

分析レポート

高齢者の免許返納率に影響する地域・生活要因の探索的分析

川口哲史

全体の流れ

背景および目的. 高齢者の免許保有と返納の地域差

高齢化による交通事故の増加と自主返納率、返納率の地域差

分析および結果. 地域別にみる返納しにくさの構造分析

相関関係、主成分分析 + クラスタリング、SHAP分析

施策案. 意思決定支援

各クラスタごとの意思決定支援案

まとめと今後の展望

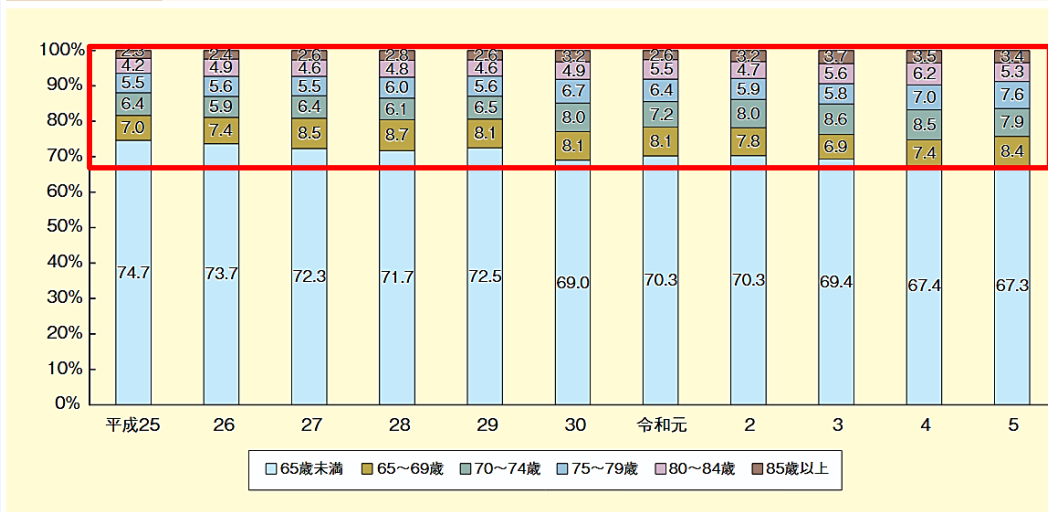
背景-高齢化による交通事故の増加と自主返納率

近年高齢化はますます高まっており、2023年で高齢者の割合は29%、2070年には約40%に近づくと予想されている

どの年齢層が死亡事故を引き起こしたのか？

特集 - 第 32 図

年齢層別交通死亡事故件数の割合の推移



注 1 警察庁資料による。

2 第1当事者が原付以上の死亡事故を計上している。

3 構成割合は小数点以下第2位を四捨五入しているため、合計しても必ずしも100とならない。

引用：R6年度交通安全白書特集

死亡事故件数を減らすための一つ的手段として
免許の自主返納が考えられる

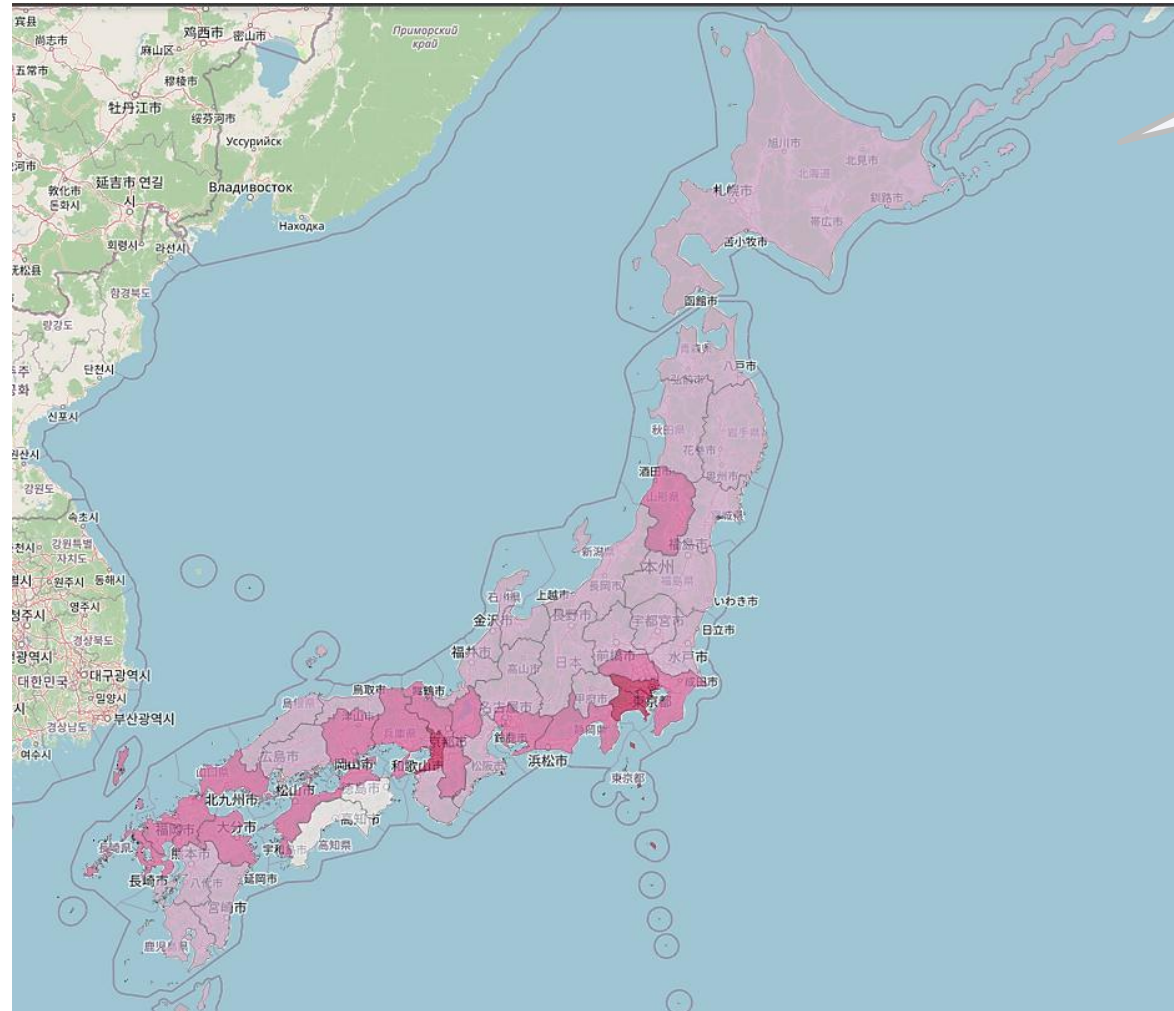
高齢者による事故件数の割合（赤枠）は徐々に増加

※高齢者：65歳以上を示す

ところが、高齢者免許保有者のうち1～3%ほどしか免許返還をしていない

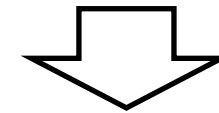
背景-返納率の地域分布

地域差の存在を視覚化し、地理的偏りはあるのか？



色が濃いほど、返納率高いことを示す

東京近辺、関西近辺、福岡近辺
は返納率が高い

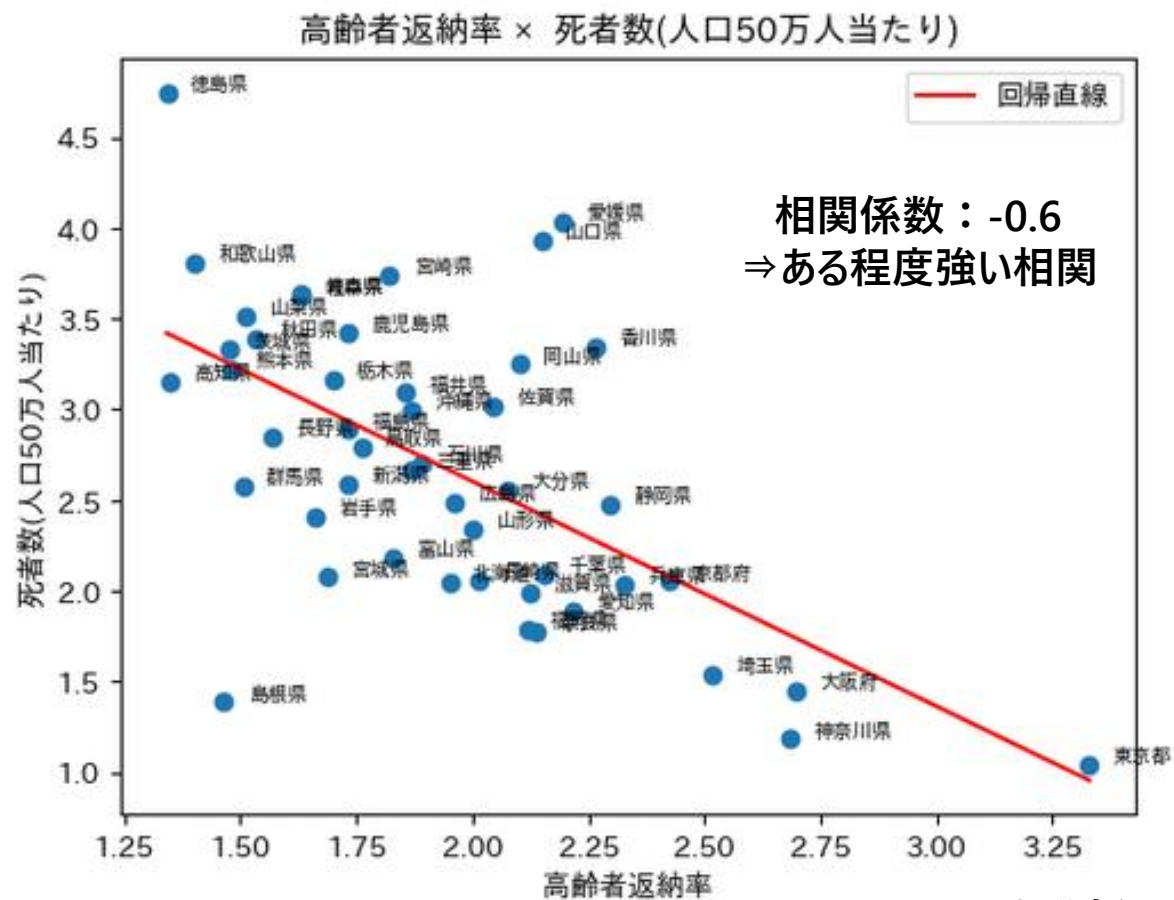


地域ごとに
極端なばらつきはありそう

図. 都道府県ごとの返納率の分布

目的-目的変数の設定

返納率と交通事故死者数を都道府県別にプロット



出典：道路の交通に関する統計（2024年）

※本分析の回帰直線は、観測範囲内での傾向（1.2～3.2%）を示すものであり
