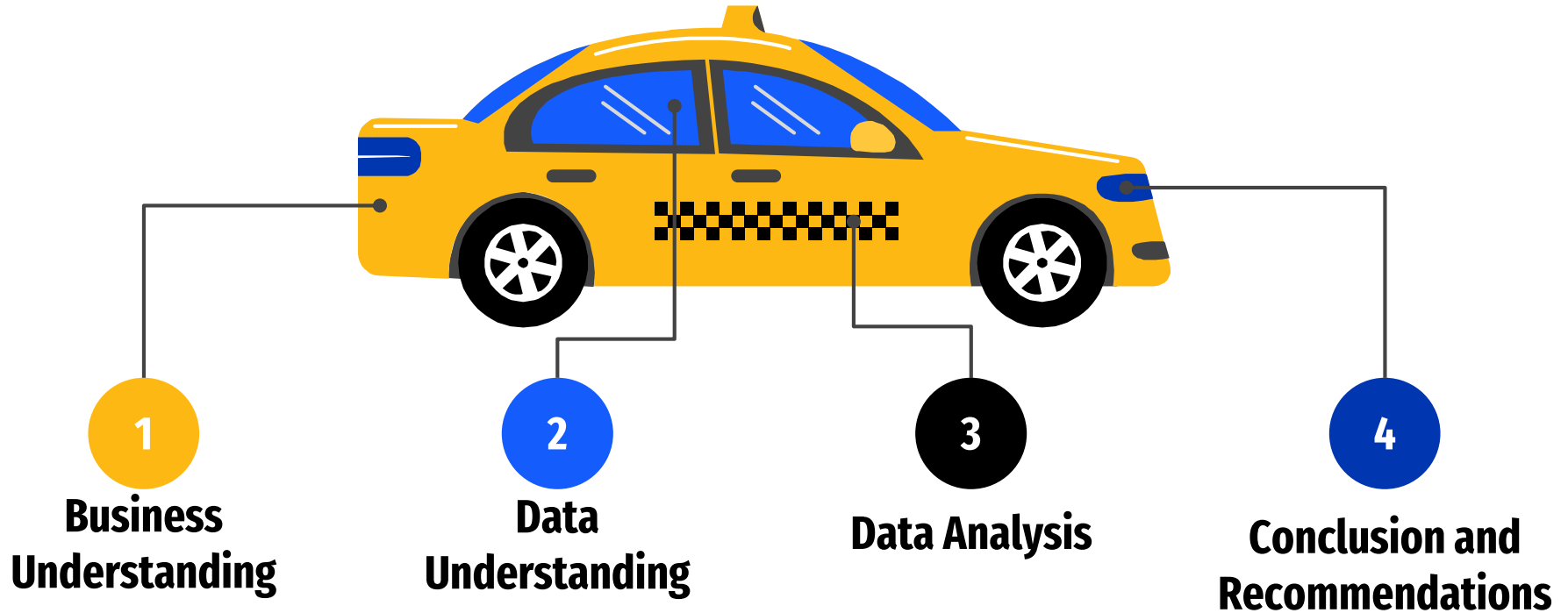


New York City Taxi and Limousine Commission (TLC) Trip Record

Gian Habli Maulana



Table of Contents



1. Business Understanding



1. Business Understanding

Background :



- Layanan taksi di New York City, yang diatur oleh Taxi and Limousine Commission (TLC), merupakan bagian integral dari sistem transportasi kota.
- Dengan populasi dan jumlah penduduk yang besar, permintaan akan layanan taksi sangat tinggi dan berfluktuasi sepanjang hari.
- Data perjalanan yang direkam oleh TLC mencakup informasi detail tentang waktu dan lokasi penjemputan serta pengantaran.
- Dengan menerapkan Exploratory Data Analysis (EDA) pada data ini, kita dapat mengeksplorasi pola-pola permintaan taksi dan memahami dinamika operasionalnya secara lebih mendalam.



1. Business Understanding

Problem Statement :

- Pada industri taksi, memahami pola permintaan pelanggan, preferensi wilayah, dan faktor-faktor yang mempengaruhi kepuasan pelanggan adalah kunci untuk mengoptimalkan operasi dan meningkatkan revenue.
- Analisis data diperlukan untuk mengidentifikasi pola permintaan pelanggan, preferensi wilayah, dan kepuasan pelanggan.
- Data ini dapat mencakup informasi tentang waktu dan lokasi penjemputan, jarak perjalanan, durasi perjalanan, tarif, metode pembayaran, dan tip.
- Dengan menggunakan teknik analisis data perusahaan dapat memperoleh wawasan yang berharga tentang perilaku pelanggan dan preferensi mereka.



1. Business Understanding

Goals:



1. Analyze Taxi Demand in New York City
2. Explore Customer Preferences and Behavior
3. Evaluate Fare Structures for NYC Taxi Services
4. Investigate Factors Affecting Tips for Taxi Drivers
5. Develop Recommendations for Improving Taxi Services in New York City



1. Business Understanding

Key Stakeholders:

- **Pemerintah Kota New York:** Bertanggung jawab dalam mengatur layanan taksi dan memastikan efisiensi sistem transportasi.
- **Taxi and Limousine Commission (TLC):** Badan pemerintah yang bertanggung jawab mengawasi layanan taksi di New York City.
- **Operator Taksi:** Perusahaan atau individu yang mengoperasikan armada taksi di New York City.



2. Data Understanding



2. Data Understanding

Overview Dataset:

- **Sumber:**
Website Resmi NYC Taxi and Limousine Commission (TLC)
<https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page>
- **Dataset:**
 1. **Dataset Utama** : NYC TLC Trip Record.csv
 2. **Dataset Pendukung**: Taxi Zone Shapefile (PARQUET)
- **Dimensi:**
 1. **Baris** : 69.600
 2. **Kolom** : 34



2. Data Understanding

Overview Dataset:

#	Column	Non-Null Count	Dtype
0	VendorID	69600 non-null	int64
1	lpep_pickup_datetime	69600 non-null	object
2	lpep_dropoff_datetime	69600 non-null	object
3	store_and_fwd_flag	65240 non-null	object
4	RatecodeID	65240 non-null	float64
5	PULocationID	69600 non-null	int64
6	DOLocationID	69600 non-null	int64
7	passenger_count	65240 non-null	float64
8	trip_distance	69600 non-null	float64
9	fare_amount	69600 non-null	float64
10	extra	69600 non-null	float64
11	mta_tax	69600 non-null	float64
12	tip_amount	69600 non-null	float64
13	tolls_amount	69600 non-null	float64
14	ehail_fee	0 non-null	float64
15	improvement_surcharge	69600 non-null	float64
16	total_amount	69600 non-null	float64
17	payment_type	65240 non-null	float64
18	trip_type	65230 non-null	float64

#	Column	Non-Null Count	Dtype
19	congestion_surcharge	65240 non-null	float64
20	PUOBJECTID	69397 non-null	float64
21	PUShape_Leng	69397 non-null	float64
22	PUShape_Area	69397 non-null	float64
23	PUzone	69397 non-null	object
24	PUborough	69397 non-null	object
25	PULongitude	69397 non-null	float64
26	PULatitude	69397 non-null	float64
27	DOOBJECTID	68912 non-null	float64
28	DOShape_Leng	68912 non-null	float64
29	DOShape_Area	68912 non-null	float64
30	DOzone	68912 non-null	object
31	DOborough	68912 non-null	object
32	DOLongitude	68912 non-null	float64
33	DOLatitude	68912 non-null	float64

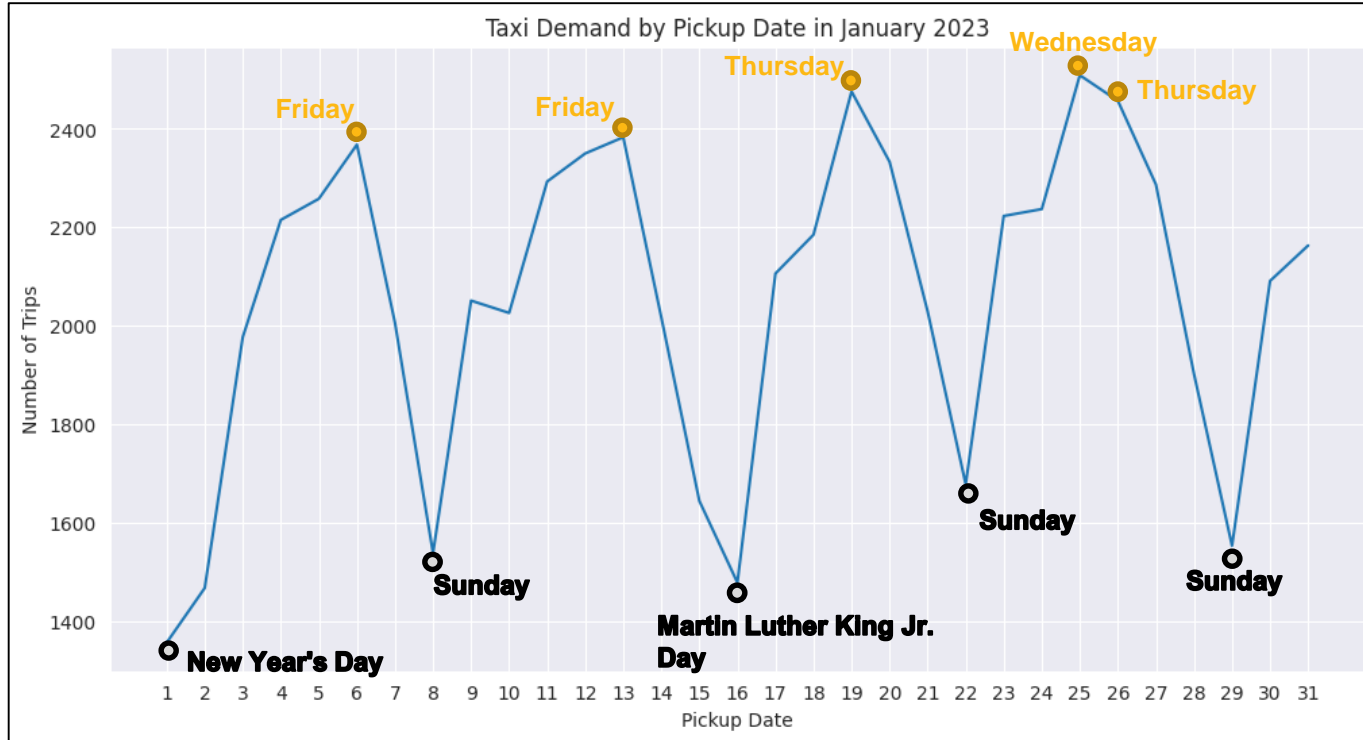


3. Data Analysis



3. Data Analysis Taxi Demand Patterns

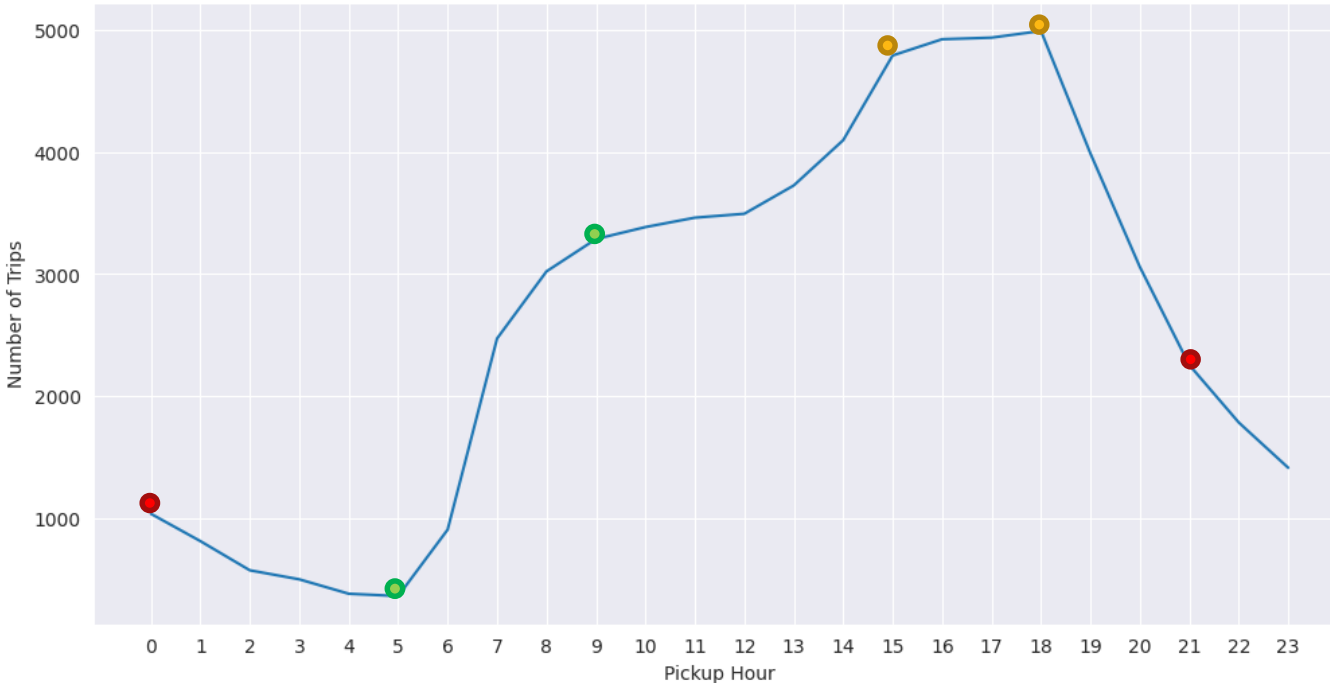
1. What are the peak dates for taxi demand?



