## 情報統計第1回

2022年8月2日 神奈川工科大学



#### 櫻井 望

国立遺伝学研究所 生命情報・DDBJセンター

### スケジュール

	2日(火) データの 見える化	3日(水) 検定の これだけは	4日(木) 分散分析と多変 量解析の雰囲気	5日(金) データ準備 発表会
1限				13 補足自習(課題、質問)
2限	<ul><li>1 ガイダンス</li><li>PC環境準備、</li><li>データの見える化</li></ul>	5 区間推定、分布とその使い方	9 分布の仲間と、分散分析	14 自習(課題、質 問)
3限	2 統計の基本と用語	6 t検定	10 相関、主成分分析	15 発表会
4限	3 プログラミングの基礎	7 検定で注意すること	11 他の多変量解 析	
5限	4 自習(課題検討、復習)	8 自習(課題検討、復習)	12 自習(課題検討、復習)	

## 授業で使うサイト

https://github.com/nsaku/kait2022/wiki

## 要パスワードサイト ID: PW:

## ガイダンス

## 統計の知識は



## 大学で

- ・実験の計画
- ・データの評価

## 仕事で

管理栄養士さん

- 調査研究等の設計
- ・データの評価
- 品質管理·問題解決

## 生活の中で

- 正しい情報を見抜く
- 話に説得力が出る

## 学習すること

- データの見える化
- 検定のこれだけは
- 分散分析
- ・多変量解析の雰囲気

講義のメッセージ

# 統計情報をうのみにしない

そのための力を身につけよう

### スケジュール

	2日(火) データの 見える化	3日(水) 検定の これだけは	4日(木) 分散分析と多変 量解析の雰囲気	5日(金) データ準備 発表会
1限				13 補足自習(課題、質問)
2限	<ul><li>1 ガイダンス</li><li>PC環境準備、</li><li>データの見える化</li></ul>	5 区間推定、分布とその使い方	9 分布の仲間と、分散分析	14 自習(課題、質 問)
3限	2 統計の基本と用語	6 t検定	10 相関、主成分分析	15 発表会
4限	3 プログラミングの基礎	7 検定で注意すること	11 他の多変量解 析	
5限	4 自習(課題検討、復習)	8 自習(課題検討、復習)	12 自習(課題検討、復習)	

## 課題のやり方

- 1班 1~5人くらい
- データを集めて解析
- 発表会

## 課題内容

- 1. 統計データを公開しているサイト(厚生労働省など)から、データを取得して、
- 2. 統計結果を公開しているサイト(都道府県ランキングなど)から、考察されている情報を得て、さらに元データを入手して、
- 3. 独自のアンケートを作成して、データを収集し、
- 4. 1~3に代わるもので、

1~4のいずれか



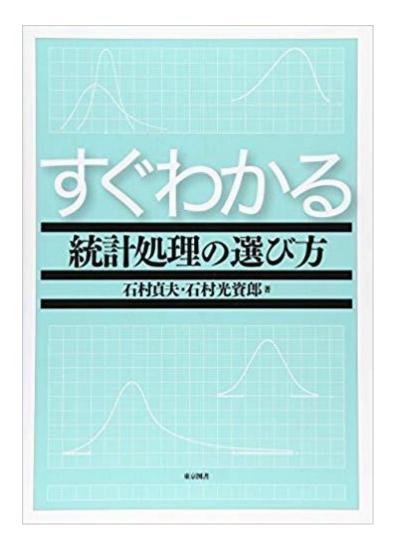
独自のグラフ等を作成し、統計的に解析して、結果・ 考察を発表する

## 参考図書は



## 参考図書は





## 第一回

データの見える化

## 学習目標

### 以下について確認します

- 統計学とは?
- 色々なグラフ

パレート図、ヒストグラム、折れ線グラフ、円グラフ、 帯グラフ、レーダーチャート、ガントチャート

### ● 統計サイト

厚生労働省、都道府県ランキング、アメリカ農務省、 WHO(世界保健機関)

# 



## 統計学を学ぶことへのイメージ

## 統計つて?

コトバンク

コトバンクを検索

Google はこの広告の表示を停止しました

この広告の表示を停止 広告表示設定 ①

統計(読み)トウケイ

#### デジタル大辞泉の解説

#### とう-けい【統計】

ました

この広告の表示を停止 広告表示設定 ①

Google はこの広告の表示を停止し [名] (スル)集団の個々の構成要素の分布を調べ、その集団の属性を数量的に 把握すること。また、その結果を数値や図表で表現したもの。「統計をとる」 「統計を出す」「就業人口を統計する」