極端な返納率に外挿することは想定していない

回帰式： $y = 5.08 - 1.24x$

返納率1%上昇 → 人口50万人あたり死亡者数 -1.24人

人口換算すると...

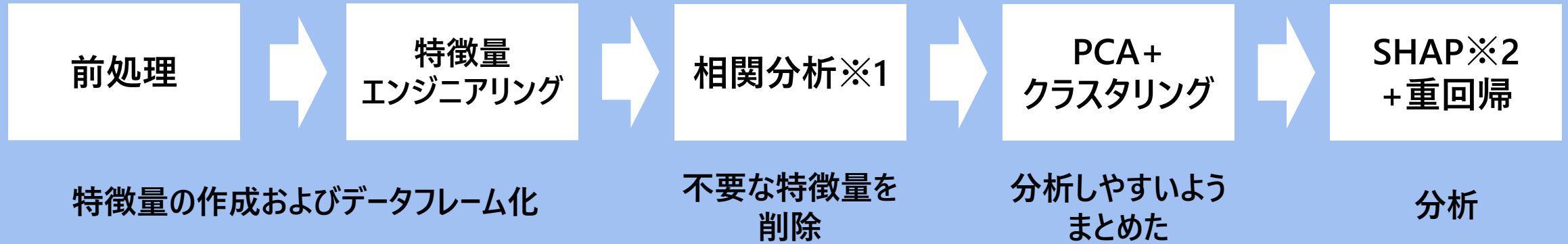
人口1億2000万人/50万人 = 240
⇒ $240 \times 1.24 \div 300$ (人)

返納率が1%上がれば、全国で年間約300人の命
が救える※（年間死者数約2600人,2024年）

目的変数を高齢者の免許返納率に設定し、都道府県別に分析する

分析-分析フロー

データ分析の流れ



※1：詳しい部分は割愛
※2：特徴量の寄与度を表す指標



意思決定支援

分析-特徴量の作成

「返納しやすさ」は生活圏内での移動手段・支援体制・医療アクセスなど複数の生活環境や地域特性に左右されていると考え、特徴量を作成

公共交通の 利便性

- ・車通勤率（都道府県別, proxy※1）
- ・バス利用率（都道府県別, proxy※1）
- ・鉄道利用率（都道府県別, proxy※1）

地形

- ・可住地面積比率（人が住める面積/総面積）
- ・離島の有無

医療・買い物施設 へのアクセス

- ・病院・診療所合計（人口10万人当たり）
- ・食料品アクセス困難人口率※2

都市構造

- ・政令指定都市の数（人口50万人以上の都市）

家族構成と 高齢者の役割

- ・三世代世帯率
- ・65歳以上のみの世帯率
- ・高齢者単身率
- ・高齢者就業率

産業

- ・外国人観光来客数
- ・第一次産業比率（農林水産業の比率）

※1 代理変数(proxy)：本来測りたい概念を直接観測できないときに代わりに用いる変数。

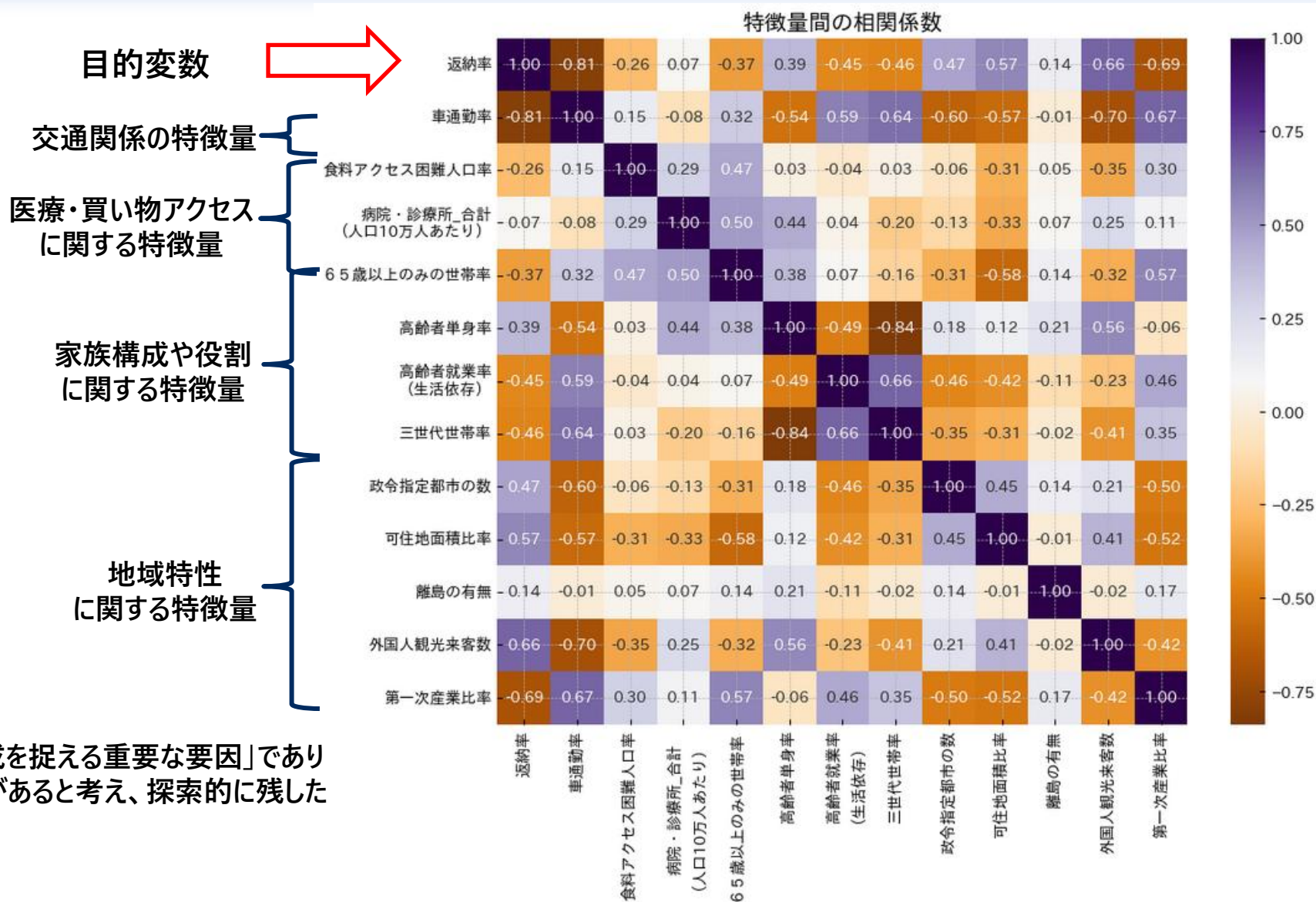
※2食料品アクセス困難人口とは、店舗まで500m以上かつ自動車利用が困難な65歳以上高齢者を指す。店舗は、食肉、鮮魚、野菜・果実小売業、百貨店、総合スーパー、食料品スーパー、コンビニエンスストア、ドラッグストアが含まれる。

