

## IDENTIFIKASI PENGELOMPOKAN TITIK PENJEMPUTAN DAN TITIK PENGANTARAN PERJALANAN TAKSI MENGGUNAKAN ALGORITMA DBSCAN

Anisa Salsabila<sup>1\*</sup>, Lizda Iswari<sup>2</sup>

<sup>1,2</sup> Universitas Islam Indonesia, Indonesia

E-mail: [anisa.salsabila@students.uii.ac.id](mailto:anisa.salsabila@students.uii.ac.id)

**Abstract:** The rapid growth of urban areas in recent years has resulted in a strong interconnection and mutual influence between the daily travel routines of city residents and the use of public transportation. The satisfaction of passenger services is closely tied to the efforts of taxi fleets, starting from the pickup process to the delivery to the destination. The waiting time for taxis becomes a crucial factor for passengers in choosing the right pickup point in urban environments. Accurately identifying the history of taxi passenger demand can assist taxi fleet managers in allocating resources, especially in urban areas. In this study, the DBSCAN algorithm is employed to identify cluster patterns emerging from taxi pickup and drop-off points based on taxi travel data. The data used is sourced from Kaggle, with a specific focus on taxi journeys in Brooklyn. At passenger pickup points, potential areas during weekdays are observed in the Boerum Place and Adams Street areas, while on holidays, they are found on Flatbush Avenue. On the other hand, at passenger drop-off points during weekdays, potential areas are identified on Myrtle Avenue, Flatbush Avenue, and Lafayette Avenue, located in Downtown Brooklyn and Fort Greene. On holidays, drop-off points tend to occur in Williamsburg.

**Keywords:** Taxi Trips, Pick-up Point, Drop-off Point, DBSCAN, Potential Areas

**Abstrak:** Pesatnya pertumbuhan wilayah perkotaan dalam beberapa tahun terakhir mengakibatkan keterkaitan dan saling mempengaruhi antara rutinitas perjalanan harian penduduk kota dan penggunaan transportasi umum. Kepuasan pelayanan kepada penumpang merupakan hal yang sangat terkait dengan usaha armada taksi, dimulai dari proses penjemputan hingga pengantaran ke tujuan. Waktu menunggu taksi menjadi faktor penting bagi penumpang dalam memilih titik penjemputan yang tepat di lingkungan perkotaan. Mengidentifikasi sejarah permintaan penumpang taksi secara akurat dapat membantu pengelola armada taksi dalam mengalokasikan sumber daya terutama di wilayah perkotaan. Dalam penelitian ini, algoritma DBSCAN digunakan untuk mengidentifikasi pola cluster yang muncul dari titik penjemputan dan titik pengantaran penumpang berdasarkan data perjalanan taksi. Data yang digunakan berasal dari Kaggle, dan fokus penelitian ini adalah perjalanan taksi di Kota Brooklyn. Pada titik penjemputan penumpang, area potensial pada hari kerja terjadi di kawasan Boerum Place serta Adams Street dan pada hari libur terdapat di Flatbush Avenue. Sedangkan pada titik pengantaran penumpang, area potensial pada hari kerja terjadi di Myrtle Avenue, Flatbush Avenue, dan Lafayette Avenue yang terletak di Downtown Brooklyn dan Fort Greene. Pada hari libur, titik pengantaran cenderung terjadi di Williamsburg.

**Kata Kunci :** Perjalanan Taksi, Titik Penjemputan, Titik Pengantaran, DBSCAN, Area Potensial

## PENDAHULUAN

Setengah dari 7.6 miliar penduduk dunia tinggal pada daerah perkotaan dan setiap tahunnya jumlah penduduk akan bertambah sekitar 83 juta jiwa yang setara dengan empat kali lipat jumlah penduduk Kota *New York* (*United Nations, Department of Economic and Social Affairs, Population Division, 2017*). Dengan pesatnya angka perkembangan daerah perkotaan tersebut, mengakibatkan hubungan antara rutinitas perjalanan sehari-hari penduduk kota dan transportasi umum saling berkaitan erat (Jian, Li, & Yu, 2021). Armada taksi merupakan transportasi umum yang sering digunakan oleh penduduk kota. Dengan meningkatnya persaingan global dan globalisasi, mobilitas perkotaan menggunakan armada taksi berhubungan dengan sangat relevan (Vizuete-Luciano, Guill'en-Pujadas, Alaminos, & Merigo-Lindahl, 2023).

Usaha armada taksi berkaitan erat dengan kepuasan pelayanan yang dimulai dari penjemputan penumpang taksi hingga pengantaran ke tempat tujuan penumpang. Kelebihan utama menggunakan taksi adalah layanan yang dapat membawa penumpang dari titik penjemputan hingga titik akhir pengantaran. Hal tersebut jarang ditemukan pada transportasi umum lainnya yang cenderung beroperasi antar stasiun atau halte dengan rute yang terbatas (Ulak, Yazici, & Aljarrah, 2020). Taksi menjadi salah satu alat transportasi yang esensial untuk mengakomodasi perjalanan penduduk perkotaan (Jian, Li, & Yu, 2021).