3. Data Analysis Taxi Demand Patterns

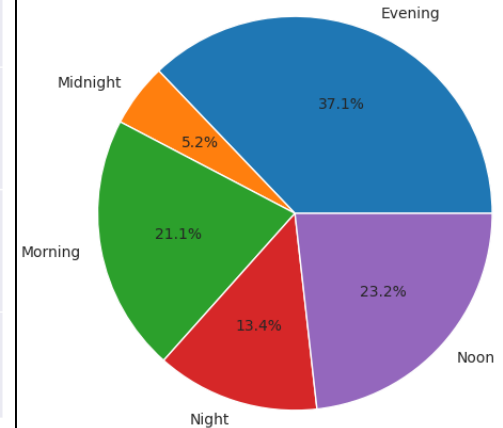
2. What are the peak hours for taxi demand?

Taxi Demand by Pickup Hour in January 2023



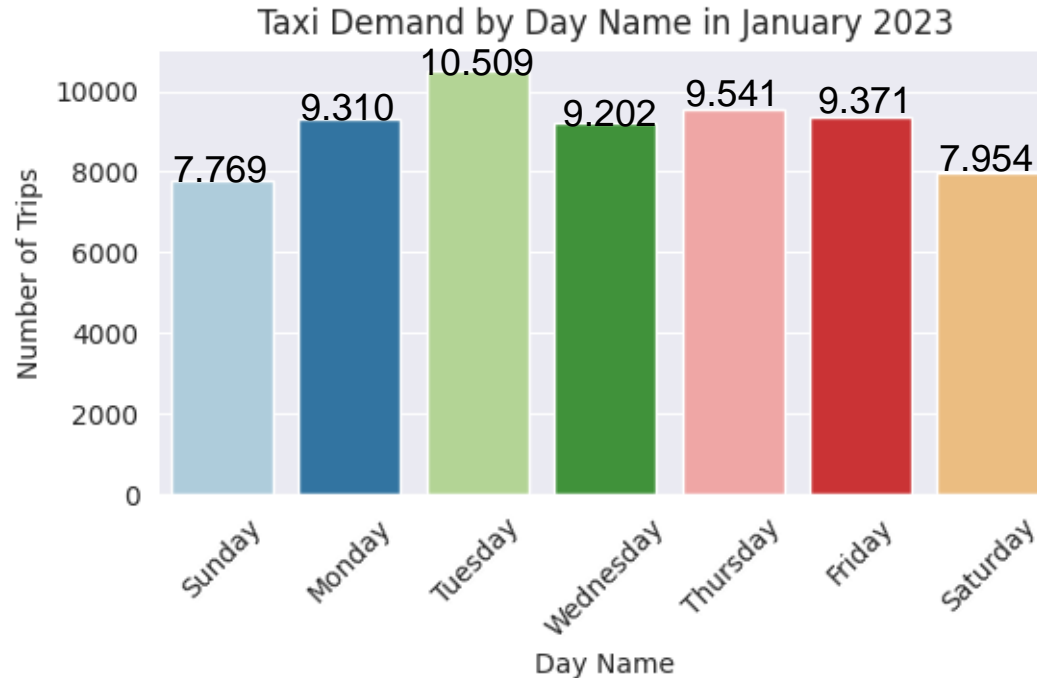
1. Midnight : 00.00 - 05.00
2. Morning : 05.00 - 11.00
3. Noon : 11.00 - 15.00
4. Evening : 15.00 - 20.00
5. Night : 20.00 - 24.00

Taxi Demand by Pickup Time Category in January 2023

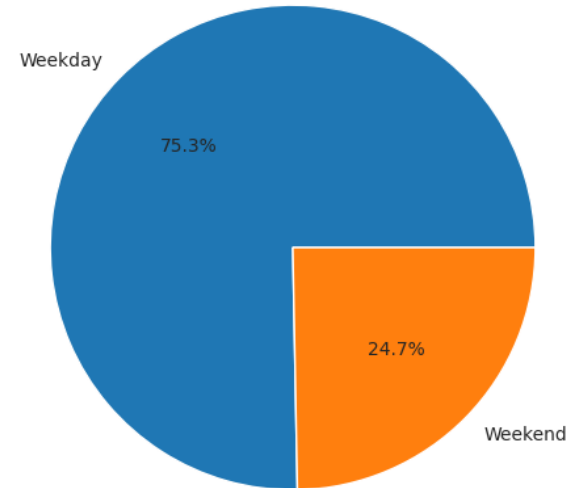


3. Data Analysis Taxi Demand Patterns

3. How does demand vary across different days of the week?

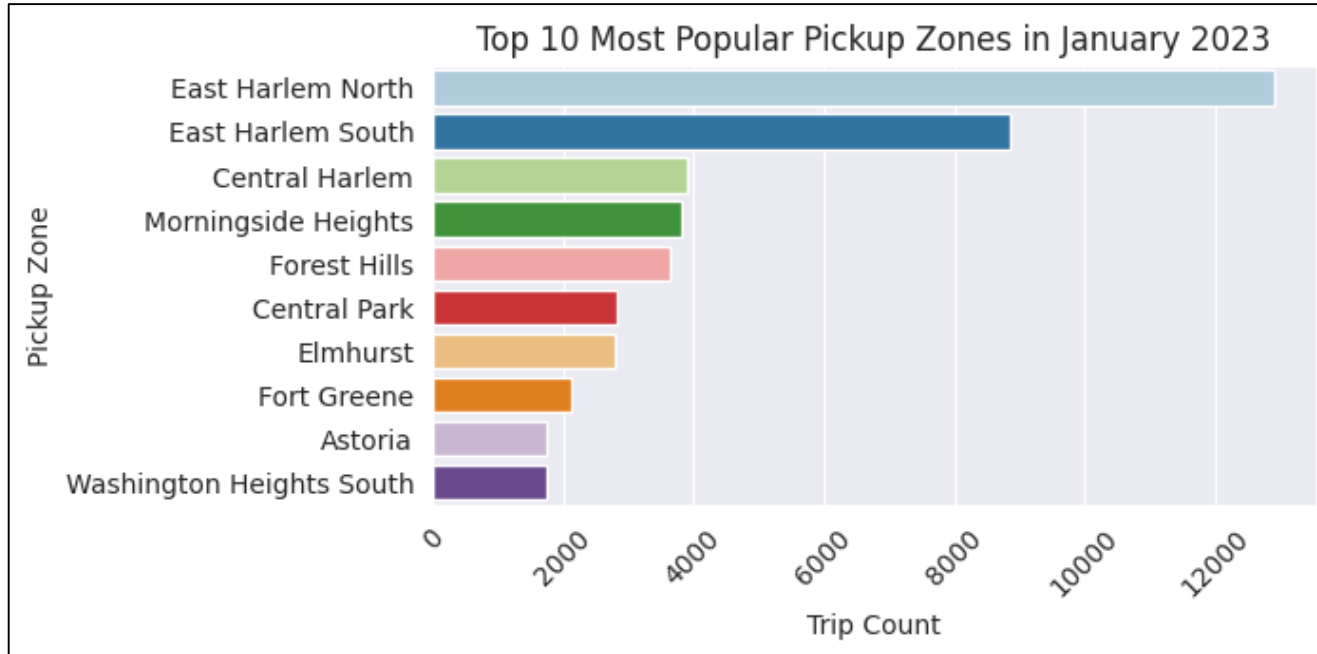


Taxi Demand by Day Category in January 2023



3. Data Analysis Taxi Demand Patterns

4. What are the most popular pickup locations?

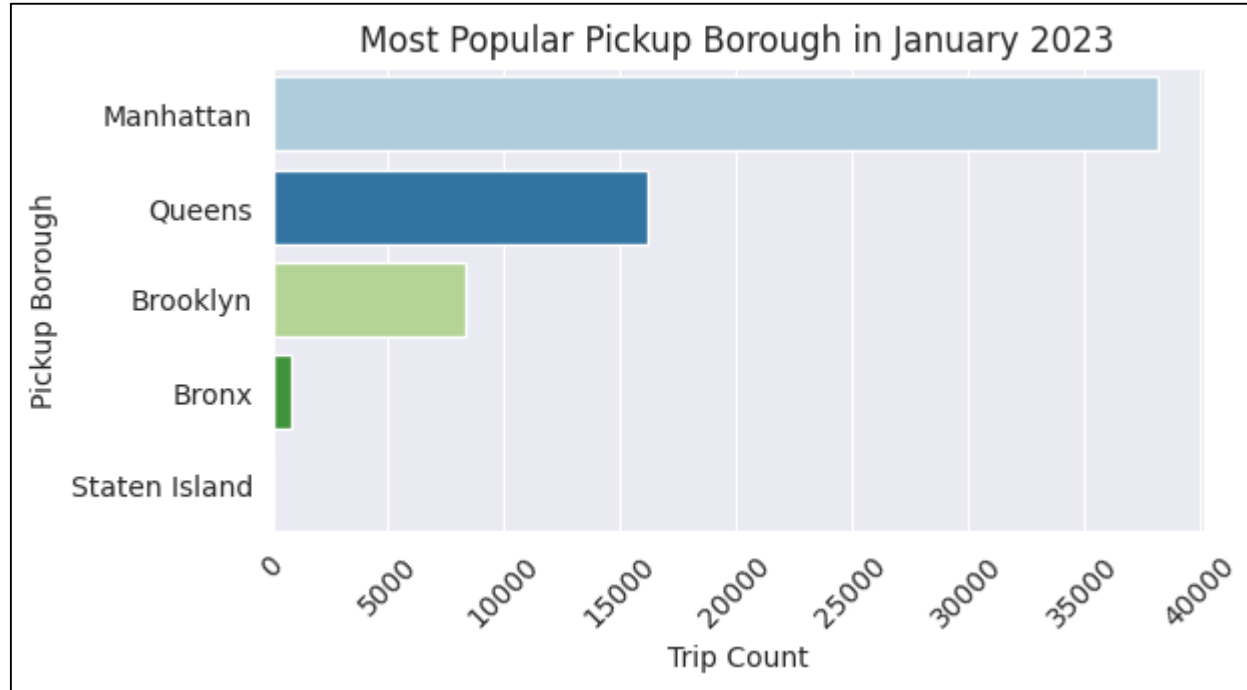


PUzone	trip_count
East Harlem North	12897
East Harlem South	8864
Central Harlem	3905
Morningside Heights	3818
Forest Hills	3645
Central Park	2838
Elmhurst	2789
Fort Greene	2120
Astoria	1748
Washington Heights South	1747