分析-相関係数

①目的変数を高齢者
自主返納率（=返納率）と設定

②多重共線性を減らすためにバス
利用率や鉄道利用率などの相関
係数（>0.7）が高い組み合わせ
を削除

※高齢者単身率×三世代世帯率は「家族構成を捉える重要な要因」であり
返納行動に対して異なる影響を与える可能性があると考え、探索的に残した



分析-PCA+クラスタリングの最適化手法

特徴量が多いので、主成分分析とクラスタリングを使って、都道府県を分類したい

表. クラスタリングの最適クラスタ数の手法

観点	エルボー法	シルエットスコア
意味	クラスタ数が増えるほどSSEが下がるが、その**減少の鈍化点（ひじ）**を見る	各データ点がどれだけ適切なクラスタに属しているかのスコア（-1～1）
メリット	<ul style="list-style-type: none">・視覚的にわかりやすい・計算が軽い	<ul style="list-style-type: none">・スコアに基づく明確な評価・クラスタの質も評価可能
デメリット	<ul style="list-style-type: none">・主観が入る（どこが「ひじ」か不明瞭なことも）・クラスタ数が少ないと誤誘導されやすい	<ul style="list-style-type: none">・計算コストが高い・高次元データだと誤差大きめになることがある
適している場面	おおまかなクラスタ数を探る初期段階	正確にクラスタの分離性と凝集度を見たいとき
推奨クラスタ数の選び方	ひじの位置	最大のスコア点

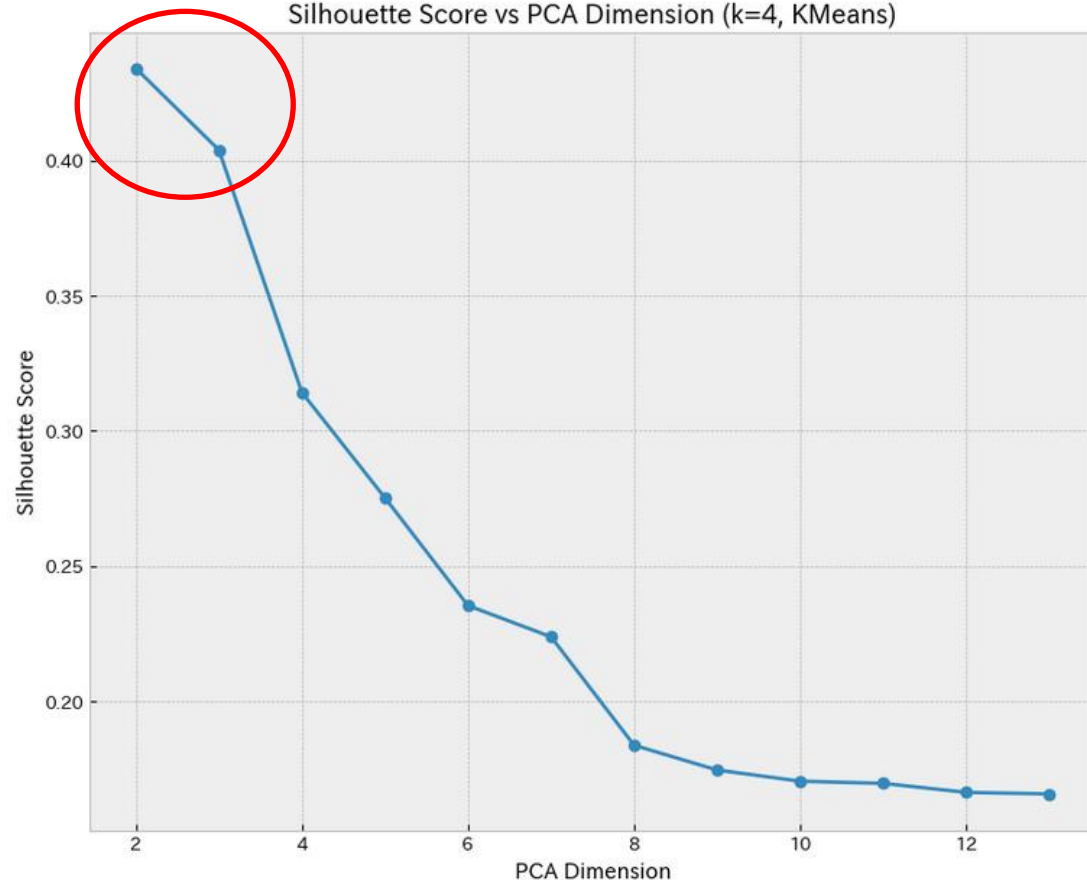


シルエットスコアを使用して、最適化

分析-PCA + クラスタリング

PCA次元を変えながら、クラスタ品質を比較

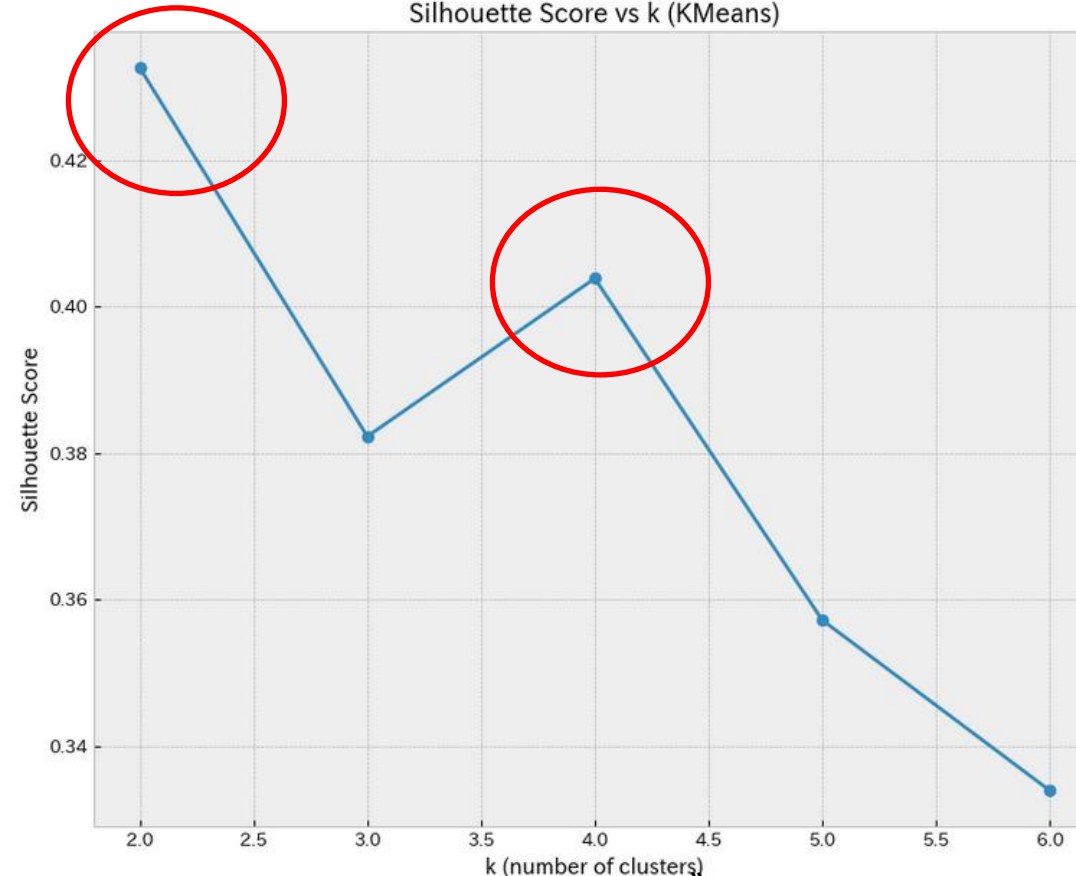
Silhouette Score vs PCA Dimension (k=4, KMeans)



クラスタリングが4の時、
PCAはk=2, 3 (赤丸) あたりがよさそう

次元を絞った上で、kを変化させて再評価

Silhouette Score vs k (KMeans)