出典 小学館 / デジタル大辞泉について 情報 | 凡例

#### 百科事典マイペディアの解説

#### 統計【とうけい】

多数の構成要素(統計単位)からなる集団において、各要素の観察によって得た数値(統計資料)を処理して集団の性 質・傾向を明らかにすること。また統計資料をもいう。集団を一時点でとらえる静態統計と、一定期間でとらえる動態統 計に分けられ、また統計調査の主体により<u>官庁統計</u>と民間統計がある。前者は統計法により規制され、特に重要なものは 指定統計として扱われる。→統計学

→関連項目グラフ | 大量観察法

出典 株式会社平凡社 / 百科事典マイペディアについて 情報

#### 大辞林 第三版の解説

#### とうけい【統計】

(名) スル (statistics)

集団現象を数量的に把握すること。一定集団について、調査すべき事項を定め、その集団の性質・傾向を数量的に表すこ と。「-をとる」

出典 三省堂 / 大辞林 第三版について 情報

#### 日本大百科全書(ニッポニカ)の解説

統計

とうけい

statistics英語

Statistikドイツ語

statistiqueフランス語

ました

この広告の表示を停止 広告表示設定 ①

Google はこの広告の表示を停止し 統計とは、社会現象の量を反映する数字であり、とくに社会集団の状況を数字 によって表現したものである。しかし、現代の統計学における統計的方法の急 速な進歩とその普及に伴って、より一般的には、自然現象や抽象的な数値の集 団をも含めて、いっさいの集団的現象を数字で表したものを統計とよんでい る。 [泉 俊衛]

#### 統計の本質 目次を見る

統計の本質とは何かが問われるのは、主として狭義の意味における統計、つま り、社会的集団の状況を語る数字としての統計についてである。自然現象や単 に抽象的な数値の集団にかかわる数字については、それらが統計としてもつ意 義や特質は初めからとくに問題にならないからである。

統計の本質は、それがまず社会に実在する固有の事実と結び付き、同時に社会的存在としての集団についての数字デー タであることである。たとえば、ある人の賃金20万円、ある世帯の月収30万円などと、それが固有の事実に結び付き、ま た社会現象とみられるものであっても、それが単一の個体についての数字データであるとき、それはまだ統計とはよばな い。それらが含まれた集団、つまり、労働者や世帯の具体的なある一定の集団についての数字データ、同種の事例(個 体) を集めた集団についての数字が統計である。統計は、統計調査における統計集団の構成(単位、標識、特定の時点な い1時期と提託かい1節囲の相定)に其づいてというとれる。その場合、統計集団をいかに構成するか。とれ目体的にい

## 統計つて?

集団の状況を 数値で表したもの



目的:集団の〇〇を知りたい

## 統計学

- ・データを集める
- 解析する
- ・解釈する



ための方法論

結果:集団の〇〇がわかった!

## データを集める準備 Google アカウント を作る

- この講義専用
- アンケートを作るのに使う
- プログラミングサイトpaiza.io の登録用に使う

(注意) 授業が終わったら必ずログアウト

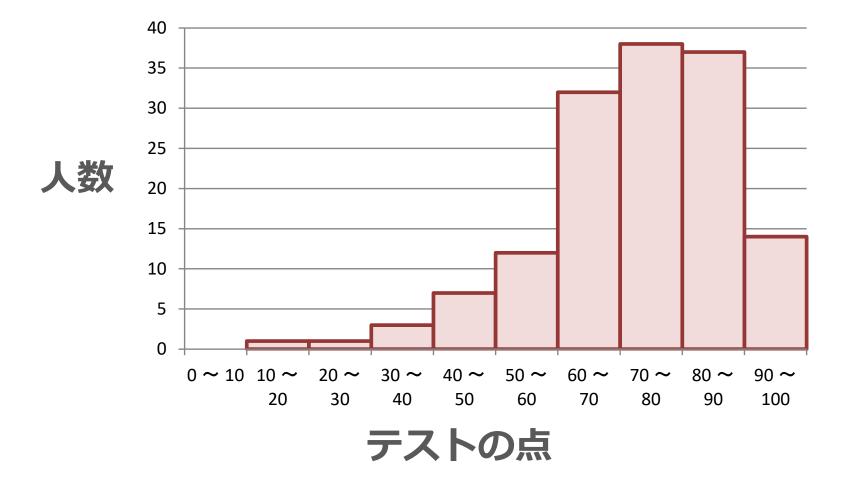
## 身長データをとってみる

授業のサイト

- → リンク公開用(要PW)
- → 作ったアンケートの 公開用Googleドキュメント



## ヒストグラム (度数分布図)



## ヒストグラムの描き方

- 1. 区間の数を決める データ数の平方根が目安。四捨五入して整数に する
- 区間の幅を決める データの最大値-最小値(データの範囲)を、区 間数で割る。四捨五入して、測定の刻み(最小 単位)になるようにする
- 3. 区間の境界を決める 一つ目の区間の開始が、最小値- (測定の刻み ÷2) になるように
- 4. それぞれの区間に入った数を数える

