

Perancangan dan Analisa Visualisasi Data Penumpang Transjakarta dengan Metode Four-Step Kimball menggunakan Power BI

Aldis Tamara Putri Iskandar, Teddy Siswanto, Syandra Sari

Jurusan Sistem Informasi, Fakultas Teknologi Industri

Universitas Trisakti

Jakarta, Indonesia

065002100021@std.trisakti.ac.id, teddysiswanto@trisakti.ac.id, syandra_sari@trisakti.ac.id

Abstract-The Special Capital Region (DKI) of Jakarta, as Indonesia's economic hub, faces significant challenges in managing public transportation due to high population mobility. Transjakarta, a Bus Rapid Transit (BRT) system, plays a crucial role in addressing the transportation needs of Jakarta and its surrounding areas. This study aims to address the complexity of managing Transjakarta passenger data by applying the Four-step Kimball method in building a data warehouse and visualizing the data using Power BI. Utilizing passenger data from 2019 to 2023, the research employed the Extract, Transform, and Load (ETL) process through Pentaho Data Integration to develop an OLAP database based on the Star Schema. The visualizations include passenger trends, distribution by route type, and route performance. Additionally, the ARIMA prediction method is applied to forecast passenger numbers, supporting strategic planning such as fleet allocation and route development. The findings provide a data-driven solution to assist Transjakarta management in improving service quality and operational efficiency in Jakarta's public transportation system.

Keywords: Transjakarta, visualisasi data, Four-step Kimball, Power BI, OLAP, ARIMA.

Abstrak-Daerah Khusus Ibukota (DKI) Jakarta sebagai pusat perekonomian Indonesia menghadapi tantangan besar dalam mengelola transportasi publik akibat tingginya mobilitas masyarakat. Transjakarta, sebagai moda transportasi *Bus Rapid Transit* (BRT), memiliki peran penting dalam memenuhi kebutuhan transportasi warga Jakarta dan sekitarnya. Penelitian ini bertujuan untuk mengatasi kompleksitas pengelolaan data penumpang Transjakarta melalui penerapan metode *Four-step Kimball* dalam pembangunan *data warehouse* serta visualisasi menggunakan Power BI. Dengan data penumpang dari tahun 2019 hingga 2023, penelitian ini menggunakan proses *Extract, Transform, and Load* (ETL) melalui Pentaho Data Integration untuk menghasilkan basis data OLAP berbasis *Star Schema*. Visualisasi yang dihasilkan mencakup tren jumlah penumpang, distribusi penumpang per jenis trayek, serta performa trayek tertentu. Selain itu, metode prediksi ARIMA digunakan untuk memperkirakan jumlah penumpang, mendukung perencanaan strategis seperti alokasi armada dan pengembangan rute. Hasil penelitian ini memberikan solusi berbasis data yang membantu manajemen Transjakarta dalam meningkatkan kualitas pelayanan dan efisiensi operasional transportasi publik di Jakarta.

Kata Kunci: Transjakarta, data visualization, *Four-step Kimball*, Power BI, OLAP, ARIMA.

1. Pendahuluan

DKI Jakarta, sebagai pusat perekonomian Indonesia, memiliki tingkat kepadatan penduduk yang sangat tinggi dan perkembangan yang pesat, yang mendorong kemajuan sektor transportasi, termasuk Transjakarta [1][12]. Transjakarta, yang beroperasi sejak 2004, adalah moda transportasi berbasis *Bus Rapid Transit* (BRT) dan non-BRT, melayani wilayah Jakarta dan kawasan Bodetabek, dengan jangkauan 251,2 km dan 273 halte. Pada 2023, Transjakarta mencatatkan rekor jumlah penumpang harian hingga 1,17 juta, dengan armada yang terus berkembang.

Untuk menghadapi tantangan dalam mengelola data penumpang yang semakin kompleks, Transjakarta

memerlukan sistem visualisasi yang mampu menyajikan informasi mengenai persebaran penumpang berdasarkan trayek, waktu, dan hari perjalanan. Hal ini akan membantu tim operasional dan manajemen dalam menganalisis pola penggunaan layanan, mengidentifikasi trayek yang padat, dan merencanakan peningkatan kapasitas layanan di masa depan.

Metode *Four-step Kimball* digunakan untuk mempermudah analisis data penumpang, dengan langkah-langkah mulai dari pemilihan proses, identifikasi fakta, hingga perancangan skema data yang terstruktur. Proses ETL (*Extract, Transform, Load*) dilakukan menggunakan Pentaho Data Integration, yang

memungkinkan pengelolaan data kompleks menjadi lebih terstruktur dan siap dianalisis. Hasil analisis ini kemudian divisualisasikan melalui Power BI untuk mempermudah pemahaman dan pengambilan keputusan.

Penerapan analisis prediksi jumlah penumpang dengan metode ARIMA dapat membantu tim manajemen merencanakan strategi operasional yang lebih proaktif. Prediksi lonjakan penumpang memungkinkan penambahan armada bus atau peningkatan frekuensi keberangkatan di koridor tertentu. Dengan *dashboard* Power BI yang interaktif, hasil prediksi ini dapat divisualisasikan, mempermudah manajemen dalam melihat tren dan mengambil keputusan cepat untuk meningkatkan kualitas layanan Transjakarta.

2. Tinjauan Pustaka

A. Data Warehouse

Data warehouse adalah kumpulan data yang bersifat subject-oriented, terintegrasi, bervariasi dari waktu ke waktu, dan non-volatile, yang dapat mendukung proses pengambilan keputusan manajemen [5][11].

B. ETL

Secara umum, proses ETL melibatkan pengambilan data dari berbagai sumber, membersihkan data agar memiliki format yang konsisten, dan kemudian menyimpannya sebagai sumber data baru di data warehouse [10][4].