3. Data Analysis Taxi Demand Patterns

4. What are the most popular pickup locations?



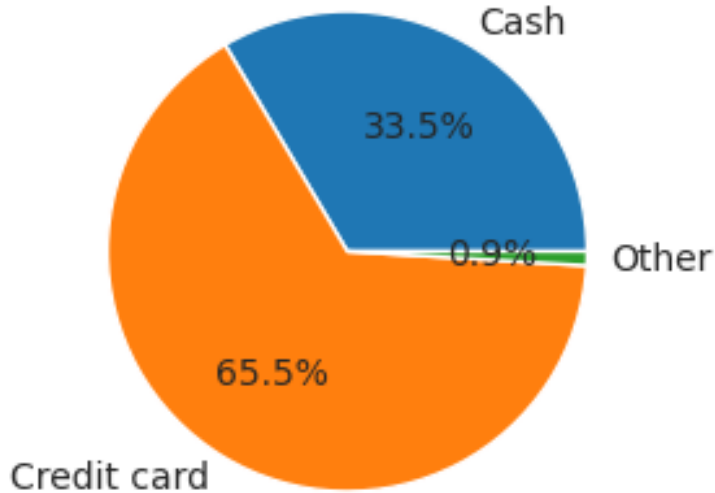
PUBorough	trip_count
Manhattan	38244
Queens	16225
Brooklyn	8341
Bronx	838
Staten Island	8



3. Data Analysis Customer Preferences and Behavior

1. What are the most common payment types and trip types?

Number of Trips by Payment Types in January 2023

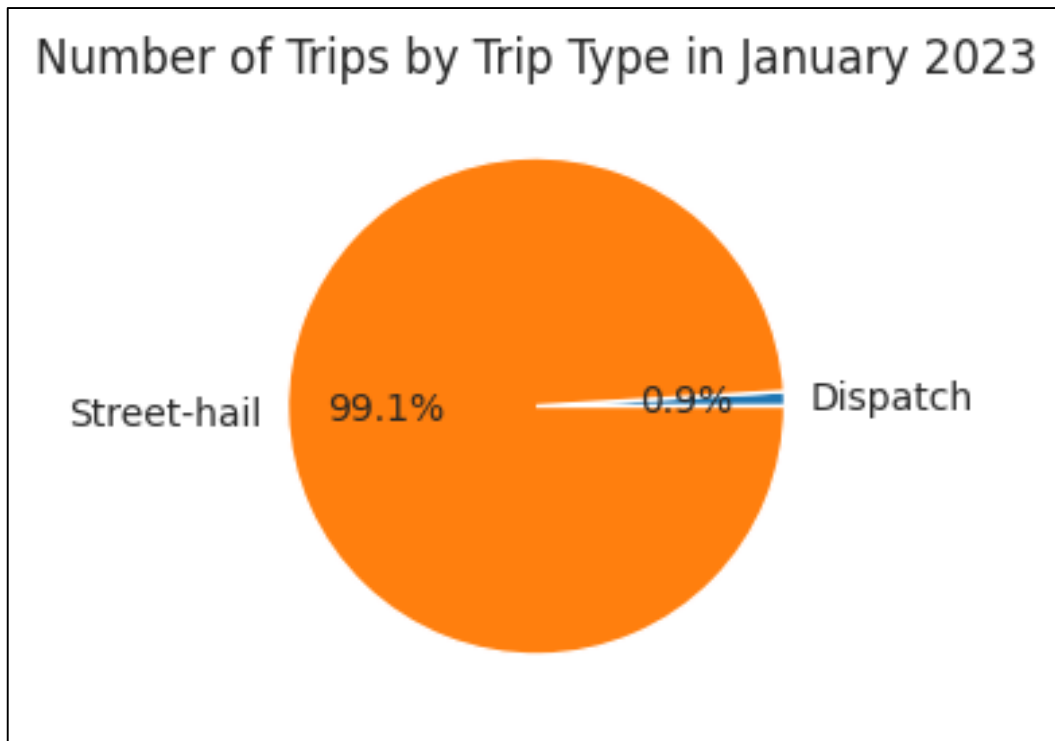


payment_type	trip_count
Cash	21335
Credit card	41721
Dispute	30
No charge	140
Unknown	430



3. Data Analysis Customer Preferences and Behavior

1. What are the most common payment types and trip types?



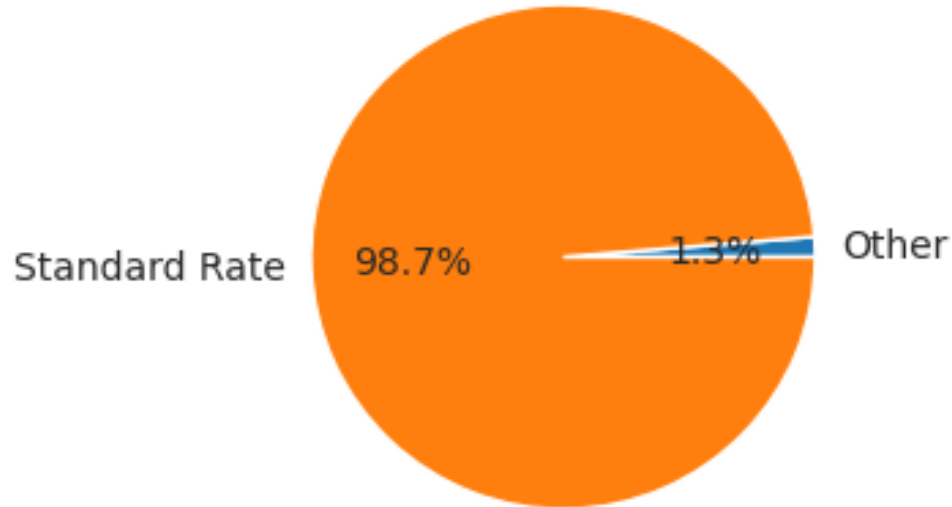
trip_type	trip_count
Dispatch	600
Street-hail	63056



3. Data Analysis Customer Preferences and Behavior

2. How do customers' preferences vary by rate, trip distance and duration?

Number of Trips by Trip Rate in January 2023

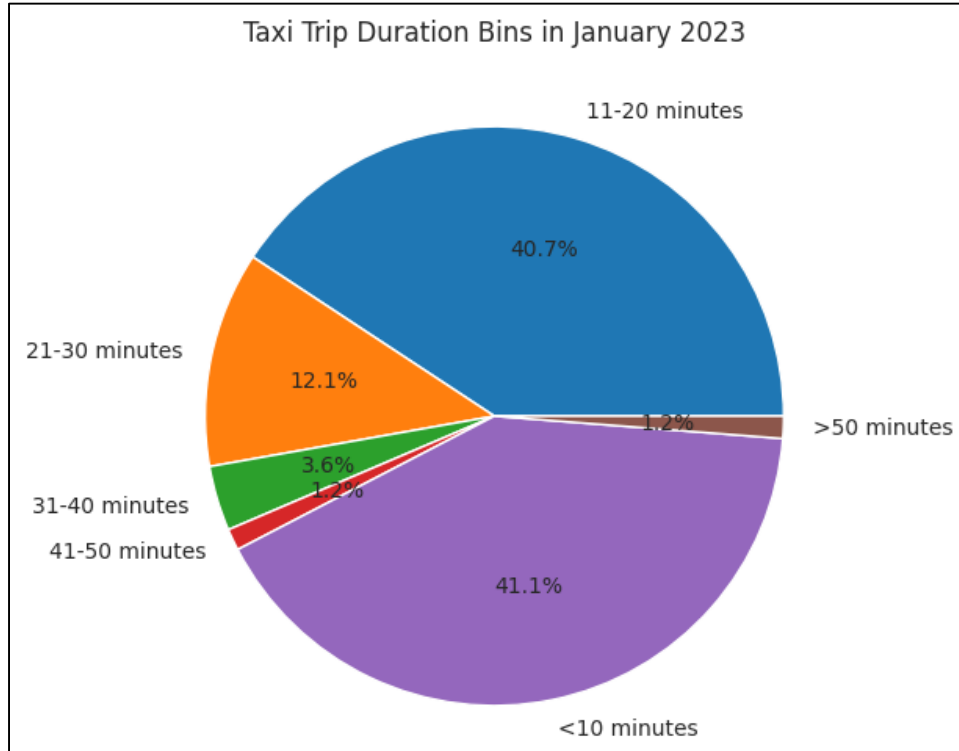


Rate	trip_count
JFK Airport	115
Nassau or Westchester	3
Negotiated Fare	693
Newark Airport	9
Standard Rate	62836



3. Data Analysis Customer Preferences and Behavior

2. How do customers' preferences vary by rate, trip distance and duration?

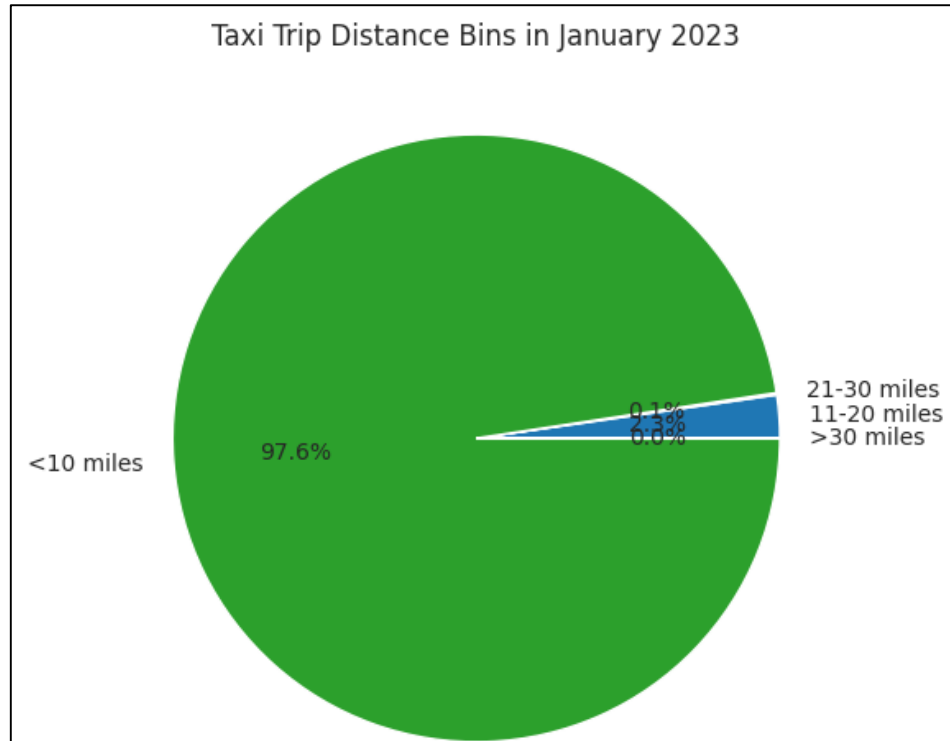


trip_duration_bins	trip_count
11-20 minutes	25932
21-30 minutes	7671
31-40 minutes	2311
41-50 minutes	767
<10 minutes	26188
>50 minutes	787