目的に応じて、細かなアレンジは可

## 身長データの ヒストグラムを 作ってみる



## 様々な見える化手法

- ヒストグラム、パレート図
- 折れ線グラフ
- 棒グラフ
- 円グラフ
- 帯グラフ
- ・レーダーチャート
- ガントチャート

## 厚生労働省 統計

## で検索



↑ ホーム

Google カスタム検索

Q検索

テーマ別に探す

報道・広報

政策について

厚生労働省について

統計情報・白書

所管の法令等

申請・募集・情報公開

↑ ホーム > 統計情報・白書 > 各種統計調査

#### 各種統計調査

- 統計調査実施のお知らせ(直近10件まで) 統計調査実施予定 最近公表の統計資料

- 年報等で公表・提供しているもの(直近5件まで)
- 月報で公表・提供しているもの(直近10件まで)

- □ 厚生労働統計一覧 □ 統計要覧一覧 □ 統計情報をご利用の方へ

#### 統計調査実施のお知らせ(直近10件まで)

▶ 実施のお知らせ一覧へ

- 2021年8月17日掲載
- ▶ 令和3年パートタイム・有期雇用労働者総合実態調査にご協力ください
- 2021年8月6日掲載
- ▶ 令和3年度雇用均等基本調査のお願い
- 2021年8月4日掲載
- ▶ 令和3年社会福祉施設等調査及び介護サービス施設・事業所調査へのご協力を お願いします

- か 統計情報・白書
- 各種統計調查
- 統計調査実施のお知らせ
- 統計調查実施予定
- ▶ 最近公表の統計資料
- ▶ 統計調査公表予定
- ▶ 厚生労働統計一覧
- 統計要覧一覧
- ▶ 統計情報をご利用の方へ
- ▶ 統計について学ぼう



Google カスタム検索

Q検索

テーマ別に探す

報道・広報

政策について

厚生労働省について

統計情報・白書

所管の法令等

申請・募集・情報公開

↑ ホーム > 統計情報・白書 > 各種統計調査 > 厚生労働統計一覧

#### 厚生労働統計一覧

1.人口・世帯

- ② 2.保健衛生 ◎ 3.社会福祉 ◎ 4.介護・高齢者福祉 ◎ 5.社会保険

- 8.賃金 9.労働時間
- 10.福利厚生 □ 11.人材開発
- 12.労働災害・労働安全衛牛・労働保険 13.労使関係 14.その他

厚生労働省で実施している主な統計調査や業務統計について、その調査内容、調査対象、調査周期、公表予定、実施 担当部局及び集計結果表等の搭載場所等をみることができます。

- 厚生労働統計調査名英訳名称一覧はこちら
- ▶ 厚牛労働統計調査・業務統計等体系図(分野別・対象別一覧表)はごちら
- ▶ 厚生労働統計調査・業務統計等体系図(ポイント) [XSLX形式:217KB] はこちら

\* 印は業務統計

#### 1.人口・世帯

出生・死亡や人口の移動などによる人口変動や世帯の活動などに関するデータを提供しています

- 1.1.人口 1.2.世帯 1.3.縦断調査 (パネル調査)

- 統計情報・白書
- 各種統計調査
- 統計調査実施のお知らせ
- 統計調查実施予定
- 最近公表の統計資料
- ▶ 統計調查公表予定
- ▼ 厚生労働統計一覧
- ▶ 地域児童福祉事業等調査
- か 統計要覧一覧
- 統計情報をご利用の方へ
- 統計について学ぼう
- 統計関連サイトリンク

#### 1.人口・世帯

出生・死亡や人口の移動などによる人口変動や世帯の活動などに関するデータを提供しています



#### 1.1.人口

統計・調査名	統計・調査内容
▶ <u>人口動態調査</u> NEW 9月7日	出生・死亡・婚姻・離婚及び死産の人口動態事象を把握 本調査は、統計法に基づく基幹統計『人口動態統計』の作成を目的と する統計調査
▶ <u>人口動態職業・産業別</u> <u>統計</u>	国勢調査年の4月1日から翌年3月31日までの1年間で発生した人口動態事象(出生・死亡・死産・婚姻・離婚)について職業(死亡については産業も含む)を調査し、人口動態事象と社会経済的属性との関連を明らかにする
▶ <u>人口動態統計特殊報告</u> NEW 7月30日	人口動態調査を基に、特定のテーマについてとりまとめたもの
▶ <u>生命表</u> NEW 7月30日	ある期間における死亡状況(年齢別死亡率)が今後変化しないと仮定 したときに、各年齢の者が1年以内に死亡する確率や平均してあと何 年生きられるかという期待値などを死亡率や平均余命などの指標(生 命関数)によって表したもの

