

# 1 EBPMを動かすための準備 -視点と思考を整える

## データ利活用の実践に向けて

EBPMにおけるデータ利活用は、単なる統計処理や数値の整理ではなく、「意思決定のプロセスそのもの」です。データを活かすということは、数字を扱う技術よりも、根拠に基づいて考え、判断し、行動する力を育てることを意味します。

EBPMは特別な専門知識を持つ人だけの取り組みではありません。日々の業務の中で、誰もが小さな工夫から始めることができます。例えば、既存データを活用して現状を見える化する、過去の施策を振り返る、住民の声を整理して仮説を立てる——これらも全てEBPMの実践です。

本ガイドブックでは、データ利活用を「難しい作業」ではなく、「チームで考え、進める意思決定の道具」として捉えることを提案します。自治体の現場において、分析のプロセスやツールを一部の専門職に委ねるのではなく、目的を定めた上で、企画・総務・現場部門など、多様な職員が協働しながら進めていくことが重要です。

データ利活用を“組織の一部の活動”ではなく、“チーム全体で共有し、成長させる文化”として根づかせること——それが、本書が目指すEBPM実践の第一歩です。



# 1 EBPMを動かすための準備 - 視点と思考を整える

## EBPMとは - データで考える自治体の新しいあたりまえ

### 01 EBPMの意味と目的

EBPM (Evidence-Based Policy Making) とは、「根拠に基づく政策立案」を意味します。政策の企画・実施・評価を行う際に、感覚や前例だけに頼るのではなく、統計データや調査結果などの「客観的な根拠」をもとに判断する考え方です。

その目的は、より効果的で説明責任のある自治体運営を実現し、住民一人ひとりの生活の質を高めることにあります。



### 02 自治体におけるEBPMの広がり

近年、中央官庁や地方公共団体のあらゆる分野でEBPMの取り組みが進んでいます。例えば、人口減少対策、空き家対策、医療・福祉施策、地域交通、教育分野など、複雑化する課題の解決には「現状をデータで正確に把握し、限られた資源をどこに重点配分するか」を見極める力が求められています。

総務省などの中央官庁によるオープンデータの進展により、誰もが統計データを使って政策を検討できる環境が整いつつあります。

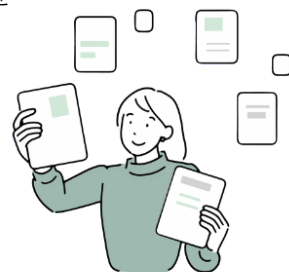
EBPMはすでに特別なことではなく、自治体にとって当たり前の手法として定着しつつあります。



### 03 既存データの活用が第一歩

EBPMを始めるために、必ずしも新しい調査や高度な分析が必要なわけではありません。自治体の現場には、すでに多くの「使えるデータ」が蓄積されています。

国や地方公共団体が実施する公的統計を活用することで、地域の現状を定量的に把握し、課題の優先度や施策の方向性を客観的に整理することが可能です。

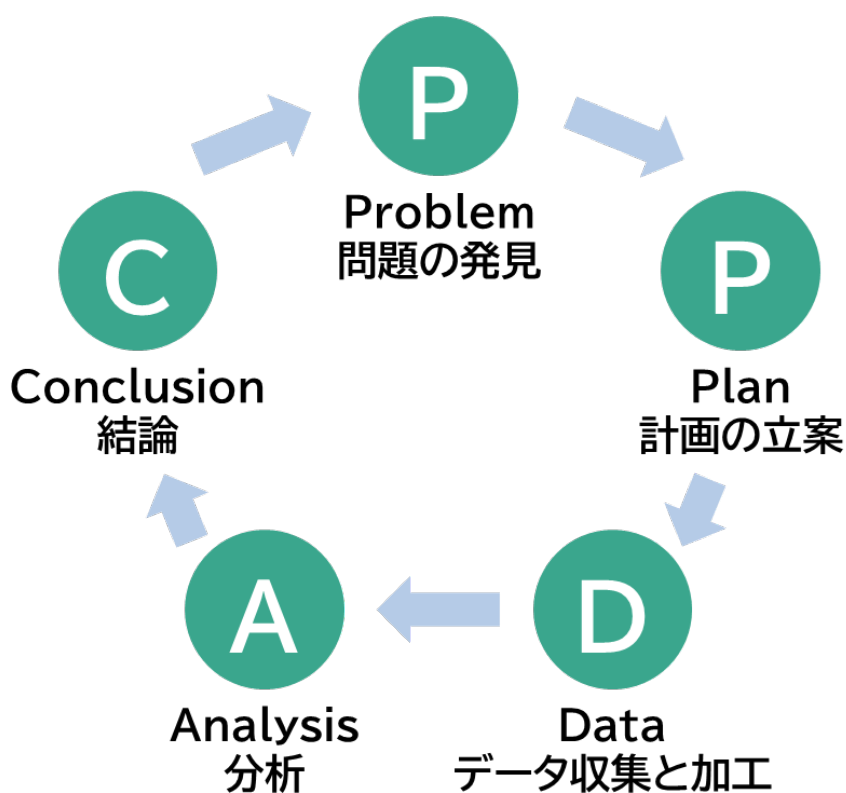


# 1 EBPMを動かすための準備 -視点と思考を整える

## データ利活用の基礎サイクル

EBPMを実践するにあたっては、課題の整理から政策判断に至るまでを一連のプロセスとして捉えることが重要です。

福祉、防災、産業振興、教育など、地方公共団体が直面する政策課題は多岐にわたりますが、データ利活用の基本的な考え方は共通しています。EBPMに有効な枠組みとして、本ガイドブックでは「PPDACサイクル」を用います。



