

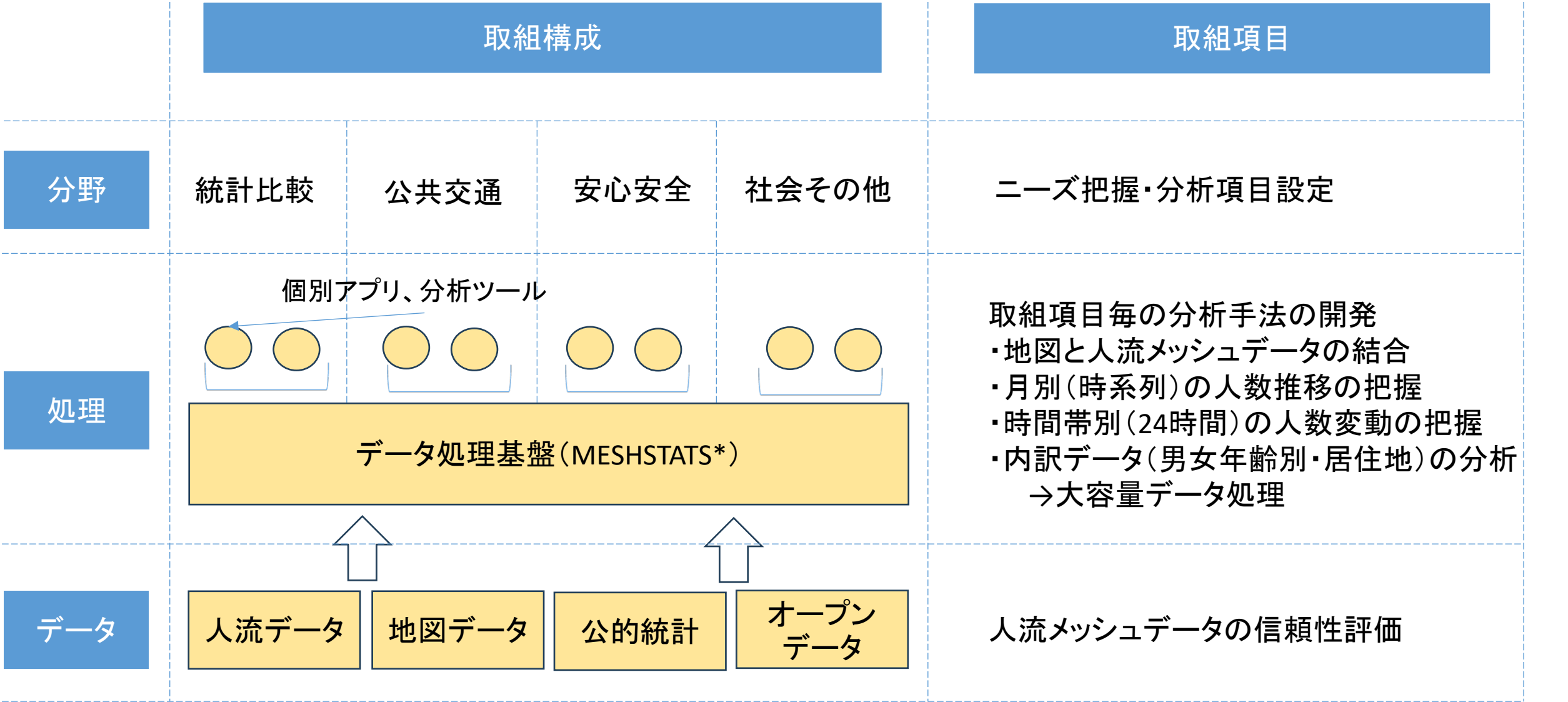
人流データ利活用研究会 R6年度実績

2025年3月25日
総務省統計品質管理推進室

研究会の概要

研究会名	人流データ利活用研究会		
研究会の 目的	民間企業が保有する人流データ等を活用した試行的な取組を実施することにより、各府省の公的統計におけるビッグデータの利活用を促進することを目的として、ビッグデータ等の利活用に関する産官学協議のための連携会議(以下「連携会議」という。)の下で、人流データ利活用研究会(以下「研究会」という。)を開催する。		
研究会の 参加者	【有識者】 岩崎 学（統計数理研究所） 菅 幹雄（法政大学経済学部） 下野 寿之（山梨大学大学院 総合研究部） 佐藤 彰洋（横浜市立大学データサイエンス学部） 所属はすべて令和7年3月現在のもの	【関係機関】 統計研究研修所 【事務局】 統計品質管理推進室 ※月1回程度定例研究会を実施	

R6年度の取組概要



*大量のメッシュ統計を世界規模で取り扱うことができるデータ処理基盤

R6年度の取組項目一覧

- 人流データ利活用研究会を7回実施した
- 昨年に引き続き、日本人宿泊者数推計を毎月行い、複数の人流データでの推計結果の比較も実施した
- 居住区分や性年齢別の内訳のある人流データを活用し、複数の分野で分析や推計を実施した(下表)

