DBMS** (DataBase Management System)
 - 以下のような機能を持ち、データベースを効率よく安全に管理するためのシステム
 - データの一貫性：複数ユーザで共有でき、同時にデータを操作しても矛盾が起きない
 - データの整合性：重複データや不正なデータが登録・更新されない
 - データの独立性：データとプログラムを分けて管理する
 - データの機密性：ユーザ認証を行い、データによりアクセス権を設定する
 - データの可用性：障害に備えてバックアップを取得し、障害発生時に復旧を行う
- **関係データベース=リレーショナルデータベース**
 - 複数の表を組み合わせてデータを管理するDBMS。幅広く使われている

関係データベースは、行と列からなる表を複数もっている

- **表(テーブル)**：データを格納する表。**行(レコード)**と**列(フィールド)**で構成される

成績表



成績表の図解。表の列は「生徒番号」「教科」「点数」で、行は「20240001」「20240001」「20240002」である。列と行の方向を示す矢印と、行と列の範囲を示す点線枠が描かれている。

生徒番号	教科	点数
20240001	英語	63
20240001	数学	55
20240002	英語	71

生徒表

生徒番号	氏名	入学年
20230058	水川あさせ	2023
20240001	新はじめ	2024
20240002	山森みずき	2024

関係データベースは、行と列からなる表を複数もっている

- **主キー**：テーブル内で行を特定する項目。テーブル内で主キーのデータは重複しないようにする。複数の項目で主キーとするものを**複合キー(連結キー)**という
- **外部キー**：他の表と関係づけるための項目

成績表

主キー(複合キー)		
生徒番号	教科	点数
20240001	英語	63
20240001	数学	55
20240002	英語	71

外部キー

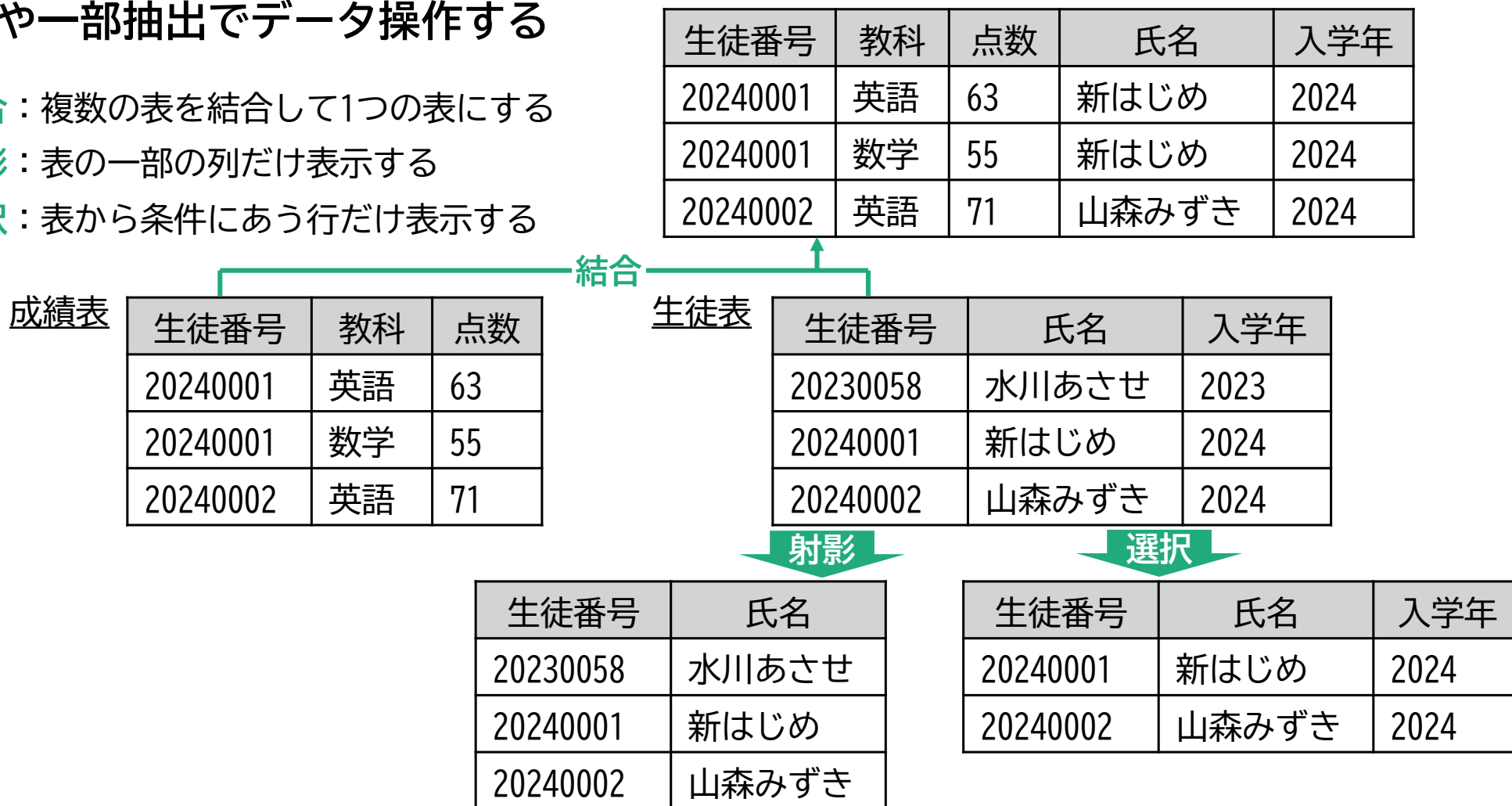
生徒表

主キー		
生徒番号	氏名	入学年
20230058	水川あさせ	2023
20240001	新はじめ	2024
20240002	山森みずき	2024

詳細をもつテーブル

結合や一部抽出でデータ操作する

- **結合**：複数の表を結合して1つの表にする
- **射影**：表の一部の列だけ表示する
- **選択**：表から条件にあう行だけ表示する



「データベース」の要点

「データベース」の要点1

- **ビッグデータ**：人やモノ、社会、自然などから収集される大量のデータのこと
- **オープンデータ**：行政機関などから提供される無償の公開データで、コンピュータが扱いやすく、二次利用可能なルールが設定されたもの
- **テキストマイニング**：文章を単語や文節で区切り、それらの出現傾向や他との関係性などを分析する手法
- **構造化データ**：集計や分析がしやすいようにデータ間の関係性を使って構造化されたデータ
例) 表にまとめられたデータ
- **非構造化データ**：文書や画像、音声、動画などの構造化されていないデータ
例) SNSでのやり取り、写真
※情報技術の発達・普及により、非構造化データが急激に増加している

「データベース」の要点2

- **データベース**：データを整理して、検索や集計などの活用ができるようにしたもの
- **データベース管理システム=DBMS**：データベースを効率よく安全に管理するためのシステム
- **関係データベース=リレーショナルデータベース**：複数の表を組み合わせてデータを管理するDBMS
- **表(テーブル)**：データを格納する表。**行(レコード)**と**列(フィールド)**で構成される
- **主キー**：テーブル内で行を特定する項目。テーブル内で主キーのデータは重複しないようにする。複数の項目で主キーとするものを**複合キー(連結キー)**という
- **外部キー**：他の表と関係づけるための項目

成績表

	主キー(複合キー)		列
	生徒番号	教科	点数
行	20240001	英語	63
	20240001	数学	55

外部キー

生徒表

主キー

生徒番号	氏名	入学年
20240001	新はじめ	2024
20240002	山森みずき	2024

詳細をもつテーブル