3. Data Analysis Customer Preferences and Behavior

2. How do customers' preferences vary by rate, trip distance and duration?

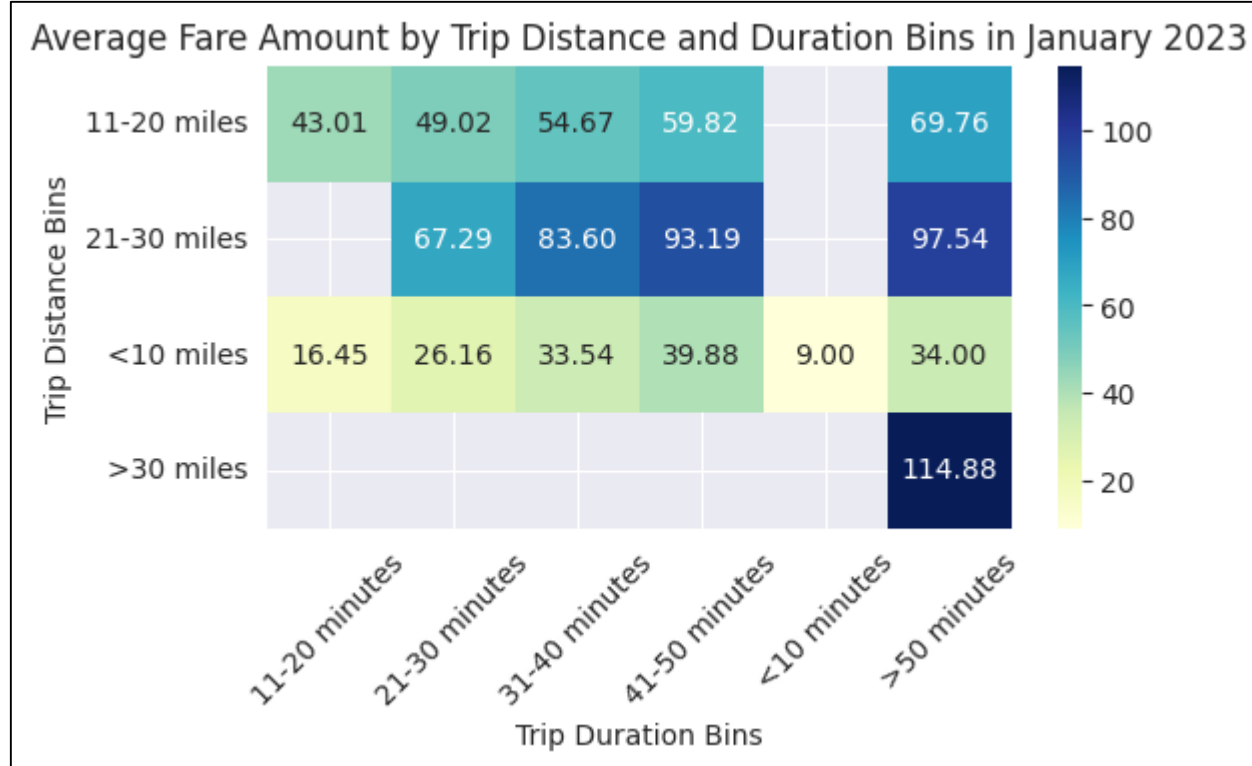


trip_distance_bins	trip_count
11-20 miles	1456
21-30 miles	69
<10 miles	62121
>30 miles	10



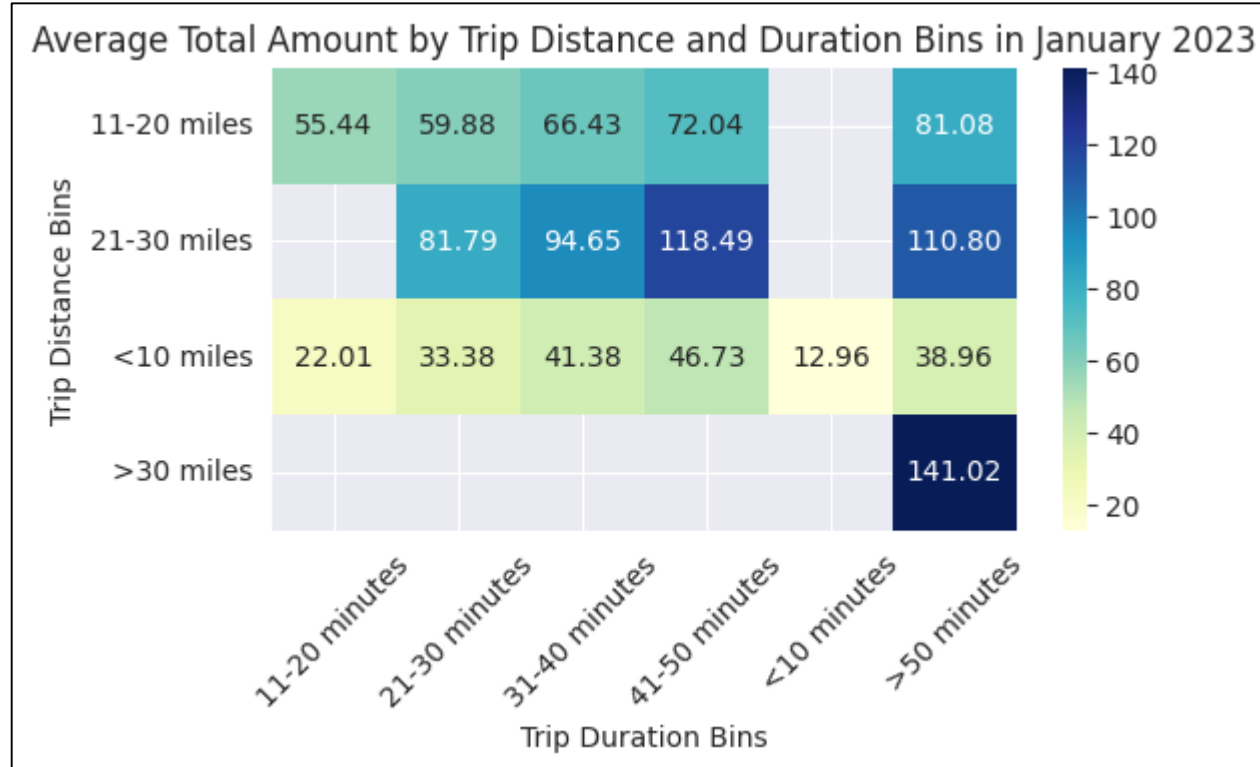
3. Data Analysis Fare Structures for NYC Taxi Services

1. How do fare amounts vary with trip distance and trip duration?



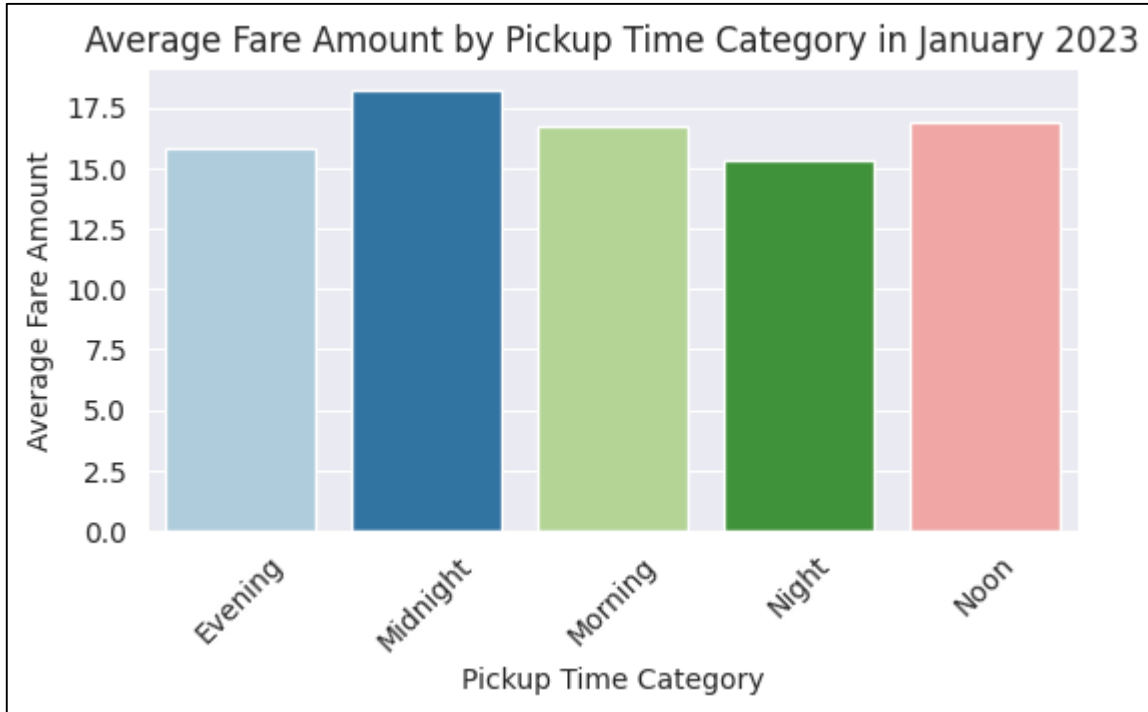
3. Data Analysis Fare Structures for NYC Taxi Services

2. How do total amounts vary with trip distance and trip duration?



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?



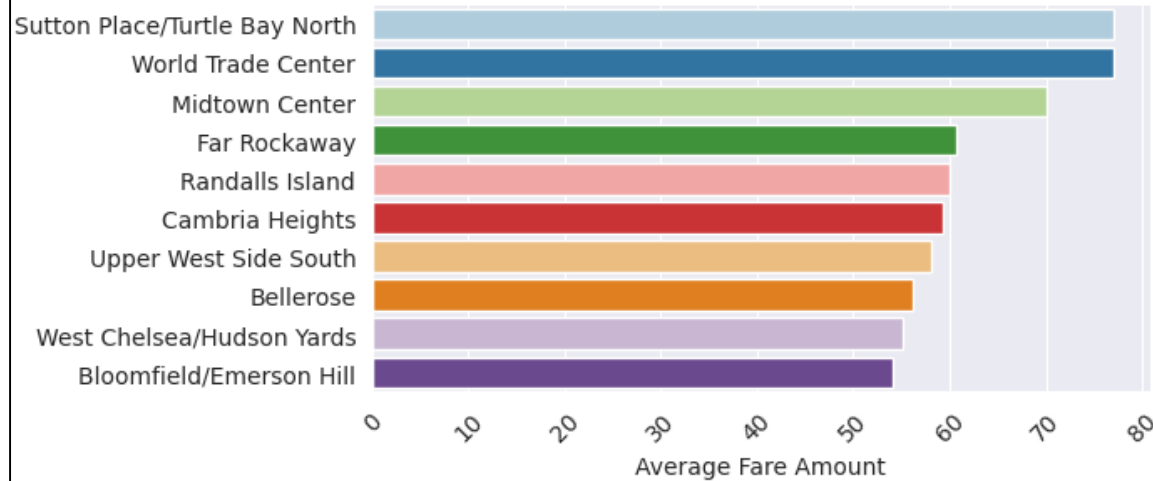
pickup_time_category	fare_amount
Evening	15.799997
Midnight	18.203393
Morning	16.737771
Night	15.279014
Noon	16.889210



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?

