



連載



情報の授業をしよう！

本コーナー「情報の授業をしよう！」は、小学校や中学校で情報活用能力を育む内容を授業で教えている先生、高校で情報科を教えている先生や、大学初年次で情報科目を教えている先生が、「自分はこの内容はこういう風に教えている」というノウハウを紹介するものです。情報のさまざまな

内容について、他人にどうやって分かってもらうか、という工夫やアイディアは、読者の皆様にもきっと役立つことと思います。そして「自分も教え方の工夫を紹介したい」と思われた場合は、こちらにご連絡ください。

(E-mail : editj@ipsj.or.jp)



高等学校教科情報における Python を利用した統計学習 —問題解決能力の向上を目指して—

阿部百合 | 二松学舎大学附属柏中学校・高等学校

高等学校教科情報における統計学習の意味

データサイエンス教育が注目され、統計の必要性が高まっている。現行の高等学校学習指導要領では次のように統計の内容が含まれている。

数学科：科目『数学Ⅰ』の中で高校1年次に「データの分析」として標準偏差や相関係数までを学習する。

情報科：科目『社会と情報』では「情報社会における問題解決」の中で扱う。科目『情報の科学』では「問題解決とコンピュータの活用」に含む。

筆者の勤務校では高校1年次に『社会と情報』を学習しており、情報科の問題解決学習の中で、文系の生徒でも受け入れやすいよう統計の授業を工夫してきた。

その特徴は、多くの高校で行われているような、

表計算を用いたデータ処理（数学科では計算演習とされる部分）ではなく、Python プログラミングによるデータ処理である。その考え方と授業内容を紹介する。

実施した授業の目的は①生徒が統計データの利便性、有用性を理解する②生徒が必要性和状況によって（紙と鉛筆または ICT のどちらを使うか、ICT を使う場合には、どのアプリケーションを使うかといった）道具の使い分けができるようにすべきであることを理解することとした。

今回の実践のねらいは、身近な問題に対して、表面だけをみて真の原因を究明せず不満を言うのではなく、データ収集・分析を行い、説得性の高い問題解決案を提示する能力を生徒に身に付けさせることである。

授業全体の構想と各回の流れ

2018年度は授業時数50分×6コマ（不連続、週2コマずつ3週間）で計画し、実際には5コマで実施した。5コマ実施となったのは、授業時数が確保できなかったためである。生徒が自分たちの身近な問題をテーマとして設定し、データ収集、分析、まとめ発表も含めて5コマで実施した。

授業時数と単元分けを図-1に示す。

図-1中の「必要」とは、単元内容を学習する際に最低限必要であると著者が考える時数である。これだけの時数は確保しないと、一部内容を実施できない、もしくは、無理に実施すると平均的能力の生徒の中で授業についていけないと感じる者が発生すると思われる。「実施」は2018年度に、実際に行った授業時数である。()内のPython利用時数は、実習の中で、次節以降で述べるように、生徒が実際にPythonを利用した授業時数である。「不足」は「必要」－「実施」の時数である。

2018年度は、実習に充てる時間が大幅に不足しており、授業時数に関しては大いに改善の余地がある状況であった。そこで2019年度は、この不足時数を補うように数学科と情報科のシラバスを作成し、全11コマ（うち3コマ数学）で、今秋実施する予定である。

図-2に、2018年度の授業単元と題材を示す。

単元分けは、知る、分かる、使えるの3ステップとした。また、各回の授業は、「用語解説→実習、事例紹介→小テスト」の流れを基本とした。ただし後半の回では、小テストの代わりにグループによる生

徒ら自身の問題発見とその分析、発表準備とした。最終回は、グループ発表と相互評価、自己評価（ふりかえり）の時間とした。

1時間目から3時間目では、記述統計と推測統計、数学の教科書の定義を用いて代表値（最頻値・中央値・平均・分散・相関係数）、データの表現方法としてグラフの種類とその特性を学んだ。さらに、具体的な統計データからデータの特徴や傾向を読み解き、Pythonや表計算ソフトを使った統計処理を実習した。