PPDACは単なる分析手順ではなく、「問いを起点に、根拠をそろえ、判断と説明につなげる」ための運用フレームです。この運用フレームを基準に考えることで、データを場当たりの扱っただけではなく、課題解決のプロセスの中に位置づけて活用することが可能になります。特に自治体の実務では、結論そのものだけでなく、「なぜその判断に至ったのか」を説明できることが重要です。前提条件や用いたデータ、言える範囲と限界を共有できてこそ、組織内外の合意形成につながります。

# 1 EBPMを動かすための準備 -視点と思考を整える

## EBPMに役立つ「統計の基礎」

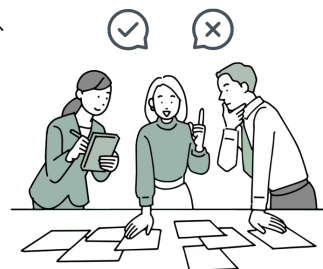
### なぜ統計の基礎が必要なのか

前ページまでに、データ利活用の基本的な流れとしてPPDACサイクルを用いた分析プロセスを紹介しました。これらを自治体の実務で使用していくためには、データを正しく読み取り、考えるための基礎的な知識が欠かせません。

EBPMでは、必ずしも高度な分析手法を使う必要はありませんが、

- データの見方
- 数値の意味の捉え方
- 比較結果の読み取り方

といった統計の基本的な考え方を理解しているかどうかで、判断の質は大きく変わります。



ここでは、EBPMを実践する上での前提として、統計の基礎となる考え方と最低限の知識を整理します。

### 統計とは

「統」はまとめる、「計」ははかるという意味を持ち、集めた情報を整理し数量的に把握することを指します。

### 統計データの分析とは

集めた統計データを整理し、比較や要約を通じて、全体の傾向や特徴を捉えることです。すなわち個々の数値を見るのではなく、統計データ全体から意味のある情報を引き出し、判断や説明に役立てることが目的です。

「統計」と聞くと、専門的で難しいイメージを持つかもしれませんが、初歩的な段階では、複雑な数式の理解や高度な分析手法を身につけている必要はありません。

EBPMの実践において重要なことは、難しい計算ができることよりも、結果を正しく解釈できることです。統計データを扱う際の基本的な考え方や注意点を理解することで、EBPMの実践は、十分に可能です。

# 1 EBPMを動かすための準備 -視点と思考を整える

## データとは

### 01 データとは

データとは、事実や状況を記録した情報を指します。効率的にデータを活用する上では、単に情報があるだけでなく、機械が読み取り、処理できる形であるかどうか重要です。

近年、自治体端末の性能向上や事務業務の電子化により、取得・活用できるデータが急速に増えています。EBPMを実践するためには、「データとは何か」「どのような種類のデータがあるのか」を共通理解として押さえておく必要があります。

### 02 データの基本的な分類

世の中に存在するデータは、大きく「構造化データ」と「非構造化データ」に分けられます。

#### 構造化データとは

行と列からなる表形式で整理できるデータで、数値や文字が一定のルールに従って並んでいるものを指します。

例：統計表、業務システムのデータ、アンケート結果など

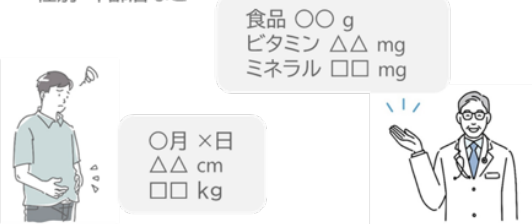
#### 非構造化データとは

文章、画像、音声、動画など、そのままでは表形式に整理できないデータを指します。文字起こしやタグ付け、分類などの前処理を行うことで、分析に活用できる形へ変換可能です。

例：文章、画像、音声、動画データなど

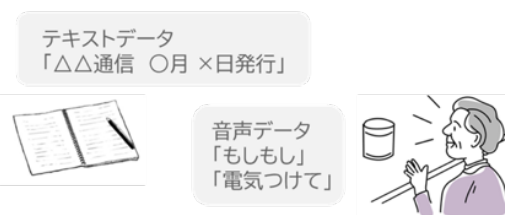
#### 構造化データの例

- 表やデータベース向き
- 例：身長・体重、検査値、売上・在庫、来店日・時刻、性別・年齢層など



#### 非構造化データの例

- 形が決まっていないデータ
- 例：文章、画像、音声、動画など



EBPMにおいては、どのような課題に対して、どのような種類のデータが必要なのかを意識しながら、適切に使い分けることが重要です。

# 1 EBPMを動かすための準備 -視点と思考を整える

## データとは

### 03 機械判読性の重要性

データ利活用の出発点となるのが、「機械判読性」です。

#### 機械判読性とは

データが、xlsx形式やCSV形式\*などの形式で整理され、検索、集計、比較といった処理を機械的に行える状態にあることです。

\*CSV：「，（カンマ）」で区切ったデータのこと。

例えば、PDFや画像として公開されている資料は、人が目で読むことはできますが、そのままのデータ形式では集計や分析に適していません。一方で、同じ内容が、xlsx形式のデータとして提供されていれば、分析や可視化、他のデータとの結合が容易になります。

