

47都道府県の空き家構成に対する主成分分析

著者	旅する書齋
所属	ARIA × STELLA × NOX × LIBER
区分	不動産・地域の公開データ分析
対象	2023年住宅・土地統計調査確報+2026年地価公示
公開日	2026-09-01
版	公開版

要旨

47都道府県の空き家を、賃貸用・売却用・二次的住宅・その他の四つへ分け、構成比として主成分分析しました。最初の二つの連続軸で構成比変動の96.32%を要約できましたが、その軸は固定的な地域分類でも政策の優先順位でもありません。

キーワード

公開データ、二次分析、定義、分母、再現性、限界

1 研究目的

空き家率が近い地域でも内訳はどれほど異なり、四つの構成比を少数の軸へ要約すると、どの地域差が見えるのか。

47都道府県の空き家を、賃貸用・売却用・二次的住宅・その他の四つへ分け、構成比として主成分分析しました。最初の二つの連続軸で構成比変動の96.32%を要約できましたが、その軸は固定的な地域分類でも政策の優先順位でもありません。

2 データと定義

対象期間・資料：2023年住宅・土地統計調査確報+2026年地価公示

空き家率は総住宅数に占める空き家総数の割合です。構成比は、その空き家総数の中で賃貸用、売却用、二次的住宅、その他が何割かを示します。分母が違うため、空き家率が同じでも内訳は一致しません。

政策、管理、仲介、利活用の検討では、総量と内訳を並べる必要があります。賃貸募集在庫、別荘・セカンドハウス、長期に使われていない可能性のある住宅は、同じ追加調査では扱えません。

3 方法

1. 別の実装

47都道府県の四区分、空き家率、住宅地変動率を整形。

2. CLR変換

合計100%の閉鎖性に対応するため、四構成比を中心化対数比へ変換。

3. scikit-learn PCA

四構成比を三つの独立軸へ要約し、説明率と負荷量を保存。

4. Spearman相関

主成分得点と空き家率・その他空き家率・住宅地変動率を順位相関で照合。

5. 1県除外感度

47通りの除外で説明率と外部相関の範囲を確認。

6. 別の実装

散布図、主成分得点、代表地域構成を日英で固定。

各方法は、分母、単位、時点、欠損、除外条件を確認してから適用しました。因果関係や将来値の推定に必要な条件がない場合、その方法は主要結果に使っていません。

4 結果

1. 二つの軸で、構成比変動の96.32%を要約

第1軸は78.59%を説明し、二次的住宅の構成が大きい側と、賃貸・売却の構成が大きい側を分けます。第2軸は17.73%で、『その他の空き家』の構成が大きい側を表します。

2. 東京70.16%、長野26.86%、島根66.48%

東京は空き家の70.16%が賃貸用、長野は26.86%が二次的住宅、島根は66.48%が賃貸・売却・二次的住宅を除く空き家です。同じ『空き家』でも、問いと必要な追加資料が変わります。

3. 地価変動率との関連はあるが、因果ではない

地価を主成分の入力には使わず、外部記述値として照合しました。第2軸と2026年住宅地変動率のSpearman相関は-0.477です。年次が異なり、人口・所得・都市構造を統制していないため、空き家構成が地価を動かしたとは言えません。