分野	タイトル	目的・用途	33回	34回	35回	36回	37回	38回	39回	成果等
			5/29	7/24	9/25	10/23	12/4	1/22	3/5	
統計 比較	都道府県別宿泊者数試算	日本人宿泊者数の先行推計		○	○	○	○	○	○	宿泊者推計を継続実施
	人流データによる居住人口評価	エリア・時間帯別、居住（滞在）人口の把握		○	○					
公共 交通	バス輸送効率性評価方式の検討	路線バスの経路別、潜在利用者数の推計	○	○	○					2024年統計関連学会連合大会報告 和文雑誌投稿予定
	公共交通アクセシビリティ指標の分析	公共交通アクセス容易な面積と人口の把握	○							
安心 安全	市消防施設に関する分析	消防署（消防・救急車数）の適正配置検討		○						
	能登半島地震に関する人流分析	避難者数および復帰状況の把握						○	○	
社会 その他	求人広告データを活用した労働市場分析	滞在人口と求人数の関係把握、賃金への影響			○	○				2024年統計関連学会連合大会報告 英文ジャーナル投稿予定
	地域の総余暇時間の作成方法検討	総余暇時間の都市部と農村部での差異把握			○	○	○	○		統計研究研修所研究報告会
その他	意見交換	人流データの活用状況把握	JTB				DIM			

*DIM: 株式会社ドコモ・インサイトマーケティング

人流データを用いた宿泊者数推計の概要

- 総務省では、ビッグデータを活用した試行的な取組の一環として、携帯端末の位置情報から作成された「メッシュ型人流データ」を用いて、都道府県単位の宿泊者数の推計に関する研究を実施

研究概要

- 人流データから宿泊者数の近似値を算出し、観光庁「宿泊旅行統計調査」の値をベンチマークにして、直近の日本人宿泊者数を推計
※ 宿泊施設が存在する500mメッシュにおける宿泊者数の近似値と、宿泊者数の実績値（「宿泊旅行統計調査」（第2次速報値））とを複数のモデルで回帰分析し、最適な回帰式により直近の宿泊者数を推計

今年度は、携帯基地局のデータを基にした、NTTドコモ「モバイル空間統計」の人流データを使用（2019年10月から推定の直近月まで）

参考URL：[【公式】モバイル空間統計 | 位置情報などのビッグデータを利用した人口統計情報](#)

本研究の訴求点

- 推計に用いるデータは、メッシュ別人流データ、国勢調査（メッシュ別常住人口）、オープンストリートマップ（宿泊施設位置）、宿泊旅行統計（公表値）のみ
- 複数の推計モデル（実数・前年同月・対数前年同月）を設定し、その中から最適なモデルを選定
- 人流データと実績値（「宿泊旅行統計調査」）の関係から、最適な回帰分析区間を自動設定（自動区分線形回帰）
- 最適化した推計モデルにより、公的統計とは異なる表章（集計地域・期間、年齢・性別、居住地別等）による指標作成も可能



宿泊施設が存在する500mメッシュにおける早朝時間帯の滞在者数を人流データから算出

⇒ 宿泊者数の近似値
≡ 午前4時滞在者数（非居住者に限定）

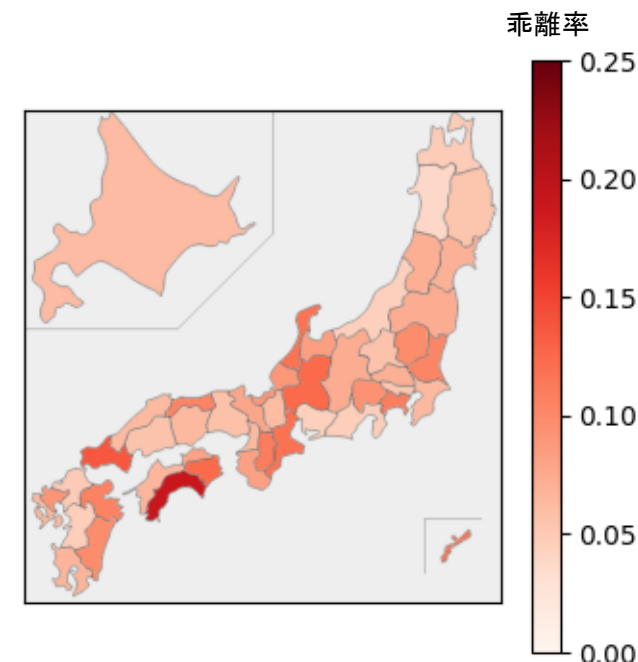
非居住者…該当メッシュと同一または近隣市町村に居住していない人

当該近似値を、過去の日本人延べ宿泊者数の実績（宿泊旅行統計調査の公表値）に回帰し、最も当てはまりの良い回帰式により直近月の宿泊者数を予測



複数のモデル設定と自動区分線形回帰

- 図は統計値と推計値間の年間平均乖離率（絶対値）を都道府県別に示したもの
 - 全国の平均は8.0%（47都道府県の単純平均）
 - 5%未満...7県
 - 10%未満...35県
 - 15%未満...46県
 - 人口規模が大きい県では安定している傾向
 - 北海道（6.2%）
 - 東京都（5.4%）
 - 愛知県（4.5%）
 - 大阪府（6.4%）
 - 福岡県（4.9%）
- その他各県の詳細は次スライド