C. OLAP

OLAP (*Online Analytical Processing*) adalah teknologi analisis data multidimensi untuk mendukung pengambilan keputusan strategis, berbeda dengan OLTP yang fokus pada transaksi. OLAP digunakan untuk analisis kompleks melalui operasi seperti *roll-up*, *drill-down*, dan *slicing*[6].

D. Four-step Design Process Methodology Kimball

Metode Kimball adalah pendekatan pengembangan Data Warehouse yang menggunakan pemodelan dimensi dengan proses *bottom-up*. Metode ini terdiri dari empat langkah: Dalam perancangan *data warehouse* OLAP ini, digunakan metodologi desain Kimball yang terdiri dari empat langkah utama, yaitu:[5]

1. *Identify the business process*
2. *Declare the grain*
3. *Identify the dimensions table*
4. *Identify the facts*

E. Pentaho Data Integration

Pentaho Data Integration (PDI) adalah alat ETL untuk mengumpulkan, membersihkan, dan memuat data ke sistem target. Dengan antarmuka grafis dan fitur drag-and-drop, PDI memudahkan integrasi data dalam strategi Business Intelligence [5].

F. Analisis Time Series

Analisis *Time Series* mempelajari pola data dalam rentang waktu tertentu untuk memprediksi tren, pola musiman, dan fluktuasi. Dalam bisnis, ini membantu memproyeksikan kebutuhan pasar dan meramalkan pendapatan. Model seperti ARIMA dan SARIMA menggunakan data historis untuk prediksi, menjadikannya alat efektif dalam perencanaan bisnis berbasis data [18].

G. ARIMA

ARIMA (*Autoregressive Integrated Moving Average*) adalah metode analisis time series untuk data dengan tren dan pola musiman sederhana. Digunakan untuk peramalan penjualan, analisis risiko, dan prediksi jumlah penumpang, dengan parameter optimal (p , d , q) ditentukan melalui analisis ACF dan PACF [18].

H. SARIMA

SARIMA (*Seasonal ARIMA*) adalah pengembangan ARIMA yang menambahkan komponen musiman (*Seasonal AR*, *Seasonal Differencing*, dan *Seasonal MA*) dengan parameter P , D , dan Q , serta periode musiman (s). SARIMA cocok untuk menangkap pola musiman seperti fluktuasi jumlah penumpang [18].

I. Power BI

Power BI dirancang untuk mengubah dan menganalisis data atau informasi yang tidak terhubung menjadi wawasan yang lebih terintegrasi, sehingga dapat digunakan untuk analisis data dan pelaporan [7].

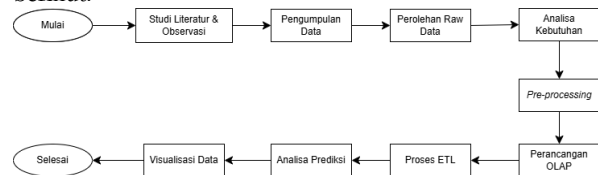
J. Visualisasi Data

Visualisasi data adalah proses menyajikan data dalam bentuk grafis atau gambar untuk membantu pemahaman dan analisis informasi yang terkandung di dalamnya. Tujuan dari visualisasi data adalah untuk mengungkap pola, tren, dan wawasan dari data yang mungkin sulit diidentifikasi hanya dengan melihat angka atau tabel [9].

3. Metodologi

A. Metodologi Penelitian

Pada tahap ini dijelaskan mengenai teknik analisis untuk penelitian ini yaitu menggunakan metode perancangan *data warehouse* dengan pendekatan *Four-step Kimball* untuk menghasilkan visualisasi data penumpang Transjakarta secara interaktif menggunakan Power BI. Metodologi ini melibatkan langkah-langkah sebagai berikut:



Gambar 1. Alur Penelitian

B. Studi Literatur & Observasi

Tahapan ini mencakup penelusuran referensi teori seperti jurnal, buku, dan artikel ilmiah yang mendukung penelitian. Referensi tersebut memperkuat landasan teori terkait visualisasi data penumpang Transjakarta dengan metode *Four-step Kimball*, serta memberikan wawasan tentang penerapan metode ini dalam analisis data transportasi umum, khususnya untuk layanan Transjakarta.

A. Perolehan Raw Data

Data raw yang digunakan dalam penelitian ini diperoleh dari PT Transportasi Jakarta, mencakup informasi penumpang Transjakarta dari 2020 hingga 2024. Data ini akan diolah menjadi *database* OLAP dengan metode *Four-step Kimball*.

B. Analisa Kebutuhan

Transjakarta memerlukan penerapan *Business Intelligence* untuk menganalisis dan memvisualisasikan

perbandingan jumlah penumpang serta trayek dari 2020 hingga 2024, berdasarkan data bulanan. Data yang saat ini ada dalam bentuk *spreadsheet* Excel akan dipindahkan ke *database* untuk memudahkan analisis. Setelah itu, data akan diolah dan divisualisasikan menggunakan Power BI untuk memperjelas perbandingan jumlah penumpang dan trayek.

C. Pre-processing

Pada proses *pre-processing*, langkah pertama yang dilakukan adalah penggabungan dua dataset yang berbeda, yaitu data penumpang untuk periode 2020-2023 dan data penumpang untuk tahun 2024, menjadi satu *dataset* utuh menggunakan *pd.concat*. Setelah itu, dilakukan penyaringan data dengan menghapus baris yang memiliki jumlah penumpang bernilai 0, karena dianggap tidak relevan atau tidak valid untuk analisis lebih lanjut. Terakhir, *dataset* yang telah difilter tersebut disimpan dalam file CSV menggunakan *to_csv* untuk digunakan dalam analisis selanjutnya.

D. Perancangan OLAP

Dalam perancangan *data warehouse* OLAP ini, digunakan metodologi desain Kimball yang terdiri dari empat langkah utama, yaitu:

1. *Identify the business process*
2. *Declare the grain*
3. *Identify the dimensions table*
4. *Identify the facts*

E. Proses ETL

Pada tahap ETL, data yang telah dinormalisasi akan diolah menggunakan Spoon dari Pentaho Data Integration dan MySQL di XAMPP.

