

Optimasi Layanan TransJakarta

Berbasis Analisis Kepadatan, Pola Pembayaran, dan Efisiensi
Koridor

Adi Nugraha





Daftar isi



Profil Perusahaan



Latar Belakang Perusahaan



Data Understanding



Data Cleaning



Feature Engineering



Exploratory Data Analysis 1



Exploratory Data Analysis 2



Kesimpulan & Solusi





Profil Perusahaan

TransJakarta

TransJakarta adalah sistem BRT terbesar di Indonesia yang beroperasi di DKI Jakarta dan sekitarnya sejak 2004. Dengan jaringan yang luas dan banyak halte, TransJakarta sangat penting bagi mobilitas harian masyarakat. Setiap hari, jutaan orang menggunakan layanan ini untuk berbagai kebutuhan, seperti bekerja, bersekolah, dan rekreasi.

Latar Belakang Analisa TransJakarta

Kinerja Operasional

- Kapasitas penumpang
- Frekuensi bus beroperasi pada suatu rute
- Waktu Tempuh Bus

Fasilitas dan Infrastruktur

- Halte
- Armada
- Jalur khusus
- Koridor


Dampak sosial ekonomi

- Kemacetan
- Mobilitas
- Lingkungan



Data Understanding

Tahap awal dalam proses analisis data atau proyek sains data yang bertujuan untuk memahami karakteristik, struktur, dan kualitas data yang akan digunakan. Ini melibatkan pengumpulan, eksplorasi, dan pemeriksaan data untuk mengidentifikasi pola, dan potensi masalah yang mungkin ada.



Data Understanding

Deskripsi Kolom

Terdapat 22 kolom pada dataset transjakarta:

- transID: ID transaksi unik untuk setiap transaksi.
- payCardID: Identifikasi utama pelanggan, yaitu kartu yang digunakan sebagai tiket untuk masuk dan keluar.
- payCardBank: Nama bank penerbit kartu pelanggan.
- payCardName: Nama pelanggan yang tertanam di dalam kartu.
- payCardSex: Jenis kelamin pelanggan yang tertanam di dalam kartu.
- payCardBirthDate: Tahun lahir pelanggan.
- corridorID: ID koridor/rute sebagai kunci untuk pengelompokan rute.
- corridorName: Nama koridor/rute yang mencantumkan titik awal dan akhir setiap rute.
- direction: 0 untuk Pergi, 1 untuk Pulang. Menunjukkan arah perjalanan.
- tapInStops: ID halte masuk untuk mengidentifikasi nama halte tempat pelanggan masuk.
- tapInStopsName: Nama halte masuk tempat pelanggan melakukan tap masuk.
- tapInStopsLat: Garis lintang (latitude) dari halte masuk.
- tapInStopsLon: Garis bujur (longitude) dari halte masuk.
- stopStartSeq: Urutan halte, misalnya halte pertama, halte kedua, dll., yang terkait dengan arah perjalanan.
- tapInTime: Waktu pelanggan melakukan tap masuk (tanggal dan waktu).
- tapOutStops: ID halte keluar untuk mengidentifikasi nama halte tempat pelanggan keluar.
- tapOutStopsName: Nama halte keluar tempat pelanggan melakukan tap keluar.
- tapOutStopsLat: Garis lintang (latitude) dari halte keluar.
- tapOutStopsLon: Garis bujur (longitude) dari halte keluar.
- stopEndSeq: Urutan halte, misalnya halte pertama, halte kedua, dll., yang terkait dengan arah perjalanan.
- tapOutTime: Waktu pelanggan melakukan tap keluar (tanggal dan waktu).
- payAmount: Jumlah biaya yang dibayarkan pelanggan. Beberapa perjalanan mungkin gratis, beberapa lainnya berbayar.

Type Data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 37900 entries, 0 to 37899
Data columns (total 22 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   transID                37900 non-null  object
1   payCardID              37900 non-null  int64
2   payCardBank            37900 non-null  object
3   payCardName            37900 non-null  object
4   payCardSex             37900 non-null  object
5   payCardBirthDate       37900 non-null  int64
6   corridorID             36643 non-null  object
7   corridorName           35970 non-null  object
8   direction              37900 non-null  float64
9   tapInStops             36687 non-null  object
10  tapInStopsName          37900 non-null  object
11  tapInStopsLat           37900 non-null  float64
12  tapInStopsLon           37900 non-null  float64
13  stopStartSeq            37900 non-null  int64
14  tapInTime               37900 non-null  object
15  tapOutStops             35611 non-null  object
16  tapOutStopsName         36556 non-null  object
17  tapOutStopsLat          36556 non-null  float64
18  tapOutStopsLon          36556 non-null  float64
19  stopEndSeq              36556 non-null  float64
20  tapOutTime              36556 non-null  object
21  payAmount               36893 non-null  float64
dtypes: float64(7), int64(3), object(12)
memory usage: 6.4+ MB
```

Data Missing

```
0
tapOutStops    2289
corridorName   1930
tapOutStopsLat 1344
stopEndSeq     1344
tapOutTime     1344
tapOutStopsLon 1344
tapOutStopsName 1344
corridorID     1257
tapInStops     1213
payAmount      1007
payCardID      0
transID        0
payCardName    0
payCardBank    0
payCardBirthDate 0
payCardSex     0
stopStartSeq   0
tapInStopsName 0
direction      0
tapInStopsLon  0
tapInStopsLat  0
tapInTime      0
dtype: int64
```

Duplicate Data

```
[ ] data.duplicated().sum()
np.int64(0)
```

Data Cleaning

Data cleaning, atau pembersihan data, adalah proses mengidentifikasi dan memperbaiki kesalahan, ketidakakuratan, atau inkonsistensi dalam suatu kumpulan data. Tujuannya adalah untuk meningkatkan kualitas data sehingga data tersebut dapat diandalkan untuk analisis, pelaporan, dan pengambilan keputusan.

Data Cleaning

Setelah Cleaning

```
Sebelum cleaning, missing payAmount: 940
Sebelum cleaning, missing payAmount: 0
```

```
Sebelum cleaning, missing corridorID: 1257
Sebelum cleaning, missing corridorName: 1930
```

```
count
corridorID
M7B      298
JAK.48A  240
M1H      165
JAK.48B  148
3H        89
dtype: int64
```

```
Missing corridorID: 132
Missing corridorName: 885
Setelah Cleaning -----
Missing corridorID: 0
Missing corridorName: 0
```

```
Sebelum cleaning, missing tapOutTime: 1297
Setelah cleaning, missing tapOutTime: 0
```

```
Sebelum cleaning, missing tapOutStops: 912
Setelah cleaning, missing tapOutStops: 0
```

```
Sebelum cleaning, missing tapInStops: 1146
Setelah cleaning, missing tapInStops: 0
```

Sebelum Cleaning

```
0
transID      0
payCardID    0
payCardBank  0
payCardName  0
payCardSex   0
payCardBirthDate 0
corridorID   0
corridorName 0
direction    0
tapInStops   0
tapInStopsName 0
tapInStopsLat 0
tapInStopsLon 0
stopStartSeq 0
tapInTime    0
tapOutStops  0
tapOutStopsName 0
tapOutStopsLat 0
tapOutStopsLon 0
stopEndSeq   0
tapOutTime   0
payAmount    0
dtype: int64
```


Feature Engineering

Proses memilih, memanipulasi, dan mengubah data mentah menjadi fitur yang dapat digunakan dalam model machine learning.

Feature Engineering

```
Contoh hasil kolom age: 0    15
1    26
2    31
3    45
4    41
Name: age, dtype: int64
Contoh hasil kolom duration: 0    0 days 00:39:09
1    0 days 00:57:17
2    0 days 00:51:49
3    0 days 00:43:25
4    0 days 00:39:28
Name: duration, dtype: timedelta64[ns]
Contoh hasil kolom duration_hours: 0    0.652500
1    0.954722
2    0.863611
3    0.723611
4    0.657778
Name: duration_hours, dtype: float64
```

```
rushHour
Evening Rush Hour    17112
Morning Rush Hour    16530
Non Rush Hour        1785
Name: count, dtype: int64
```

```
dayCategory
Weekday    32009
Weekend     3418
Name: count, dtype: int64
```

```
serviceCategory
Transjakarta    20604
Mikrotrans     12968
Royaltrans      1646
Wisata           209
Name: count, dtype: int64
```

```
ageCategory
Dewasa    28963
Remaja     5745
Lansia      719
Name: count, dtype: int64
```