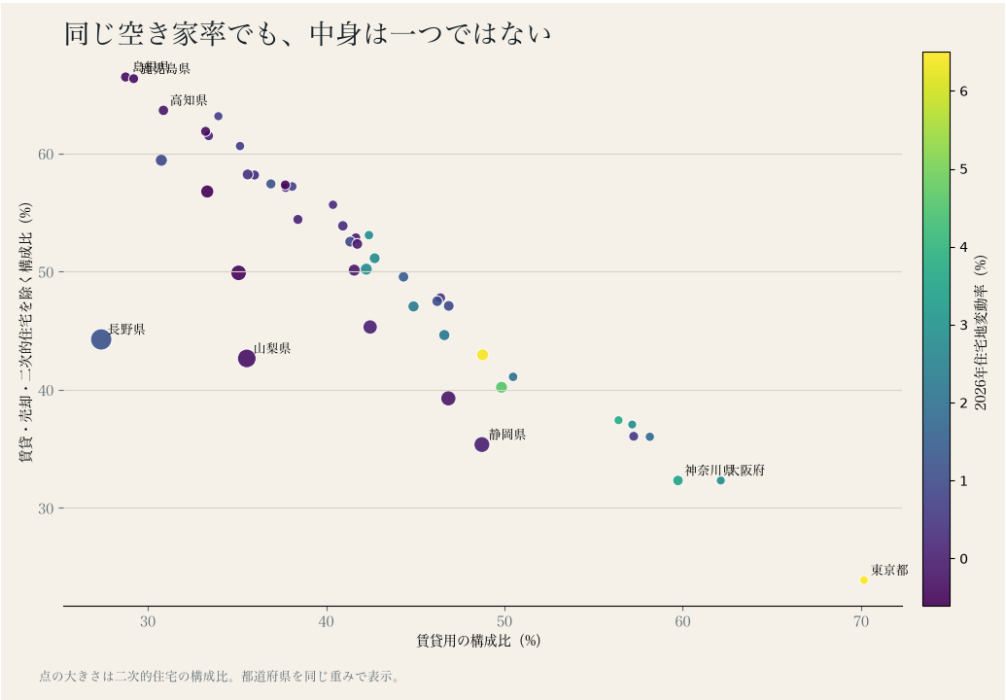


図1 図1 | 賃貸用とその他空き家の構成

点の大きさは二次的住宅の構成比、色は2026年住宅地変動率です。

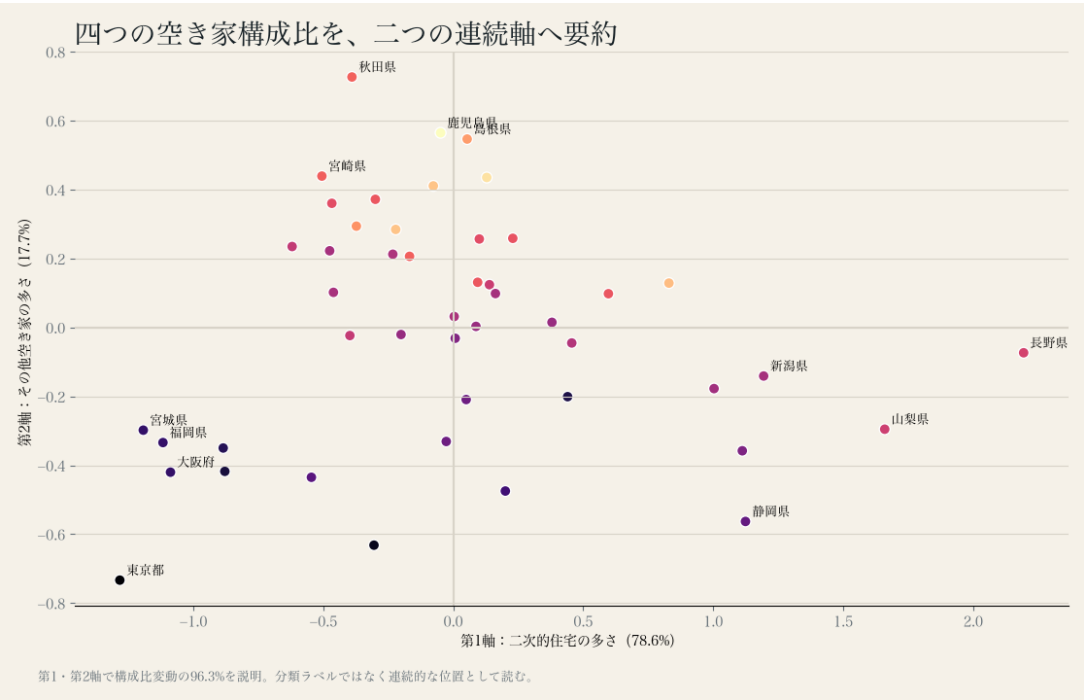


図2 図2 | 空き家構成比の二つの主成分
最初の二軸で96.32%を要約しますが、固定的な分類ではありません。

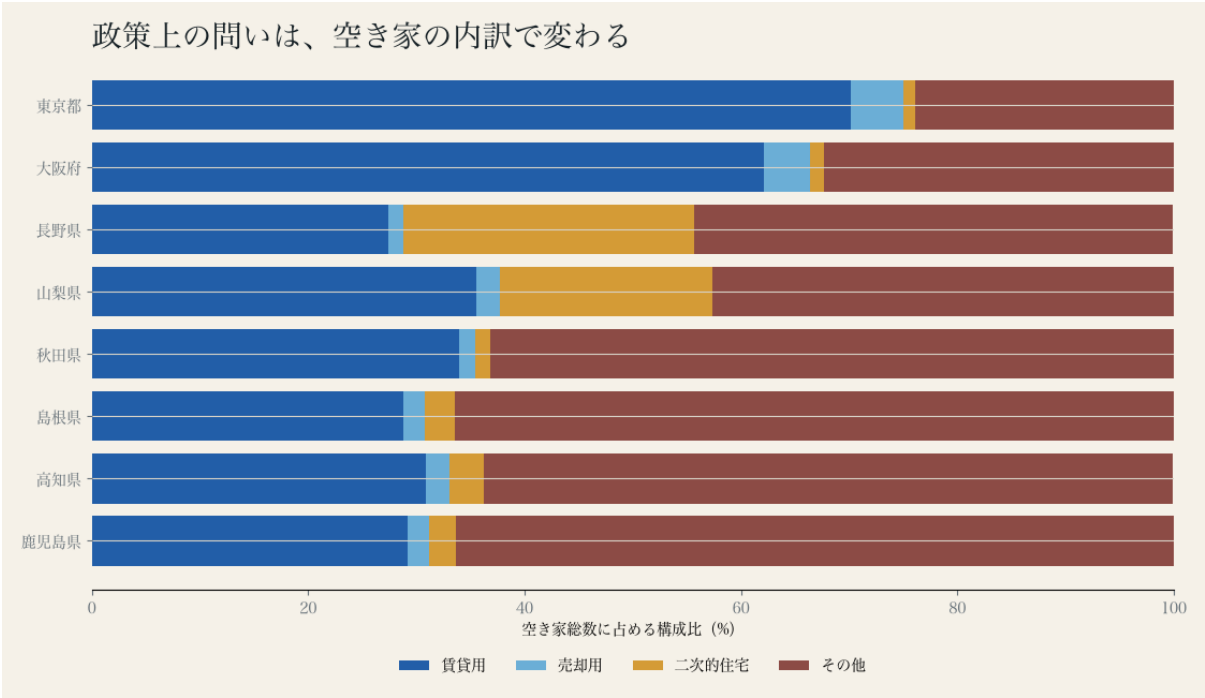


図3 図3 | 代表8地域の空き家内訳
同程度の空き家率でも、賃貸用・二次的住宅・その他の比重は異なります。

5 考察

2 | 合計100%のデータを、そのまま相関へ入れない

四つの構成比は合計がおおむね100%で、一つが増えると他は機械的に減ります。そのまま通常の主成分分析へ入れると、この閉じた制約が相関として現れます。そこで、各構成比へ対数を取り、地域内の幾何平均との差を表す中心化対数比（CLR）へ変換しました。

すべての都道府県で四区分が正の値だったため、ゼロ置換はしていません。都道府県は戸数や人口で重み付けせず、地域構造の47観測として等しく扱います。

3 | 第1軸は二次的住宅、第2軸はその他空き家

第1軸の二次的住宅の係数は+0.842で、賃貸用-0.300、売却用-0.434と反対方向です。長野、山梨、新潟、静岡などが高い側に位置します。これは『良い・悪い』の軸ではなく、構成の対比です。