1. *Extract*: Mengambil data penumpang dan trayek dari file CSV di MySQL (XAMPP) ke dalam Spoon.
2. *Transform*: Mengubah data mentah menjadi terstruktur dengan pembersihan, modifikasi, dan perubahan tipe data untuk memudahkan analisis.
3. *Load*: Menyimpan hasil transformasi ke dalam *database* OLAP melalui koneksi JDBC ke MySQL.

F. Analisa Prediksi

Analisa prediksi dilakukan dengan memproses data, mengubah kolom tanggal menjadi indeks, dan memvisualisasikan pola awal. Stasioneritas diuji menggunakan *Dickey-Fuller Test*, dengan differencing jika diperlukan. Model ARIMA, SARIMA, dan SARIMAX digunakan untuk memprediksi jumlah penumpang, diikuti visualisasi hasil. Prediksi digabungkan dengan data asli dan disimpan dalam file Excel untuk analisis lebih lanjut di Power BI.

G. Pembuatan Visualisasi

Pada tahap ini, data penumpang dan trayek Transjakarta divisualisasikan menggunakan Power BI, menampilkan grafik perbandingan bulanan, top 5 penumpang terbanyak dan terendah berdasarkan trayek, serta pie chart penumpang berdasarkan jenis trayek. Fungsi slicer memungkinkan pemilihan grafik berdasarkan tahun, jenis, dan trayek. Menggunakan model *Star Schema*, visualisasi ini memudahkan analisis data yang dinamis, mendukung pengambilan keputusan, mempermudah input data, dan memfasilitasi perbandingan secara efisien.

4. Hasil dan Pembahasan

A. Perolehan Raw Data

Raw data yang digunakan pada penelitian ini adalah raw data yang diperoleh dari PT. Transportasi Jakarta yaitu data penumpang Transjakarta dari bulan Januari 2020 sampai dengan bulan Desember 2024 dan data rute Transjakarta yang berbentuk file Microsoft Excel. Data penumpang berisi kolom tahun, bulan, kode trayek, trayek, jenis, dan jumlah penumpang, sedangkan data rute berisi kolom jenis, kode trayek, dan trayek.

Tahun	Bulan	Kode Trayek	Trayek	Jenis	Jumlah Penumpang
2020	1	101	101	101	101
2020	2	102	102	102	102
2020	3	103	103	103	103
2020	4	104	104	104	104
2020	5	105	105	105	105
2020	6	106	106	106	106
2020	7	107	107	107	107
2020	8	108	108	108	108
2020	9	109	109	109	109
2020	10	110	110	110	110
2020	11	111	111	111	111
2020	12	112	112	112	112
2021	1	113	113	113	113
2021	2	114	114	114	114
2021	3	115	115	115	115
2021	4	116	116	116	116
2021	5	117	117	117	117
2021	6	118	118	118	118
2021	7	119	119	119	119
2021	8	120	120	120	120
2021	9	121	121	121	121
2021	10	122	122	122	122
2021	11	123	123	123	123
2021	12	124	124	124	124
2022	1	125	125	125	125
2022	2	126	126	126	126
2022	3	127	127	127	127
2022	4	128	128	128	128
2022	5	129	129	129	129
2022	6	130	130	130	130
2022	7	131	131	131	131
2022	8	132	132	132	132
2022	9	133	133	133	133
2022	10	134	134	134	134
2022	11	135	135	135	135
2022	12	136	136	136	136
2023	1	137	137	137	137
2023	2	138	138	138	138
2023	3	139	139	139	139
2023	4	140	140	140	140
2023	5	141	141	141	141
2023	6	142	142	142	142
2023	7	143	143	143	143
2023	8	144	144	144	144
2023	9	145	145	145	145
2023	10	146	146	146	146
2023	11	147	147	147	147
2023	12	148	148	148	148
2024	1	149	149	149	149
2024	2	150	150	150	150
2024	3	151	151	151	151
2024	4	152	152	152	152
2024	5	153	153	153	153
2024	6	154	154	154	154
2024	7	155	155	155	155
2024	8	156	156	156	156
2024	9	157	157	157	157
2024	10	158	158	158	158
2024	11	159	159	159	159
2024	12	160	160	160	160

Gambar 2. Raw Data Penumpang

Jenis	Kode Trayek	Trayek
101	101	101
102	102	102
103	103	103
104	104	104
105	105	105
106	106	106
107	107	107
108	108	108
109	109	109
110	110	110
111	111	111
112	112	112
113	113	113
114	114	114
115	115	115
116	116	116
117	117	117
118	118	118
119	119	119
120	120	120
121	121	121
122	122	122
123	123	123
124	124	124
125	125	125
126	126	126
127	127	127
128	128	128
129	129	129
130	130	130
131	131	131
132	132	132
133	133	133
134	134	134
135	135	135
136	136	136
137	137	137
138	138	138
139	139	139
140	140	140
141	141	141
142	142	142
143	143	143
144	144	144
145	145	145
146	146	146
147	147	147
148	148	148
149	149	149
150	150	150
151	151	151
152	152	152
153	153	153
154	154	154
155	155	155
156	156	156
157	157	157
158	158	158
159	159	159
160	160	160

Gambar 3. Raw Data Rute

B. Pre-processing

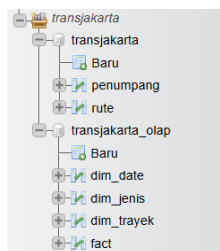
Pada proses ini dilakukan *pre-processing*, atau penggabungan data jumlah penumpang Transjakarta dari tahun 2020-2024 dengan tahun 2024 menggunakan python pada tools Jupyter Notebook dan menghasilkan file berupa CSV.