```
array(['1T', 'B14', 'T21', 'D32', 'S31', 'B13', 'D31', '1K', '6P', 'S12'],
      dtype=object)
```

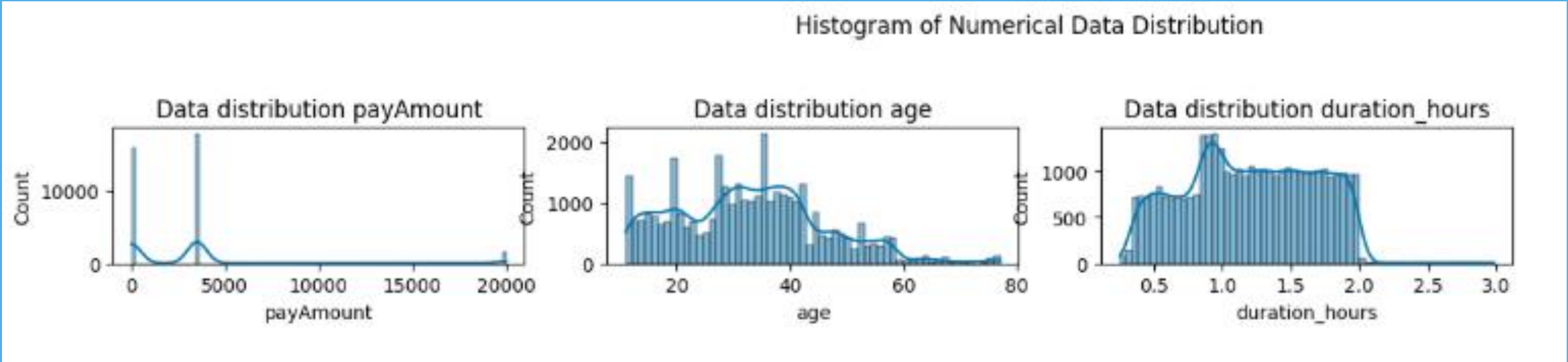


Exploratory Data Analysis 1

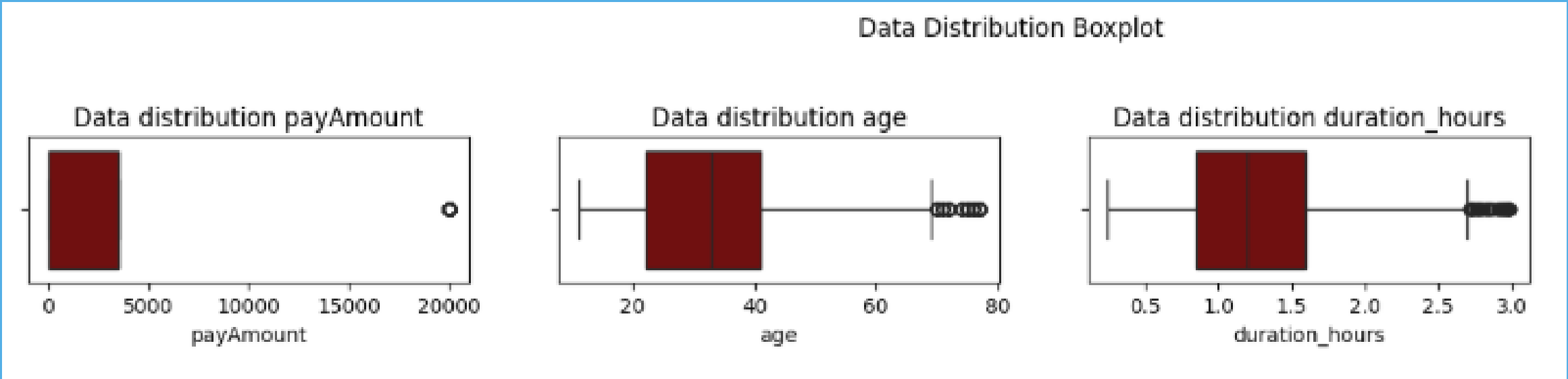
Eksplorasi Umum Untuk Memulai Analisa



Histogram



Outlier



Uji Normalitas

RESULTS FOR payCardBirthDate:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0529, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9766, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 933.5304, p-value: 0.0000 -> Not Normal

RESULTS FOR tapInStopsLat:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0537, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9885, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 473.7209, p-value: 0.0000 -> Not Normal

RESULTS FOR tapInStopsLon:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0473, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9919, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 392.6098, p-value: 0.0000 -> Not Normal

RESULTS FOR stopStartSeq:
Kolmogorov-Smirnov - Statistic: 0.8559, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.1335, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.8844, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 6680.5478, p-value: 0.0000 -> Not Normal

RESULTS FOR tapOutStopsLat:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0498, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9878, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 543.9942, p-value: 0.0000 -> Not Normal

RESULTS FOR tapOutStopsLon:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0465, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9934, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 263.0427, p-value: 0.0000 -> Not Normal

RESULTS FOR stopEndSeq:
Kolmogorov-Smirnov - Statistic: 0.9594, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.1039, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9427, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 3084.9326, p-value: 0.0000 -> Not Normal

RESULTS FOR payAmount:
Kolmogorov-Smirnov - Statistic: 0.5505, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.3771, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.5106, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 23207.7485, p-value: 0.0000 -> Not Normal

RESULTS FOR age:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0529, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9766, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 933.5304, p-value: 0.0000 -> Not Normal

RESULTS FOR year:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: nan, p-value: nan -> Not Normal
Shapiro-Wilk - Statistic: 1.0000, p-value: 1.0000 -> Normal
D'Agostino Pearson - Statistic: nan, p-value: nan -> Not Normal

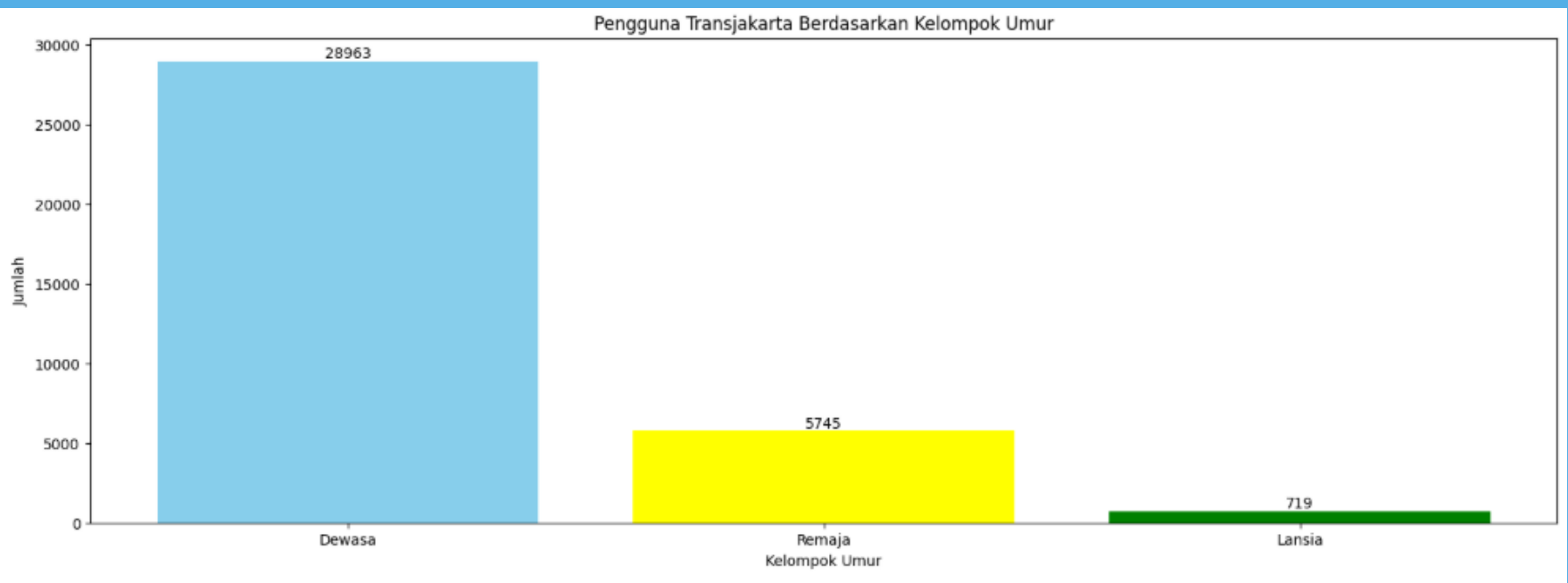
RESULTS FOR hour:
Kolmogorov-Smirnov - Statistic: 1.0000, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.2114, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.8548, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 155987.1164, p-value: 0.0000 -> Not Normal

RESULTS FOR duration_hours:
Kolmogorov-Smirnov - Statistic: 0.6271, p-value: 0.0000 -> Not Normal
Lilliefors Test - Statistic: 0.0506, p-value: 0.0010 -> Not Normal
Shapiro-Wilk - Statistic: 0.9702, p-value: 0.0000 -> Not Normal