Pengoperasian taksi merupakan proses jangka panjang dimana banyak penumpang yang dijemput dan diturunkan secara terus menerus dengan jangka waktu tertentu. Satu perjalanan memungkinkan sopir taksi mengarah ke suatu tempat di mana sulit untuk menemukan penumpang selanjutnya. Hal tersebut membuat taksi berjalan secara kosong tanpa penumpang (Gao, Jiang, & Xu, 2018). Meskipun beberapa sopir taksi memperoleh penghasilan yang cukup, namun banyak sopir taksi yang gagal mendapatkan penghasilan karena tidak mendapatkan penumpang yang cukup. Sedikit sopir taksi yang berkeliling kota selama berjam-jam untuk mencari penumpang. Sementara, banyak sopir taksi yang memilih untuk menunggu di area yang sibuk seperti di luar mal besar atau bandara. Bahkan ketika mendapatkan penumpang, sopir taksi harus menunggu lama ataupun berkeliling mencari penumpang. Hal tersebut menyebabkan kekeliruan pengelolaan taksi karena penempatan taksi yang kurang efektif (Malhan, 2017). Selain itu, mayoritas penumpang yang ingin menggunakan layanan taksi harus menunggu lama atau mencari

taksi terdekat. Waktu tunggu tersebut menjadi hal yang krusial bagi penumpang taksi dalam memilih lokasi penjemputan yang ramai di lingkungan perkotaan (You, et al., 2021).

Efisiensi layanan taksi di kota besar tidak hanya mempengaruhi kenyamanan perjalanan masyarakat, tetapi juga lalu lintas perkotaan dan keuntungan bagi pengelola taksi (Gao, Jiang, & Xu, 2018). Oleh sebab itu, dengan mengidentifikasi sejarah lokasi permintaan taksi secara akurat dan memberikan rekomendasi lokasi bagi sopir taksi dapat membantu pengelola taksi dapat mengalokasikan sumber daya taksi di daerah perkotaan yang potensial. Sehingga, sopir taksi pun dapat menemukan penumpang secara cepat dan mengurangi waktu tunggu penumpang untuk menggunakan layanan taksi melalui hasil identifikasi lokasi potensial (Zhang, et al., 2020).

Dalam mengidentifikasi lokasi permintaan taksi, data geospasial menjadi data yang penting dan dapat memberikan informasi untuk mengeksplorasi lingkungan spasial perkotaan. Salah satu metode pionir yang dapat digunakan untuk mengidentifikasi pola tersebut adalah *Density-Based Spatial Clustering of Application with Noise* (DBSCAN). Algoritma DBSCAN sebagai pembelajaran tanpa pengawasan dapat mengidentifikasi kelompok yang berbeda pada data. Algoritma tersebut akan membentuk area yang memiliki tingkat kepadatan tinggi ke dalam sebuah *cluster* dan menemukan *cluster* dalam bentuk sembarang melalui data spasial yang di dalamnya terdapat titik *noise* yang tidak termasuk dalam pengelompokan *cluster* (Sugi Almantara, Sri Ariyani, & Alit Swamardika, 2020).

Dengan memanfaatkan algoritma DBSCAN terhadap data perjalanan taksi, penelitian ini bertujuan untuk mengidentifikasi *cluster* yang terbentuk dari titik penjemputan dan titik pengantaran penumpang. Penelitian ini menggunakan *dataset* yang didapatkan melalui Platform Kaggle yang berjudul “*New York City Taxi Trip Duration*” dengan catatan waktu perjalanan pada bulan April dan bulan Mei tahun 2016. Penelitian akan berfokus pada perjalanan taksi di Kota Brooklyn.

## **METODE**

### **1. Data Preprocessing**

Penelitian ini menggunakan data perjalanan taksi yang diperoleh pada platform Kaggle. *Data cleaning* merupakan langkah untuk menangani atau membersihkan elemen yang tidak standar ataupun tidak relevan dari sekumpulan data. Elemen

tersebut diantaranya *missing value*, *outlier*, *duplicate*, dan sejenisnya (Chicco, Oneto, & Tavazzi, 2022). Proses ini juga akan menganalisis karakteristik dari masing-masing atribut dan hubungan antara satu atribut dengan atribut lainnya. Data yang digunakan adalah catatan perjalanan taksi pada bulan April dan bulan Mei yang berjumlah 135.000 baris.

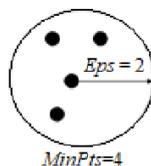
**Tabel 1.** Atribut Data Perjalanan Taksi

No	Atribut	Keterangan	Contoh Atribut
1	<i>id</i>	Kode unik untuk setiap perjalanan	id3504673
2	<i>vendor_id</i>	Kode untuk penyedia taksi terkait dengan catatan perjalanan	1 atau 2
3	<i>pickup_datetime</i>	Tanggal dan waktu saat penjemputan penumpang	2016-04-06 19:32:31
4	<i>dropoff_datetime</i>	Tanggal dan waktu saat tiba di pengantaran penumpang	2016-04-06 19:39:40
5	<i>passenger_count</i>	Jumlah penumpang di dalam kendaraan taksi	4
6	<i>pickup_longitude</i>	Garis bujur saat penjemputan penumpang	-74.010040
7	<i>pickup_latitude</i>	Garis lintang saat penjemputan penumpang	40.719971
8	<i>dropoff_longitude</i>	Garis bujur saat tiba di pengantaran penumpang	-74.012268
9	<i>dropoff_latitude</i>	Garis lintang saat tiba di pengantaran penumpang	40.706718
10	<i>store_and_fwd_flag</i>	Tanda ini menunjukkan apakah catatan perjalanan disimpan di memori taksi sebelum dikirim ke vendor sebab taksi tidak memiliki koneksi ke server	Y atau N
11	<i>trip_duration</i>	Durasi perjalanan dalam hitungan detik	429