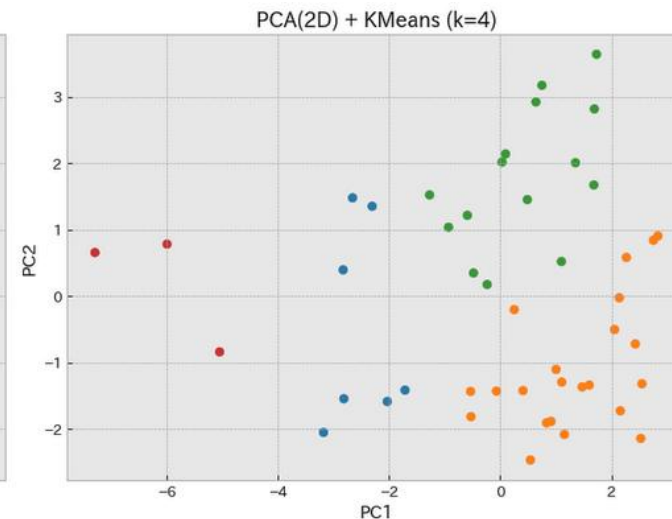
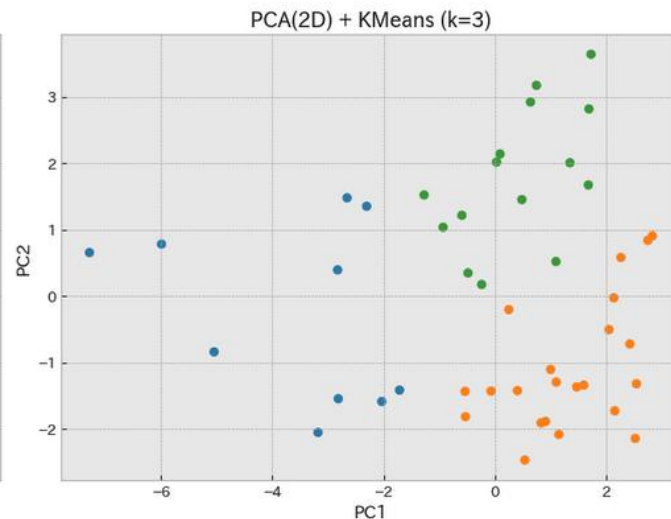
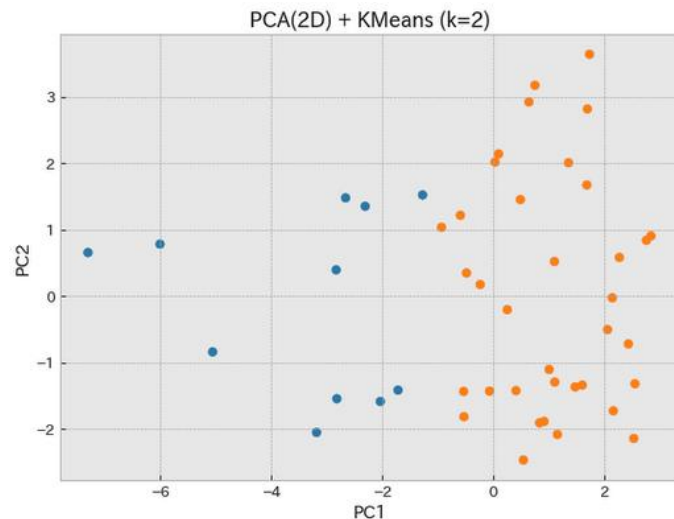


PCA k=2, 3のいずれでも
クラスタリングはk=2, 4 (赤丸) あたりがよさそう
(画像は、PCA k=3)

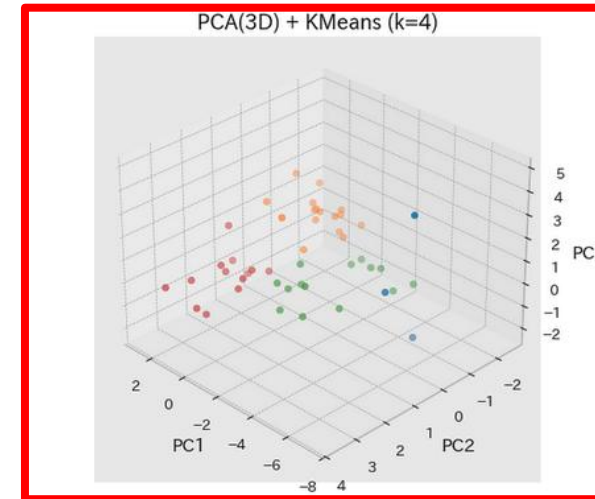
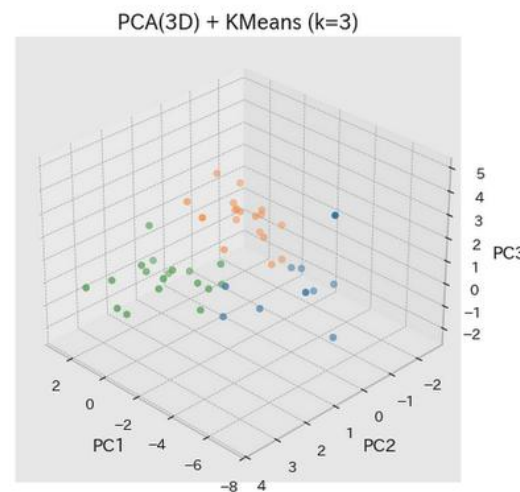
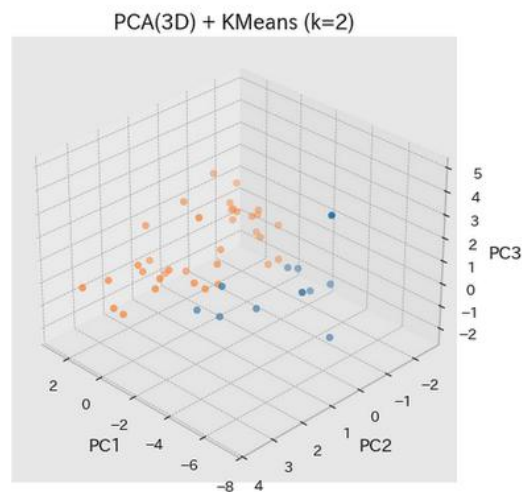
分析-PCA + クラスタリング

主成分分析（PCA）は2 or 3、クラスタリングは2~4でそれぞれ実施、およびプロット

PCA2次元



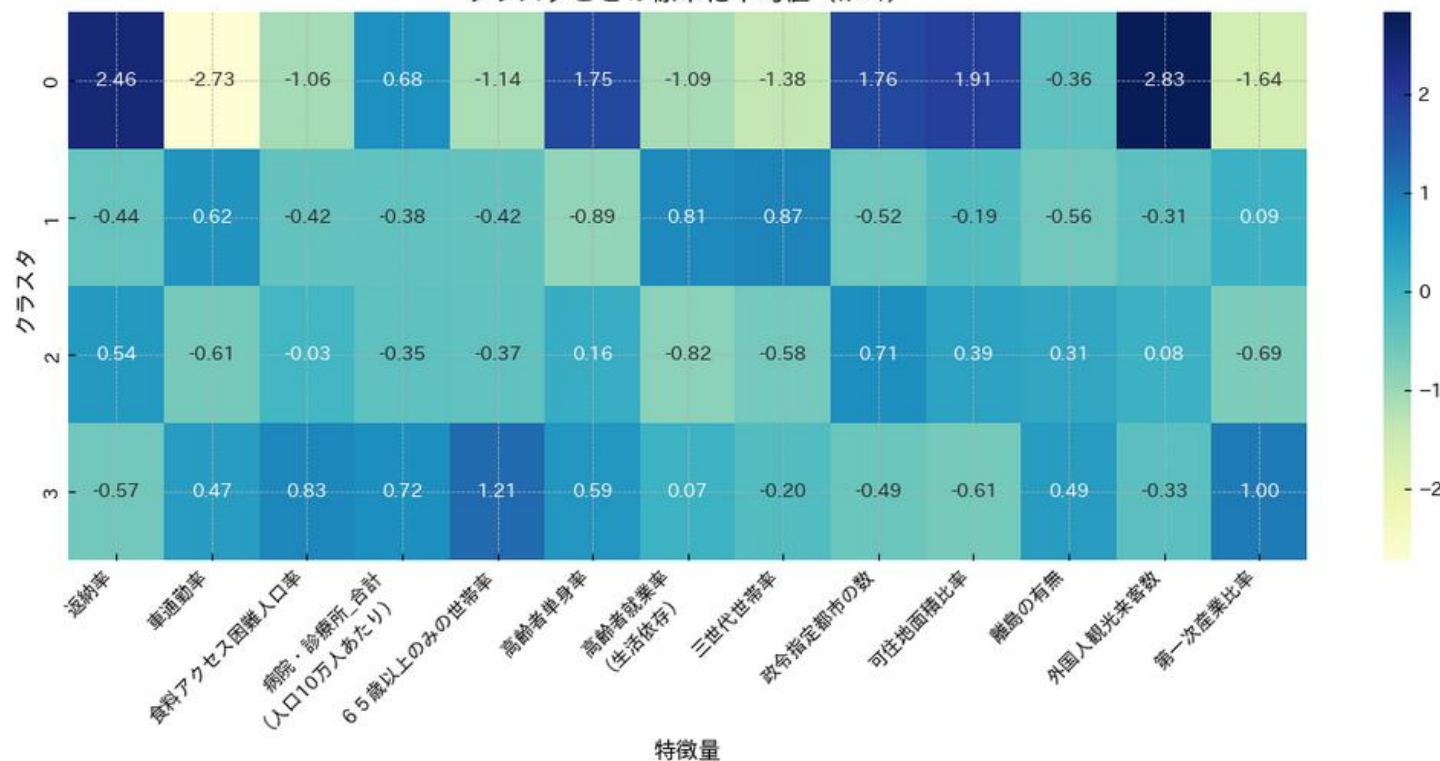
PCA3次元



最もきれいに分かれたPCA : 3, クラスタリング : 4で進めることとした

分析-クラスタリング (k=4)

クラスタごとの標準化平均値 (k=4)