Tahun	Bulan	Kode Trayek	Trayek	Jenis	Jumlah Penumpang
2020	1	101	101	101	101
2020	2	102	102	102	102
2020	3	103	103	103	103
2020	4	104	104	104	104
2020	5	105	105	105	105
2020	6	106	106	106	106
2020	7	107	107	107	107
2020	8	108	108	108	108
2020	9	109	109	109	109
2020	10	110	110	110	110
2020	11	111	111	111	111
2020	12	112	112	112	112
2021	1	113	113	113	113
2021	2	114	114	114	114
2021	3	115	115	115	115
2021	4	116	116	116	116
2021	5	117	117	117	117
2021	6	118	118	118	118
2021	7	119	119	119	119
2021	8	120	120	120	120
2021	9	121	121	121	121
2021	10	122	122	122	122
2021	11	123	123	123	123
2021	12	124	124	124	124
2022	1	125	125	125	125
2022	2	126	126	126	126
2022	3	127	127	127	127
2022	4	128	128	128	128
2022	5	129	129	129	129
2022	6	130	130	130	130
2022	7	131	131	131	131
2022	8	132	132	132	132
2022	9	133	133	133	133
2022	10	134	134	134	134
2022	11	135	135	135	135
2022	12	136	136	136	136
2023	1	137	137	137	137
2023	2	138	138	138	138
2023	3	139	139	139	139
2023	4	140	140	140	140
2023	5	141	141	141	141
2023	6	142	142	142	142
2023	7	143	143	143	143
2023	8	144	144	144	144
2023	9	145	145	145	145
2023	10	146	146	146	146
2023	11	147	147	147	147
2023	12	148	148	148	148
2024	1	149	149	149	149
2024	2	150	150	150	150
2024	3	151	151	151	151
2024	4	152	152	152	152
2024	5	153	153	153	153
2024	6	154	154	154	154
2024	7	155	155	155	155
2024	8	156	156	156	156
2024	9	157	157	157	157
2024	10	158	158	158	158
2024	11	159	159	159	159
2024	12	160	160	160	160

Gambar 4. Pre-processing

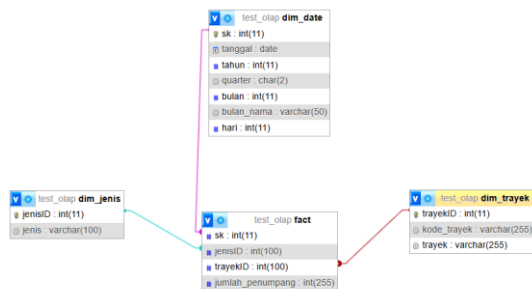
C. Proses ETL Database OLAP

Setelah melewati tahap *pre-processing*, selanjutnya yaitu melakukan tahap ETL (*Extract, Transform, dan Load*) untuk membuat *database* OLAP. Pada tahap ini menggunakan tools Pentaho Data Integration untuk melakukan proses ETL dan phpMyAdmin di Xampp untuk menyimpan data



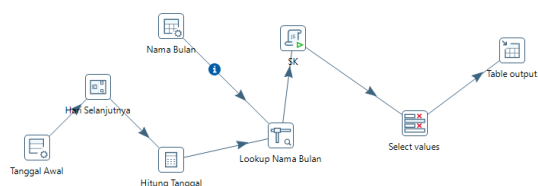
Gambar 5. Database

Model dimensional yang sesuai untuk database OLAP ini, berdasarkan data dan tabel yang tersedia, adalah skema bintang atau *star schema*.

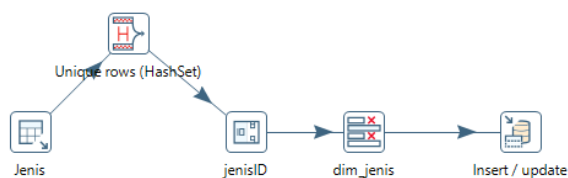


Gambar 6. Star Schema

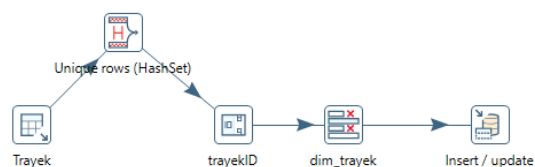
Setelah pembuatan database selesai, langkah selanjutnya adalah melakukan proses ETL (*Extract, Transform, dan Load*) untuk membangun tabel dimensi dan fakta.



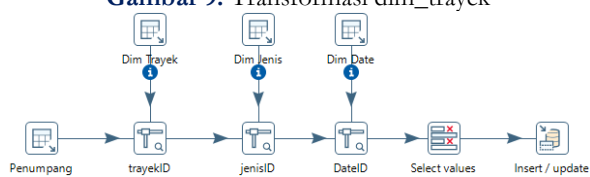
Gambar 7. Transformasi dim_date



Gambar 8. Transformasi dim_jenis



Gambar 9. Transformasi dim_trayek



Gambar 10. Transformasi fact

D. Analisa Prediksi

Pada tahap analisis statistik dan pra-pemrosesan, data dirangkum menggunakan nilai rata-rata, median, deviasi standar, dan statistik lainnya. Visualisasi data dilakukan untuk mengidentifikasi tren, musiman, dan anomali. Uji *Augmented Dickey-Fuller* (ADF) diterapkan untuk menguji apakah data bersifat stasioner atau tidak. Jika data tidak stasioner, proses differencing diterapkan untuk mencapai stasioneritas.

Results of Dickey-Fuller Test:

Test Statistic -9.473117e+00
p-value 4.056047e-16
#Lags Used 0.000000e+00
Number of Observations Used 4.600000e+01
Critical Value (1%) -3.581258e+00
Critical Value (5%) -2.926785e+00
Critical Value (10%) -2.601541e+00
dtype: float64
result : Stationary

Gambar 11. Uji Stationeritas

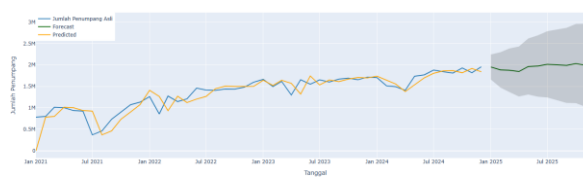
Pemodelan dilakukan dengan menggunakan pendekatan ARIMA standar melalui fungsi `auto_arima` untuk menentukan parameter optimal (p, d, q) secara otomatis.