図-2中の「データ収集と分析」「分析結果と提案の発表」は生徒4人ずつのグループで実施した。各グループが身近な問題を発見し（ブレインストーミングによるタブレットへの書き出しや、ミーティングによる決定）、アンケート調査やWebから生データを取得した。アンケート調査は、生徒がグループごとに作成したアンケートを、お互いに回答しあうことで行った。その後、取得したデータをPythonやGoogleスライドの機能、Excelを使って分析した。さらに、GoogleのClassroomを使って配布したテンプレートを利用して発表資料を作成し、発表した（アンケート調査や分析ツールについては付録を参照）。

実習でPythonを利用した理由は、プログラミングでも計算できることを実感させるためである。（マウス操作が苦手な生徒は意外とたくさんいる。彼らはマウス操作によるオートフィルを使うExcelでの計算に苦戦する。むしろキーボード操作でプログラムを書き、Pythonで計算する方が簡単と感じる生徒も結構存在する）。また、生徒は、計算というものは、数学の時間は紙と鉛筆で計算し実生活では役に立たないもの、情報の時間は表計算ソフト（Excel）

単 元	統計実習時数(うちPython利用時数)		
	必要	実施	不足
記述統計と推測統計	1(0)	1(0)	0(0)
代表値とその計算	2.5(2)	1(1)	1.5(1)
データの表現方法と読み取り	1.5(1)	1(0.5)	0.5(0.5)
データの整理方法	1(1)	0(0)	1(1)
データ収集と分析、資料作成	1.5(1)	0.7(0.3)	1.8(1.7)
相関関係と因果関係	1.7(0.7)	0.3(0.2)	1.4(0.5)
いくつかの分析手法	0.3(0.3)	0(0)	0.3(0.3)
分析結果と提案の発表	1.5(0)	1(0)	0.5(0)
合計	11(6)	5(2)	6(4)

■図-1 授業時数と単元分け

単 元	実施時間	題材、内容
記述統計と推測統計	1時間目	検品、視聴率、リーグ戦の結果
代表値とその計算	2時間目	沼の魚の体長、テストの点数
データの表現方法と読み取り	3時間目	お菓子の箱、ニュースのグラフ画像
データ収集と分析、資料作成	4時間目	アンケート実施*、テンプレートによる資料作成(googleスライドで配布)
相関関係と因果関係	4時間目	身長と体重、四角形の面積と辺の長さ
分析結果と提案の発表	5時間目	グループごとに成果発表、相互評価

■図-2 授業の単元と題材

を使う面倒なもの、普段の生活ではスマホの計算アプリや電卓ですもの、と分けて考えている。現実に直面する問題解決では、限られた時間内で成果を出すことが求められており、利用できる ICT 手段は、臨機応変にすべて利用すべきである。時間の制約を受けながら成果を求められる場面として、小テストを位置づけた。小テストは、知識や理解を確認するためでなく、時間内で問題解決をするためには、自分にとってどのような手段が最適であるかを生徒一人ひとりが考える時間である。

統計学習と生徒の発表の環境

統計学習の実施環境について述べる。

授業はコンピュータを使用して進めた。小テスト、グループでのデータ収集と分析ではタブレットの使用も許可した。なお、本校ではコンピュータとタブレットは1人1台使用できる。ソフトウェアについては、単元「代表値とその計算」「データの表現方法と読み取り」(図-1)においては全員 Python を使用した(小テストのときは Python 以外も使用可)。単元「データ収集と分析、資料作成」のグループでの学習の際は Python に限定せずほかの計算ソフトの使用も可とした(図-3)。

Python 実習の環境

ここでは Python をどのように使うかと処理するデータの入力方法について述べる(図-4)。具体的

な統計処理のプログラムは次節で詳しく述べる。

また、Python に限らずプログラミングを取り入れた授業では、その環境設定をハードルに感じる方が多い。「Python を使ってみたいけれど、どうやって Python を使える環境を準備したらよいのか分からない」、「Python のインストールが難しそう」と言われることがある。そこで、今回はインストールの手順から説明する。

事前準備：

Python を利用するために必要な事前準備は、生徒の使用するコンピュータに Python をインストールするだけである。インストールは生徒に各々させたいところであるが、セキュリティの都合上復元ソフトやフィルタリングをかけているので教員が行った(参考：生徒にインストールからさせたこともある。以前行ったさまざまなソフトウェアを使った問題解決の授業の際には、必要なグループの数だけ Sikuli や Processing をインストールさせた)。