図：年間平均乖離率（絶対値） 2024年

乖離率＝推計値／統計値－1

- 統計値は観光庁「宿泊旅行統計調査」の第2次速報を基に算出
- 推計に使用した人流データはNTTドコモ「モバイル空間統計」

都道府県別、2024年の年間平均乖離率

	平均	標準偏差		平均	標準偏差		平均	標準偏差
北海道	6.2%	0.07	石川県	11.6%	0.08	岡山県	6.4%	0.05
青森県	4.9%	0.04	福井県	8.9%	0.07	広島県	5.5%	0.05
岩手県	5.5%	0.04	山梨県	9.3%	0.08	山口県	13.7%	0.09
宮城県	7.0%	0.05	長野県	7.4%	0.06	徳島県	12.4%	0.07
秋田県	3.8%	0.03	岐阜県	12.4%	0.09	香川県	8.1%	0.07
山形県	7.1%	0.08	静岡県	4.6%	0.04	愛媛県	6.6%	0.07
福島県	7.3%	0.04	愛知県	4.5%	0.04	高知県	18.9%	0.16
茨城県	10.3%	0.09	三重県	11.9%	0.07	福岡県	4.9%	0.05
栃木県	9.9%	0.06	滋賀県	6.1%	0.06	佐賀県	9.1%	0.10
群馬県	5.6%	0.04	京都府	7.8%	0.06	長崎県	6.5%	0.05
埼玉県	7.0%	0.06	大阪府	6.4%	0.04	熊本県	4.7%	0.03
千葉県	6.4%	0.07	兵庫県	5.7%	0.03	大分県	10.6%	0.09
東京都	5.4%	0.03	奈良県	11.0%	0.15	宮崎県	9.7%	0.06
神奈川県	10.9%	0.10	和歌山県	8.2%	0.09	鹿児島県	6.6%	0.05
新潟県	4.5%	0.03	鳥取県	10.4%	0.07	沖縄県	11.4%	0.10
富山県	8.3%	0.05	島根県	6.3%	0.05			