Best model: ARIMA(0,1,0)(1,0,0)[12]
Total fit time: 3.615 seconds

SARIMAX Results						
Dep. Variable:	y	No. Observations:	48			
Model:	SARIMAX(0, 1, 0)x(1, 0, 0, 12)	Log Likelihood	-631.617			
Date:	Tue, 28 Jan 2025	AIC	1267.233			
Time:	07:27:24	BIC	1270.934			
Sample:	01-01-2021 - 12-01-2024	HQIC	1268.626			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.S.L12	0.3542	0.102	3.486	0.000	0.155	0.553
sigma2	2.213e+10	1.0e-12	1.23e+22	0.000	2.21e+10	2.21e+10
Ljung-Box (L1) (Q):	1.28	Jarque-Bera (JB):	10.20			
Prob(Q):	0.26	Prob(JB):	0.01			
Heteroskedasticity (H):	0.16	Skew:	-0.71			
Prob(H) (two-sided):	0.00	Kurtosis:	4.78			

Gambar 12. Menentukan model SARIMA

Model ini diikuti oleh prediksi jumlah penumpang untuk beberapa periode ke depan dengan menampilkan hasil prediksi beserta *confidence interval* dalam visualisasi interaktif.



Gambar 13. Visualisasi Forecast

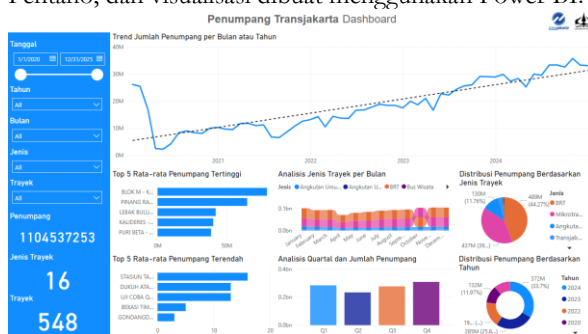
Setelah memodelkan dan memprediksi jumlah penumpang menggunakan ARIMA dan SARIMA, hasil prediksi digabungkan dengan data asli untuk melihat perbandingan.

2024-06-01	1768855.0	NaN	NaN
2024-07-01	1880390.0	NaN	NaN
2024-08-01	1841833.0	NaN	NaN
2024-09-01	1811551.0	NaN	NaN
2024-10-01	1930318.0	NaN	NaN
2024-11-01	1819079.0	NaN	NaN
2024-12-01	1955703.0	NaN	NaN
2025-01-01	NaN	1.925489e+06	1.951691e+06
2025-02-01	NaN	1.932171e+06	1.883738e+06
2025-03-01	NaN	1.930693e+06	1.876709e+06
2025-04-01	NaN	1.931020e+06	1.846904e+06
2025-05-01	NaN	1.930948e+06	1.963159e+06
2025-06-01	NaN	1.930964e+06	1.974585e+06
2025-07-01	NaN	1.930960e+06	2.014089e+06
2025-08-01	NaN	1.930961e+06	2.000433e+06
2025-09-01	NaN	1.930961e+06	1.989707e+06
2025-10-01	NaN	1.930961e+06	2.031773e+06
2025-11-01	NaN	1.930961e+06	1.992374e+06
2025-12-01	NaN	1.930961e+06	2.040764e+06

Gambar 14. Hasil Prediksi Data

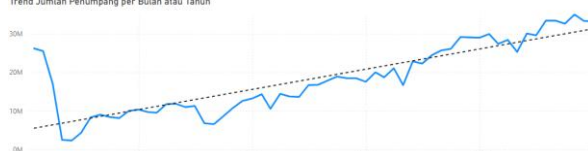
E. Visualisasi Data

Pada tahap ini dilakukan visualisasi *dashboard* dari hasil data yang telah melalui proses ETL menggunakan tools Pentaho, dan visualisasi dibuat menggunakan Power BI.



Gambar 15. Dashboard Penumpang Transjakarta

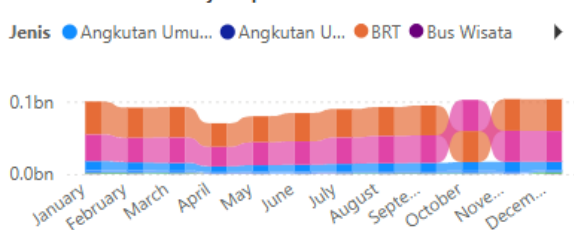
Gambar 16 menampilkan tren jumlah penumpang Transjakarta dari tahun ke tahun dalam bentuk line chart.



Gambar 16. Tren Jumlah Penumpang

Gambar 17 ini *ribbon chart*, ditampilkan distribusi jumlah penumpang berdasarkan jenis layanan, seperti Angkutan Umum, BRT, Bus Wisata, dan Mikrotrans.

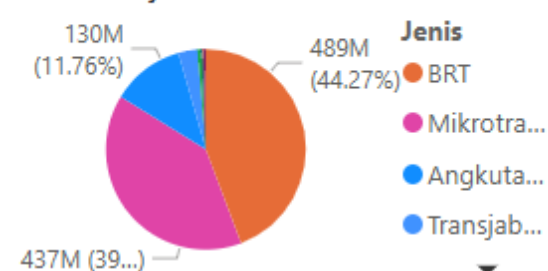
Analisis Jenis Trayek per Bulan



Gambar 17. Analisis Jenis Trayek per Bulan

Gambar 18 menggunakan *pie chart* untuk menampilkan persentase distribusi penumpang berdasarkan jenis layanan.