D'Agostino Pearson - Statistic: 5192.7761, p-value: 0.0000 -> Not Normal

Demografi Pengguna



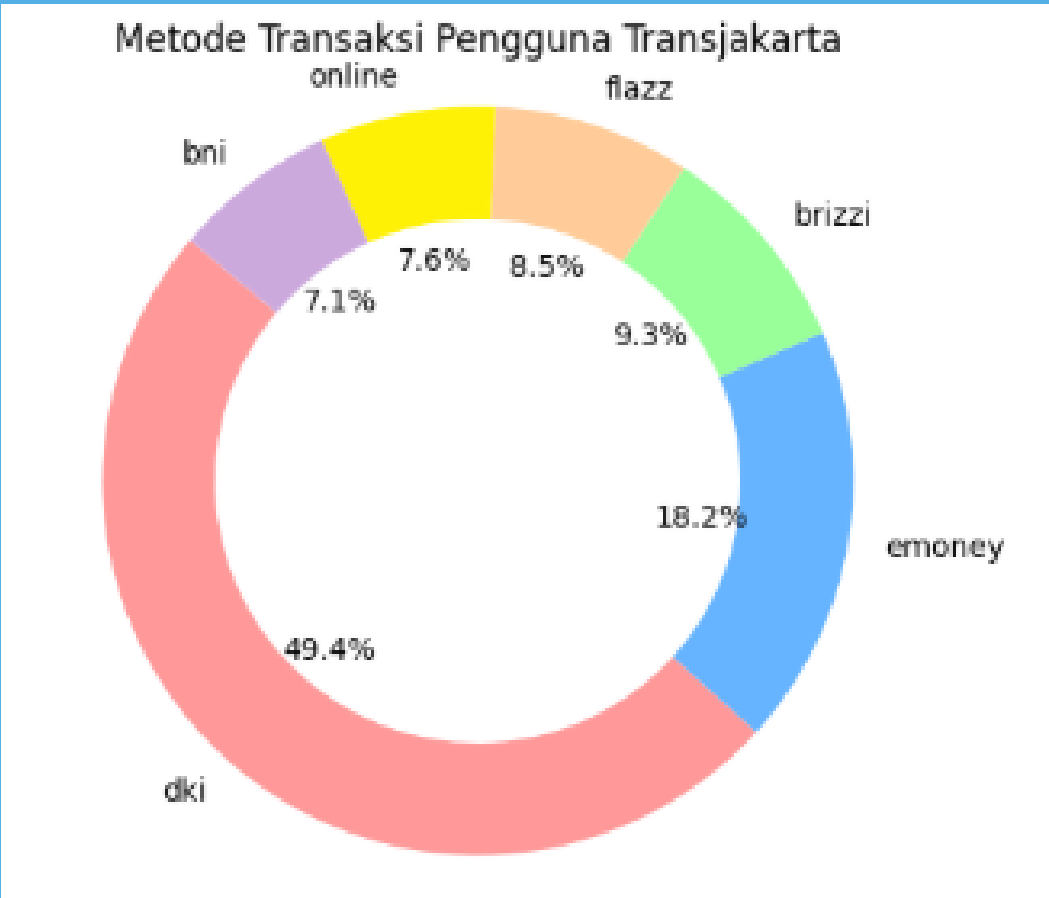
Berdasarkan Kelompok Umur



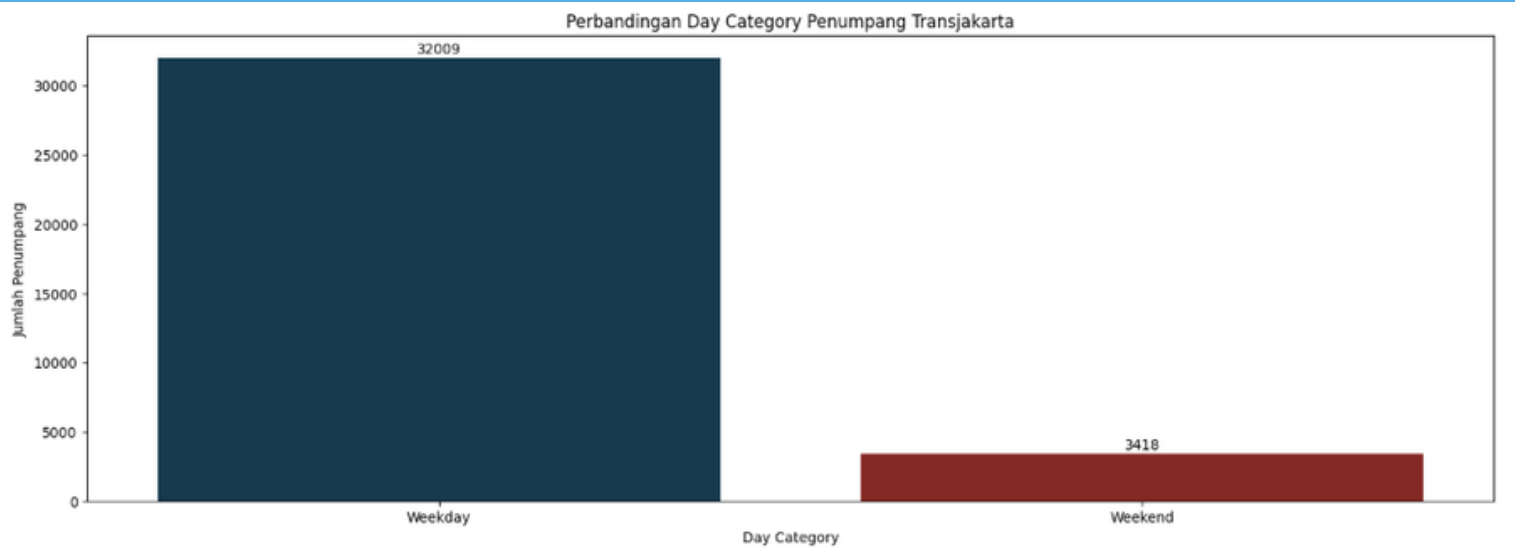
Berdasarkan Jenis Kelamin



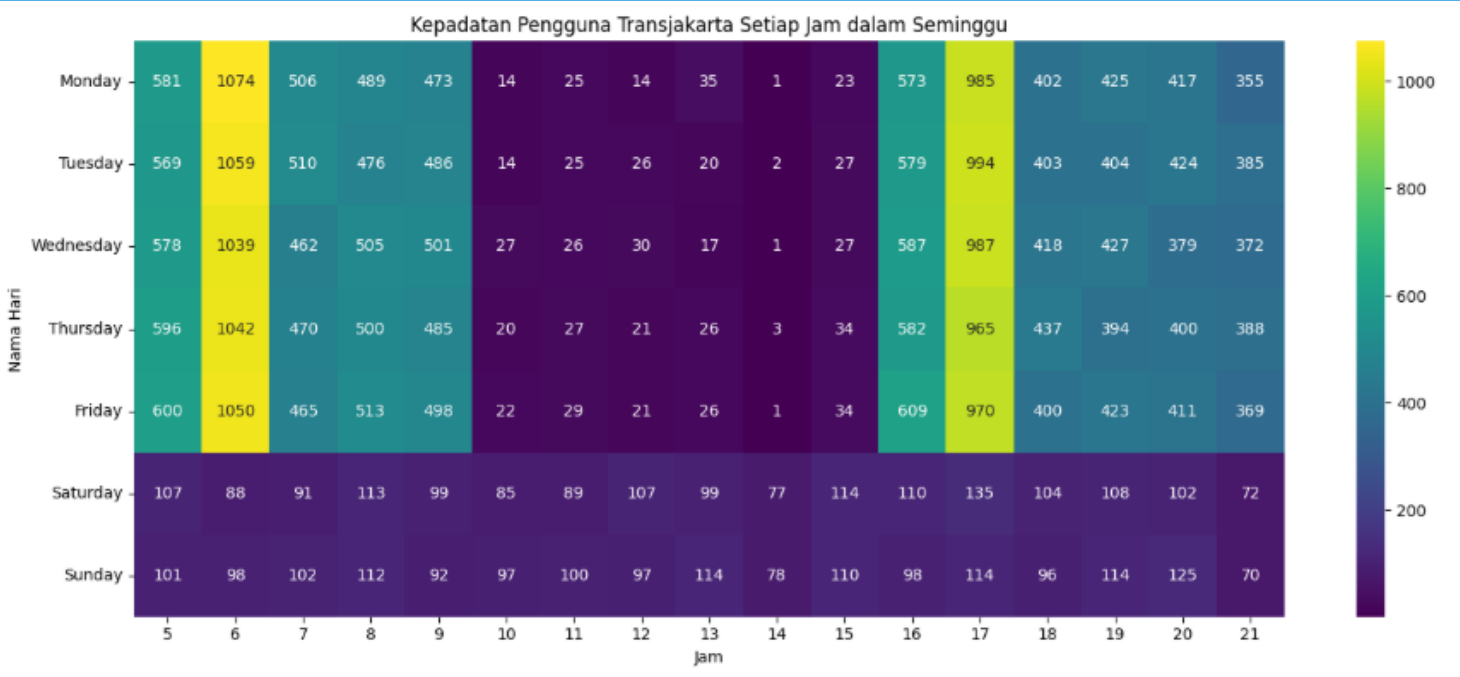
Jenis Metode Pembayaran



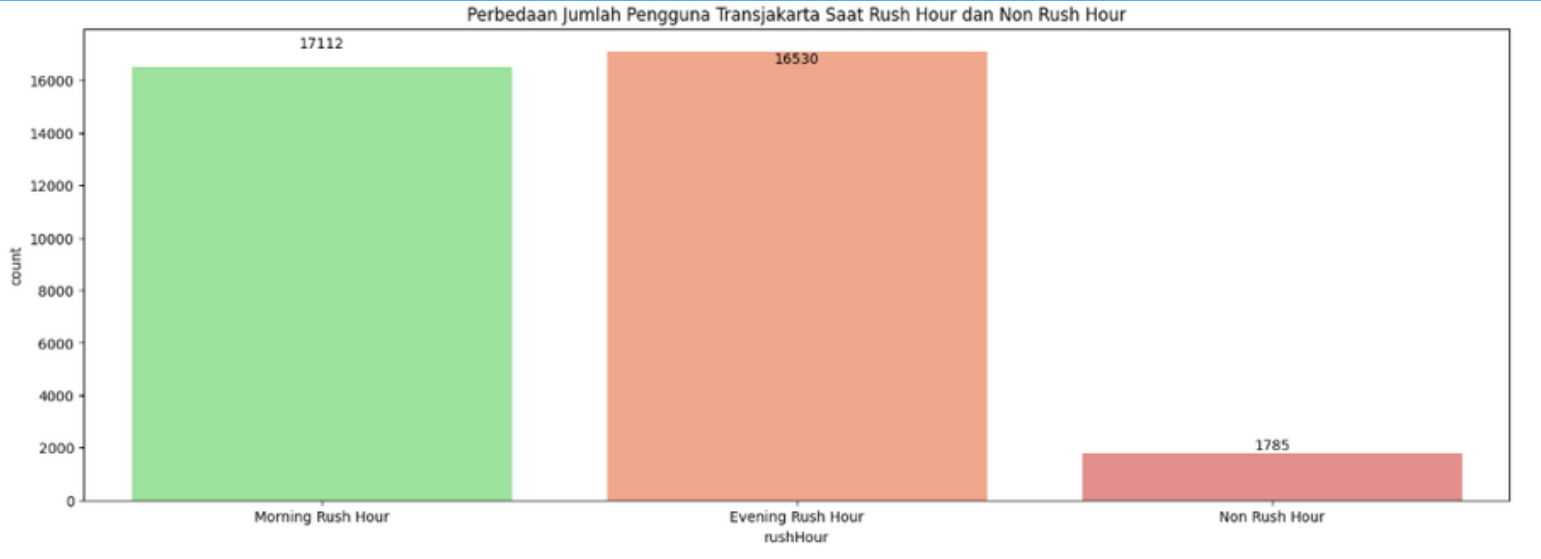
Days Category Pengguna



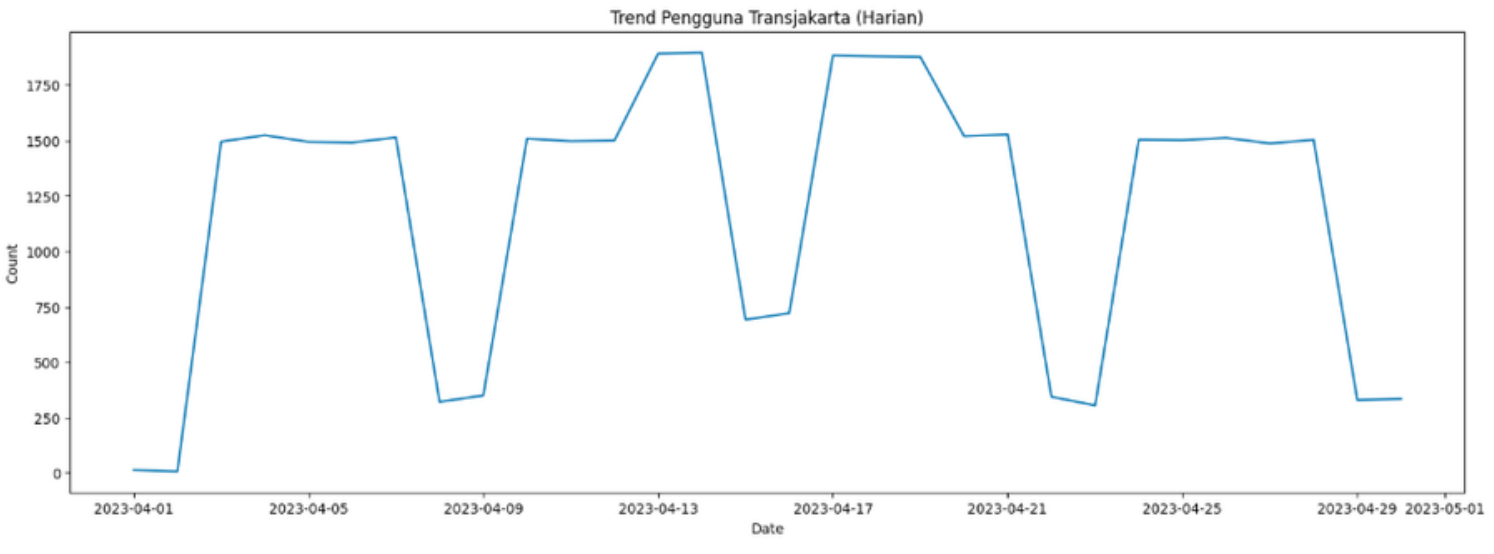
Setiap Jam dalam Seminggu



Pengguna Saat Jam Sibuk & Tidak



Trend Harian





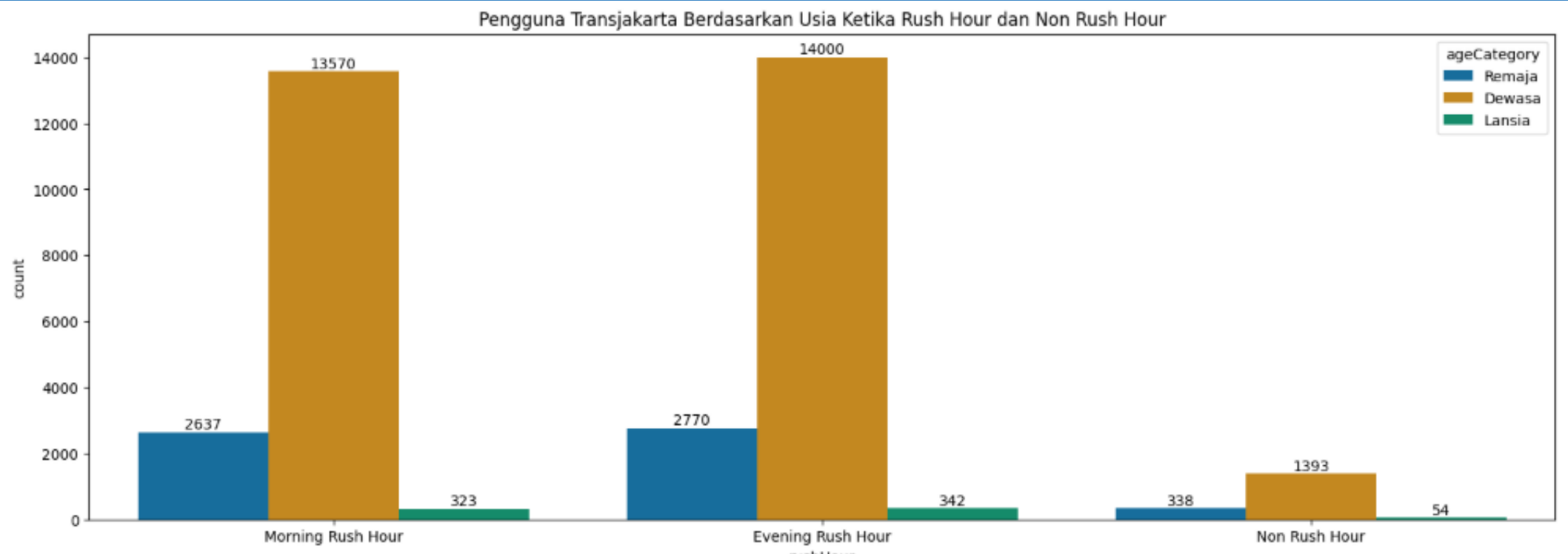
Exploratory Data Analysis 2

Eksplorasi Lebih Dalam

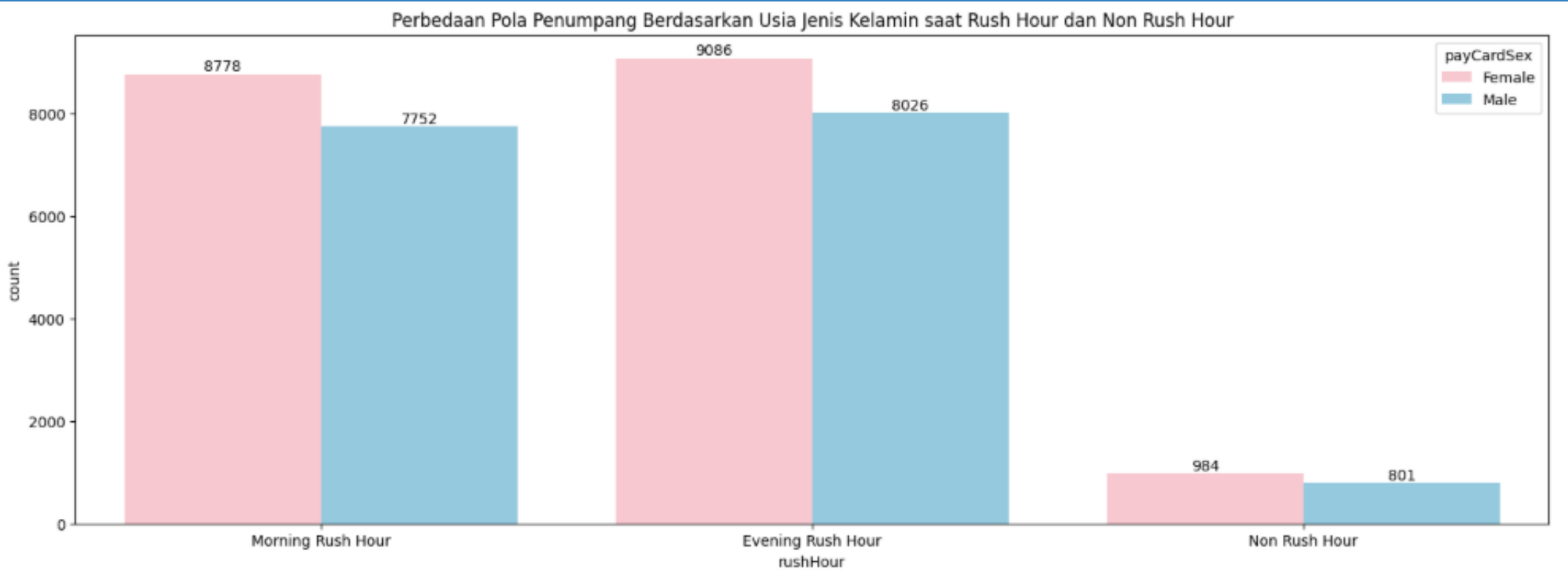


Pola Pengguna Transjakarta Ketika Rush Hour & Non Rush Hour

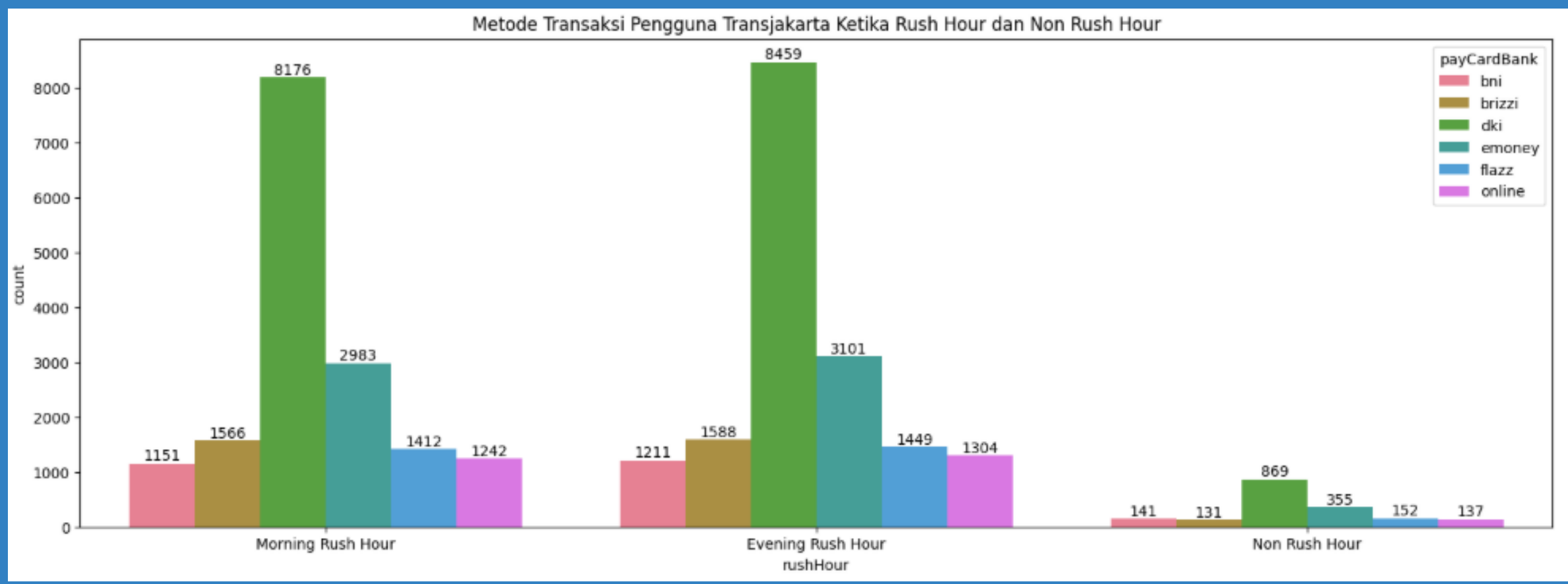
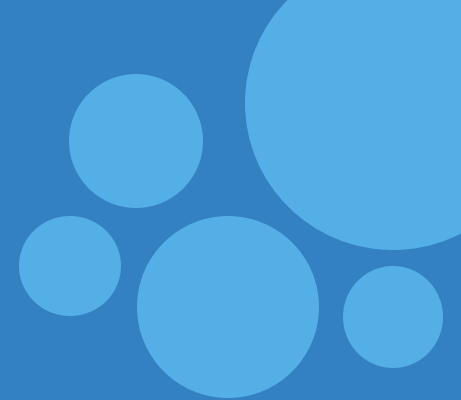
Berdasarkan Usia



Jenis Kelamin

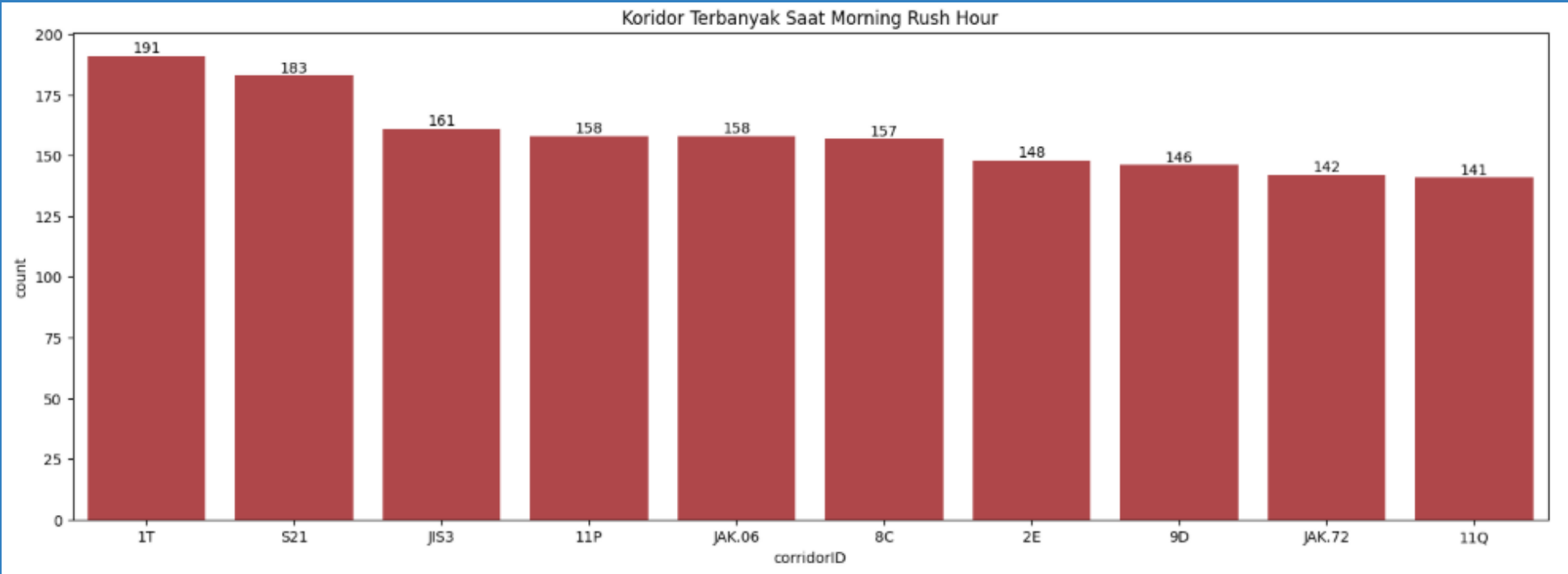


Pola Pengguna Transjakarta Ketika Rush Hour & Non Rush Hour

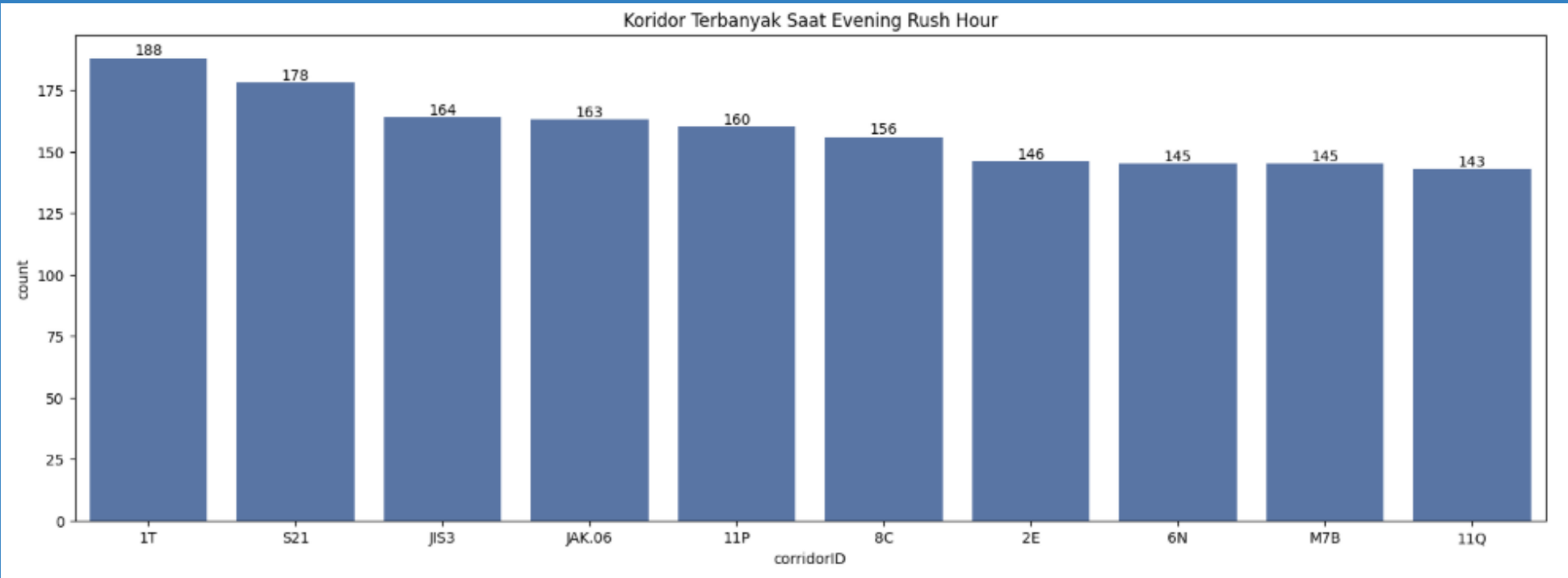


Jenis Metode Transaksi

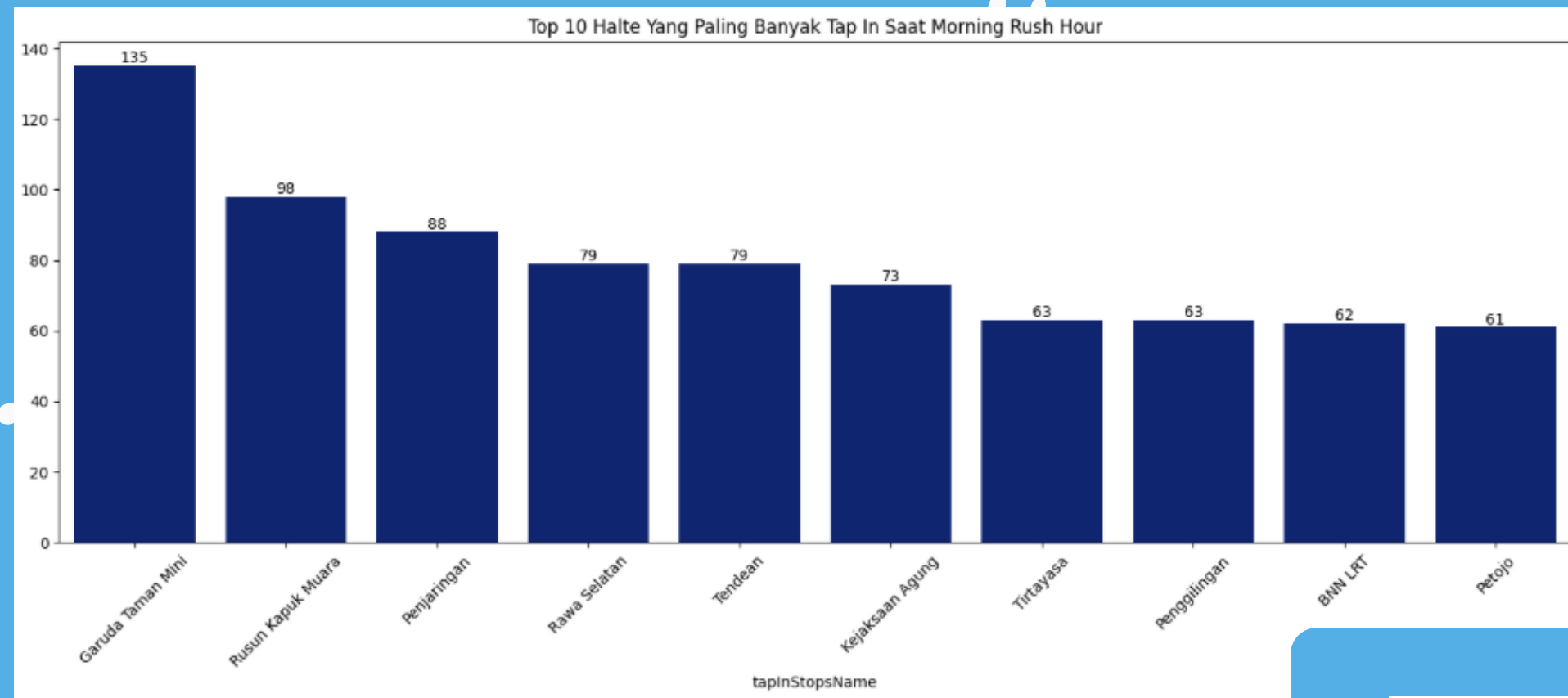
Pola Pengguna Transjakarta Berdasarkan Koridor Terpadat



Morning Hours

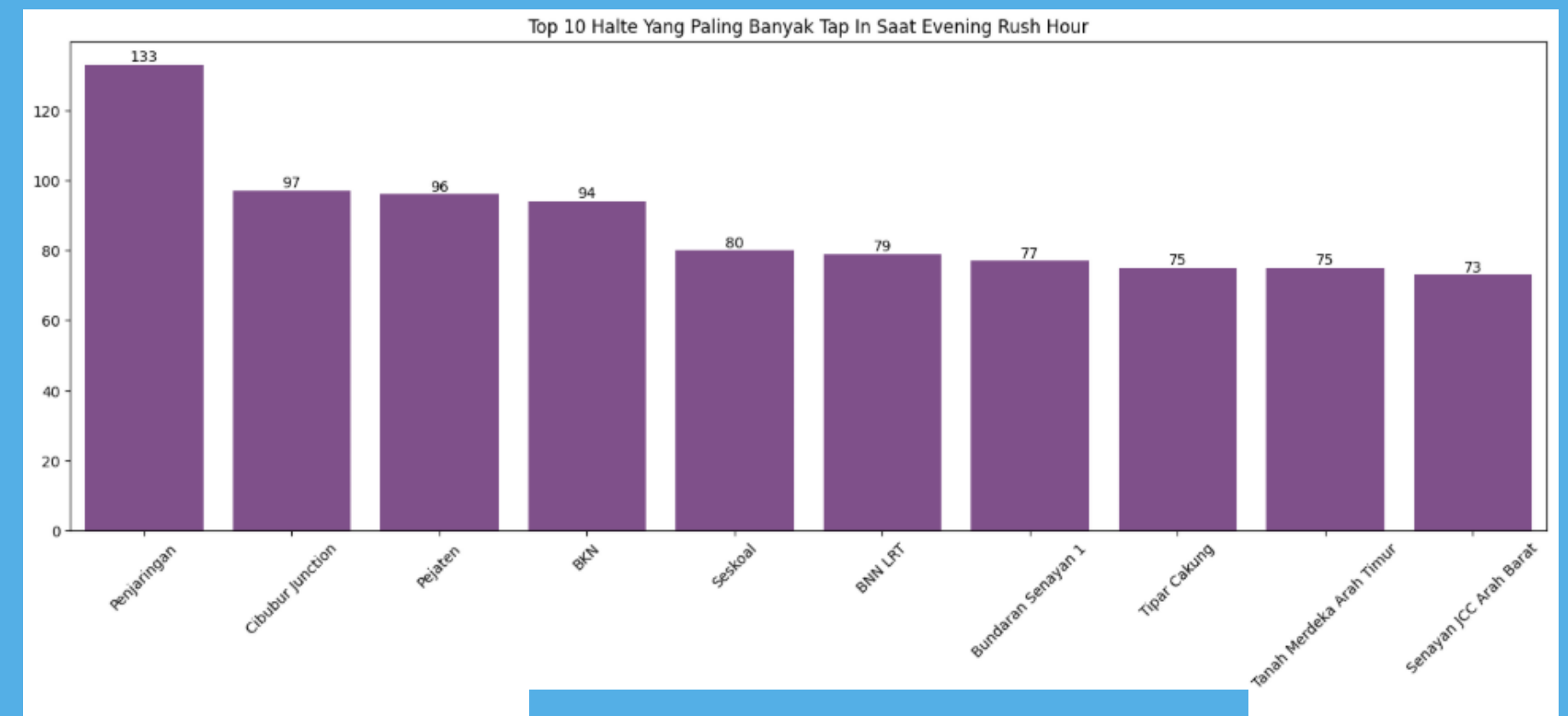


Evening Hours