御意見募集やパブリックコメントはこちら ■ 国民参加の場

テーマ別に探す

報道・広報

厚生労働省

政策について

厚生労働省について

統計情報·白書

所管の法令等

申請·募集·情報公開

ホーム〉統計情報・白書〉各種統計調査〉厚生労働統計一覧〉人口動態調査

#### 人口動態調査

#### ■お知らせ

人口動態統計の調査票について、平成16年、18年、21~29年に3 都道府県からの報告漏れがあることがわかりました。 再集計後の数値については。「人口動態統計(確定数)の概況」をご覧ください。

- 平成29年人□動態調査で追加集計する統計表について(募集結果)
- ◆ 人□動熊調査における外国人統計に関する意見募集(募集結果)。

#### 調査の概要

- 調査の目的
- 調査の根拠法令
- ∌ 抽出方法
- 調査票
- 調査の方法

- 調査の沿革
- 調査の対象
- 調査事項
- 調査の時期

#### ■調査の結果

結果の概要 New 7月22日

集計:推計方法

- 正誤情報
- 利活用事例

- ⇒用語の解説
- 利用上の注意
- 統計表一覧(政府統計の総合窓口e-Statホームページへ移動し

ます)



- 各種統計調査
- ▶ 統計調査実施のお知らせ
- ▶最近公表の統計資料
- ▶ 厚生労働統計一覧
- ▶統計要覧一覧
- ▶統計情報をご利用の方へ
- ▶統計について学ぼう。
- ▶ 統計関連サイトリンク
- 白書、年次報告書。



オーダーメード集計 匿名データ提供



政府統計の 統一ロゴタイプ



文字サイズの変更 様態 大 特大 Goo 御意見募集やパブリックコ

■ 月報(概数)

所管の法:

テーマ別に探す 報道・広報 政策について 厚生労働省について 統計情報・白書 ホーム〉統計情報・白書 > 各種統計調査 > 厚生労働統計一覧 > 人口類形調査 > 結果の概要

人口動態調査

結果の概要

人口動態調査の結果

■ 確定数 ■ 月報年計(概数)

■ 速報 ■ 報告書

⑥その他人口動態統計関連の公表物

■ 我が国の人口動態 ■ 年間推計

#### ■人口動態統計(確定数)の概況

●月報年計(機数)に修正を加えた確定数です。毎年、調査年の翌年9月頃に公表しています。公表時期については<u>。こちら</u>

● <u>都道府県からの報告漏れ</u>による再集計を反映した平成16~29年(2004~2017年)の確定数・保管統計表・保管統計表 (都道府県編)の各統計表を● <u>e-Stat</u>に掲載しました。

(注意) 概況の過去数値については、平成20年(2018年) 以降の規況は再集計による過去数値の修正をおごなっていますが、平成16~28年(2004~2017年)の概況については、過去数値の修正をおこなっていません。口 終告漏れによる再集針をおこなった過去数値を確認される場合は、平成20年以降の規況をご覧ください。

- 令和元年 平成30年 平成29年 平成28年
- 平成27年 平成26年 平成25年 平成24年 平成23年
- 平成22年 平成21年 平成20年 平成19年 平成18年
- 平成17年 平成16年 平成15年 平成14年 平成13年
- → 平成12年 → 平成11年 → 平成10年 → 平成9年 → 平成8年
- <u>平成7年</u>

#### ページの一番下に…

#### ■人口動態統計(報告書)

- ◆人□動態統計の報告書です。毎年、調査年の翌年3月に刊行しています。

#### 我が国の人口動態

- ◆人□動態統計の主な内容をグラフ化したものです。
- 2 平成30年我が国の人口動態(平成28年までの動向)[1,522KB]
- ●グラフデータ及び統計表をxls形式でダウンロードできます。
- 図 人口・出生(P6~14)[148KB] 図 死亡・乳児死亡(P15~25)[221KB] 図 自然増減(P26~27)[61KB]
- 図 死産・周産期死亡(P28~29)[48KB] 図 婚姻・離婚(P30~36)[156KB] 図 特殊報告・平均寿命(P37~39)[70KB]
- 図 統計表(P42~56)[226KB]

