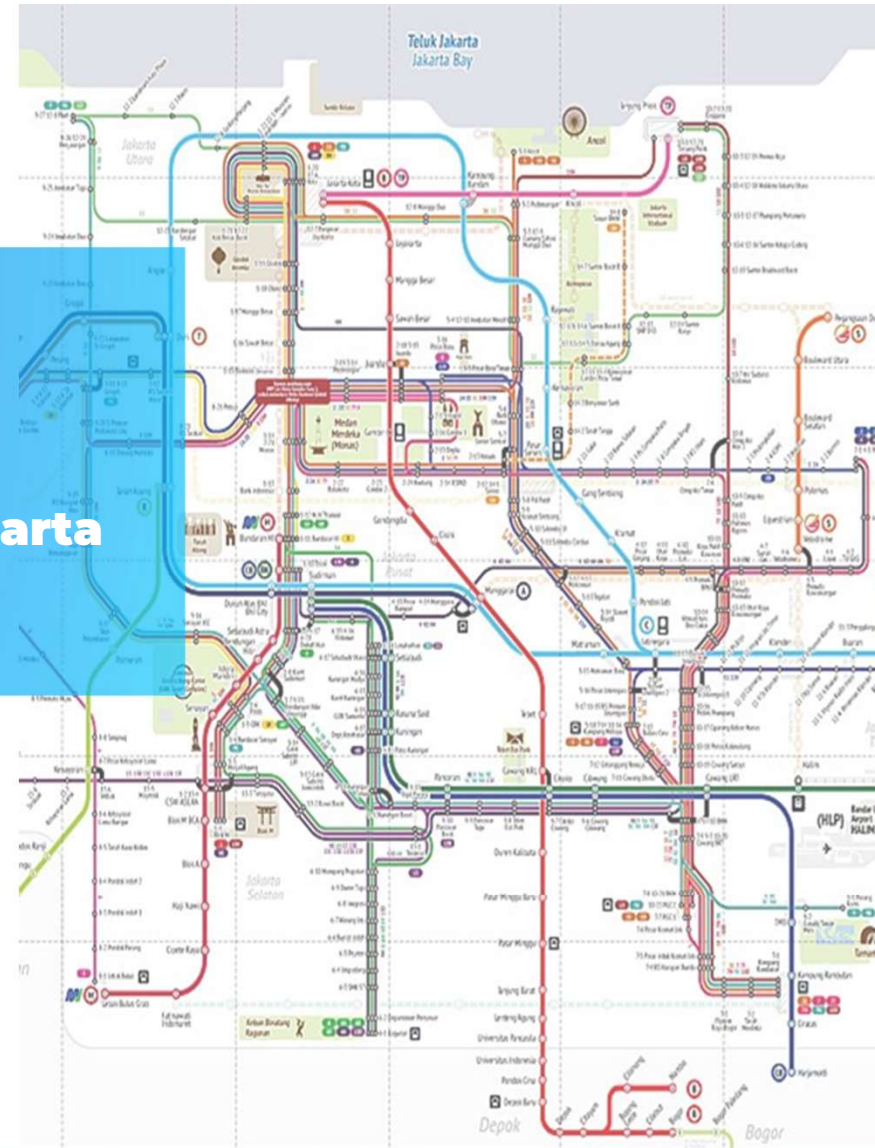


CAPSTONE PROJECT MODULE 2

Exploratory Data Analysis Public Transportation Transaction Transjakarta

RIZKI NUGRAHA - JCDS ONLINE LEARNING



OUTLINE

1

LATAR BELAKANG

2

BUSINESS PROBLEM

3

GOALS

4

BUSINESS QUESTIONS

5

DATA ANALYSIS

6

***CONCLUSION &
RECOMMENDATION***



LATAR BELAKANG

Tentang TransJakarta

TransJakarta merupakan sebuah perusahaan yang bergerak di jasa transportasi yang sudah beroperasi sejak tahun 2004 di Indonesia. Perusahaan ini dirancang sebagai moda transportasi massal pendukung aktivitas ibu kota yang sangat padat. Saat ini TransJakarta memiliki 252 halte yang tersebar dalam 14 koridor dan beroperasi 24 jam

Context - Audience

Disini audience yang diharapkan adalah manajemen TransJakarta



BUSINESS PROBLEM & GOALS

Business Problem

Bagaimana Perusahaan terus dapat meningkatkan service level terhadap pelanggan TransJakarta untuk dapat meningkatkan pendapatan Perusahaan serta pelayanan terhadap pelanggan

Goals

Perusahaan ingin meningkatkan service level terhadap pelanggan TransJakarta, baik dari sisi peningkatan layanan armada, peningkatan layanan infrastruktur pada halte-halte yang ada, meningkatkan keamanan perjalanan, ataupun penambahan rute/koridor perjalanan untuk menampung kapasitas penumpang dan meningkatkan tingkat kepuasan pelanggan terhadap TransJakarta.



