

TDPF 行政データ整備モデル事業 成果報告会

データ整備の必要性和マニュアルの使い方

令和4年3月29日
東京都デジタルサービス局

TDPF
Tokyo Data Platform

1. 本事業の概要と目的

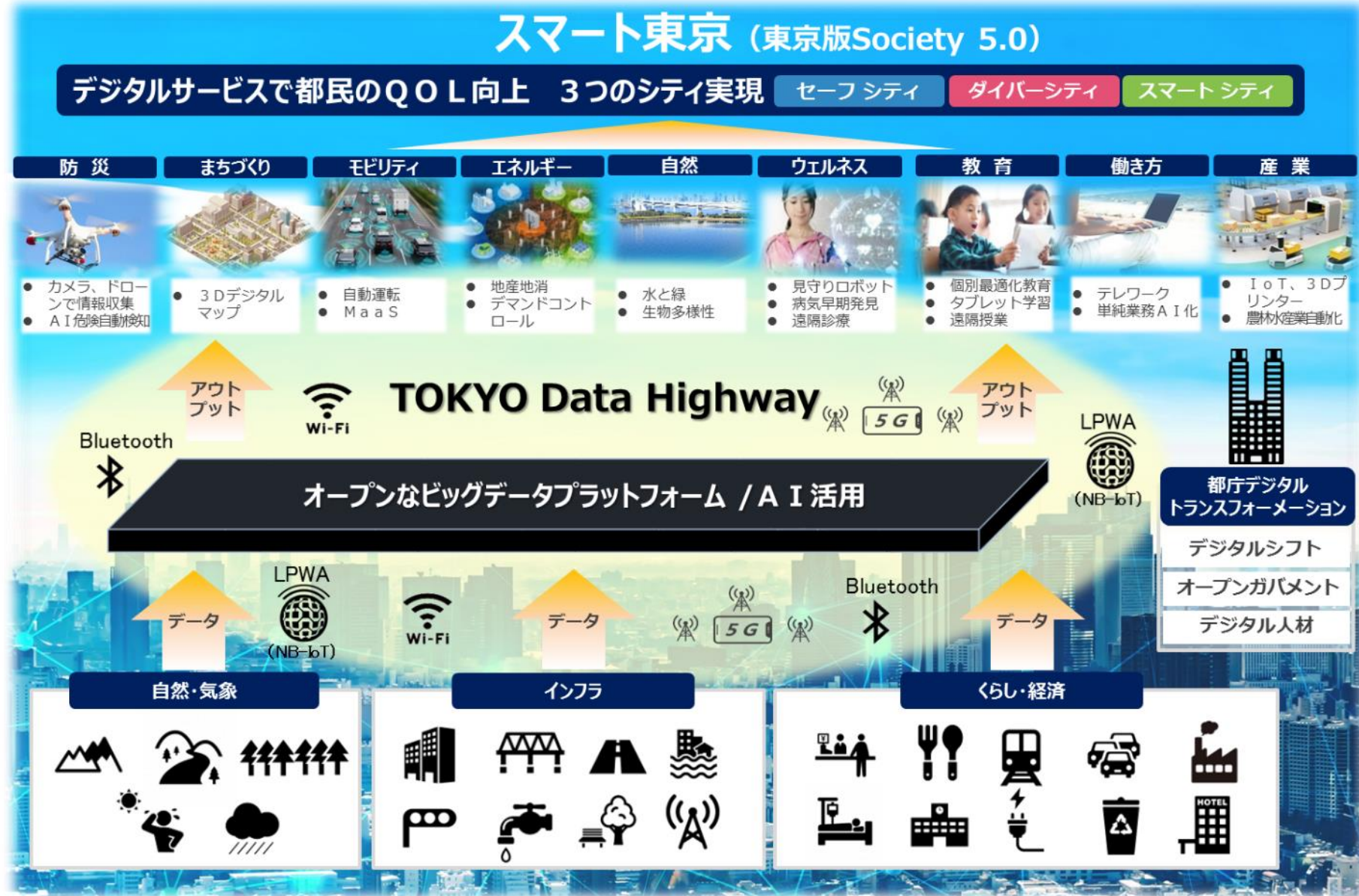
2. データの価値 ～なぜ使いやすいデータが必要なのか～

3. データ活用のプロセス

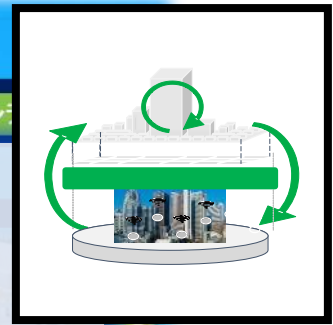
4. データ整備・公開のプロセス

5. マニュアルの利用について

はじめに



東京都のデータ利活用の取組



TDPFの事業概要とデータ整備事業

まずデータ流通推進から取組、段階的な事業拡大を検討

イメージ

TDPF

早期に実施

①データ
流通推進

データ

データ

対価等

対価等

比較的早期に実施

②データ整備

整備前データ

整備の対価等

整備データ等

ニーズ等踏まえ、
将来的な実施を検討

③データ分析・
コンサル等

データ提供に
係る助言等

データ活用に
係る助言等

対価等

対価等

データ利用者
サービス
都民・来訪者等
対価等

事業概要

①

データ流通推進

1.データラ
イブラリ

- 官民の各種データを一元的に検索・利用可能な基盤を提供

2.データ流
通プラット
フォーム

- 官民の各種データの取引条件を定め、データの仲介機能を提供

3.ネット
ワーキング

- データ活用事例の共有や、提供者の課題提示や利用者のニーズを収集する勉強会、イベント開催

②

データ
整備

4.データ整備
支援

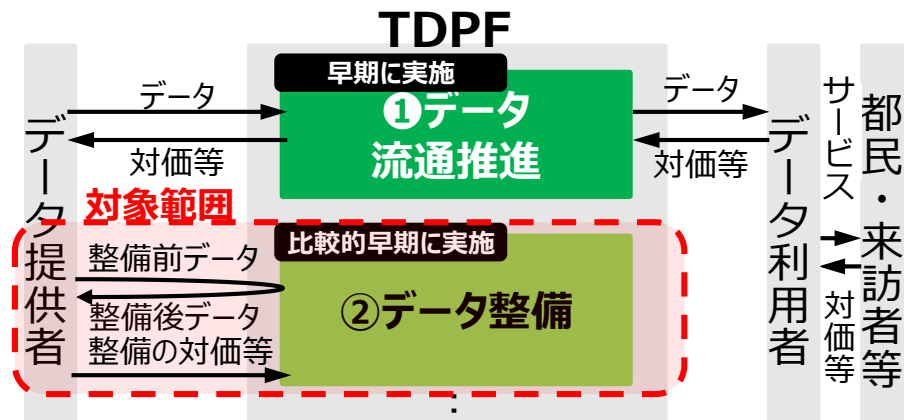
- 紙データ等のデジタル化の支援や、データのクレンジング等を実施

行政データ整備モデル事業 事業概要

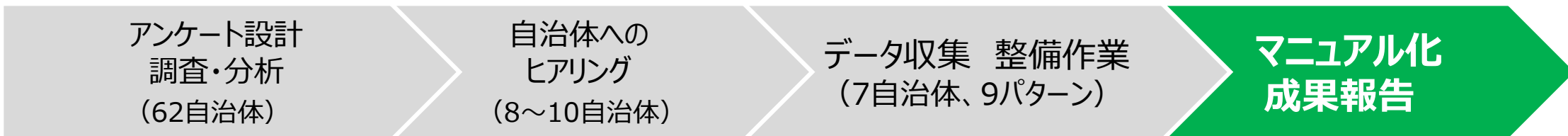
アンケート等を通して選定した行政データの整備を実施し、マニュアル化→公開

事業目的

- 自治体のデータ保有に関する課題認識や、データの質向上
- 機械判読可能な形式へのデータ整備手法をモデル化し、TDPFを通じた区市町村のデータ利活用を促進



進捗



アンケート結果とヒアリング内容を踏まえ、データ整備モデルを絞り込み

対象データ

下記データ（緑色）の整備モデルのマニュアル化を実施

地域・年齢別人口	介護サービス事業所の一覧	医療機関の一覧	食品等 営業許可・届出一覧
イベントや工事に伴う 交通規制	行政情報公開請求の 処理状況 (統計処理したデータ)	バリアフリートイレ情報	避難関連施設情報

