

医学歯学分野におけるデータサイエンス教育の本学のこれまでの取り組み

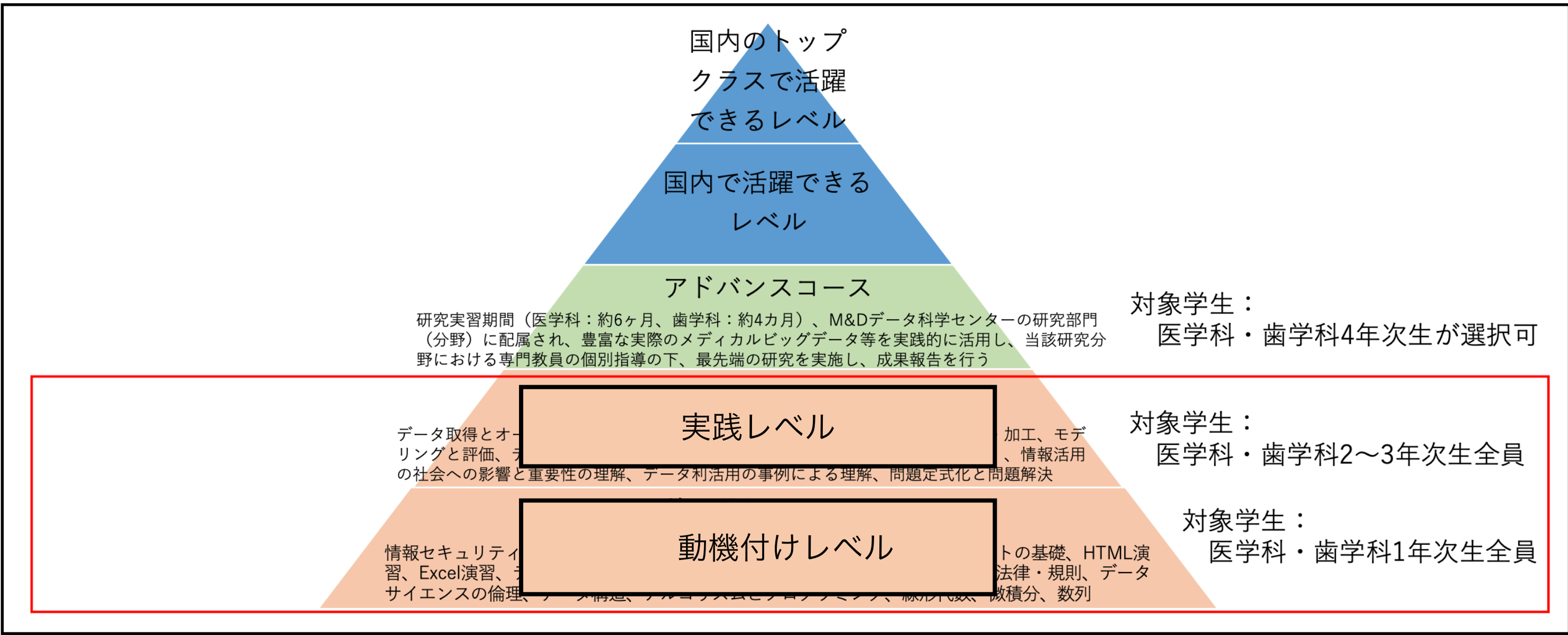
東京医科歯科大学
統合教育機構
須藤毅顕

本学における事業

事業名：医学・歯学分野における数理・データサイエンス・AI教育の開発

将来、

- ・保健医療分野におけるAI技術研究を自ら進めることができる医学・歯学研究者
- ・企業などの技術者と共同で保健医療分野におけるAI技術を推進する能力を持った医師、歯科医師になるための**素養を身に付ける**ことができるモデルカリキュラムを開発する。



動機付けレベル

：データサイエンスの重要性と理解、情報**セキュリティ**、情報**倫理**、**PC・インターネットの基礎**、HTML・Excelの基礎、データサイエンスの倫理、線形代数・微積分・数列、**プログラミング**の導入、など

実践レベル

：**データ取得・前処理・加工**、データ構造、**モデリングと評価**、**統計**の基礎、アルゴリズムと**プログラミング**、問題定式化と問題解決、など

アドバンスコース

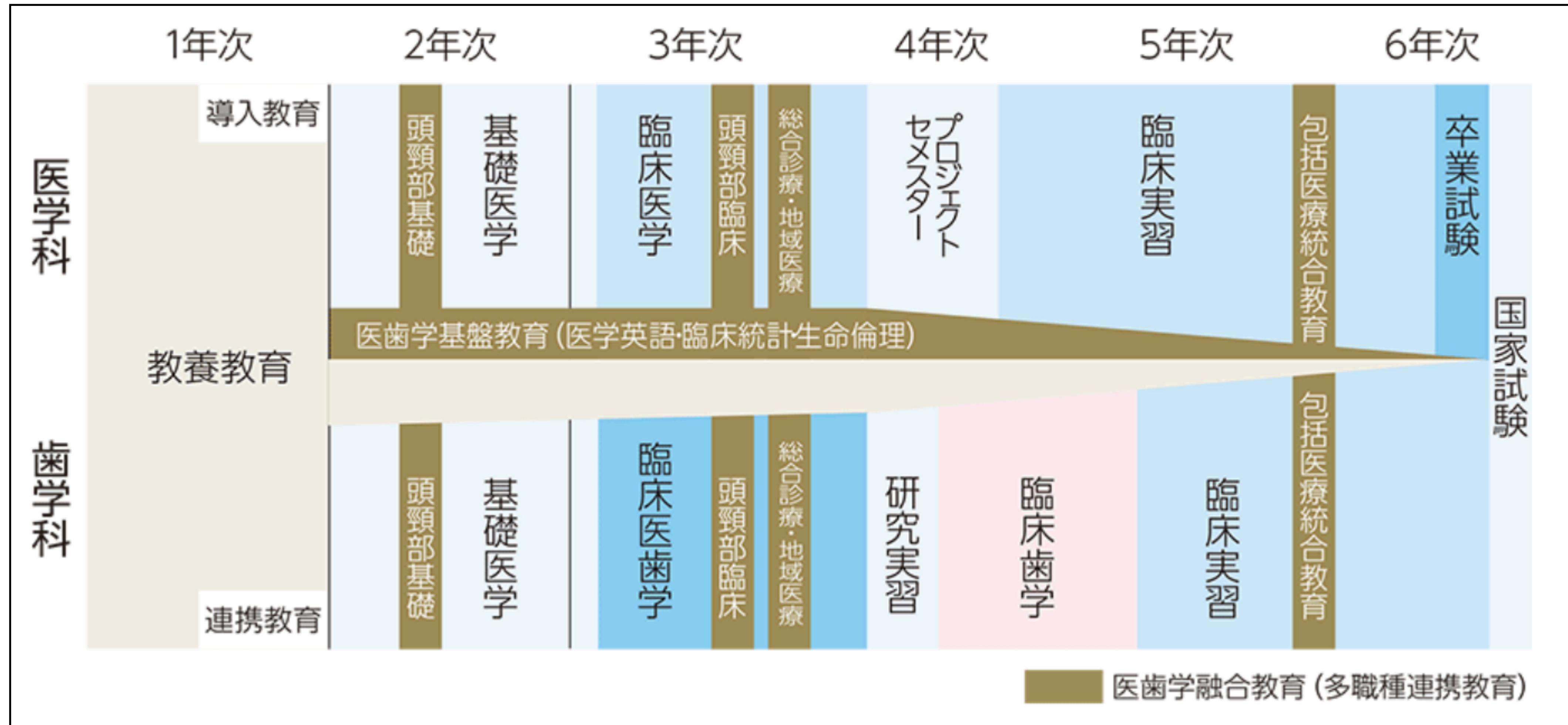
：専門分野での医療データの実践的学習

データサイエンス・AIの理解とプログラミングの導入

医学部・歯学部の特異性

・濃密なカリキュラム

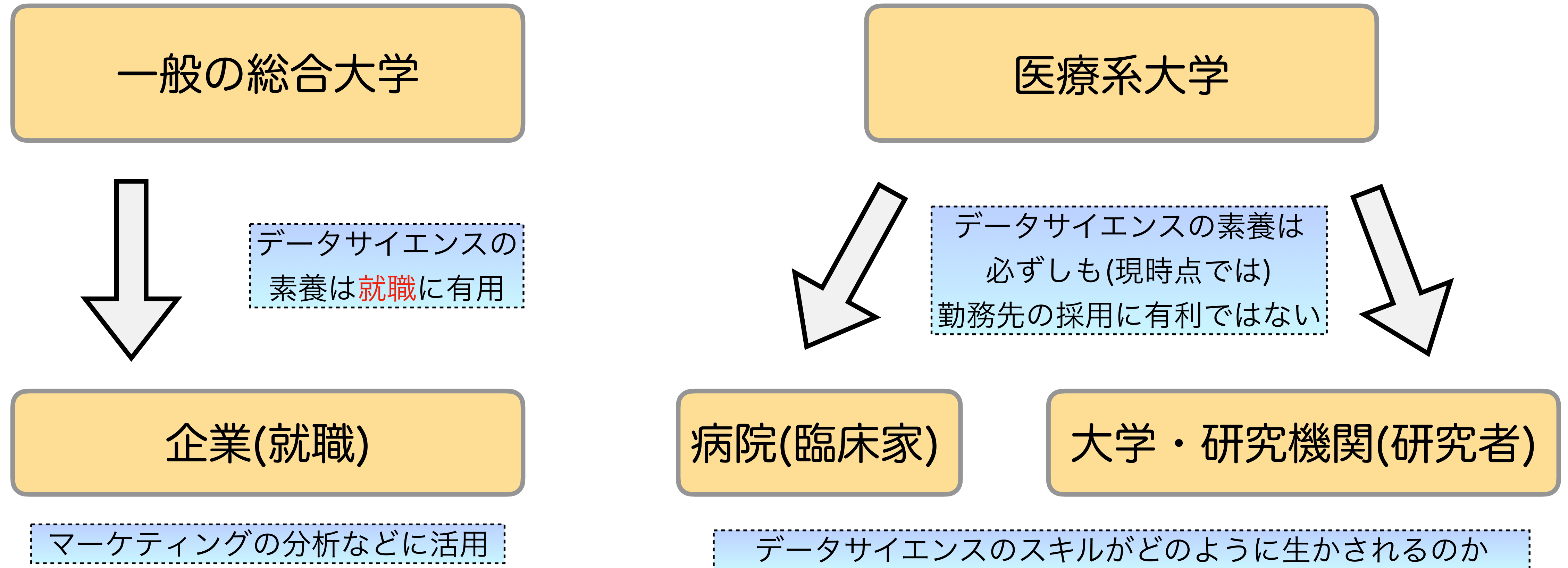
(本学ホームページより)



教養部 (1年)、基礎医学→臨床医歯学→臨床実習→国家試験

医学部・歯学部の特特殊性

- ・ 会社への就職が少なく、多くが臨床家もしくは研究者(大多数は臨床家)



データサイエンスの必要性、ゴール設定の再検討

時代の潮流

医学

プレシジョンメディスン

遺伝子検査
オーダーメイド医療

デジタル医療

オンライン診療
ロボティクス(診療補助、遠隔操作)
VR、AR
CAdE、CAdx

放射線分野 病理分野

AIによる画像診断

- ・ X線写真
肺がん、乳がん、虫歯など
- ・ CT、MRI
肝がん、肺がん、認知症、
冠動脈疾患、肺結節など

歯学

デジタルデンティストリー

3Dプリンター
口腔内スキャナー
CAD/CAM

VR、AR
CAdE、CAdx

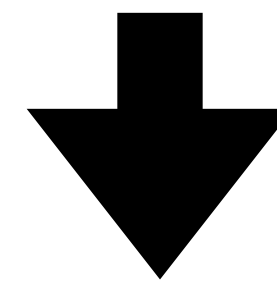
コンピューターの性能の向上に伴い、医療もデータサイエンスの応用へシフトしている

本事業の目的

将来、

- ・保健医療分野におけるAI技術研究を自ら進めることができる医学・歯学研究者
- ・企業などの技術者と共同で保健医療分野におけるAI技術を推進する能力を持った医師、歯科医師になるための**素養を身に付ける**ことができるモデルカリキュラムを開発する。

= 医師・歯科医師(その他医療職(コメディカル))
とAIのダブルメジャーの育成を目指す



既存の科目(教養教育での数学/物理学/統計学および情報処理、
学部教育以降での統計分析/疫学教育)を最大限に生かしつつ、
プログラミングを利用したデータ解析が出来る様な知識と基礎的技術を身につける
(プログラミング言語は人気、有用性からPython、Rを使用予定)