Python のインストールには pip を使って Python 本体のみを入れる方法と、パッケージと呼ばれる Python 本体と Python でよく使われるプログラムを集めたライブラリ、コードを読みやすくしてくれるエディタなどのセットを入れる方法がある。パッケージはさまざまあるが、Anaconda (<https://www.anaconda.com/>) がインストールの手順も分かりやすく便利である。特に、インストール時のハードルとなりやすいパスの設定(コンピュータのどこに Python を置くか)が、Anaconda ではパスの自動設定のチェックボックスにチェックを入れるだけで済む。

種類	具体的名称
PCハードウェア生徒機	FUJITSU FMVD1504Vモデル intel CORE i3
PCのOS	Windows10 ver.xxx
タブレット端末	iPad9.7インチcellular A1823モデル
統計処理に使用したソフトウェア	Python 3.7.0 (Anaconda3 5.3.0 64bit) Excel 2016 Googleスライド内Googleスプレッドシート
発表資料作成に使用したソフトウェア	Googleスライド

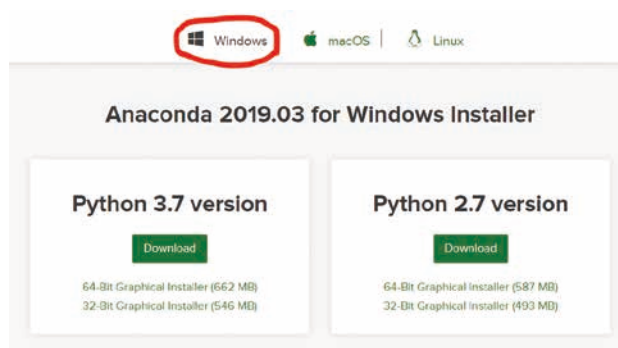
■図-3 統計学習と生徒の発表環境

種類	具体的名称
環境	エディタ Windows PowerShell に直打ち
	スクリプト言語 Python
	シェル Windows PowerShell
使用コマンド	Python, import
利用ライブラリ	statistics, math, numpy, matplotlib.pyplot
関数名	mean(), pvariance(), stdev(), corrcoeff(), grid(), scatter(), show()
データの入力方法	変数を用意して配列として数値を手入力 例) test = [0, 1, 2]

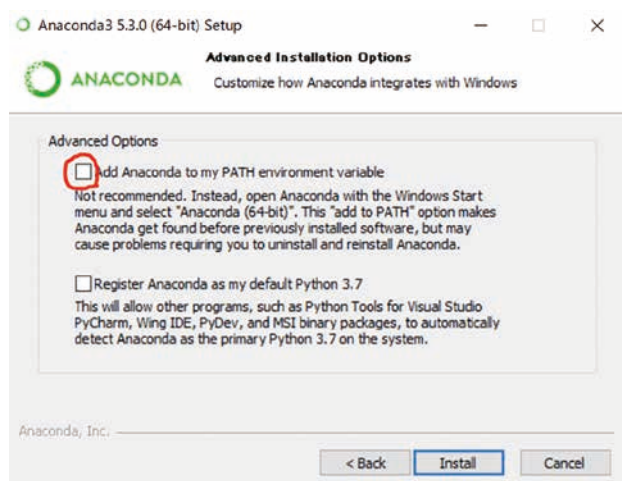
■図-4 Python 実施環境と使用方法

Anaconda のサイトとインストール手順は英語であるが英語をハードルに感じる必要はない。サイトでは“Download”の文字を探して進み、使用するコンピュータの OS (高校ではたいてい Windows (参考図-3)) 用のパッケージをダウンロードすればよい。デフォルトでは Mac 用のインストーラが表示されている。Windows 用をインストールする場合は、図-5 の赤丸で示した Windows をクリックすると Windows 用のインストーラが表示される。

インストールはダウンロードしたファイルを開き、デフォルトのまま“Next”や“I agree”をクリックして進めればよい。ただし、パスの自動設定のチェックボックスにはデフォルトでチェックが入っていないため、ここだけは自分でチェックを入れる必要がある (図-6)。



■図-5 Anaconda ダウンロード画面 (Windows 用表示)