整備モデルと協力自治体

自治体のデータ整備の取組状況に応じた、整備モデルのマニュアル化を実施

整備モデル	対象データ	協力自治体	マニュアルの内容
①画像・PDFから変換	地域・年齢別人口	板橋区	PDFのみでしか取り出せないデータがある自治体向けのマニュアル
②エクセルデータからの変換 (緯度経度クレンジング含まず)	介護サービス事業所の一覧	八王子市	エクセルデータ化まではできているが、エクセルデータのフォーマット統一を目的とした自治体向けのマニュアル
	医療機関の一覧	板橋区	
	医療機関の一覧	三鷹市	
	避難関連施設情報	八王子市	
③エクセルデータからの変換 (緯度経度クレンジング含む)	介護サービス事業所の一覧	三鷹市	エクセルデータの中で、住所等のフォーマット統一ができしており、緯度経度情報を追加し正規化するマニュアル
	医療機関の一覧	八王子市	
	バリアフリートイレ情報 (※)	港区	
④複数区のデータ変換・統合	地域・年齢別人口	品川区 目黒区 渋谷区	複数の自治体のオープンデータを変換・統合し、民間事業者の中で流通しやすい形式にするためのマニュアル

1. 本事業の概要と目的

2. データの価値 ～なぜ使いやすいデータが必要なのか～

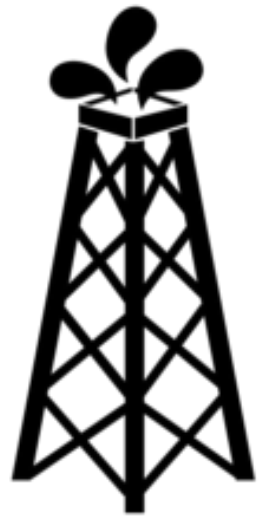
3. データ活用のプロセス

4. データ整備・公開のプロセス

5. マニュアルの利用について

データは21世紀の石油

データを活用可能な状態に精製して流通する



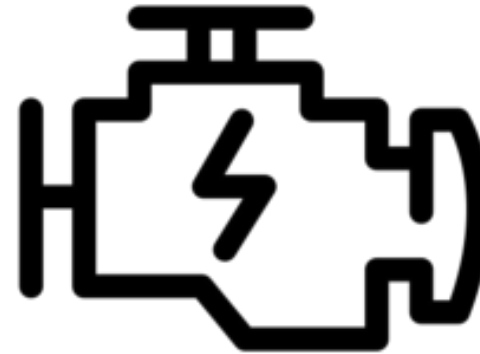
Raw Data

原油



Refined Data

ガソリン



Analytics / ML / AI

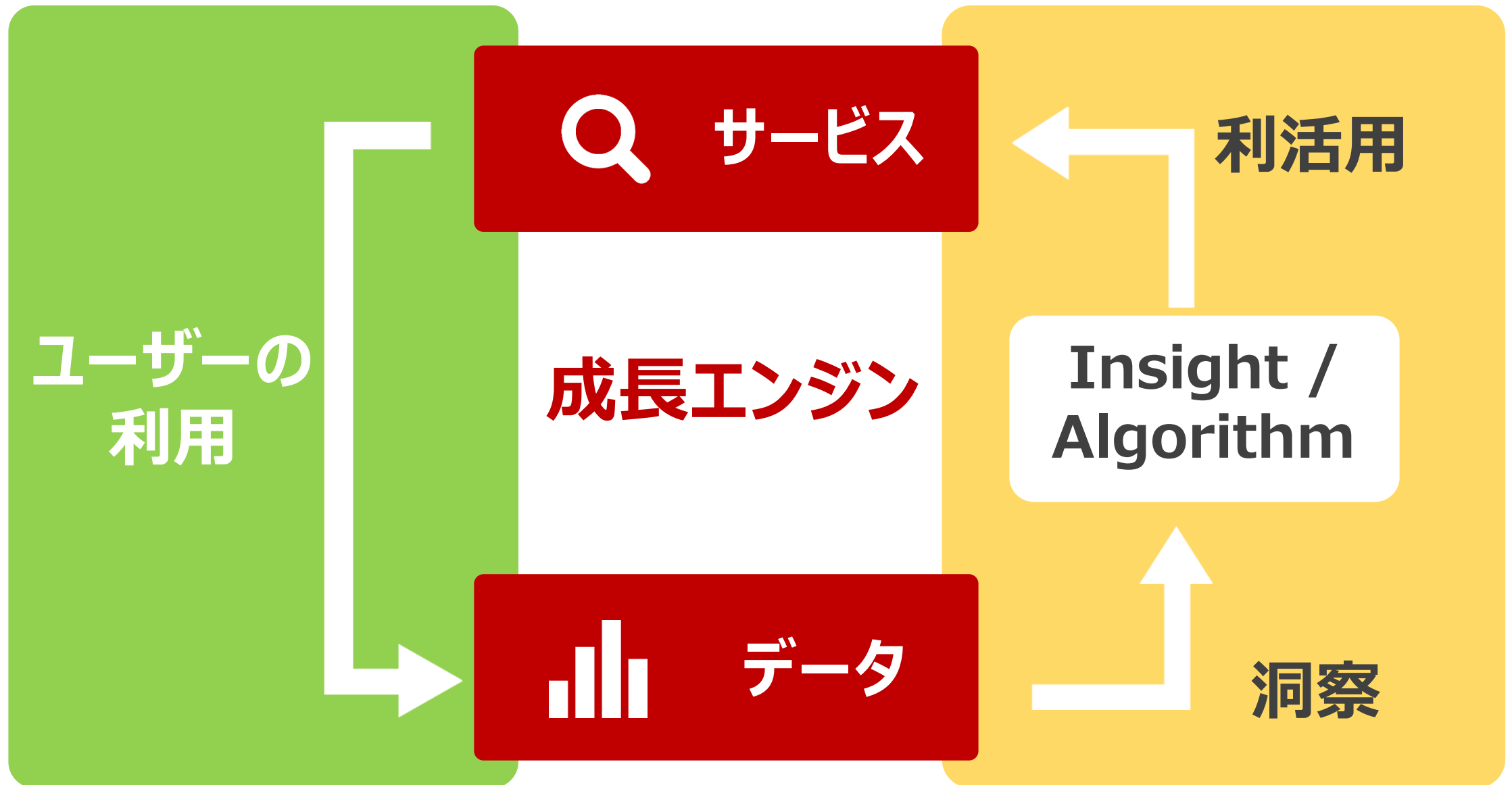
エンジン



Data Product

自動車

データは無限の再生可能エネルギー



データの活用（整備）に向けた課題

キレイなデータを公開したほうがいいことはわかっているが・・・

スキルがない・・・
人材がない・・・

ノウハウの継承が
難しい・・・

どのデータから
公開すれば・・・

作業が増える・・・



行政データの民間ニーズ・活用例

本事業で整備したデータのニーズは、QOLの向上に直結