## 2. Clustering

Titik penjemputan dan pengantaran penumpang taksi akan dikelompokkan dengan menggunakan metode *clustering*. Fitur yang dibutuhkan untuk melakukan *clustering* dengan *Density-Based Spatial Clustering of Application with Noise* (DBSCAN) adalah *pickup\_longitude*, *pickup\_latitude*, *dropoff\_longitude*, dan *dropoff\_latitude*. Algoritma ini akan memisahkan kelompok untuk posisi penjemputan dan kelompok untuk posisi penurunan penumpang taksi. Pada *cluster* titik penjemputan, atribut yang digunakan adalah *pickup\_longitude* dan *pickup\_latitude*. Dengan memasukkan parameter DBSCAN, yaitu *Eps* dan *MinPts*, algoritma akan

mengidentifikasi *core point* dan menghitung titik-titik yang menjadi *border point* dengan jarak yang ditetapkan pada *Eps*. Titik-titik yang memenuhi batasan yang sudah diatur oleh *MinPts* akan dimasukkan dalam sebuah *cluster*. Namun, jika tidak memenuhi batasan tersebut, maka titik-titik tersebut akan dianggap sebagai *noise*. (C. & H., 2022)



**Gambar 1.** Pembentukan *Cluster* Berdasarkan Parameter  
(Hossain et al., 2021)

Setelah algoritma DBSCAN membentuk *cluster*, salah satu metode yang digunakan untuk menilai kualitas *cluster* adalah dengan memeriksa nilai *silhouette coefficient*. Nilai *silhouette coefficient* akan mendekati 1 jika objek dalam *cluster* terbentuk dengan rapat dan terpisah secara signifikan dari *cluster* tetangga lainnya. Sebaliknya, jika *cluster* terbentuk dengan jarang dan berdekatan dengan *cluster* tetangga, nilai *silhouette coefficient* akan mendekati negatif 1 (Suyanto, 2017).

Penelitian yang dilakukan oleh (Sari & Primajaya, 2019) menerapkan algoritma DBSCAN pada data yang diperoleh dari Dinas Pertanian Kabupaten Karawang. Data tersebut mencakup informasi kecamatan, luas tanam, luas panen, dan organisme pengganggu tanaman. Hasilnya menunjukkan adanya dua *cluster* wilayah kecamatan dengan perbedaan karakteristik, seperti tingkat curah hujan, jumlah serangan hama, jenis hama, dan luas lahan yang memengaruhi produksi pertanian. Pengaturan parameter yang optimal adalah *Eps* 325 dan *MinPts* 2 dengan evaluasi performa menggunakan *average silhouette width* yang menghasilkan skor 0,74.

Metode serupa juga diterapkan pada data yang diperoleh dari situs web resmi *World Health Organization* (WHO). Penelitian ini menggunakan algoritma DBSCAN untuk mengelompokkan negara-negara dengan pola kasus Covid-19 serupa. Hasil pengelompokan dijadikan rekomendasi untuk penanganan kasus Covid-19 dengan memperhatikan negara-negara dalam satu *cluster*. Algoritma DBSCAN diuji dengan variasi parameter *Eps* mulai 0,1 hingga 0,2 dan *MinPts* yaitu 3 dan 4. Pengujian tersebut menghasilkan parameter terbaik dengan *Eps* 0,2 dan *MinPts* 3 yang

menghasilkan 3 *cluster* dengan nilai *silhouette index* sebesar 0,3624 (Nurhaliza & Mustakim, 2021).

Pada penelitian terdahulu, terdapat hal yang serupa setelah menerapkan algoritma DBSCAN terhadap masing-masing dataset dengan permasalahan yang berbeda. Penerapan algoritma DBSCAN dapat menjawab rumusan masalah pada penelitian terdahulu dengan mengelompokkan atribut yang digunakan pada masing-masing penelitian. Pada penelitian ini, algoritma DBSCAN digunakan pada data perjalanan taksi untuk mengidentifikasi pengelompokan dari titik penjemputan dan titik pengantaran yang terjadi di Kota Brooklyn. *Cluster* yang terbentuk akan diidentifikasi dan akan menghasilkan karakteristik yang dapat membantu pengelola taksi dalam mengalokasikan sopir taksi pada area titik penjemputan dan pengantaran yang potensial untuk menemukan penumpang selanjutnya.

## HASIL DAN PEMBAHASAN

### Hasil

#### 1. *Data Preprocessing*

Penelitian berfokus pada catatan perjalanan taksi di wilayah Kota Brooklyn. Untuk itu, atribut *pickup\_longitude* dan *pickup\_latitude* dapat digunakan untuk membuat atribut baru berupa *sub\_urban* yang berisi kota lainnya di sekitar Kota New York. Setelah dilakukan pengecekan *missing values*, terdapat 126 baris yang harus dibersihkan. Setelah data dilakukan pengecekan terhadap *duplicate*, *output* mencetak 0 yang berarti tidak ada data *duplicate*. Data yang tidak relevan ataupun *outlier* untuk perjalanan taksi di Kota Brooklyn akan dihilangkan.

```

id                0
vendor_id         0
pickup_datetime  0
dropoff_datetime  0
passenger_count  0
pickup_longitude  0
pickup_latitude   0
dropoff_longitude 0
dropoff_latitude  0
store_and_fwd_flag 0
trip_duration     0
sub_urban         126
dtype: int64
    
```

**Gambar 2.** *Missing Value Pada Dataset*

Pada atribut *pickup\_datetime* dan *dropoff\_datetime*, terdapat catatan berupa waktu penjemputan dan waktu pengantaran pada hari Senin hingga Minggu. Dilakukan

pengelompokan berdasarkan hari kerja dan hari libur yang bertujuan untuk mengetahui catatan perjalanan berdasarkan karakteristik waktu. Pengelompokan tersebut dibagi berdasarkan hari kerja (Senin – Jumat) dan hari libur (Sabtu dan Minggu). Perjalanan taksi di wilayah *Brooklyn* juga didominasi pada hari kerja.