■図-6 Anaconda インストールの注意すべき画面

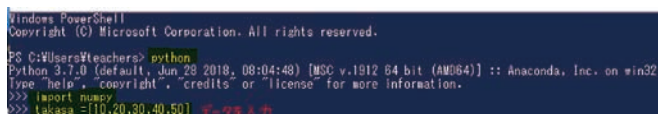
Python の実行方法：

Python の起動および、プログラムの記述と実行はすべて Windows PowerShell で行う (本学ではターミナルやコマンドプロンプトをセキュリティ上使用できないため Windows PowerShell を使用した)。

Windows PowerShell は Windows コンピュータに標準装備されている。Windows のスタートボタンから Windows PowerShell フォルダの中で一番上にある Windows PowerShell を起動すればよい。

Windows PowerShell を起動後、Python と入力し Enter キーを押せば Python が実行できる。今回のプログラムはごく短いものであるため保存しない。そのためエディタは使用せず、一行ずつプログラムを直打ちしその都度実行する (図-7)。

ここで、Anaconda パッケージには、開発環境、エディタとして Jupyter Notebook があり、すぐに使うことができるのに、なぜそれを使わないのか、と疑問に思われる方もいるだろう。Jupyter Notebook やほかの開発環境をお使いの方はご存知のように、色分けしてくれたり自動でインデントを入れてくれたりタイピングミスがある場合は修正してくれたりするので、こういった開発環境の利用は大変便利である。しかし、GUI 操作は一見便利のように感じられるが、生徒 43～45 人に対して教員 1 人の場合、その操作を説明することはなかなか厄介である。たとえば、戻るボタン 1 つ説明するのに、「画面の右上にある U の字を横向きにした先に矢印がついているマーク」とか「1 回作業、話している人はやめて、先生の手元が映っているモニタ見て」とか「モニタに映っている説明 (あらかじめ用意した、操作画面のスクリーンショットにボタンの説明書きを加えておいた資料等) を見て」といったことであ



■図-7 Windows PowerShell での Python 実行の様子
黄色の文字は入力内容、白い文字はコンピュータ側の実行による出力、赤字は筆者が追記したコメントである

る。厄介だけでなく、こういった GUI の説明をすることはナンセンスである。開発環境の使い方を説明することが授業の目的ではない。そもそも、同じソフトウェアであっても、バージョンは頻繁に変わる上、その都度操作性が変わることなど日常茶飯事である。たまたま授業で扱うことになった、たった1つのアプリケーションの操作をていねいに説明したところで、生徒が社会に出ていくときにはほとんど意味がない（これは高校の情報の授業は某ソフトウェアの使い方の授業と誤解している人々に特に言いたい）。

したがってそういった操作の説明をできる限り排除しようとした結果、今回のように開発環境を使わない CUI による操作に落ち着いた。なお、今後もっと複雑なプログラムを書こうとしたときや、生徒のグループ学習によるデータ処理の際には Jupyter Notebook のようなエディタを使用することも検討している。コードの保存が必要になること、書く量が増えるのでスペルミスのチェック機能が欲しいこと、複数の人で取り組むので誰でもが見やすいインデントが欲しいことがその理由である。

データ入力方法：

データ数が多いほど統計処理の意味も大きくなり、ICT を利用するメリットも大きくなるので本来であれば大量のデータを扱いたい。しかし、大量のデータを扱うためにはデータベースの利用やデータフレームの設計が欠かせない。データベース作成や外部ファイルからの読み込みはエラーが発生するリスクが格段に増える。今回はプログラミング初学の生徒全員が Python を使って代表値の計算と分析結果の可視化（グラフ表示）ができることを優先した。そのため、データ数を少なくし手入力とした。

単元と Python 実習の関係

それぞれの単元の中で Python の利用を計画した（図-8）。図-8 では、Python を利用する単元を黒字

で、使用しない単元を灰色の文字で表している。また、その単元で使用する関数も記載した。

しかし実際には、図-1 で示したように授業時数不足と難易度から、Pandas パッケージを利用した「データの整理方法」、SQLite3 を利用した「データ収集と分析、資料作成」でのデータベースと外部ファイルの利用、seaborn パッケージを利用した「いくつかの分析手法」は実施しなかった。

実施した Python 実習の中から、代表値の計算と散布図の表示のプログラムを1つずつ紹介する。

代表値の計算：

Python の実行は前段で述べた通り Windows Power Shell を利用して行う。図-9 のスクリプトの1行目にあるように、Python と入力し Enter キーを押すと、そこから Python の実行環境になる。以後1行書くごとに Enter キーを押して実行していく。図-9 のプログラムは、2つのクラスの小テストの点数の平均点と標準偏差を計算し、比較・考察するためのものである。2, 3 行目では統計処理のためのライブラリをインポートしている。生徒には、図書館で専門書を借りてくると同じように、使いた