町丁別・年齢別人口統計データセット

飲食チェーン店、小売チェーン店の出店計画
チラシのコンテンツや配信時間などの最適化



避難関連施設のデータセット

復旧支援や避難経路案内等に利活用
避難所の開設や混雑状況の把握



行政データの民間ニーズ・活用例

他にも行政データの活用例や、更なるニーズも多くある

食品関係営業台帳（届出）

酒卸の**運送ルート**計画
飲食店の**出店計画**



学校給食の献立情報

給食向けの**農作物戦略**の立案



議会の議事録

都市計画向けの**事業者提案**



献立情報には、**アレルギー情報**などがあると参考になるので嬉しい、という意見もありました。

都知事杯オープンデータ・ハッカソン「Demo Day」

ゴミ分別さん

ゴミ分別を画像からAIが分析、通知するサービス



子育てエンジニアズ

保育園の空き状況をマップ表示するサービス



Code for OTA

最適な観光・デートプランを支援するアプリ



ToDCS

昼食難民とキッチンカーをマッチングするサービス



ふわラボ

飲食店の出店支援サービス



ハッカソンの活動



使われ方の理解



データの公開へ

都知事杯オープンデータ・ハッカソン
Demo Day

<https://www.youtube.com/watch?v=rLHNp7Pvt3M>



1. 本事業の概要と目的

2. データの価値 ～なぜ使いやすいデータが必要なのか～

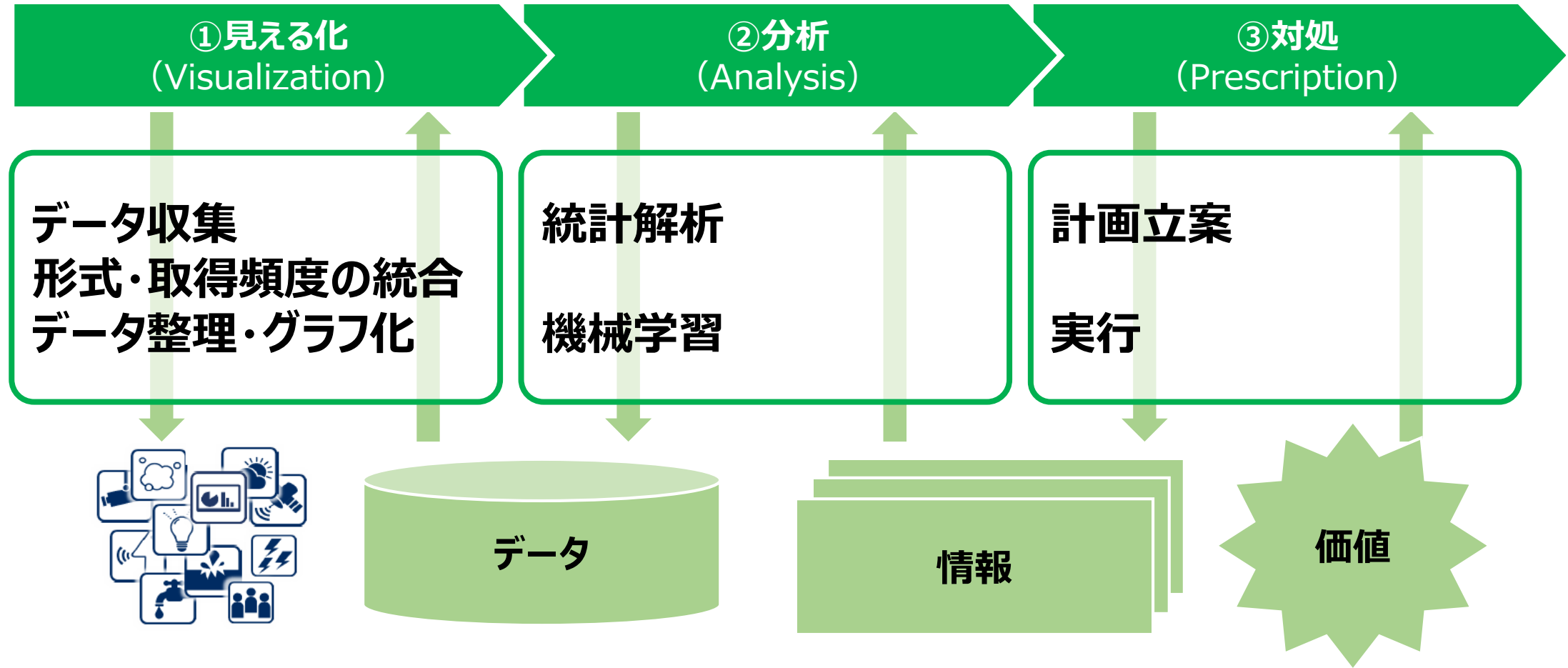
3. データ活用のプロセス

4. データ整備・公開のプロセス

5. マニュアルの利用について

データ活用のプロセスについて

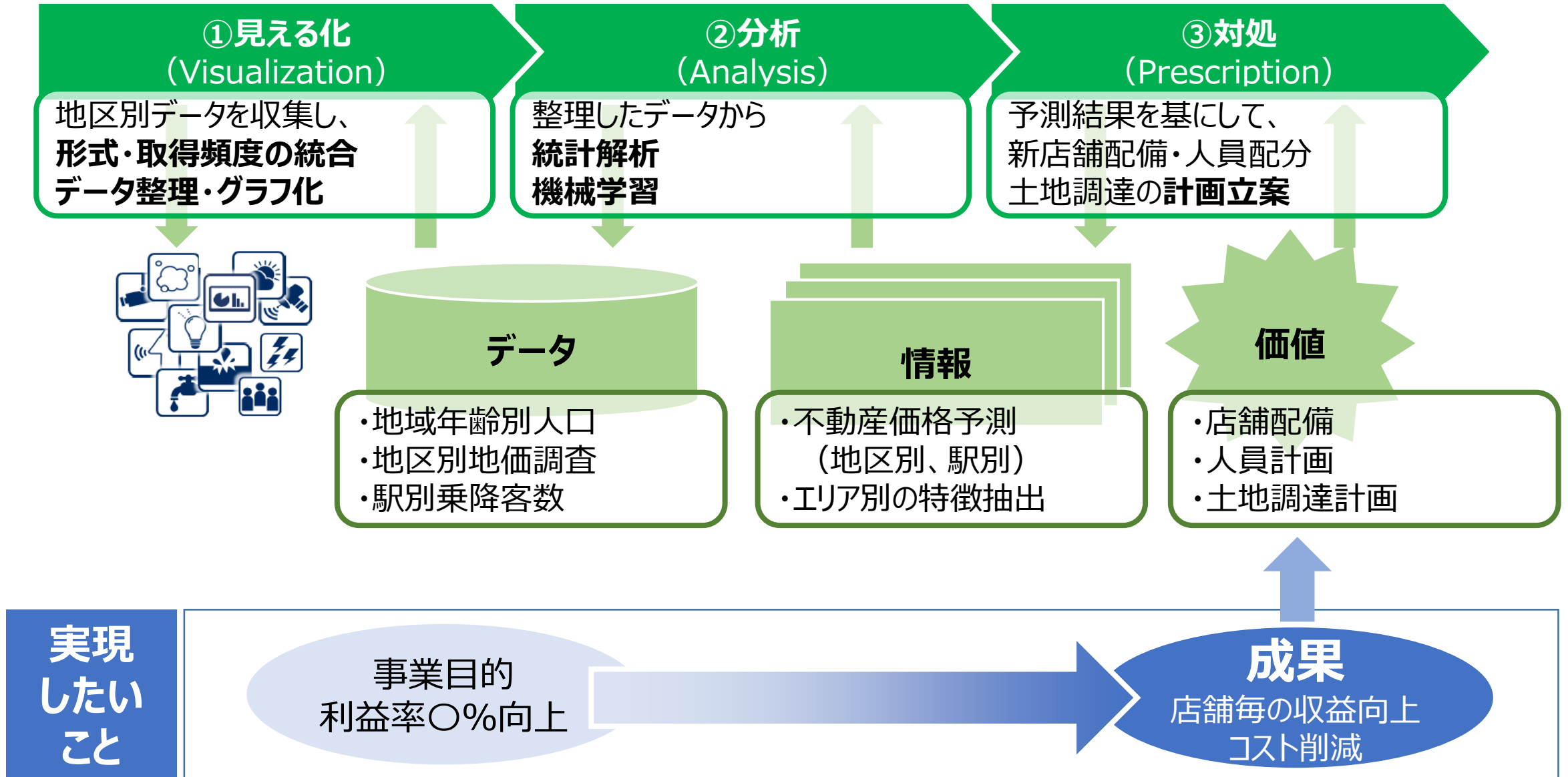
作成協力
日本電気株式会社
本橋様



➡ **地域・年齢別人口のデータを用いてデータ活用のプロセスを試行**
次頁以降：目黒区内人口と国交省の地価調査データを使用した、地価および地価差分予測

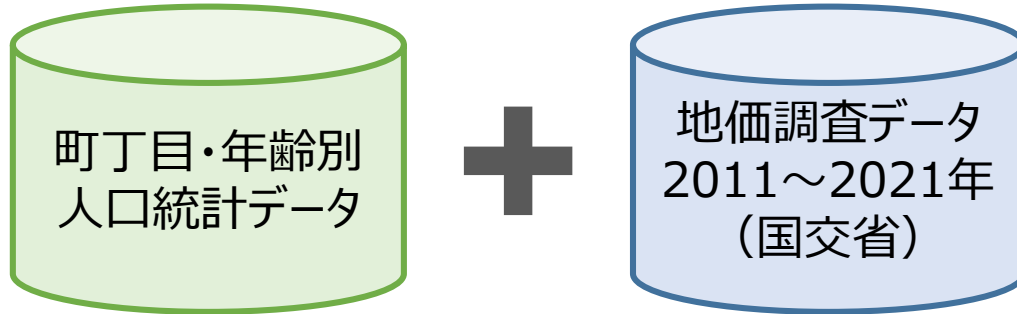
データ活用のプロセスの試行

作成協力
日本電気株式会社
本橋様



データ活用のプロセスの試行 見える化