BUSINESS QUESTIONS

Berdasarkan Latar Belakang di atas, sebagai seorang *data analyst* kita akan mencoba membantu Perusahaan untuk mengetahui informasi yang bisa digunakan oleh Perusahaan untuk dapat menjawab *business problem* yang ada di atas guna pengambilan keputusan kedepan berdasarkan dataset yang kita punya, maka kita akan mencoba menjawab beberapa *business question* :

1. Koridor mana saja yang menjadi koridor dengan tingkat transaksi tertinggi?
2. Koridor mana saja yang menjadi koridor dengan tingkat transaksi terendah?
3. Halte mana saja yang paling sering digunakan oleh pelanggan TransJakarta sebagai tempat Tap In dan Tap Out?
4. Kapan waktu terpadat penggunaan TransJakarta?
5. Kapan waktu terlowong penggunaan TransJakarta?
6. Siapa yang paling banyak menggunakan TransJakarta ? (by gender dan by usia)



DATASET

```
Jumlah baris dan kolom di dataset df adalah (37900, 22)
<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
...
20  tapOutTime            36556 non-null  object
21  payAmount             36893 non-null  float64
dtypes: float64(7), int64(3), object(12)
memory usage: 6.4+ MB

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

Kita memiliki dataset berjumlah:

37900 baris dan 22 kolom.

Source analysis: by Jupyter Notebook



DATASET

```
<class 'pandas.core.frame.DataFrame'>
Index: 32555 entries, 0 to 37899
Data columns (total 22 columns):
#   Column                Non-Null Count  Dtype
---  -
0   transID                32555 non-null  object
1   payCardID              32555 non-null  int64
2   payCardBank            32555 non-null  object
3   payCardName            32555 non-null  object
4   payCardSex             32555 non-null  object
5   payCardBirthDate       32555 non-null  int64
6   corridorID             32555 non-null  object
7   corridorName           32555 non-null  object
8   direction              32555 non-null  float64
9   tapInStops             32555 non-null  object
10  tapInStopsName          32555 non-null  object
11  tapInStopsLat           32555 non-null  float64
12  tapInStopsLon           32555 non-null  float64
13  stopStartSeq           32555 non-null  int64
14  tapInTime               32555 non-null  object
15  tapOutStops             32555 non-null  object
16  tapOutStopsName         32555 non-null  object
17  tapOutStopsLat          32555 non-null  float64
18  tapOutStopsLon          32555 non-null  float64
19  stopEndSeq             32555 non-null  float64
20  tapOutTime              32555 non-null  object
21  payAmount              32555 non-null  float64
dtypes: float64(7), int64(3), object(12)
memory usage: 5.7+ MB
```

Source analysis: by Jupyter Notebook

Setelah melakukan cleansing kita memiliki dataset berjumlah :

32555 baris dan 22 kolom.



DATASET

1. transID: Unique transaction id for every transaction
2. payCardID: Customers main identifier. The card customers use as a ticket for entrance and exit.
3. payCardBank: Customers card bank issuer name
4. payCardName: Customers name that is embedded in the card.
5. payCardSex: Customers sex that is embedded in the card
6. payCardBirthDate: Customers birth year
7. corridorID: Corridor ID / Route ID as key for route grouping.
8. corridorName: Corridor Name / Route Name contains Start and Finish for each route.
9. direction: 0 for Go, 1 for Back. Direction of the route.
10. tapInStops: Tap In (entrance) Stops ID for identifying stops name
11. tapInStopsName: Tap In (entrance) Stops Name where customers tap in.
12. tapInStopsLat: Latitude of Tap In Stops
13. tapInStopsLon: Longitude of Tap In Stops



DATASET

- 14. stopStartSeq: Sequence of the stops, 1st stop, 2nd stops etc. Related to direction.
- 15. tapInTime: Time of tap in. Date and time
- 16. tapOutStops: Tap Out (Exit) Stops ID for identifying stops name
- 17. tapOutStopsName: Tap out (exit) Stops Name where customers tap out.
- 18. tapOutStopsLat: Latitude of Tap Out Stops
- 19. tapOutStopsLon: Longitude of Tap Out Stops
- 20. stopEndSeq: Sequence of the stops, 1st stop, 2nd stops etc. Related to direction.
- 21. tapOutTime: Time of tap out. Date and time
- 22. payAmount: The number of what customers pay. Some are free. Some not.



DATA ANALYSIS

1. Koridor mana saja yang menjadi koridor dengan tingkat transaksi tertinggi dan terendah?

Koridor tertinggi