Top 10 Most Expensive Pickup Zones by Average Fare Amount in January 2023

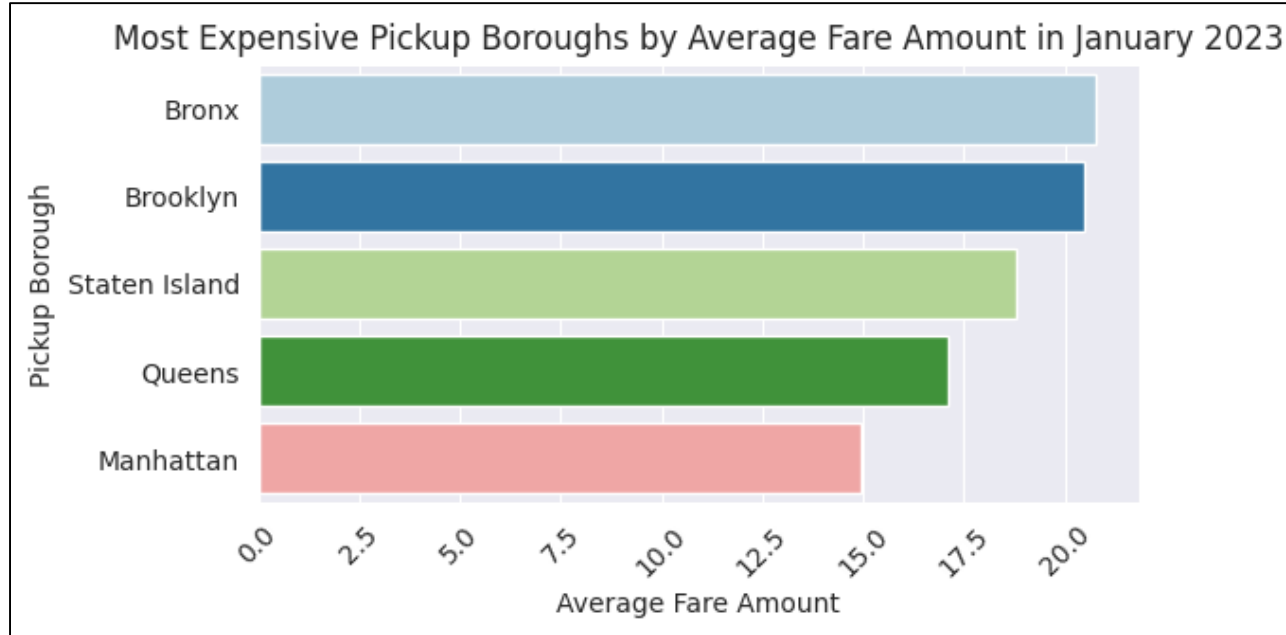


PUzone	fare_amount
Sutton Place/Turtle Bay North	77.000000
World Trade Center	77.000000
Midtown Center	70.000000
Far Rockaway	60.673333
Randalls Island	60.000000
Cambria Heights	59.301000
Upper West Side South	58.065000
Bellerose	56.200000
West Chelsea/Hudson Yards	55.150000
Bloomfield/Emerson Hill	54.100000



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?

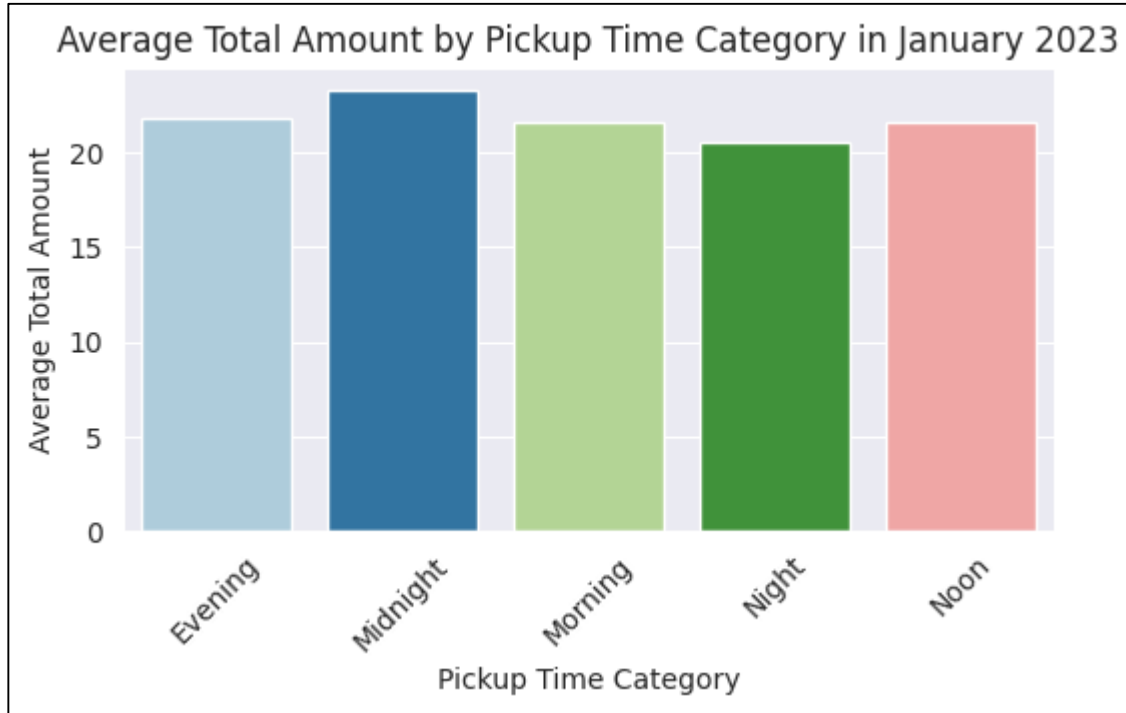


PUborough	fare_amount
Bronx	20.781838
Brooklyn	20.484256
Staten Island	18.775000
Queens	17.120862
Manhattan	14.950049



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?

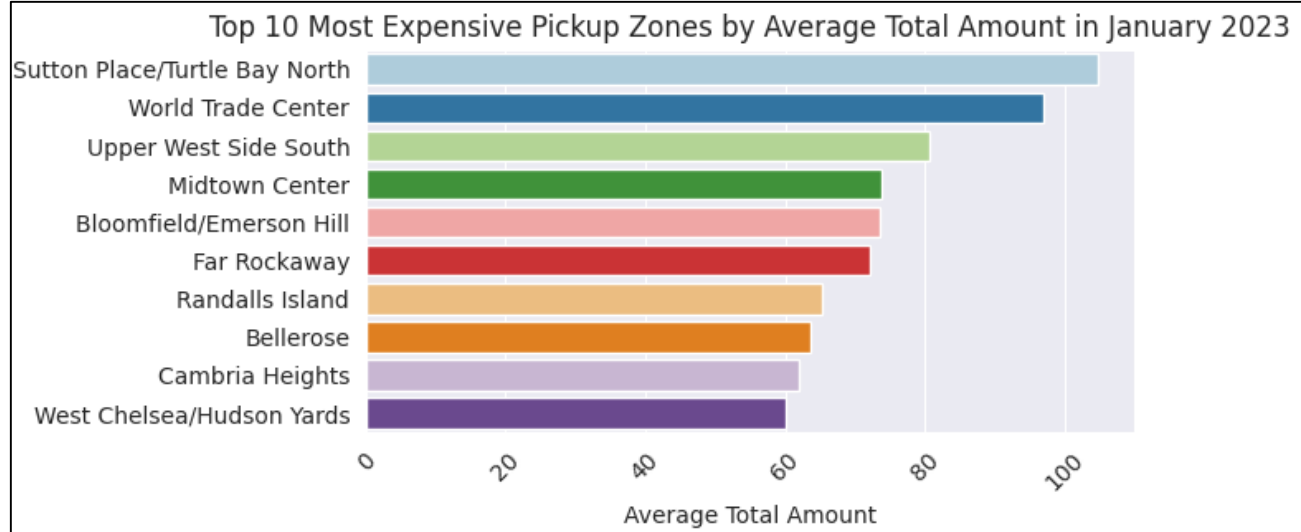


pickup_time_category	total_amount
Evening	21.825025
Midnight	23.308776
Morning	21.573580
Night	20.589067
Noon	21.577776



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?

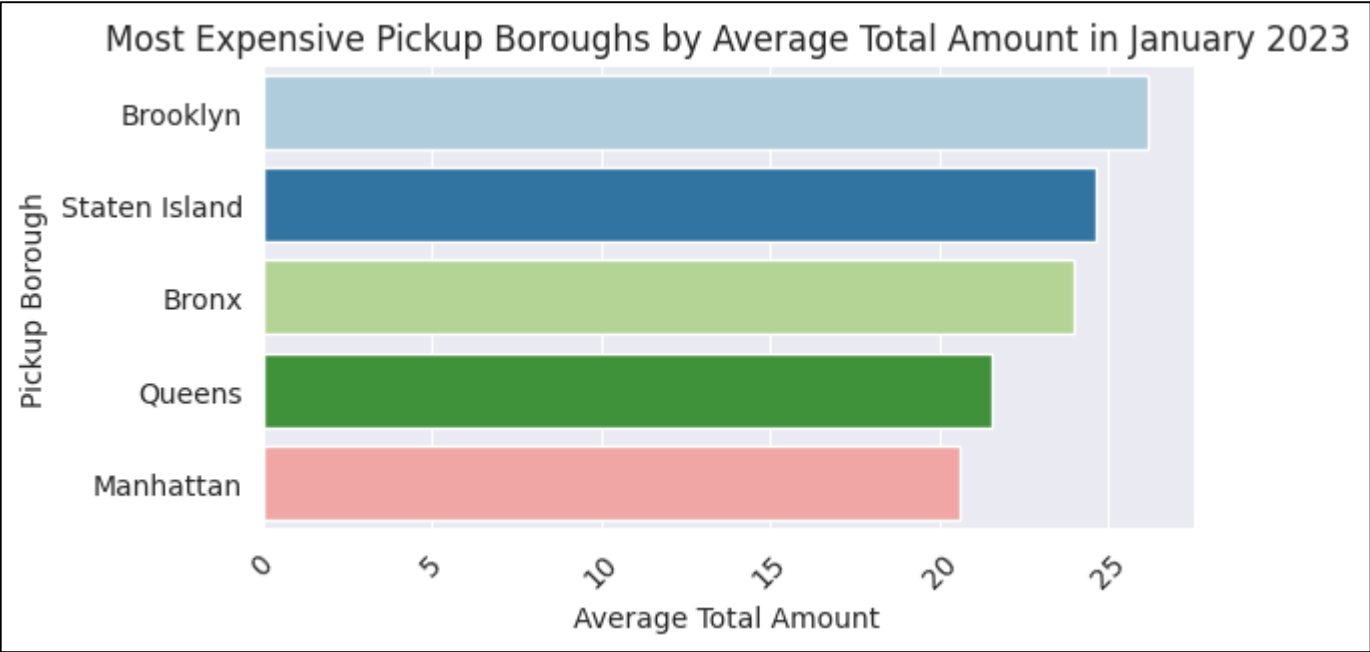


PUzone	total_amount
Sutton Place/Turtle Bay North	104.760000
World Trade Center	96.900000
Upper West Side South	80.800000
Midtown Center	73.750000
Bloomfield/Emerson Hill	73.700000
Far Rockaway	72.273333
Randalls Island	65.300000
Bellerose	63.700000
Cambria Heights	62.001000
West Chelsea/Hudson Yards	60.020000



3. Data Analysis Fare Structures for NYC Taxi Services

3. Are there specific times or locations where fare adjustments could increase revenue?

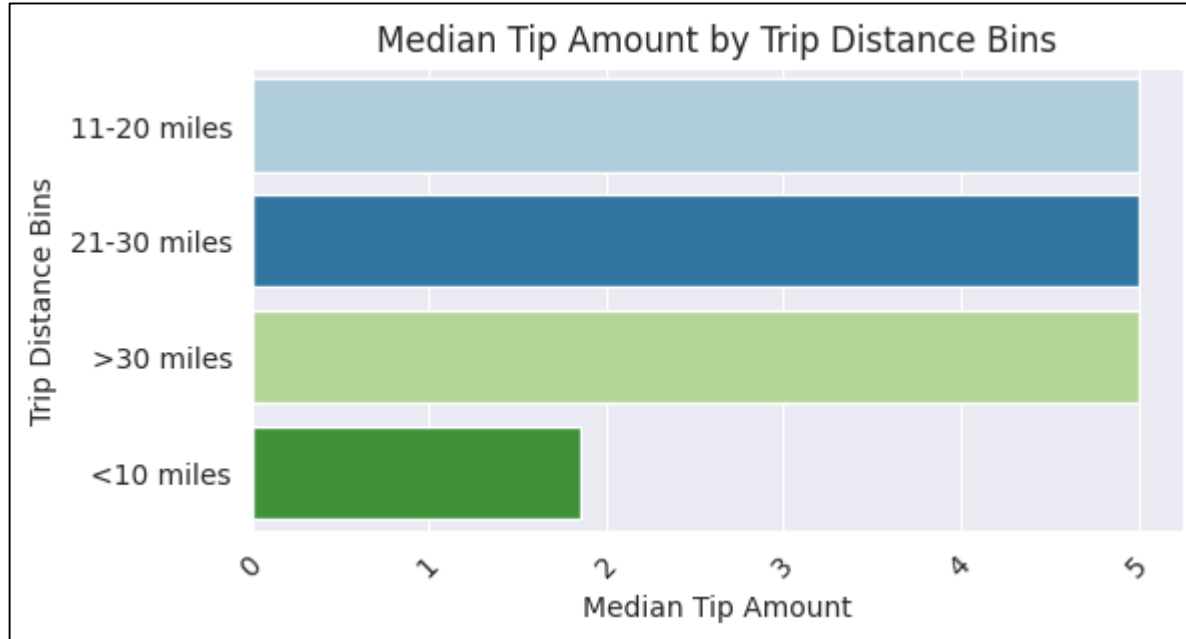


PUBorough	total_amount
Brooklyn	26.176685
Staten Island	24.607500
Bronx	23.948126
Queens	21.525015
Manhattan	20.625278



