

日本の家族構成と 行動分析を介した 発展的な産学融合

立正大学データサイエンス学部
2023年12月8日（金）
白川 清美



データサイエンスセンタープロジェクト

1. プロジェクト名：

「日本の家族構成とその行動分析」

2. プロジェクトの期間

2021年10月1日～2024年9月30日

3. 概要

社会生活基本調査や国勢調査の公的統計調査のデータに基づいた無償労働時間の男女間格差や少子化状況の分析及びその結果の可視化により、根本的な要因を導出する。



連携企業

- プロジェクト実施による企業との連携

- 1. 株式会社タクミイノベーションテクノロジー

- 生活時間調査データの分析
- SASを使ったデータ分析の講義の実施
 - 機会学習ができるSAS VIYA等

- 2. 株式会社プラージュ

- 生活行動データ可視化アプリ
 - 「Life Trackerの開発（デモ可能）」
- データ取得システムの構築（開発中）



連携企業

- プロジェクト実施による企業との連携

- 1. 株式会社タクミイノベーションテクノロジー

- 生活時間調査データの分析
- SASを使ったデータ分析の講義の実施
 - 機会学習ができるSAS VIYA等

- 2. 株式会社プラージュ

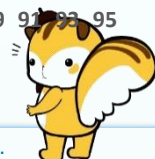
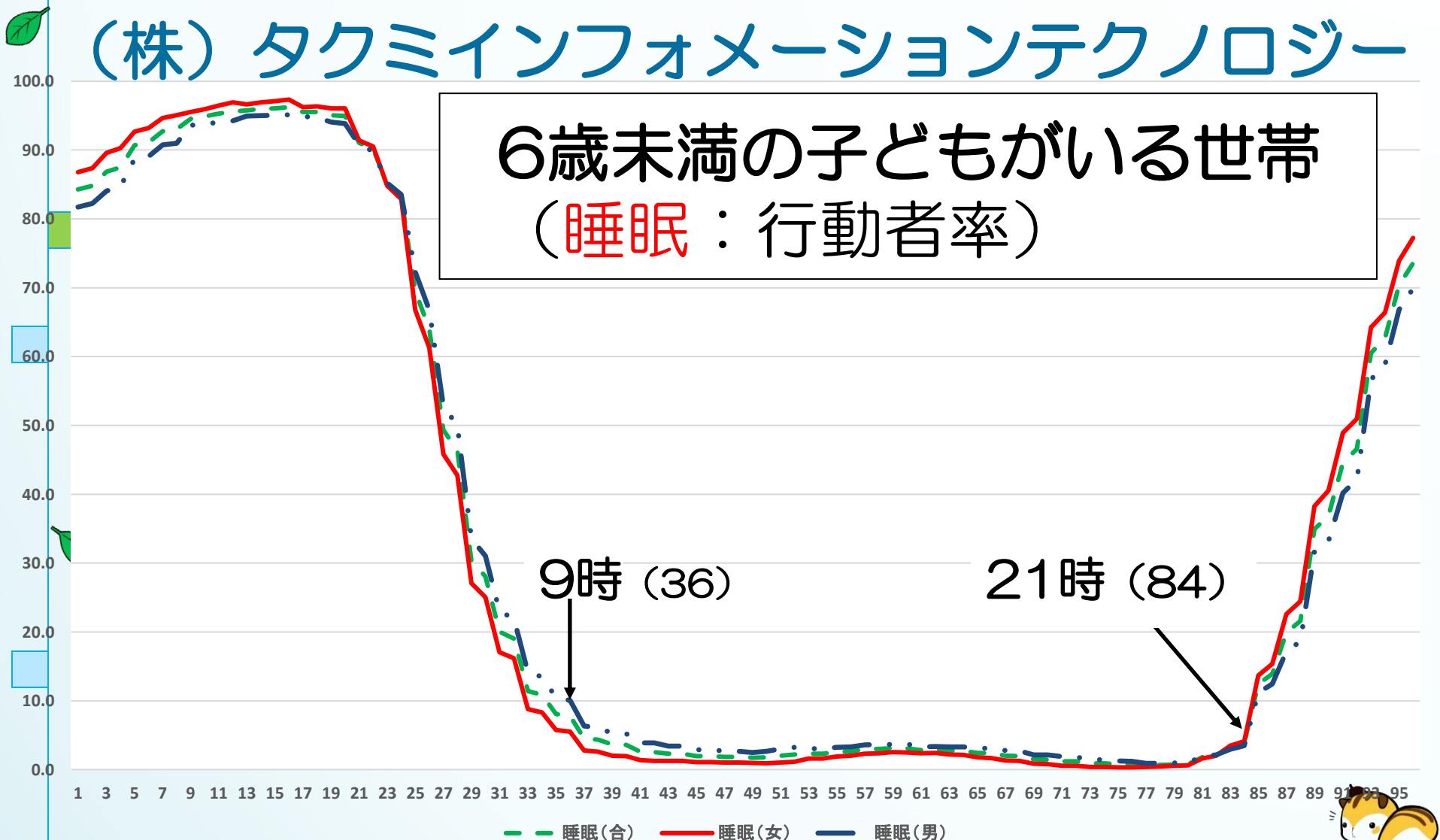
- 生活行動データ可視化アプリ
 - 「Life Trackerの開発（デモ）」
- データ取得システムの構築（開発中）



協働分析

(株) タクミインフォメーションテクノロジー

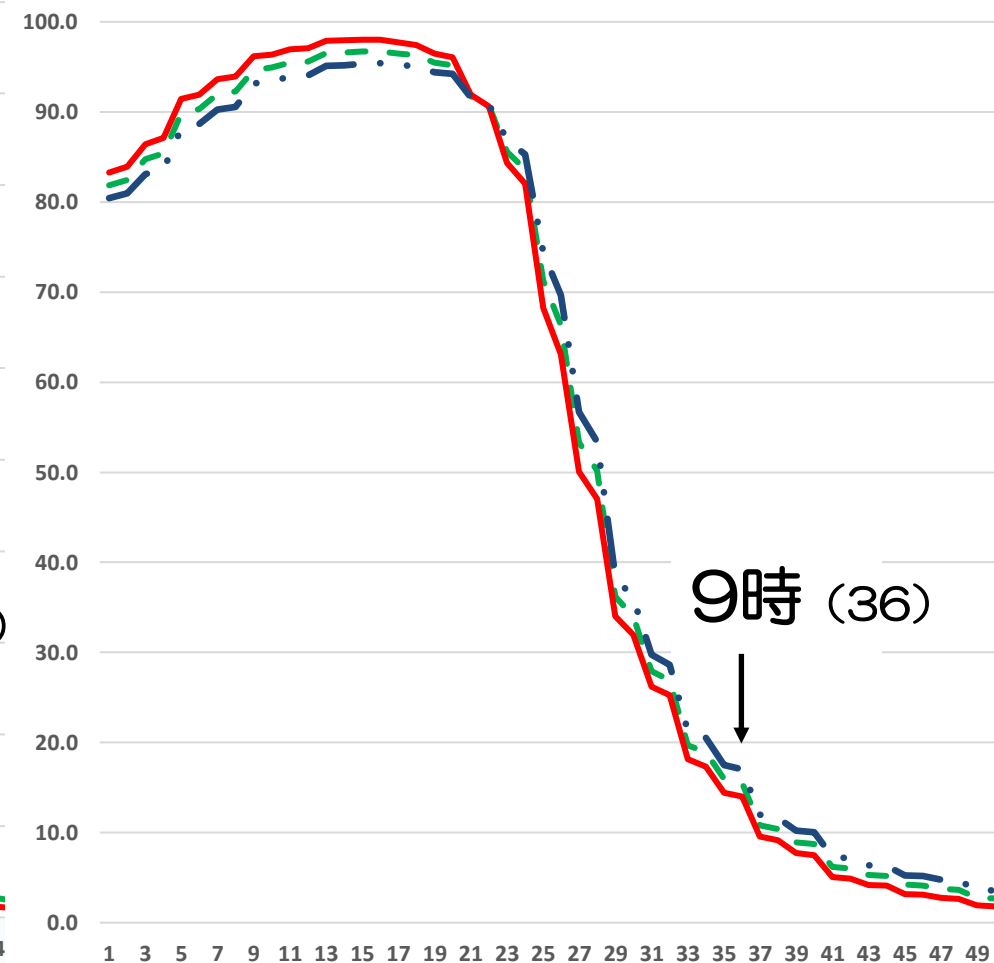
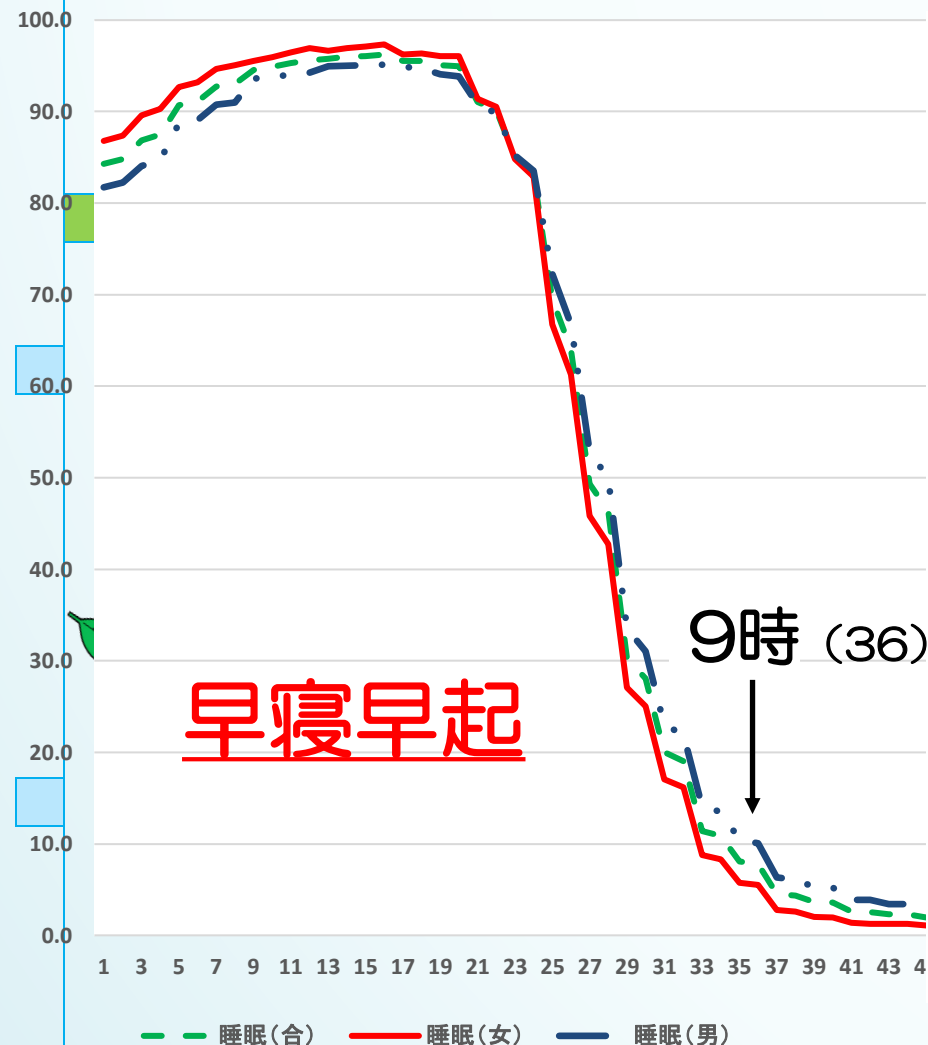
6歳未満の子どもがいる世帯
(睡眠：行動者率)



