

# 地域の特徴を踏まえたポスト・コロナにおける タクシー利用の需要予測

外部変数を追加した時系列データの予測手法検討

2023 年 5 月

デロイト トーマツ ファイナンシャルアドバイザー合同会社



## 1. はじめに

### 1.1 社会的背景

COVID-19 の蔓延により、世界中の人々の移動は制限され、観光業・飲食業・小売業といった様々な業種に多大な影響を与えました。今回のパンデミックのように、予測不可能な事態が生じた際や収束した後の需要を正確に予測することは、ビジネス上のリスクを軽減するうえで重要な課題であると考えられます。

### 1.2 分析概要

本稿では、統計分析手法である「需要予測」をビジネスに応用するための取り組みとして、COVID-19 の影響を受けたタクシー業界を題材に、COVID-19 が収束した後の「ニューヨークにおけるタクシー利用回数」の予測分析を行いました。

データには、ニューヨーク市政府が開示している「タクシー利用回数」のオープンデータを用いました。分析には、株式会社 Meta の開発した時系列分析用のオープンソースライブラリである「Prophet」(URL : [Prophet | Forecasting at scale. \(facebook.github.io\)](https://facebook.github.io/prophet/)) を用いました。

Prophet は、外れ値や欠損値の処理といったデータの前処理をせずとも、予測精度の高い分析を可能とします。加えて、予測モデルにドメイン知識として外部変数を追加することができるため、より柔軟性の高い分析を行うことができます。

## 2. ニューヨーク市の行政区ごとの概要と現状分析

### 2.1 行政区ごとの概要

ニューヨーク市の行政区ごとのタクシー利用客を分析するにあたり、各行政区の特色を比較しました。ニューヨーク市は5つの行政区によって構成され(図1参照)、行政区ごとの特色は下記のとおりです。

図1：ニューヨーク行政区の地図



1. Staten Island (特色：自由の女神が存在するリゾート地として有名)
2. Queens (特色：ニューヨーク市で最大の面積を持ち、JFK 空港が存在する)
3. Brooklyn (特色：ニューヨーク市で最も人口が多い住宅街)
4. Manhattan (特色：ニューヨーク市で最も人口密度が高く、ウォール街などが存在するビジネス街)
5. Bronx (特色：ニューヨーク市の最北端に位置し、ラップやヒップホップ発祥の地として有名)

### 2.2 現状分析

表1は、ニューヨーク市の行政区ごとの1日あたりの平均タクシー利用回数の推移を表しています。表1より、Manhattanは1日あたりの平均タクシー利用回数が最も多く、その次にQueensが多いことが読み取れます。一方、BrooklynやBronx、Staten Islandは1日あたりの平均タクシー利用回数があまり多くないことが読み取れます。

表 1：ニューヨーク市の行政区ごとの 1 日あたりの平均タクシー利用回数の推移

年度	Staten Island	Queens	Brooklyn	Manhattan	Bronx
2019	10.54	16203.84	2881.88	209880.05	487.29
2020	9.87	3857.38	1081.77	61374.96	403.16
2021	10.20	5768.07	996.25	76512.11	358.99
2022	8.66	9398.14	800.78	96644.41	141.53

図 2 は、ニューヨーク市の行政区ごとの 1 日あたりの平均タクシー利用回数の変化率と、変化率を地図上に可視化しています。

図 2：ニューヨーク市の行政区ごとの 1 日あたりの平均タクシー利用回数の変化率と地図上への可視化

(地図上の色が薄い青ほど変化率がマイナスであることを表し濃い青ほど変化率がプラスであることを表している)

各行政区ごとの1日あたりの平均タクシー利用回数の変化率

年	Staten Island	Queens	Brooklyn	Manhattan	Bronx
2020	-6.36%	-76.19%	-62.46%	-70.76%	-17.26%
2021	3.34%	49.53%	-7.91%	24.66%	-10.96%
2022	-15.10%	62.93%	-19.62%	26.31%	-60.58%



図 2 より、1 日あたりの平均タクシー利用回数の変化率において、COVID-19 の影響を受けた行政区と、あまり影響を受けていない行政区が存在することが読み取れます。具体的には、Queens や Brooklyn、Manhattan は 2020 年の変化率がマイナスかつ変化率が大きいことから、COVID-19 の影響を大きく受けていることが読み取れます。