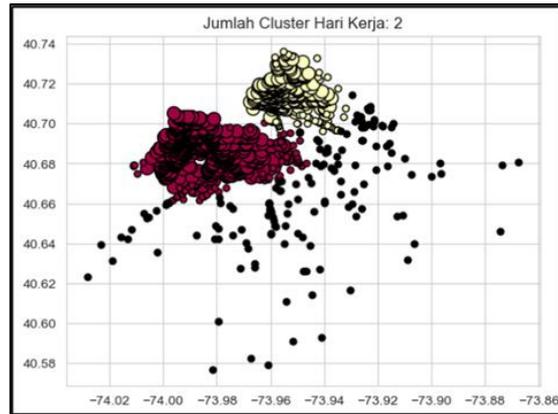
## 2. Clustering

### a. Titik Penjemputan

Setelah data melewati proses *data preprocessing*, data siap untuk masuk ke tahap *clustering* menggunakan algoritma *Density-Based Clustering of Application with Noise (DBSCAN)*. Untuk itu, diperlukan proses iterasi untuk menemukan parameter terbaik yang akan digunakan pada *clustering* data perjalanan taksi. Pada titik penjemputan hari kerja, iterasi parameter dilakukan dengan rentang nilai Eps yaitu rentang 0.006 hingga 0.008 dan MinPts 21 hingga 30. Berdasarkan 50 iterasi, terdapat parameter terbaik dengan nilai *silhouette coefficient* sebesar 0.546 yaitu Eps = 0.0075 dan MinPts = 29.

**Tabel 2.** Iterasi Parameter Titik Penjemputan Hari Kerja

<i>Eps</i> \ <i>MinPts</i>	0.006	0.0065	0.007	0.0075	0.008
21	0.533	0.541	0.542	0.493	0.503
22	0.413	0.542	0.541	0.486	0.502
23	0.406	0.538	0.541	0.486	0.495
24	0.402	0.537	0.544	0.485	0.495
25	0.387	0.533	0.543	0.483	0.492
26	0.526	0.527	0.543	0.481	0.490
27	0.525	0.527	0.539	0.544	0.485
28	0.524	0.527	0.539	0.544	0.485
29	0.524	0.533	0.535	0.546	0.485
30	0.520	0.531	0.538	0.540	0.485

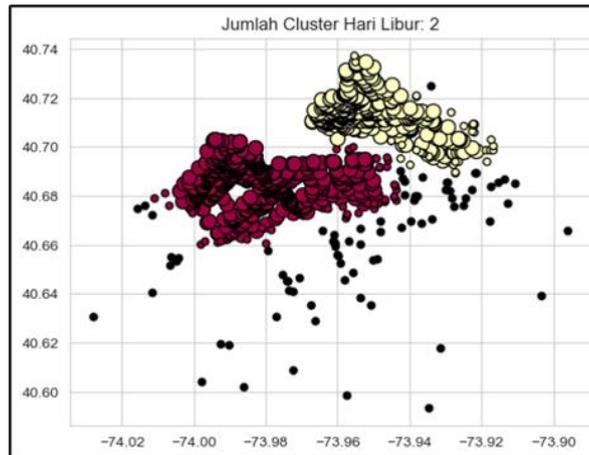


**Gambar 3.** Cluster Titik Penjemputan Hari Kerja

Selanjutnya, dilakukan iterasi parameter untuk titik penjemputan hari libur. Iterasi parameter dilakukan dengan rentang nilai Eps yaitu rentang 0.006 hingga 0.008 dan MinPts 20 hingga 30. Berdasarkan 55 iterasi, terdapat parameter terbaik dengan nilai *silhouette coefficient* sebesar 0.524 yaitu Eps = 0.008 dan MinPts = 25.

**Tabel 3.** Iterasi Parameter Titik Penjemputan Hari Libur

<i>Eps</i> / <i>MinPts</i>	0.006	0.0065	0.007	0.0075	0.008
20	0.388	0.470	0.517	0.521	0.443
21	0.452	0.461	0.517	0.521	0.441
22	0.451	0.453	0.517	0.523	0.520
23	0.451	0.449	0.517	0.523	0.521
24	0.449	0.448	0.469	0.523	0.521
25	0.511	0.448	0.450	0.523	0.524
26	0.511	0.455	0.450	0.522	0.523
27	0.513	0.511	0.449	0.519	0.523
28	0.512	0.510	0.455	0.514	0.522
29	0.331	0.511	0.454	0.515	0.520
30	0.312	0.511	0.454	0.518	0.519