0=平均、0から離れるほどその度合いが大きいことを示す

クラスタ0

大都市型。返納率が**超高い**。公共交通機関が発展しており、高齢者単身率も高く、観光客も多い

クラスタ1

返納率**低め**。車通勤率や高齢者就業率や三世代世帯率が高いことから、**日常生活や仕事で使用し、家族同居が多い**。

クラスタ2

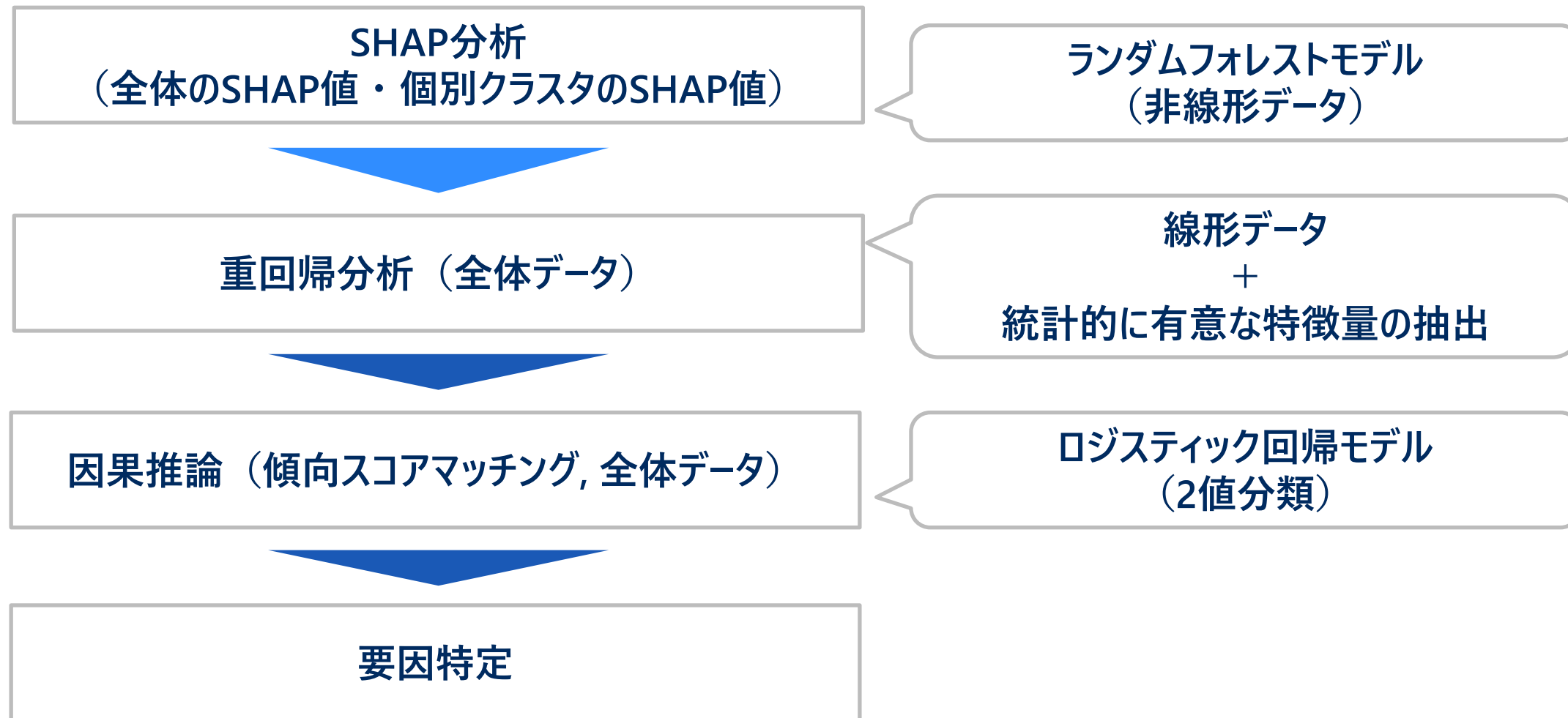
返納率高め。日常生活で車がなくても生活できる。政令指定都市の数は多いものの、就業率は低めである。**中核都市型**

クラスタ3

返納率**低め**。**日常生活で車がないと不便な地域**に在住。高齢者のみの世帯が多い。また、第一次産業比率も高い

分析-要因分析の流れ

網羅性を重視した特徴量選択フロー



分析-使用するモデル（SHAP）

SHAPで特徴量の寄与度をそれぞれ分析→学習モデルはランダムフォレスト

選定理由

理由	説明
非線形を捉えられる	返納率に対して、単純な相関関係ではなく“しきい値”や“組み合わせ効果”がある場合に強い
前処理がほぼ不要	標準化やスケーリングなしで使える （構造が複雑なデータでもOK）
変数の重要度を出しやすい	.feature_importances_ や SHAPとの相性が抜群
解釈可能なツリー系モデル	単純な回帰よりも柔軟で、しかも重みの“方向”も見やすい

その他モデルとの比較

モデル	特徴	SHAPとの相性
線形回帰	単純な相関分析に近いが 複雑な関係を捉えにくい	△（SHAPの意味が薄くなる）
決定木	直感的だが過学習しやすい	○
ランダムフォレスト	精度・汎用性・解釈性のバランス◎	◎
LightGBM / XGBoost	精度高いが少しテクニカル	◎（大規模データなら）

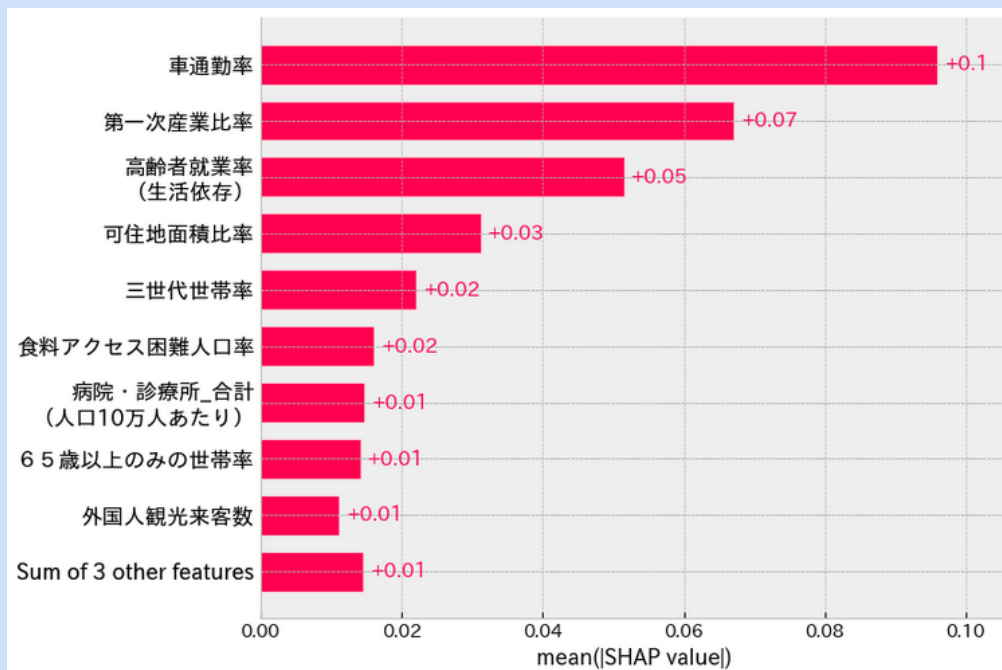
◎：とてもいい ○：いい △：少しよくない

分析-SHAP値の全体傾向

目的変数を返納率としたSHAP値の結果

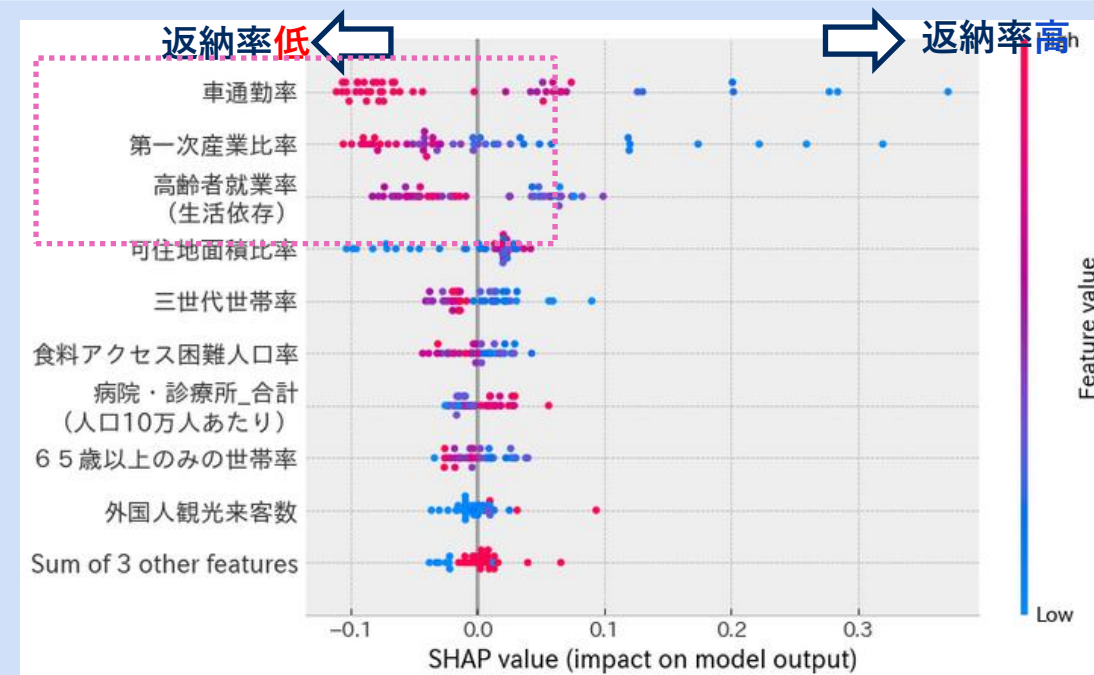
決定係数 R^2 : 0.933
MAE (平均絶対誤差): 0.075
RMSE (二乗誤差の平方根): 0.103