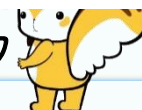
子どもがいる世帯比較（睡眠：行動者率）

6歳未満

それ以外



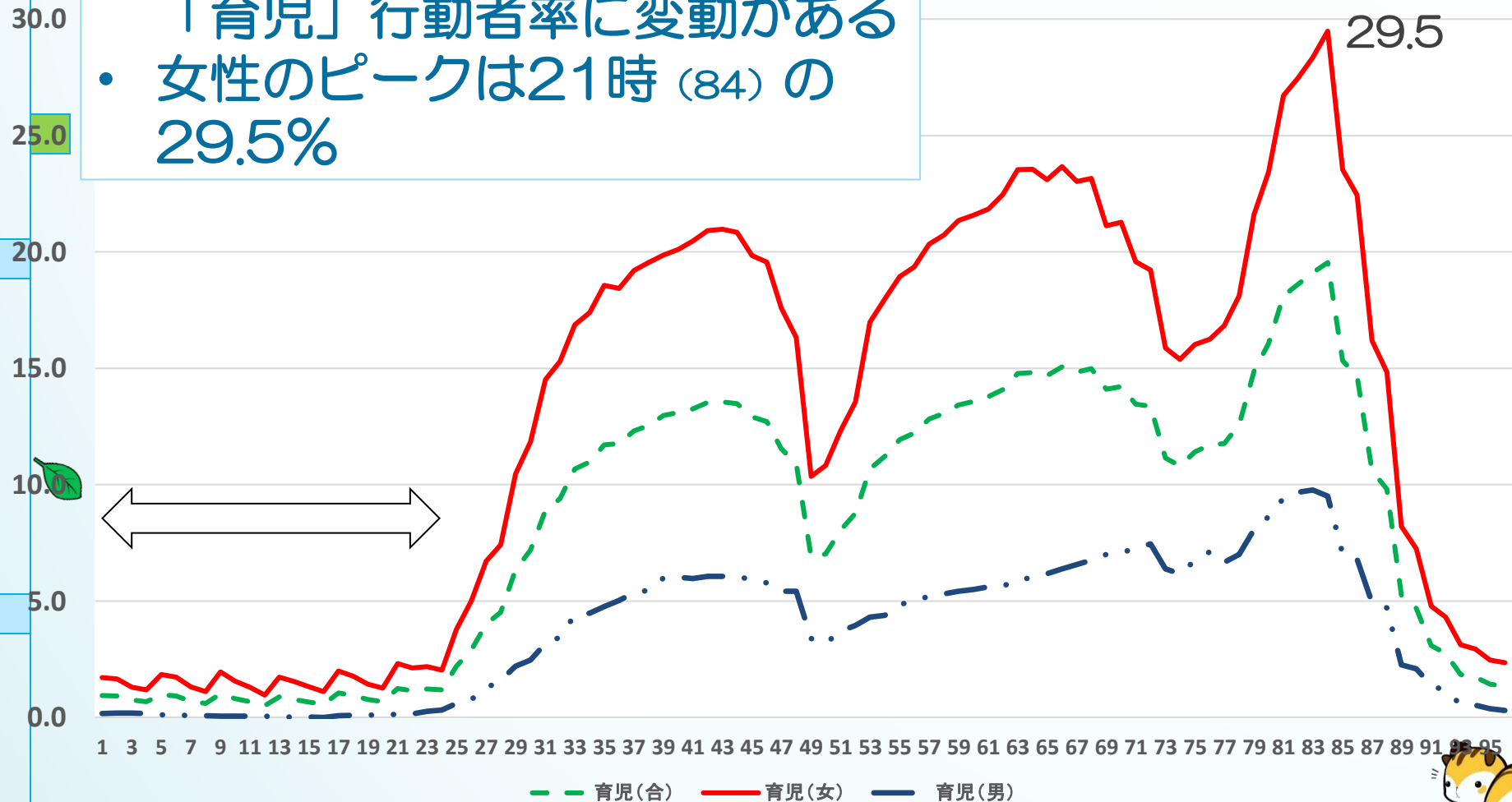
2016年社会生活基本調査 匿名データ



6歳未満の子どもがいる世帯

(育児：行動者率)

- 0時 (1) から6時 (24) の間にも「育児」行動者率に変動がある
- 女性のピークは21時 (84) の29.5%



6歳未満の子どもがいる世帯

- 総時間を用いた分析では、その負担度が示せない

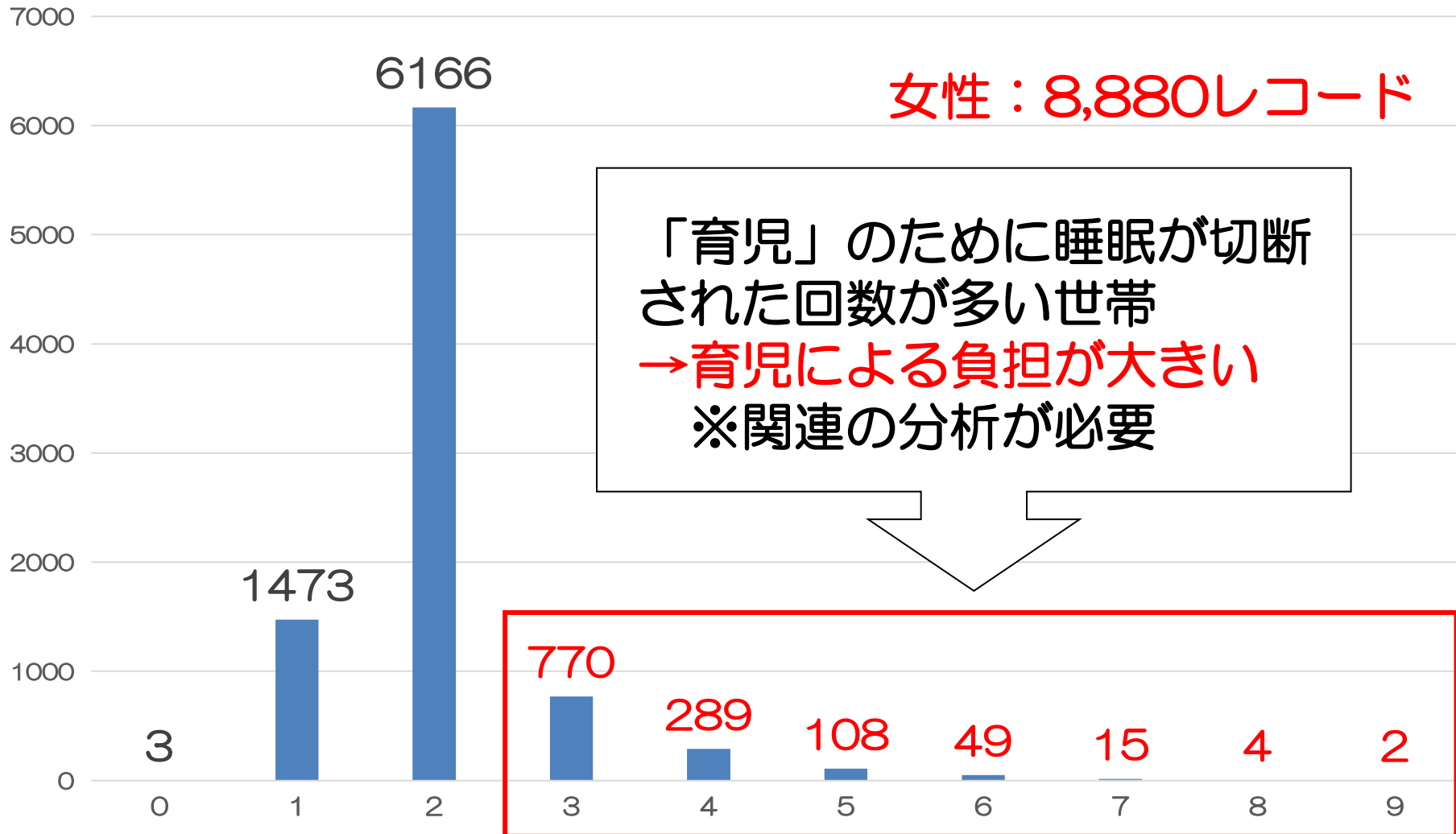
睡眠時間		育児時間	
平均	7.940	平均	3.144
標準誤差	0.018	標準誤差	0.036
中央値（メジアン）	8.000	中央値（メジアン）	2
最頻値（モード）	8.000	最頻値（モード）	0
標準偏差	1.684	標準偏差	3.415
分散	2.835	分散	11.660
尖度	3.449	尖度	0.989
歪度	0.404	歪度	1.206
範囲	23.250	範囲	18
データの個数	8,880	データの個数	8,880



6歳未満の子どもがいる世帯

2016年社会生活基本調査 匿名データ

(睡眠：切断回数)



匿名データを利用した講義

- タクミインフォメーション
テクノロジーとの協働

1. 国勢調査の概要

- 調査項目間の整合性

2. 匿名データの取得

- 申請

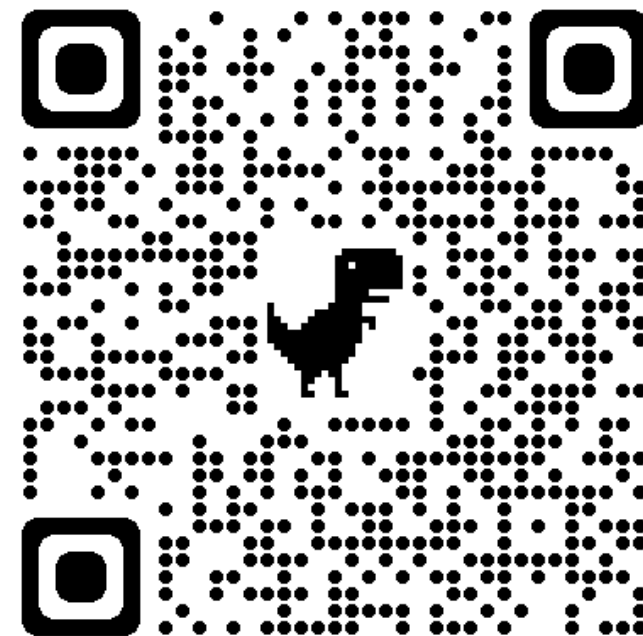
3. SASプログラム

- 起動、終了、コード作成



国勢調査の概要


<https://www.stat.go.jp/data/kokusei/2020/index.html>



1. 国勢調査の調査票

➤ 家族構成が大事

秘 基幹統計調査


 政府統計

国勢調査調査票

国勢調査は、統計法に基づき政府が実施する統計調査です。
 秘密の保護には万全を期していますので、ありのままを記入してください。


令和2年10月1日
 総務省統計局

世帯について
 (調査票が2枚以上にわたる場合は1枚目のみに記入してください)

1 世帯員の数	総数	男	女
・ふだん住んでいる人 全員の人数を書いてください	<input type="text"/> <input type="text"/> <input type="text"/> 人	<input type="text"/> <input type="text"/> <input type="text"/> 人	<input type="text"/> <input type="text"/> <input type="text"/> 人