一方、Staten Island と Bronx は 2020 年の 1 日あたりの平均タクシー利用回数の変化率が、他の行政区に比べて小さいことから、COVID-19 の影響をあまり受けていないことが読み取れます。

### 3. 予測結果

本稿では、ニューヨーク市の行政区の中でも、COVID-19の影響を強く受けた「Queens」と「Manhattan」、「Brooklyn」に焦点を当て、コロナ後のタクシー利用回数どのように変化するか予測分析を行いました。その際、行政区ごとの特色に応じて、予測モデルに空港や地下鉄の利用者数などの外部変数を追加し、予測精度が向上するかを検証しました。外部変数を追加した場合と追加していない場合における、予測精度の結果は表 2 に記しています。

#### 3.1 学習と予測に使用したデータ期間

表 2 は、行政区ごとの「学習・予測に使用したデータの期間」と「データ単位」を記しています。

Manhattan と Brooklyn については、日次におけるタクシー利用回数を予測しています。「2020 年 2 月 29 日～2022 年 9 月 30 日」のデータに基づき、「2022 年 10 月 01 日～2022 年 11 月 30 日」の日次のタクシー利用回数を予測しました。

一方、Queens については、分析の都合上、月次におけるタクシー利用回数を予測しています。「2019 年 2 月～2022 年 1 月」のデータに基づき、「2022 年 2 月 01 日～2022 年 11 月」の月次タクシー利用回数を予測しました。

表 2：行政区ごとの学習・予測に用いたデータ期間とデータの間隔

行政区	使用したデータの期間		データ単位
	学習	予測	
Queens	2019 年 2 月～2022 年 1 月	2022 年 2 月～2022 年 11 月	月次
Manhattan	2020 年 2 月 29 日～2022 年 9 月 30 日	2022 年 10 月 01 日～2022 年 11 月 30 日	日次
Brooklyn	2020 年 2 月 29 日～2022 年 9 月 30 日	2022 年 10 月 01 日～2022 年 11 月 30 日	日次

#### 3.2 予測精度の指標

本稿では、予測精度の指標として、平均絶対パーセント誤差(以下、MAPE)を使用しました。MAPE は「予測値と正解値の差を、正解値で割った値」の絶対値の総和を、データ数で除算した値であり、値が小さいほど精度が高いとされる指標です。主に時系列予測で用いられる指標であり、誤差をパーセントで評価できるため、スケールの異なる予測にも対応可能となっています。

予測精度は MAPE を用いて、下記のように定義しました。

予測精度の指標 =  $1 - \text{MAPE}$

表 3：「外部変数を追加しない場合」と「外部変数を追加した場合」の予測精度の比較

行政区	外部変数を追加しない場合	外部変数を追加した場合	外部変数
Queens	47.51%	96.34%	空港利用者数
Manhattan	90.35%	91.82%	地下鉄利用者数
Brooklyn	76.80%	75.32%	地下鉄利用者数

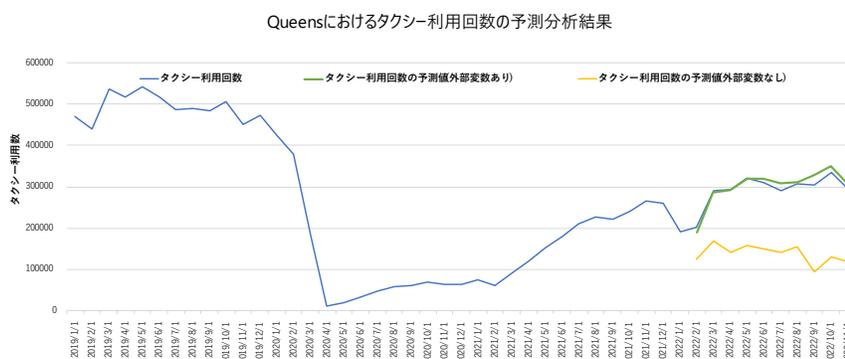
### 3.3 Queens の予測結果

Queens は JFK 空港が存在し、空港からタクシーを利用する客が多く存在すると考えられるため、「タクシー利用回数」と「空港利用者数」には関連があると想定しました。したがって、予測モデルの外部変数に「空港利用者数」を追加し、分析を行いました。