EBPMを進める上では、「データが存在するかどうか」だけでなく、「分析に使える形になっているか」という点を常に意識することが重要です。



### 04 オープンデータ

こうした機械判読性を前提として、近年、重要性が高まっているのが「オープンデータ」です。

#### オープンデータとは

中央官庁や地方公共団体などが保有するデータのうち、誰もがインターネット等を通じて自由に利用、加工、再配布できる形で公開されているデータのことです。

オープンデータは、利用目的を問わず使用でき、機械判読可能な形式で提供されています。

また、無償で利用できるため、自治体内部だけでなく、民間や研究機関、市民も含めた幅広い主体によるデータ利活用が可能です。



# 1 EBPMを動かすための準備 -視点と思考を整える

## データとは

### 04 オープンデータ

中央官庁等が提供するオープンデータ活用サービスとして、代表的なものは以下です。

名称	内容	出典
 政府統計の総合窓口	<ul style="list-style-type: none"><li>人口推計や国勢調査・家計調査・消費者物価指数・労働力調査</li></ul>	総務省統計局
	<ul style="list-style-type: none"><li>国や民間企業等が公表している統計データがグラフ等に加工されている</li></ul>	総務省統計局
 地域経済分析システム	<ul style="list-style-type: none"><li>産業構造や人流など、国や民間企業が保有するビッグデータを集約</li></ul>	経済産業省・内閣官房
Japan Dashboard	<ul style="list-style-type: none"><li>経済・財政・人口と暮らしに関する指標を収録</li><li>市町村別データも提供</li></ul>	デジタル庁
	<ul style="list-style-type: none"><li>不動産に関する情報を地図上で閲覧できるツール</li></ul>	国土交通省
	<ul style="list-style-type: none"><li>国立情報学研究所が運営しているデータベースサービス</li></ul>	国立情報学研究所

これらのサービスを活用することで、地方公共団体は一からデータを収集することなく、信頼性の高い統計データを活用することができます。



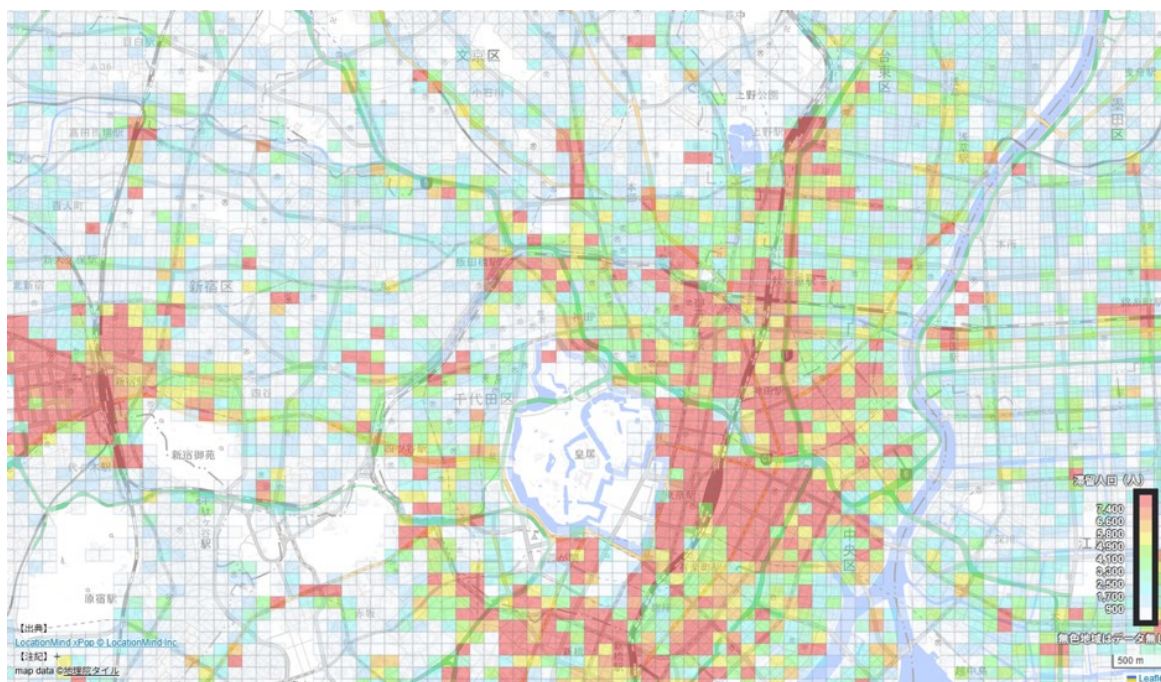
# 1 EBPMを動かすための準備 -視点と思考を整える

## データとは

### 05 具体的な活用例

公共交通分野では、e-StatやRESASから取得した統計データと、バスや鉄道の利用実態データを組み合わせて分析することが考えられます。

図のように、複数のデータを重ねて見ることで、利用者減少の背景や地域ごとの特徴を客観的に把握し、施策検討につなげることが可能になります。



オープンデータで取得できる情報の例

RESASで取得した滞留人口メッシュ（2024年・平日・午前8時・東京都千代田区周辺）

出典：[https://resas.go.jp/town-planning-staying-](https://resas.go.jp/town-planning-staying-mesh/)

[mesh/?pref=13&city=13101&tab=0&year=2024&month=all&hour=8&day=weekday&mesh=125&lat=35.694005388960925&lng=139.75364685058597&zoom=14&opacity=0.4](https://resas.go.jp/town-planning-staying-mesh/?pref=13&city=13101&tab=0&year=2024&month=all&hour=8&day=weekday&mesh=125&lat=35.694005388960925&lng=139.75364685058597&zoom=14&opacity=0.4)

このように、データ活用の基本的な考え方やオープンデータの仕組みは、公共交通分野に限らず、地方公共団体が取り組む、様々な政策課題に共通するものです。



# 1 EBPMを動かすための準備 -視点と思考を整える

## データの前処理

### 01 前処理とは

前処理とは、分析や可視化を行う前に、データを使える状態に整える作業のことです。多くの場合、データは分析目的ではなく、業務や記録のために収集され、複数の人や組織、システムを経由して蓄積されるため、形式や内容にばらつきがあります。