学生に何を求めるか

全体の底上げ

- ・ AIで何ができるのかを皆がわかっている
- ・ プログラミングに慣れ親しむ(抵抗感をなくす)
- ・ 卒後にどのような医療系データサイエンティストの道があるかを知る

一部のデータサイエンスに興味のある学生の萌芽

- ・ 興味を抱いた学生がさらに極められるような道筋を提示する
- ・ 大学院、医局入局後にプログラミングの素養が既に身についていることで研究内容の充実化を図る

目指す学生の将来像

卒業後の進路

```
graph TD; A[卒業後の進路] --> B[臨床家]; A --> C[研究者・教育者];
```

臨床家

- ・ デジタル機器、IoTなどを適切に活用し、医療に応用出来る
- ・ 臨床情報を集約し、ビッグデータとして自ら解析しエビデンスと比較、エビデンスを構築出来る
- ・ 研究機関と対等に議論しながらエビデンスを構築出来る

研究者・教育者

(医療政策、感染症対策などの医系技官も含む)

- ・ オーダーメイド医療に必要なデータ解析を自ら実施出来る
- ・ 大量の臨床情報やゲノムデータなどを適切に扱える
- ・ 基礎研究にてデータサイエンティストと対等に議論しながら共同研究が実施出来る

本事業のロードマップ

卒前教育

数学

微積分
数列
線形代数

統計学

確率
推定・検定
回帰分析

情報処理

CUI・GUI
インターネット
セキュリティ

生化学

実験手技
生命現象の理解
実験データの分析

専門教育

医学

医学科、検査、看護

歯学

歯学科、口腔保健、口腔工学

卒後教育、実践

ゲノム解析

スパコン
データ加工
教師なし学習

→

エクソーム解析
メタゲノム解析
シングルセル解析

AI利活用

教師あり学習
深層学習
教師なし学習

大規模臨床研究

コホート研究
ケースコントロール研究
メタ解析

デジタル医療

VR、AR
ロボティクス(診療補助、遠隔操作)
CAdE、CAdx

本事業のロードマップ

卒前教育

数学

統計学

情報処理

生化学

プログラミングでのデータ処理に数学や統計の知識が必須 → 数学や統計処理にプログラミングを導入・改変

将来どのように生かされるかが学生に伝わるべき → 動機付けの為の新設科目設置(2021年度)

プログラミング

データ加工
データベース
自動化、作業効率化
機械学習

専門教育

医学

医学科、検査、看護

歯学

歯学科、口腔保健、口腔工学

卒後教育、実践

ゲノム解析

スパコン
データ加工
教師なし学習

→

エクソーム解析
メタゲノム解析
シングルセル解析

AI利活用

教師あり学習
深層学習
教師なし学習

大規模臨床研究

コホート研究
ケースコントロール研究
メタ解析

デジタル医療

VR、AR
ロボティクス(診療補助、遠隔操作)
CAdE、CAdx

データサイエンス教育のシームレスな連携・移行を目指す

実装に向けての取り組み

- **既存のデータサイエンス教育の整理**

数理・データサイエンス教育強化拠点コンソーシアムによるリテラシーレベルとの照らし合わせ
→ 必要なスキルセットの整備

- **既存のデータサイエンス教育の改変**

プログラミングの追加導入
→ 必要な整備（プラットフォームの統一など）

- **新設科目「医療とAI・ビッグデータ入門」の整備**

後述（演題3）

数理・データサイエンス教育強化拠点コンソーシアムによるリテラシーレベルのモデルカリキュラム

各大学はこの学習内容およびスキルセットを
参考にカリキュラムを作成している

1.社会におけるデータ・AI活用	学修内容	スキルセット(キーワード)
1-1. 社会で起きている変化	社会で起きている変化を知り、数理・データサイエンス・AIを学ぶことの意義を理解する AIを活用した新しいビジネス/サービスを知る	・ビッグデータ、IoT、AI、ロボット ・データ量の増加、計算機の処理性能の向上、AIの非連続的变化 ・第4次産業革命、Society5.0、データ駆動型社会 ・人間の知的活動とAIの関係性、複数技術を組み合わせたAIサービス ・データを起点としたものの見方、人間の知的活動を起点としたものの見方
1-2. 社会で活用されているデータ	どんなデータが集められ、どう活用されているかを知る	・調査データ、実験データ、人の行動ログデータ、機械の稼働ログデータなど ・1次データ、2次データ、データのメタ化 ・構造化データ、非構造化データ（文章、画像/動画、音声/音楽など）
1-3. データ・AIの活用領域	さまざまな領域でデータ・AIが活用されていることを知る	・データ・AI活用領域の広がり（生産、消費、文化活動など） ・研究開発、調達、製造、物流、販売、マーケティング、サービスなど ・仮説検証、知識発見、原因究明、計画策定、判断支援、活動代替、新規生成など
1-4. データ・AI活用のための技術	データ・AIを活用するために使われている技術の概要を知る	・データ解析：予測、グルーピング、パターン発見、最適化、シミュレーションなど ・データ可視化：複合グラフ、2軸グラフ、多次元の可視化、関係性の可視化、地図上の可視化、挙動・軌跡の可視化、リアルタイム可視化など ・非構造化データ処理：言語処理、画像/動画処理、音声/音楽処理など ・特化型AIと汎用AI、今のAIで出来ることと出来ないこと、AIとビッグデータ ・認識技術、ルールベース、自動化技術
1-5. データ・AI活用の現場	データ・AIを活用することによって、どんな価値が生まれているかを知る	・データサイエンスのサイクル（課題抽出と定式化、データの取得・管理・加工、探索的データ解析、データ解析と推論、結果の共有・伝達、課題解決に向けた提案） ・流通、製造、金融、サービス、インフラ、公共、ヘルスケア等におけるデータ・AI活用事例紹介
1-6. データ・AI活用の最新動向	データ・AI活用における最新動向 (ビジネスモデル、テクノロジーを知る)	・AI等を活用した新しいビジネスモデル（シェアリングエコノミー、商品のレコメンデーションなど） ・AI最新技術の活用例（深層生成モデル、敵対的生成ネットワーク、強化学習、転移学習など）

2.データリテラシー	学修内容	スキルセット(キーワード)
2-1. データを読む	データを適切に読み解く力を養う	・データの種類(量的変数、質的変数) ・データの分布(ヒストグラム)と代表値(平均値、中央値、最頻値) ・代表値の性質の違い(実社会では平均値=最頻値でないことが多い) ・データのばらつき(分散、標準偏差、偏差値) ・観測データに含まれる誤差の扱い ・打ち切りや脱落を含むデータ、層別の必要なデータ ・相関と因果(相関係数、疑似相関、交絡) ・母集団と標本抽出(国勢調査、アンケート調査、全数調査、単純無作為抽出、層別抽出、多段抽出) ・クロス集計表、分割表、相関係数行列、散布図行列 ・統計情報の正しい理解(誇張表現に惑わされない)
2-2. データを説明する	データを適切に説明する力を養う	・データ表現(棒グラフ、折線グラフ、散布図、ヒートマップ) ・データの図表表現(チャート化) ・データの比較(条件をそろえた比較、処理の前後での比較、A/Bテスト) ・不適切なグラフ表現(チャートジャンク、不必要な視覚的要素) ・優れた可視化事例の紹介(可視化することによって新たな気づきがあった事例など)
2-3. データを扱う	データを扱うための力を養う	・データの集計(和、平均) ・データの並び替え、ランキング ・データ解析ツール(スプレッドシート) ・表形式のデータ(csv)

3.データ・AI活用における留意事項	学修内容	スキルセット(キーワード)
3-1. データ・AIを扱う上での留意事項	データ・AIを利活用する上で知っておくべきこと	・ELSI(Ethical, Legal and Social Issues) ・個人情報保護、EU一般データ保護規則(GDPR)、忘れられる権利、オプトアウト ・データ倫理:データのねつ造、改ざん、盗用、プライバシー保護 ・AI社会原則(公平性、説明責任、透明性、人間中心の判断) ・データバイアス、アルゴリズムバイアス ・AIサービスの責任論　・データ・AI活用における負の事例紹介
3-2. データを守る上での留意事項	データを守る上で知っておくべきこと	・情報セキュリティ:機密性、完全性、可用性 ・匿名加工情報、暗号化、パスワード、悪意ある情報搾取 ・情報漏洩等によるセキュリティ事故の事例紹介

4. オプション	学修内容	スキルセット(キーワード)
4-1. 統計および数理基礎	数学基礎および統計基礎を学ぶ	・確率、順列、組み合わせ ・線形代数(ベクトル、ベクトルの基本的な演算、ノルム、行列とベクトルの積、行列の積、内積) ・1変数関数の微分と積分 ・指数関数、対数関数 ・集合、ベン図
4-2. アルゴリズム基礎	アルゴリズム基礎を学ぶ	・アルゴリズムの表現（フローチャート、アクティビティ図） ・並び替え（ソート） ・探索（サーチ）
4-3. データ構造とプログラミング基礎	データ構造とプログラミング基礎を学ぶ	・数と表現、計算誤差、データ量の単位、文字コード、配列 ・変数、代入、繰り返し、場合に応じた処理
4-4. 時系列データ解析	時系列データ解析の概要を知る	・時系列データ（トレンド、周期、ノイズ） ・季節調整、移動平均
4-5. テキスト解析	自然言語処理の概要を知る	・形態素解析、単語分割、ユーザ定義辞書、n-gram、文章間類似度 ・かな漢字変換の概要
4-6. データハンドリング	大規模データをハンドリングする力を養う	・データベース(リレーショナルデータベース、SQL) ・プログラミング(Python) ・データクレンジング:外れ値、異常値、欠損値の処理 ・データの抽出 ・データの結合　・名寄せ
4-7. データ活用実践（教師あり学習）	データ利活用プロセス（教師あり学習）を体験し、データを使って考える力を養う	・教師あり学習による予測 例)売上予測、罹患予測、成約予測、離反予測など ・データの収集(分析に必要なデータの確認、対象となるデータの収集) ・データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) ・データの分析(単回帰分析、重回帰分析、ロジスティック回帰分析、モデルの評価) ・データ分析結果の共有、課題解決に向けた提案
4-8. データ活用実践（教師なし学習）	データ利活用プロセス（教師なし学習）を体験し、データを使って考える力を養う	・教師なし学習によるグルーピング 例)顧客セグメンテーション、店舗クラスターリング ・データの収集(分析に必要なデータの確認、対象となるデータの収集) ・データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) ・データの分析(階層クラスターリング 非階層クラスターリング) ・データ分析結果の共有、課題解決に向けた提案

既存のデータサイエンス関連科目の整備

教養部(1年)

前期：

医学科、月曜 3 限目、**数学**1M(中口) 医学科、木曜 1 限目、**情報処理**(α)(徳永/前澤/永並)

歯学科、水曜 1 限目、**数学**1D(中口) 歯学科、木曜 2 限目、**情報処理**(β)(徳永/前澤/永並)

医学科歯学科、火曜3,4限目(共通領域セミナー科目)、**R・講義と実習によるバイオサイエンス1**(服部)、**R・離散構造**(徳永)、**S・数学基礎**(中口)

歯学科、後期 2 限 ✕ 4 日、**基礎情報医歯学**1-4(木下、青木、田畑、須永)

後期：

医学科歯学科、水曜 4 限目、**数学 II** (α (中林)、 β (徳永)、 γ (中口))

医学科歯学科、金曜 3 限目、**S・データサイエンス入門**(中口)、**R・講義と実習によるバイオサイエンス2**(服部)、**S・プログラミングで学ぶ幾何学デザイン**(徳永)

医学科、火曜終日、**医学導入**(データサイエンスの講義あり)

医学科歯学科2年 木曜 1、2 限目、教養/医歯学基盤教育(1~15、1~16、予備 2 日間) -**臨床統計1**(能登、徳永) -

医学科歯学科3年 木曜 1、2 限目、教養/医歯学基盤教育(1~15、1~15、予備 2 日間) -**臨床統計2**(能登) -

医学科歯学科4年 木曜 1、2 限目、教養/医歯学基盤教育(1~6) -**臨床統計3**(能登) -

歯学科2年 木曜 2、3 限目 8 コマ、**感染と生体防御**(バイオインフォマティクス講義)

木曜 3 コマ、**情報科学演習**(中口、水口、金澤)