3. Data Analysis Tips for NYC Taxi Services

1. How do trip distance, trip duration, fare amount, and speed affect tip amounts?



Kruskal-Wallis test results:
Statistic: 369.69190393934093
p-value: 8.118075877605254e-80

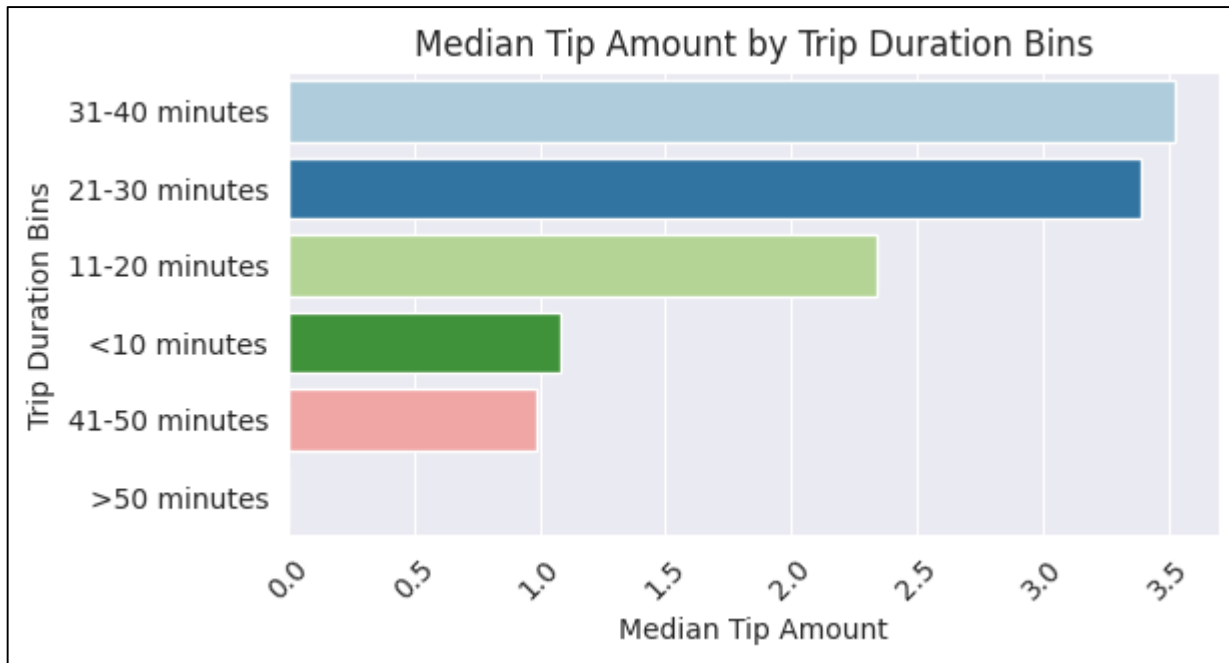
Conclusion:

There is **statistically significant difference** in median tip amount among different trip distance bins.



3. Data Analysis Tips for NYC Taxi Services

1. How do trip distance, trip duration, fare amount, and speed affect tip amounts?



```
Kruskal-Wallis test results:  
Statistic: 4025.3229727652865  
p-value: 0.0
```

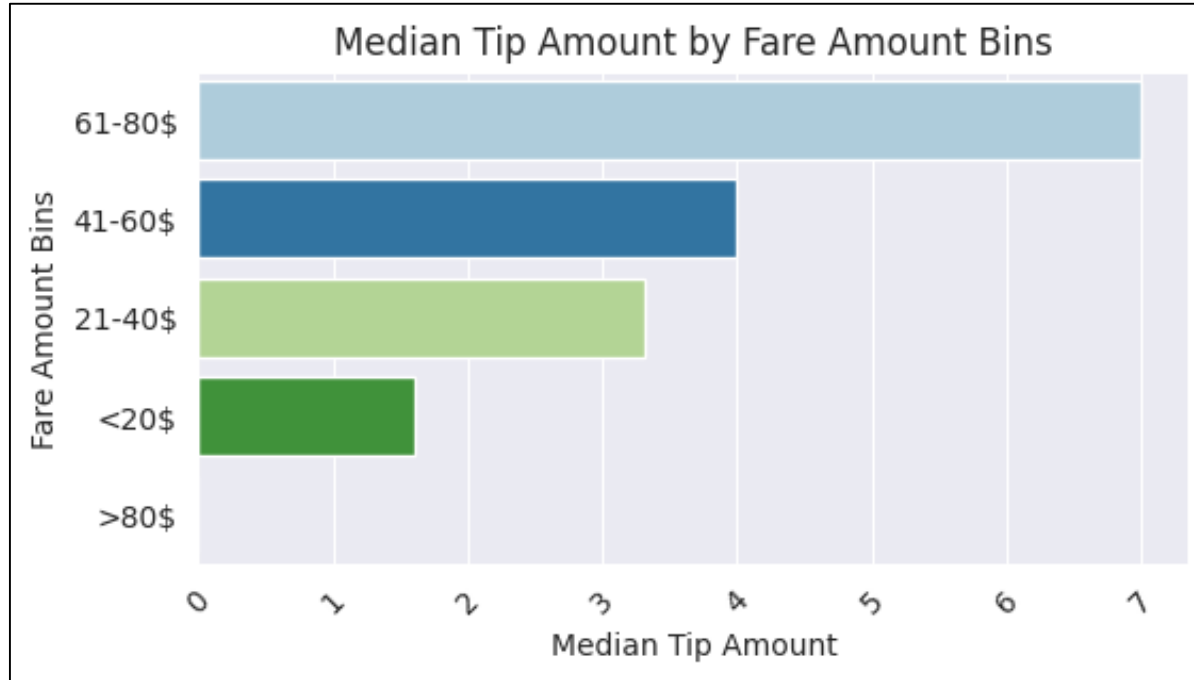
Conclusion:

There is **statistically significant difference** in median tip amount among different trip duration bins.



3. Data Analysis Tips for NYC Taxi Services

1. How do trip distance, trip duration, fare amount, and speed affect tip amounts?



Kruskal-Wallis test results:
Statistic: 2932.3227697687707
p-value: 0.0

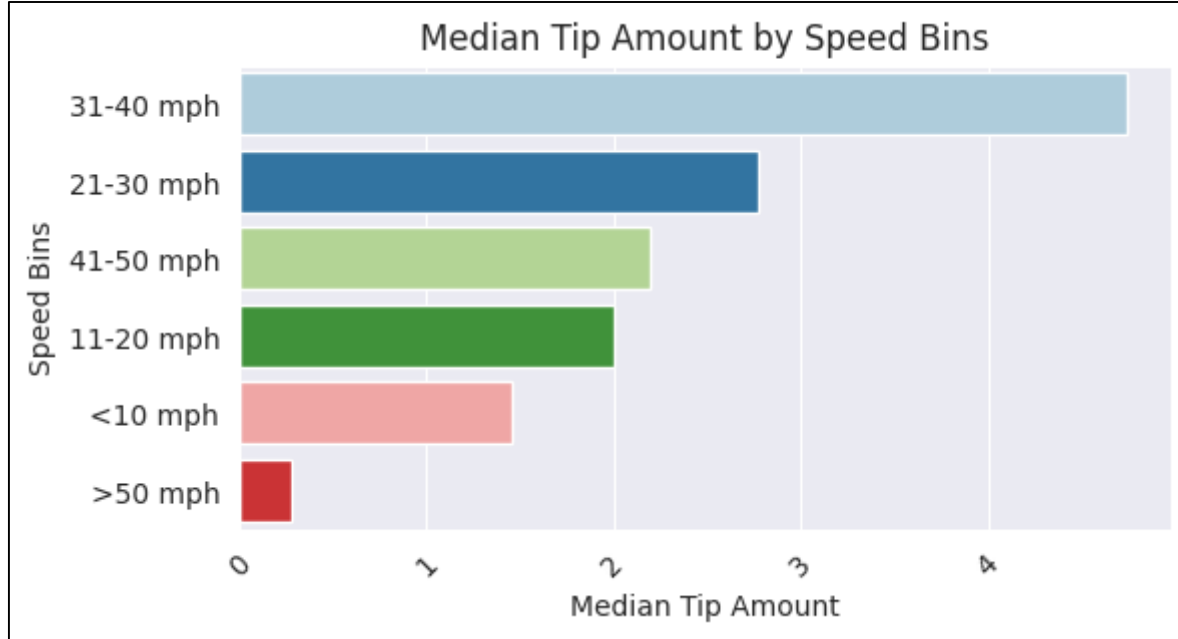
Conclusion:

There is **statistically significant difference** in median tip amount among different fare amount bins.



3. Data Analysis Tips for NYC Taxi Services

1. How do trip distance, trip duration, fare amount, and speed affect tip amounts?



```
Kruskal-Wallis test results:  
Statistic: 1036.6697125020673  
p-value: 6.911427616629211e-222
```

Conclusion:

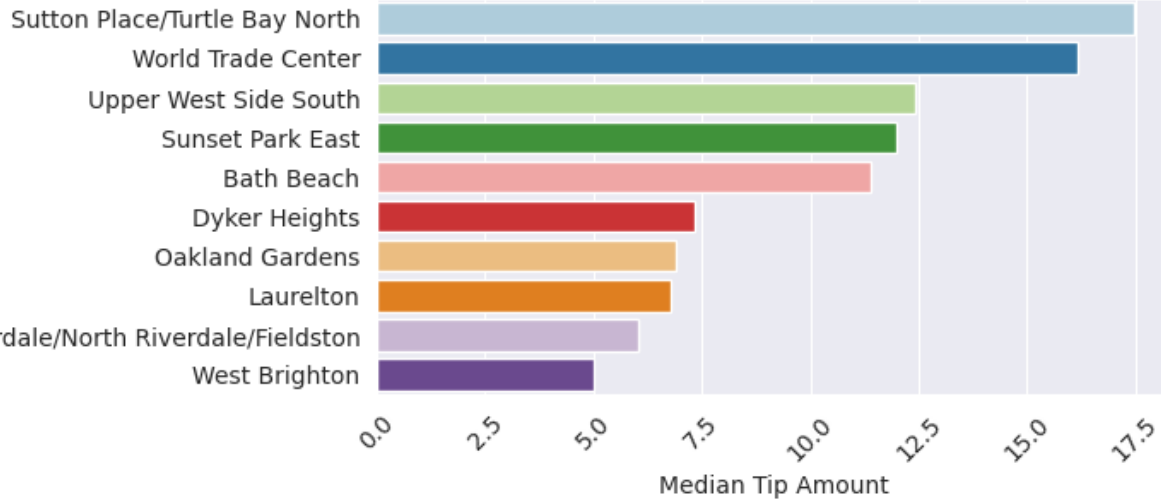
There is **statistically significant difference** in median tip amount among different speed bins.