2 住居の種類

都道府県・市区町村営の賃貸住宅	都市再生機構・公社等の賃貸住宅	民営の賃貸住宅	給与住宅(社宅・公務員住宅など)	住宅に間借り	会社等の独身寮・寄宿舎	その他
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

記入は黒の鉛筆で
 数字は右づめに 

4 人

数字の記入例

たて線1本 すきまをあける とじる

1 2 3 4 5 6 7 8 9 0

はねない 上につきぬける 角をつける

○ 黒の鉛筆で記入し、間違えた場合は、消しゴムできれいに消してください。

○ 記入欄が○の場合は、当てはまる○を●のようにぬりつぶしてください。

○ 数字を記入する場合は、わくの中に右づめで書いてください。

「調査票の記入」



調査事項間の整合性

「調査票情報」にない「氏名」がなぜ必要か？

世帯員全員について（世帯員ごとに記入してください）

3 氏名及び男女の別	1 (氏名)	2 (氏名)	3 (氏名)	4 (氏名)
・ふだん住んでいる人を もれなく書いてください	男 女 <input type="radio"/> <input type="radio"/>	男 女 <input type="radio"/> <input type="radio"/>	男 女 <input type="radio"/> <input type="radio"/>	男 女 <input type="radio"/> <input type="radio"/>
4 世帯主との続き柄	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者
・世帯主の配偶者(妻又は夫)の 祖父母・兄弟姉妹はそれぞれ 祖父母・兄弟姉妹に含めます	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>
・孫の配偶者は孫に 兄弟姉妹の 配偶者は兄弟姉妹に含めます	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他

のしかた」を参照して 黒い太わ

「調査票の記入のしかた」

世帯について
(調査票が2枚以上にわたる場合は1枚目のみに記入してください)

第1面

1 世帯員の数 総数 男 女
・ふだん住んでいる人
全員の人数を書いて
ください 4人 2人 2人

2 住居の種類
都道府県・ 都市再生機構・ 民 営 の 給与住宅 住宅に 会社等の
持ち家 市区町村営 公 社 等 の 賃貸住宅 賃貸住宅 (住宅・公営) 間借り 施設等 寄 宿 舎 その他

6ページを参照

世帯員全員について（世帯員ごとに記入してください）

3 氏名及び男女の別	1 (氏名)	2 (氏名)	3 (氏名)	4 (氏名)
・ふだん住んでいる人を もれなく書いてください	統計 秋代 男 女 <input type="radio"/> <input type="radio"/>	統計 太郎 男 女 <input type="radio"/> <input type="radio"/>	統計 花子 男 女 <input type="radio"/> <input type="radio"/>	統計 大介 男 女 <input type="radio"/> <input type="radio"/>
4 世帯主との続き柄	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者	世帯主 世帯主 又は の 子 子の 世帯主 世帯主 代表者 配偶者 配偶者 の 父母 又は の 父母 代表者 配偶者
・世帯主の配偶者(妻又は夫)の 祖父母・兄弟姉妹はそれぞれ 祖父母・兄弟姉妹に含めます	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>
・孫の配偶者は孫に 兄弟姉妹の 配偶者は兄弟姉妹に含めます	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他	孫 祖父母 兄弟 他の 住み込み 姉妹 親族 の 雇人 その他

回答者が誤って記入
の間違いをする場合、
「記入者誤差」とい
い、調査事項間の整
合性を確認する

家族属性の整合性

世帯員全員について（世帯員ごとに記入してください）

3 氏名及び男女の別	1	2	3	4
・ふだん住んでいる人を もれなく書いてください	(氏名) 男 女	(氏名) 男 女	(氏名) 男 女	(氏名) 男 女
4 世帯主との続き柄	世帯主 又は 代表者	世帯主 又は 代表者	世帯主 又は 代表者	世帯主 又は 代表者
・世帯主の配偶者（妻又は夫）の 祖父母・兄弟姉妹はそれぞれ 祖父母・兄弟姉妹に含めます ・孫の配偶者は孫に 兄弟姉妹の 配偶者は兄弟姉妹に含めます	世帯主の 配偶者の 父母 孫 祖父母 兄弟 姉妹 他 の親族 住み込み の雇人 その他	世帯主の 配偶者の 父母 孫 祖父母 兄弟 姉妹 他 の親族 住み込み の雇人 その他	世帯主の 配偶者の 父母 孫 祖父母 兄弟 姉妹 他 の親族 住み込み の雇人 その他	世帯主の 配偶者の 父母 孫 祖父母 兄弟 姉妹 他 の親族 住み込み の雇人 その他
5 出生の年月	明治 大正 昭和 平成 令和 西暦 年 月	明治 大正 昭和 平成 令和 西暦 年 月	明治 大正 昭和 平成 令和 西暦 年 月	明治 大正 昭和 平成 令和 西暦 年 月
6 配偶者の有無	未婚（幼児など を含む） 配偶者 あり 死別 離別	未婚（幼児など を含む） 配偶者 あり 死別 離別	未婚（幼児など を含む） 配偶者 あり 死別 離別	未婚（幼児など を含む） 配偶者 あり 死別 離別
7 国 籍	日本 外国 (国名)	日本 外国 (国名)	日本 外国 (国名)	日本 外国 (国名)

- 調査事項をクロスしながら、それぞれの記入内容が正しいか、チェックプログラムで確認



家族属性の整合性の確認

- 世帯の3世代構成の調査の確認方法
 1. 各世代の年齢範囲を明確に定義する必要
 2. 性別と年齢から収入を推定するためには、収入と年齢、性別の関係についての統計データが必要
 3. 収入の推定は非常に複雑
 - 地域、教育、職歴など多くの要因が影響を与えるため、正確な推定は困難



整合性チェックのための仮定

※事例であり、実際の仮定とは異なる

- 仮定を設けてシンプルな表を作成
 - 具体的には、以下のような仮定を設定
 1. 祖父、祖母は60歳以上、
 2. 父、母は30歳から59歳、
 3. 子供は0歳から29歳
 4. 収入
 - 男性の収入は女性の収入よりも高い
 - 年齢が高いほど収入が高い



整合性の表を作成

- 各世代の年齢範囲、性別、および推定収入
- 推定収入が最も高い家庭の世帯主を選定
 - 収入の推定には単純な数値を使用
 - 実際の収入とは一致しない

世代	性別	年齢範囲	推定収入 (万円/年)
祖父	男	60歳以上	400
祖母	女	60歳以上	300
父	男	30歳 - 59歳	600
母	女	30歳 - 59歳	500
子供	男	0歳 - 29歳	200
子供	女	0歳 - 29歳	150

家族構成のチェック

Pythonコードの場合

```
import pandas as pd
# データフレームの作成
data = {
    'Gender': ['M', 'F', 'M', 'F', 'M', 'F'],
    'Age': [65, 70, 40, 35, 10, 15]
}
df = pd.DataFrame(data)
# GenerationとSex列の生成
df['Generation'] = pd.cut(df['Age'], bins=[-1, 29, 59, 100],
    labels=['Child', 'Parent', 'Grandparent'], right=True)
df['Sex'] = df['Gender'].map({'M': 'Male', 'F':
    'Female'}).fillna('Unknown')
# データフレームの出力
print(df)
```



SASコード

Pythonからの変換

```
data generation;  
  input Gender $ Age;  
  datalines;  
M 65  
F 70  
M 40  
F 35  
M 10  
F 15;  
run;
```



data classify_generation;

set generation;

if Age >= 60 then Generation = 'Grandparent';
else if 30 <= Age < 60 then Generation = 'Parent';
else if 0 <= Age < 30 then Generation = 'Child';
else Generation = 'Unknown';

if Gender = 'M' then Sex = 'Male';
else if Gender = 'F' then Sex = 'Female';
else Sex = 'Unknown';

run;

proc print data=classify_generation;

run;



課題により理解度の向上

家族構成のチェックプログラムの修正

OBS	Gender	Age	Generation	Sex
1	M	65	Grandparent	Male
2	F	70	Grandparent	Fema
3	M	40	Parent	Male
4	F	35	Parent	Fema
5	M	10	Child	Male
6	F	15	Child	Fema

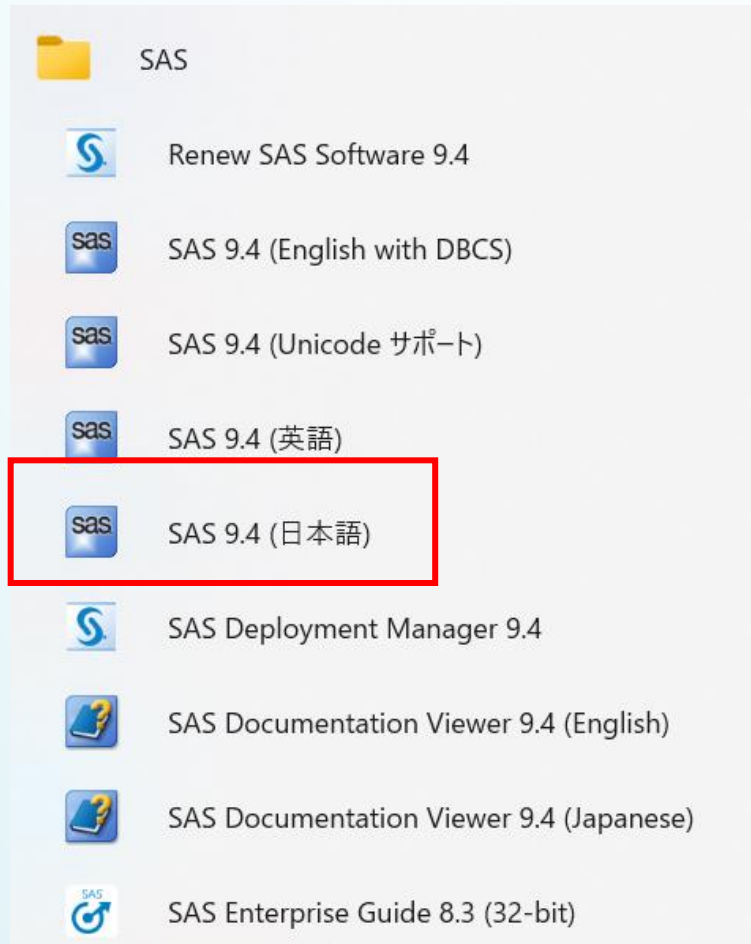


OBS	Gender	Age	Generation
1	M	65	Grandfather
2	F	70	Grandmother
3	M	40	Father
4	F	35	Mother
5	M	10	Son
6	F	15	Daughter

SASコードで説明しているが、Python、EXCEL VBA、Rでも可とする。

SAS社認定プログラム参加

個人PCにもSASインストール



指定講義を履修すること
により、SASのスキルの
証明がSAS社から貰える



学生向け「SAS認定プログラム」

➤ DS学部単位でのSASライセンスを取得し、
教育に活かす

-SAS Academic Specialization-

SAS共同認定資格プログラム「SAS Academic Specialization」とは

近年、ビックデータから価値ある情報を生み出し、様々な意思決定において、データに基づいて合理的な判断ができるデータサイエンスに係わる知識・技術（数学、統計、データ解析等）を有し、統計やデータ解析等を実施できる統計ソフトを扱うスキルを持った人材が求められています。

そこで本学は、様々な企業で導入されているSASソフトウェア（Statistical Analysis System/Software：統計解析システム/ソフトウェア）を用いて現実世界のデータ等を解析する能力を身につけることで、SASソフトウェアに関するスキルを修得できる共同認定資格プログラム「SAS Academic Specialization」を導入しました。