#### ■人□動態統計の年間推計

- ●月報(概数)と速報の公表数値を用いた推計です。



平成30年

### 我が国の人口動態

Vital statistics in Japan

平成28年までの動向 Trends up to 2016



厚生労働省政策統括官(統計·情報政策担当)

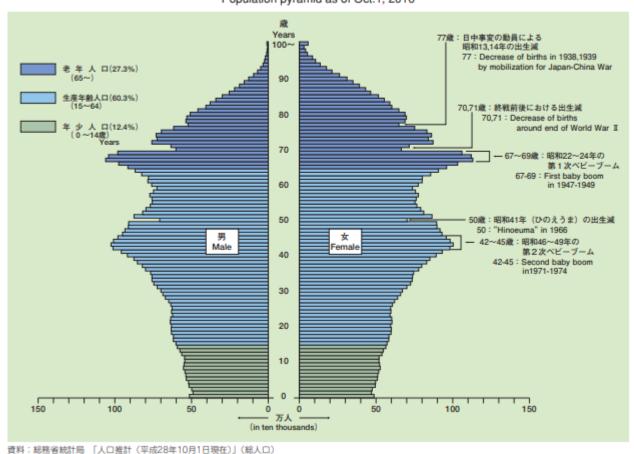
DIRECTOR-GENERAL FOR STATISTICS AND INFORMATION POLICY,

## ヒストグラム

#### 人 口 Population

平成28年の総人口は1億2693万人 老年人口は27.3%

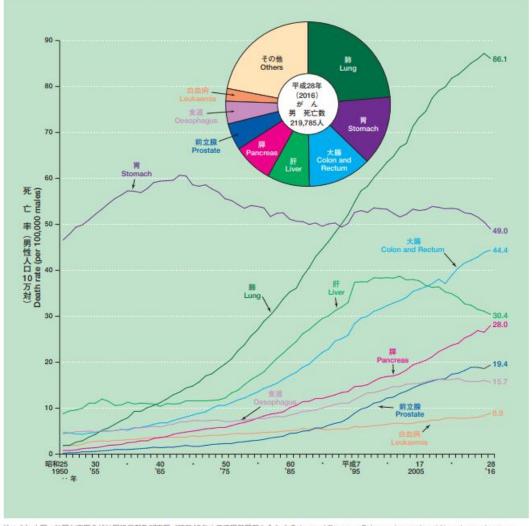
我が国の人口ピラミッドー平成 28 年 10 月 1 日現在ー Population pyramid as of Oct.1, 2016



# 折れ線グラフ円グラフ

### 部位別にみたがんの死亡率の年次推移, 男一昭和25~平成28年一

Trends in death rates for cancers by site, Male, 1950-2016

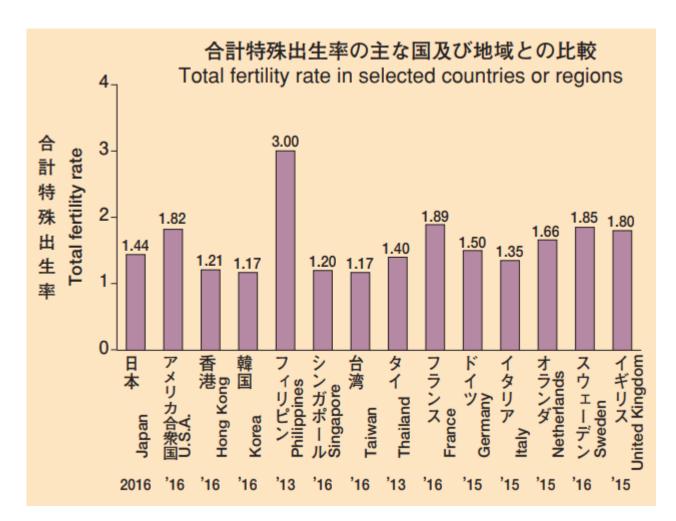


p.20

注: 1) 大腸・結腸と直腸 S 状結腸移行部及び直腸 (昭和42年まで直腸肛門部を含む。) Colon and Réctum・Colon and rectosigmoid junction and rectum
2) 肝・肝及び肝内胆管 (昭和32年まで胆のう及び肝外胆管を含む。) Liver・Liver and intrahepatic bile ducts