第2軸はその他空き家の係数が+0.854で、秋田、鹿児島、島根、宮崎、高知が高い側です。『その他空き家率』と強く関連するのは、同じその他空き家数を別の分母で使っているためでもあり、独立した発見として過大評価しません。

4 | 地域を箱へ固定せず、連続的な位置として扱う

クラスター数を先に決めて『都市型』『地方型』と名付けると、分析者の都合で境界を作りやすくなります。この分析では主成分得点を連続軸として示し、必要なときだけ個別地域の原表へ戻ります。

一県を順に除外しても、最初の二軸が説明する割合は94.88%から97.52%の範囲でした。ただし、軸の向きや係数は標本に依存し、別年の調査へそのまま固定できません。

5 | 地価との照合は、仮説の最初の資料に限定する

2026年住宅地変動率は主成分分析へ入れず、後から色と相関で重ねました。第1軸とのSpearman相関は-0.330、第2軸とは-0.477でした。一県ずつ除外した第2軸相関は、おおむね-0.51から-0.44の範囲です。

住宅データは2023年、地価変動は2026年で、時点がそろいません。人口減少、世帯数、所得、観光、都市規模、供給なども統制していないため、方向や原因を結論にしません。

6 | 会議では、内訳ごとに次の資料を変える

賃貸用の構成が大きい地域では、募集期間、家賃、管理状態、建築時期を確認しています。二次的住宅が大きい地域では、季節利用、別荘地、管理費、交通条件が必要です。その他空き家が大きい地域では、腐朽・破損、所有者、接道、相続、除却・改修可能性など別の資料が要ります。