このプログラムでは、修了要件を満たした学生に対して、本学とSAS Institute Inc.が共同で、一定のSASスキルを修得したことを認定する認定証を交付します。

【対象学生】 全学部の学生（大学院生の参加も可）

【修了要件】 以下の条件①、条件②のどちらも満たすこと

条件①：対象の3科目の単位（計6単位）を取得

* 選択必修: 以下のいずれか1科目2単位

- 統計学実習III ... 2単位

- 統計学実習IV ... 2単位

* 選択必修: 以下から2科目4単位

- 統計学IV ... 2単位

- サービスデータサイエンス ... 2単位

- 計量マーケティング ... 2単位

条件②：SASソフトウェアを用いてデータ解析を行い「論文（SASを用いた研究課題や実践課題）」を提出し、審査に合格

SAS Studioでの実行

- GUIで散布図の作成

SAS® Studio

ファイルとフォルダ

タスクとユーティリティ



ヒートマップ

ヒストグラム

折れ線グラフ

モザイクプロット

円グラフ

散布図

系列プロット

マップ

統計量

線形モデル

ハイパフォーマンス統計量

生存時間分析

プロセスフロー 1 x SASHELP.IRIS x
設定 コード/結果 分割

データ

表示

情報

データ

SASHELP.IRIS

フィルタ: (なし)

役割

*X 軸: (1 項目)

PetalLength

*Y 軸: (1 項目)

PetalWidth

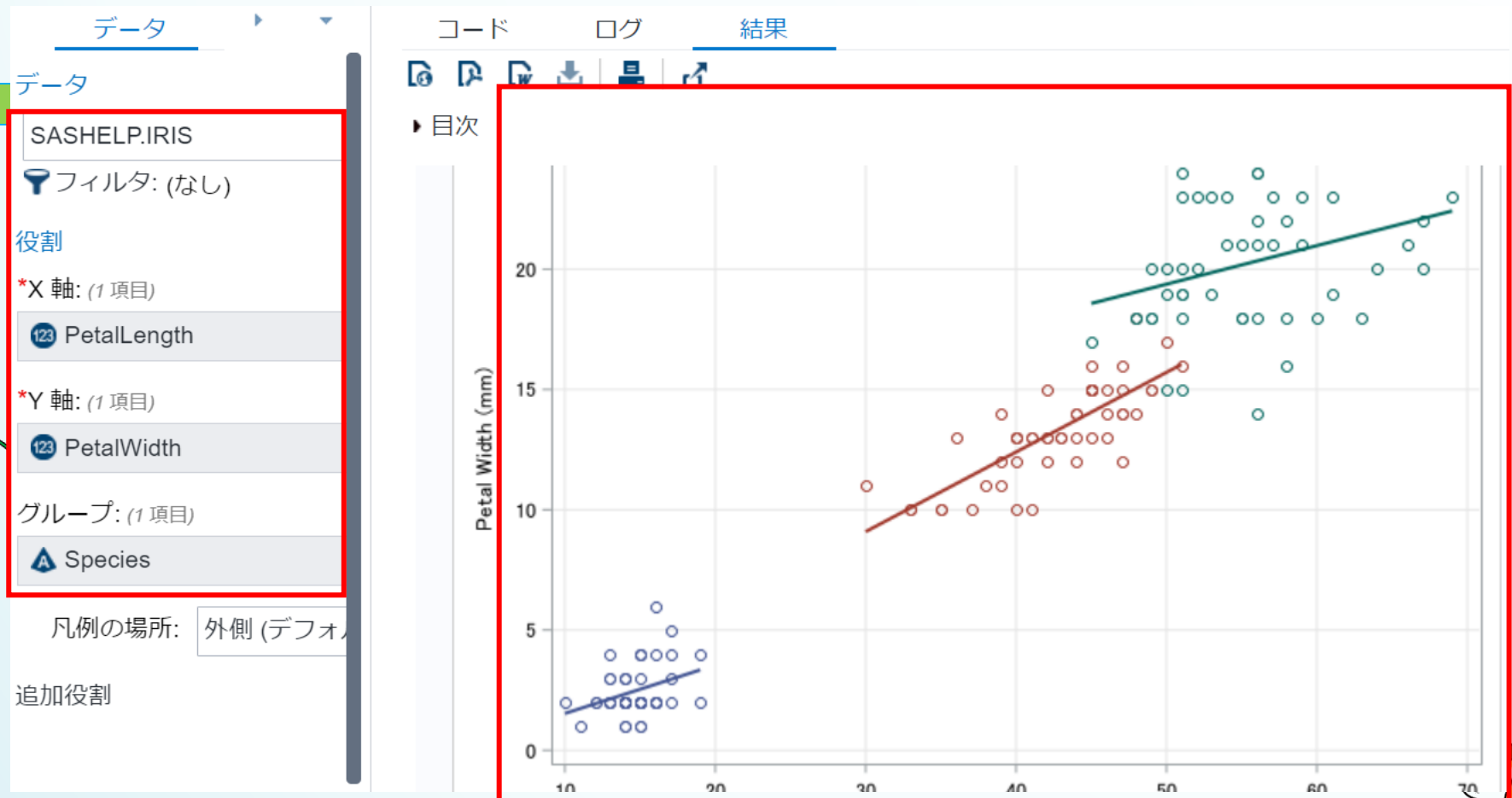
グループ: (1 項目)