<sup>3)</sup> 肺←気管、気管支及び肺 Lung←Trachea, bronchus and lung

### 棒グラフ



## **p.14**

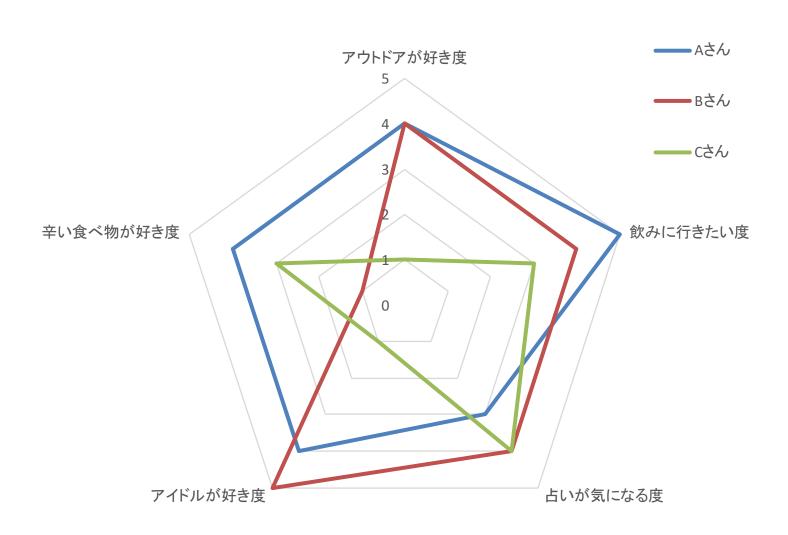
## ● 帯グラフ

### 都道府県別にみた年齢3区分別人口割合及び老年人口指数一平成28年一

Percent distribution of population by 3 age groups and aged dependency ratio, by prefecture, 2016

in	年少人口 千人 0~14歳 thousands Years	生産年齢人口 15~64	老年人口 65~	老年人口指数 Aged dependency ratio
全 国 Total	126,933 12.4	60.3	27.3	45.2
北海道 Hokkaido	5,352 11.2	58.9	29.9	50.9
青 森 Aomori	1,293 11.2	57.8	31.0	53.7
岩 手 Iwate	1,268 11.6	57.2	31.1	54.4
宮 城 Miyagi	2,330 12.2	61.3	26.4	43.1
秋 田 Akita	1,010 10.3	55.0	34.7	63.1
山 形 Yamagata	1,113 11.9	56.5	31.5	55.8
福 島 Fukushima	1,901 11.9	58.7	29.5	50.2
茨 城 Ibaraki	2,905 12.4	60.0	27.6	46.0
栃 木 Tochigi	1,966 12.7	60.6	26.7	44.1
群 馬 Gunma	1,967 12.5	59.1	28.3	47.9
埼 玉 Saitama	7,289 12.4	62.1	25.5	41.0
千 葉 Chiba	6,236 12.2	61.2	26.5	43.4

# レーダーチャート



## ガントチャート

			8月					L	Ĺ						L					L								9)	月							Ĺ						L								
			1	2 3	3 4	5	6	7 {	3 9	10	11	12	13 1	14 1	5 16	17	18	19 2	20 2	22	23	24 .	25 2	26 27	28	29 3	30 3	11	1 2	3	4	5 t	i 7	8	9 1	0 11	12	13	14 1	5 16	5 17	18	19	20	21	22 2	23 2	4 25	26	27
			木 :	金土	: 日	月.	火力	k オ	: 金	<u>±</u>	日	月,	火	Κą	金力	<u>±</u>	日.	月;	く <sub>力</sub>	木	金.	<u>±</u> ,	<u>ال</u>	月火	水.	木 <u>5</u>	金ュ	t E	1月	火.	水石	金力	<u>±</u>	日,	<u>月</u> ッ	火水	木	金	<u> </u>	3 月	火	水	木	金	<u>±</u>	日月	引火	く水	木	金
非稼働日				*								<u>ا</u>	k ×	* *												*	*	k   *				*		1	*				*	<b>*</b> *	*	*	*		<b>*</b>	* *				
工程	開始日	終了日																																																
	_		* *	k *	*	* :	* *	*	*	*	*	<b>*</b> }	k *	<b>*</b>	*	*	*	* *	: *	*	*	* 1	<b>k</b> *																											
①情報収集		8月20日																																																
																					*	* *	<b>k</b> *	k *	*	* *	k																							
②内容詳細設計	8月21日	9月4日																																																
																							I						*	*	* *	*	*	* 7	<b>*</b> *	*	*	*	* *	*										
③スライド作成																																																		
授業	9月17日	9月23日												T									T	T,																										

### 都道府県別の統計

### https://todo-ran.com/

都道府県別統計とランキングで見る県民性

トップ 国土・インフラ

社会・政治

産業・経済

文化・くらし・健康

娯楽・スポーツ

(i) X

その他

店舗分布

あなたの自動車保険、高すぎ!!