このため、分析に使用できる状態にデータを加工したり、分析に使用できるデータであるかを確認したりしないまま分析に用いると、意図しない誤解や誤った結論に繋がる恐れがあります。この工程は、こうしたリスクを防ぐとともに、EBPMや根拠ある判断を支える上で、重要な通過点と言えます。

※よく似た言葉に「データの加工」がありますが、全くの別物です。  
加工については次のページで紹介します。

### 02 前処理の必要性

前処理を行わずに分析を進めると、次のような問題が生じやすくなります。

- 実態と異なる傾向や数値が示される
- 同じデータを使っているのに、分析結果が人によって異なる
- 結果の理由を説明できず、判断に使えない
- 意思決定の根拠として信頼されない

**EBPMでは、なぜ、その結果を導いたのかを説明できることに加えて、他の人が見ても同じ理解に至ることが重要です。**  
そのため、分析の前提として前処理には丁寧さが求められます。



# 1 EBPMを動かすための準備 -視点と思考を整える

## データの事前処理

### 03 前処理の具体例

前処理の段階で行う作業の具体例を紹介します。

データの前処理	例
欠損値（空欄）の確認	<ul style="list-style-type: none"><li>・未入力、集計対象外、取得漏れなどの確認</li><li>・「0なのか、未記入なのか」を区別する</li></ul>
表記ゆれ・分類の統一	<ul style="list-style-type: none"><li>・例：「男性／男／M」などの表記を統一</li><li>・地域名・事業名・年度表記の揺れを修正</li></ul>
データ型の整理	<ul style="list-style-type: none"><li>・数値として扱うべき項目が文字列になっていないか</li><li>・日付形式が年度・年月日で混在していないか</li></ul>
異常値の確認	<ul style="list-style-type: none"><li>・明らかに実態と合わない数値（入力ミスなど）がない</li><li>・修正・除外・注記の判断を行う</li></ul>
分析目的に応じた加工	<ul style="list-style-type: none"><li>・年度別・地域別に集計し直す</li><li>・指標化（割合、1人当たり、前年差など）を行う</li></ul>

これらは高度な統計知識がなくても、データを丁寧に見ることで実施できる作業です。どれだけ高度な分析を行っても、前処理が不十分であれば、データの信頼性が失われ、その結果は判断の根拠になり得ません。

**前処理は、成果が見えにくく、とても手間のかかる工程です。**  
しかし、EBPMを実践につなげる重要なポイントであることを覚えておきましょう。

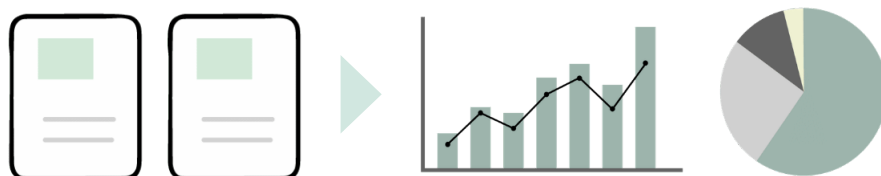


# 1 EBPMを動かすための準備 -視点と思考を整える

## データの加工

### 01 データ加工とは

前処理によって整えられたデータを、分析や可視化、判断に使える形へ変換する作業のことです。データは、そのままの形では、不必要な情報が含まれていたり、粒度が細かすぎたり、あるいは粗すぎたりすることもあり、そのままでは比較や傾向の把握が難しい場合があります。そのため、前処理を終えた後、目的に応じて整理や変換を行うことで、データは初めて意味のある情報として活用できるようになります。



### 02 データ加工の必要性

EBPMでは、「データがあること」よりも、「判断に使える形で示されていること」が特に重要です。

加工を行わないままでは、次のような状況に陥ります。

- 数値の大小は分かっていても、違いの意味が分からない
- 年度や地域ごとの比較ができない

データの加工は、問いや課題、目的に応じて、データの見方を整える工程であり、EBPMにおける「分析」と「判断」をつなぐ役割を果たします。



# 1 EBPMを動かすための準備 -視点と思考を整える

## データの加工

### 03 データ加工の具体例

データ加工の段階で行う具体例を紹介します。

データ加工	例
集計	<ul style="list-style-type: none"><li>・日別データを月別・年度別にまとめる</li><li>・個票データを地域別・区分別に集約する</li></ul>
指標化	<ul style="list-style-type: none"><li>・実数を割合、1人当たり、1件当たりなどに変換する</li><li>・単純な件数では見えにくい差を明確にする</li></ul>
分類・区分の設定	<ul style="list-style-type: none"><li>・年齢層、時間帯、距離帯などに分ける</li><li>・分析の視点に合わせてグループ化する</li></ul>
比較しやすい形への変換	<ul style="list-style-type: none"><li>・年度間の増減、前年差、平均との差を算出する</li><li>・他地域や過去との違いを把握できる形にする</li></ul>

これらは単なる見せ方の工夫ではなく、問いに対して、どのような情報が必要かを考え、判断に必要な視点から見ることができるように、データを加工する工程です。

前処理でデータを整え、加工によって意味のある形に変換することで、分析結果と判断は、初めてつながります。

# 1 EBPMを動かすための準備 -視点と思考を整える

## 基本統計量

### 01 基本統計量とは

データを活用する際、いきなり高度な分析手法を用いるのではなく「まずデータそのものがどのような特徴を持っているのか」を把握することから始めます。その作業が「基本統計量」の確認にあたります。

#### ポイント

基本統計量は、分析をする前の作業として、次のような点を確認するために用いられます。

- 全体としてどの程度の水準や分布を持つデータであるのか
- 値のばらつきが大きいのか/小さいのか
- 一部の極端な値に含まれていないか/一部の値が全体に影響を与えていないか