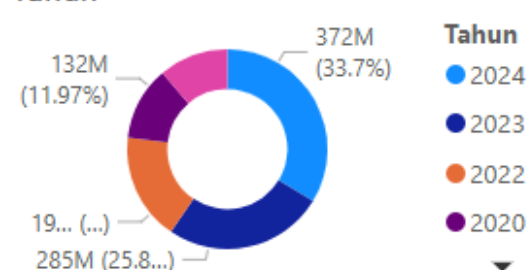
Distribusi Penumpang Berdasarkan Jenis Trayek



Gambar 18. Distribusi Penumpang Per Jenis Trayek

Gambar 19 menggunakan *pie chart* untuk menampilkan persentase distribusi penumpang berdasarkan tahun.

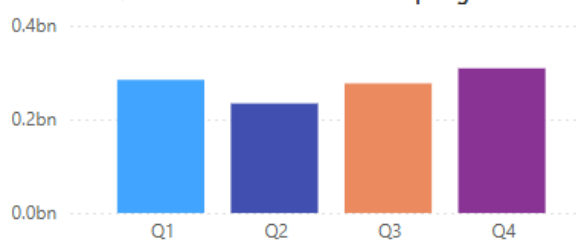
Distribusi Penumpang Berdasarkan Tahun



Gambar 19. Distribusi Penumpang Per Tahun

Gambar 20 ini menggunakan *stacked column chart*, jumlah penumpang dibagi berdasarkan kuartal (Q1, Q2, Q3, dan Q4).

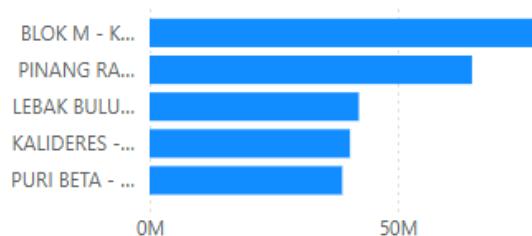
Analisis Quartal dan Jumlah Penumpang



Gambar 20. Analisis Quartal dan Jumlah Penumpang

Gambar 21 ini menggunakan *clustered bar chart* untuk menampilkan lima trayek dengan rata-rata jumlah penumpang tertinggi.

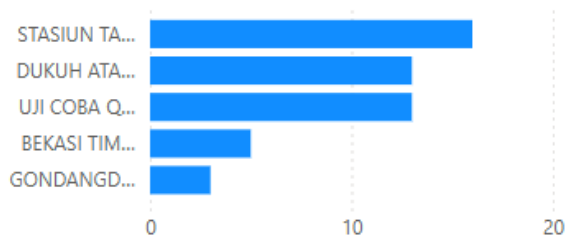
Top 5 Rata-rata Penumpang Tertinggi



Gambar 21. Top 5 Rata-rata Penumpang Tertinggi

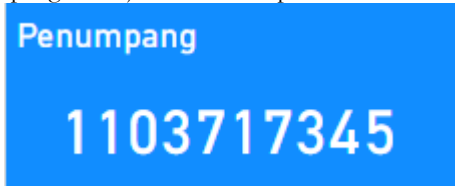
Gambar 22 ini juga menggunakan *clustered bar chart* untuk menampilkan lima trayek dengan rata-rata jumlah penumpang terendah.

Top 5 Rata-rata Penumpang Terendah



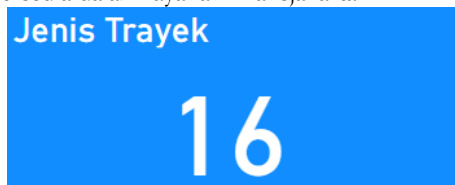
Gambar 22. Top 5 Rata-rata Penumpang Terendah

Gambar 23 ini menunjukkan total akumulasi jumlah penumpang Transjakarta selama periode.



Gambar 23. Card Jumlah Penumpang

Gambar 24 ini menampilkan jumlah total jenis trayek yang tersedia dalam layanan Transjakarta.

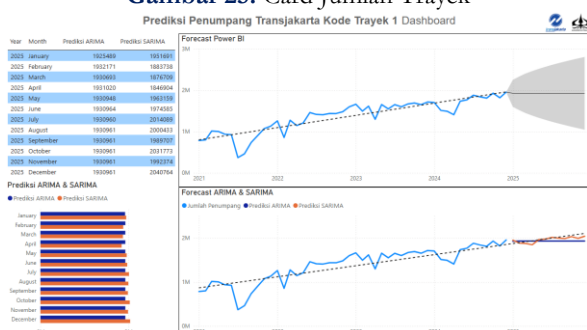


Gambar 24. Card Jumlah Jenis Trayek

Gambar 25 pada card ini menunjukkan total jumlah trayek atau rute yang beroperasi selama periode tertentu.

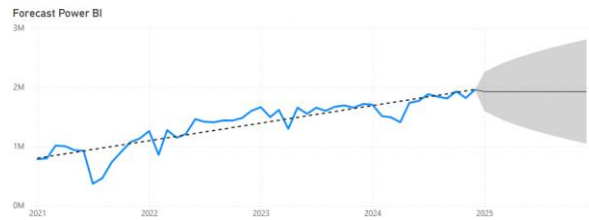


Gambar 25. Card Jumlah Trayek



Gambar 26. Dashboard Prediksi Penumpang Transjakarta kode trayek 1 "Blok M - Kota"

Gambar 27 menampilkan prediksi jumlah penumpang yang dihasilkan oleh algoritma *forecasting* bawaan Power BI.



Gambar 27. Forecast Power BI

Gambar 28 membandingkan prediksi antara model ARIMA dan SARIMA dengan menggunakan *line chart* dan garis tren.



Gambar 28. Forecast ARIMA & SARIMA

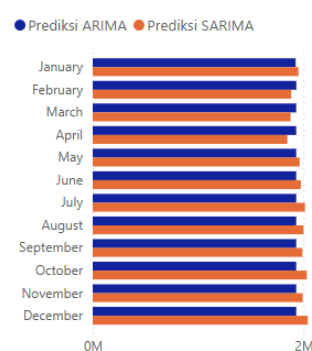
Gambar 29 menyajikan hasil prediksi jumlah penumpang Transjakarta kode trayek 1 "Blok M - Kota" untuk tahun 2025 menggunakan model ARIMA dan SARIMA.