医学科3年 前期、**衛生学実習**15コマ、**公衆衛生学**12コマ

基本は数学、情報処理、統計学がメイン（他の医療系大学も共通）

リテラシーレベルとの照らし合わせによる検討

- ・ 恐らく類似科目が既に多くの大学でも存在しているので応用が容易
- ・ コマ数の制限があるので、既存の授業内容を改変する工夫が必要

数学、統計

2 .データリテラシー	学修内容	スキルセット(キーワード)
2-1. データを読む	データを適切に読み解く力を養う	・ データの種類(量的変数、質的変数) ・ データの分布(ヒストグラム)と代表値(平均値、中央値、最頻値) ・ 代表値の性質の違い(実社会では平均値=最頻値でないことが多い) ・ データのばらつき(分散、標準偏差、偏差値) ・ 観測データに含まれる誤差の扱い ・ 打ち切りや脱落を含むデータ、層別の必要なデータ ・ 相関と因果(相関係数、疑似相関、交絡) ・ 母集団と標本抽出(国勢調査、アンケート調査、全数調査、単純無作為抽出、層別抽出、多段抽出) ・ クロス集計表、分割表、相関係数行列、散布図行列 ・ 統計情報の正しい理解(誇張表現に惑わされない)
2-2. データを説明する	データを適切に説明する力を養う	・ データ表現(棒グラフ、折線グラフ、散布図、ヒートマップ) ・ データの図表表現(チャート化) ・ データの比較(条件をそろえた比較、処理の前後での比較、A/Bテスト) ・ 不適切なグラフ表現(チャートジャンク、不必要な視覚的要素) ・ 優れた可視化事例の紹介(可視化することによって新たな気づきがあった事例など)
2-3. データを扱う	データを扱うための力を養う	・ データの集計(和、平均) ・ データの並び替え、ランキング ・ データ解析ツール(スプレッドシート) ・ 表形式のデータ(csv)

情報処理

3.データ・AI活用における留意事項	学修内容	スキルセット(キーワード)
3-1. データ・AIを扱う上での留意事項	データ・AIを利活用する上で知っておくべきこと	・ ELSI(Ethical, Legal and Social Issues) ・ 個人情報保護、EU一般データ保護規則(GDPR)、忘れられる権利、オプトアウト ・ データ倫理:データのねつ造、改ざん、盗用、プライバシー保護 ・ AI社会原則(公平性、説明責任、透明性、人間中心の判断) ・ データバイアス、アルゴリズムバイアス ・ AIサービスの責任論 ・ データ・AI活用における負の事例紹介
3-2. データを守る上での留意事項	データを守る上で知っておくべきこと	・ 情報セキュリティ:機密性、完全性、可用性 ・ 匿名加工情報、暗号化、パスワード、悪意ある情報搾取 ・ 情報漏洩等によるセキュリティ事故の事例紹介

数学、統計

4. オプション	学修内容	スキルセット(キーワード)
4-1. 統計および数理基礎	数学基礎および統計基礎を学ぶ	・ 確率、順列、組み合わせ ・ 線形代数(ベクトル、ベクトルの基本的な演算、ノルム、行列とベクトルの積、行列の積、内積) ・ 1変数関数の微分と積分 ・ 指数関数、対数関数 ・ 集合、ベン図
4-2. アルゴリズム基礎	アルゴリズム基礎を学ぶ	・ アルゴリズムの表現（フローチャート、アクティビティ図） ・ 並び替え（ソート） ・ 探索（サーチ）
4-3. データ構造とプログラミング基礎	データ構造とプログラミング基礎を学ぶ	・ 数と表現、計算誤差、データ量の単位、文字コード、配列 ・ 変数、代入、繰り返し、場合に応じた処理
4-4. 時系列データ解析	時系列データ解析の概要を知る	・ 時系列データ（トレンド、周期、ノイズ） ・ 季節調整、移動平均

4. オプション	学修内容	スキルセット(キーワード)
4-5. テキスト解析	自然言語処理の概要を知る	・ 形態素解析、単語分割、ユーザ定義辞書、n-gram、文章間類似度 ・ かな漢字変換の概要
4-6. データハンドリング	大規模データをハンドリングする力を養う	・ データベース(リレーショナルデータベース、SQL) ・ プログラミング(Python) ・ データクレンジング:外れ値、異常値、欠損値の処理 ・ データの抽出 ・ データの結合 ・ 名寄せ
4-7. データ活用実践（教師あり学習）	データ活用プロセス（教師あり学習）を体験し、データを使って考える力を養う	・ 教師あり学習による予測 例)売上予測、罹患予測、成約予測、離反予測など ・ データの収集(分析に必要なデータの確認、対象となるデータの収集) ・ データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) ・ データの分析(単回帰分析、重回帰分析、ロジスティック回帰分析、モデルの評価) ・ データ分析結果の共有、課題解決に向けた提案
4-8. データ活用実践（教師なし学習）	データ活用プロセス（教師なし学習）を体験し、データを使って考える力を養う	・ 教師なし学習によるグルーピング 例)顧客セグメンテーション、店舗クラスターリング ・ データの収集(分析に必要なデータの確認、対象となるデータの収集) ・ データの加工(データクレンジング、サンプリング、簡単な説明変数の作成) ・ データの分析(階層クラスターリング 非階層クラスターリング) ・ データ分析結果の共有、課題解決に向けた提案

課題はプログラミングの導入と動機付けの工夫

既存の数理・データサイエンス科目の改変

教養教育

数学

微積分
数列
線形代数

統計学

確率
推定・検定
回帰分析

情報処理

CUI・GUI
インターネット
セキュリティ

生化学

実験手技
生命現象の理解

プラットフォーム(Spyder)
の統一による効率性の向上

既存のデータサイエンスに
プログラミングの要素を追加

教養の1年生に対する
動機付けを期待

全学生のPCにPython
インストール(Anaconda)

Pythonによる演習
Rによる演習

Rによる演習

Pythonによる演習

どのように医学・歯学のデータサイエンスに結びついていくのか、
実例を通してイメージさせる工夫

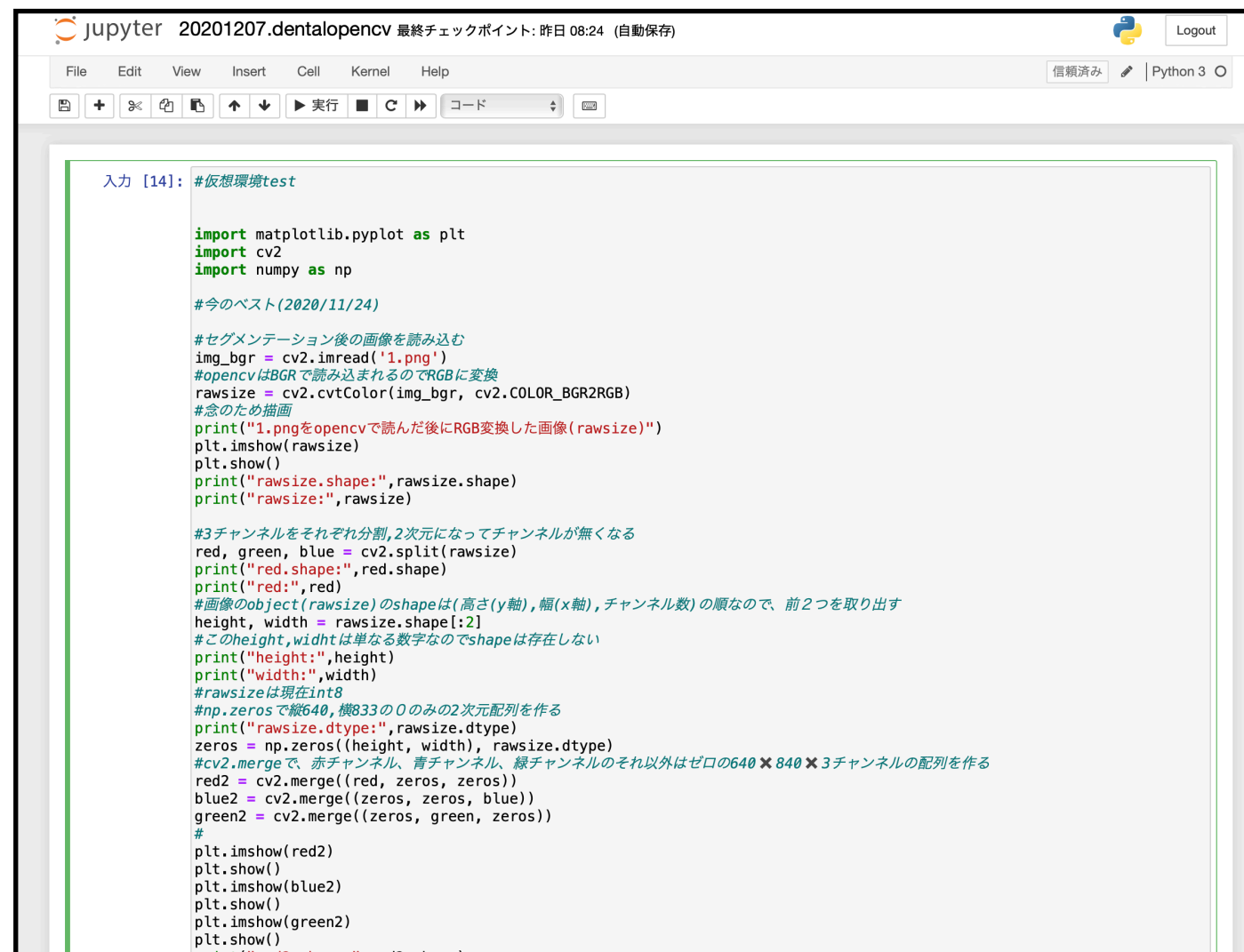
プログラミング

専門教育

多くの実行環境が存在 (python)