Species



SAS Studioでの実行

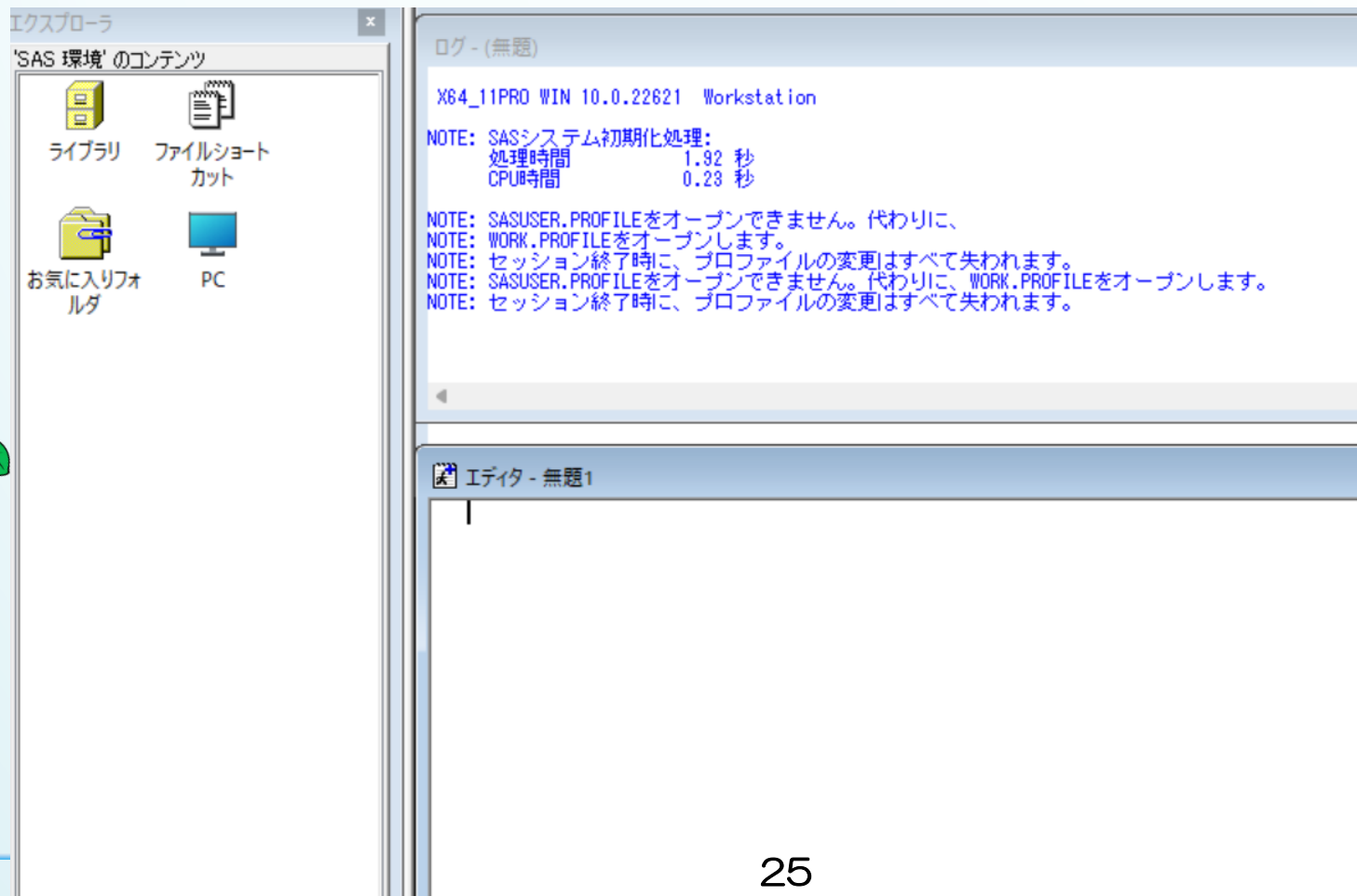
- GUIでirisの散布図



SAS 9.4での実行

起動時の画面

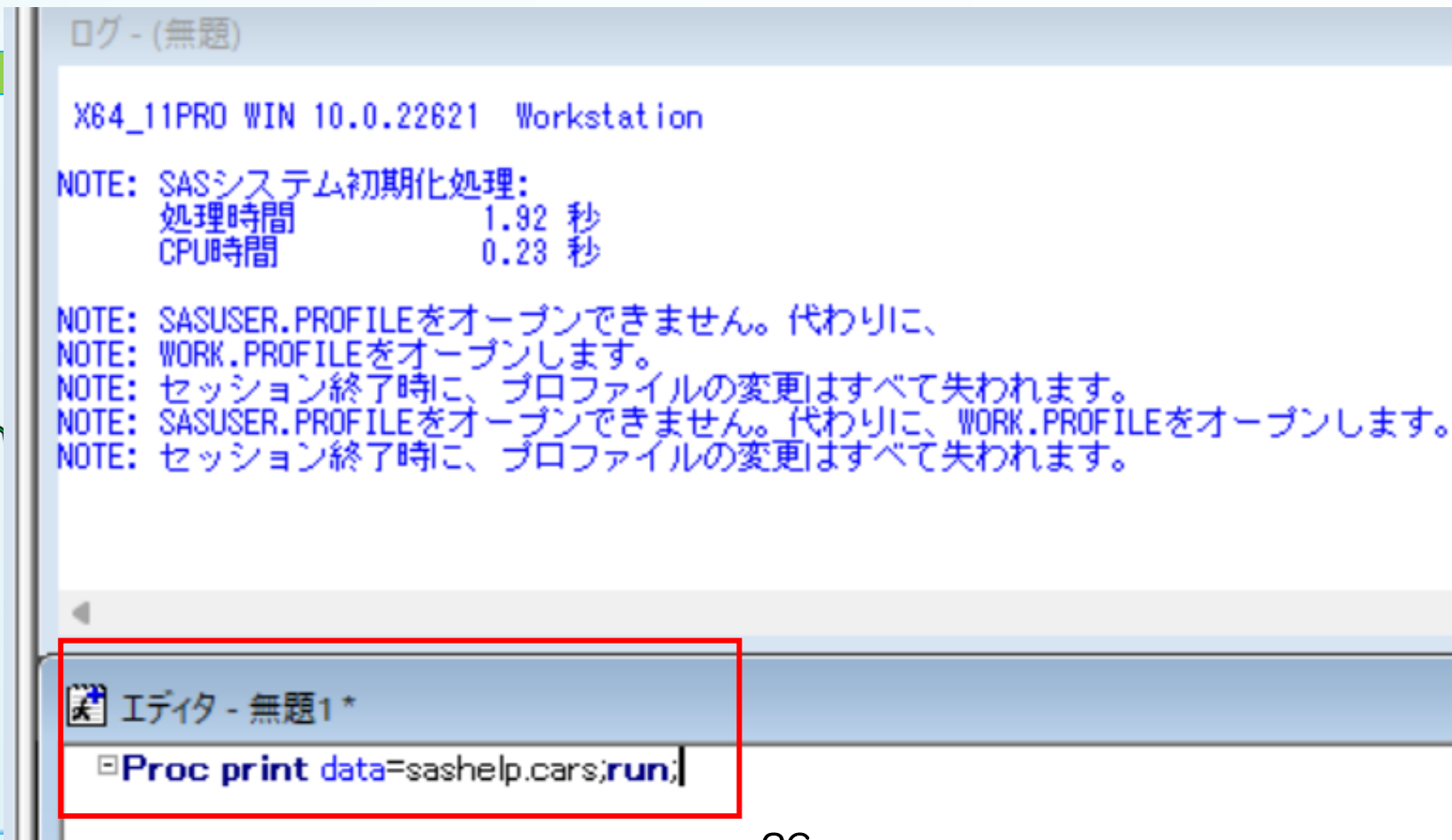
エディタの画面にコードを入力



SAS 9.4での実行

- コード入力

Proc print data=sashelp.cars;run;



The screenshot displays the SAS 9.4 Workstation interface. The top window is the 'Log - (無題)' (Log - (Untitled)) window, which contains the following text:

```
X64_11PRO WIN 10.0.22621 Workstation
```

NOTE: SASシステム初期化処理:
処理時間 1.92 秒
CPU時間 0.23 秒

NOTE: SASUSER.PROFILEをオープンできません。代わりに、
NOTE: WORK.PROFILEをオープンします。
NOTE: セッション終了時に、プロファイルの変更はすべて失われます。
NOTE: SASUSER.PROFILEをオープンできません。代わりに、WORK.PROFILEをオープンします。
NOTE: セッション終了時に、プロファイルの変更はすべて失われます。

The bottom window is the 'エディタ - 無題1*' (Editor - Untitled1*) window, which contains the SAS code:

```
Proc print data=sashelp.cars;run;
```

The editor window title bar and the code itself are highlighted with a red rectangle.



SAS 9.4での実行

- コード入力

Proc print data=sashelp.cars;run;

実行結果

結果ビューア - SAS Output

SAS システム

OBS	Make	Model	Type	Origin	DriveTrain	MSRP	Invoice	EngineSize	Cylinders	HorsePower
1	Acura	MDX	SUV	Asia	All	\$36,945	\$33,337	3.5	6	
2	Acura	RSX Type S 2dr	Sedan	Asia	Front	\$23,820	\$21,761	2.0	4	
3	Acura	TSX 4dr	Sedan	Asia	Front	\$26,990	\$24,647	2.4	4	
4	Acura	TL 4dr	Sedan	Asia	Front	\$33,195	\$30,299	3.2	6	
5	Acura	3.5 RL 4dr	Sedan	Asia	Front	\$43,755	\$39,014	3.5	6	
6	Acura	3.5 RL w/Navigation 4dr	Sedan	Asia	Front	\$46,100	\$41,100	3.5	6	

エディタ - 無題1*

Proc print data=sashelp.cars;run;



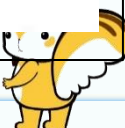
SAS 9.4での実行

- irisのデータを使って散布図の作成

```
ods graphics / reset width=6.4in height=4.8in  
imagemap;
```

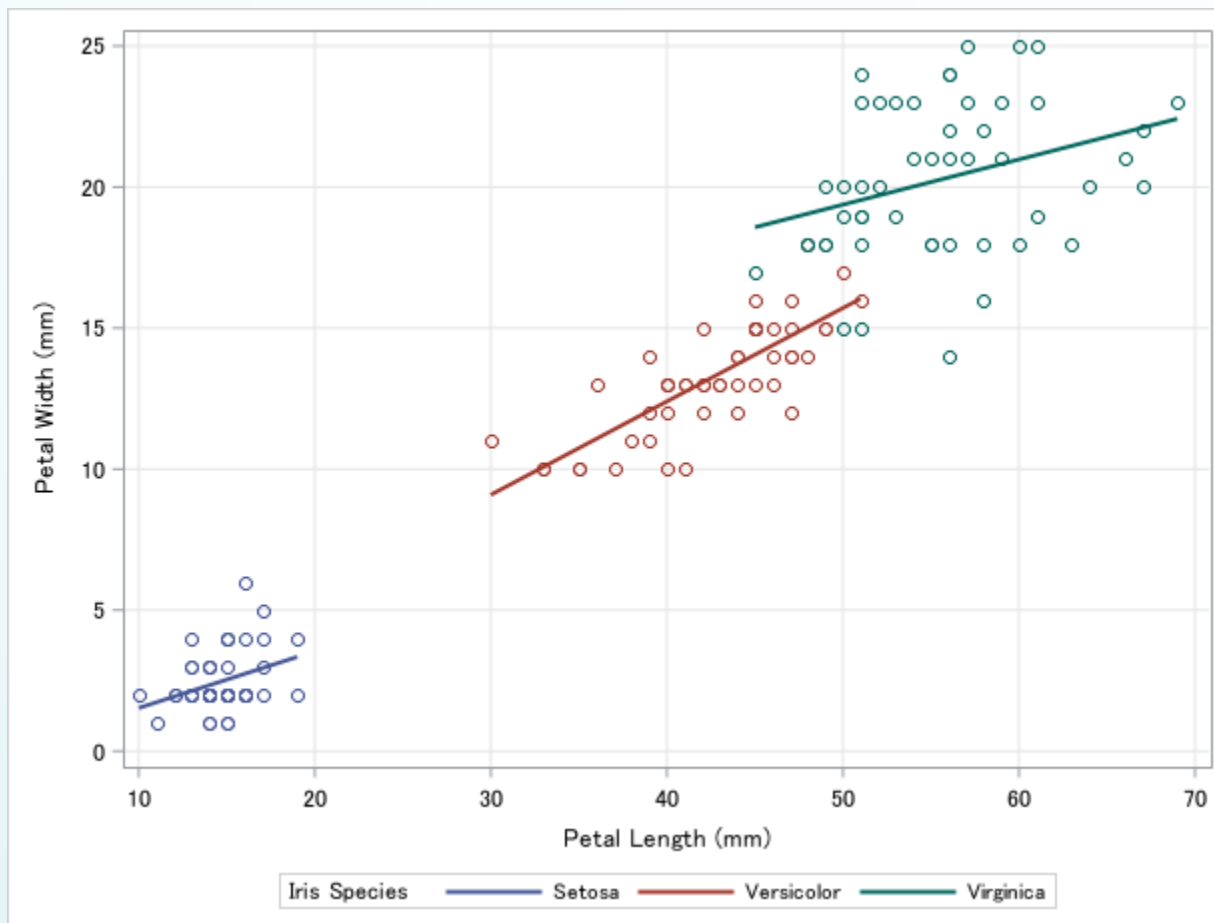
```
proc sgplot data=SASHELP.IRIS;  
    reg x=PetalLength y=PetalWidth / nomarkers  
    group=Species;  
    scatter x=PetalLength y=PetalWidth /  
    group=Species;  
    xaxis grid;  
    yaxis grid;  
run;  
ods graphics / reset;
```

- SepalLength : アヤメのがくの長さ (単位: mm)
- SepalWidth : アヤメのがくの幅 (単位: mm)
- PetalLength : アヤメの花弁の長さ (単位: mm)
- PetalWidth : アヤメの花弁の幅 (単位: mm)



irisの散布図

reg x=PetalLength y=PetalWidth / nomarkers
group=Species;



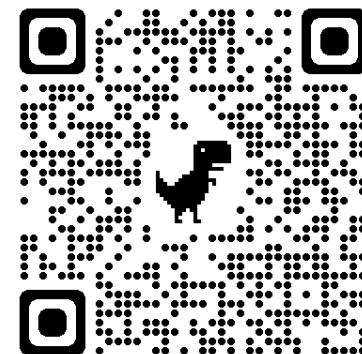
匿名データの利用開始

- 独立行政法人統計センターのサイトから資料を取得

国勢調査（最終更新日：2020/3/31）

総務省統計局のホームページにおいて、「[匿名データの作成・提供及びオーダーメイド集計](#)」が掲載されています。ご参考にしてください。

年次	提供年次数	調査の概要	データレイアウト・符号表	レコード数	地域区分	【参考】項目名一覧
平成12年 (2000年)	4年分	調査ページ (総務省統計局)	Excel:114KB	約124万レコード	・都道府県別 ・人口50万以上の各市区とその他の別	CSV:2KB
平成17年 (2005年)		調査ページ (総務省統計局)	Excel:105KB	約124万レコード		CSV:2KB



匿名データの取得（申請等）

使ってみよう！ミクロデータ

匿名データについて

- 1.匿名データとはどのようなものでしょうか？
- 2.現在、統計センターで利用できる匿名データはどの統計調査のものでしょうか？
- 3.匿名データは、どのようなファイル構成となっていますか？
- 4.匿名データは具体的にどのような加工（匿名化措置）がされているのでしょうか？
- 5.匿名データのデータ形式等を教えてください。
- 6.匿名データを利用するに当たり、参照すべき資料はありますか？
- 7.符号表に「△」などの2バイト文字がありましたが、データ内に2バイト文字があるのでしょうか？

匿名データの利用上の注意

- 毎回、USBスティックを受け取り、匿名データを利用した分析を開始
- 講義終了時、USBスティックを返却

USBスティックの保存内容

2010kokucho	2023/11/25 9:25	ファイル フォルダー
2011shakai_B	2023/11/25 9:26	ファイル フォルダー
2011shakai_J	2023/11/25 16:51	ファイル フォルダー
2011shakai_S	2023/11/25 17:24	ファイル フォルダー
2015kokucho	2023/11/25 2:15	ファイル フォルダー
2016shakai_B	2023/11/25 17:23	ファイル フォルダー
2016shakai_J	2023/11/25 17:08	ファイル フォルダー
2016shakai_S	2023/11/25 17:23	ファイル フォルダー
集計結果及び公表値等資料	2023/11/24 17:59	ファイル フォルダー
 (別添 2) データ提供時配布資料.doc	2020/02/04 14:29	Microsoft Word 97-2...