#### メリット

基本統計量を確認することには、次のようなメリットがあります。

- **データの全体像や傾向を短時間で把握できる**

平均値、中央値やばらつきを確認することで、データがどの程度の品質水準を持つデータであり、どのような分布をしているのかを大まかに理解することができます。

- **一部の極端な値やデータの偏りに左右されない判断が可能になる**

平均値だけでなく中央値や分布をあわせて見ることで、外れ値の影響を受けた誤解やミスリードを防ぐことができます。

- **次に行う分析や可視化の方向性を適切に定められる**

データのばらつきや分布の特徴を把握することで、どの指標（基本統計量）に注目すべきか、どのような可視化の手法や分析手法が適しているのかを判断しやすくなります。

# 1 EBPMを動かすための準備 - 視点と思考を整える

## 基本統計量

### 02 基本統計量の主な種類と考え方

#### ①位置を示す基本統計量（代表値）

位置を示す基本統計量は、データの中心的な位置を表す指標で、代表値とも呼ばれます。代表値の中の代表的なものとして、平均値、中央値、最頻値があります。

基本統計量	説明
平均値	全ての値を足し合わせてデータの個数で割った値で、全体的な水準を把握する際によく用いられます。一方で、極端に大きい値や小さい値が含まれている場合、その影響を強く受けるという特徴があります。
中央値	データを小さい順に並べたときに中央に位置する値で、外れ値の影響を受けにくい指標です。データの実態に近い「典型的な値」を把握したい場合には、平均値と併せて確認することが重要です。
最頻値	データの中で最も多く出現する値を示します。「最もよく現れる値」を把握する際に有効ですが、分布によっては複数存在する場合もあり、必ずしも全体の中心を示すとは限りません。

#### ②ばらつきを示す基本統計量

位置を示す指標だけでは、データの散らばり具合を把握することはできません。そこで用いられるのが、ばらつきを示す基本統計量です。代表的なものとして、分散と標準偏差があります。

基本統計量	説明
分散	各データが平均値からどの程度離れているかを示す指標です。しかし、計算過程で二乗を用いるため、元のデータとは異なる単位になります。
標準偏差	標準偏差は、平均値の周りにどの程度データが散らばっているかを示す指標です。分散（各データの平均からのずれを二乗して平均したもの）の平方根をとることで、元のデータと同じ単位に揃えた値で表されます。値が大きいほどばらつきが大きいことを意味します。平均値が同じであっても、標準偏差が異なれば、データの性質は大きく異なります。

# 1 EBPMを動かすための準備 -視点と思考を整える

## 基本統計量

### 02 基本統計量の主な種類と考え方

#### ③分布の形を示す基本統計量

基本統計量には、分布の形そのものを把握するための指標もあります。

基本統計量	説明
歪度	「わいど」と読み、分布が左右どちらに偏っているかを示し、0に近いほど左右対称であることを意味します。
尖度	「せんど」と読み、分布の中心への集中度合いを示す指標です。

これらの指標は、必ずしも日常的に使うものではありませんが、分析結果の解釈や、適切な分析手法を選択する際の参考情報として役立ちます。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 基本統計量

### 03 Excelを用いた基本統計量の算出方法

基本統計量は、Microsoft社製のExcelを用いて容易に算出することができます。Excelには「分析ツール」というアドイン機能があり、これを利用することで、複数の基本統計量を一括で確認することが可能です。 ※Microsoft Excel2013 以降のバージョンにて利用可能。

#### STEP 01

左上のバーの「ファイル」を開き、その後左下の「オプション」を開きます。



#### STEP 02

「アドイン」を選択します。画面下部の「管理」で「Excelアドイン」を指定して「設定」をクリックし、「分析ツール」にチェックを入れます。



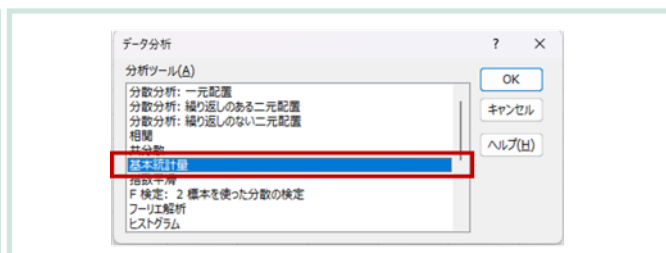
#### STEP 03

設定後、左上のバーの「データ」タブをクリックし開くと、右上に「データ分析」が表示されるようになります。



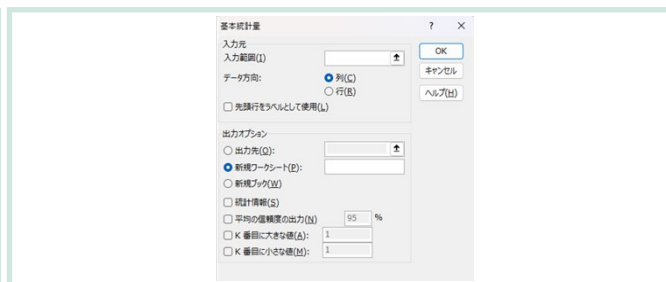
#### STEP 04

「データ分析」をクリックし、「基本統計量」を選択します。



#### STEP 05

入力範囲に対象となるデータを指定し、出力先を設定すると、平均値、中央値、最頻値、分散、標準偏差などの基本統計量が一覧で算出されます。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 基本統計量

### 03 Excelを用いた基本統計量の算出方法

個別に算出する場合には、AVERAGE（平均値）、MEDIAN（中央値）、MODE（最頻値）、VAR（分散）、STDEV（標準偏差）といった関数を用いることも可能です。目的やデータの規模に応じて、分析ツールと関数を使い分けることが重要です。