プラットフォームの候補

Jupyter notebook



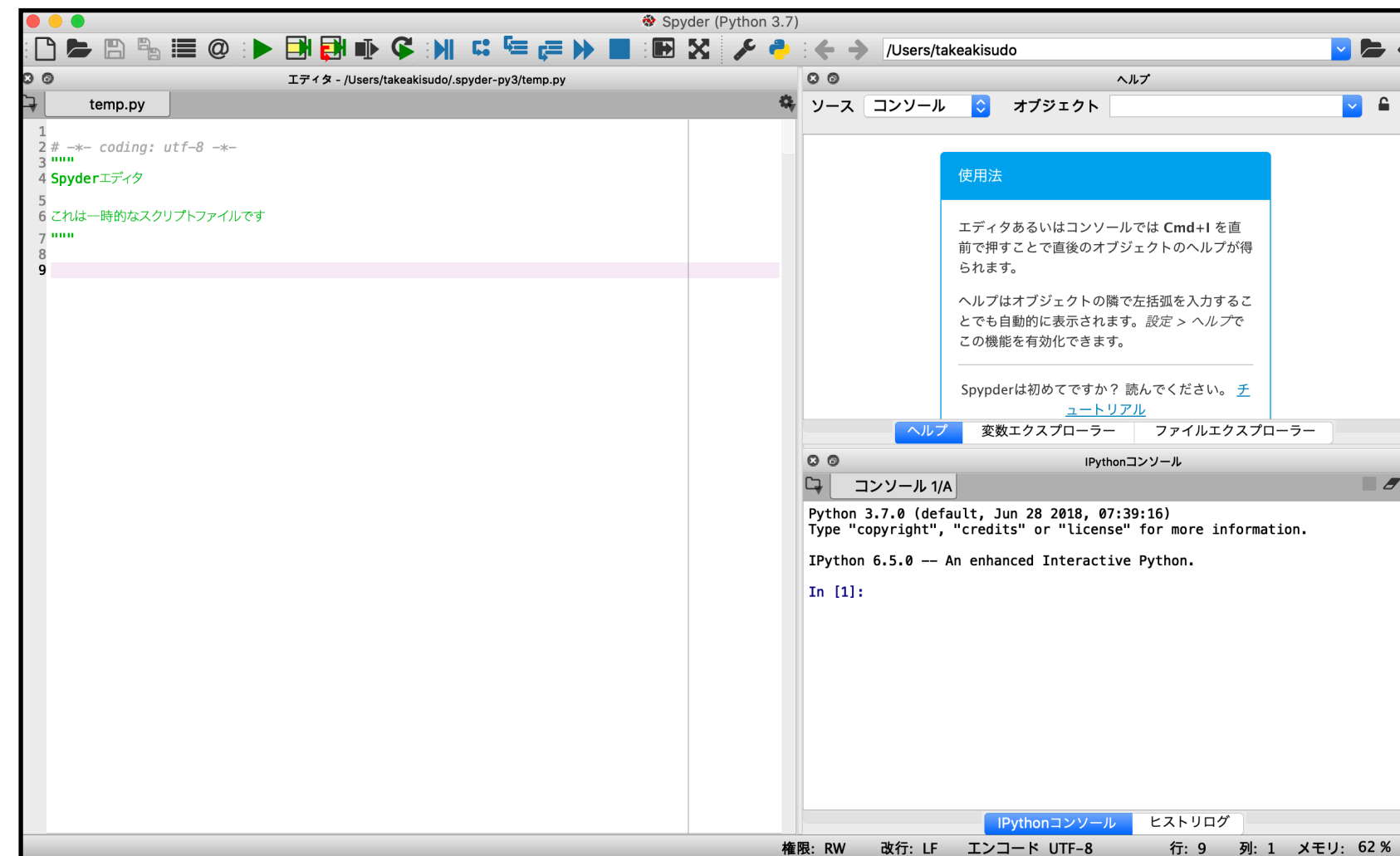
利点：見やすい、使用している参考書が多い

欠点：text editorとしては使えない

拡張子が.ipynbで少し追加の説明が必要か

pythonファイル(.py)は実行できない

Spyder

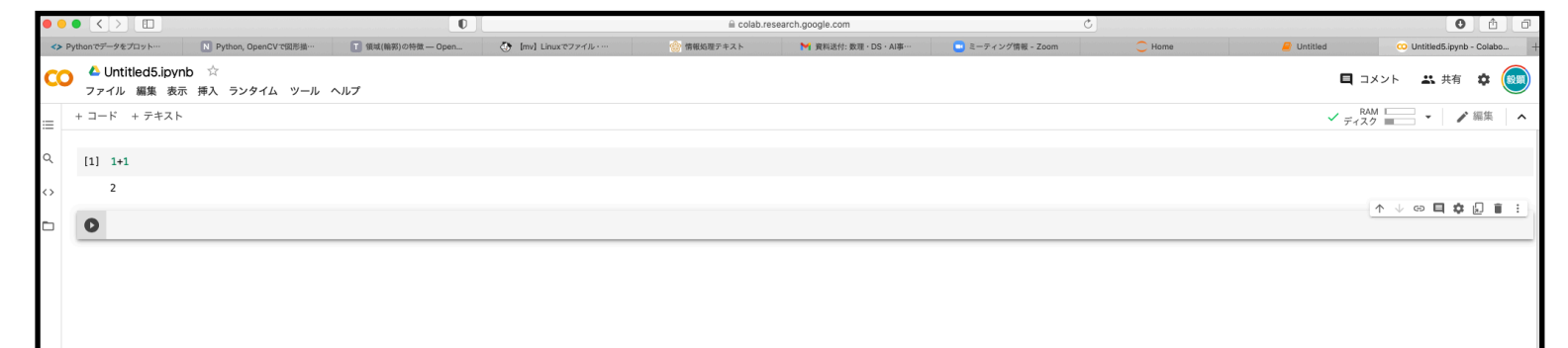


利点：そのまま書くことも、pythonのファイルも実行できる。

図など見やすい、Rstudioに似てる

欠点：デフォルトがpython(text editorとしてはどうか)

Google Colaboratory



利点：見やすい、

Jupyter notebookよりスペックが良い(GPU)