乖離率＝推計値／統計値－1
・ 統計値は観光庁「宿泊旅行統計調査」の第2次速報を基に算出
・ 推計に使用した人流データはNTTドコモ「モバイル空間統計」

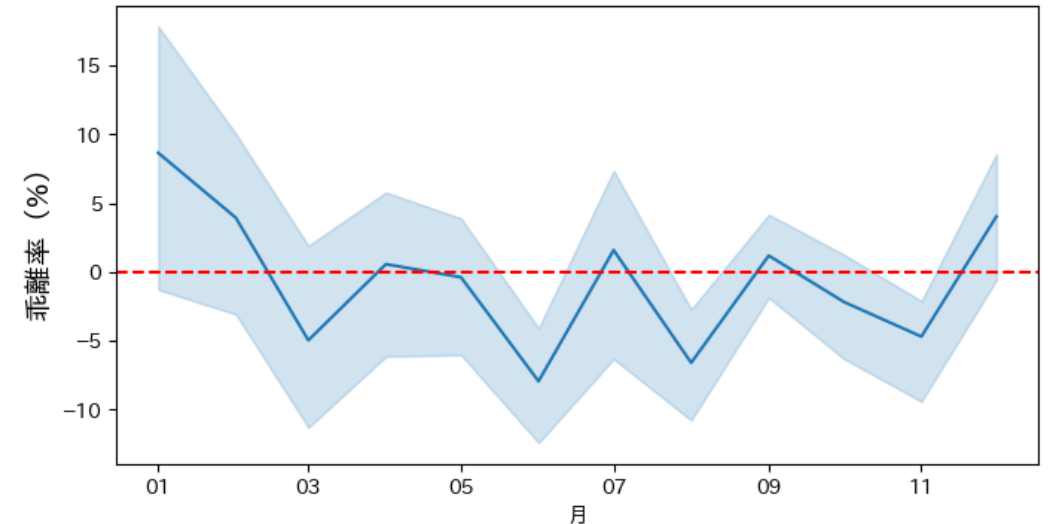
標準偏差は%表記する前のもの

- 図は、月別に47都道府県の乖離率の単純平均を示したもの

■ 平均乖離率：青の実線

- 全体として、過小推計になった月、地域が多かった

✓ 回帰推計に用いているデータ(2019年10月一直近月)がコロナ禍の時期を含むことも1つの要因として考えられる



平均、標準偏差ともに%換算した値

図：月別平均乖離率 全国平均2024年

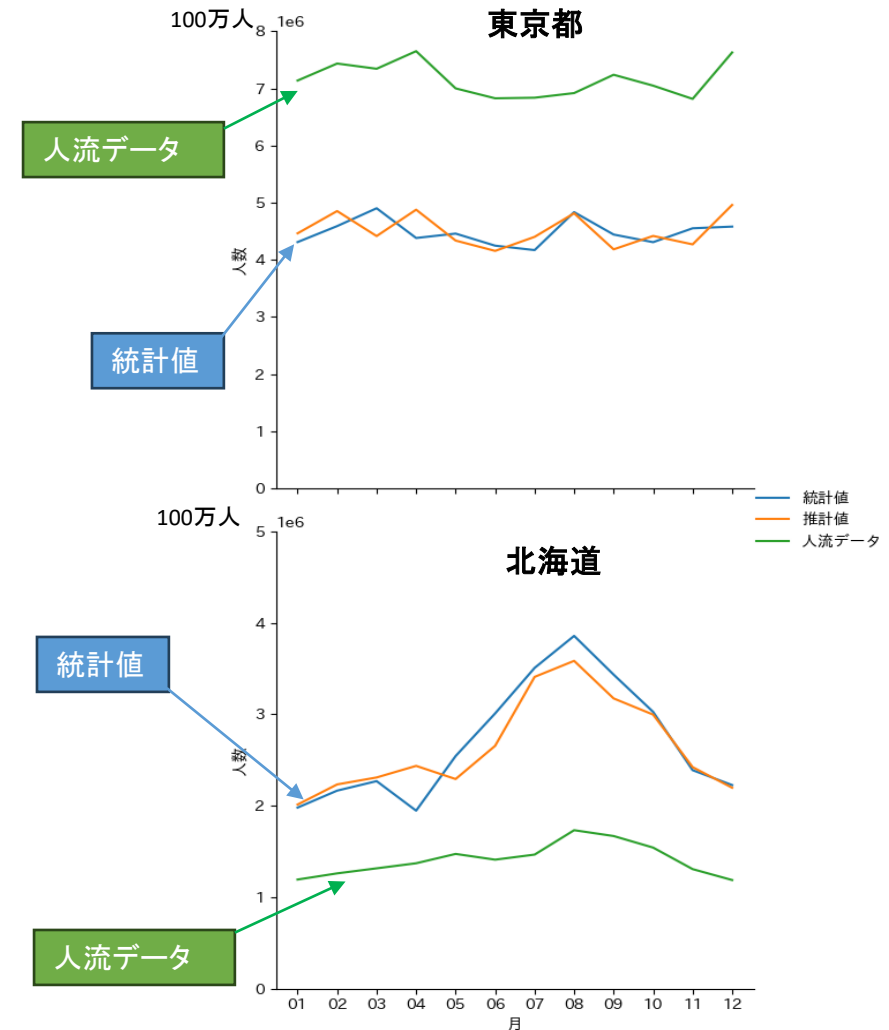
- 青の実線：平均乖離率
- 影をつけた部分は四分位範囲(第3四分位ー第1四分位)

乖離率 = 推計値 / 統計値 - 1

- 統計値は観光庁「宿泊旅行統計調査」の第2次速報を基に算出
- 推計に使用した人流データはNTTドコモ「モバイル空間統計」

事例：東京都と北海道の推計結果 2024年

- 図は2024年、東京都と北海道の各データの値の月次推移を示したもの
- 統計値(青)と人流データ値(緑)の関係
 - 東京都: 統計値 < 人流データの値
 - 北海道: 統計値 > 人流データの値
 - 全国的にみても、東京都と大阪府以外では人流データの値のほうが小さい
- 使用した人流データにおいて:
 - ✓ 拡大推計や人数の秘匿(10人未満メッシュ)の影響がある
 - ✓ 年齢層が限定(15歳以上)
 - ✓ 同一および隣接市町村の滞在者を本推計では除外
 - ✓ すべての宿泊施設を指定できていない etc.
- ただし統計値と人流データの変化のトレンドは、両地域でおおむね一致している



青:「宿泊旅行統計調査」(第2次速報値)から算出した日本人宿泊者数
橙: 推計値
緑: 宿泊施設があるメッシュの人流データ値(午前4時、月間、非居住)