3. Data Analysis Tips for NYC Taxi Services

2. How do pickup zone affect tip amounts?

Top 10 Pickup Zones with the Highest Median Tip Amount



```
Kruskal-Wallis test results:  
Statistic: 8.406546336206892  
p-value: 0.49375115402430325
```

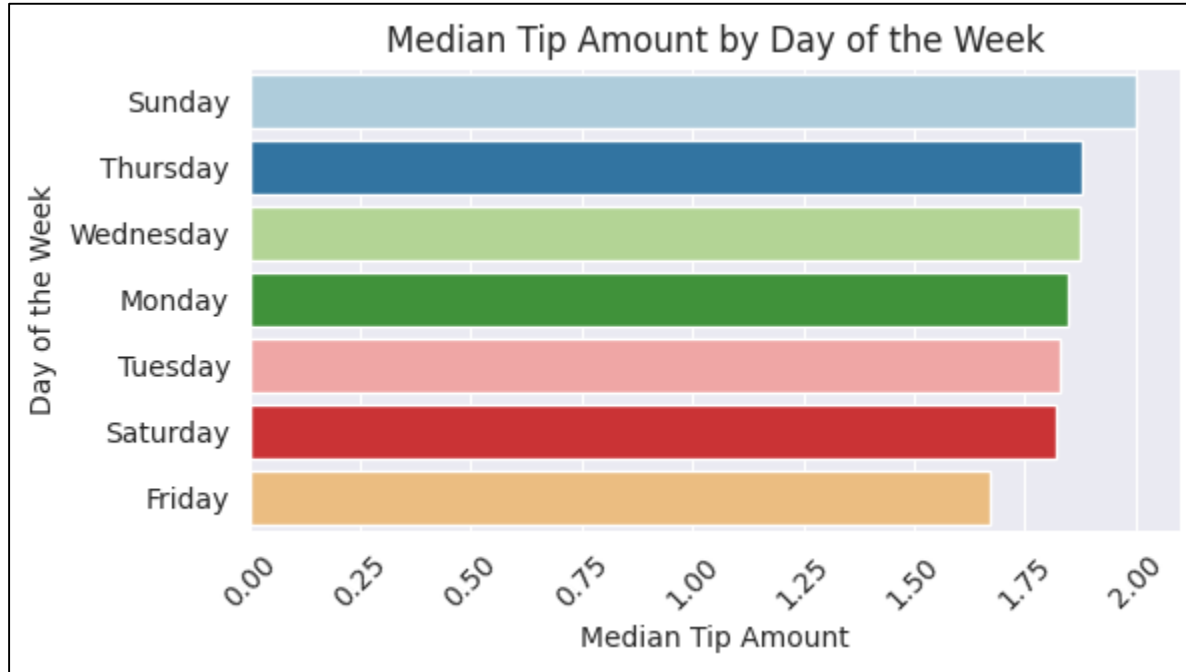
Conclusion:

There is **no statistically significant difference** in median tip amount among different pickup zones.



3. Data Analysis Tips for NYC Taxi Services

3. How do times (hours, days) affect tip amounts?



```
Kruskal-Wallis test results:  
Statistic: 26.468314149110263  
p-value: 0.00018207500753749176
```

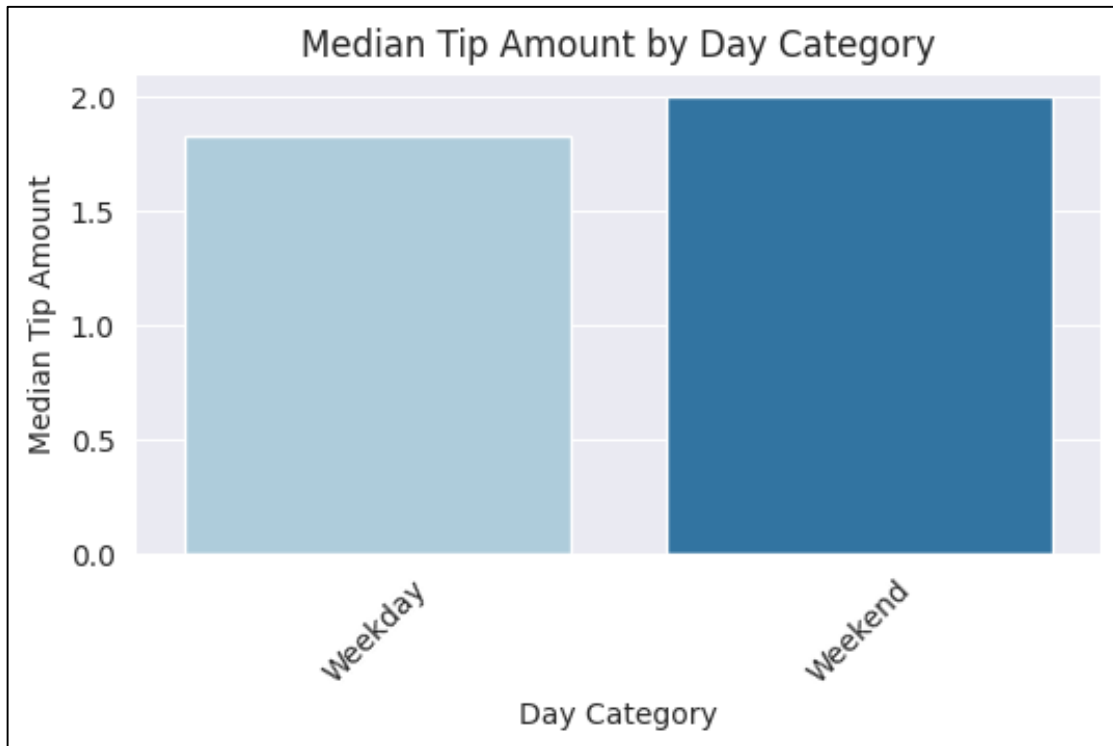
Conclusion:

There is **statistically significant difference** in median tip amount among different day names.



3. Data Analysis Tips for NYC Taxi Services

3. How do times (hours, days) affect tip amounts?



Mann-Whitney U test results:
Statistic: 372166976.5
p-value: 0.01605410381090577

Conclusion:

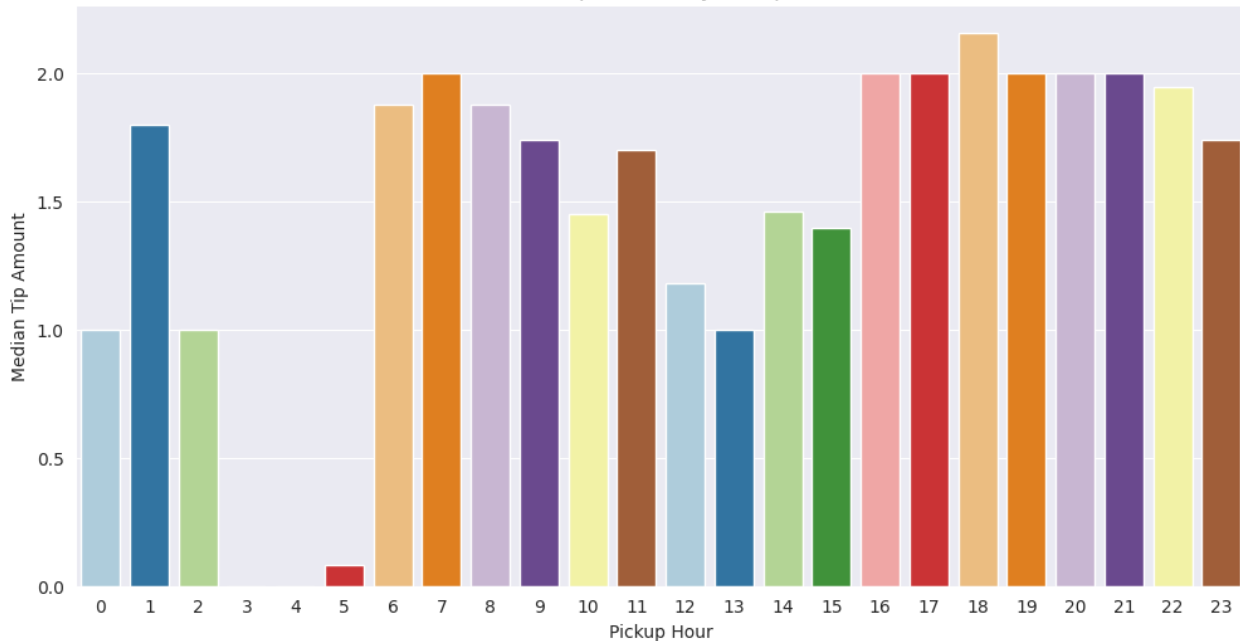
There is **statistically significant difference** in median tip amount among different day categories.



3. Data Analysis Tips for NYC Taxi Services

3. How do times (hours, days) affect tip amounts?

Median Tip Amount by Pickup Hour



```
Kruskal-Wallis test results:  
Statistic: 261.86025850187883  
p-value: 2.123991964587294e-42
```

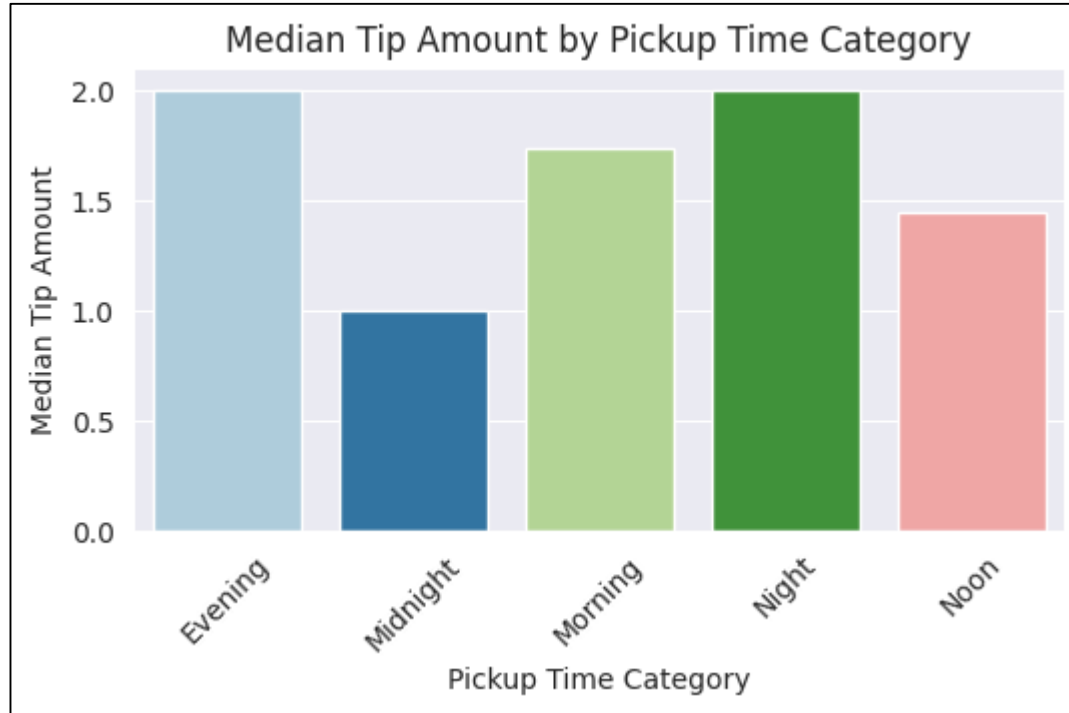
Conclusion:

There is **statistically significant difference** in median tip amount among different pickup hours.