```
corridorName
Cibubur - Balai Kota          371
Ciputat - CSW                 355
Harmoni - Jakarta International Stadium  318
Pulo Gadung - Monas           317
Kampung Rambutan - Pondok Gede  308
Kalideres - Bundaran HI via Veteran  304
Rusun Pondok Bambu - Walikota Jakarta Timur  303
Kebayoran Lama - Tanah Abang    297
Rusun Rawa Bebek - Kodamar      289
Pasar Minggu - Tanah Abang     285
Name: count, dtype: int64
```

Koridor terendah

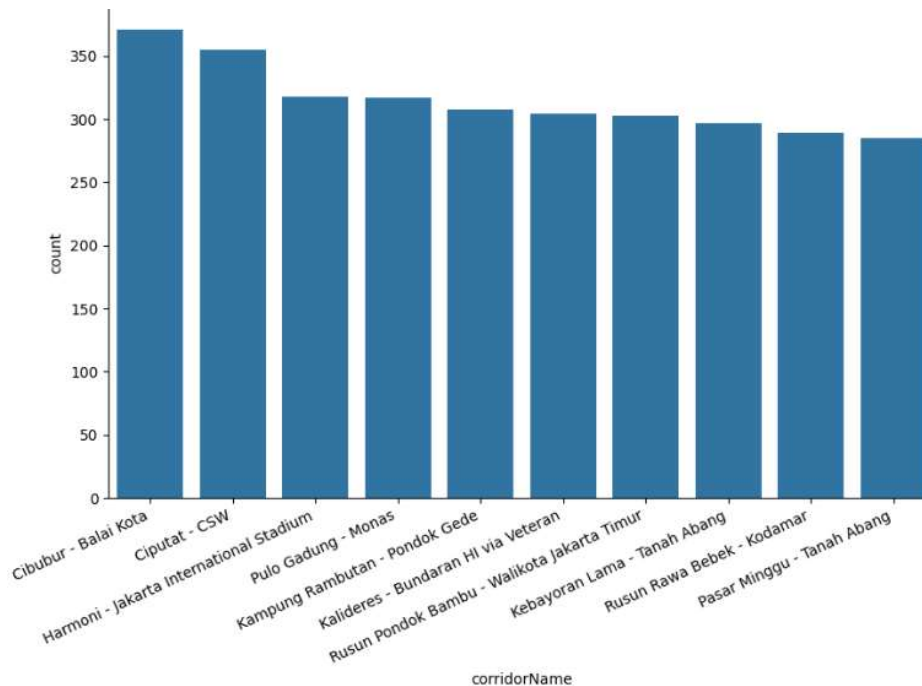
```
corridorName
Rusun Pinus Elok - Rusun Pulo Gebang    52
St. LRT Pegangsaan Dua - JIS           50
Stasiun Manggarai - Blok M             49
Pondok Kelapa - BKN                    45
Grogol - Pos Pengumben via Slipi       37
Gondangdia - Balai Kota                37
Senen - Tanah Abang                    36
Term. Pulo Gadung - Lampiri            28
Tanah Abang - Kebayoran Lama via Pos Pengumben  22
Kampung Rambutan - Blok M              17
Name: count, dtype: int64
```

Source analysis: by Jupyter Notebook



DATA ANALYSIS

Koridor tertinggi



Source analysis: by Jupyter Notebook



Dari visualisasi diatas terlihat disini bahwa Koridor Cibubur - Balai Kota menjadi koridor paling banyak transaksi yaitu sebanyak 371 transaksi disusul Ciputat - CSW sebanyak 355 transaksi kemudian Harmoni - Jakarta International Stadium sebanyak 318 transaksi.

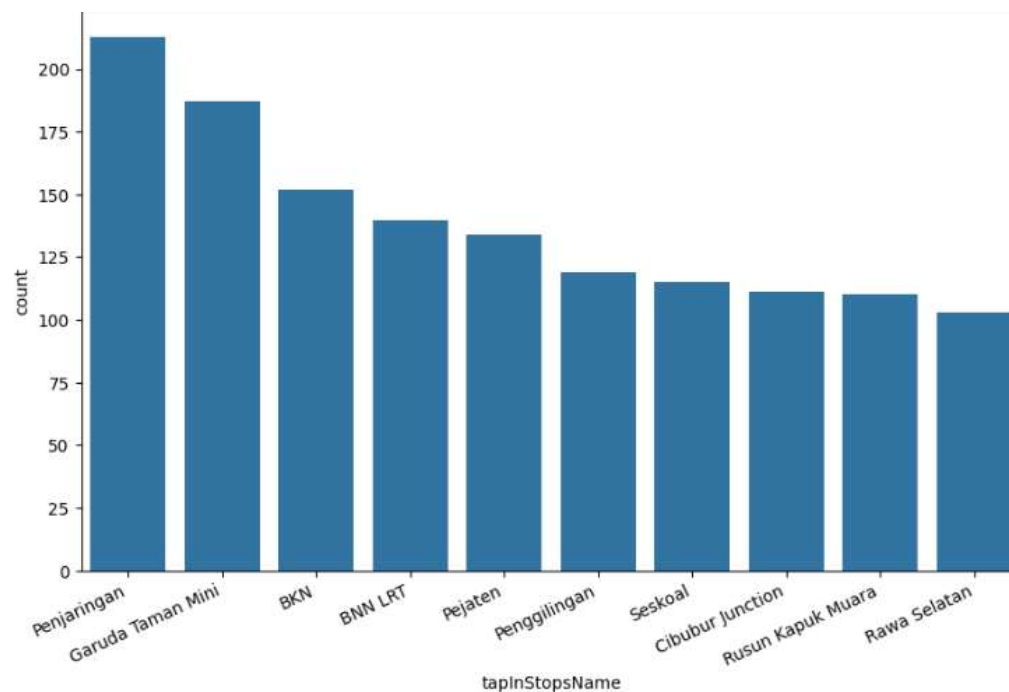
Insight:

- Armada pada koridor ini harus selalu prima agar tidak menghambat perjalanan TransJakarta dan tetap menjaga tingkat penggunaan oleh pelanggan.