研究による主な知見と今後の取り組み

【研究による知見】

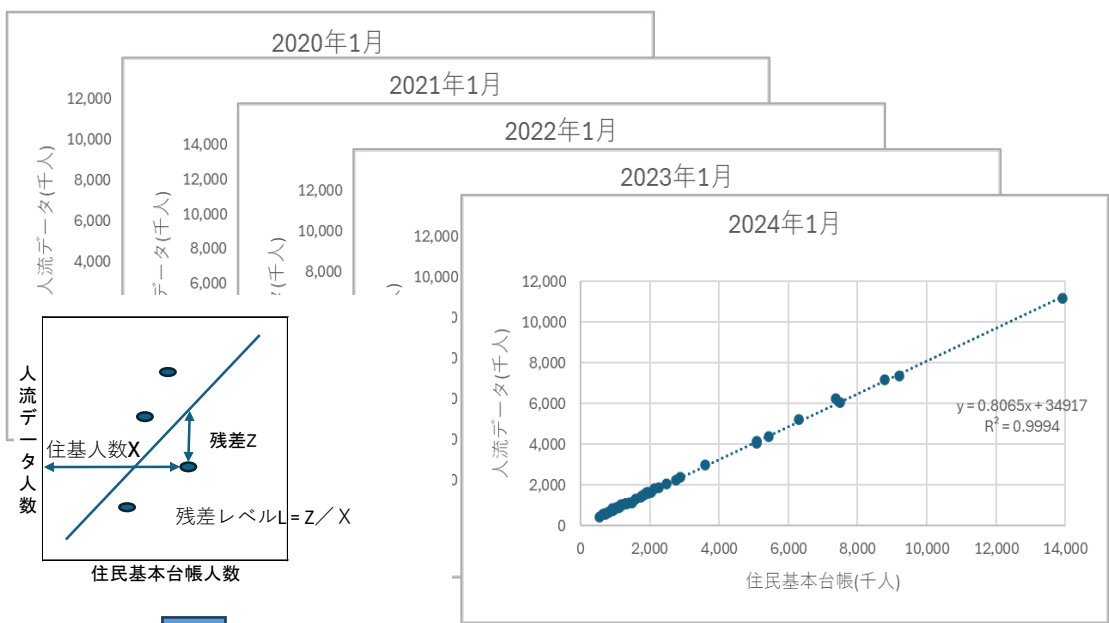
- 人流データから得られる夜間人口について、「宿泊旅行統計調査」の宿泊者数をベンチマークとして検証
 - 人流データから得られた夜間人口と統計調査の値には強い相関があり、先行予測も人口規模が大きい地域を中心に、一定の精度で実現可能であることを確認
 - 一連の処理の自動化により、対象月末日から半月程度で、全国レベルで予測値の算出が可能(統計の公表は約2か月後)
 - ただし、人数の値そのものは統計の宿泊者数と人流データの値に大きな乖離がある
 - 人流データはそのデータサイズの大きさから、様々な対象について集計できる可能を持っている。一方で、個人の行動理由や目的までは一般的に捕捉できず、データ上で対象の特定化が難しいケースは多い。集計目的に即した合理的な仮定をおいたもとで、推論のプロセス(推計や補正等)が必要となる。

【今後の取り組み】

- 人流データを用いた日本人宿泊者数の推計は一定の成果が得られた。今後は、性年齢等の属性別の動向、および訪日外国人の動向にも領域を広げ、宿泊を中心とした様々な活動を例にとりながら、データの妥当性や効果的な推計方法を検討していく

参考

石川県避難者数推計_能登半島地震に関する人流分析（個人研究）



大規模災害時の避難者数を算出する方法を研究

- 年毎に都道府県単位の住民基本台帳人数と人流データの居住者数を相関分析し、都道府県ごとの残差を算出
- 石川県における2024年と例年との残差の差分を能登半島地震の影響と仮定した
- 残差の差分から割り戻して、石川県の避難者数を推計した
 - 公表された避難者数と同水準の数値になった

(使用データ)

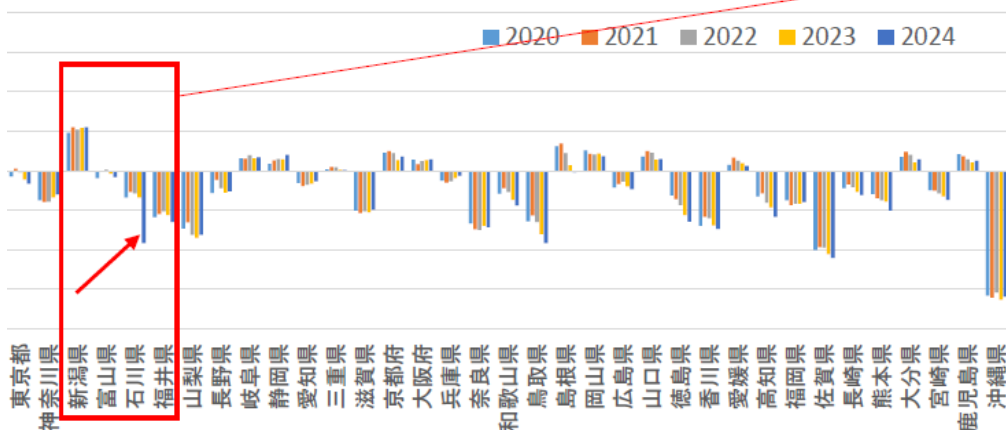
人流データ: NTTドコモ「モバイル空間統計」

居住人口: 住民基本台帳

避難者数: [令和6年能登半島地震における避難所運営の状況\(内閣府防災R6.4.15\)](#)