匿名データを利用した講義の内容

詳細：

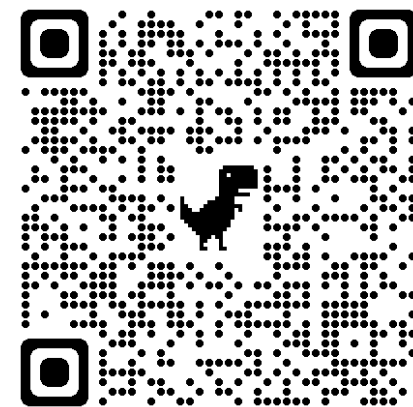
「公的統計匿名データを利用したデータ

サイエンス講義のための取組み」

データサイエンス研究 創刊号，

2022 年3 月 63 頁～77 頁

を参照



なお、この時点では「R」を利用したプログラム作成を計画していたが、企業連携により「SAS」に変更している。



SAS、Python以外の集計方法

- EXCELにおける集計の限界（全数読み込み不可）

	男	女	総数
総数	509,798	538,777	1,048,575

- Powerpivotでそれを打破（全数読み込み集計）

A	B	C	D	E
	41			
	999	3830	4304	8,134
	42			
	999	6278	7175	13,453
	43			
	100	3345	3777	7,122
	999	4727	5336	10,063
	44			
	999	5380	6044	11,424
	45			
	999	5106	5681	10,787
	46			
	201	2773	3139	5,912
	999	4817	5327	10,144
	47			
	999	6745	7072	13,817
	総計	605432	640362	1,245,794

ピボットテーブルのフィールド

アクティブ すべて

レポートに追加するフィールドを選択してください:

検索

2015kokucho

☒ F1

☐ F2

☐ F3

☒ F4

☒ F5

☐ F6

☐ F7

☐ F8

☐ F9

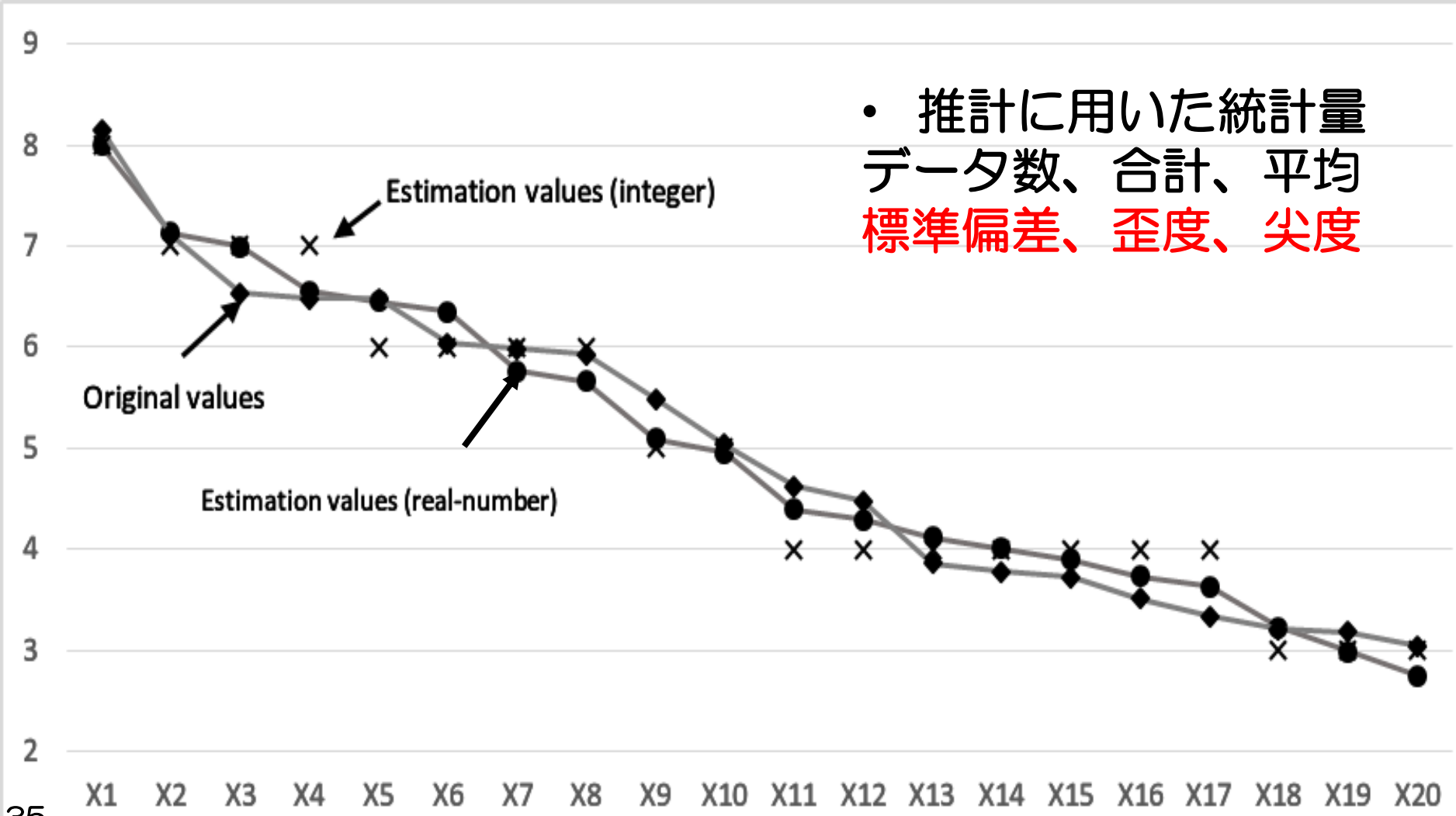
☐ F10

☐ F11

匿名データ利用時のリスク管理

- 基本統計量を公開することのリスク

➤ 推計値(実数)はMS-EXCELソルバー機能による詳細推計



- 推計に用いた統計量
データ数、合計、平均
標準偏差、歪度、尖度

元データを推計するためのDB

$$x_1 + x_2 + \dots + x_{99} + x_{100} = 100$$

$$x_{100} \leq x_{99} \leq \dots \leq x_2 \leq x_1$$

すべての数値パターンを作成：約1.9億パターン → **0が多い行列（78.2%）**

度数	変数1	変数2	変数3	変数4	...	変数98	変数99	変数100	合計	平均	SD
1	100	0	0	0	...	0	0	0	100	-	-
2	99	1	0	0	...	0	0	0	100	50	69.3
2	98	2	0	0	...	0	0	0	100	50	67.9
3	98	1	1	0	...	0	0	0	100	33.3	56.0
.
.
98	2	2	1	1	...	1	0	0	100	1.02	0.14
98	3	1	1	1	...	1	0	0	100	1.02	0.20
99	2	1	1	1	...	1	1	0	100	1.01	0.10
100	1	1	1	1	...	1	1	1	100	1.00	0.00



度数別にテーブルを作成：
拡張性がなく検索が非効率



効果的なデータの格納、抽出と拡張性の向上

0以外のセルを「行」と「列」と「数値」に変換

Table1			Table2			Table3		
行	列	値	行	度数	平均	行	度数	SD
1	1	100	1	2	50.0	1	2	70.7
2	1	99	1	3	33.3	1	3	57.7
3	1	98	1	4	25.0	1	4	50.0
3	2	2	1	5	20.0	1	5	44.7
4	1	98	1	6	16.7	1	6	40.8
4	2	1	1	7	14.3	1	7	37.8
4	3	1	1	8	12.5	1	8	35.4
.
.
.



拡張性や検索速度
が向上

スパース行列から0を除去した
数値群（MM：Market Matrix形式）

度数20の数値パターン抽出結果

➤ 度数20の元データ

	X_1	X_2	標準偏差	歪度	尖度
元データ	8.156	7.116	1.531	0.346	-1.004

➤ 度数20のテーブルから近似データを抽出

	X_1	X_2	抽出パターン数	標準偏差	歪度	尖度
組合せ1	8	8	14	1.522	-0.0369 - 1.195	0.398 - 1.653
組合せ2	8	7	8			

➤ DBから抽出した推計に用いるデータ

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
8	7	7	7	6	6	6	6	5	5
X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}
4	4	4	4	4	4	4	3	3	3
標準偏差	1.521771821			歪度	0.367745		尖度	-1.00681	

講義での演習

4値のデータで散らばりを理解する

1. 100値のDBを基に4値のデータベースを作成
 2. $0 \leq X_4 \leq X_3 \leq X_2 \leq X_1 \leq 100$
- 基本統計量作成（拡張性のあるテーブル設計）
2値、3値、4値別に作成

X ₁	X ₂	X ₃	X ₄	Total	average
100	0	0	0	100	25
99	1	0	0	100	25
98	2	0	0	100	25
98	1	1	0	100	25
97	3	0	0	100	25
97	2	1	0	100	25
...	100	25
25	25	25	25	100	25

Pythonコード

```
import pandas as pd
patterns = []
# x1 from 100 to 25
for x1 in range(100, 24, -1):
    # x2 from 0 to x1
    for x2 in range(min(x1+1, 101-x1)):
        # x3 from 0 to x2
        for x3 in range(min(x2+1, 101-x1-x2)):
            # x4 from 0 to x3
            for x4 in range(min(x3+1, 101-x1-x2-x3)):
                total = x1 + x2 + x3 + x4
                # Ensure total is 100
                if total == 100:
                    patterns.append([x1, x2, x3, x4])
# Convert list of patterns to pandas DataFrame
df = pd.DataFrame(patterns, columns=['x1', 'x2', 'x3', 'x4'])

# Save DataFrame to Excel file
df.to_excel(r'f:\Py_pattern4.xlsx', index=False)

print(f"The end !")
```



2値から4値への拡張

- 2値、3値、4値の合計がそれぞれ100の場合、平均、分散、標準偏差を計算（赤字に注目）
- ※このプログラムもChatGPTが対応（黄色い部分）

X1	X2	X3	X4	Total	Mean2	Mean3	Mean4	VAR2	VAR3	VAR4	SD2	SD3	SD4
100	0	0	0	100	50	33.3	25	2500	2222.2	1,875.0	50	47.1	43.3
99	1	0	0	100	50	33.3	25	2401	2156.2	1,825.5	49	46.4	42.7
98	2	0	0	100	50	33.3	25	2304	2091.5	1,777.0	48	45.7	42.1
97	3	0	0	100	50	33.3	25	2209	2028.2	1,729.5	47	45.0	41.5
97	3	0	0	100	50	33.3	25	2209	2028.2	1,729.5	47	45.0	41.5
96	4	0	0	100	50	33.3	25	2116	1966.2	1,683.0	46	44.3	41.0
98	1	1	0	100	0	33.3	25	0	2090.8	1,776.5	0	45.7	42.1
97	2	1	0	100	0	33.3	25	0	2026.8	1,728.5	0	45.0	41.5
96	3	1	0	100	0	33.3	25	0	1964.2	1,681.5	0	44.3	41.0
96	2	2	0	100	0	33.3	25	0	1963.5	1,681.0	0	44.3	41.0
96	2	1	1	100	0	0	25	0	1984.6	1,680.5	0	0	40.9
26	25	25	24	100	0	0	25	0	0	0.5	0	0	0.7
25	25	25	25	100	0	0	25	0	0	0	0	0	0

ACCESSを使った近似値の検索

ACCESSを使って、度数、標準偏差、歪度、尖度の値を指定した検索



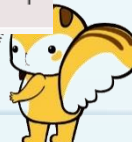
度数:	<input type="text" value="4"/>
標準偏差:	<input type="text" value="4.2172772"/>
歪度:	<input type="text" value="0.1023652"/>
尖度:	<input type="text" value="-1.137644"/>