特徴量の重要度ランキング (目的変数にどれだけ影響を与えたか)



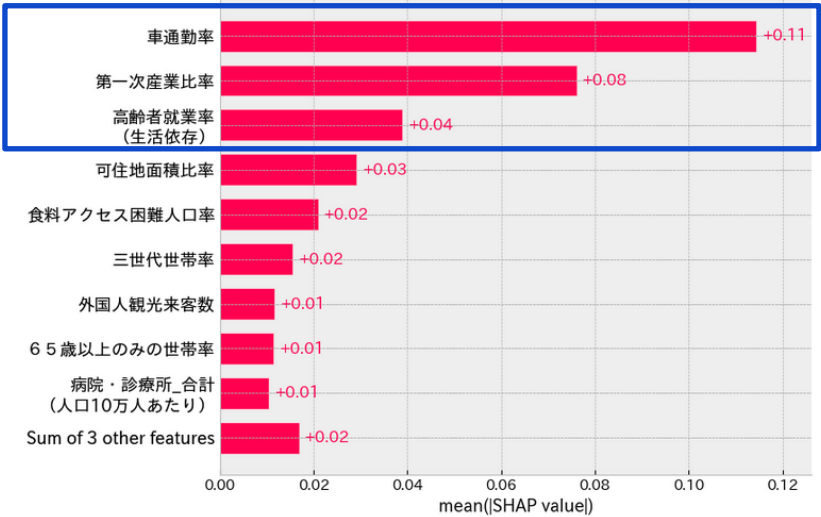
車通勤率 > 第一次産業比率 >
高齢者就業率 > 可住地面積比率

各特徴量が返納率の上げ下げにどの程度効いているのか？



車通勤率が**高い**ほど、返納率**低い**傾向
第一次産業比率が**高い**ほど、返納率**低い**傾向
高齢者就業率が**高い**ほど、返納率が**低い**傾向

分析-SHAP分析まとめ



SHAP value（左図は全体のSHAP）のうち
上位3位をそれぞれ表にまとめた

表. 各モデルの精度とSHAP値

	全体	クラスタ0 (大都市)	クラスタ1 (田舎、家族同居)	クラスタ2 (地方都市)	クラスタ3 (田舎、高齢者多)
精度 (決定係数)	0.94	0.84	0.84	0.84	0.90
SHAP value (上位3位)	<div>・車通勤率</div> <div>・第一次産業比率</div> <div>・高齢者就業率</div>	<div>・政令指定都市の数</div> <div>・食料アクセス困難人口率</div> <div>・高齢者就業率</div>	<div>・離島の有無</div> <div>・65歳以上のみの世帯率</div> <div>・病院・診療所_合計</div> <div>・第一次産業比率</div>	<div>・65歳以上のみの世帯率</div> <div>・第一次産業比率</div> <div>・車通勤率</div>	<div>・第一次産業比率</div> <div>・高齢者就業率</div> <div>・外国人観光来客数</div> <div>・食料アクセス困難人口率</div>

※1：各クラスタの分析結果は付録に記載
※2：赤字は地域特性の特徴量
※3：同率3位の特徴量は複数記載

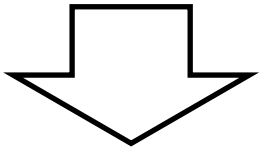
分析-重回帰分析

表. 各モデルの精度とSHAP値

	全体	クラスタ0 (大都市)	クラスタ1 (田舎、家族同居)	クラスタ2 (地方都市)	クラスタ3 (田舎、高齢者多)
精度 (決定係数)	0.94	0.84	0.84	0.84	0.90
SHAP value (上位3位)	<ul style="list-style-type: none">・車通勤率・第一次産業比率・高齢者就業率	<ul style="list-style-type: none">・政令指定都市の数・食料アクセス困難人口率・高齢者就業率	<ul style="list-style-type: none">・離島の有無・65歳以上のみの世帯率・病院・診療所_合計・第一次産業比率	<ul style="list-style-type: none">・65歳以上のみの世帯率・第一次産業比率・車通勤率	<ul style="list-style-type: none">・第一次産業比率・高齢者就業率・外国人観光来客数・食料アクセス困難人口率

※同率3位の特徴量は複数記載

モデル精度を高く保ちつつ、SHAP値で関連する特徴量を選定した



重回帰分析で統計的に有意な特徴量の抽出

分析-重回帰分析（全体データ）

サンプル数に対する説明変数の数を考慮し、モデルの安定性と前提条件（残差の正規性など）を満たすため特徴量をSHAPの上位6つに絞った

OLS Regression Results

Dep. Variable:

返納率

R-squared:

0.706

Model:

OLS

Adj. R-squared:

0.662

Method:

Least Squares

F-statistic:

15.99

Date:

Mon, 11 Aug 2025

Prob (F-statistic):

2.82e-09

Time:

16:49:46

Log-Likelihood:

5.7082

No. Observations:

47

AIC:

2.584

Df Residuals:

40

BIC:

15.53

Df Model:

6

Covariance Type: nonrobust

	coef	std err	t	P> t	[0.025	0.975]
const	1.9323	0.034	57.026	0.000	1.864	2.001
車通勤率	-0.2511	0.059	-4.244	0.000	-0.371	-0.132
高齢者就業率（生活依存）	0.0209	0.050	0.416	0.680	-0.081	0.123
可住地面積比率	0.0386	0.044	0.868	0.390	-0.051	0.129
三世代世帯率	0.0078	0.052	0.151	0.881	-0.096	0.112
第一次産業比率	-0.0880	0.049	-1.787	0.081	-0.188	0.012
食料アクセス困難人口率	-0.0274	0.038	-0.727	0.472	-0.104	0.049

Omnibus:

1.977

Durbin-Watson:

2.346

Prob(Omnibus):

0.372

Jarque-Bera (JB):

1.364

Skew:

0.165

Prob(JB):

0.506

Kurtosis:

2.234

Cond. No.

3.98

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified

表. モデル精度

指標	値	評価
R-squared	0.706	モデルは返納率の約70%を説明
Adj. R-squared	0.662	自由度調整済みモデルでも70%に近い精度
F-statistic	19.31 (p値 = 2.82e-9)	モデル全体が統計的に有意。少なくとも1つの変数が有意。
Condition No.	3.98	問題なし



モデル精度問題なし

P値<0.05が統計的に有意と判断

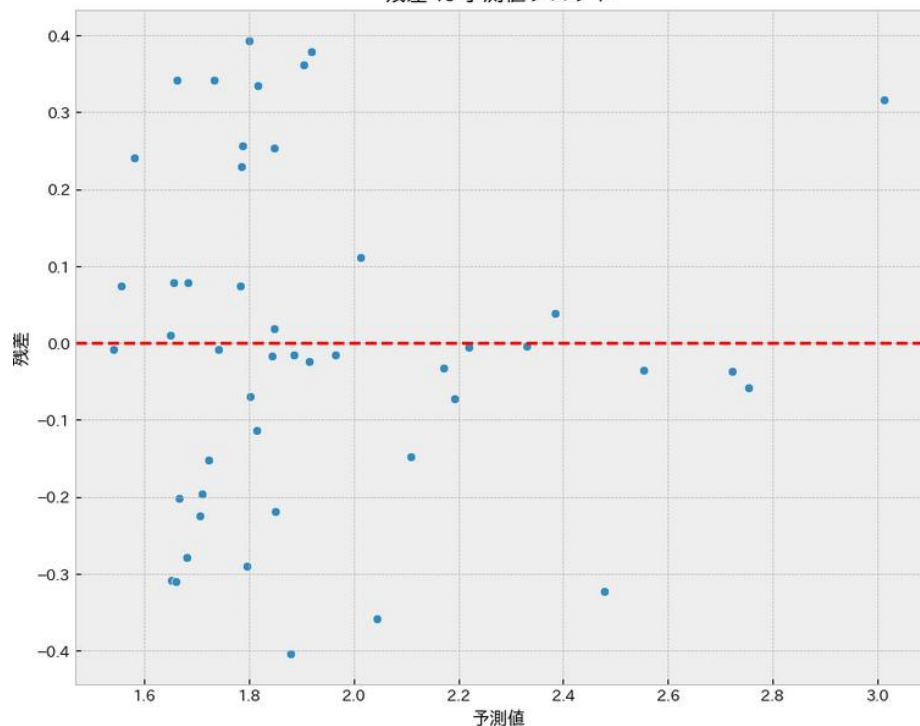
- P値より、
- ・車通勤率は統計的に有意
 - ・第一次産業比率は統計的にわずかに有意でないので残す