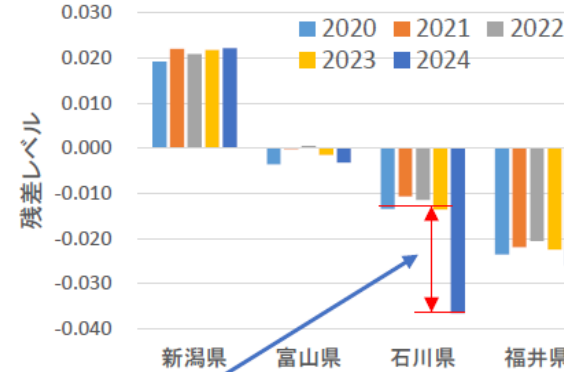
残差

都道府県別残差レベル



石川県以外では例年で残差レベルは一定水準

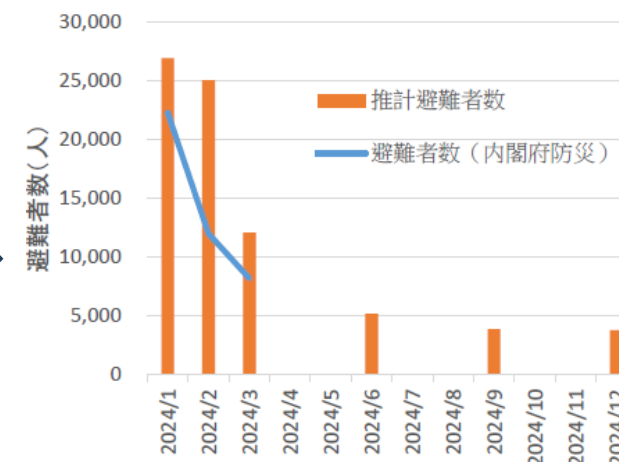
都道府県別残差レベル



能登半島地震の影響

推計

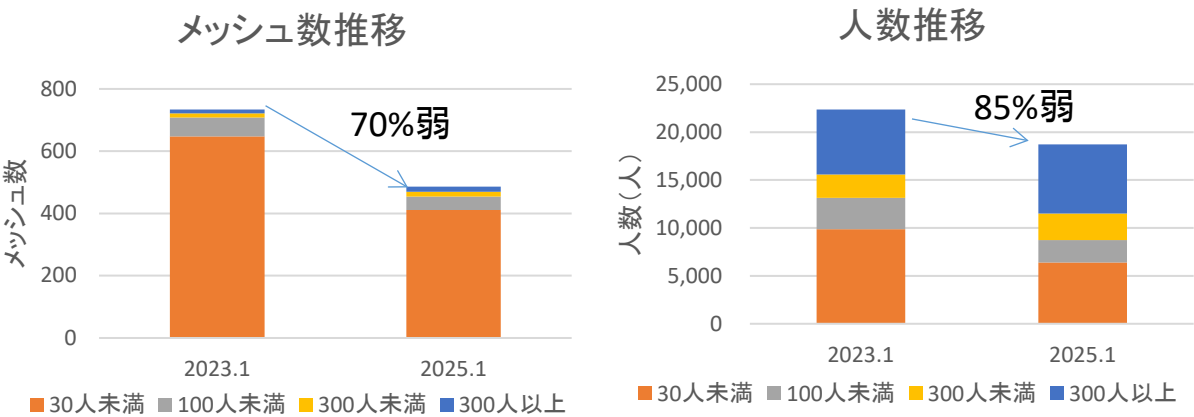
避難者数推移



輪島市避難者状況_能登半島地震に関する人流分析（個人研究）

2025年1月時点での輪島市の被災者はあまり復帰・帰還できていない(2023年1月との居住エリアの比較)

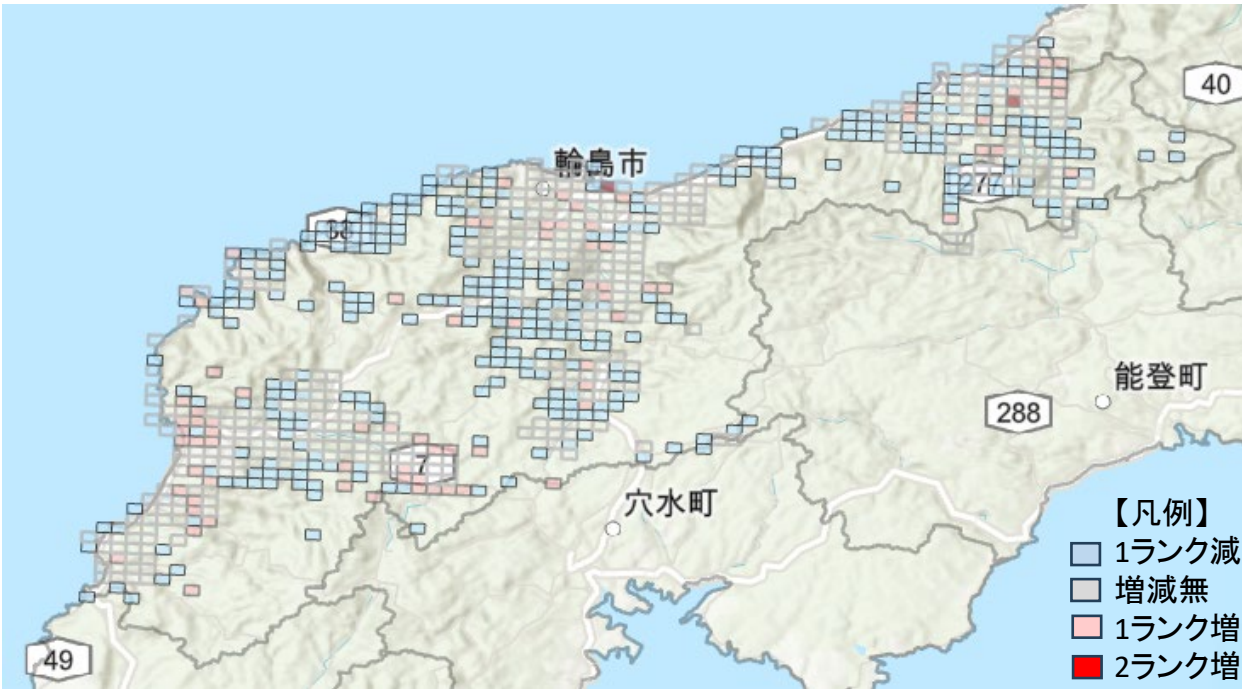
- 居住エリアはメッシュ数ベースで70%弱、人数ベースで85%弱である
- 30人未満のメッシュの約半分がメッシュ無(10人未満)となっており、海岸沿いと山間部で顕著である
- 人数が増加しているメッシュも多数あり、仮設住宅等が設置された場所と考えられる