以下の表は、Excelの基本統計量の関数を表にまとめたものです。

項目	意味	Excelの関数
平均値	全ての個別データを足しあげて、その合計をデータの個数で割り算した値	AVERAGE
中央値	データを大きさの順に並べて、ちょうど真ん中にくる値	MEDIAN
最頻値	データがもっとも集中しているデータ値	MODE
標準偏差	データの散らばり度合いを表す値。値が大きいほど、データにばらつきがある	STDEV
分散	データの散らばり度合いを表す値。標準偏差の二乗	VARP
範囲	データの取る範囲 最大値-最小値	(MAX-MIN)
最小値	データの最も小さい値	MIN
最大値	データの最も大きい値	MAX
合計	データの合計	SUM
標本数	データの数	COUNT

基本統計量は、それ自体が政策の結論を示すものではありません。  
しかし、データ分析の前提となる重要な情報です。  
データの特徴を正しく理解しないまま分析を進めると、  
誤った解釈やミスリードにつながるおそれもあるので注意しましょう。



# 1 EBPMを動かすための準備 -視点と思考を整える

## データの可視化

### 01 データの可視化とは

文字や数値の集合であるデータを、グラフや図を用いて「目で見て分かる形」に変換することを指します。

データを活用するためには、数値をそのまま並べて確認するだけでは不十分です。データは、適切に可視化することで初めて、その特徴や傾向、違いが直感的に理解できるようになります。

#### 使用場面

データの可視化は、分析プロセスの中ではさまざまな場面で用いられます。

- 課題設定の段階：現状把握
- 分析の段階：結果の解釈
- 結論を示す段階：判断の根拠の説明

数値だけでは捉えにくい特徴や変化を理解したり、意思決定につなげたりするための基礎となります。



# 1 EBPMを動かすための準備 -視点と思考を整える

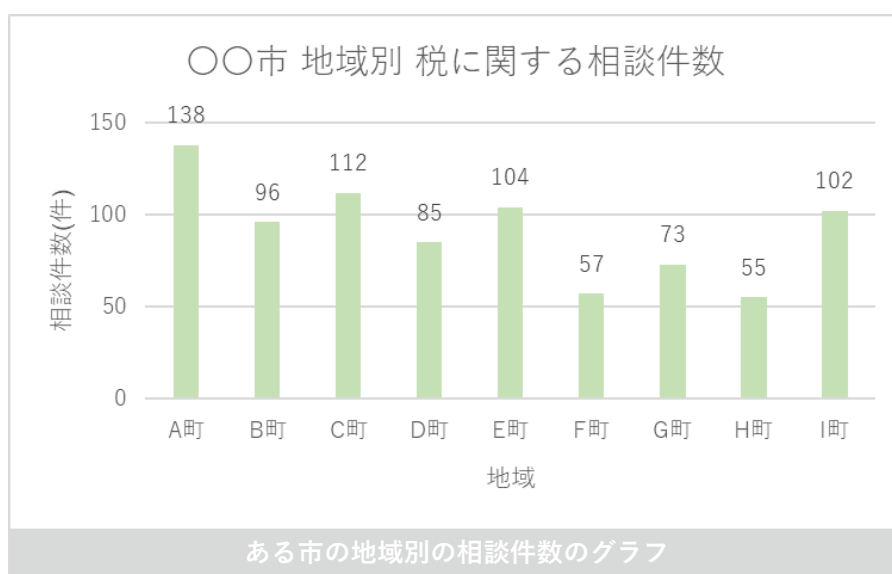
## データの可視化

### 02 主な可視化手法と活用の考え方

可視化の種類には、さまざまなものがありますが、ここでは主要な可視化の手法を紹介します。

#### ①棒グラフ

項目ごとの数量の違いを比較する際に最も基本的な用いられる可視化の手法です。同一の指標について、複数の地域、年度、属性、政策対象などを比較したい場合に適しています。



#### 活用場面

政策対象ごとの地域別の件数、年度別の実績値や利用回数などを比較することで、どの項目が多いのか、差がどの程度あるのかを直感的に把握できます。

#### 留意点

数量の大小を正確に伝えるため、軸の目盛りや原点の設定には注意が必要です。

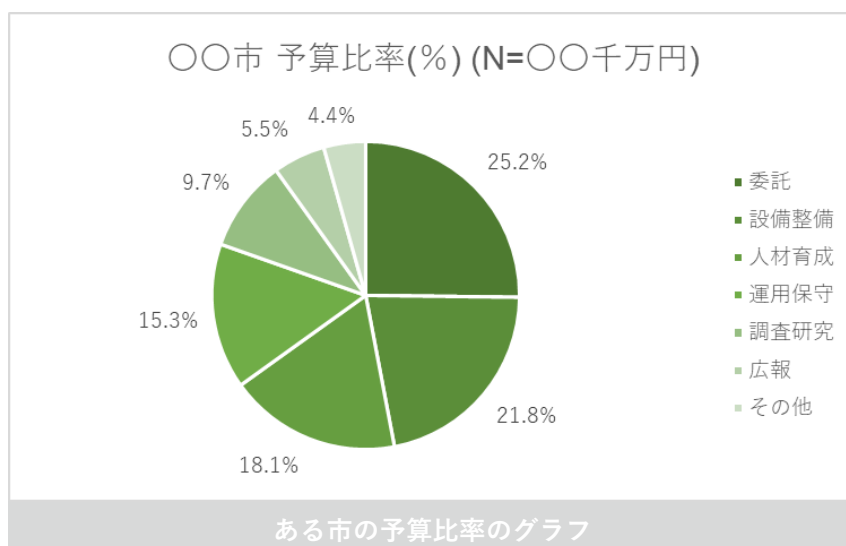
# 1 EBPMを動かすための準備 -視点と思考を整える

## データの可視化

### 02 主な可視化手法と活用の考え方

#### ②円グラフ

円グラフは、全体に占める各項目の割合を示す際に用いられる可視化の手法です。「全体の構成比」を把握したい場合に適しており、内訳のイメージを共有するときに有効です。



#### 活用場面

属性別の構成割合、政策別の予算配分、回答選択肢の割合などを示す際に用いられます。ただし、項目数が多い場合や、差が小さい場合には比較が難しくなるため、使用する場面を選ぶ必要があります。