**Gambar 4.** Cluster Titik Penjemputan Hari Libur

Setelah mendapatkan parameter terbaik, parameter tersebut akan dimasukkan pada algoritma DBSCAN dan menghasilkan *cluster* seperti gambar berikut. Sesuai dengan hasil iterasi pada titik penjemputan hari kerja dengan menggunakan parameter  $Eps = 0.0075$  dan  $MinPts = 29$  menghasilkan 2 *cluster* besar yang ditandai dengan warna kuning dan merah. Pada titik penjemputan hari libur, dengan menggunakan parameter  $Eps = 0.008$  dan  $MinPts = 25$  menghasilkan 2 *cluster* dengan warna kuning dan merah. Titik-titik hitam merupakan titik *noise* dan tidak masuk ke dalam *cluster*. Setelah jumlah titik pada masing-masing *cluster* dicetak, pada titik penjemputan hari kerja terdapat 148 titik *noise* yang dilabeli dengan -1. Sedangkan pada titik penjemputan hari libur terdapat 78 titik *noise*. Berikut merupakan jumlah anggota *cluster* titik penjemputan pada hari kerja dan hari libur.

cluster	
0	911
1	400
-1	148

**Gambar 5.** Jumlah Anggota Cluster Penjemputan Hari Kerja

cluster	
0	607
1	438
-1	78

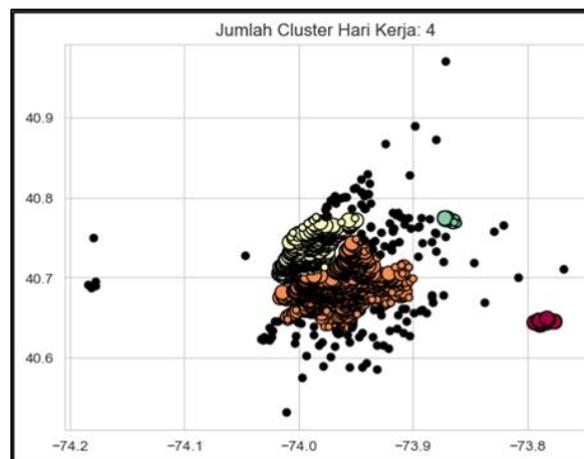
**Gambar 6.** Jumlah Anggota Cluster Penjemputan Hari Libur

b. Titik Pengantaran

Pada titik pengantaran hari kerja, iterasi parameter dilakukan dengan rentang nilai Eps yaitu rentang 0.006 hingga 0.008 dan MinPts 2 hingga 10. Berdasarkan 45 iterasi, terdapat parameter terbaik dengan nilai *silhouette coefficient* sebesar 0.321 yaitu Eps = 0.008 dan MinPts = 8.

**Tabel 4.** Iterasi Parameter Titik Pengantaran Hari Kerja

<i>Eps</i> / <i>MinPts</i>	0.006	0.0065	0.007	0.0075	0.008
2	0.024	0.047	0.055	0.096	0.139
3	0.052	0.062	0.090	0.122	0.177
4	0.063	0.078	0.133	0.166	0.172
5	0.153	0.147	0.228	0.231	0.203
6	0.168	0.179	0.182	0.247	0.235
7	0.004	0.122	0.121	0.306	0.310
8	0.039	0.014	0.106	0.318	0.321
9	0.208	0.089	0.105	0.316	0.318
10	0.206	0.211	0.226	0.098	0.316

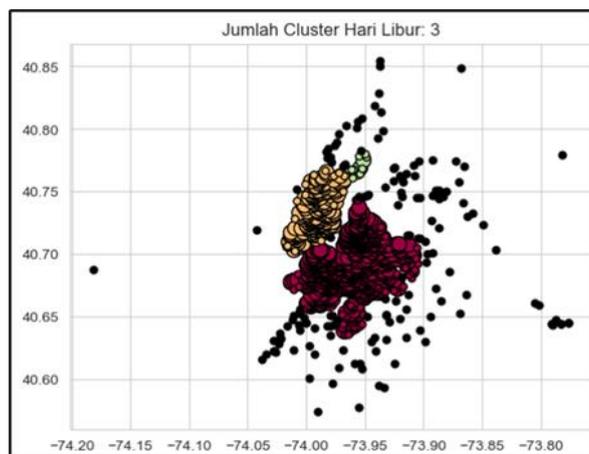


**Gambar 7.** Cluster Titik Pengantaran Hari Kerja

Selanjutnya, dilakukan iterasi parameter untuk titik pengantaran hari libur. Iterasi parameter dilakukan dengan rentang nilai Eps yaitu rentang 0.006 hingga 0.008 dan MinPts 2 hingga 10. Berdasarkan 45 iterasi, terdapat parameter terbaik dengan nilai *silhouette coefficient* sebesar 0.258 yaitu Eps = 0.008 dan MinPts = 7.