ト付与には一定の条件があります。※3保険加入に伴うポイントの付与は、 エコシステムによる募集経費の削減効果等を楽天会員に還元する制度です

Rakuten 楽天損保

インターネット割引

詳しくはこちら〉

Square



都道府県別統計を比較し た都道府県ランキング。 1419 ランキング掲載中

### 都道府県 ベスト&ワースト

各都道府県の1位と47位 だけを一覧表にまとめま した。県民性が一目で分 かります。

### 都道府県比較

東京vs大阪、埼玉vs千葉 vs神奈川など任意の都道 府県の似たとこ、似てい ないところを一覧表にま とめました。

### □ 著者について

著者: 久保哲朗 プロフィール メール: odomon@ amail.com

> (i) X Square



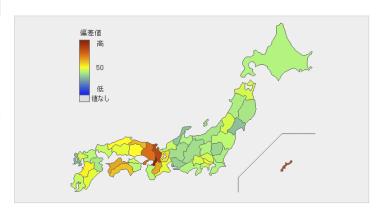
様々な都道府県別統計を比較したランキング。県民性をデータと都道府県ランキングで表します。 最終更新日:2021-9-1

### 最新ランキング

お笑い芸人出身地 [2021年 第一位 大阪府]

▲ いいね!

**ツ**ツイート







□ 記事を探す 検索から探す(googleサイト内検索) Google 提供 サイト内検索 ▶ カテゴリ別全記事一覧 ▶ 新着順全記事一覧 □ テーマ別ランキング ▶ 東西対立型ランキング 東西で高低が分かれるランキング ▶ 都市地方型ランキング 都市と地方の格差が大きいランキング ▶ 東京突出型ランキング