表 3 より、Queens は外部変数を追加した場合、大幅に予測精度が向上することが読み取れます。また、図 3 より、外部変数を追加しない場合のタクシー利用回数の予測値(黄色)は、タクシー利用回数の実測値(青色)と乖離していることが読み取れます。一方、外部変数を追加しない場合のタクシー利用回数の予測値(緑色)は、タクシー利用回数の実測値(青色)と近似しています。

図 3：空港利用者数を外部変数に追加した場合の、Queens の予測分析結果

(青色は「タクシー利用回数の実測値」、緑色は「タクシー利用回数の予測値(外部変数を追加した場合)」、黄色は「緑色は「タクシー利用回数の予測値(外部変数を追加しない場合)」を表しています。)



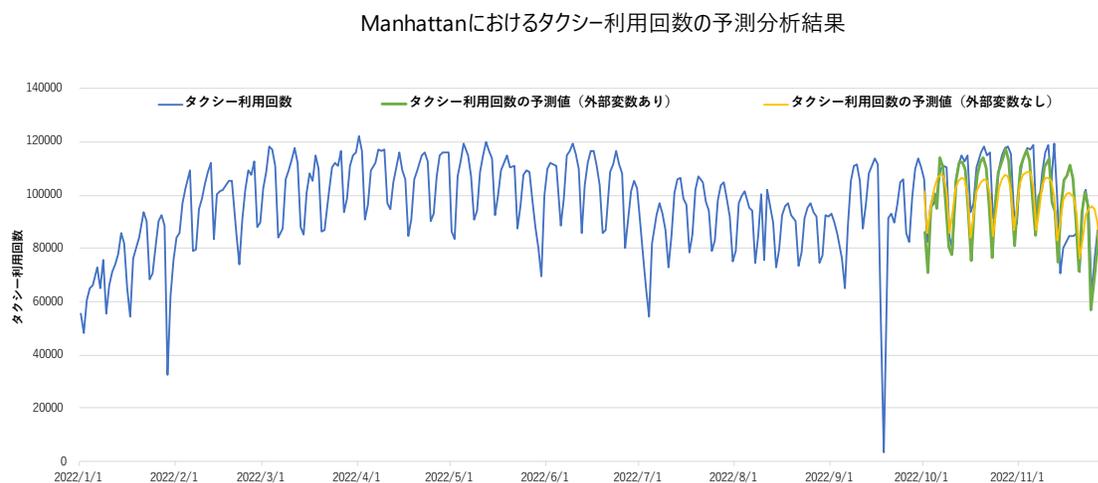
### 3.4 Manhattan の予測結果

Manhattan はアメリカを代表するビジネス街であるため、通勤のために地下鉄を利用する人々が多いと考え、時系列予測モデルの外部変数に「地下鉄利用者数」を追加しました。

表 3 より、Manhattan は外部変数を追加した場合、予測精度が少し向上することが読み取れます。また、図 4 より、外部変数を追加しない場合のタクシー利用回数の予測値(黄色)よりも、外部変数を追加しない場合のタクシー利用回数の予測値(緑色)のほうが、タクシー利用回数の実測値(青色)に近似していることが読み取れます。

図 4：地下鉄利用者数を外部変数に追加した場合の、Manhattan の予測分析結果

(青色は「タクシー利用回数の実測値」、緑色は「タクシー利用回数の予測値(外部変数を追加した場合)」、黄色は「緑色は「タクシー利用回数の予測値(外部変数を追加しない場合)」を表しています。)