- Perusahaan juga bisa melakukan perencanaan untuk jadwal perawatan berkala pada masing-masing koridor dengan melihat data di atas.

DATA ANALYSIS

2. Halte mana saja yang memiliki Tap In dan Tap Out tertinggi? (top 10)

TAP IN TRANSACTION



Source analysis: by Jupyter Notebook

- Penjaringan 213
- Garuda Taman Mini 187
- BKN 152
- BNN LRT 140
- Pejaten 134
- Penggilingan 119
- Seskoal 115
- Cibubur Junction 111
- Rusun Kapuk Muara 110
- Rawa Selatan 103

Insight:

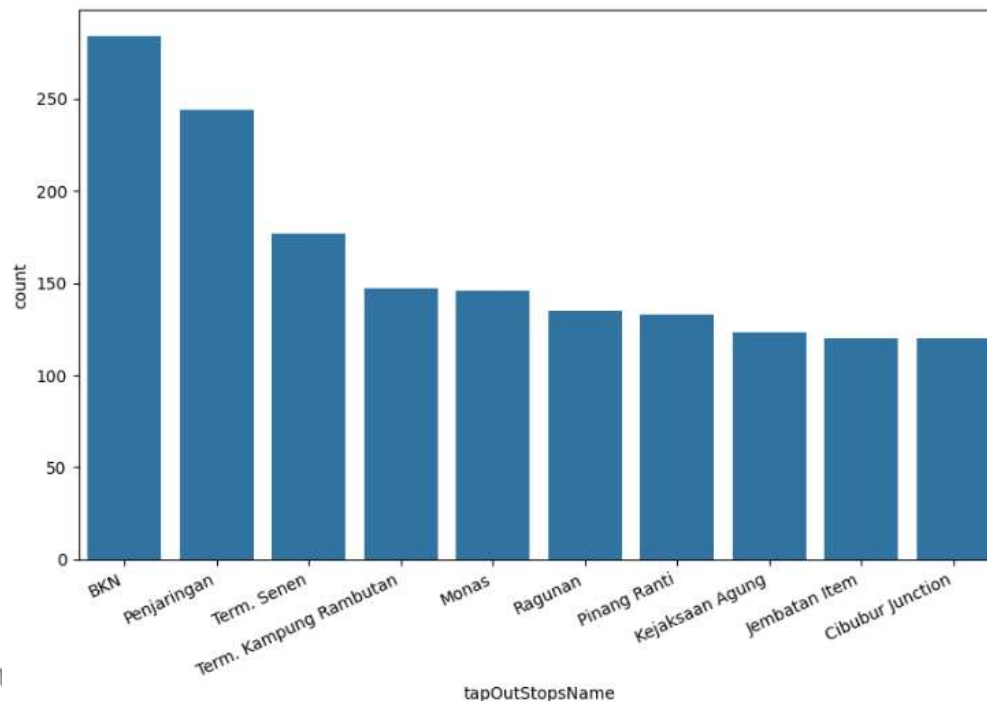
- ☐ Perusahaan harus meningkatkan pelayanan infrastruktur pada halte-halte tersebut, seperti mesin tap, monitor layar (untuk memantau kedatangan bus), tempat duduk untuk menunggu, dll.
- ☐ Tingkat keamanan juga harus ditingkatkan pada halte-halte yang memiliki intensitas penumpang yang tinggi ini



DATA ANALYSIS

2. Halte mana saja yang memiliki Tap In dan Tap Out tertinggi? (top 10)

TAP OUT TRANSACTION



Source analysis: by Jupyter Notebook

- BKN 284
- Penjaringan 244
- Term. Senen 177
- Term. Kampung Rambutan 147
- Monas 146
- Ragunan 135
- Pinang Ranti 133
- Kejaksaan Agung 123
- Jembatan Item 120
- Cibubur Junction 120

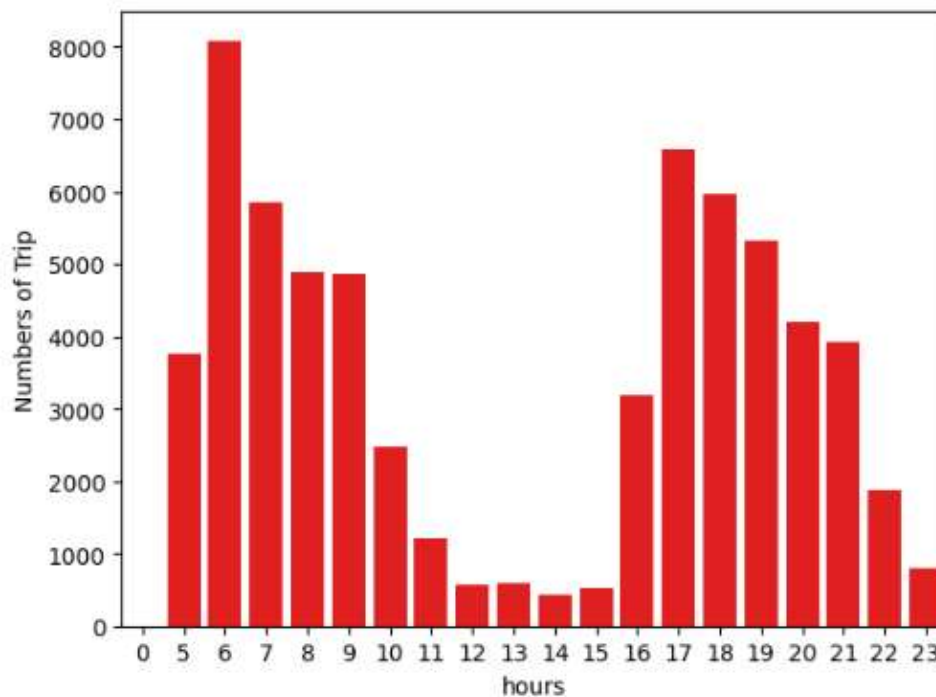
Insight:

- ☐ Perusahaan harus meningkatkan pelayanan infrastruktur pada halte-halte tersebut, seperti mesin tap, monitor layar (untuk memantau kedatangan bus), tempat duduk untuk menunggu, dll.