#### 留意点

比率を表す際には、必ず全体の総数である「N数」を併記する必要があります。

# 1 EBPMを動かすための準備 -視点と思考を整える

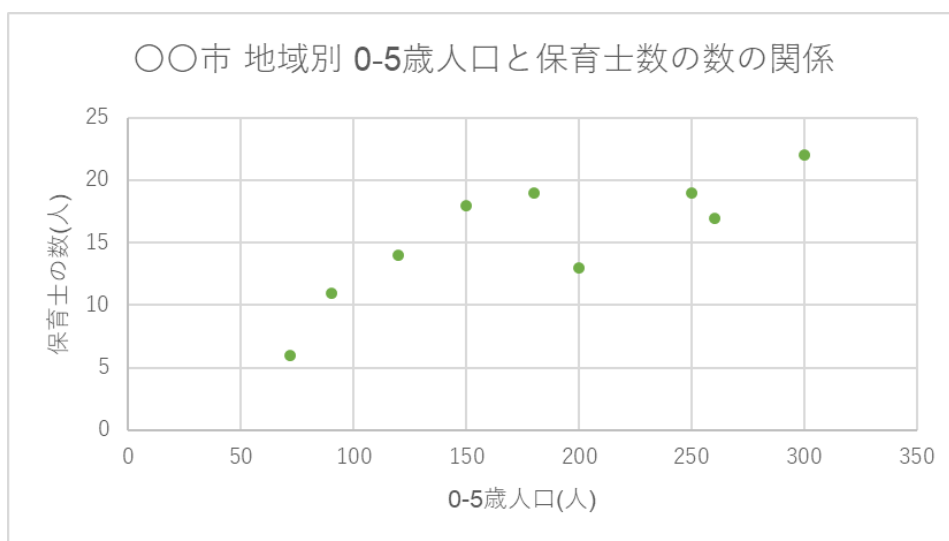
## データの可視化

### 02 主な可視化手法と活用の考え方

#### ③ 散布図

散布図は、二つの数値データの関係性を把握するための可視化手法です。

横軸と縦軸に異なる指標を取り、データを点として配置することで、両者の間に関係がありそうか否かを視覚的に確認できます。



#### 活用場面

人口規模と利用件数、支出額と成果指標、施策投入量と結果指標など、二つの量的データの関係を探索する際に用いられます。

散布図は因果関係を示すものではありませんが、次に検討すべき視点を見つけるための有効な手がかりとなります。

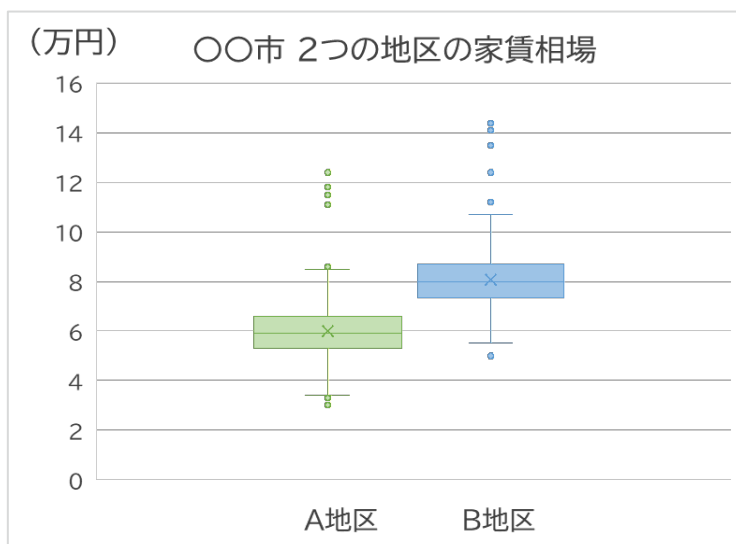
# 1 EBPMを動かすための準備 -視点と思考を整える

## データの可視化

### 02 主な可視化手法と活用の考え方

#### ④箱ひげ図

箱ひげ図は、データの分布やばらつきを把握するための可視化の手法です。中央値や四分位数 \*を用いて、データがどの範囲に集中しているか、外れ値が存在するかを視覚的に示します。



※四分位数とはデータを小さい順に並べて、4等分したものです。

#### 活用場面

複数の地域、複数の組織や政策対象ごとの数値分布を比較する場合、箱ひげ図を用いることで、平均値だけでは分からない、ばらつきの違いや極端な値の存在を把握することができます。基本統計量と併せて確認することで、データへの理解がより深まります。

# 1 EBPMを動かすための準備 -視点と思考を整える

## データの可視化

### 03 可視化を行う際の留意点

可視化は非常に有効な手段である一方、表現方法によっては誤解を招く可能性もあります。

#### 留意点

可視化を行う際は以下の点に注意してください。

- 軸の切り方
- 軸の目盛りの設定
- グラフの色使い
- グラフの恣意的な強調

データ分析においては、「分かりやすさ」だけでなく、「正確に伝えること」を重視する必要があります。

読み手が誤った解釈をしないかという視点を持ち、データの実態を誠実に表現しているかを確認することが重要です。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 自治体で行われる統計解析例