Year	Month	Prediksi ARIMA	Prediksi SARIMA
2025	January	1925489	1951691
2025	February	1932171	1883738
2025	March	1930693	1876709
2025	April	1931020	1846904
2025	May	1930948	1963159
2025	June	1930964	1974585
2025	July	1930960	2014089
2025	August	1930961	2000433
2025	September	1930961	1989707
2025	October	1930961	2031773
2025	November	1930961	1992374
2025	December	1930961	2040764

Gambar 29. Tabel Prediksi ARIMA dan SARIMA

Gambar 30 menampilkan prediksi jumlah penumpang bulanan berdasarkan model ARIMA dan SARIMA dengan menggunakan *Clustered Bar Chart*.

Prediksi ARIMA & SARIMA



Gambar 30. Prediksi ARIMA & SARIMA

5. Kesimpulan

Berdasarkan penelitian "Perancangan Visualisasi Data Penumpang Transjakarta dengan Metode *Four-step Kimball* Menggunakan Power BI" yang telah dilakukan, dapat disimpulkan hal-hal berikut:

1. Terdapat tren peningkatan jumlah penumpang Transjakarta dari tahun ke tahun setelah penurunan tajam pada awal tahun 2020, yang kemungkinan disebabkan oleh pandemi COVID-19 berdasarkan hasil dari visualisasi.
2. Proses analisis prediksi kode trayek 1 "Blok M - Kota" menggunakan model SARIMA membantu dalam mengidentifikasi pola dan tren dalam data penumpang. Hasil prediksi menunjukkan adanya penurunan pada 4 bulan kedepan dan kenaikan pada 8 bulan setelahnya. Model SARIMA yang digunakan menunjukkan pola data aktual dan prediksi yang saling berdekatan yang artinya memiliki nilai akurat yang cukup dalam analisa.
3. Rute dengan jumlah penumpang tertinggi adalah "Blok M - Kota", diikuti "Pinang Ranti - Pluit" dan "Lebak Bulus - Pasar Baru", sementara rute dengan penumpang terendah termasuk "Stasiun Tanah Abang - Stasiun Angke". Jenis trayek BRT mendominasi (44,27%), diikuti oleh Mikrotrans dan Angkutan Umum Integrasi, sedangkan Transjabodetabek memiliki porsi terkecil. Tahun 2024 mencatat jumlah penumpang tertinggi pascapandemi. Distribusi kuartalan cukup merata, dengan Q3 dan Q4 sedikit lebih tinggi.

6. Daftar Pustaka

- [1] S. SAHARA and D. Yuliana, "Analisis Tingkat Kepuasan Pelanggan Dalam Penerapan Sistem Boarding Pass Di Gate Keberangkatan Terminal Terpadu Pulo Gebang," *LOGISTIK*, vol. 14, no. 1, pp. 44–56, Apr. 2021, doi: <https://doi.org/10.21009/logistik.v14i1.20507>.
- [2] Drs. M.N. Nasution, M.S.Tr., APU., *Manajemen Transportasi*, 4th ed. Bogor: Ghalia Indonesia, 2016, p. 364.
- [3] P. Y. Rianti and R. W. Tuti, "Kualitas Pelayanan Transjakarta Busway di DKI Jakarta," *SWATANTRA*, vol. 15, no. 02, Apr. 2018.
- [4] D. Prastyo, "Perancangan Sistem Data Warehouse Menggunakan Four-step Methodology Kimball dengan Talend Open Studio dan Posgre SQL," *Uksw.edu*, 2021, doi: <https://repository.uksw.edu/handle/123456789/22502>.
- [5] A. V. Winona, T. Siswanto, and D. Sugiarto, "Dashboard Design for New Employee Candidate Program at XYZ Bank Using Tableau," *Intelmatix*, vol. 4, no. 1, pp. 31–37, Mar. 2024, doi: <https://doi.org/10.25105/itm.v4i1.17654>.
- [6] K. Salim, L. Damayanti, M. Puspita, S. Liujaya, and A. S. Girsang, "Data warehouse using Kimball approach in computer maniac," *IOP Conference Series: Materials Science and Engineering*, vol. 725, p. 012099, Jan. 2020, doi: <https://doi.org/10.1088/1757-899x/725/1/012099>.
- [7] R. Y. Sifa, "Visualisasi Data Pengunjung dan Peminjaman Buku di Perpustakaan Daerah Menggunakan Power BI," *TEKNOFILE : Jurnal Sistem Informasi*, vol. 2, no. 3, pp. 142–151, 2024, doi: <https://doi.org/10.30829/jipi.v8i1.13843>.
- [8] A. Hidayat, Z. Halim, and F. N. Hasan, "Implementasi Business Intelligence Untuk Menganalisis dan Memvisualisasikan Data Penumpang Bus Transjakarta Menggunakan Tableau," *Kesatria : Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 4, no. 3, pp. 771–779, 2023, doi: <https://doi.org/10.30645/kesatria.v4i3.224>.
- [9] M. Hidayat, "Perancangan Dashboard Untuk Visualisasi Data Jumlah Penumpang Transjakarta," *INSANtek*, vol. 4, no. 1, pp. 32–36, May 2023, doi: <https://doi.org/10.31294/insantek.v4i1.2222>.
- [10] Y. O. H. Kristanto, "Proses ETL (Extract Transformation Loading) Data Warehouse untuk Peningkatan Kinerja Biodata dalam Menyajikan Profil Mahasiswa dari Dimensi Asal Sekolah," *RESEARCH FAIR UNISRI*, vol. 3, no. 1, Jan. 2019, doi: <https://doi.org/10.33061/rsfu.v3i1.2615>.
- [11] P. Agus and I. Made, "Business Intelligence Based on Kimball Nine-Steps Methodology for Monitoring the Feasibility of Goods in Market," *International Journal of Advances in Data and Information Systems*, vol. 4, no. 2, pp. 135–144, Sep. 2023, doi: <https://doi.org/10.25008/ijadis.v4i2.1301>.
- [12] A. Pristanto, O. Trifena, D. Aryani, and S. Sahara, "Analisis Kepuasan Pelanggan Terhadap Fasilitas yang Disediakan oleh Pihak Transjakarta," *Jurnal Manajemen Riset Inovasi*, vol. 1, no. 3, pp. 09–17, May 2023, doi: <https://doi.org/10.55606/mri.v1i3.1162>.
- [13] M. I. Salman and R. W. P. Pamungkas, "Analisis Business Intelligence Data Penjualan Elektronik 2014 - 2017," *Nusantara Journal of Multidisciplinary Science*, vol. 1, no. 12, pp. 712–722, 2014, doi: <https://doi.org/10.30865/mib.v6i3.4338>.
- [14] S. Anardani, M. N. L. Azis, and M. Y. Asyari, "The Implementation of Business Intelligence to Analyze Sales Trends in the Indofishing Online Store Using Power BI," *Brilliance: Research of Artificial Intelligence*, vol. 3, no. 2, pp. 300–305, Dec. 2023, doi: <https://doi.org/10.47709/brilliance.v3i2.3232>.
- [15] G. Ritchie and S. Hariyanto, "Implementation of Business Intelligence In Analyzing Data Using Tableau at PT Global Bintan Permata," *bit-Tech*, vol. 6, no. 1, pp. 40–50, Aug. 2023, doi: <https://doi.org/10.32877/bt.v6i1.875>.
- [16] N. E. Warestika, D. Sugiarto, and T. Siswanto, "Business Intelligence Design for Data Visualization and Drug Stock Forecasting," *Intelmatix*, vol. 1, no. 1, Jan. 2021, doi: <https://doi.org/10.25105/itm.v1i1.7407>.