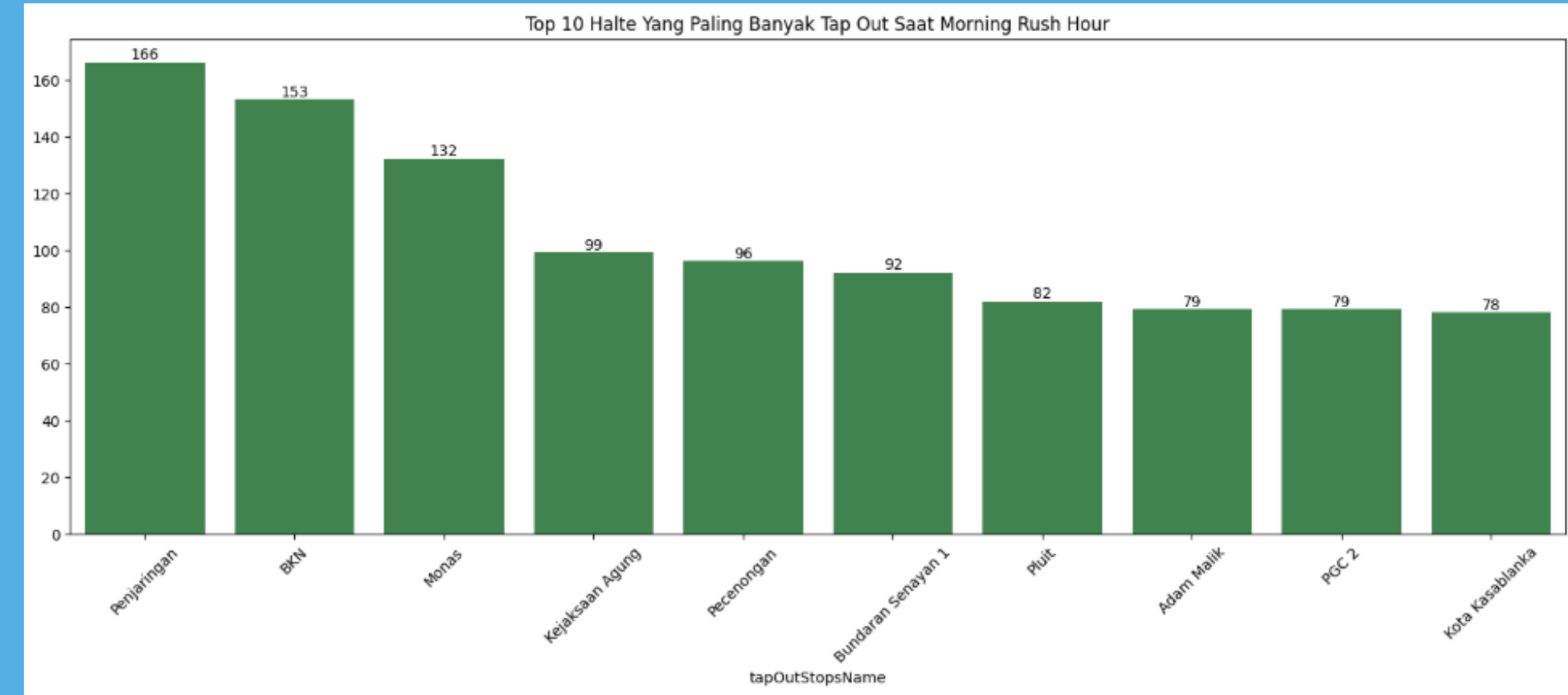
Morning Hours

Pola Pengguna
TransJakarta
Berdasarkan Halte
Terpadat saat Tap In



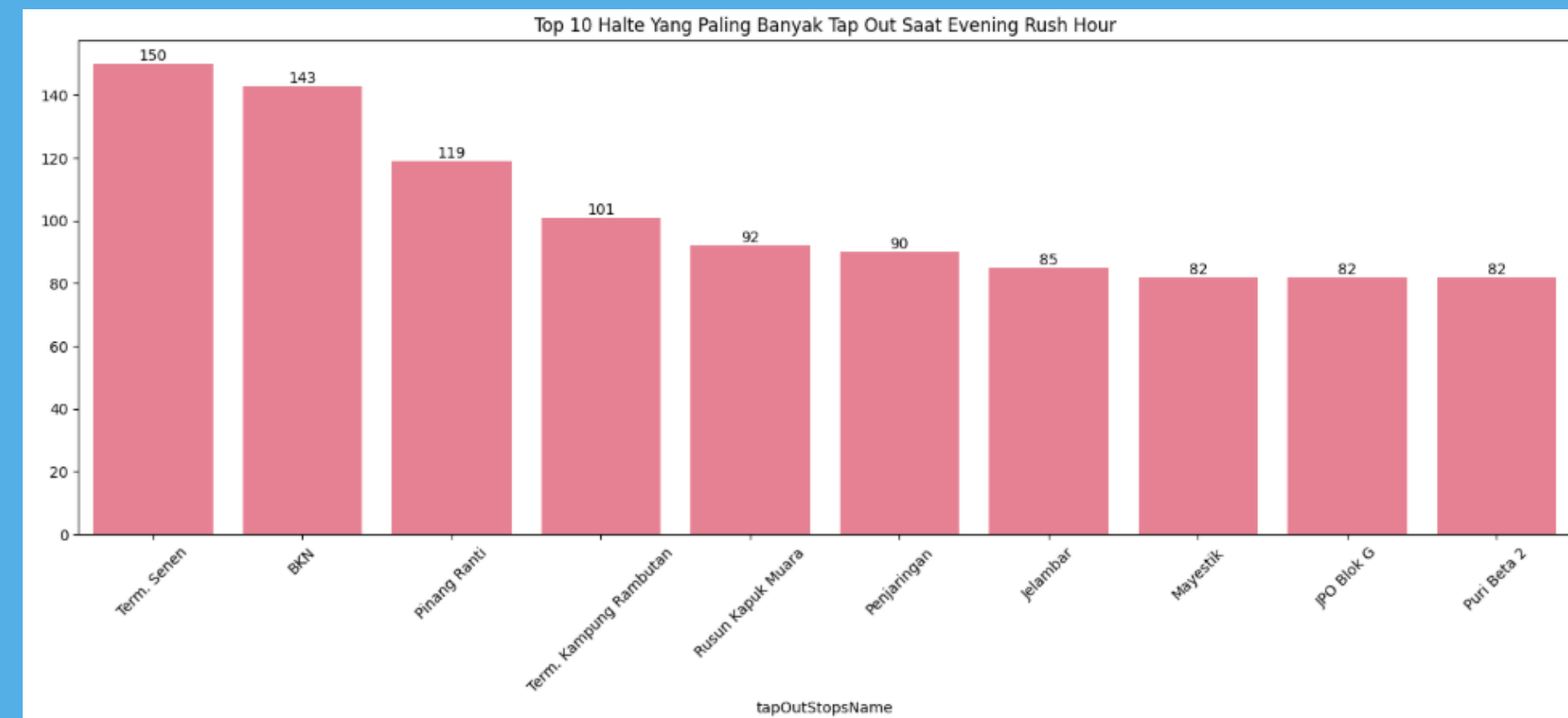
Evening Hours

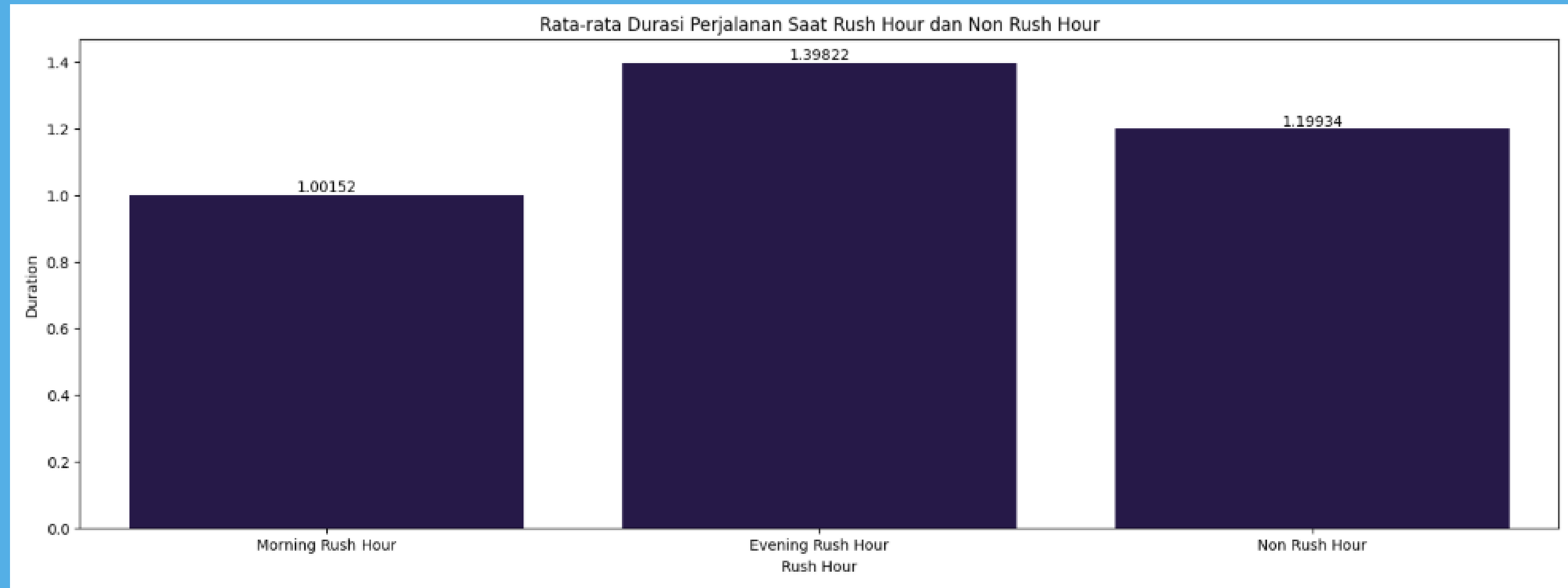
Morning



Pola Pengguna
TransJakarta
Berdasarkan Halte
Terpadat saat Tap Out

Evening





Durasi Perjalanan

TransJakarta

Saat Jam Sibuk & Tidak

Kesimpulan

- Kelompok umur pengguna transjakarta yaitu, remaja, dewasa dan lansia.
- Pengguna transjakarta didominasi oleh perempuan.
- Jenis transaksi yang paling banyak digunakan adalah kartu dki.
- Kepadatan pengguna transjakarta terjadi saat weekday dan terjadi saat evening rush hour pukul 17.00.
- Menunjukkan bahwa sebagian besar pengguna saat rush hour adalah kelompok usia produktif, yaitu 19- 59 tahun dengan didominasi pengguna berjenis kelamin perempuan
- Koridor mana saja yang mengalami kepadatan saat rush hour ? Koridor yang mengalami kepadatan saat rush hour yaitu 1T, S21, dan JIS3.
- Halte yang mengalami tap in paling banyak garuda taman mini.
- Halte dengan tap out paling banyak halte penjaringan