分析-正規性の確認（全体）

重回帰モデルの結果を解釈する前に前提条件を確認する必要がある

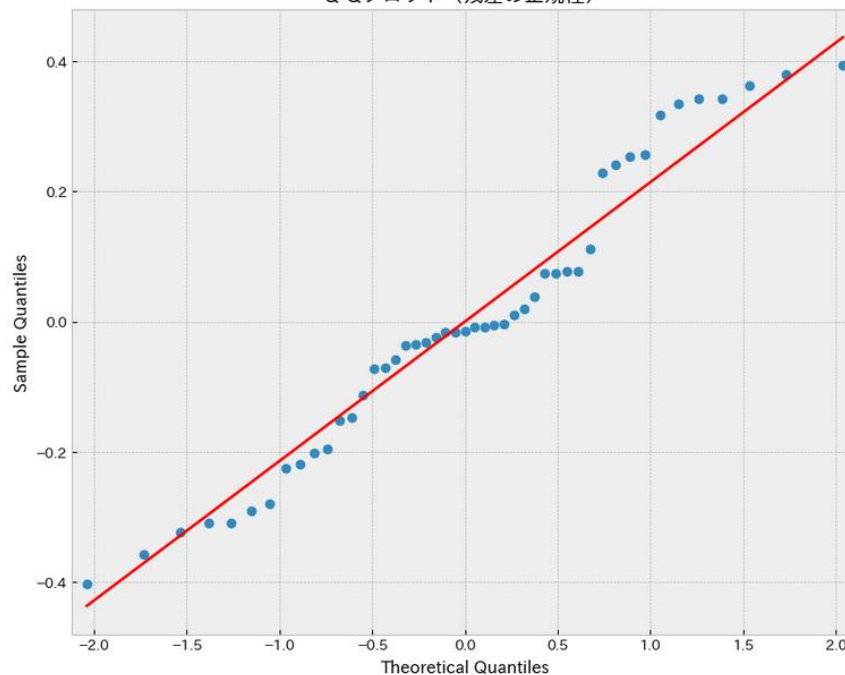
ランダムに散らばっているのか？

残差 vs 予測値 プロット



赤線上に点が並んでいるのか？

Q-Qプロット（残差の正規性）



P値 > 0.05か？

Shapiro-Wilk 検定統計量: 0.9546722919898334
p値: 0.06614255321485304
→ 正規性を仮定してよい（帰無仮説を棄却できない）

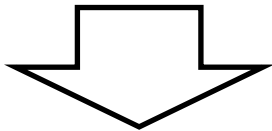
モデルの推定結果の信頼性（p値や信頼区間の妥当性）を担保するために
誤差項（残差）が正規分布に従うことが理想 ⇒ **問題なし**

分析-重回帰分析結果まとめ

表. 重回帰分析モデルにおける精度と統計的に有意な特徴量

	全体	クラスタ0 (大都市)	クラスタ1 (田舎、家族同居)	クラスタ2 (地方都市)	クラスタ3 (田舎、高齢者多)
精度 (自由度調整 済み決定係 数)	0.67	データ不足	0.69	0.38	0.32
統計的に有意 な特徴量	・車通勤率 ・ 第一次産業比率	--- (データ不足)	・離島の有無 ・65歳以上のみの世帯率 ・病院・診療所_合計	なし	なし

※赤字は統計的にわずかに有意ではないが、可能性として記載



全体とクラスタ1の特徴量の因果推論を実施したが有効な結果は得られなかった

※詳細はAppendix3へ

分析-結果まとめ

主成分分析とクラスタリング

- 主成分分析： $n=3$
- クラスタリング($k=4$)：
 - クラスタ0：車が必要だが、息子夫婦や孫と一緒に住んでおり、生活には比較的困らない高齢者のいる県
 - クラスタ1：返納率は最低、生活で一番車が必要な高齢者のいる県
(単身率は比較的高め、食料アクセス困難高め、65歳以上のみの世帯率高)
 - クラスタ2：都市圏の高齢者。高齢者単身化が顕著で、公共交通機関が潤っているので車不要。当然、返納率も高い
 - クラスタ3：返納率低め。日常生活で車がないと不便な地域に在住。高齢者のみの世帯が多い。また、第一次産業比率も高い

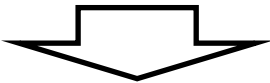
要因分析（SHAP+重回帰分析+傾向スコアマッチング）

- SHAP値では、返納数に寄与する上位3位の特徴量を見出すことができた
- 重回帰分析による統計的に有意な特徴量は一定数見出せた
- 傾向スコアマッチングでは共変量調整が不十分であり、不確実性がある（因果的解釈を試みたが、現行データでは十分でなく、今後の課題）

施策案-意思決定支援

前章において、確実性の高いSHAP値のみの結果を用いて意思決定支援を行う
ただし、統計的に有意な特徴量に関する政策は優先的に行うものとする

	全体	クラスタ0 (大都市)	クラスタ1 (田舎、家族同居)	クラスタ2 (地方都市)	クラスタ3 (田舎、高齢者多)
SHAP value (上位3位)	・車通勤率 ・第一次産業比率 ・高齢者就業率	・政令指定都市の数 ・食料アクセス困難人口率 ・高齢者就業率	・離島の有無 ・65歳以上のみの世帯率 ・病院・診療所_合計 ・第一次産業比率	・65歳以上のみの世帯率 ・第一次産業比率 ・車通勤率	・第一次産業比率 ・高齢者就業率 ・外国人観光来客数 ・食料アクセス困難人口率
統計的に有意な特徴量	・車通勤率 ・第一次産業比率	--- (データ不足)	・離島の有無 ・65歳以上のみの世帯率 ・病院・診療所_合計	なし	なし



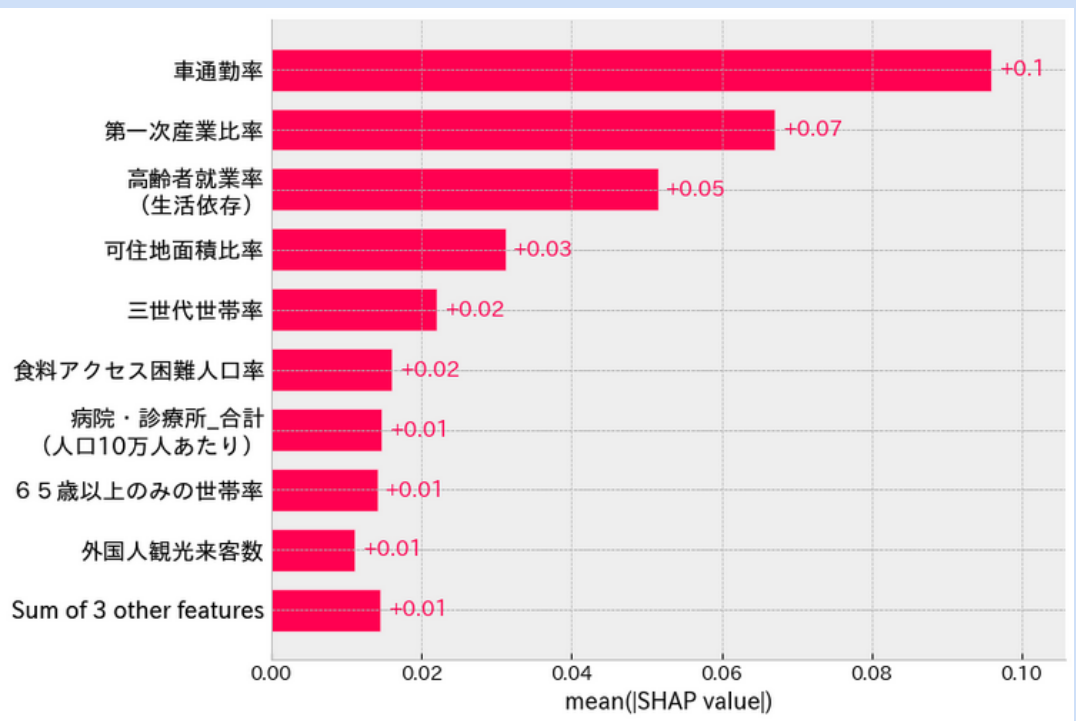
本分析では、「死亡事故件数を減らすための一つの手段として、免許の自主返納が一つの手段」と背景にて述べた。
よって、本章では返納率を上げるための政策だけでなく、返納率が上がらなくても事故リスクを下げられる手段も含めた「複合的アプローチ」に基づいて政策を提言する