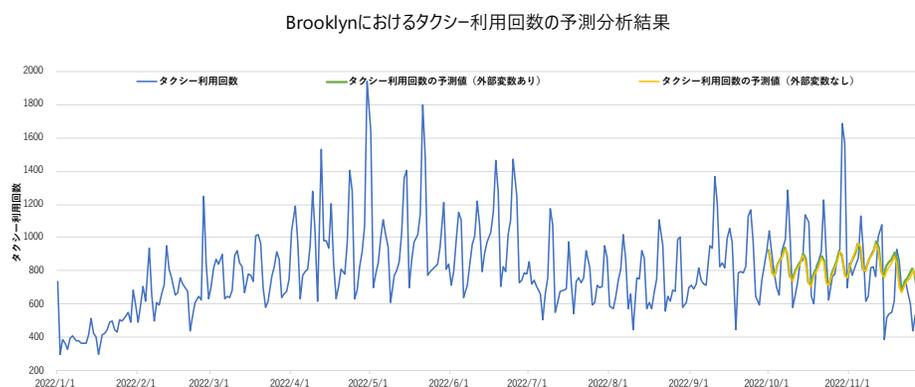


### 3.5 Brooklyn の予測結果

表 3 より、Brooklyn における予測精度については、外部変数を追加した場合/追加しない場合のどちらも低い傾向にあります。また、表 3 より Brooklyn は外部変数を追加した場合、予測精度が少し低下することが読み取れます。図 5 より、外部変数を追加しない場合のタクシー利用回数の予測値(黄色)と外部変数を追加しない場合のタクシー利用回数の予測値(緑色)のどちらも、タクシー利用回数の実測値(青色)と乖離していることが読み取れます。

図 5：地下鉄利用者数を外部変数に追加した場合の、Brooklyn の予測分析結果

(青色は「タクシー利用回数」、緑色は「タクシー利用回数の予測値(外部変数を追加した場合)」、黄色は「タクシー利用回数の予測値(外部変数を追加しない場合)」を表しています。)



#### 4. 考察

表 3 より、予測モデルに外部変数を追加した場合、Queens と Manhattan の分析結果の予測精度が向上したことが読み取れます。特に Queens は、予測精度が大幅に向上しており(47.51% vs. 96.34%)、外部変数によって地域ごとの特色を反映させることができたことを示唆しています。

一方、Brooklyn は他の行政区とは異なり、予測モデルに外部変数を追加した場合、予測精度が低下しました。Manhattan と Brooklyn はどちらも、予測モデルの外部変数には「地下鉄利用者数」を使用しましたが、Manhattan は予測精度がやや向上し、Brooklyn は予測精度がやや低下するという、反対の結果が得られました。

Manhattan はビジネス街であるため、地下鉄の利用者数が多いと考えられます。一方、Brooklyn は住宅街であるため、地下鉄の利用者数はあまり多くないと考えられます。そのため、Brooklyn の予測モデルの外部変数に「地下鉄の利用者数」を用いることは、あまり適しておらず、予測精度が低下したと考えられます。

したがって、Brooklyn の予測モデルの外部変数には、「地下鉄利用者数」ではなく、たとえば「車通勤の人数」や「昼間人口」などを追加すれば、予測精度はより向上すると考えられます。ただし、そのようなオープンデータが公開されていなかったため、今回は分析を行っておりません。

## 5. おわりに

本稿では、「ニューヨークの行政区におけるタクシー利用回数」のオープンデータを用いて、COVID-19 が収束した後のタクシー利用回数の予測分析を行いました。その際、ニューヨークにおける行政区の特色に応じて、予測モデルに外部変数を追加することで、COVID-19 後のタクシー利用回数を正確に予測することができました。

今回の分析では、ニューヨークのタクシー利用客数を題材としましたが、地域ごとの特色に応じた「需要予測」としても応用することができます。

例えば、分析対象の地域を変更し、日本の関東地域における「年間旅行者数」を予測したいとします。東京は Manhattan と同様にビジネス街であるため、外部変数には「電車の利用者数」といった要素が大きく影響し、千葉などは東京のベッドタウンであるため「昼間人口」が外部変数として用いることができると考えられます。

