

2024年11月

第1回 国土交通省 地理空間情報データチャレンジ～国土数値情報編～ アイデア部門

ユーザー名: kageto

# 東京都のファミリー層向け不動産需要の地域分析

## 1. 背景と課題

東京都の不動産市場は多様な顧客層を抱える一方で、特にファミリー層向け物件の需要が増加傾向にある。ファミリー層は安定した住環境や子育てに適した地域を重視するため、地域ごとの需要特性を定量的に評価し、その地域がどの程度ファミリー層に適しているかを把握することが重要である。これは、効果的な不動産戦略を策定する上で欠かせない要素となる。

しかし、従来の不動産価格予測モデルは、物件の物理的特性と地域特性を直接組み合わせて需要を予測することが一般的であった。一方で、物件特性をトピックモデルによって抽象化し、「ファミリー層向け」といった特定のトピックを抽出して地域特性との関係性を評価する手法は、十分に研究されていない。

本レポートでは、物件データからファミリー層向け不動産需要を把握し、人口予測データと組み合わせることで、東京都内の地域ごとにファミリー層への投資戦略の初期指針として有用な知見を提供する。

## 2. 提案するアイデア

本レポートでは、以下の手法を用いて分析を行う。

### 1. 物件データのトピック抽出

物件の設備、仕様などのデータを基にトピック分析を実施し、ファミリー層向け物件の特性を識別する。

### 2. 階層ベイズモデルによる需要影響の分析

抽出した物件トピックがファミリー層の需要に与える影響を、階層ベイズモデルを用いて分析し、地域ごとの需要傾向を予測する。

### 3. 人口予測データとの統合分析

ファミリー層向け不動産需要は、将来的な若年層人口の動向に大きく影響されると考えられる。若年層(本レポートでは0歳～19歳と定義)の人口が増加する地域では、将来的にファミリー層の増加が期待でき、不動産需要も高まると予想される。一方、若年層人口が減少する地域では、ファミリー層の減少に伴い、不動産需要も低下する可能性がある。そこで、国土数値情報の人口予測データ<sup>1</sup>を用いて若年層人口の将来動向を分析し、現在のファミリー層向け需要との関連性を明らかにする。このアプローチにより、各地域の不動産需要の現状と将来展望を総合的に評価し、投資適性や整備の必要性について提言を行う。

## 3. データの前処理

前処理では、トピックモデルおよび階層ベイズモデルの構築に必要なデータの準備を行った。データは、民間企業から提供された物件データであり、物件の設備や仕様などの情報が含まれる。まず、この物件特性データに基づいてトピック抽出を行い、物件ごとの特徴を集約した後、得られたトピック情報を用いて価格予測モデルを構築する。

### 3.1 前処理と各変数の概要

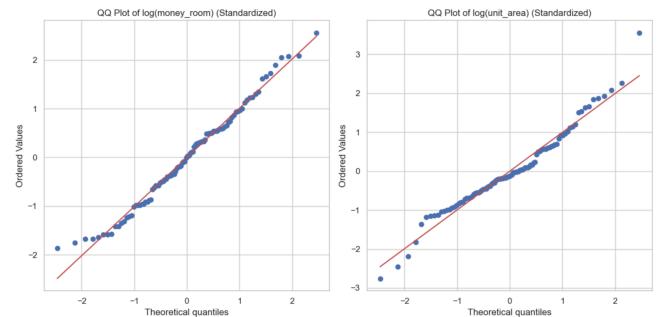
## 1. 設備情報の展開

設備情報(building\_tag\_idおよびstatuses)はスラッシュ区切りの形式で提供されているため、これを分割し各タグをバイナリ変数化した。

## 2. 価格と専有面積の対数変換と標準化

物件価格(money\_room)と専有面積(unit\_area)は、裾が重い分布を示すため対数変換を行い、正規分布に近似させた後、標準化を実施した。

以下に対数変換と標準化を実施した平米単価のQ-Qプロットを示す。



このプロットから、対数変換により正規分布にある程度近似できていることが確認できる。

最終的に、83796件のデータを用いてモデル構築と分析を行った。

## 4. 手法

本レポートでは、トピックモデルを用いたトピック抽出と、階層ベイズモデルによる地域別のトピック効果の推定を行う。

### 4.1 トピック抽出

設備・仕様データから物件特性に関するトピックを抽出するためには、ProdLDAを用いた。ProdLDAの詳細については、Srivastava & Sutton (2017) の論文<sup>2</sup>を参照されたい。