1. データの統合（行政データ+オープンデータ結合）

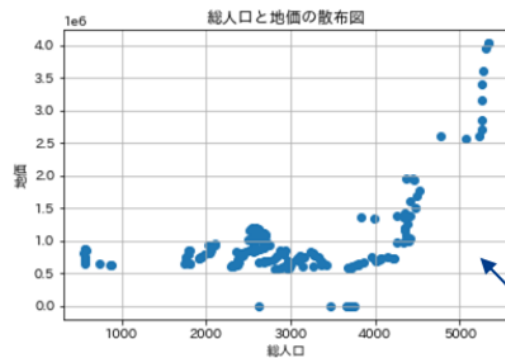


区名	年月日	住民種別	地域コード	町丁名	年齢区分	年齢区分	属性	数値	備考	KEY_CODE
42	2011-08-01	NaN	112	新堀二丁目	0.0	999.0	個人	679	NaN	13110001002
168	2011-08-01	NaN	121	青葉台一丁目	0.0	999.0	個人	2355	NaN	13110002001
504	2011-08-01	NaN	142	大橋二丁目	0.0	999.0	個人	3635	NaN	13110004002
588	2011-08-01	NaN	152	上田原二丁目	0.0	999.0	個人	4259	NaN	13110005002
630	2011-08-01	NaN	153	上田原三丁目	0.0	999.0	個人	4775	NaN	13110005003
...
39564	2021-08-01	NaN	206	陣文台六丁目	0.0	999.0	個人	1784	NaN	13110018006
39942	2021-08-01	NaN	322	緑が丘二丁目	0.0	999.0	個人	3238	NaN	13110022002
40068	2021-08-01	NaN	332	自由が丘二丁目	0.0	999.0	個人	2639	NaN	13110023002
48362	2021-08-01	NaN	361	八雲一丁目	0.0	999.0	個人	2548	NaN	13110028001
48614	2021-08-01	NaN	372	東が丘二丁目	0.0	999.0	個人	3103	NaN	13110027002

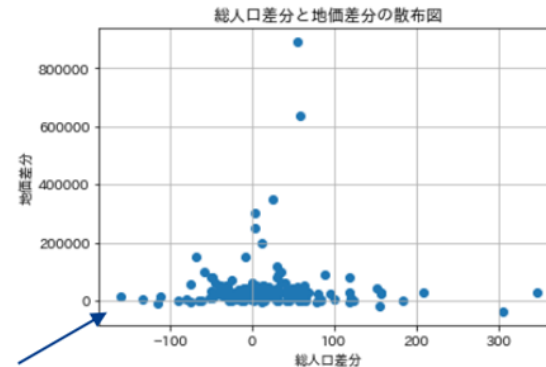
分析し易いように加工・整理
※地価調査データ欠損の影響で
88丁目→16丁目に減少

2. データ見える化（丁目別人口と地価の相関）

人口・地価



人口差分・地価差分 (前年比)



点の総数：176個
(16丁目×11年 (2011~2021))

正の相関あり (相関係数：0.56)

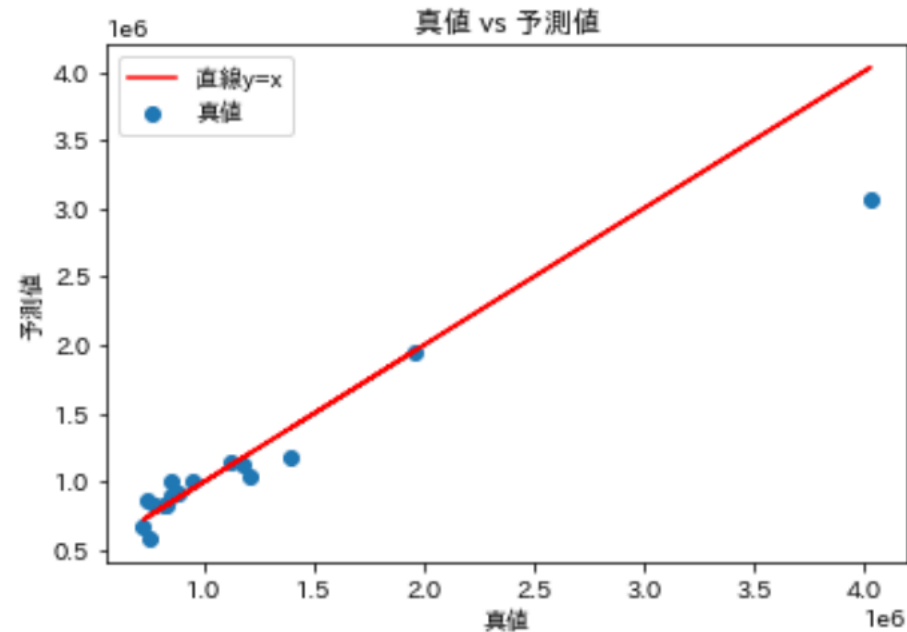
相関なし (相関係数：0.21)