### 01 回帰分析とは

回帰分析とは、ある数値（目的変数）が、別の数値（説明変数）とどのような関係にあるかを分析し、予測や傾向把握を行う手法です。

分析対象となる結果を「目的変数」といい、目的変数に関連すると考えられる事象や値のことを「説明変数」といいます。



「〇〇が増えると△△はどう変わるのか？」  
「この条件のとき、結果はどのくらいになりそうか？」  
といった問いに答えるために使います。

### 公共交通における具体例

ある地方公共団体で「バスの利用者数が年々減っている原因は何か」を知りたいと考えました。

分析したい問い

「どの要因が、バスの利用者数に影響しているのか？」  
「条件が変わると、利用者数はどのくらい変化するのか？」

回帰分析で分かること  
(イメージ)

- ・運行本数が10本増えると、利用者数は平均で〇人増える
- ・雨が降ると、利用者が増える傾向にある
- ・人口の減少が、利用者数減少に強く影響している

このように回帰分析では、「どの要因が、どのくらい影響しているか」を数値で示すことができます。

なお、回帰分析は要因と結果の関係を数量的に整理する手法ですが、回帰分析で有意な関係が確認されたとしても、それが因果関係であるとは限りません。因果を論じるためには、追加の検討や分析設計が必要となります。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 自治体で行われる統計解析例

### 01 相関分析とは

相関分析とは、2つの事象がどのような関係性を持つかを把握する手法です。

相関係数と呼ばれる値を計算することで、関係性の強さを把握することができます。相関係数は、2つのデータの直線的な関係性の強さを数値化したもので、-1から1の値をとります。一方が増加すればもう一方も増加する関係を「正の相関がある」と呼び、このとき相関係数は1に近い値になります。

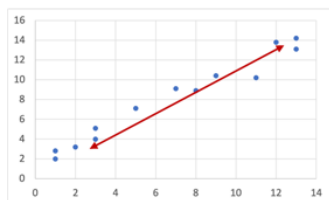
逆に、一方が増加すれば、もう一方が減少する関係を「負の相関がある」と呼び、このとき相関係数は-1に近い値になります。

- +1 に近い → 強い正の相関（両方とも増える）
- -1 に近い → 強い負の相関（一方が増えると他方は減る）
- 0 に近い → ほとんど関係がない

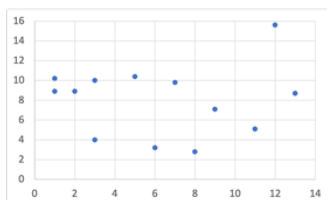
### 具体例

- 相関係数 0.8 → 強い正の相関
- 相関係数 -0.6 → 負の相関がある
- 相関係数 0.1 → ほぼ無関係

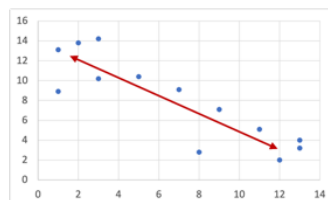
#### 相関のパターンと散布図



正の相関がある場合  
(相関係数は1に近い)



相関がない場合  
(相関係数は0に近い)



負の相関がある場合  
(相関係数は-1に近い)

相関分析で分かることは、あくまで「関係があるかどうか」のみです。

- 相関があることは分かっても、原因・結果は分からない
- 「Aが原因でBが起きた」とは言えない

この場合を「相関は因果を意味しない」といいます。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 自治体で行われる統計解析例

### 02 単回帰分析・重回帰分析

回帰分析は、データを使って「ある結果が、どの要因とどのような関係にあるか」を明らかにする手法でした。回帰分析と一口に言っても、考える要因の数によって、方法を使い分けます。

例えば、

「運行本数だけが利用者数に与える影響を知りたい場合」と、

「運行本数に加えて、人口や気温など複数の要因を同時に考えたい場合」とでは、手法を使い分けます。

#### 単回帰分析

1つの目的変数を1つの説明変数で予測するものを「単回帰分析」といいます。その予測を行う2つのデータの関係性 $y=ax+b$ という一次方程式の形で表せます。これは「回帰」で使われる最も基本的なモデルです。

$$\text{単回帰分析} : y = ax + b$$

#### 重回帰分析

2つ以上（2次元以上）の説明変数を持つものを「重回帰分析」といいます。適切な変数を複数選択することで、誤差の少ない予測式を立てることができます。

$$\text{重回帰分析} : y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + b_0$$

単回帰分析は、1つの要因と結果の関係を確認するための基本的な方法です。しかし実際の政策分析では、複数の要因が同時に影響していることが多いため、注目したい要因の「純粋な効果」を推定するには、他の要因を統制できる重回帰分析が有効です。分析の目的と前提に応じて、手法を選択することが重要です。



# 1 EBPMを動かすための準備 -視点と思考を整える

## 自治体で行われる統計解析例

### 03 主成分分析とは

たくさんの変数（項目）を、情報をなるべく失わずに、少ない軸にまとめる分析手法で、項目が多くて分かりにくいデータを、「特徴をよく表す少数のものさし（軸）」に整理することができます。

#### 主成分分析でできること

- 多くの項目を2～3軸に要約して可視化する
- データの全体傾向やグループを把握する
- 似た特徴をもつデータを整理する

#### 具体例

##### 例：バス路線の特徴を整理したい場合

ある地方公共団体で、路線ごとに次のようなデータがあるとします。

#### データ

- 利用者数
- 運行本数
- 平均乗車時間
- 高齢者利用率
- 通学利用率

項目が多く、路線の特徴が分かりにくい。



#### 主成分分析をすると

- 第1主成分：  
利用の多さ・規模
- 第2主成分：  
高齢者向け／通学向け

「路線の性格」を主成分という概念を用いて少ないものさし（軸）で説明できるようになります。