### 4.2 地域別トピック効果の推定

次に、抽出されたトピックをもとに地域ごとのトピック効果を推定し、物件価格に与える地域ごとの影響を考慮した階層ベイズモデル構築を行った。

### 4.3 モデル構成

物件*i*の価格 $y_i$ に対する線形予測子を以下に定義する。

$$y_i \sim N(\alpha_{r[i]} + \beta_{ua} u_i + \sum_{k=1}^K x_{i,k} \cdot \beta_{r[i],k} \sigma^2)$$

- $\alpha_{r[i]}$  地域毎のインセプトで、ある物件*i*の地域*r*の基本的な価格を示す。
- $\beta_{ua}$  ある物件*i*の平米数に対してグローバルに適用される係数を示す。一般にファミリー向け物件は広い平米数を有し、平米数が広いほど賃料も上昇する傾向にあるため、平米数の影響を調整する必要がある。平米数に起因する賃料への

影響を統計的に制御し、ファミリー層向け需要の純粋な効果を抽出することを目的として導入した。

また、ファミリー層向けのトピックへの多重共線性は事前分布の分散を1に狭める正則化や8万件以上のデータ数によって影響を低減できていると仮定する。

- $u_i$ ある物件 $i$ における平米数を示す。
- $\beta_{r[i],k}$ ある物件 $i$ 地域 $r$ におけるトピック $k$ の影響度であり、各トピックが地域特有の需要傾向にどの程度関わるかを示す。
- $x_{i,k}$ 物件 $i$ におけるトピック $k$ の強度を示し、トピックモデルによって得られる値である。

#### 4.4 パラメータの事前分布

モデル内の各パラメータに以下の事前分布を設定した。

- インターセプトの事前分布  
 $\alpha_{r[i]} \sim N(0, 1)$
- 地域毎のトピック効果の事前分布  
 $\beta_{r[i],k} \sim N(\mu_{\beta_k}, \sigma_{\beta_k}^2)$
- 平米数効果の事前分布  
 $\beta_{ua} \sim N(\mu_{\beta_{ua}}, \sigma_{\beta_{ua}}^2)$
- 観測データの分散の事前分布  
 $\sigma \sim Exponential(1)$

### 5. 実験結果

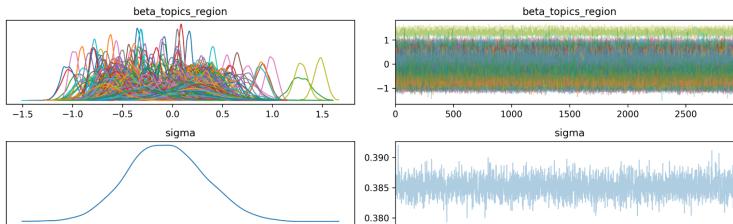
物件特性に関するトピックを抽出するために、設備や仕様データをもとにトピック数を5、7、10、13、15、17、20の7パターンでモデルを構築し、各モデルについてCoherenceスコアおよび20%のテストデータを用いたPerplexityスコアを評価した。

トピック数	coherence	perplexity
5	0.689	14.739
7	0.623	14.836
10	0.699	14.127
13	<b>0.724</b>	14.017
15	0.713	14.013
17	0.708	16.572
20	0.651	<b>13.965</b>

本レポートではファミリー層向けの特性を抽出することを主たる目的としているため、解釈性の指標であるCoherenceスコアが高いモデルを採用することが望ましい。上記の結果より、最も高いCoherenceスコア(0.724)を示したトピック数13のモデルを最適なモデルとして採用することとした。