欠点：text editorとしては使えない

拡張子が.ipynbで少し追加の説明が必要か

pythonファイル(.py)は実行できない

googleのアカウントが必要

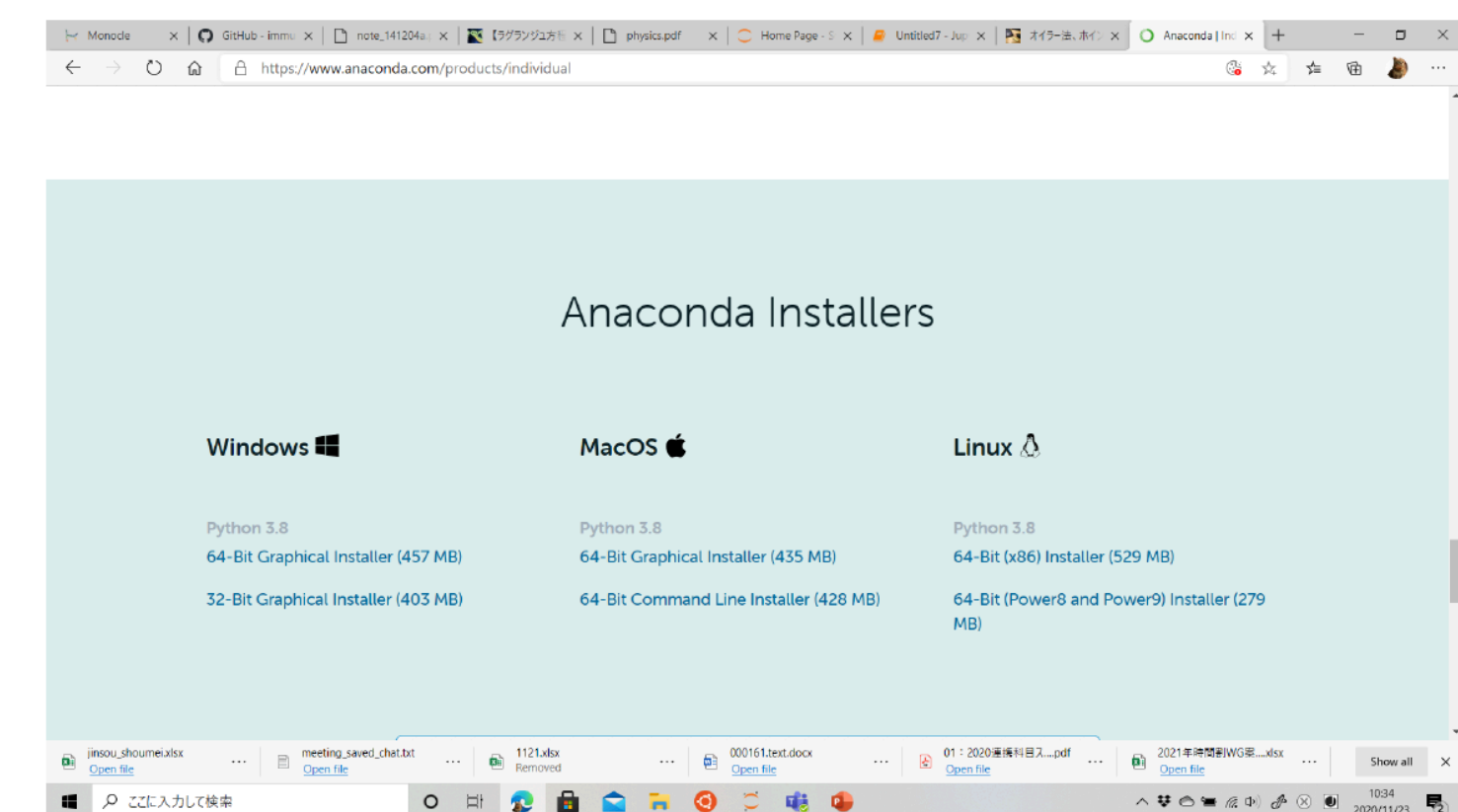
プラットフォームの統一の検討

教養部 1 年時の情報処理の限られた時間で如何にインストールのトラブルを無くすか

情報科学演習

東京医科歯科大学 教養部 数学科
中林 潤
nakab.las@tmd.ac.jp

Anaconda Installerのダウンロード



自分の環境に合ったインストーラーをダウンロードする

それぞれの授業でインストールについての授業に割く時間を削減する

異なる開発環境での授業による学生の混乱を防ぐ

ANACONDA NAVIGATORの利用

教養部 1 年時の情報処理の限られた時間で如何にインストールのトラブルを無くすか

<https://www.anaconda.com>

Anaconda Navigator

JupyterLab

Jupyter Notebook

Qt Console

Spyder

VS Code

Glueviz

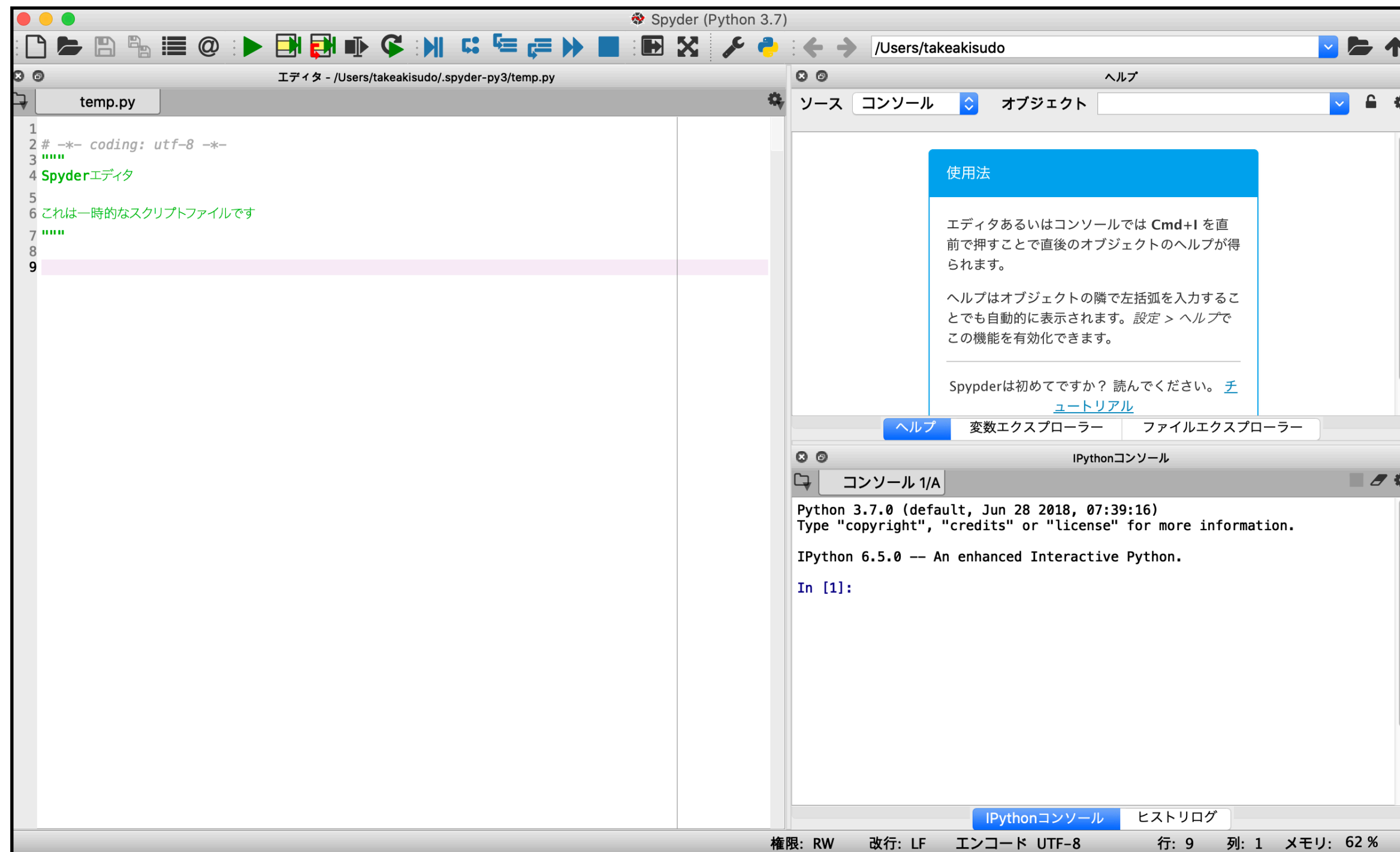
Orange 3

RStudio

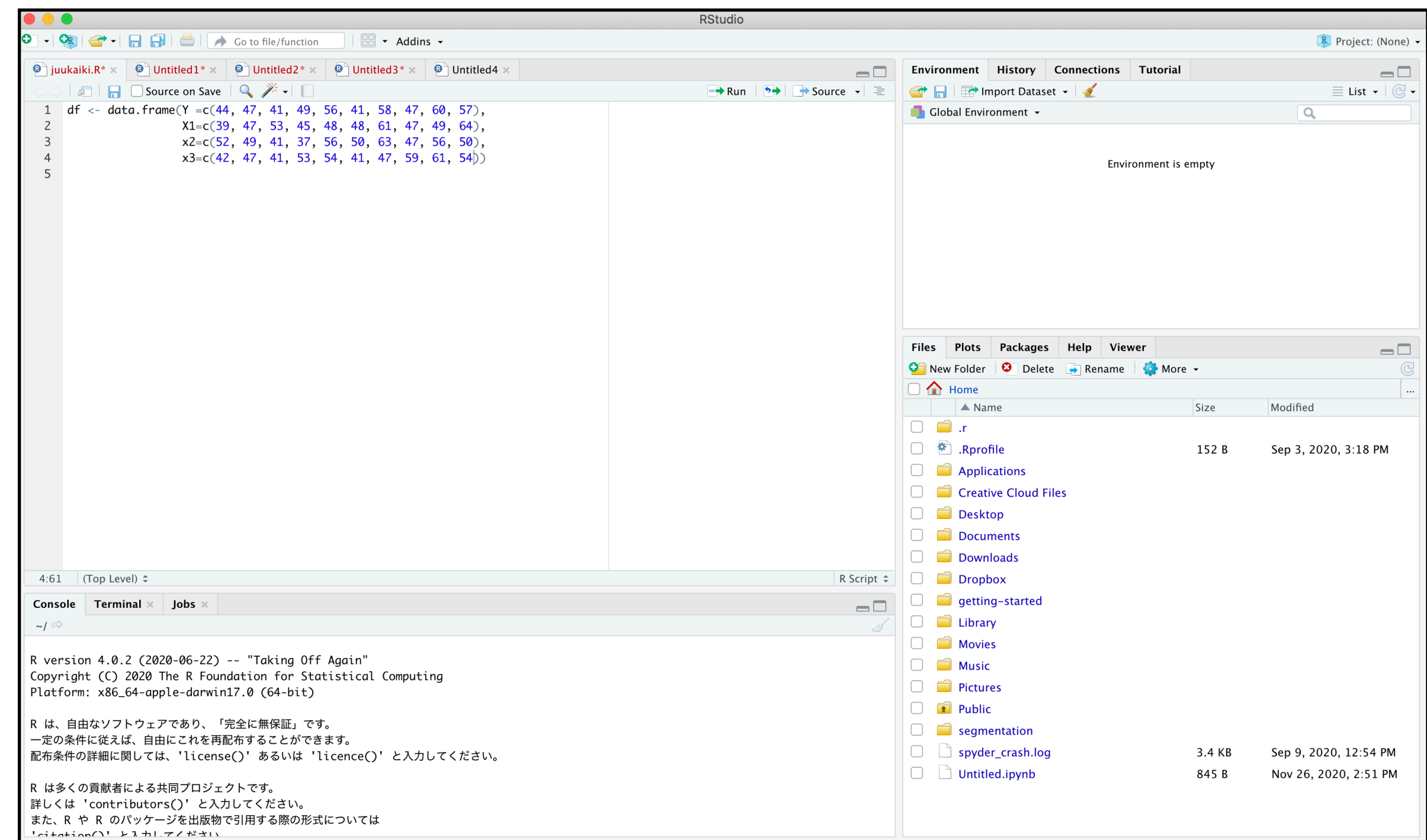
Anaconda(ANACONDA NAVIGATOR)の統一により各授業のプログラミングの導入をスムーズにする

SpyderによるRstudioへのスムーズな移行を期待

Spyder



RStudio



.pyを実行出来る、Rstudioへの移行がスムーズに出来ることを期待してSpyderの方針

既存科目のプログラミングの実践・導入の検討

必修

数学Ⅰ, 数学Ⅱ α 、 β 、 γ

対象：医学科歯学科1年生,約190人(数学Ⅰ)

医学科歯学科1年生,各約20人(数学Ⅱ)

使用言語：R,SciLab(数学Ⅰ)

→pythonの導入を検討中

情報科学

対象：保健衛生学科、口腔保健学科1年生,必修約120人

使用言語：十進BASIC、HTML

→ pythonの(導入)を検討中

臨床統計

対象：医学科歯学科2-4年生,約190人

使用言語：R (EZR)

統計処理をRを用いて実践

情報処理

対象：医学科歯学科1年生,約190人

使用言語：HTML → pythonの(導入)を実施予定

情報科学演習

対象：歯学科2年生,約50人

使用言語：python(Jupyter notebook)

選択

臨床統計実習(中級)

対象：医学科歯学科2-4年生

使用言語：python(デモ)

機械学習(教師あり学習、教師なし学習)

を用いたバイオインフォマティクス全般

S：データサイエンス入門 (2020年度で終了予定)

S：実用統計学

R：離散構造

→ pythonの(導入)を検討中

数学・情報処理でのプログラミングの実践・導入

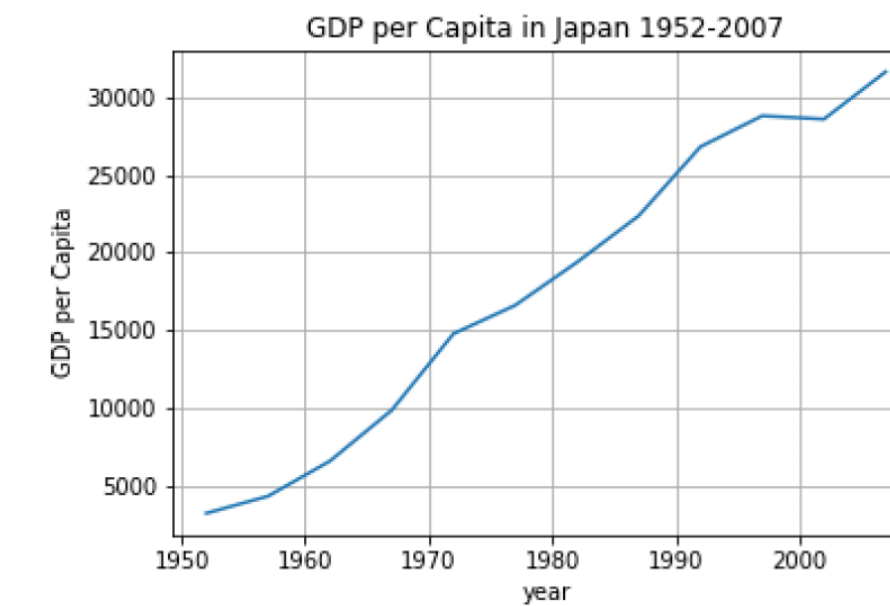
データの加工、統計処理、可視化の実践

(情報科学演習(中林潤先生)より)

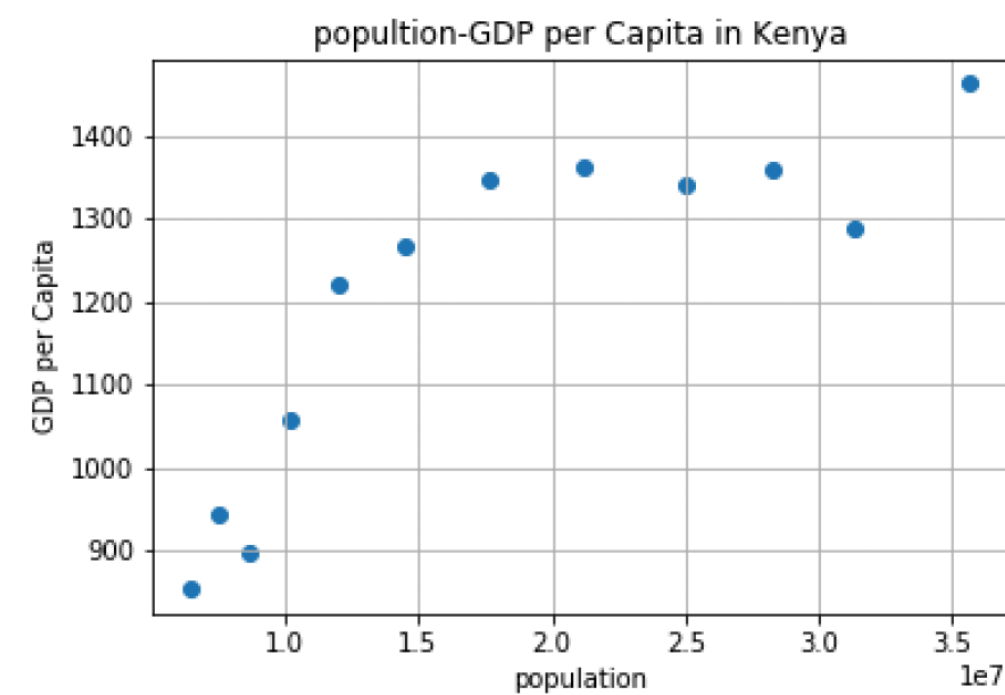
今日の課題

- gdppercapi_table.csvのデータを読み込み
- 日本 (Japan) における1952年～2007年までのGDP per capita (gdpPercap) の折れ線グラフ
- ケニア (Kenya) における人口 (pop) とGDP per capitaの散布図
- アフリカ諸国 (continentがAfrica) における2007年度人口の棒グラフ
- AsiaとEurope諸国の2007年度GDP per capitaの箱ひげ図
- パワーポイントのスライド1枚につき図一つを張り付けて (計4枚) 提出
- plt.savefig('file名.png')
- パワーポイントのファイル名は学籍番号__ローマ字の名前

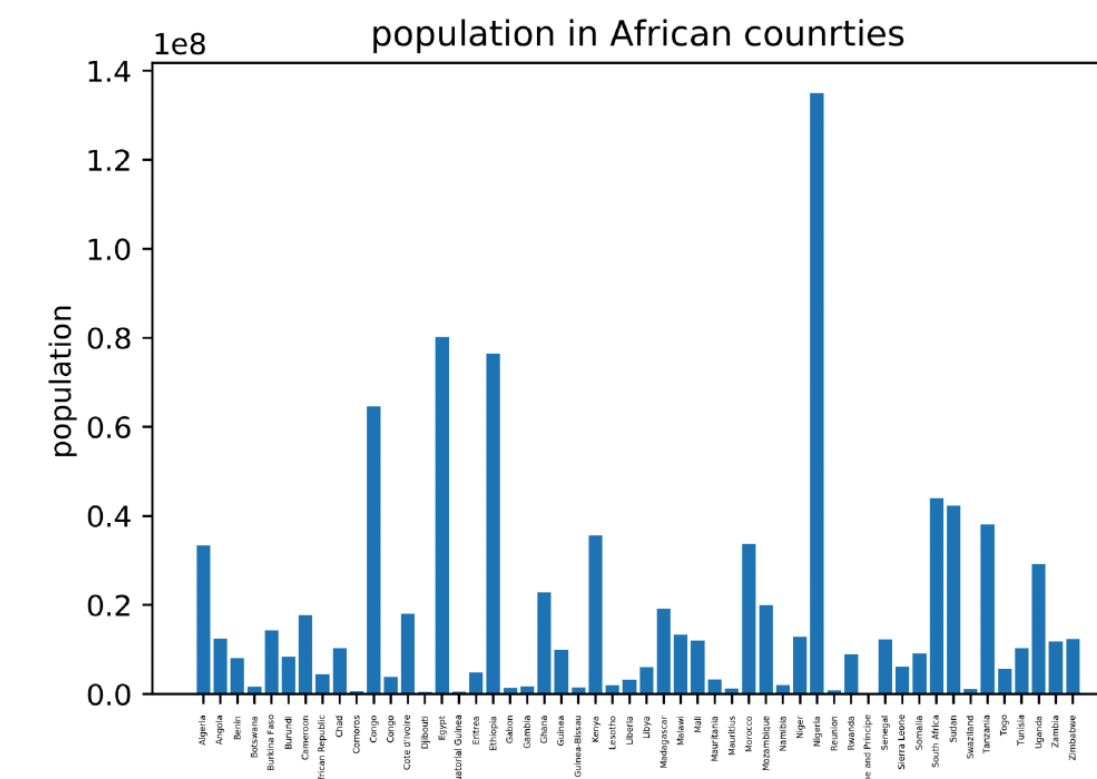
前回の課題について



前回の課題について



前回の課題について



コンピューターリテラシーの理解とプログラミングの実践

(情報処理HP(徳永伸一先生)より)

リンク集	
目次	
1. パソコンに関する基本事項	
1.1 ハードウェアの概要	
1.2 ソフトウェアの概要	
1.3 パソコンの起動と終了	
2. パソコンの基本操作	
2.1 初期画面の観察	
2.2 マウス操作	
2.3 ウィンドウの基本操作	
2.4 複数ウィンドウの操作	
3. インターネットとWWW	
3.1 インターネットの基礎知識	
3.2 WWWブラウザの使用法・活用法	
5. Eメール	
5.1 Eメール（電子メール）の仕組み	
5.2 Webメールの設定と使用法	
5.3 Webメール使用上の注意（オートコンプリート他）	
6. キーボード操作とタイピング	
6.1 キーボード操作の基本	
6.2 タイピング練習	
7. パソコンの内部構造とデータ構造とデ	