EXCELソルバーでの推計

EXCELのソルバーやシナリオを使って分析

A	B	C	D	E	F	G	H	I	J	K
活用事例（推計には標準偏差、歪度、尖度を利用）										
		推計に用いる4つの値								
		シナリオ	X_1	X_2	X_3	X_4	度数	合計	平均	
元データ	実値	3	30,967	20,934	16,899	2,013	4	70,813	17,703	
合計を 100に 変換	百分比	3	43.731	29.56	23.86	2.843	4	100	25	
			X_1	X_2	X_3	X_4	度数	合計	平均	
	推計	3	30.84	26.7	22.9	19.5	4	100	25	
<div>シナリオ番号の選択</div> <div>ソルバーの実行</div>										
>	活用事例	T_SD_Skew_Kurt_R	T_SD_Skew_Kurt_R2	T_SD_Skew_Kurt_R3	Scenario	+				



連携企業

- プロジェクト実施による企業との連携
- 1. 株式会社タクミイノベーションテクノロジー
 - 生活時間調査データの分析
 - SASを使ったデータ分析の講義の実施
 - 機会学習ができるSAS VIYA等
- 2. 株式会社プラージュ
 - 生活行動データ可視化アプリ
 - 「Life Trackerの開発（デモ可能）」
 - データ取得システムの構築（開発中）



Life Trackerによる多種多様な行動分析

アプリの開発：株式会社プラージュ

- 社会生活基本調査の仕様から大幅な機能拡張
 1. 生活時間調査20の行動 → 行動対象を拡大中
 2. 24時間96コマ → 1440コマ（1分刻み）
 3. 夫婦のみの世帯 → 夫婦と子供2人
 4. アバターと行動リストを連結
 - 喜怒哀楽を追加（全年齢対象）
 - 1. 日本語 → 英語を追加
 - 2. 2つの行動比較（2画面）
 - 夫や妻からの視点からの行動分析



行動データの作成 (EXCEL)

連続した行動：まとめて入力が可能（日英可）
EXCELでデータ作成：多種多様なパターンを
作成、比較できる

You can press the button after entering the daily data for one person.

ActData OutPut

【Prerequisites for households surveyed : 世帯の前提条件】

Year (調査年)	2023
Family composition1 (家族構成)	4
Family composition2	4
City (居住している都市)	Tokyo

【Action Entry Requirements : 行動の入力条件】

Husband/Wife/Child (夫/妻/子)	Husband
Age (年齢)	36-45
Day of the week (週)	Weekday

check	Time (HH:MM)		Activity(行動)	
	Beginning	End	English	OR Japanese
1	0:00	7:00	Sleeping in the nursery	
2	7:00	8:00	Personal hygiene	子供の世話(怒)
3	8:00	9:00	Meals (at home)	睡眠(子供部屋)
4	9:00	20:00	working(sad) for test	睡眠 (親の部屋)
5	20:00	21:00	Dining at work	入浴
6	21:00	21:51	Meals (at home)	洗顔・手洗い・歯磨き・身の回りの
7	21:51	22:30	Bathing	家の仕事の合間の休憩
8	22:30	23:59	Sleeping with parents	仕事中の休憩
×				学校での休憩
×				朝まで仕事 (ワーク/リフレッシュ)



行動データ管理表 (EXCEL)

行動を管理

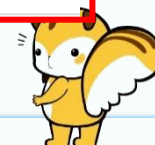
A	B	C	D	E	F
Code	Action Category	行動カテゴリ(JP)	display(EN)	display(JP)	
R001	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Nursing and caregiving	子供の世話(怒)	
R002	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Sleeping in the nursery	睡眠(子供部屋)	
R003	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Sleeping with parents	睡眠(親の部屋)	
R004	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Bathing	入浴	
R005	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Personal hygiene	洗顔・手洗い・歯磨き・身の回りの	
R006	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Work breaks(at home)	家の仕事の合間の休憩	
R007	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Work breaks	仕事中の休憩	
R008	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Breaks in school institutions	学校での休憩	
R009	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Rest and relax,Short naps	部屋で休養、くつろぎ(リビング)	
R010	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Rest and relax (at bedroom)	休養、くつろぎ(寝室)	
R011	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Outpatient and hospitalization	通院・受診	
R012	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Medical treatment(illness)	病気での療養	
R013	Sleep & Rest & Recuperate	睡眠・休憩・療養・身辺ケア	Other medical treatment	その他の療養	
T001	during transit	移動系	Commuting to school	通学・通園	
T002	during transit	移動系	Commuting to work	通勤(出勤)	
T003	during transit	移動系	Returning home	通勤(帰宅中)	
T004	during transit	移動系	Returning home	通学(帰宅中)	
T005	during transit	移動系	Pick up and drop off of	子供の送迎	
T006	during transit	移動系	Driving to work	車での通勤	
T007	during transit	移動系	Travel by public transport	公共交通機関での移動	
H001	Hobbies & Recreation	趣味・娯楽	Creative activities	その他の創作活動	
H002	Hobbies & Recreation	趣味・娯楽	Making sweets (at home)	菓子作り	
H003	Hobbies & Recreation	趣味・娯楽	Participating in competitions	コンテスト参加のための活動	
H004	Hobbies & Recreation	趣味・娯楽	Sewing as a hobby	趣味での裁縫	
H005	Hobbies & Recreation	趣味・娯楽	Use computers as a hobby	趣味でパソコンを使用	
H006	Hobbies & Recreation	趣味・娯楽	Computer games	ゲーム(PC、ゲーム専用機)	
H007	Hobbies & Recreation	趣味・娯楽	Playing table games	テーブルゲーム	
H008	Hobbies & Recreation	趣味・娯楽	pissed at me for playing too	ゲームしながら怒られ	
H009	Hobbies & Recreation	趣味・娯楽	Gardening	園芸(趣味)	
H010	Hobbies & Recreation	趣味・娯楽	Driving (fun)	ドライブ(趣味)	
H011	Hobbies & Recreation	趣味・娯楽	General indoor leisure	一般屋内レジャー	
H012	Hobbies & Recreation	趣味・娯楽	Art or music performance	芸術活動、音楽演奏	



EXCELファイルでのデータ作成（改訂後）

- 夫と妻の属性と1440コマの行動を登録
- 2つのファイル（夫の視点、妻の視点）で同時に表示

	1	2	3	4	5	6	7	8	9
1	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	平日	夫	東京都	睡眠	睡眠
2	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	土曜	夫	東京都	睡眠	睡眠
3	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	日曜	夫	東京都	睡眠	睡眠
4	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	平日	妻	東京都	睡眠	睡眠
5	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	土曜	妻	東京都	睡眠	睡眠
6	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	日曜	妻	東京都	睡眠	睡眠
7	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	平日	子供1	東京都	0時から1分刻み	
8	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	土曜	子供1	東京都		
9	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	日曜	子供1	東京都	子供の睡眠	子供の睡眠
10	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	平日	子供2	東京都	親との睡眠	親との睡眠
11	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	土曜	子供2	東京都	親との睡眠	親との睡眠
12	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	日曜	子供2	東京都	親との睡眠	親との睡眠



EXCELファイルでのデータ作成（改訂後）

- 夫の視点は、夫が入力しているので、喜怒哀楽を含む細かい行動入力が可能

	1	2	3	4	5	6	7	8	472	473	474
1	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	平日	夫	東京都	睡眠	通勤・出勤	ゴミ捨て	ゴミ捨て
2	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	土曜	夫	東京都	睡眠	睡眠	睡眠	睡眠
3	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	日曜	夫	東京都	睡眠	通勤・出勤	通勤・出勤	ゴミ捨て
4	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	平日	妻	東京都	睡眠	家事	家事	家事
5	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	土曜	妻	東京都	睡眠	睡眠	睡眠	睡眠
6	2020	夫婦と子供二人の世帯	夫からの視点	25~35歳	日曜	妻	東京都	睡眠	家事	家事	家事
7	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	平日	子供1	東京都	子供の睡眠	身の回りの	身の回りの	身の回りの
8	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	土曜	子供1	東京都	子供の睡眠	子供の睡眠	子供の睡眠	子供の睡眠
9	2020	夫婦と子供二人の世帯	夫からの視点	6~10歳	日曜	子供1	東京都	子供の睡眠	身の回りの	身の回りの	身の回りの
10	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	平日	子供2	東京都	親との睡眠	休養・くつ	休養・くつ	休養・くつ
11	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	土曜	子供2	東京都	親との睡眠	親との睡眠	親との睡眠	親との睡眠
12	2020	夫婦と子供二人の世帯	夫からの視点	0~5歳	日曜	子供2	東京都	親との睡眠	休養・くつ	休養・くつ	休養・くつ



夫婦間の視点の相違 1

- 1分刻みにしたことで、夫の行動に「ごみ捨て」が表示可能

2020 東京都

夫婦と子供二人の世帯

夫からの視点

平日 ▾

07:29~通勤・出勤

07:30~家事

ごみ捨て

家事

身回りの用事

休養・くつろぎ(乳幼児)

夫: 25-35歳 妻: 25-35歳 子供1: 6-10歳 子供2: 0-5歳

07:48:18

2020 東京都

夫婦と子供二人の世帯

妻からの視点

07:30~通勤・出勤

07:30~洗濯

通勤・出勤

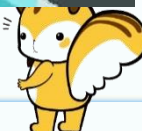
洗濯

身回りの用事

休養・くつろぎ(乳幼児)

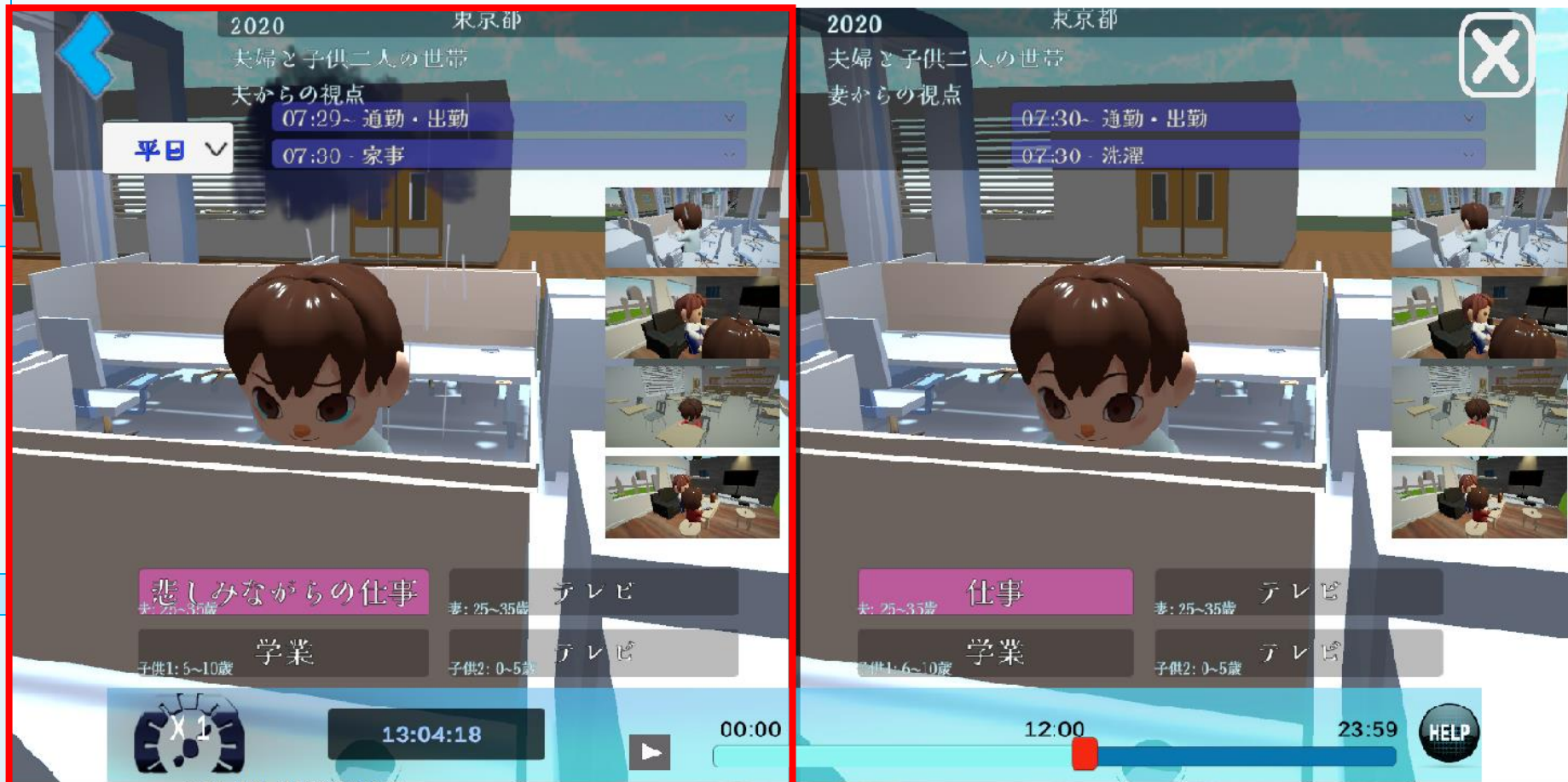
夫: 25-35歳 妻: 25-35歳 子供1: 6-10歳 子供2: 0-5歳