3. 統計解析（数値予測①：地価を予測）

予測結果：

真値に近い予測モデルが作成可能
決定係数 0.89

※決定係数
予測式の真値への当てはまり具合を表す
数値で1.0が最大



真値と予測値の関係グラフ

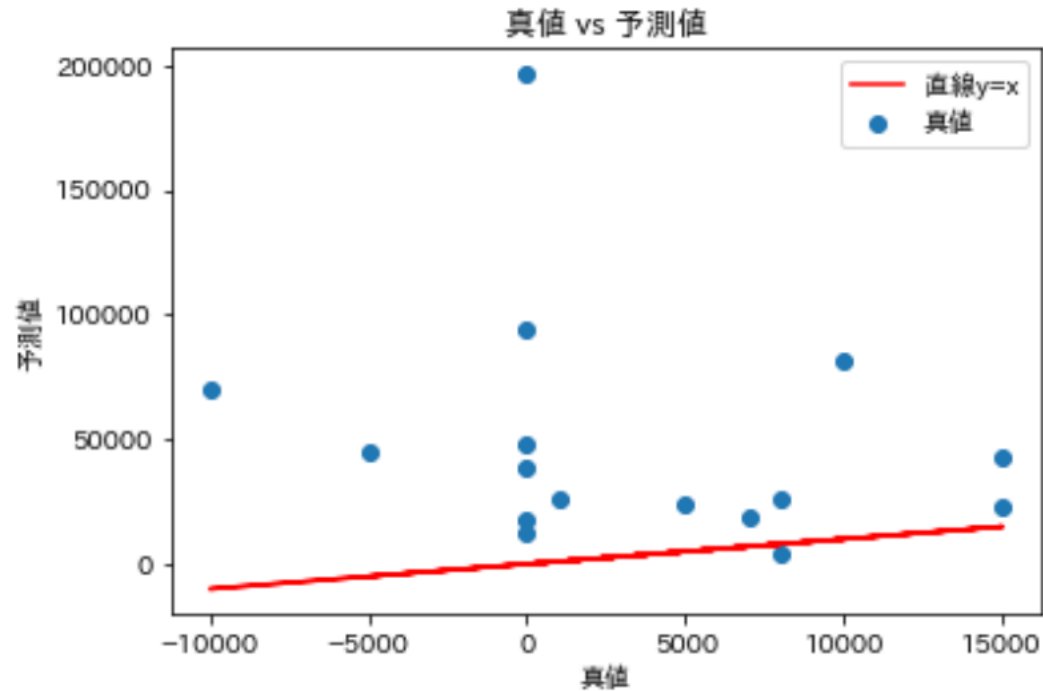
	説明変数	重要度
0	駅からの距離	0.246575
1	数量_前年	0.187215
2	町丁名	0.180365
3	年度	0.132420
4	地積	0.111872
5	数量	0.077626
6	数量_前年差分	0.047945
7	駅名	0.015982

特徴量重要度（最大値1.0）

4. 統計解析（数値予測②：地価差分を予測）

評価価格_1年前差分（地価）

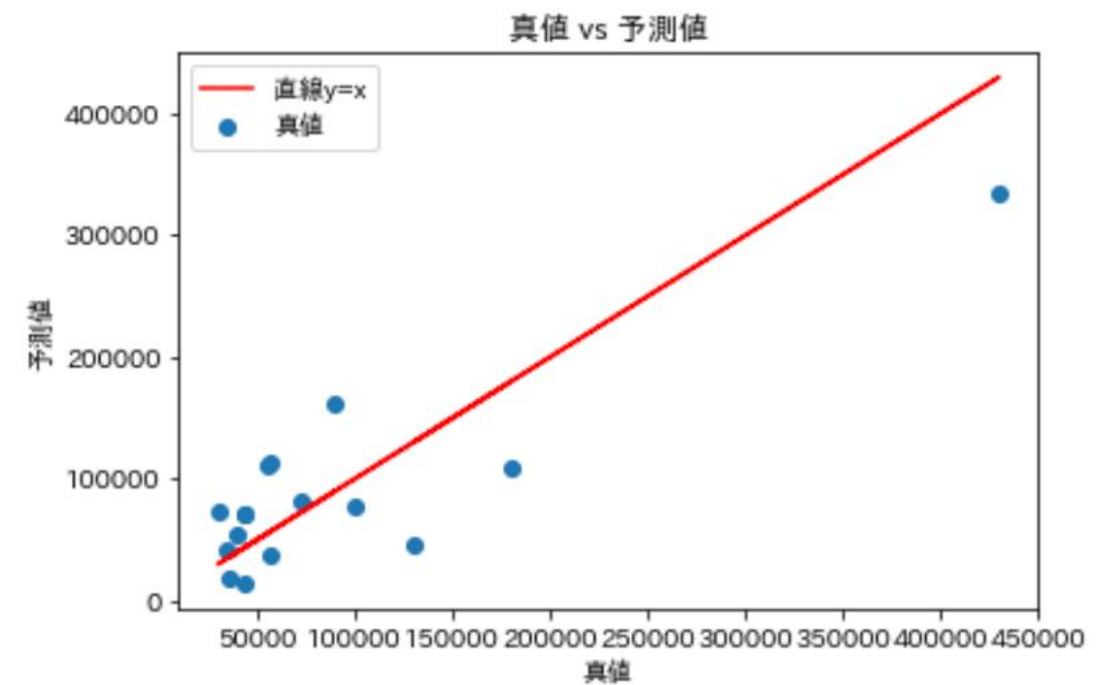
適切な予測モデル作成が困難



真値 と 予測値 の関係グラフ

評価価格_3年前差分（地価）

ある程度適切なモデルを作成可能



真値 と 予測値 の関係グラフ

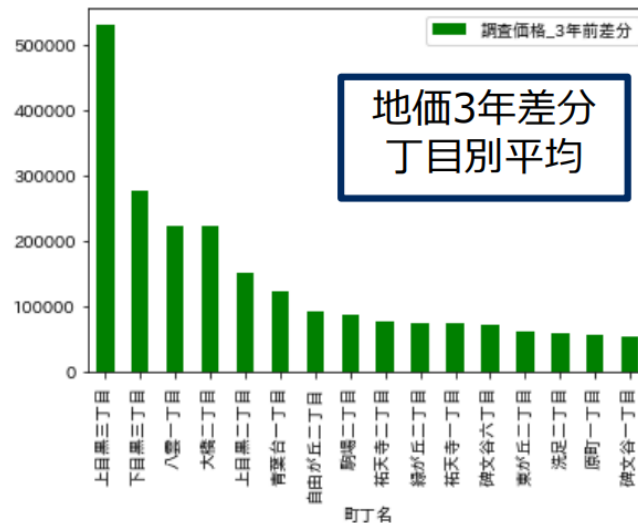
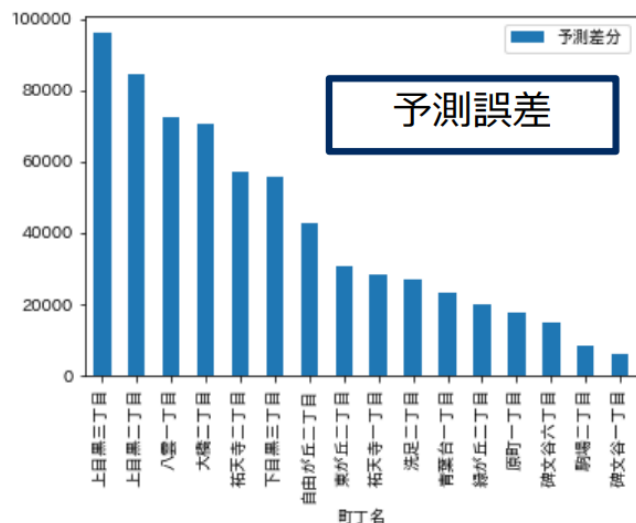
→ 時系列的な説明変数を増やしたことによる改善がみられる