施策案-SHAP値の全体傾向（再掲）

目的変数を返納率としたSHAP値の結果

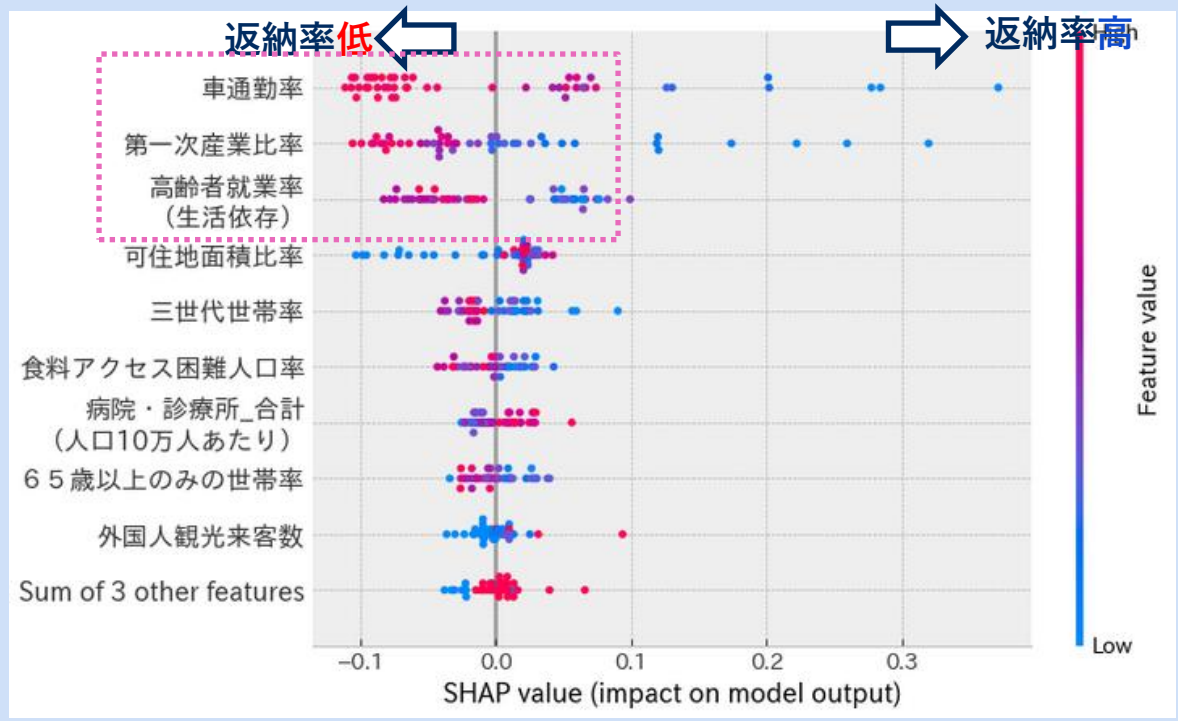
決定係数 R^2 : 0.933
MAE (平均絶対誤差): 0.075
RMSE (二乗誤差の平方根): 0.103

特徴量の重要度ランキング（目的変数にどれだけ影響を与えたか）



車通勤率 > 第一次産業比率 >
高齢者就業率 > 可住地面積比率

各特徴量が返納率の上げ下げにどの程度効いているのか？



車通勤率が**高**いほど、返納率**低**い傾向
第一次産業比率が**高**いほど、返納率**低**い傾向
高齢者就業率が**高**いほど、返納率が**低**い傾向

施策案-全体データへの意思決定支援案

	SHAP値
全体	◎車通勤率が 高い ほど、返納率 低い 傾向 ◎第一次産業比率が 高い ほど、返納率 低い 傾向 高齢者就業率が 高い ほど、返納率が 低い 傾向

◎：統計的に有意な特徴量



都市部・田舎部に問わず一定の傾向がみられるので、全国的な政策とし、統計的に有意な特徴量を優先

第一次産業比率が高いエリア

背景仮説

- ・都市部では第一次産業比率が低いので、免許返納率は高い
- ・田舎部では第一次産業比率が高く、仕事に免許が必要

施策案

- ・農業スマートシティ（補足ページ）

車通勤率が高いエリア

背景仮説

- ・車通勤を前提とした都市設計
- ・一部に郊外住宅地や工業地域あり
- ・公共交通はあるが、通勤には向かない

施策案

- ・高齢者リモート就労移行支援
- ・車が必要な通勤層に限定した“免許更新支援制度（免許返納制度をより厳しくした制度）”
- ・自動運転サービス（補足ページ）
- ・公共交通機関の増加を含めた地域の再編成⇒リ・デザイン（補足ページ）

高齢者就業率が高いエリア

背景仮説

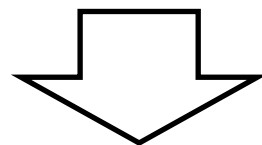
- ・都市部とはいえ、東京都や神奈川県でも人口の少ない過疎地がある（箱根や奥多摩のほう）

施策案

- ・郊外高齢者向けの移動スーパーなど
 - ・買い物支援
- ⇒移動販売、オンライン注文＋配達支援

施策案-意思決定支援（全体データ）

- ・車通勤率が高い地域ほど、免許返納率は低い⇒仕事に必要
- ・65歳以上のみの世帯率が高い地域ほど、免許返納率は低い⇒生活するために車が必要
- ・第一次産業比率が高い地域ほど、免許返納率は低い⇒農水産業で作物や魚の運搬に車が必要



これらの特徴量はいずれも高齢者が
車を手放しにくい要因

免許制度の見直し（**現実的対応策**）と技術的支援（**将来的対応策**）の2本立てで包括的に対応可能

- ・免許制度：免許返納制度の促進・PRとサポートカー限定免許の義務化
⇒地域によらず、死亡事故を減らすことが可能（現実的対応策）
- ・技術的アプローチ：自動運転システムの開発促進および地域公共交通のリ・デザイン化
⇒公共交通機関の増加により、免許の返納率向上、死亡事故の減少が可能（将来的な対応策）

※上記案の引用スライドはAppendix5へ

施策によって返納率が向上すれば、高齢者ドライバー数の減少を通じて
死亡者数の減少に寄与する可能性が高い

施策案-その他意思決定支援案まとめ

表. 意思決定支援案まとめ（全体データおよび各クラス）

クラスター\項目	政策領域1： 第一次産業比率高い	政策領域2： 車通勤率高い	政策領域3： 高齢者就業率高い	その他特徴的政策領域
全体◎	<ul style="list-style-type: none"> ・農業スマートシティ ・セーフティサポートカーの義務化 ・自動運搬サービス 	<ul style="list-style-type: none"> ・公共交通機関の強化 ・リモートワーク推進 ・免許更新条件強化 ・地域公共交通のり・デザイン 	<ul style="list-style-type: none"> ・高齢者向け移動支援 ・在宅勤務促進 	—
クラスター0	—	—	—	<ul style="list-style-type: none"> ・交通インフラ利用支援 ・コミュニティバス、シニアカーシェア ・食料配送、移動販売
クラスター1◎	<ul style="list-style-type: none"> ・安全運転訓練 ・農業スマートシティ ・セーフティサポートカーの義務化 ・自動運搬サービス 	—	<ul style="list-style-type: none"> ・高齢者医療 ・訪問診療拡充 ・オンライン診療拡充 	—
クラスター2	（全体と共通）	（全体と共通）	—	<ul style="list-style-type: none"> ・交通インフラ利用支援 ・コミュニティバス、シニアカーシェア
クラスター3	（全体と共通）	—	（全体と共通）	<ul style="list-style-type: none"> ・食料配送、移動販売 ・観光業の強化

※◎から優先的に施策を実施
 ※各クラスターの支援案はAppendix4へ

その他追加検討項目

追加検討

以下の項目については、短期的な施策が解決困難であったり、解釈に不確かさが残るため、追加検討が必要と考えた

- ・高齢者就業率：⇒都市設計・産業構造・ライフスタイルに関わる長期的な変革が必要なので、追加検討
- ・離島の有無：統計的に有意としても確認されているが、解釈に不確かさが残るため追加検討

本施策案はあくまで構造的分析をもとにした意思決定支援案であり、実行可能性の検証には、地域ごとの予算状況・既存施策・制度整備の進捗等を踏まえた追加検討が必要である。

全体まとめ

返納率が向上しない構造的特性とその意思決定支援について述べた

課題提起

高齢化が進んでおり、高齢者による死亡事故比率は増加傾向

⇒死亡事故件数を減らすための一つの手段として、免許の自主返納が一つの手段として挙げられる

分析結果

返納率が低い地域：車依存度が高く、農業従事者や高齢世帯が多い

返納率が高い地域：公共交通の充実・都市化が進んでいる

意思決定支援

現実的対応：免許返納制度の促進やサポートカー限定免許の義務化によって、死亡事故を減らす

中長期的対応：技術的アプローチ（自動運転技術の促進やスマートシティ化）や地域公共交通の再整備



施策によって返納率が向上すれば、高齢者ドライバー数の減少を通じて死亡者数の減少に寄与する可能性が高い

※仮に返納率が1%向上すれば、年間死亡者数は約300人減少する見込み

今後の展望

- ・データ量が足りず、回帰分析モデルの信頼性が弱くなったので、データ量を増やして改めて分析
- ・特徴量「離島の有無」がおそらく擬相関なので、さらなる構造的要因の探索と要因の特定
- ・特徴量「高齢者就業率」を使って、どう政策に反映していくかの追加検討

参考データ

データ	URLもしくは発行元
特徴量	
車依存率/バス依存率/鉄道依存率	国勢調査（R2、利用交通手段）
病院・診療所_合計（人口50万人当たり）	医療施設調査（R5）
食料アクセス困難人口率	国勢調査（R2年に基づく推計結果、農林水産政策研究所）
可住地面積比率	統計でみる都道府県のすがた2024 自然環境
三世代世帯率/65歳以上のみの世帯率/高齢者 単身率/高齢者就業率	国勢調査（R2、人口等基本集計、就業状態等基本集計）
政令指定都市の数	総務省、指定都市一覧
第一次産業比率	国勢調査（R2、就業状態等基本集計）
外国人観光来客数	日本政府観光局