00:00 12:00 23:59



夫婦間の視点の相違 2

- 喜怒哀楽は本人のみが入力できる情報
- 誰もが楽しめるデータの可視化



夫婦間の視点の相違 3

- 悲しみから怒りへの気持ちの変化を表現
- 表現の追加は可能



最新の日本語・英語切り替えバージョン







2023 Gotanda Tokyo

Household with a couple and two children
Wife has a second job; household income is 8 million yen

Weekday ✓

12:41~ Main job(outside of the home) ✓
12:00~ General sports or exercise ✓

Main job(outside of the home)
Husband: 25-35

Education in school institutions
Child1: 6-10

General sports or exercise
Wife: 25-35

Communication via PC, Using SNS
Child2: 0-5

X1





14:03:53

2023 Gotanda Tokyo

Household with a couple and two children
Wife has a second job; household income is 8 million yen

English ✓

12:41~ Main job(outside of the home) ✓
12:00~ General sports or exercise ✓

Main job(outside of the home)
Husband: 25-35

Education in school institutions
Child1: 6-10

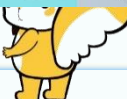
General sports or exercise
Wife: 25-35

Communication via PC, Using SNS
Child2: 0-5

12:00

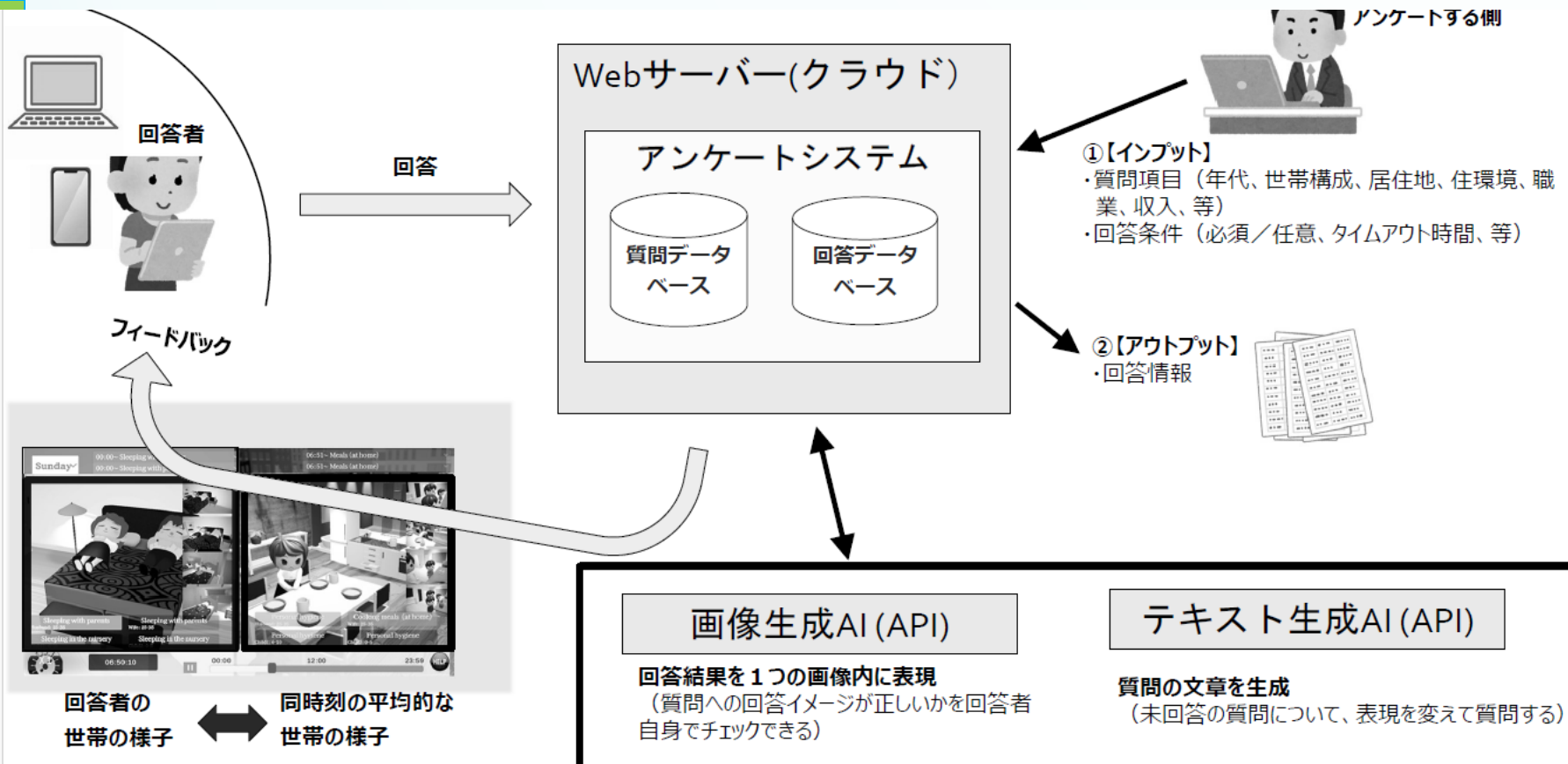
23:59

HELP



今後の展開のイメージ


1. メタバーズやアバターの活用
2. マルチリンガルに対応
3. グローバルな領域での調査が可能




画面イメージ（アンケート画面）

デモ画面（開発中）


WEBアンケート (DEMO)

 えーと初めまして、あなたの名前を教えてくださいな！
(Hi there! May I know your name, please?)

我只懂中文

 知道了
請告訴我你的名字


ラーメン丸です

 ラーメン丸さんですね
ステキな名前ですね
やはりお若いんですよね？

はい いいえ


我只懂中文

WEBアンケート (DEMO)


 今、どの辺りにお住まいですか？

北海道 東北 ...

沖縄 ?

 北海道ですね 今は寒そうですね
市区町村まで教えてもらえますか？

no

 I got it ! By the way, how many people are living here now?

1 2 3

4 ... ?

no



企業連携により期待される成果

1. データ分析の経験を得る
 - 企業との連携
2. システムはデータ分析のためのツール
3. SASのスキル習得による差別化
 - SAS VIYAでの機械学習



データ分析を核とした起業を実現



【その他の企業融合】 スポーツデータサイエンス次世代プリンター育成

オリンピックの男女100m走において、日本人のメダリストがない。個々の体格の格差を補うため、これまでの走法をスポーツデータサイエンスに基づき、日本人に最適な走法を導出し、この種目のメダリストを育成する。

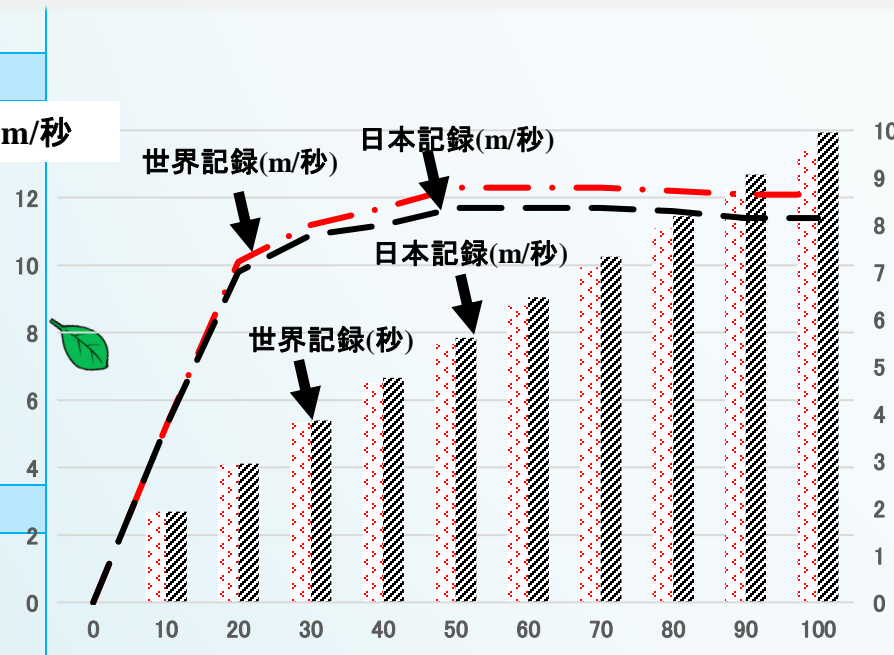


図 世界記録と日本記録における区間ごとの秒速と区間タイム

＜スポーツデータサイエンス＞
 体格的に勝る外国人に勝つためには、加速域から最高疾走時までの区間を縮め、最高疾走時の区間を伸ばす走法をデータに基づき解決する

局面	区間	状況
1次加速	スタートから20mの区間	クラウチングスタートの姿勢から、上体を徐々に起こし、ピッチを速めながら急速に加速
2次加速	20～50mの区間	ストライドを増加させながら、最大疾走速度に到達
最大疾走	50～80mの区間	ストライドと姿勢を保ちながらスピードを維持
減速	80～100mの区間	最大疾走のストライドと姿勢を保てず減速

表 100m走における4つの局面

オプトジャンプネクストでの計測

立正大学では、10mのオプトジャンプネクストを使って、100m走の科学的なトレーニングを実施中



デモンストレーション

1. スクリーンショットでは表現できないところを、「Life Tracker」で可視化
 - ✓ 必要に応じて、動画ファイルを作成
2. 今後の展開
 - 生成AIによるデータの取得

※時間次第で、デモを実施予定





立正大学

「モラリスト×エキスパート」を育む。

ご清聴ありがとうございました。

質問等がございましたら、以下のメールアドレスに連絡願います。

shirakawa.kiyomi@rissho-univ.jp