データ活用のプロセスの試行 考察

作成協力
日本電気株式会社
本橋様

5. 考察（地価3年前差分の予測結果）

真値-予測値 誤差	所属する丁目	傾向	提案
4万円以内	碑文谷一丁目, 駒場二丁目, 碑文谷六丁目, 原町一丁目, 緑ヶ丘二丁目, 青葉台一丁目, 洗足二丁目, 祐天寺一丁目, 東が丘二丁目	予測誤差が小さく、今後の地価の変動を予測しやすい。	地価変動予測を元にした企業の新規店舗配備計画、引っ越し付帯サービスのマーケティングなど
4万～6万円	自由が丘二丁目, 下目黒三丁目, 祐天寺二丁目	予測誤差がやや大きい。地価の変動は比較的小さい。	他地区データや粒度の細かいデータの追加による予測モデル精度向上
6万円以上	大橋二丁目, 八雲一丁目, 上目黒二丁目, 上目黒三丁目	予測誤差が大きく、地価の変動を予測しづらい。	丁目別に特化したデータ収集・予測分析



丁目別人口や、今後整備しうるデータは、**民間企業向けのマーケティング基礎データ**等として、用いることができる可能性が高い。

提供された丁目別人口については、対象の地域の限定性があったものの、**場所及び時間の粒度の細かさを活かしたマーケティング分析・需要予測**などに用いることができる可能性が高い。

一方、居住人口だけでわかる産業影響は限定的であるため、今後、多様なデータを加えた分析ができると、さらに発展する。

1. 本事業の概要と目的

2. データの価値 ～なぜ使いやすいデータが必要なのか～

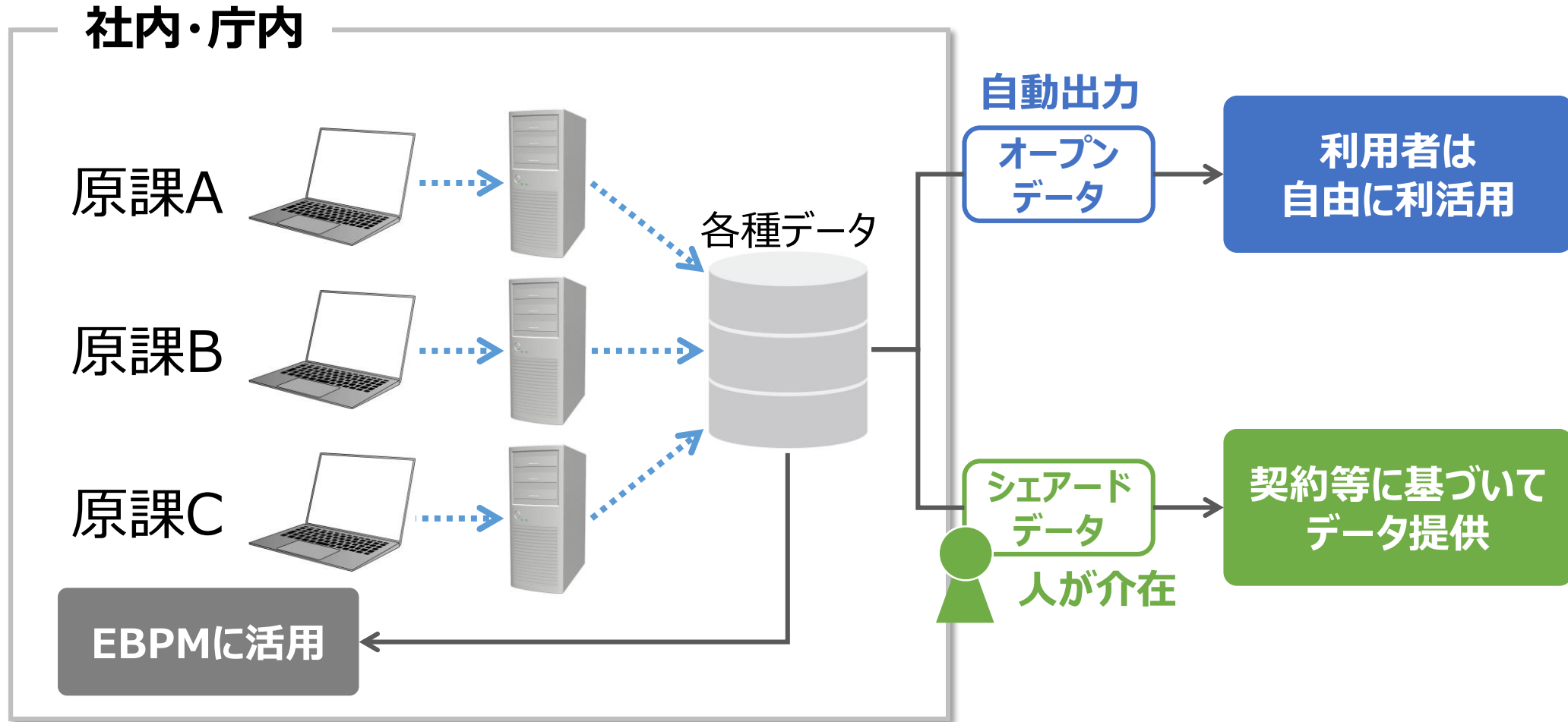
3. データ活用のプロセス

4. データ整備・公開のプロセス

5. マニュアルの利用について

データ整備→公開の理想

入力時点でデータ利活用を想定したフォーマットでのデジタル化が望ましい



しかし、システム更改や業務改善に向けたシステム構築には時間が必要

現時点でデータを作るときの心掛け 1/2

機械判読性の高いデータ提供のために、Excelファイルのレイアウトに気をつける

1つのセルには1データ

1セルにデータが複数あると、後で利用しにくいから控える。

✖

品名	
仕入額	273 (平成27年度) , 434 (平成28年度) , 549 (平成29年度) , 626 (平成30年度) , 741 (平成31年度)
出荷額	173 (平成27年度) , 1234 (平成28年度) , 3449 (平成29年度) , 3728 (平成30年度) , 3911 (平成31年度)

○

年度	仕入額	出荷額
平成27年度	273	173
平成28年度	434	1234
平成29年度	549	3449
平成30年度	626	3728
平成31年度	741	3911

✖

企業番号	売上金額 (単位:円)
総計	1189388 391445 (305942)
A	59394 89565 (89882)
B	154138 59469 (45922)
C	95522 44098 (35210)
D	95883 22258 (29297)
E	593993 37952 (35916)
F	27456 15289 (14820)
G	119985 115499 (11029)
H	84819 39817 (32997)

○

企業番号	売上金額	費用総額
総計	1189388	391445
A	59394	89565
B	154138	59469
C	95522	44098
D	95883	22258
E	593993	37952
F	27456	15289
G	119985	115499
H	84819	39817

結合はやめよう

複数のセルを結合するのも、使いにくい原因になります。

✖

品名	品名	品名
品名		

○

品名	品名	品名
品名	品名	品名

✖

品名	品名	品名
品名		

○

品名	品名	品名
品名	品名	品名

数値と文字列は分ける

数値と文字は混ぜるな危険。文字列として認識されます。

✖

サンプル	単価	前回増分	生産台数
サンプル1	10,000円	130	12000
サンプル2	9,190円	▲30	29,000
サンプル3	8,000円	▲30	37,000
サンプル4	7,545円	500	49,000
SUM関数	0	630	0
計算式	=VALUE	=VALUE	121000

○

サンプル	単価	前回増分	生産台数
サンプル1	10000	130	12000
サンプル2	9190	-30	29000
サンプル3	8000	-30	37000
サンプル4	7500	500	49000
SUM関数	34950	80	121000
計算式	=VALUE	=VALUE	121000

↑黄色の行は関数で合計を表示した例

✖

仕入台数	在庫台数
サンプル1	5,000 1,300
サンプル2	4,000 1,800
サンプル3	3,000 1,000
サンプル4	2,000 1,500