- ☐ Tingkat keamanan juga harus ditingkatkan pada halte-halte yang memiliki intensitas penumpang yang tinggi ini.



DATA ANALYSIS

3. Kapan waktu terpadat dan terlowong penggunaan TransJakarta?



Jam tersibuk ada pada jam 05.00-07.00 & jam 17.00-19.00,

Jam terlowong ada pada jam 12.00 - 15.00 .

Insight:

Perusahaan harus memastikan ketersediaan armada yang memadai pada jam-jam tersebut serta meningkatkan keamanan di halte dan di armada

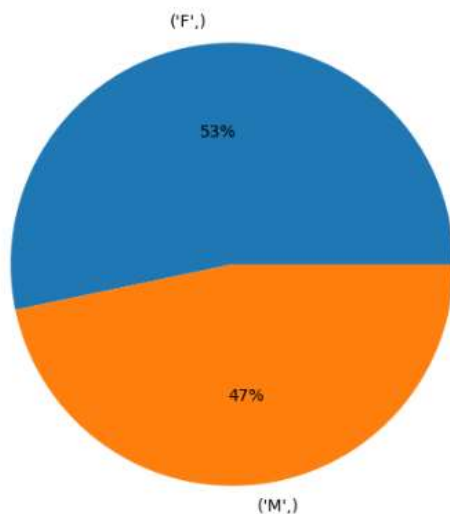
Source analysis: by Jupyter Notebook



DATA ANALYSIS

4. Siapa yang paling banyak menggunakan TransJakarta ? (by gender dan by usia)

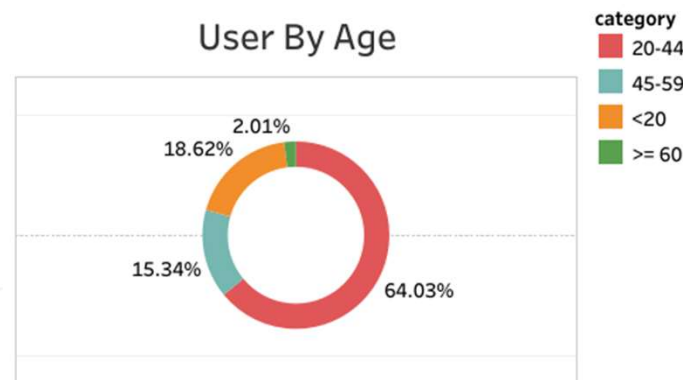
User By Gender



Source analysis: by Jupyter Notebook



User By Age



Source analysis: by tableau

Mayoritas pengguna TransJakarta adalah Perempuan sebesar 53% (17.350 transaksi) dan Laki-laki 47% (15.205 transaksi) dan berada pada rentang umur 20-44 tahun.

Insight:

Perusahaan harus meningkatkan tingkat keamanan baik di halte maupun di dalam armada, bisa dengan menempatkan petugas keamanan di tiap-tiap armada yang beroperasi

CONCLUSION & RECOMENDATION

KESIMPULAN

- Koridor Cibubur - Balai Kota menjadi koridor paling banyak transaksi yaitu sebanyak 371 transaksi disusul Ciputat - CSW sebanyak 355 transaksi kemudian Harmoni - Jakarta International Stadium sebanyak 318 transaksi.
- Sedangkan, Koridor Kampung Rambutan - Blok M menjadi koridor paling sedikit transaksi yaitu sebanyak 17 transaksi disusul Tanah Abang - Kebayoran Lama via Pos Pengumben 22 transaksi dan Term. Pulo Gadung - Lampiri 28 transaksi.