- [17] Q. D. H. B. Sitepu, S. Sutarmanto, and M. A. P. Siregar, "Metode Autoregressive Integrated Moving Average (Arima) dalam Memprediksi Jumlah Penumpang Kereta Api Kota Binjai," *Jurnal Arjuna Publikasi Ilmu Pendidikan Bahasa dan Matematika*, vol. 2, no. 2, pp. 69–85, Jan. 2024, doi: <https://doi.org/10.61132/arjuna.v2i2.621>.
- [18] D. G. Taslim and I. M. Murwantara, "A Comparative Study of ARIMA and LSTM in Forecasting Time Series Data," *2022 9th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Aug. 2022, doi: <https://doi.org/10.1109/icitacee55701.2022.9924148>.
- [19] I. Fitriani, P. R. Arum, and S. Amri, "Pemodelan Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk Meramalkan Volume Angkutan Barang Kereta Api di Pulau Jawa Tahun 2021: Memberikan prediksi volume angkutan barang kereta api di pulau jawa untuk masa mendatang," *Journal Of Data Insights*, vol. 2, no. 1, pp. 26–35, 2021, Accessed: Jan. 27, 2025. [Online]. Available: <http://103.97.100.158/index.php/jodi/article/view/167>
- [20] A. Febiola *et al.*, "Perbandingan Metode ARIMA dan SARIMA Dalam Peramalan Jumlah Penumpang Bandara Provinsi Kepulauan Bangka Belitung," *Jambura Journal of Mathematics*, vol. 6, no. 2, Aug. 2024, doi: <https://doi.org/10.37905/jjom.v6i2.25081>.
- [21] K. Y. Iman, "Metode Seasonal Autoregressive Integrated Moving Average (SARIMA) untuk Memprediksi Jumlah Penumpang Kereta Api di Pulau Sumatera - Repository Unja," *Unja.ac.id*, Jul. 2021, doi: <https://repository.unja.ac.id/23229/1/Skripsi%20Full%20Teks.pdf>.
- [22] P. Utomo and A. Fanani, "Peramalan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA)," *Jurnal Algebra*, vol. 1, no. 2, pp. 169–178, 2020, doi: <https://doi.org/10.29080/algebra.v1i2.6>.
- [23] S. Cania, D. M. Putri, and I. D. Rianjaya, "Penerapan Model Seasonal Autoregressive Integrated Moving Average (SARIMA) pada Jumlah Penumpang Kereta Api di Sumatera Barat," *JOSTECH Journal of Science and Technology*, vol. 3, no. 2, pp. 209–220, Oct. 2023, doi: <https://doi.org/10.15548/jostech.v3i2.6880>.
- [24] N. Nurhayati, Z. Razi, and W. Apriani, "PEMODELAN DAN PREDIKSI CURAH HUJAN MENGGUNAKAN IMBAL HASIL HETEROSKEDASTIS DI KABUPATEN BANDUNG," *Jurnal Lebesgue Jurnal Ilmiah Pendidikan Matematika Matematika dan Statistika*, vol. 5, no. 3, pp. 2303–2319, Dec. 2024, doi: <https://doi.org/10.46306/lb.v5i3.860>.
- [25] P. Zainal, Y. Angraini, and A. Rizki, "Penerapan Metode Generalized Auto-Regressive Conditional Heteroscedasticity untuk Peramalan Harga Minyak Mentah Dunia," *Xplore*, vol. 12, no. 1, pp. 12–21, Jan. 2023, doi: <https://doi.org/10.29244/xplore.v12i1.1096>.
- [26] A. D. W. Sumari, M. B. Musthafa, N. Ngatmari, and D. R. H. Putra, "Comparative Performance of Prediction Methods for Digital Wallet Transactions in the Pandemic Period," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 642–647, Aug. 2020, doi: <https://doi.org/10.29207/resti.v4i4.2024>.