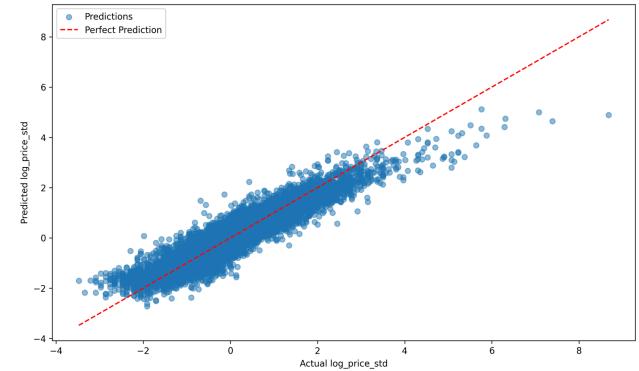
#### 5.1 トピック数13における階層ベイズモデルの予測精度

変分推論を用いて階層ベイズモデルのパラメータ推定結果の分布を以下に示す。



beta\_topics\_region(地域ごとのトピック効果)およびsigma(残差の標準偏差)に関するパラメータ分布が安定していることが確認できた。

また、テストデータに対して平米単価の予測を行い、実測値と予測値を比較した散布図を図に示す。



MSE: 0.146, R<sup>2</sup>: 0.854

若干の外れ値は見られるものの、データ点の多くは理想的な予測線(45度線)に沿って分布しており、モデルの予測が実際のデータと良好に一致していることがわかる。

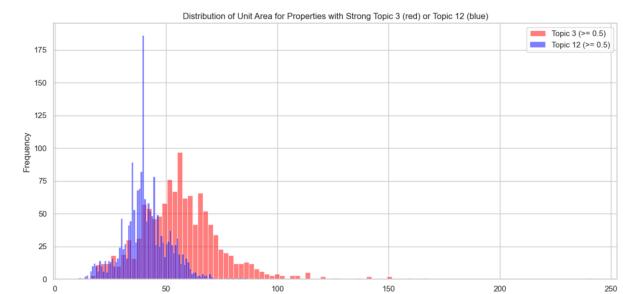
標準化されたベースラインモデル(MSE = 1.0)と比較すると、MSEが0.146に低減しており、モデルによって予測精度が大幅に向かっていることが示された。さらに、決定係数(R<sup>2</sup>)は0.854を示しており、モデルがデータの分散を約85.4%説明していることから、モデルの精度が高いことがわかる。

#### 5.2 トピックに対応するワードクラウドの結果

トピック分布に対応したワードクラウドを以下に示す。



特に、Topic 3とTopic 12は広めの間取りが特徴であり、ファミリー層向けの特性を持つと考えられるため、これら二つのトピックがどのように異なるかを検討する。各トピックの分布が0.5以上(半分以上の成分を持つ)である物件のみを抽出し、Topic 3とTopic 12について平米数の分布を比較するためのヒストグラムを図に示す。



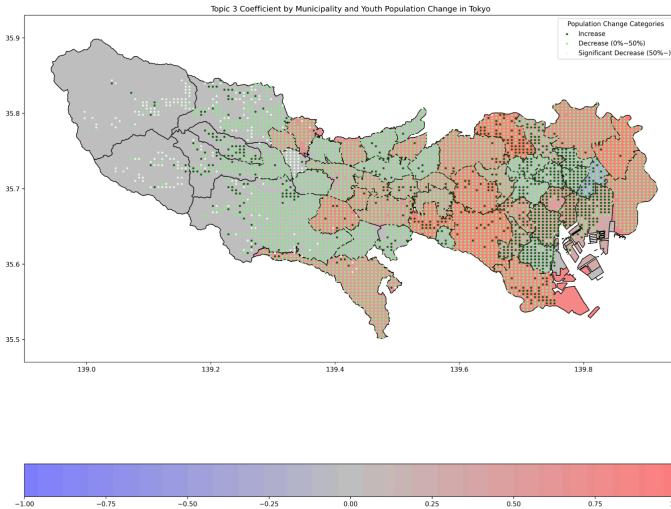
ヒストグラムの結果から、Topic 3は全体的に広い平米数の物件に関連していることが示され、ファミリー層向けの特性を持つと推察される。一方で、Topic 12は比較的狭い平米数の分布を示しており、子供のいない二人世帯や同棲カップル、あるいは広めの部屋を希望する単身者層に適した特性であると考えられる。

### 5.3 東京都地図上に人口予測データとTopicのマッピング

階層ベイズモデルにより推定されたTopic 3の係数が高いほど、ある地域におけるファミリー層の需要が高く、低いほど需要が低いと予測されるため、Topic 3の係数と国土交通省の人口予測データを、国土数値情報の行政区域データ<sup>3</sup>に基づいて東京都地図上にマッピングする。

Topic 3の係数については、信頼区間をHDI(Highest Density Interval)の3%~97%で評価し、0を含む場合は有意性がないと判断して係数を0に置き換えた。0を灰色、正の値を赤、負の値を青でスケーリングして色分けを行った。

さらに、人口予測データでは、2020年から2050年にかけての0歳~19歳の人口変動に基づいて3つのカテゴリに分類し、視覚化する。具体的には、人口が増加する地域を濃い緑、0~50%減少する地域を薄い緑、50%を超えて減少する地域を白で示す。これらのデータは、MESH\_IDから算出した緯度・経度に基づき、各メッシュにドットをプロットする形で表示する。



全体的に東京都内のファミリー向け物件の需要は増えており、若者の人口とある程度の相関関係があるように見える。

本分析では、家族向け物件の需要と若年層人口の動向に基づき、対象地域を以下の4つに分類した。

- ①ファミリー向け物件の需要が増える & 若者の人口が増える
- ②ファミリー向け物件の需要が増える & 若者の人口が減る
- ③ファミリー向け物件の需要が減る & 若者の人口が増える
- ④ファミリー向け物件の需要が減る & 若者の人口が減る