- Halte dengan Tap In dan Tap Out terbanyak ada pada Halte Penjaringan 452 total transaksi (213 Tap In & 244 Tap Out) dan Halte BKN dengan total transaksi 436 (152 Tap In dan 284 Tap Out)
- Jam tersibuk ada pada jam 05.00-07.00 dan jam 17.00-19.00, sedangkan waktu terlowong ada di jam 12.00 - 15.00.
- Mayoritas pengguna TransJakarta adalah Perempuan sebesar 53% (17.350 transaksi) dan Laki-laki 47% (15.205 transaksi).



CONCLUSION & RECOMENDATION

REKOMENDASI

- Mengoperasikan lebih banyak armada atau dilakukan penambahan armada pada koridor-koridor yang memiliki tingkat transaksi terbanyak.
- Mengalihkan sebagian armada pada koridor dengan intensitas transaksi yang tergolong rendah ke koridor-koridor yang memiliki intensitas transaksi tinggi.
- Melakukan peningkatan infrastruktur atau revitalisasi pada halte-halte dengan tingkat kepadatan yang tinggi guna menghindari antiran atau penumpukan penumpang serta tetap bisa menjaga kenyamanan pengguna.
- Meningkatkan tingkat keamanan baik pada halte maupun di dalam armada. Hal ini dikarenakan ada beberapa halte yang padat dan juga sesuai hasil analisa dilihat mayoritas pengguna berjenis kelamin perempuan.



A background image showing a group of people in business attire huddled together with their hands clasped in the center, symbolizing teamwork and success. The image is partially obscured by a blue gradient box on the right.

Thanks!