7. パソコンの内部構造とデータ構造（未完成）

■あらためて「コンピューター」を開き、パソコンの内部構造を見る。

7.1 ドライブ、フォルダとファイル

- ・ **ドライブ**：記憶装置などのハードウェアに対応。Aドライブ、Cドライブという風にアルファベットを付けて呼ぶ約束になっている。
 - 以下のドライブに対応するアイコンを「マイコンピュータ」のウィンドウ内で確認せよ。
 - フロッピーディスク（FD）ドライブ・・・（あれば）A・Bドライブが割り当てられる。以前はFDが外部記憶メディアの主流だったが、最近ではほとんど用いられなくなった。
 - ハードディスク（HD）ドライブ・・・HDは磁気ディスクの一種で、最も中心的な役割を担う記憶装置。Cドライブは通常HDドライブで、OSはここにインストールされている。実習室のパソコンではCのみHDドライブ（市販のパソコンの大半と同じ）。
- 【注意】：大学のパソコンにおいては、原則として **ファイルはすべてUSBフラッシュメモリ等の外部記憶メディアに保存すること**。作業用一時的なファイルの置き場所が必要なときはCドライブの「ドキュメント」を利用せよ。ただし教養部のパソコンにおいてはCドライブはHDD保護ソフトによって保護されているため、再起動すると保存ファイルが消えてしまうことに注意。
- 光学ディスクドライブ・・・DVD-ROM/R/RAM、CD-ROM/R/RW等の光学ディスクを扱うドライブ。Dドライブが割り当てられることが多い（ハードディスクドライブが複数あればE以降に繰り下がる）。
- ネットワークドライブ・・・ネットワークにより実現される仮想ドライブ（実体は別のパソコンやサーバ内のHDなど）。実習室のパソコンでは現在設定されていない。
- その他・・・USBメモリ、カードリーダ等は接続すると新しいドライブとして表示される
- ・ **フォルダ**：記憶装置内の仮想的な"入れ物"。フォルダの中にフォルダを作りができ、階層構造をなす。 **ディレクトリ**と呼ばれることもある。
 - **エクスプローラー**（Internet Explorerではない！）を起動して全体の構造を確認せよ。エクスプローラーの起動法：
 - 方法1：タスクバー上のアイコンをクリック（アイコンが無ければ不可）
 - 方法2：「スタート」ボタン上で右クリック→メニューから「エクスプローラ」を選択。
- ・ **ファイル**：コンピュータが1度に扱うデータの塊。
 - フォルダ内のファイルをクリックして選択すると、そのファイルの詳細情報がウィンドウ左側に表示される。

(情報科学HP(徳永伸一先生)より)

★右側のフレーム内に表示されているのは「十進BASICのホームページ」内のコンテンツで、著者は白石和夫先生です。

プログラミング実習について

プログラミング言語として、文教大学の白石和夫先生が開発された「十進BASIC」（以下「十進BASIC」と略記）を用い、テキストとして白石先生が作成された「十進BASICによるJIS Full BASIC入門」（右側のフレームに表示）と徳永による補足テキスト（このフレーム内に表示）を併用します。
〔注意〕以下文中の「テキスト」とは右のフレームに表示された「（仮称）十進BASICによるJIS Full BASIC入門」を指すものとする。

授業の進行について

- ・ 十進BASICの実行ファイルは「basic.exe」です。このファイルのアイコンまたはインストール時に作成されたショートカットアイコンをダブルクリックして起動します。
- ・ 別途指示された場合を除き、すべての作業はこの補足テキストの指示に従って進めること。具体的には「十進BASICによるJIS Full BASIC入門」内の**サンプルプログラム（「例＊」など）を順次入力・実行**しながら内容を理解してもらいます。さらに毎回の学習内容に応じた課題が与えられるので、それを終えてプログラムを提出し、合格（OK）の返信を受信したら終了（★授業中に返信できないことが多いですが、返信が届き次第、確認するようにしてください）。
- ・ 右のテキストおよびこの補足テキストを読みながらの独習を基本とします。授業の冒頭で

5 データの入力と出力

5.1 文字列

5.1.1 文字列定数

プログラム中に文字列を記述する場合、その前後を" "でくくります。
" "でくくった文字列を文字列定数といいます。
文字列定数に"を含めるときは、"を重ねて""とします。
例 54

```
10 PRINT "Say ""Hello!""  
20 END
```

5.1.2 文字列変数

文字列を値としてもつ変数を文字列変数といいます。
BASICでは、文字列変数には、末尾に\$(ドル記号)を持つ名前を用いる約束になっています。

5.1.3 文字列連結

文字列には、連結の演算が定義されています。文字列の連結を文字列と文字列の間に&を書いて表します。

例 55

```
10 LET s$="Brown"  
20 LET s$="Mr. " & s$  
30 PRINT s$  
40 END
```

5.1.4 STR\$関数

組込関数STR\$(x)は数値xを文字列化します。数値の前後には空白を含めません。
次のプログラムは、入力された数値を2進法に変換し上位桁が左にくるように表示します。

例 56

```
10 INPUT n  
20 LET s$ = ""  
30 DO UNTIL n=0  
40   LET r = MOD(n,2)  
50   LET s$ = STR$(r) & s$  
60   LET n = (n-r)/2  
70 LOOP  
80 PRINT s$  
90 END
```

文字を1個も含まない文字列を空文字列といいます。空文字列は""で表されます。
上のプログラムは、最初、s\$を空文字列としておき、入力された数値を2で割った余りを求めるたびに、それを文字列化し、s\$の左に連結しています。

数学・情報処理でのプログラミングの実践・導入

数式の可視化の実践

(数学Ⅰ(中口悦史先生)より)

WebClass
FLA20003 数学Ⅰ・Ⅱ 2020 M1D1MT1

教材成績出席その他コースログアウト

お知らせがあります。

タイムライン

C

さらに過去の記録を取得

数学Ⅱ r 2020

数学Ⅱ r 授業資料 (教員用)

詳細

資料

数学Ⅱ・物理学Ⅱ クラス分け

数学Ⅱ・物理学Ⅱ クラス分け (検査含む)

詳細

利用回数 1

資料

数学Ⅰ 期末試験

数学Ⅰ 期末試験 リハーサル 試験問題

詳細

資料

利用可能期間 2020/09/09 09:20 - 2020/09/09 12:20

数学Ⅰ 期末試験 リハーサル 答案提出

詳細

試験

利用可能期間 2020/09/09 09:00 - 2020/09/09 11:59

数学Ⅰ 期末試験 リハーサル 解答例

詳細

資料

利用可能期間 2020/09/09 12:00 - 2020/10/31 23:59

数学Ⅰ 第14回

第14回講義ノート (9/07 MT講義後修正)

詳細

資料

利用回数 1

数学Ⅰ 第13回

第13回講義ノート (8/31 M講義後修正)

詳細

資料

利用回数 2

第13回演習+解答例

詳細

資料

数学Ⅰ 附録

数学Ⅰ 高校数学からの補足

詳細

資料

数学ソフトウェア紹介

ソフトウェア「R」のインストール方法等

詳細

資料

利用回数 1

ソフトウェア「SciLab」のインストール方法等

詳細

資料

利用回数 1

数学Ⅰ 第12回

第12回講義ノート (8/31 MT講義後再修正)

詳細

資料

利用回数 1

第12回演習+解答例 (8/06 解説動画 (30分) を追加)

詳細

資料

第12回演習+解答例の(2)の図を描画するSciLabプログラム

詳細

資料

数学Ⅰ 第11回

第11回講義ノート (7/20 M講義後訂正)

詳細

資料

第11回演習+解答例

詳細

資料

第11回演習+解答例の(3)の図を描画するSciLabプログラム

詳細

資料

利用回数 1

数学Ⅱ r 2020

数学Ⅱ・物理学Ⅱ クラス分け

数学Ⅰ 期末試験

数学Ⅰ 第14回

数学Ⅰ 附録

数学ソフトウェア紹介

数学Ⅰ 第12回

数学Ⅰ 第11回

数学Ⅱ r 2020

数学Ⅱ・物理学Ⅱ クラス分け

数学Ⅰ 期末試験

数学Ⅰ 第14回

数学Ⅰ 附録

数学ソフトウェア紹介

定理 (実二次形式のグラフ) 判定定理

実二次形式 $w = \sum_{i,j=1}^n a_{ij} x_i x_j$ は、
係数行列 $A = (a_{ij})$ の 固有値 が、

① すべて正 のとき、 $x=0$ で 極小 かつ 最小

② すべて負 のとき、 $x=0$ で 極大 かつ 最大

③ その他 のとき、 $x=0$ で 極大 とも 極小 ともない。
(正のものも負のものもあって、
0 かつ 0 なければ、 $x=0$ は 鞍点。) ④ p.248 ~ p.249

(4) 二次曲線・二次曲面

(a) 二次曲線 $ax^2 + 2bxy + cy^2 + dx + ey + f = h$

$\Leftrightarrow x$ と y の二次関数の $z=h$ の 等位線。
level curve.

↓ 変数変換

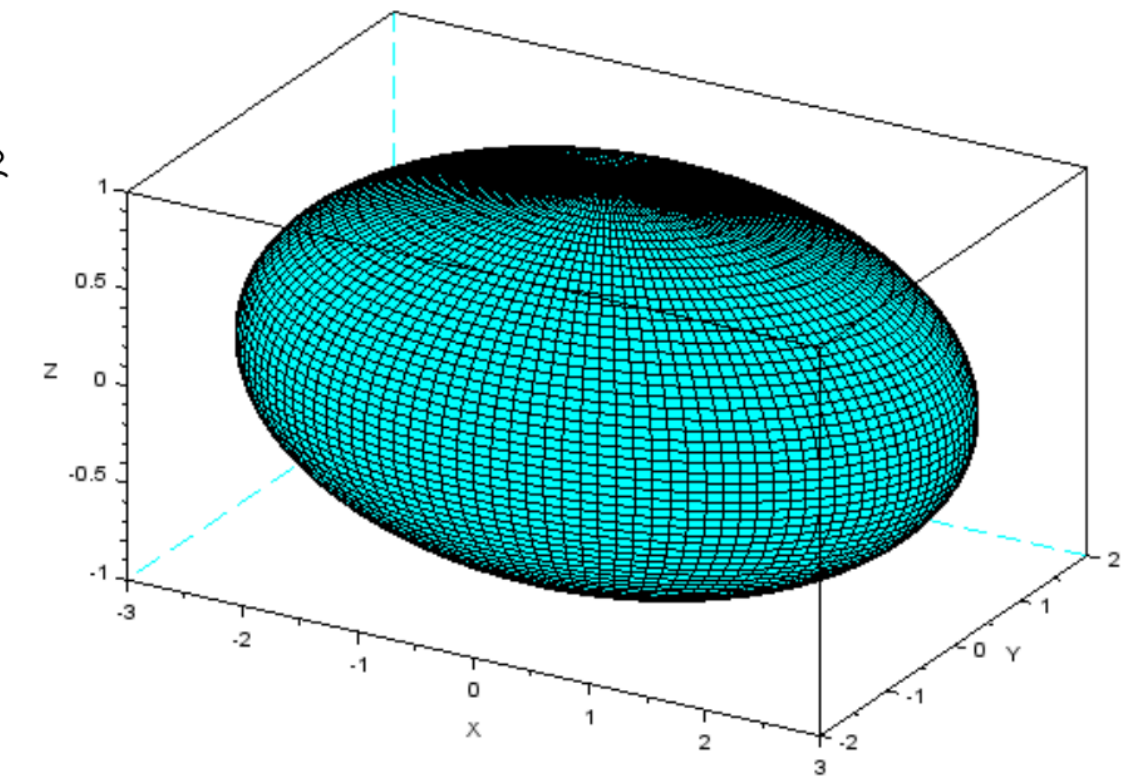
標準形 $ps^2 + qt^2 = h$ ellipse 楕円 ($pq > 0$)

あるいは、 $ps^2 = h + rt$ あるいは $qt^2 = h + rs$ 双曲線 ($pq < 0$)

($q=0$) 放物線 ($p=0$) parabola.