○

仕入台数	在庫台数
サンプル1	5000 1300
サンプル2	4000 1800
サンプル3	3000 1000
サンプル4	2000 1600

不要な空白/改行はNG

スペースや改行は、正しくデータが読み取れない原因に。

✖

分類	総数	事業所数	企業数
合計	900	450	450
A	200	100	100
B	300	150	150
C	400	200	200

○

分類	総数	事業所数	企業数
合計	900	450	450
A	200	100	100
B	300	150	150
C	400	200	200

✖

業別名	出荷台数	在庫台数
業別 A-1	429	756
業別 A-2	321	649
業別 A-3	384	439
業別 A-4	408	775

○

業別名	出荷台数	在庫台数
業別 A-1	429	756
業別 A-2	321	649
業別 A-3	384	439
業別 A-4	408	775

現時点でデータを作るときの心掛け 2/2

日頃よりExcelを正しく使うことがデータ整備・利活用の第一歩

1シート1テーブル

扱いやすいように1シートに表は1つ。複数ある時は分割。

✖ 1シートに複数の表を掲載

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

○ 1つの表を1シートに分割

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

データを分断しない

不必要な表の分離は機械判読を阻害します。

✖ データが分断

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

○ 空白列を削除した状態

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

✖ 表を折り返し

品名	単価	数量	金額
りんご	1000	1000	100000
バナナ	500	2000	100000
みかん	800	1250	100000
梨	1500	666	100000
ぶどう	2000	500	100000

項目名はそのままだ

項目名が省略されると機械が自動で判読できなくなります。

✖ 同じ名称を空白で省略

品名	単価	数量
りんご	429	756
	321	648
	384	438
	408	775

○ 省略せずに入力した状態

品名	単価	数量
りんご	429	756
りんご	321	648
りんご	384	438
りんご	408	775

オブジェクトは使わない

オブジェクトは扱いにくいいため、セル入力を基本としましょう。

✖ オブジェクトを使用している

名称	内容	数量
商品 A	サンプル1	100
	サンプル2	200
	サンプル3	300

○ オブジェクトを削除した状態

名称	内容	数量
商品 A	サンプル1	100
商品 A	サンプル2	200
商品 A	サンプル3	300

どのようなデータから着手すべきか

他のデータと比較や結合できるデータは、相乗的に価値が出る

見やすい

推計人口

(平成26年12月1日現在)

平成26年12月19日
静岡県企画広報情報統計局 統計調査課

1 概要
平成26年12月1日現在の静岡県の人口は、**3,695,868人** となり、前月と比べ、**1,266人の減少**となった。
内訳をみると、自然動態が734人の減少（出生2,341人、死亡3,075人）、社会動態が532人の減少（転入8,448人、転出8,978人）となっている。
なお、前年12月1日（3,713,549人）と比べ、17,681人の減少となっている。

2 人口と世帯の推移

区分	人口（日本人・外国人）及び世帯数				前5年間の純増減
	総数	男	女	世帯数	
H 2. 10 .1	3,870,840	1,808,951	1,861,889	1,117,893	96,...
国 勢 7. 10 .1	3,737,689	1,841,947	1,895,742	1,204,189	88,3...

XLSファイル

市区町名	H26.12.1人口総数(人)	男(人)	女(人)	世帯数(世帯)	人口総数(日本人)(人)	男(日本人)(人)	女(日本人)(人)
県計	3,695,868	1,818,026	1,877,842	1,435,013	3,634,890	1,788,845	1,846,045
熱海市	37,558	16,756	20,802	19,316	37,319	16,681	20,638
伊東市	69,069	32,350	36,739	30,926	68,110	31,914	36,196
下田市	23,337	11,227	12,110	10,524	23,155	11,164	11,991
伊豆市	31,737	15,055	16,682	12,266	31,610	15,015	16,595
伊豆の国市	48,721	23,278	25,443	19,327	48,205	23,045	25,160
東伊豆町	13,041	6,148	6,893	5,812	12,994	6,128	6,866
河津町	7,542	3,585	3,957	3,016	7,519	3,581	3,938
南伊豆町	8,834	4,206	4,628	3,603	8,803	4,197	4,606
松崎町	6,926	3,254	3,672	2,895	6,905	3,245	3,660
西伊豆町	8,481	3,993	4,488	3,734	8,422	3,977	4,445
沼津市	193,661	94,581	99,080	80,225	191,080	93,457	97,623
三島市	110,508	54,046	56,462	45,691	109,665	53,684	55,981
富士宮市	131,255	63,727	67,528	49,667	129,596	63,727	65,869
富士市	246,232	121,182	125,050	100,540	246,232	121,182	125,050
御殿場市	87,072	42,274	44,798	33,000	87,072	42,274	44,798
裾野市	52,496	25,479	27,017	20,000	52,496	25,479	27,017
函南町	38,021	19,600	20,421	15,466	38,021	19,600	20,421
清水町	32,565	16,083	16,482	12,565	31,289	15,603	15,686

二次利用しやすい

CSVファイル



地域・年齢別人口
や各種施設情報は
活用例が多い

自治体のデータ整備の取組状況に応じた、整備モデルのマニュアル化を実施

整備モデル	対象データ	協力自治体	マニュアルの内容
①画像・PDFから変換	地域・年齢別人口	板橋区	PDFのみでしか取り出せないデータがある自治体向けのマニュアル
②エクセルデータからの変換 (緯度経度クレンジング含まず)	介護サービス事業所の一覧	八王子市	エクセルデータ化まではできているが、エクセルデータのフォーマット統一を目的とした自治体向けのマニュアル
	医療機関の一覧	板橋区	
	医療機関の一覧	三鷹市	
	避難関連施設情報	八王子市	
③エクセルデータからの変換 (緯度経度クレンジング含む)	介護サービス事業所の一覧	三鷹市	エクセルデータの中で、住所等のフォーマット統一ができしており、緯度経度情報を追加し正規化するマニュアル
	医療機関の一覧	八王子市	
	バリアフリートイレ情報 (※)	港区	
④複数区のデータ変換・統合	地域・年齢別人口	品川区 目黒区 渋谷区	複数の自治体のオープンデータを変換・統合し、民間事業者の中で流通しやすい形式にするためのマニュアル

1. 本事業の概要と目的

2. データの価値 ～なぜ使いやすいデータが必要なのか～

3. データ活用のプロセス

4. データ整備・公開のプロセス

5. マニュアルの利用について

実際にマニュアル使ってみました ～スキルセット～

下記スキルセットで、PowerQueryを使ってデータ整備を実施

Excelスキル中級

- フィルターを使いこなせる
- グラフをキレイに作れる
- ピボットテーブルを使える
- △ 関数は（ググれば）何とか使える
- × マクロ組めない
- × PowerQuery知りませんでした

