

トイデータセットのデータの解説書

地域環境・動態データ項目

ファイル名：RegionalEnvironment.csv

地域環境・動態データは国土数値情報及び国勢調査（250mメッシュ）のデータ群を加工し、各データ項目を作成した。

データ項目

#	項目名	単位	型	備考	データソース
1	KEY_CODE	-	str	1次～5次メッシュコード（10桁）	5次メッシュ（250mメッシュ） （総務省統計局e-stat）
2	災害危険区域	-	bool	メッシュの重心が災害危険区域に含まれていれば真	災害危険区域データ：データ基準年令和3年度版 （国土数値情報）
3	津波浸水想定区域	-	bool	メッシュの重心が津波浸水想定区域に含まれていれば真	津波浸水想定データ：データ基準年令和4年度版 （国土数値情報）
4	小学校区	-	str	メッシュの重心の小学校区	小学校区データ：データ基準年令和3年度版 （国土数値情報）
5	小学校までの距離	m	float	上記学校までの距離	学校（ポイント）データ：データ基準年令和3年度版 （国土数値情報）
6	中学校区	-	str	メッシュの重心の中学校区	中学校区データ：データ基準年令和3年度版 （国土数値情報）
7	中学校までの距離	m	float	上記学校までの距離	学校（ポイント）データ：データ基準年令和3年度版 （国土数値情報）
8	二次医療圏	-	str	メッシュの重心の二次医療圏名	医療圏データ：データ基準年令和2年度版 （国土数値情報）
9	二次医療圏コード	-	int	二次医療圏コード	医療圏データ：データ基準年令和2年度版 （国土数値情報）
10	近隣の警察署	-	str	メッシュの重心から最も近い警察署	警察署データ：データ基準年平成24年度版 （国土数値情報）
11	警察署までの距離	m	float	上記までの距離	上記より加工
12	近隣のバス停	-	str	メッシュの重心から最も近いバス停	バス停データ：データ基準年令和4年度版 （国土数値情報）
13	バス停までの距離	m	float	メッシュの重心から最も近いバス停までの距離	上記より加工
14	人口	人	int		令和2年国勢調査－世界測地系（250mメッシュ）

トイデータセットのデータの解説書

地価公示データ項目

ファイル名 : Hausing.geojson

地価公示データは国土数値情報のデータを加工し、各データ項目を作成した。

データ項目

#	項目名	単位	型	備考	データソース
1	Geometry	-	geometry. Point		-
2	公示価格	円/m2	int		地価公示データ : データ基準年令和5年度版 (国土数値情報)
3	地積	m2	int		同上
4	間口比率	-	float		同上
5	建蔽率	-	float		同上
6	容積率	-	float		同上
7	利用現況	-	str		同上
8	建物構造	-	str		同上
9	水道・ガス・下水有無	-	bool		同上
10	前面道路の方位	-	str		同上
11	前面道路の幅員	m	float		同上
12	駅までの距離	m	int		同上

トイデータセットのデータの解説書

250mメッシュデータ項目

ファイル名 : Polygons.geojson

メッシュデータは総務省統計局 (e-stat) の5次メッシュ (250mメッシュ) より加工し作成した。

データ項目

#	項目名	単位	型	備考	データソース
1	Geometry	-	geometry. Polygon		-
2	KEY_CODE		str	1次~5次メッシュコード (10桁)	5次メッシュ (250mメッシュ) (総務省統計局e-stat)
3	MESH1_ID	-	str		同上
4	MESH2_ID	-	str		同上
5	MESH3_ID	-	str		同上
6	MESH4_ID	-	str		同上
7	MESH5_ID	-	str		同上
8	OBJ_ID	-	int		同上

トイデータセットのデータの解説書

分析ノートブックの概要

トイデータセットを使用した簡易的な分析ノートブックを公開します。公示価格を予測するモデルの作成を想定していますが、予測対象を地価公示データ以外の任意の地点データに変更することで、他ユースケースへの応用も可能です。

提供する分析ノートブック

LiniearRegression_OLS.ipynb

→最小2乗法を用いた線形回帰モデル（Pythonで実行可能）

LiniearRegression_縮小推定.ipynb

→リッジ回帰（縮小推定）を用いた線形回帰モデル（Pythonで実行可能）

LiniearRegression_変数選択.ipynb

→多重共線性を考慮し、変数選択した、最小2乗法を用いた線形回帰モデル（Pythonで実行可能）

LinearRegression_OLS.R

→最小2乗法を用いた線形回帰モデル（Rで実行可能）

LiniearRegression_縮小推定.R

→リッジ回帰（縮小推定）を用いた線形回帰モデル（Rで実行可能）