楕円面

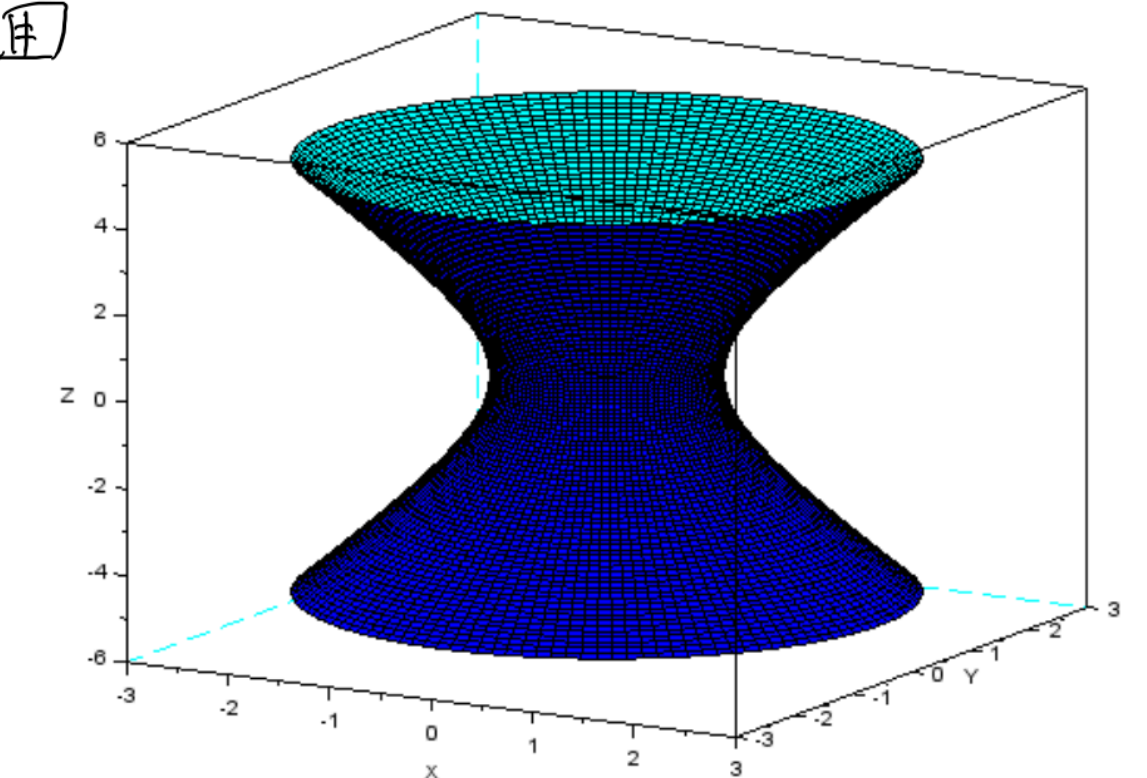
$$\frac{x^2}{3^2} + \frac{y^2}{2^2} + \frac{z^2}{1^2} = 1$$



(Scilab 6.0.2 で作成)

一葉双曲面

$$x^2 + y^2 - \frac{z^2}{2^2} = +1$$



(Scilab 6.0.2 で作成)

数学・情報処理でのプログラミングの実践・導入

(臨床統計実習(中林潤先生)より)

実践例の提示によるモチベーションの向上

データ

GSE67310

Nature 2016 8;534(7607):391-5

• 教師用データ

MEF細胞 73サンプル

Neuron 33サンプル

• D2 induced 98サンプル

D2 intermediate 20サンプル

D5 early neuron 16サンプル

D5 early myocyte 2サンプル

D5 failed reprogram 13サンプル

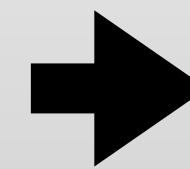
D5 intermediate 24サンプル

D22 failed reprogram 15サンプル

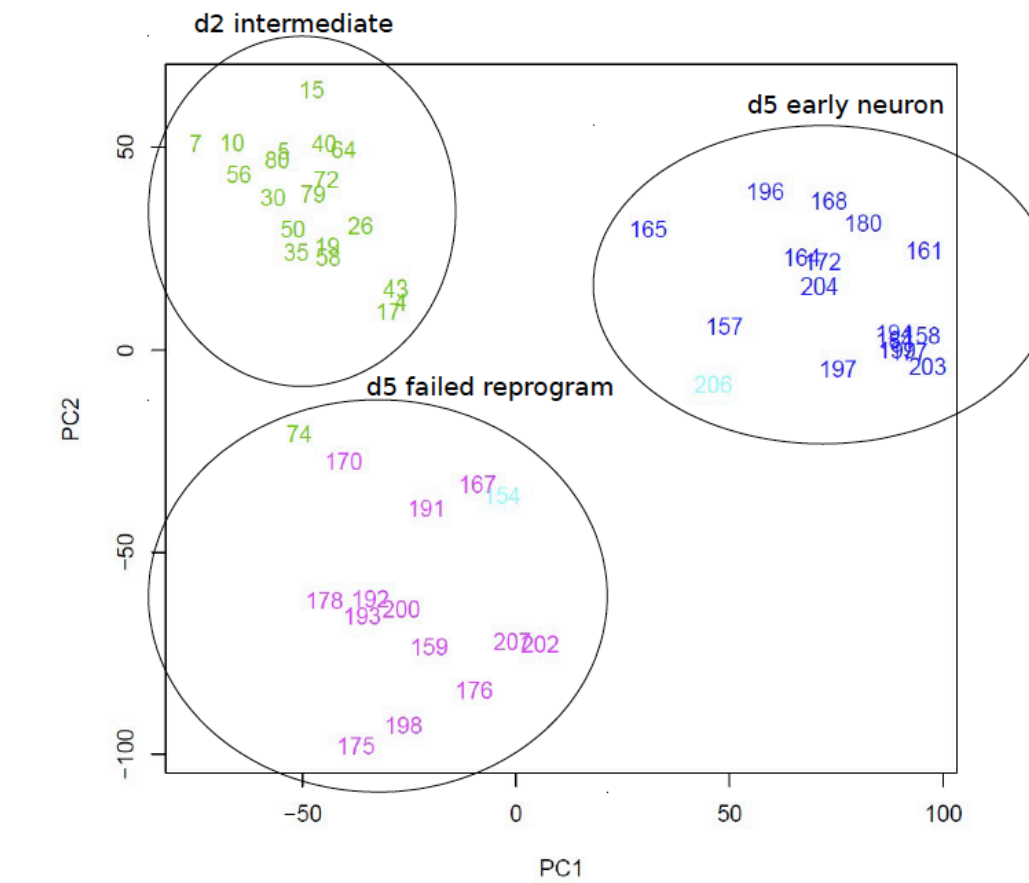
fibroblast 13サンプル

myocyte 89サンプル

22527遺伝子FPKM値



主成分分析



Dataset

• GSE76312

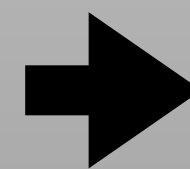
Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia.

Giustacchini A et al. *Nat Med.* 2017 Jun 23(6):692-702

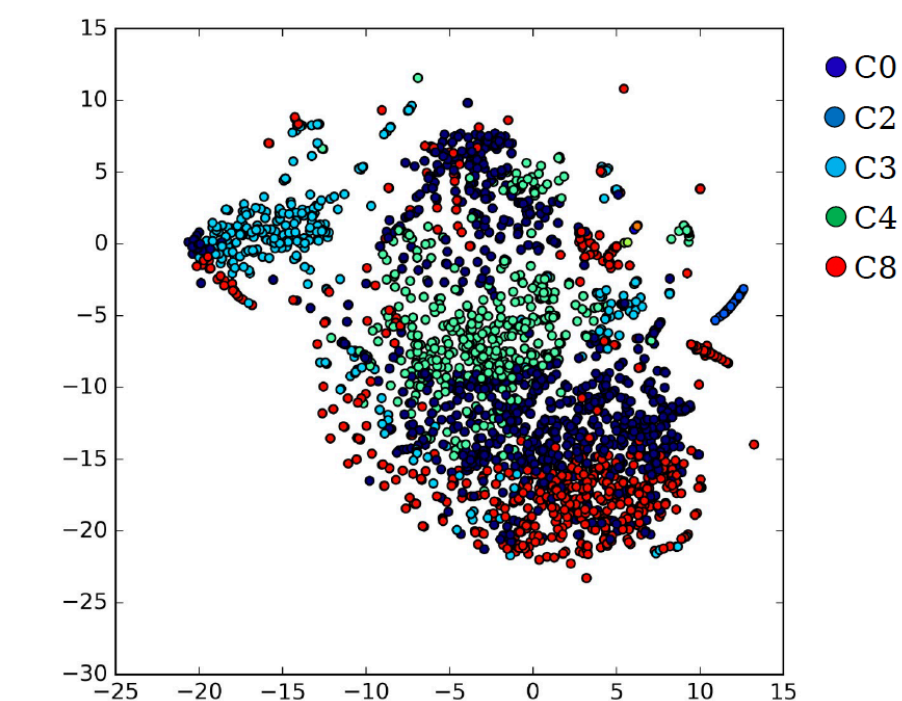
• Lin-CD34+CD38- cells

sample	number
Normal HSC	232
CML	1613
CML blast crisis	351
K562	91

23384 genes 2287 cells



CML DPGMM



数学・情報処理でのプログラミングの実践・導入

(数学Ⅰ・Ⅱr(中口悦史先生)より)

数学Ⅰ 2020 授業概要

2020 日程と授業内容予定 (各回の内容は進行状況によって変更することがあります)

回	M	D	T	授業内容予定
第 1 回	5/11	5/13	5/11	ガイダンス, 数の体系, 数列と級数
第 2 回	5/18	5/20	5/18	数ベクトルとユークリッド空間, 写像
第 3 回	5/25	5/27	5/25	行列の表記と演算
第 4 回	6/01	6/03	6/01	正則行列, 行列の階数と行基本変形
第 5 回	6/08	6/10	6/08	連立一次方程式の交代定理と解構造
第 6 回	6/15	6/17	6/15	正方行列の固有値と固有ベクトル
第 7 回	6/22	6/24	6/22	正方行列の固有空間と対角化
第 8 回	6/29	7/01	6/29	実関数の微積分, 曲線の接ベクトル
第 9 回	7/06	7/08	7/06	二次曲面と実対称行列
第 10 回	7/13	7/15	7/13	多変数関数と連続性, 陰関数と曲面
第 11 回	7/20	7/22	7/20	偏微分と全微分, 曲面の勾配
第 12 回	7/27	7/29	7/27	多変数の極値問題
第 13 回	8/31	9/02	8/31	重積分, 積分可能性, 累次積分
第 14 回	9/07	9/09	9/07	重積分の変数変換, ガウス積分
第 15 回	9/14	9/16	9/14	総合演習

7/27(M)と7/29(D)は後期「数学Ⅱ」「物理学Ⅱ」のクラス分け説明&希望調査の予定

教科書: 上村・坪井『数学入門Ⅱ』(東京化学同人)

副読本: 椎名・姫野・保科『データサイエンスのための数学』(講談社)

※ WebClass に毎回の講義ノートを順次掲載予定 (教科書とは進行が異なるので注意)。

成績評価と出席確認について:

授業15回のうち3分の2回以上, つまり10回以上出席し, 出席した回の授業内演習すべてを提出したもののみに期末試験の受験資格を与え (病欠や忌引き等は多少配慮する), その上で, 期末試験を受験したもののみを, 評価の対象とする。期末試験は筆記によって行う予定。期末試験のほかに, 授業内演習への取組みなども総合して評価する。

出席確認は, 出席管理システムZoom 参加者リストと, 演習 (リアクションペーパー) 提出を併用する。

担当教員:

中口悦史 教養部数学分野 居室: ヒポクラテスホール4階

オフィスアワー: 月・水 12:00~13:00 その他, 16:30以降なども随時対応

電子メールでは随時対応 E-mail: nakaguti.las@tmd.ac.jp

数学Ⅱ(r) (数理最適化) ガイダンス

授業担当: 中口悦史 (教養部数学)

日程と授業内容予定 (改訂2版)

1. (10/07) 数理最適化の概要

2. (10/14) 無制約非線形最適化

3. (10/21) 制約付き非線形最適化

4. (10/28) 線形計画問題 (生産計画)

5. (11/04) 線形計画問題 (続)

6. (11/11) 整数計画問題 (在庫管理)

7. (11/18) 整数計画問題 (続)

8. (11/25) (休講)

9. (12/02) 0-1整数計画問題

10. (12/09) ネットワーク計画問題

11. (12/16) ネットワーク計画問題 (続)

12. (01/06) グループワーク 2コマ

13. (01/13) (休講)

14. (01/20) グループワーク 2コマ

15. (01/27 or 02/03) 発表会・まとめ

教科書・参考書 (シラバスから修正)

教科書: 福島雅夫『新版数理計画入門』(朝倉書店, 2011年) 寒野善博『最適化手法入門』 (講談社, 2019年)

参考書: 寒野善博『最適化手法入門』(講談社, 2019年) 福島雅夫『新版数理計画入門』 (朝倉書店, 2011年) 並木誠『Pythonによる数理最適化入門』 (朝倉書店, 2018年) 藤澤克樹・梅谷俊治『応用に役立つ50の最適化問題』 (朝倉書店, 2009年) 追加!