その他情報

社会人歴：17年

通信キャリア（民間）で
15年ほど勤務

企画畑に従事

プログラムは書けない

実際にマニュアル使ってみました ～整備前～

Excel 2016 画面キャプチャ。タイトル: totizinkou (1).xlsx - Excel。ユーザー: 長船 光祐。

6. 住民基本台帳による町丁目別世帯数及び男女別人口

町丁目	平成31年1月1日現在				令和2年1月1日現在				町丁目	平成31年1月1日現在			
	世帯数	人口			世帯数	人口				世帯数	人口		
		総数	男	女		総数	男	女			総数	男	女
総数	309 133	#####	#####	#####	#####	#####	#####	#####	東新町2丁目	2 199	4 505	2 232	2 270
									上坂元1丁目	1 203	1 986	965	1 020
坂元1丁目	4 671	7 345	3 582	3 763	4 742	7 487	3 593	3 894	上坂元2丁目	2 711	4 599	2 208	2 390
坂元2丁目	3 792	5 722	2 812	2 910	3 810	5 754	2 815	2 939	上坂元3丁目	1 391	2 190	1 053	1 130
坂元3丁目	3 124	5 411	2 713	2 698	3 207	5 519	2 745	2 774	清水町	3 767	6 465	3 192	3 270
坂元4丁目	3 174	5 051	2 551	2 500	3 207	5 071	2 565	2 506	蓮沼町	3 011	5 113	2 524	2 580
加賀1丁目	1 941	5 102	2 450	2 652	1 948	5 113	2 474	2 639	大原町	2 313	4 500	2 227	2 270
加賀2丁目	2 105	4 529	2 168	2 361	2 112	4 533	2 159	2 374	泉町	1 841	3 252	1 673	1 570
大山東町	3 755	5 848	2 860	2 988	3 765	5 824	2 874	2 950	宮本町	3 116	5 309	2 649	2 660
大山金井町	3 520	5 519	2 812	2 707	3 616	5 599	2 811	2 788	志村1丁目	2 877	5 188	2 485	2 700
熊野町	2 647	4 425	2 228	2 197	2 711	4 496	2 263	2 233	志村2丁目	2 104	4 308	2 090	2 210
中丸町	3 573	6 205	3 157	3 048	3 619	6 245	3 150	3 095	志村3丁目	1 865	3 227	1 570	1 630
南町	2 790	4 603	2 347	2 256	2 815	4 630	2 345	2 285	坂下1丁目	3 212	5 778	2 716	3 060
稲荷台	1 384	2 702	1 347	1 355	1 468	2 919	1 443	1 476	坂下2丁目	2 911	5 835	2 753	3 080

実際にマニュアル使ってみました ～整備前～

The screenshot shows an Excel spreadsheet with the following data:

6. 住民基本台帳による町丁目別世帯数及び男女別人口										
平成31年1月1日現在					令和2年1月1日現在					
町丁目	世帯数		人口		世帯数	人口		世帯数	人口	
	総数	男	女	総数		男	女		総数	男
総	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####
坂元1丁目	4 671	7 345	3 582	3 763	4 742	7 48	3 582	3 763	4 742	7 48
坂元2丁目	3 792	5 722	2 812	2 910	3 810	5 754	2 812	2 910	3 810	5 754
坂元3丁目	3 124	5 411	2 713	2 698	3 207	5 519	2 713	2 698	3 207	5 519
坂元4丁目	3 174	5 051	2 551	2 500	3 207	5 071	2 551	2 500	3 207	5 071
加賀1丁目	1 941	5 102	2 450	2 652	1 948	5 113	2 450	2 652	1 948	5 113
加賀2丁目	2 105	4 529	2 168	2 361	2 112	4 533	2 168	2 361	2 112	4 533
大山東町	3 755	5 848	2 860	2 988	3 765	5 824	2 860	2 988	3 765	5 824
大山金井町	3 520	5 519	2 812	2 707	3 616	5 599	2 812	2 707	3 616	5 599
熊野町	2 647	4 425	2 228	2 197	2 711	4 496	2 228	2 197	2 711	4 496
中丸町	3 573	6 205	3 157	3 048	3 619	6 245	3 157	3 048	3 619	6 245
南町	2 790	4 603	2 347	2 256	2 815	4 630	2 347	2 256	2 815	4 630
稲荷台	1 384	2 702	1 347	1 355	1 468	2 919	1 347	1 355	1 468	2 919

セルの結合がある

表が横にもある

文字の間にスペース

実際にマニュアル使ってみました ～作業後～

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J
1	町丁目	平成31年1月	平成31年1月	平成31年1月	平成31年1月	令和2年1月	令和2年1月	令和2年1月	令和2年1月	1日現在女
2	坂元1丁目	4671	7345	3582	3763	4742	7487	3593	3894	
3	坂元2丁目	3792	5722	2812	2910	3810	5754	2815	2939	
4	坂元3丁目	3124	5411	2713	2698	3207	5519	2745	2774	
5	坂元4丁目	3174	5051	2551	2500	3207	5071	2565	2506	
6	加賀1丁目	1941	5102	2450	2652	1948	5113	2474	2639	
7	加賀2丁目	2105	4529	2168	2361	2112	4533	2159	2374	
8	大山東町	3755	5848	2860	2988	3765	5824	2874	2950	
9	大山金井町	3520	5519	2812	2707	3616	5599	2811	2788	
10	熊野町	2647	4425	2228	2197	2711	4496	2263	2233	
11	中丸町	3573	6205	3157	3048	3619	6245	3150	3095	
12	南町	2790	4603	2347	2256	2815	4630	2345	2285	
13	稻荷台	1384	2702	1347	1355	1468	2919	1443	1476	
14	仲宿	4320	7581	3730	3851	4365	7627	3732	3895	
15	氷川町	2502	4346	2108	2238	2627	4533	2207	2326	
16	栄町	2409	4136	2050	2086	2423	4129	2061	2068	
17	大山町	2739	4575	2242	2333	2756	4593	2288	2305	
18	大山西町	2713	4552	2198	2354	2870	4784	2288	2496	

実際にマニュアル使ってみました ～作業後～

Excel screenshot showing a table with columns A through J. The table contains data for various locations (町丁目) and years (平成31年1月, 令和2年1月). The formula bar shows the active cell A2 contains the text "坂元1丁目".

	A	B	C	D	E	F	G	H	I	J
1	町丁目	平成31年1月	平成31年1月	平成31年1月	平成31年1月	令和2年1月	令和2年1月	令和2年1月	令和2年1月	日現在女
2	坂元1丁目	4671	7345	3582	3763	4742	7487	359	3894	
3	坂元2丁目	3792	5722	2812	2910	3810	5754		2939	
4	坂元3丁目	3124	5411	2713	2698	3207			2774	
5	坂元4丁目	3174	5051	2551	2500	3207			2506	
6	加賀1丁目	1941	5102	2450	2652	1948			2639	
7	加賀2丁目	2105	4529	2168	2361	2112	4533	2139	2374	
8	大山東町	3755	5848	2860	2988	3765	5824	2874	2950	
9	大山金井町	3520	5519	2812	2707	3616	5599	2811	2788	
10	熊野町	2647	4425	2228	2197	2711	4496	2263	2233	
11	中丸町	3573	6205	3157	3048	3619	6245	3150	3095	
12	南町	2790	4603	2347	2256	2815	4630	2345	2285	
13	稻荷台	1384	2702	1347	1355	1468	2919	1443	1476	
14	仲宿	4320	7581	3730	3851	4365	7627	3732	3895	
15	氷川町	2502	4346	2108	2238	2627	4533	2207	2326	
16	栄町	2409	4136	2050	2086	2423	4129	2061	206	
17	大山町	2739	4575	2242	2333	2756	4593	2288	2305	
18	大山西町	2713	4552	2198	2354	2870	4784	2288	2496	

列名の説明
を統合

1シート
1データ

実際にマニュアル使ってみました ～作業結果～

作業内容

作業時間：5時間

作業内容：

ファイル取得 → PowerQueryに読み込む → 作業する → 事前準備が足りないことに気づく (3時間)
→ 事前準備作業 → PowerQueryに読み込む → 作業する → データ整備完了 (2時間)

マニュアルへの所感

- PowerQueryで**データを読み込むところまではスムーズ**
- その後の作業は、マニュアルを**そのまま踏襲すればできる、というわけではない**
- マニュアルを使い、試行錯誤することで、「Power Query」の本質を理解ができ、**データの違いによる、作業の違いを吸収できる**
- 実際のデータ整備作業としては、2時間程度（慣れれば短縮もできそう）

デジタル化社会に不可欠な
キレイなデータのために
データ整備へチャレンジをお願いします

EOF