メッシュ推移(2023.1→2025.1)

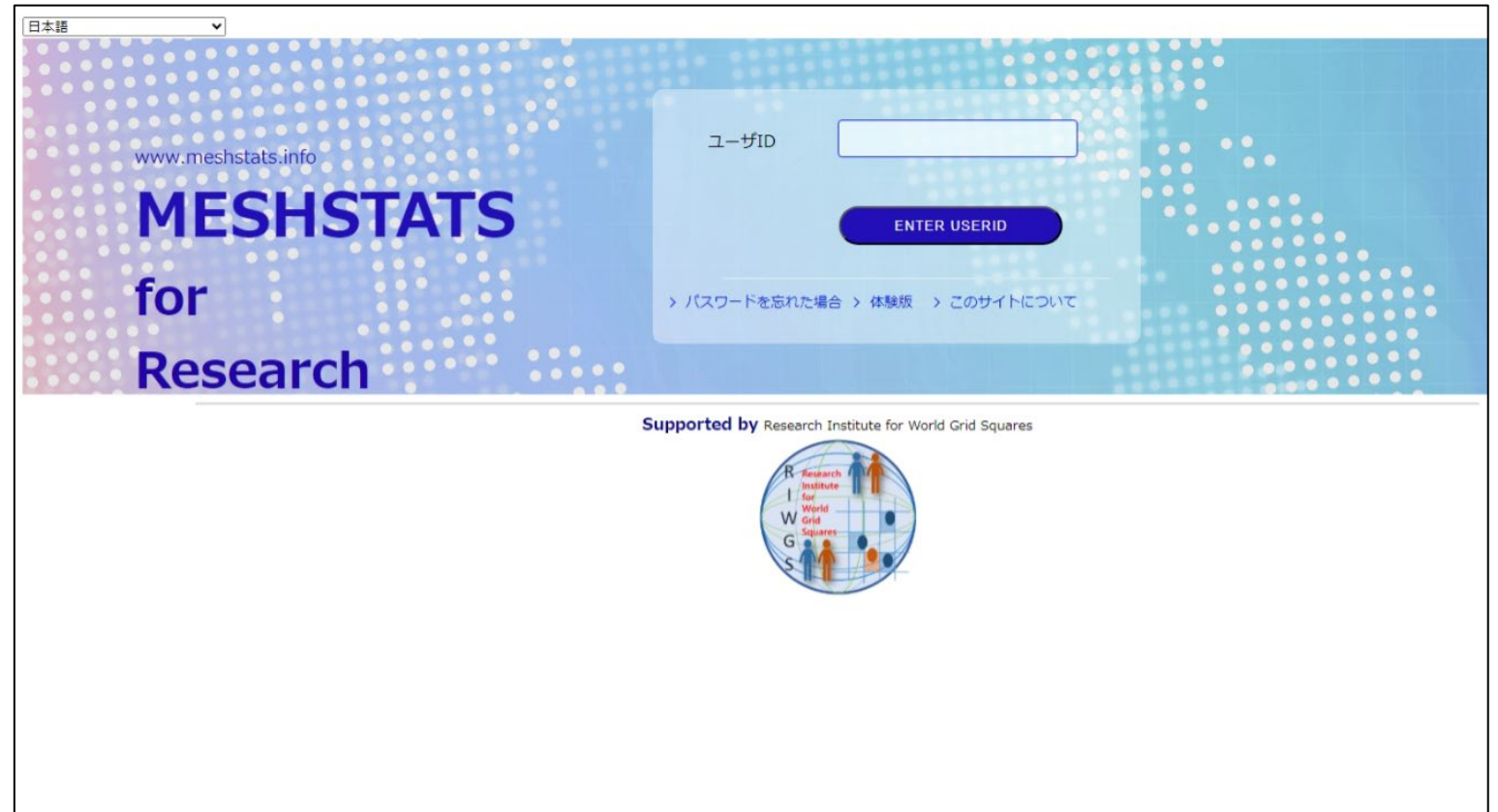
		メッシュ数(2025.1メッシュ数ランク)					
		メッシュなし	30人未満	100人未満	300人未満	300人以上	合計
メッシュ数 (2023.1メッシュ数ランク)	メッシュなし	0	52	0	0	0	52
	30人未満	300	333	14	1	0	648
	100人未満	0	26	25	8	1	60
	300人未満	0	0	4	6	4	14
	300人以上	0	0	0	1	11	12
	合計	300	411	43	16	16	786
		人数増: 80			人数減: 331		

