

SKRIPSI

**ANALISIS KEPADATAN DAN *CLUSTERING* JUMLAH PENUMPANG
TRANSJAKARTA SAAT JAM SIBUK MENGGUNAKAN METODE
*CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING***

***ANALYSIS OF DENSITY AND CLUSTERING OF TRANSJAKARTA
PASSENGER NUMBERS DURING RUSH HOURS USING THE CROSS
INDUSTRY STANDARD PROCESS FOR DATA MINING METHOD***



RIFQI ARIEF PERMANA

24010121130075

DEPARTEMEN MATEMATIKA

FAKULTAS SAINS DAN MATEMATIKA

UNIVERSITAS DIPONEGORO

SEMARANG

2025

SKRIPSI

**ANALISIS KEPADATAN DAN *CLUSTERING* JUMLAH PENUMPANG
TRANSJAKARTA SAAT JAM SIBUK MENGGUNAKAN METODE
*CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING***

***ANALYSIS OF DENSITY AND CLUSTERING OF TRANSJAKARTA
PASSENGER NUMBERS DURING RUSH HOURS USING THE CROSS
INDUSTRY STANDARD PROCESS FOR DATA MINING METHOD***

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Sarjana
Matematika (S.Mat)



RIFQI ARIEF PERMANA

24010121130075

**DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO
SEMARANG**

2025

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS KEPADATAN DAN CLUSTERING JUMLAH PENUMPANG
TRANSJAKARTA SAAT JAM SIBUK MENGGUNAKAN METODE
CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING

Telah dipersiapkan dan disusun oleh:

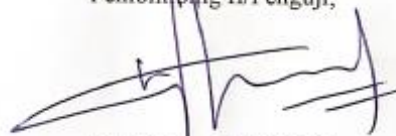
Rifqi Arief Permana

24010121130075

Telah dipertahankan di depan Tim Penguji pada tanggal, 15 Januari 2025

Susunan Tim Penguji

Pembimbing II/Penguji,



Dr. Sutrisno, S.Si., M.Sc

NIP. 198609012014041003

Penguji,



Dr. Drs. Kartono, M.Si

NIP. 19630825199031003

Ketua Departemen Matematika,



Dr. Susilo Hariyanto, S.Si., M.Si

NIP. 197410142000121001

Pembimbing I/Penguji,



Dr. R. Heru Tjahjana, S.Si., M.Si

NIP. 197407172000121001

ABSTRAK

ANALISIS KEPADATAN DAN *CLUSTERING* JUMLAH PENUMPANG TRANSJAKARTA SAAT JAM SIBUK MENGGUNAKAN METODE *CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING*

oleh

Rifqi Arief Permana

24010121130075

Transjakarta merupakan transportasi publik dengan sistem *Bus Rapid Transit* (BRT) yang banyak digunakan oleh masyarakat Jakarta. Namun, tingginya jumlah penumpang terutama saat jam sibuk pagi dan sore, sering menyebabkan kepadatan di koridor tertentu serta ketidak seimbangan distribusi armada bus. Hal ini mengakibatkan kekurangan bus di koridor dengan permintaan tinggi dan kelebihan bus di koridor dengan permintaan rendah, sehingga menurunkan efisiensi operasional. Oleh karena itu, diperlukan analisis pola jumlah penumpang untuk mengidentifikasi koridor dengan kepadatan tinggi serta mengoptimalkan jumlah bus yang beroperasi. Penelitian ini menerapkan *clustering* untuk mengelompokkan koridor berdasarkan pola jumlah penumpang menggunakan data periode Desember 2023 – Agustus 2024. Metodologi yang digunakan mengikuti tahapan CRISP - DM, meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan deployment. Pada tahap pemodelan, algoritma *K – Means* dan *K – Medoids* digunakan untuk melakukan *clustering*. Evaluasi hasil *clustering* dilakukan menggunakan tiga nilai evaluasi yaitu *Sum of Squared Error* (SSE), *Davies – Bouldin Index* (DBI), dan *Silhouette Score* untuk menilai kualitas *cluster* yang terbentuk. Hasil penelitian menunjukkan tiga cluster utama, yaitu *cluster 0* (jumlah penumpang rendah), *cluster 1* (jumlah penumpang sedang), dan *cluster 2* (jumlah penumpang tinggi). Hasil ini memberikan wawasan mengenai koridor dengan kepadatan tinggi selama jam sibuk, yang dapat menjadi dasar bagi manajemen Transjakarta dalam mengoptimalkan jumlah bus dan distribusi armada, serta berkontribusi dalam pengambilan keputusan untuk meningkatkan efisiensi layanan Transjakarta.

Kata kunci: *Clustering*, *K – Means*, *K – Medoids*, CRISP – DM, Transjakarta

ABSTRACT

ANALYSIS OF DENSITY AND CLUSTERING OF TRANSJAKARTA PASSENGER NUMBERS DURING RUSH HOURS USING THE CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING

by

Rifqi Arief Permana

24010121130075

Transjakarta is a public transportation system with a Bus Rapid Transit (BRT) model widely used by the people of Jakarta. However, the high number of passengers, especially during peak hours in the morning and afternoon, often causes congestion in certain corridors and imbalance in bus fleet distribution. This results in a shortage of buses in corridors with high demand and an excess of buses in corridors with low demand, thus reducing operational efficiency. Therefore, an analysis of passenger patterns is needed to identify corridors with high congestion and optimize the number of buses operating. This study applies clustering to group corridors based on passenger patterns using data from December 2023 to August 2024. The methodology follows the CRISP - DM stages, including business understanding, data understanding, data preparation, modeling, evaluation, and deployment. In the modeling stage, K - Means and K - Medoids algorithms are used for clustering. The clustering results are evaluated using three evaluation metrics: Sum of Squared Error (SSE), Davies - Bouldin Index (DBI), and Silhouette Score to assess the quality of the formed clusters. The results of the study show three main clusters: cluster 0 (low passenger count), cluster 1 (moderate passenger count), and cluster 2 (high passenger count). These results provide insights into corridors with high congestion during peak hours, which can serve as a basis for Transjakarta management to optimize the number of buses and fleet distribution, as well as contribute to decision-making to improve the efficiency of Transjakarta services.

Keywords: Clustering, K – Means, K – Medoids, CRISP – DM, Transjakarta