ネットワーク計画問題 教p.109~p.139.

ネットワーク構造の制約のついた最適化問題 network structure

重み付きグラフ weighted graph.

$G=(V,E,c)$ $c=c(E)$: 各辺の重み V : 節点の集合, E : 辺の集合.

グラフ graph. 接続関係の抽象モデル.

$G=(V,E)$ $V=\{v_1,v_2,\dots,v_n\}$: 節点 node の集合. $E=\{e_1,e_2,\dots,e_m\}$: 辺 edge の集合. $e_j=(v_{j_1},v_{j_2})$ 辺は2節点に結ぶ. 始点 終点

② 辺に向きがあるとき = 有向グラフ directed graph.

グラフの例

木 tree

閉路 cycle

最大閉路 maximal cycle

1036への最短路 path.

無向グラフ $G=(V,E)$ $V=\{1,2,3,4,5,6\}$ $E=\{(1,2),(1,4),(2,3),(2,4),(3,4),(3,5),(4,5),(5,6)\}$

有向グラフ. $G=(V,E)$ 辺は一方通行. $V=\{1,2,3,4,5,6\}$ $E=\{(1,2),(1,4),(2,3),(4,2),(3,4),(3,5),(4,5),(5,6)\}$

② p.110~p.111の図も参照 辺の成分は始点, 終点の順.

数学・情報処理でのプログラミングの実践・導入

バイオインフォマティクスなどの応用までの理解へ

<https://www.shoeisha.co.jp/book/detail/9784798145600>

<https://link.springer.com/article/10.1007/s40484-019-0181-x>



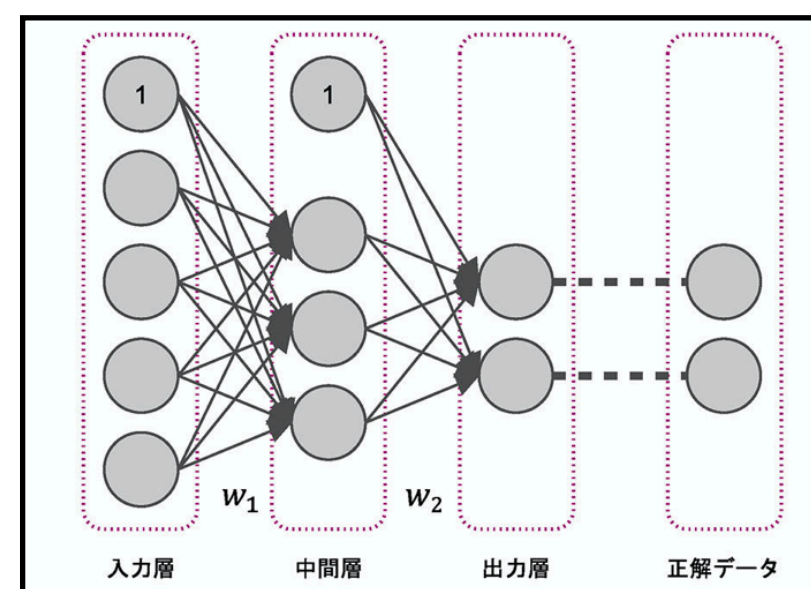
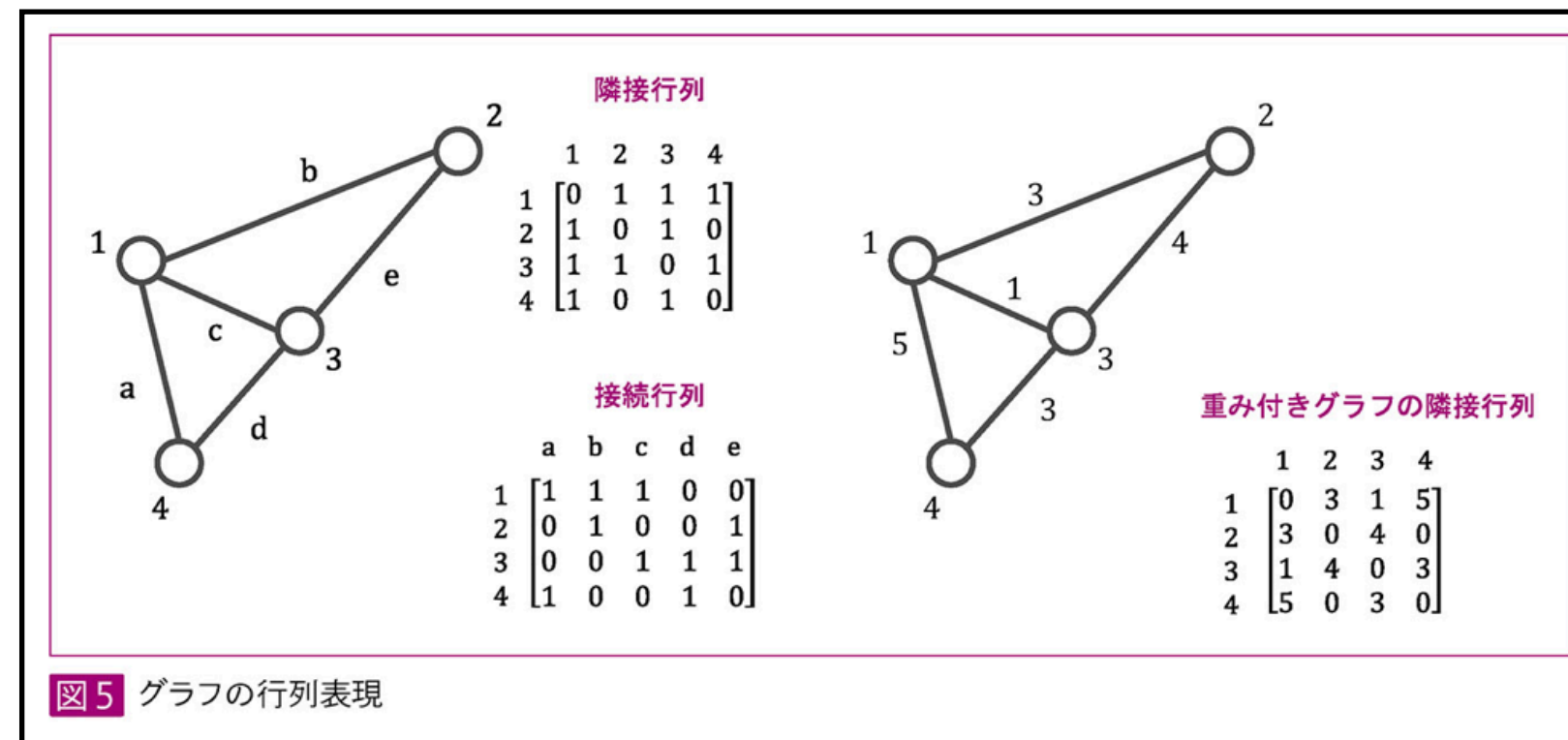
CHAPTER 5 | 重み付けと最適化プログラム

01 グラフ理論

ここではグラフ理論の基本について解説します。

POINT ♡グラフの概要

♡グラフ理論の基本



誤差関数の重みの更新式

$$w_{t+1} = w_t - \epsilon \nabla E$$

学習係数 ϵ

誤差関数の傾き

$$\nabla E = \frac{\partial E}{\partial w}$$

Quantitative Biology 2019, 7(4): 278–292
<https://doi.org/10.1007/s40484-019-0181-x>

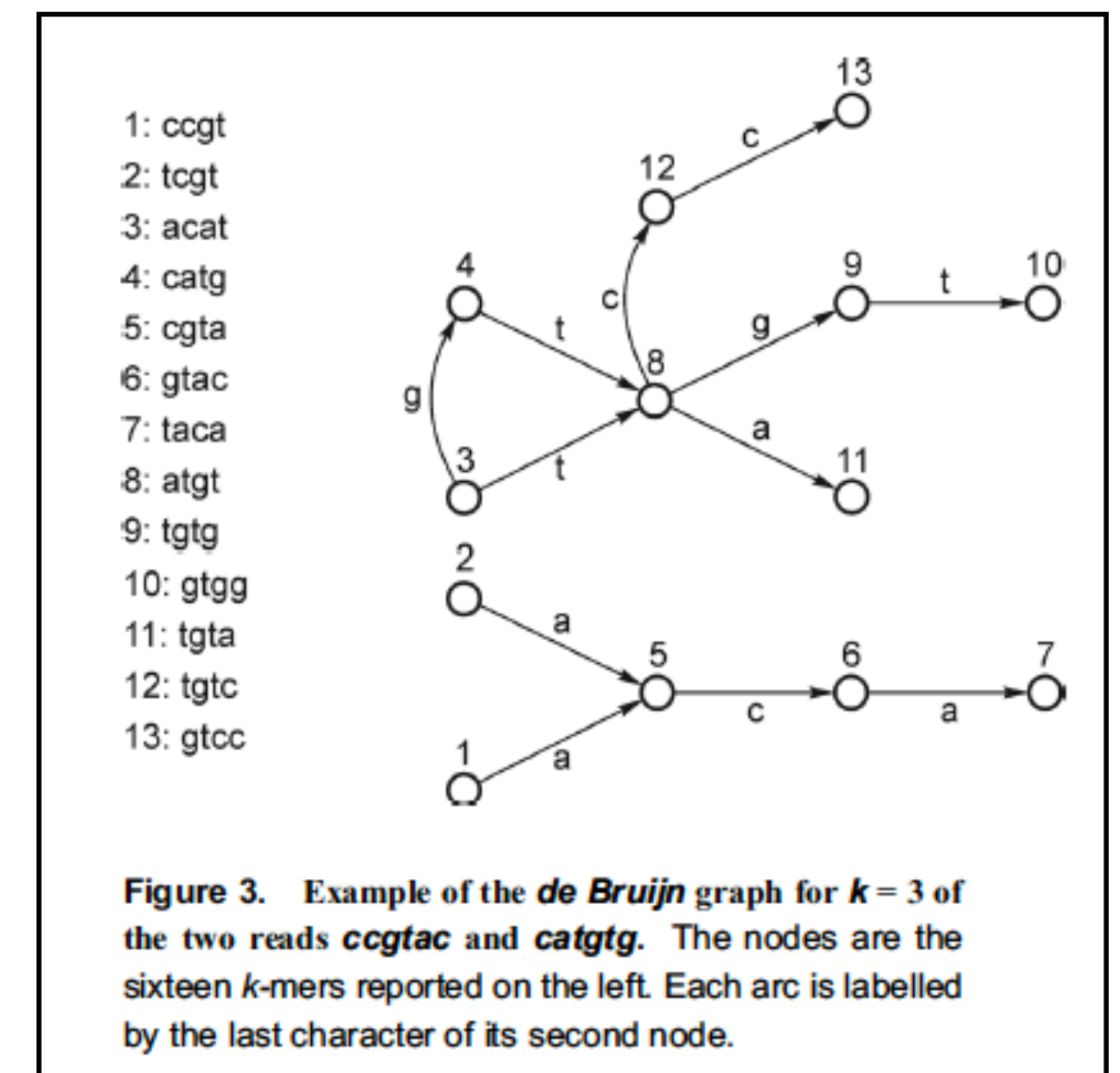
REVIEW

Overlap graphs and *de Bruijn* graphs: data structures for *de novo* genome assembly in the big data era

Raffaella Rizzi, Stefano Beretta, Murray Patterson, Yuri Pirola, Marco Previtali, Gianluca Della Vedova, Paola Bonizzoni*

Department of Informatics, Systems and Communications, University of Milan-Bicocca, Milan, Viale Sarca 336, 20126, Italy
* Correspondence: bonizzoni@disco.unimib.it

Received June 4, 2019; Revised July 3, 2019; Accepted July 26, 2019



今後の展望

2020年度

- ・ **動機付けレベル**
新設科目の設立の準備、既存の科目のプログラミング応用の検討
- ・ **実践レベル**
必要なスキルセットの整備

2021年度

- ・ **動機付けレベル**
新設科目「AI・ビッグデータ入門」実装、数学・統計処理などの改変、実装
- ・ **実践レベル**
新設科目の準備

2022年度

- ・ **実践レベル**
新設科目「AI・ビッグデータ応用」実装
- ・ **他の医療系大学との連携**
本学のモデルカリキュラムの配布
スキルコンテンツの共有

御清聴ありがとうございました