特に、①に該当する地域として、大田区、板橋区、世田谷区などが挙げられる。これらの地域では、将来的に家族層向けの需要がさらに増加することが見込まれ、安定した市場成長が期待できる。したがって、これらの地域はファミリー層が定住しやすく、ファミリー層向けの物件開発に適したエリアであると考えられる。

一方で、②に該当する江戸川区は、若年層人口が減少しているにもかかわらず、ファミリー層向け物件の需要が増加している。この状況は、ファミリー層に適した地域づくりや他地域からの誘致を積極的に行うための好機であると捉えられるが、将来的な人口減少リスクや複合災害リスクをも考慮した長期的な地域戦略が求められる。特に、災害対策の強化や防災インフラの整備は、ファミリー層が安心して定住できる環境を提供し、地域の魅力向上に寄与するため、不可欠な要素である。

③に関しては、23区内で唯一マイナスの影響が見られる墨田区が挙げられる。墨田区が実施した第7次墨田区住宅マスタープラン改

定基礎調査の第7章墨田区の住宅課題<sup>4</sup>によれば、「災害・治安への不安」「コンパクトな住宅が多く、家族向けには厳しい」「子育て世代の転出傾向がある」など、他の区にはない課題が指摘されている。また、単身世帯、外国人人口、後期高齢者の増加など、居住者の多様性が広がっていることもあり、整備がまだ整っていない可能性が示唆される。これらの要因が、家族向け物件の需要減少に影響を及ぼしていると考えられる。

最後に、④に該当する地域として、図には係数がマイナス傾向で人口減少が示されるエリアは含まれていないものの、多摩エリアがその可能性として挙げられる。この地域では、係数が0付近もしくはマイナス傾向にあり、データがさらに集まれば、ファミリー層向けの需要が一層低下する可能性が示唆される。また、若年層の人口も減少していることから、投資リスクが高い地域と考えられる。このような地域においては、新規の大規模開発や長期的な投資は控え、既存の老朽物件のリノベーションや短期賃貸物件の運営など、リスクを抑えた柔軟な戦略が求められる。

## 6. 結論

本レポートでは、物件特性データによるファミリー層向けの不動産需要予測と人口予測データを組み合わせて、地域ごとの評価を行うことで、東京都内におけるファミリー層向け不動産市場の動向を量的に分析した。この分析により、直感的に理解されていた地域特性をデータに基づいて可視化し、ファミリー層向けに適した地域が明確に示された。特に、大田区、板橋区、世田谷区などファミリー層向け需要が高く若年層人口が増加している地域では、安定した市場成長が期待でき、今後の開発や居住環境整備に向けた指針として妥当性が確認された。

また、若年層人口が減少する一方でファミリー層向け需要が増加している江戸川区では、災害対策や防災インフラの整備がファミリー層の安心定住に重要であると考えられるので、こうしたリスクを考慮した持続可能な地域戦略の策定が求められる。

一方、23区内で唯一ファミリー層向け需要が低下傾向にあると見られる墨田区については、子育て環境や住宅構造に関する課題が影響している可能性が示唆され、さらなる詳細な調査が求められる。さらに、多摩エリアについては、現時点で係数が0付近であり、データが集まればファミリー層向け需要が減少傾向に転じる可能性があるため、投資判断には慎重な検討が必要である。

本レポートの分析手法は、物件特性と人口動態データを組み合わせることで、ファミリー層向け不動産需要に基づく地域ごとの評価を可能にし、将来的な開発や投資の方向性に関する初期の指針を提供するものである。今後、追加データや他の地域・ターゲット層に対する分析を通じ、さらに充実した市場分析と地域特性に基づく戦略的な不動産開発が期待される。

## 7. 参考文献

1. 国土数値情報 (2018) 1kmメッシュ別将来推計人口 (H30国政局推計)  
[https://nlftp.mlit.go.jp/ksj/old/datalist/old\\_KsjTmplt-m1kh30.html](https://nlftp.mlit.go.jp/ksj/old/datalist/old_KsjTmplt-m1kh30.html)
2. Srivastava, A., & Sutton, C. (2017) Autoencoding Variational Inference For Topic Models  
<https://arxiv.org/abs/1703.01488>
3. 国土数値情報 (2024) 行政区域データ  
<https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-N03-2024.html>
4. 墨田区 (2022) 第7次墨田区住宅マスタープラン改定基礎調査 第7章 墨田区の住宅課題. 墨田区公式ウェブサイト  
[https://www.city.sumida.lg.jp/kurashi/zuyutaku/jyutaku\\_keikaku/jyuumasu7.files/dai7syou.pdf](https://www.city.sumida.lg.jp/kurashi/zuyutaku/jyutaku_keikaku/jyuumasu7.files/dai7syou.pdf)