3. Data Analysis Tips for NYC Taxi Services

3. How do times (hours, days) affect tip amounts?



```
Kruskal-Wallis test results:  
Statistic: 118.40309267578844  
p-value: 1.1713933440657628e-24
```

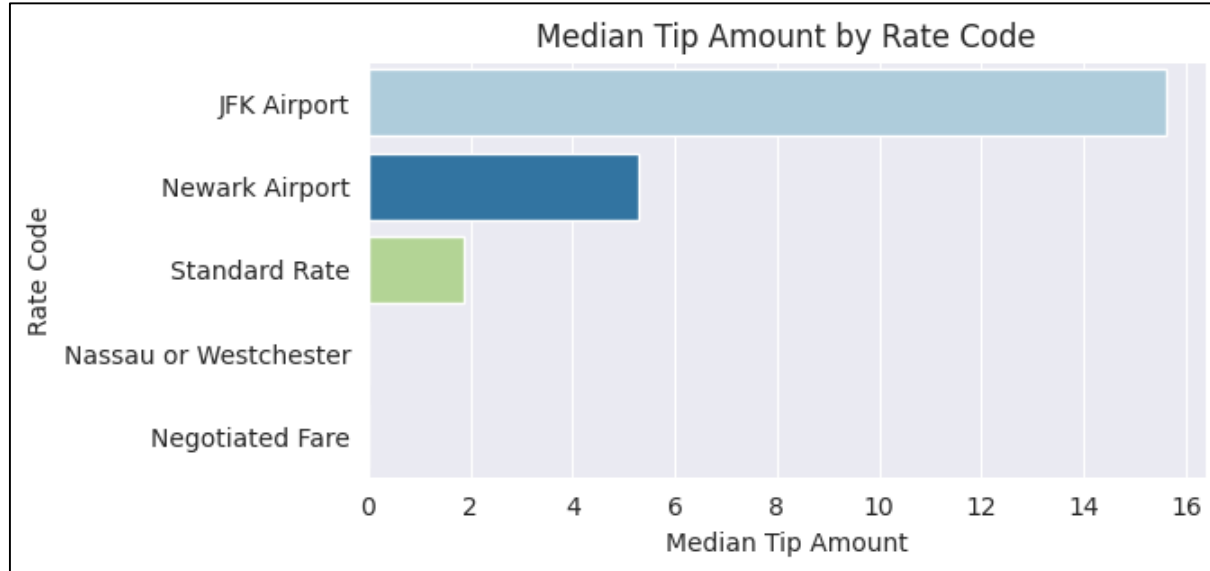
Conclusion:

There is **statistically significant difference** in median tip amount among different pickup time categories.



3. Data Analysis Tips for NYC Taxi Services

4. How do Trip Types and Rate affect tip amounts?



Kruskal-Wallis test results:
Statistic: 135.57216560106994
p-value: 2.5025275544086144e-28

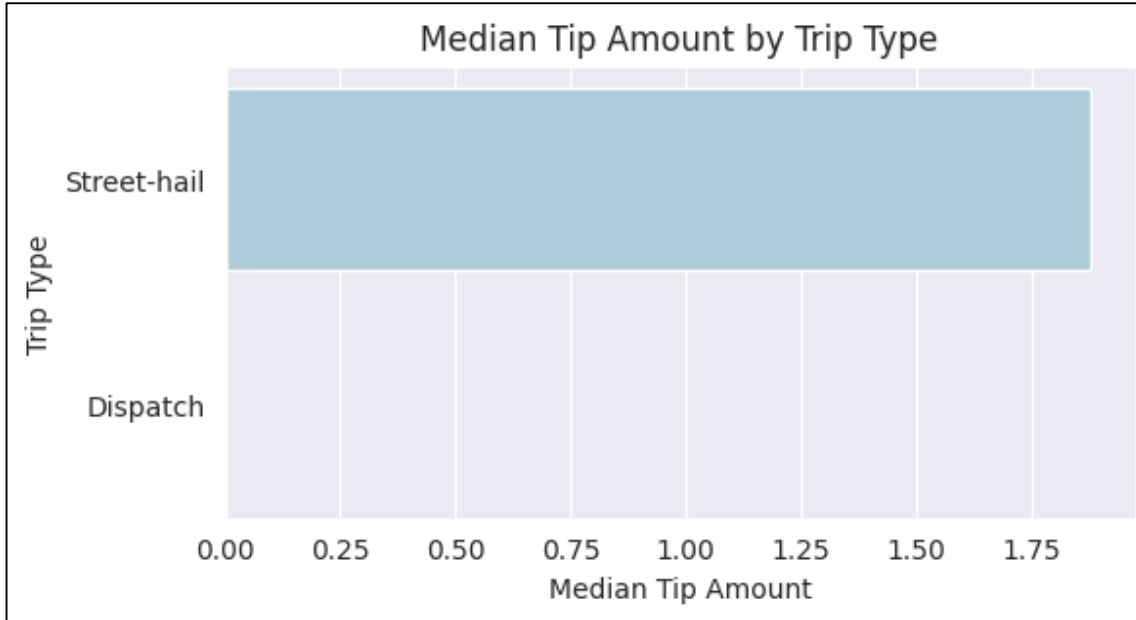
Conclusion:

There is **statistically significant difference** in median tip amount among different rate codes.



3. Data Analysis Tips for NYC Taxi Services

4. How do Trip Types and Rate affect tip amounts?



```
Mann-Whitney U test results:  
Statistic: 20538248.5  
p-value: 0.00018370989942526915
```

Conclusion:

There is **statistically significant difference** in median tip amount among different trip types.



4. Conclusion and Recommendations



01



Conclusion :

4. Conclusion and Recommendations

Conclusion :

Taxi demand in New York City

1. Tanggal puncak permintaan taksi terjadi pada weekdays, terutama Kamis dan Jumat, ditandai dengan adanya mobilitas yang meningkat selama minggu kerja sehingga membutuhkan layanan taksi.
2. Permintaan yang lebih rendah terjadi pada akhir pekan (minggu) dan hari libur nasional (hari tahun baru dan hari lahir Martin Luther King Jr.), hal ini ditandai dengan aktivitas perjalanan yang berkurang karena masyarakat Kota New York cenderung berdiam diri di rumah atau menggunakan kendaraan pribadi untuk berlibur.
3. Jam penjemputan menunjukkan dua puncak harian:
 - pagi hari selama perjalanan ke tempat kerja atau sekolah
 - sore hari selama perjalanan pulang.
4. Zona penjemputan yang paling banyak terpusat di Borough Manhattan.



4. Conclusion and Recommendations

Conclusion :

Customer Preferences and Behaviour

1. Customer Taxi di Kota New York sebagian besar menggunakan kartu kredit (lebih dari 80%) untuk pembayaran taksi, menunjukkan preferensi untuk kenyamanan dan keamanan.
2. Sebagian besar perjalanan (99,9%) adalah street-hail, hal ini karena metode pemanggilan taksi ini cukup populer dan mudah diakses.
3. Sebagian besar perjalanan adalah jarak pendek, dengan sebagian besar hanya berlangsung kurang dari 20 menit dan menempuh jarak kurang dari 10 mil. Hal ini menunjukkan bahwa taksi terutama digunakan untuk transportasi lokal di dalam Kota New York.
4. Pelanggan lebih memilih standar rate untuk perjalanan taksi, mencerminkan sebagian besar perjalanan terjadi di dalam Kota New York.



4. Conclusion and Recommendations

Conclusion :

Fare Structures for NYC Taxi Services

1. Jumlah tarif dan total jumlahnya umumnya meningkat dengan jarak dan durasi perjalanan yang meningkat.
2. Ada waktu-waktu tertentu (misalnya tengah malam) dan lokasi-lokasi tertentu (misalnya Bronx, Brooklyn) di mana penyesuaian tarif berpotensi meningkatkan pendapatan.
3. Menganalisis kategori waktu pengambilan, zona pengambilan, dan borough pengambilan dapat memberikan wawasan untuk menyesuaikan tarif demi peningkatan pendapatan.
4. Penting untuk mempertimbangkan baik jumlah tarif maupun total jumlahnya ketika mengevaluasi struktur tarif.



4. Conclusion and Recommendations

Conclusion :

Tips for NYC Taxi Services

1. Jarak perjalanan, durasi perjalanan, jumlah tarif, dan kecepatan secara signifikan mempengaruhi jumlah tip.
2. Tidak ada zona penjemputan khusus di mana tips lebih tinggi.
3. Tips lebih tinggi di akhir pekan dan selama jam tertentu dalam sehari.
4. Kode tarif dan jenis perjalanan juga memengaruhi jumlah tip.



Recommendations :



02

4. Conclusion and Recommendations

Recommendations :

Penjadwalan Kendaraan Taksi:

1. Tingkatkan jumlah taksi yang tersedia selama jam-jam sibuk pada pagi hari (05.00-11.00) dan sore hari (15.00-20.00) serta di lokasi-lokasi permintaan tinggi seperti di Borough Manhattan untuk mengurangi waktu tunggu pelanggan.
2. Kurangi jumlah taksi pada hari minggu dan libur nasional lain. Selain itu, kurangi juga jumlah armada taksi pada jam-jam permintaan rendah seperti midnight (00.00-05.00).
3. Gunakan data historis dan analisis prediktif untuk mengoptimalkan penjadwalan armada dan memastikan ketersediaan taksi yang cukup.



4. Conclusion and Recommendations

Recommendations :

Penargetan Wilayah:

1. Identifikasi zona penjemputan yang paling menguntungkan seperti pada Borough Manhattan dan fokuskan upaya pemasaran dan penjangkauan di wilayah tersebut.
2. Analisis pola permintaan dan sesuaikan strategi penargetan wilayah secara berkala untuk memaksimalkan pendapatan.
3. Pertimbangkan untuk menawarkan diskon untuk perjalanan dari wilayah yang kurang populer seperti pada Borough Staten Island dan Bronx.
4. Menawarkan insentif kepada pengemudi untuk beroperasi di area-area dengan permintaan tinggi.



4. Conclusion and Recommendations

Recommendations :

Strategi Berdasarkan Harga:

1. Memberi tarif yang lebih tinggi selama waktu-waktu sibuk seperti pada hari Kamis-Jumat pagi hari (05.00-11.00) dan sore hari (15.00-20.00) atau di lokasi-lokasi tertentu seperti di Borough Manhattan.
2. Menawarkan tarif yang lebih rendah selama waktu-waktu permintaan rendah seperti pada hari minggu dan libur nasional dan weekday pada dini hari (00.00-05.00) serta di lokasi-lokasi yang kurang populer seperti pada Borough Staten Island dan Bronx..
3. Menawarkan paket tarif untuk pelanggan setia.
4. Perkenalkan program loyalitas atau sistem poin untuk mendorong pelanggan menggunakan layanan taksi secara berulang.



4. Conclusion and Recommendations

Recommendations :

Opsi Pembayaran:

1. Terus dukung penggunaan kartu kredit sebagai metode pembayaran utama.
2. Pertimbangkan untuk memperkenalkan opsi pembayaran digital lainnya seperti dompet elektronik dan pembayaran tanpa kontak untuk meningkatkan kenyamanan dan keamanan pelanggan.
3. Pastikan bahwa sistem pembayaran aman dan mudah digunakan.
4. Pertimbangkan untuk menawarkan diskon atau insentif untuk pelanggan yang menggunakan opsi pembayaran tertentu.



4. Conclusion and Recommendations

Recommendations :

Trip Perjalanan:

1. Terus pertahankan metode pemanggilan taksi street-hail karena masih populer dan mudah diakses.
2. Pastikan bahwa sistem pemesanan mudah digunakan dan dapat diakses oleh pelanggan.
3. Pertimbangkan untuk menawarkan diskon atau insentif untuk pelanggan yang menggunakan opsi pemesanan online (Dispatch)



4. Conclusion and Recommendations

Recommendations :

Peningkatan Kualitas Pelayanan :

1. Pantau dan ukur waktu tunggu rata-rata dan waktu perjalanan untuk mengidentifikasi area yang perlu ditingkatkan.
2. Terapkan teknologi untuk mengoptimalkan rute dan mengurangi waktu perjalanan.
3. Berikan pelatihan kepada pengemudi tentang cara mengemudi yang aman dan efisien.
4. Pastikan bahwa kendaraan taksi dalam kondisi baik dan terawat dengan baik.



4. Conclusion and Recommendations

Recommendations :

Promosi dan Penawaran:

1. Tawarkan diskon atau promosi untuk perjalanan jarak jauh (seperti pada perjalan > 20 miles).
2. Buat paket tarif yang menarik untuk perjalanan pulang-pergi atau perjalanan beberapa hari.
3. Bekerjasama dengan bisnis lokal atau organisasi masyarakat untuk menawarkan diskon atau promosi khusus bagi pelanggan mereka.



Terima Kasih