**Tabel 5.** Iterasi Parameter Titik Pengantaran Hari Libur

<i>Eps</i> / <i>MinPts</i>	0.006	0.0065	0.007	0.0075	0.008
2	-0.016	-0.125	-0.036	-0.052	0.090
3	-0.004	-0.123	-0.067	-0.054	0.096
4	-0.027	-0.095	-0.075	-0.005	0.216
5	-0.101	0.014	-0.053	0.098	0.219
6	-0.016	0.023	-0.079	0.100	0.251
7	-0.075	-0.000	-0.024	0.101	0.258
8	-0.094	0.003	-0.155	0.073	0.152
9	-0.126	-0.082	0.017	-0.012	0.151
10	-0.213	-0.078	0.004	-0.049	0.074



**Gambar 8.** Cluster Titik Pengantaran Hari Libur

Setelah mendapatkan parameter terbaik, parameter tersebut akan dimasukkan pada algoritma DBSCAN dan menghasilkan *cluster* seperti gambar berikut. Sesuai dengan hasil iterasi pada titik pengantaran hari kerja dengan menggunakan parameter  $Eps = 0.008$  dan  $MinPts = 8$  menghasilkan 4 *cluster* yang ditandai dengan warna merah, kuning, hijau, dan oren. Pada titik pengantaran hari libur, dengan menggunakan parameter  $Eps = 0.008$  dan  $MinPts = 7$  menghasilkan 3 *cluster* dengan warna hijau, kuning, dan merah. Titik-titik hitam merupakan titik *noise* dan tidak masuk ke dalam *cluster*. Setelah jumlah titik pada masing-masing *cluster* dicetak, pada titik pengantaran hari kerja terdapat 146 titik *noise* yang dilabeli dengan -1. Sedangkan pada titik pengantaran hari libur terdapat 151 titik

*noise*. Berikut merupakan jumlah anggota *cluster* titik pengantaran pada hari kerja dan hari libur.

cluster	
1	864
2	430
-1	146
0	17
3	14

**Gambar 9.** Jumlah Anggota *Cluster* Penjemputan Hari Kerja

cluster	
0	766
1	210
-1	151
2	12

**Gambar 10.** Jumlah Anggota *Cluster* Pengantaran Hari Libur

### 3. *Pattern Evaluation and Knowledge*

Setelah mengidentifikasi anggota *cluster* pada titik penjemputan dan titik pengantaran, visualisasi peta terhadap anggota *cluster* menjadi hal yang berguna untuk menentukan nama-nama area dengan kepadatan titik tertentu. Memahami pola ini dapat menjadi panduan bagi sopir taksi dan pengelola taksi dalam mengelola operasional taksi. Dengan informasi ini, sopir taksi dapat mengetahui area-area yang berpotensi sebagai titik penjemputan pada hari kerja dan hari libur, sehingga mereka dapat menyesuaikan penempatan taksi untuk memberikan layanan yang optimal kepada penumpang. Sopir taksi juga dapat melihat area yang cenderung ramai dengan penumpang pada hari kerja dan hari libur.

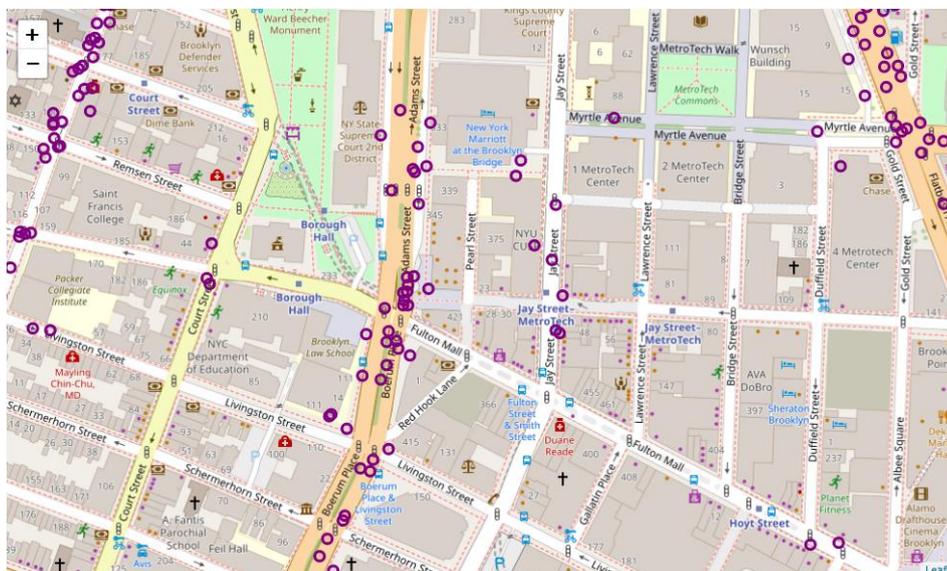
Dengan pengetahuan tentang area potensial titik penjemputan, sopir taksi dapat mengelola persediaan bahan bakar taksi dengan efisien untuk menghindari kehabisan bahan bakar. Selain itu, mereka dapat mengurangi waktu sopir taksi berkendara tanpa penumpang yang dapat meningkatkan potensi pendapatan. Melalui pemahaman tentang area potensial ini, sopir taksi juga dapat membuat keputusan cerdas terkait waktu istirahat mereka, memilih untuk beristirahat pada saat permintaan taksi rendah.

Berikut adalah visualisasi peta yang akan mempermudah pengguna dalam mengidentifikasi nama-nama jalan yang menjadi area potensial di Kota Brooklyn.