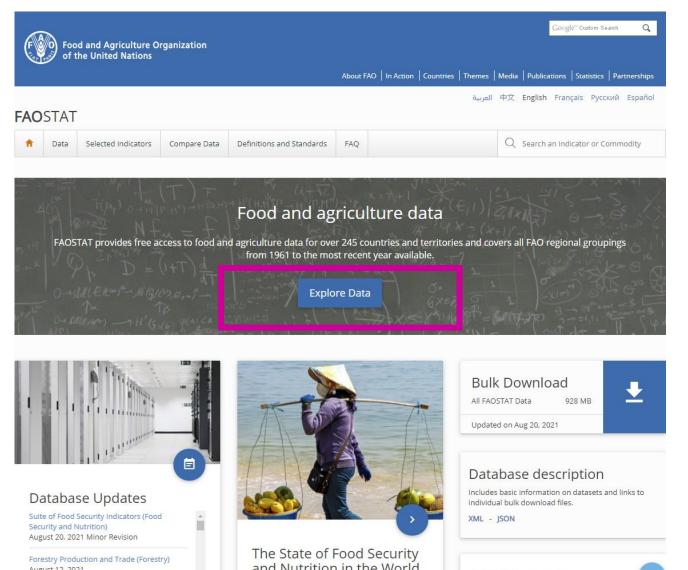


東京が突出しているランキング

▶ ダントツ型ランキング ダントツの都道府県があるランキング

### アメリカ農務省の統計

http://www.fao.org/faostat/en/







Data

Selected Indicators

Compare Data

Definitions and Standards

FAQ

Q Search an Indicator or Commodity

### Data

DOMAINS

DOMAINS TABLE

- 辮 Production
- ▶ 🎢 Food Security and Nutrition 💨 SDG indicators



- Food Balance
- ▶ ﷺTrade
- (\$) Prices
- **U**Land, Inputs and Sustainability



Annual population

**Employment Indicators** 

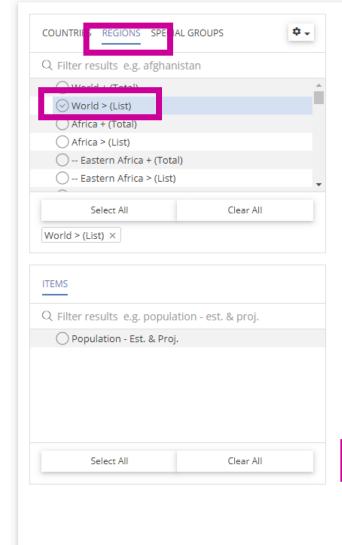
Investment SDG indicator

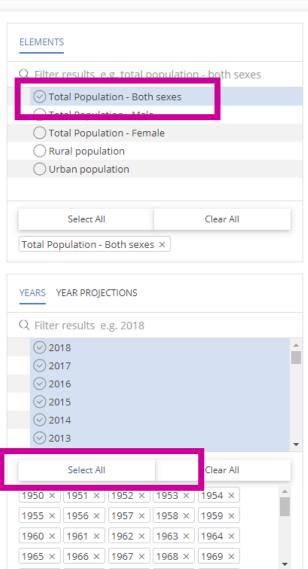


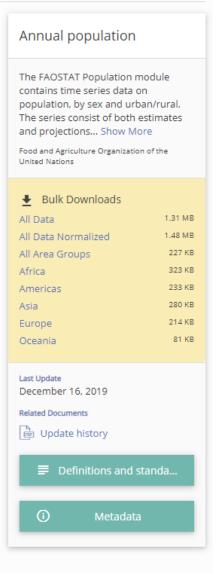
- ▶ Macro-Economic Indicators
- Climate Change
- ▶ Ø Forestry
- Discontinued archives and data series

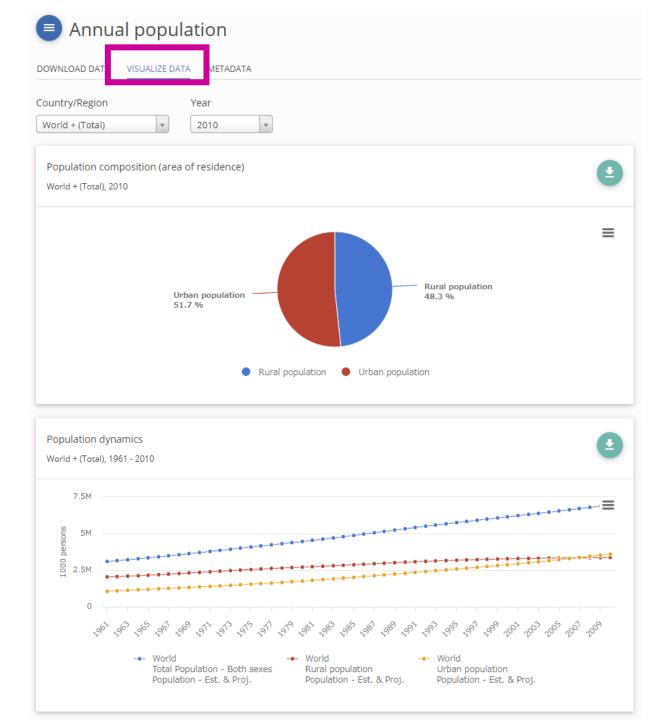


DOWNLOAD DATA VISUALIZE DATA METADATA





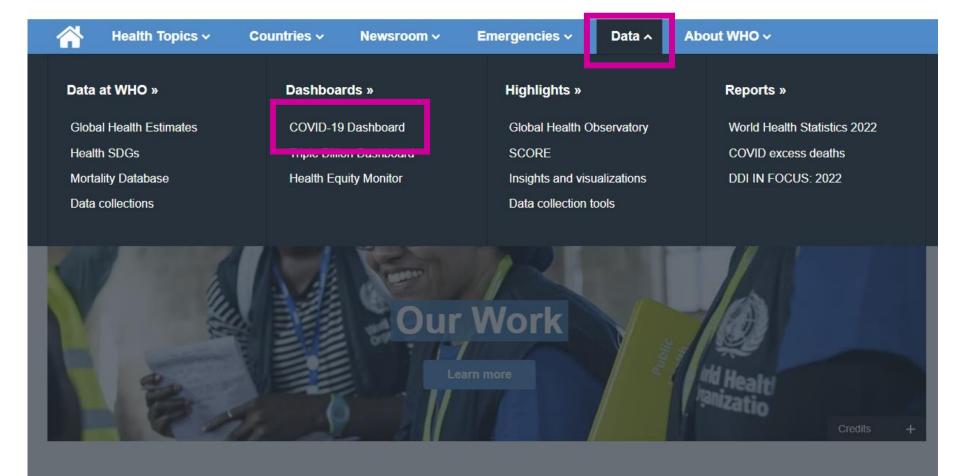




### WHOの統計

### https://www.who.int/







Globally, as of 5:12pm CEST, 25 July 2022, there have been 566,977,818 confirmed cases of COVID-19, including 6,376,503 deaths, reported to WHO. As of 18 July 2022, a total of 12,219,375,500 vaccine doses have been administered.

# 課題内容

- 1. 統計データを公開しているサイト(厚生労働省など)から、データを取得して、
- 2. 統計結果を公開しているサイト(都道府県ランキングなど)から、考察されている情報を得て、さらに元データを入手して、
- 3. 独自のアンケートを作成して、データを収集し、
- 4. 1~3に代わるもので、

1~4のいずれか



独自のグラフ等を作成し、統計的に解析して、結果・ 考察を発表する

# 

統計サイトの検索、閲覧など