- Rush hour (jam sibuk) memberikan dampak negatif pada efisiensi koridor Transjakarta. Peningkatan volume kendaraan, terutama pada jam masuk dan pulang kerja, menyebabkan kemacetan pada jalur umum yang pada akhirnya memengaruhi waktu tempuh dan keteraturan operasional Transjakarta.
- Saat jam sibuk (rush hour), waktu tempuh perjalanan Transjakarta bisa lebih lama dari biasanya karena kepadatan lalu lintas.



Solusi

- Untuk meningkatkan penggunaan Transjakarta oleh seluruh kelompok umur, dapat meningkatkan fasilitas, kemudahan akses, dan kampanye edukasi.
- Untuk mencapai keseimbangan pengguna Transjakarta antara perempuan dan laki-laki, perlu ada peningkatan keamanan dan kenyamanan untuk perempuan, perubahan paradigma sosial yang bias gender, serta kebijakan yang mendukung kesetaraan gender.
- Untuk meratakan penggunaan transaksi di Transjakarta, solusi yang bisa dilakukan antara lain meningkatkan promosi metode pembayaran, memastikan ketersediaan di halte, memberikan insentif, melatih petugas, dan mengembangkan aplikasi agar lebih mudah digunakan.
- Untuk mengatasi disparitas jumlah penumpang Transjakarta antara hari kerja dan akhir, dapat diterapkan peningkatan frekuensi layanan, penyesuaian rute, serta peningkatan fasilitas dan informasi. Selain itu, edukasi dan promosi penggunaan Transjakarta juga penting untuk meningkatkan kesadaran masyarakat.
- Untuk meratakan kepadatan di koridor Transjakarta, bisa diterapkan optimalisasi waktu operasional, penambahan armada, perluasan rute, peningkatan integrasi antar moda, serta edukasi dan sosialisasi kepada penumpang.
- Untuk meratakan distribusi "tap in" di Transjakarta, solusi meliputi peningkatan mesin tap in di halte, sosialisasi tentang "tap in" dan "tap out", serta integrasi sistem pembayaran seperti QRIS.

Solusi

- Untuk mengatasi masalah tap out Transjakarta, beberapa solusi yang dapat diterapkan adalah memastikan mesin berfungsi, sosialisasi aturan, sanksi bagi pelanggar, menyediakan petugas, dan menggunakan teknologi untuk pemantauan.
- Solusi untuk mengurangi durasi perjalanan Transjakarta saat jam sibuk adalah dengan meningkatkan frekuensi layanan dan memperbanyak armada bus.



Thank You!

Presentations are
communication tools.

Created by
Adi Nugraha