## Pembahasan

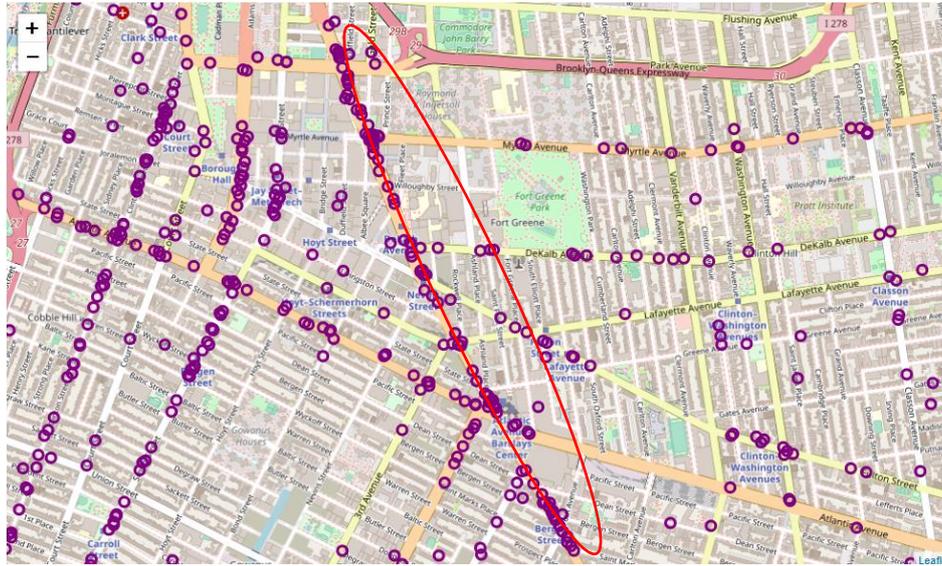
### A. Titik Penjemputan

Dengan memperbesar visualisasi peta pada area Boerum Place dan Adams Street pada hari kerja, dapat ditemukan bahwa di bagian barat laut Kota Brooklyn terdapat beberapa titik penjemputan terletak di sekitar perguruan tinggi seperti Saint Francis College, Packer Collegiate Institute, dan Brooklyn Law School. Selain itu, titik penjemputan juga terpecah di sekitar Borough Hall, sebuah bangunan pemerintahan. Di area ini, terdapat stasiun transportasi umum yang berdekatan dengan titik penjemputan Borough Hall dan Jay Street. Simbol-simbol pada visualisasi juga menandakan keberadaan fasilitas umum seperti hotel, pusat perbelanjaan, dan kantor bank di sekitar daerah tersebut.



**Gambar 11.** Area Potensial Titik Penjemputan Hari Kerja

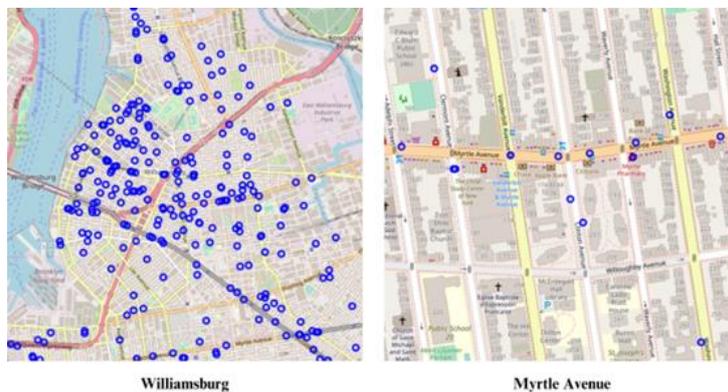
Pada hari libur, titik penjemputan banyak terdistribusi di sepanjang jalan Flatbush Avenue. Namun, titik penjemputan yang memiliki kepadatan tinggi terletak di persimpangan jalan yang melintasi Flatbush Avenue, dekat dengan pusat perbelanjaan, tempat ibadah, dan area hiburan. Flatbush Avenue, sebuah arteri utama yang melintasi Kota Brooklyn, melewati beberapa wilayah utama.



**Gambar 12.** Visualisasi Peta Titik Penjemputan Hari Libur di Flatbush Avenue

### B. Titik Pengantaran

Titik pengantaran pada hari kerja tersebar di bagian utara Kota Brooklyn yang mencakup wilayah Bushwick, Greenpoint, dan Williamsburg. Beberapa jalan utama yang sering menjadi destinasi penumpang taksi antara lain Myrtle Avenue, Flatbush Avenue, dan Lafayette Avenue yang terletak di Downtown Brooklyn dan Fort Greene. Dalam visualisasi berikut, terlihat bahwa area titik pengantaran memiliki distribusi yang signifikan pada wilayah Williamsburg. Lokasi pengantaran juga menyebar luas di bagian utara Kota Brooklyn yaitu wilayah Williamsburg sebagai tujuan populer sebab kemudahan akses ke Manhattan, kekayaan kuliner, serta kehidupan musik dan hiburan. Daerah ini menjadi tujuan pengantaran yang banyak dipilih, terutama pada hari libur.



**Gambar 13.** Visualisasi Peta Titik Pengantaran di Williamsburg dan Myrtle Avenue

## SIMPULAN

Dalam mengidentifikasi lokasi permintaan taksi, data geospasial menjadi data yang penting dan dapat memberikan informasi untuk mengeksplorasi lingkungan spasial perkotaan. Salah satu metode pionir yang dapat digunakan untuk mengidentifikasi pola tersebut adalah *Density-Based Spatial Clustering of Application with Noise* (DBSCAN). Pada titik penjemputan, parameter terbaik pada hari kerja yaitu  $Eps = 0.0075$  dan  $MinPts = 29$  yang menghasilkan 2 *cluster* besar dan 148 titik *noise*. Selanjutnya, parameter titik penjemputan pada perjalanan hari libur yaitu  $Eps = 0.008$  dan  $MinPts = 25$  yang menghasilkan 2 *cluster* dan 78 titik *noise*. Area potensial pada titik penjemputan pada hari kerja terdapat di kawasan Boerum Place dan Adams Street. Sedangkan pada hari libur, area potensial titik penjemputan terjadi di Flatbush Avenue.