構成比は追加資料の順番を決める入口で、個別住宅の危険性や活用可能性を判定するものではありません。

7 | 再現資料と限界

公式原表から作った47都道府県CSV、CLR変換後の主成分得点、負荷量、説明率、相関、1県除外感度をJSONへ保存しました。記事、日英図、研究報告書、1ページ要約は同じ結果ファイルを使います。

主成分は記述要約で、予測モデル、政策効果、個別物件評価ではありません。47都道府県を等重みで扱うため、全国の戸数構成を表す軸でもありません。

6 限界

1. 限界1

主成分は集計構成比の記述要約で、政策類型や因果モデルではありません。

2. 限界2

都道府県を等重みで扱い、戸数・人口・面積・取引量では重み付けしていません。

3. 限界3

住宅は2023年、地価は2026年で時点が異なり、方向と原因は推定できません。

4. 限界4

都道府県内の市区町村差、建築時期、腐朽・破損、所有者事情は説明できません。

7 結論

空き家率は総量の最初の資料ですが、地域課題を決めるには内訳が必要です。四つの構成比をCLR変換して主成分分析すると、二次的住宅とその他空き家を中心とする二つの連続軸で大半の構成差を要約できました。ただし、軸は分類や因果関係ではなく、そのあと確認する資料を選ぶための記述結果として使うべきです。

次に必要なデータ

建築時期、腐朽・破損、共同住宅比率、人口・世帯変化を同じ市区町村単位で接続し、都道府県内の差を確認しています。

参考文献

1. 総務省統計局：令和5年住宅・土地統計調査・調査結果。<https://www.stat.go.jp/data/jyutaku/2023/tyousake.html> (stat.go.jp)
2. e-Stat：主要統計表1・居住世帯の有無。<https://www.e-stat.go.jp/stat-search/file-download?statInfId=000040353925&fileKind=4> (e-stat.go.jp)
3. 総務省統計局：用語の解説。<https://www.stat.go.jp/data/jyutaku/2023/pdf/giy00.pdf> (stat.go.jp)
4. 国土交通省：令和8年地価公示。https://www.mlit.go.jp/tochi_fudousan_kensetsugyo/tochi_fudousan_kensetsugyo_fr4_000001_00324.html (mlit.go.jp)
5. 国土交通省：第5表・都道府県別変動率。https://www.mlit.go.jp/tochi_fudousan_kensetsugyo/content/001986180.xlsx (mlit.go.jp)

付録A 再現に必要な情報

公開記事：<https://aria-stella-studio.tabisurushosai.chatgpt.site/analysis/japan-vacancy-composition-pca-2023>

本文データ：<app/analysis/japan-vacancy-composition-pca-2023/content.json>

分析環境：103機能を実行確認済み。各テーマではデータ条件に合う方法だけを採用しました。

1. 別の実装

47都道府県の四区分、空き家率、住宅地変動率を整形。

2. CLR変換

合計100%の閉鎖性に対応するため、四構成比を中心化対数比へ変換。

3. scikit-learn PCA

四構成比を三つの独立軸へ要約し、説明率と負荷量を保存。

4. Spearman相関

主成分得点と空き家率・その他空き家率・住宅地変動率を順位相関で照合。

5. 1県除外感度

47通りの除外で説明率と外部相関の範囲を確認。

6. 別の実装
散布図、主成分得点、代表地域構成を日英で固定。

付録B 集計表

表B-1 代表8地域の空き家構成と外部指標

地域	空き家率	賃貸用%	売却用%	二次的%	その他%	住宅地変動率
東京都	10.93%	70.16%	4.89%	1.07%	23.89%	6.5%
大阪府	14.24%	62.13%	4.27%	1.25%	32.33%	2.8%
長野県	20.06%	27.39%	1.39%	26.86%	44.27%	1.2%
山梨県	20.42%	35.55%	2.18%	19.61%	42.66%	-0.3%
秋田県	15.77%	33.96%	1.44%	1.44%	63.17%	0.7%
島根県	17.05%	28.75%	2.01%	2.75%	66.48%	-0.4%
高知県	20.28%	30.88%	2.16%	3.18%	63.66%	-0.2%
鹿児島県	20.48%	29.21%	2.01%	2.44%	66.34%	-0.4%

注：空き家構成は2023年、住宅地変動率は2026年。割合は表示用に小数第2位へ丸めています。