需要予測は分析の適用範囲が広く、他にも製造業であれば、製品の特色に応じた売り上げ数の予測などにも応用することが可能です。

## 問い合わせ先：

デロイトトーマツ ファイナンシャルアドバイザー合同会社

Digital 何 嘉杭

Digital 鈴木 雅貴

〒100-8363 東京都千代田区丸の内 3-2-3 丸の内二重橋ビルディング

Tel 03-6213-1180

E-mail dtfa.koho@tohmatsu.co.jp

デロイトトーマツグループは、日本におけるデロイト アジア パシフィック リミテッドおよびデロイトネットワークのメンバーであるデロイトトーマツ合同会社ならびにそのグループ法人（有限責任監査法人トーマツ、デロイトトーマツ コンサルティング 合同会社、デロイトトーマツ ファイナンシャルアドバイザー 合同会社、デロイトトーマツ 税理士法人、DT 弁護士法人およびデロイトトーマツ コーポレート ソリューション 合同会社を含む）の総称です。デロイトトーマツグループは、日本で最大級のプロフェッショナルグループのひとつであり、各法人がそれぞれの適用法令に従い、監査・保証業務、リスクアドバイザー、コンサルティング、ファイナンシャルアドバイザー、税務、法務等を提供しています。また、国内約 30 都市に約 1 万 7 千名の専門家を擁し、多国籍企業や主要な日本企業をクライアントとしています。詳細はデロイトトーマツグループ Web サイト（[www.deloitte.com/jp](http://www.deloitte.com/jp)）をご覧ください。

Deloitte（デロイト）とは、デロイト トウシュ トーマツ リミテッド（“DTTL”）、そのグローバルネットワーク組織を構成するメンバーファームおよびそれらの関係法人（総称して“デロイトネットワーク”）のひとつまたは複数指します。DTTL（または“Deloitte Global”）ならびに各メンバーファームおよび関係法人はそれぞれ法的に独立した別個の組織体であり、第三者に関して相互に義務を課しまたは拘束させることはありません。DTTL および DTTL の各メンバーファームならびに関係法人は、自らの作為および不作為についてのみ責任を負い、互いに他のファームまたは関係法人の作為および不作為について責任を負うものではありません。DTTL はクライアントへのサービス提供を行いません。詳細は [www.deloitte.com/jp/about](http://www.deloitte.com/jp/about) をご覧ください。

デロイト アジア パシフィック リミテッドは DTTL のメンバーファームであり、保証有限責任会社です。デロイト アジア パシフィック リミテッドのメンバーおよびそれらの関係法人は、それぞれ法的に独立した別個の組織体であり、アジア パシフィック における 100 を超える都市（オ克兰ド、バンコク、北京、ベンガルール、ハノイ、香港、ジャカルタ、クアラルンプール、マニラ、メルボルン、ムンバイ、ニューデリー、大阪、ソウル、上海、シンガポール、シドニー、台北、東京を含む）にてサービスを提供しています。

Deloitte（デロイト）は、監査・保証業務、コンサルティング、ファイナンシャルアドバイザー、リスクアドバイザー、税務、法務などに関連する最先端のサービスを、Fortune Global 500®の約 9 割の企業や多数のプライベート（非公開）企業を含むクライアントに提供しています。デロイトは、資本市場に対する社会的な信頼を高め、クライアントの変革と繁栄を促し、より豊かな経済、公正な社会、持続可能な世界の実現に向けて自ら率先して取り組むことを通じて、計測可能で継続性のある成果をもたらすプロフェッショナルの集団です。デロイトは、創設以来 175 年余りの歴史を有し、150 を超える国・地域にわたって活動を展開しています。“Making an impact that matters”をパーパス（存在理由）として標榜するデロイトの約 415,000 名の人材の活動の詳細については、([www.deloitte.com](http://www.deloitte.com)) をご覧ください。

本資料は皆様への情報提供として一般的な情報を掲載するのみであり、DTTL、そのグローバルネットワーク組織を構成するメンバーファームおよびそれらの関係法人が本資料をもって専門的な助言やサービスを提供するものではありません。皆様の財務または事業に影響を与えるような意思決定または行動をされる前に、適切な専門家にご相談ください。本資料における情報の正確性や完全性に関して、いかなる表明、保証または確約（明示・黙示を問いません）をすることもありません。また DTTL、そのメンバーファーム、関係法人、社員・職員または代理人のいずれも、本資料に依拠した人に関係して直接または間接に発生したいかなる損失および損害に対して責任を負いません。

© 2023. For information, contact Deloitte Tohmatsu Financial Advisory LLC



IS 669126 / ISO 27001

Member of  
Deloitte Touche Tohmatsu Limited