単 元	利用関数	利用箇所
記述統計と推測統計		
代表値とその計算	statistics mean(), pvariance(), stdev()	平均、分散、標準偏差の算出
データの表現方法と読み取り	matplotlib.pyplot grid(), scatter(), bar(), show()	折れ線グラフ、散布図、棒グラフの描画
データの整理方法	pandas read_excel(), text()	csv形式の外部ファイルの読み込みと配列に整理
データ収集と分析、資料作成	sqlite3	データベース作成のための言語(Pythonで使える)
相関関係と因果関係	statistics corrcoef()	相関係数の算出
いくつかの分析手法	seaborn regplot(), lmpplot()	回帰モデルの可視化
分析結果と提案の発表		

■図-8 統計実習での Python の利用関数と利用箇所

```

1 python
2 import statistics
3 import math
4 testA = [70, 50, 55, 65]
5 testB = [20, 95, 35, 90]
6 statistics.mean(testA)
7 statistics.pvariance(testA)
8 statistics.pvariance(testB)
9 statistics.stdev(testA)

```

■図-9 代表値の計算のスクリプト

い分野についてすでに書かれているプログラムの集まり（パッケージ）を使わせてもらうための1行と説明している。4, 5行目では計算したいデータを入力している。6行目では借りてきたプログラム（統計処理をするときに使うプログラムの名前は、統計を示す英単語 statistics になっている）を使って平均（mean）を計算している。この1行を実行すると testA の平均点 60 が出力される。ここで生徒には、同様に testB の平均点を出力するプログラムを1行書かせる。7, 8行目では分散（pvariance）を、9行目では testA の標準偏差（stdev）を求めている。授業実施時点では、数学 I において、生徒は「データの分析」を未習であった。そのため代表値の定義は数学の教科書に従い、式も講義で確認した上で図-10のスク립トのように計算は Python に任せる授業とした。定義を確認しなければ、返ってきた出力結果の数字を見ても何も分からないからである。

散布図の表示：

代表値の計算のスク립ト同様、1行目で Python の環境に入る。

2行目では相関係数を求めるためのライブラリをインポートしている。3, 4行目では、相関を知りたい2組のデータを入力している。5行目で相関係数を求めている。6行目ではグラフ表示のためのパッケージ（matplotlib.pyplot）をインポートし、さらに、パッケージ名が長いので plt と短い呼び方で利用できるようにしている。7行目で散布図（scatter）を指定し、8行目で表示（show（））している。

グラフ表示は散布図（scatter）以外に bar や pie を指定することで棒グラフや円グラフも表示できる。

```
1 python
2 import numpy
3 takasa = [10,20,30,40,50]
4 yokohaba = [5,4,3,2,1]
5 numpy.corrcoef(takasa,yokohaba)[0,1]
6 import matplotlib.pyplot as plt
7 plt.scatter(takasa,yokohaba)
8 plt.show()
```

■図-10 相関係数の計算と散布図のスク립ト

実習では、一人ひとりプログラムを写し、全員が散布図の表示を確認した。その後グループごとのアンケートデータの処理の際に Python を使いグラフ表示までしたグループが2グループあった。

授業の結果と生徒の反応

授業をした結果、意外な生徒の反応が得られた。マウス操作よりキーボード操作の方が簡単であった。実行画面が黒く、すべてキーボード操作で進むことに「コンピュータを操作している感じがする！」、キーボード操作だけでグラフが表示される様子を「かっこいい！」と驚き喜ぶ姿があちこちで見られた。Python の利点は、プログラミング経験がなくても簡単にコードが読めるようになり、コードを書いて実行できることである。保存を必要としない、まさに計算機代わりに使う程度ならば、テキストエディタすら不要である。これはファイルの保存場所の概念ができあがっていない生徒にとってプログラム実行のハードルが格段に下がる大変ありがたい点であったと気づいた。また、コードを書く方法でプログラミングすることは、キーボード操作だけで計算ができ、グラフが表示されるので、左手でマウス操作をする、ダブルクリックができない、ポインターの位置を器用に動かせない生徒にとって、新鮮で嬉しい経験であった。