Pada titik pengantaran hari kerja memperoleh parameter terbaik yaitu  $Eps = 0.008$  dan  $MinPts = 8$  yang menghasilkan 4 *cluster* dan 146 titik *noise*. Pada titik pengantaran hari libur, parameter terbaik yang digunakan adalah  $Eps = 0.0065$  dan  $MinPts = 7$  yang menghasilkan 3 *cluster* dan 151 titik *noise*. Area potensial pada titik pengantaran pada hari kerja terdapat di kawasan Myrtle Avenue, Flatbush Avenue, dan Lafayette Avenue yang terletak di Downtown Brooklyn dan Fort Greene. Sedangkan pada hari libur, area potensial titik pengantaran terjadi di Williamsburg.

Area potensial tersebut dapat dijadikan layanan taksi sebagai bahan evaluasi untuk membaca karakteristik penumpang berdasarkan sejarah perjalanan taksi. Hal ini dapat membantu layanan taksi dalam meningkatkan operasi dan mengalokasikan armada taksi pada area yang potensial. Pada lokasi penjemputan, taksi dapat dialokasikan ke area tersebut untuk memudahkan sopir taksi dalam menemukan penumpang. Selain itu, terdapat titik penjemputan pada tempat hiburan dan tempat ibadah. Pengelola taksi dapat memperhatikan jadwal operasional tempat-tempat tersebut dengan menggunakan visualisasi temporal terhadap rentang waktu tertentu untuk lebih efisien menemukan penumpang. Selanjutnya pada titik pengantaran, pengelola dan sopir taksi dapat memperkirakan destinasi penumpang berdasarkan rentang waktu. Pada hari kerja, sopir taksi dapat melihat daerah dan jalan utama dengan mobilitas tinggi. Untuk itu, sopir taksi dapat mempelajari rute alternatif yang dapat membawa penumpang ke tujuan dengan menghindari kemacetan. Mengetahui titik pengantaran penumpang juga dapat membantu

pengelola dan sopir taksi dalam merencanakan optimalisasi armada taksi, termasuk persiapan bahan bakar.

## DAFTAR RUJUKAN

- C., M., & H., C. (2022). Parallel DBSCAN Clustering Algorithm Using Hadoop Map-reduce Framework for Spatial Data. *I.J. Information Technology and Computer Science*, 3.
- Chicco, D., Oneto, L., & Tavazzi, E. (2022). Eleven quick tips for data cleaning and feature engineering. *PLoS Comput Biol* 18(12): e1010718.
- Gao, Y., Jiang, D., & Xu, Y. (2018). Optimize taxi driving strategies based on reinforcement learning. *International Journal of Geographical Information Science*.
- Hossain, M. Z., Islam, M. J., Rahman Miah, M. W., Rony, J. H., & Begum, M. (2021). Develop a dynamic DBSCAN algorithm for solving initial parameter selection problem of the DBSCAN algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, 1.
- Jian, S., Li, D., & Yu, Y. (2021). Research on Taxi Operation Characteristics by Improved DBSCAN Density Clustering Algorithm and K-means Clustering Algorithm. 1.
- Malhan, A. (2017). ST-OPTICS: A spatial-temporal clustering algorithm with time recommendations for taxi services. *Indraprastha Institute of Information Technology, Delhi*.
- Nurhaliza, N., & Mustakim. (2021). Pengelompokan Data Kasus Covid-19 di Dunia Menggunakan Algoritma DBSCAN. *Indonesian Journal of Informatic Research and Software Engineering*, 1-8.
- Sari, B. N., & Primajaya, A. (2019). Penerapan Clustering DBSCAN untuk Pertanian Padi di Kabupaten Karawang. *Jurnal Informatika dan Komputer Vol. 4*.
- Sugi Almantara, I. P., Sri Ariyani, N. W., & Alit Swamardika, I. B. (2020). Spatial Data Analysis using DBSCAN Method and KNN classification. *International Journal of Engineering and Emerging Technology, Vol.5, No.2*.
- Suyanto, 1.-. (2017). *Data mining : untuk klasifikasi dan klasterisasi data / Suyanto*. Bandung: Penerbit Informatika.

- Ulak, M. B., Yazici, A., & Aljarrah, M. (2020). Value of convenience for taxi trips in New York City.
- United Nations, Department of Economic and Social Affairs, Population Division. (2017). *World Population Ageing 2017 (ST/ESA/SER.A/408)*.
- Vizquete-Luciano, E., Guill'en-Pujadas, M., Alaminos, D., & Merigo-Lindahl, J. M. (2023). Taxi and urban mobility studies: A bibliometric analysis. 1.
- You, L., Guan, Z., Li, N., Zhang, J., Cui, H., Claramunt, C., & Cao, R. (2021). A Spatio-Temporal Schedule-Based Neural Network for Urban Taxi Waiting Time Prediction. *ISPRS International Journal of Geo-Information*.
- Zhang, C., Zhu, F., Wang, X., Sun, L., Tang, H., & Lv, Y. (2020). Taxi Demand Prediction Using Parallel Multi-Task Learning Model. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*.