人数増減地図(2023.1→2025.1)



MESHSTATS for Research [1]とは

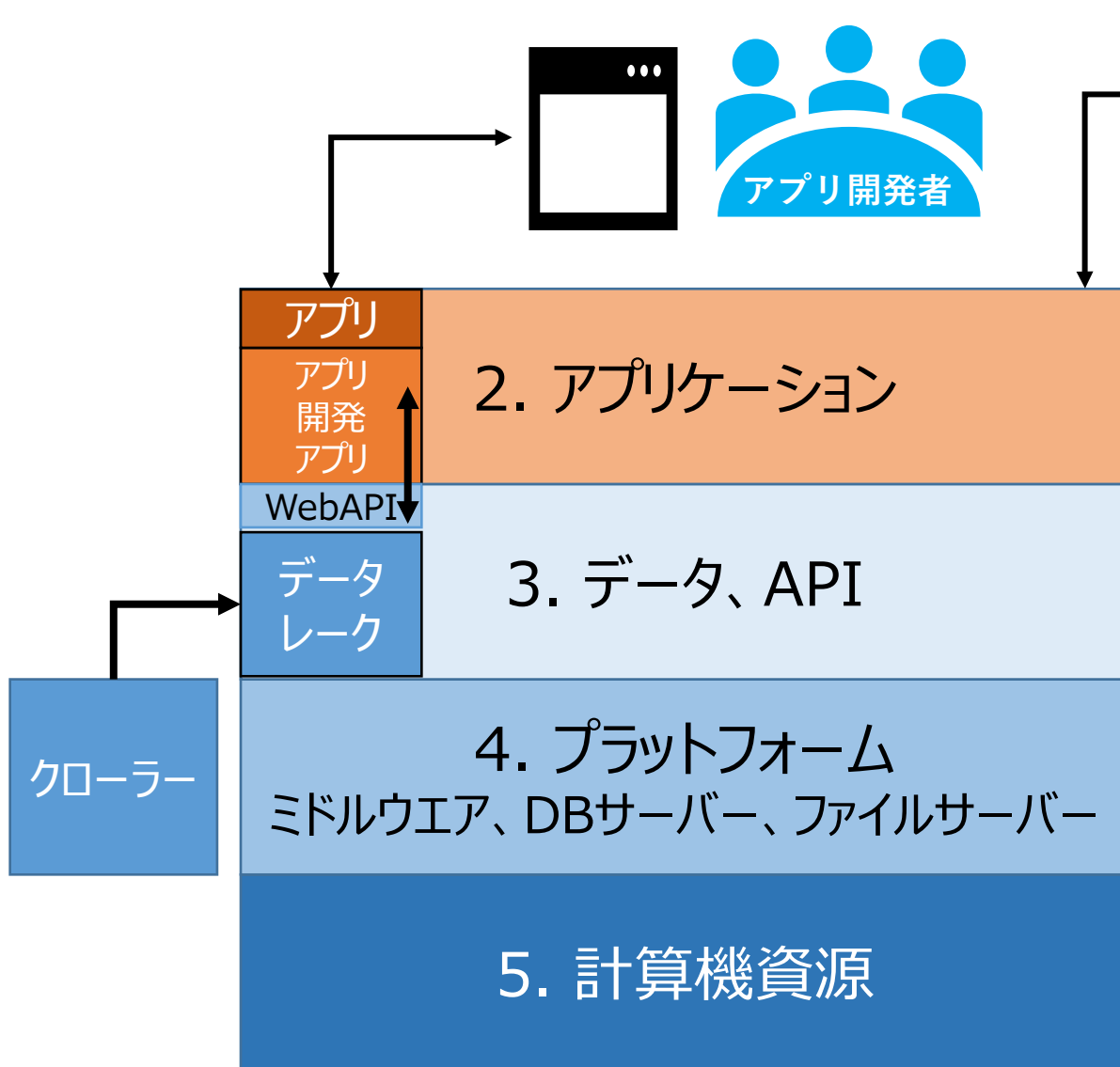
日本産業規格地域メッシュコード(JIS X0410)の独自拡張である世界メッシュコード[2]に基づくメッシュ統計をインターネット経由で活用できるクラウド型サービス



[1] MESHSTATS for Research, <https://www.meshstats.info/meshstats/>, 2025年2月28日

[2] 一般社団法人世界メッシュ研究所, <https://www.fttsus.org/worldgrids>, 2025年2月28日

MESHSTATSの概念図



1. ユースケース

- MESHSTATSは認証・ユーザー管理機能、メッシュ統計、位置情報データベース、APIを備えたメッシュ統計データ利活用のための基盤
- アプリ開発者がアプリ開発時に呼び出しすることができるWebAPIを複数搭載
- 開発者はMESHSTATS上に配置された各種WebAPIを利用してクライアントサイドスクリプティングによりデータアプリを開発することが可能
- 利用者は許可されたデータとアプリへのアクセスとメタリングが行われる