期末テストの際に実施したアンケートでは以下の囲みのようなコメントが得られた。

「自分で作業する時間が多く、頭に入りやすくてとても良かったです！！ Python などを使った計算は家でも使えそうで良かったです」

「Python が楽しかった」

「もう少しゆっくり教えてほしかった」

「アンケートをとるのが大変だった」

「もう少し発表の日を遅くしてほしかった」

今後の展望

これまで、単元の扱い、授業時数、環境設定、開発環境の利用、Python のプログラムなどについて述べてきた。今後、より内容を広げていく上での課題、テーマとして次のことが考えられる。

- (1) 数学科との連携による時間不足の解消
- (2) 開発環境の利用によるデータ設計の学習
- (3) 実データの扱いによるデータ前処理の学習

(4) 実データの扱い（問題解決学習）

(5) データベース設計とその利用

今後も、生徒の状況に合わせて、また最新の情報を求めつつ改良を重ねていきたい。

(2019 年 5 月 13 日受付)

阿部百合（正会員） abey@nisho-k.jp

二松学舎大学附属柏中学校・高等学校教員。2016 年より情報科（数学兼任）。2019 年より数学科（情報兼任）。中学数学、高校数学、情報を担当している。

2018年2学期データの分析による問題解決 生徒が取り組んだテーマ一覧

テーマ	調査方法	調査対象	分析ツール
学校満足度調査	聞き込み	クラス43人	Python
テスト勉強はいつからやるのか	聞き込み	クラス43人	googleスライド
お菓子和本物のたけのこときのこの好みの関連	聞き込み	クラス43人	googleスライド
ラーメンの人気のない味の消費量を上げよう	アンケート	クラス43人	googleスプレッドシート
学校が面倒くさいか	アンケート	クラス43人	googleスライド
1-3でうどんとそばどちらがよく食べられているか	聞き込み	クラス43人	googleスライド
ディズニーランドかディズニーシーか	聞き込み	クラス43人	googleスプレッドシート
スクール鞆のバッグ提案	Instagram	フォロワー	Excel
学校が良いか悪いか	聞き込み	クラス44人	googleスライド
就寝時間と授業の集中度の関係	Instagram	フォロワー100人	googleスライド
起きる時間と家を出る時間	聞き込み	クラス44人	googleスライド
iPhoneとAndroid利用者の違いによる偏見をなくす	聞き込み	クラス44人	googleスライド
睡眠時間と朝食の関係	聞き込み	クラス44人	Python
人気の味のラーメンを調べ販売数増を提案	LINE	クラス44人	googleスライド
県別総生産からどの県がお金持ちかを調べる	統計局データ	各都道府県	googleスライド
嫌いな科目と苦手な科目に関連があるか	聞き込み	クラス44人	googleスライド
好きな人がいるかからクリスマスの過ごし方を予測	googleform	学年全体	googleスライド
外食の頻度から経済の活性化を図る	聞き込み	クラス+知人	googleスライド
好きな人がいるかとその人と付き合っているか	聞き取り	クラス42人	googleスライド
ペットを飼っているかを知りペットとの時間を増やす	アンケート	クラス42人	googleスライド
日本の人口の遷移から今後を考える	統計局	1900年から	googleスライド
睡眠時間から授業内睡眠時間を減らすことを考える	聞き取り	クラス42人	googleスライド
船橋市の月別降水量から傘持参を考える	気象庁	1990～2010	googleスライド
YouTubeを1日どのくらい見るかを調べ成績upを狙う	Instagram	フォロワー	googleスライド
通学に自転車を使う割合	LINE	クラス44人	googleスライド
スクールバスの利用率と座れる人の割合	聞き取り	クラス44人	googleスライド
平均睡眠時間から授業集中力up	聞き込み	クラス44人	googleスライド
食堂と購買の利用状況	アンケート	クラス44人	googleスライド
帰宅時間と就寝時間から見る翌日への影響	SNS	フォロワー	Python
授業で寝ないために朝ごはんかパンか	Twitter	フォロワー	Twitter集計ツール
初回と最終回から見る漢字テストの点数を上げよう	点数一覧表	クラス44人	Python
化学の時間に寝ている男女	15分ごとに計測	クラス44人	googleスライド

©阿部百合