

**ANALISIS DATA PELANGGAN DARI PENJUALAN MOTOR**

**PT XYZ MENGGUNAKAN ALGORITMA K-MEANS**



**UMN**

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

Skripsi

**Luis Fernando**

**0000045013**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN INFORMATIKA  
UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG**

**2024**

**ANALISIS DATA PELANGGAN DARI PENJUALAN MOTOR**

**PT XYZ MENGGUNAKAN ALGORITMA K-MEANS**



**Skripsi**

Diajukan sebagai Salah Satu Syarat untuk Memperoleh  
Gelar Sarjana Komputer (S.Kom)

**Luis Fernando**

**0000045013**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN INFORMATIKA**

**UNIVERSITAS MULTIMEDIA NUSANTARA  
TANGERANG**

**2024**

## HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Luis Fernando  
Nomor Induk Mahasiswa : 00000045013  
Program studi : Sistem Informasi

Skripsi dengan judul:

### **Analisis Data Pelanggan dari Penjualan Motor PT XYZ Menggunakan Algoritma K-Means**

merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 27 Mei 2024



Luis Fernando

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

ii

Analisa Data Pelanggan..., Luis Fernando, Universitas Multimedia Nusantara

## HALAMAN PERSETUJUAN

Skripsi dengan judul

**Analisis Data Pelanggan dari Penjualan Motor  
PT XYZ Menggunakan Algoritma K-Means**

Oleh

Nama : Luis Fernando

NIM : 00000045013

Program Studi : Sistem Informasi

Fakultas : Teknik dan Informatika

Telah disetujui untuk diajukan pada

Sidang Ujian Skripsi Universitas Multimedia Nusantara

Tangerang, 15 Mei 2024

Pembimbing



Melissa Indah Fianty, S.Kom., M.MSI.  
0313019201/063252  
Ketua Program Studi Sistem Informasi



05/14/2024

Ririn Ikana Desanti, S.Kom., M.Kom.

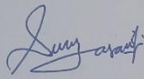
## HALAMAN PENGESAHAN

Skripsi dengan judul  
**Analisis Data Pelanggan dari Penjualan Motor  
PT XYZ Menggunakan Algoritma K-Means**

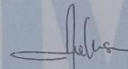
Oleh  
Nama : Luis Fernando  
NIM : 00000045013  
Program Studi : Sistem Informasi  
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Senin, 27 Mei 2024  
Pukul 15.00 s.d 17.00 dan dinyatakan  
**LULUS**  
Dengan susunan penguji sebagai berikut.

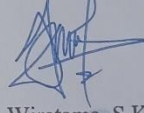
Ketua Sidang

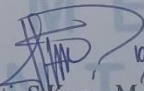
  
Suryasari, S.Kom., M.T.  
0323088301/067345  
10/6-2024

Pembimbing

  
Melissa Indah Fianty, S.Kom., M.MSI.  
0313019201/063252  
Ketua Program Studi Sistem Informasi  
10/6.24

Penguji

  
Jansen Wiratama, S.Kom., M.Kom.  
0409019301  
10/06  
2024

  
Ririn Ikana Desanti, S.Kom., M.Kom.  
10/6.24

iv

Analisa Data Pelanggan..., Luis Fernando, Universitas Multimedia Nusantara

NUSANTARA

## LEMBAR PERSETUJUAN PUBLIKASI

### KARYA ILMIAH MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Luis Fernando

NIM : 00000045013

Program Studi : Sistem Informasi

Jenjang : D3 / S1 / S2

Judul Karya Ilmiah :

**ANALISIS DATA PELANGGAN DARI PENJUALAN MOTOR PT XYZ  
MENGUNAKAN ALGORITMA K-MEANS**

Menyatakan dengan sesungguhnya bahwa saya bersedia:

Memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya di repositori Knowledge Center, sehingga dapat diakses oleh Civitas Akademika/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial dan saya juga tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.


Saya tidak bersedia, dikarenakan:

Dalam proses pengajuan untuk diterbitkan ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)\*.

Tangerang, 10 Juni 2024

Yang menyatakan,

UNIVERSITA  
MULTIMEDIA  
NUSANTARA



Luis Fernando

\* Jika tidak bisa membuktikan LoA jurnal/HKI selama 6 bulan kedepan, saya bersedia mengizinkan penuh karya ilmiah saya untuk diunggah ke KC UMN dan menjadi hak institusi UMN.

## KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya dapat menyelesaikan penulisan laporan Skripsi ini dengan judul: "Analisis Data Pelanggan dari Penjualan Motor PT XYZ Menggunakan Algoritma K-Means" yang dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Jurusan Sistem Informasi pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari sepenuhnya bahwa pencapaian ini tidak terwujud tanpa dukungan, bimbingan, dan bantuan dari berbagai pihak sepanjang perjalanan, mulai dari masa perkuliahan hingga pada tahap penyusunan tugas akhir ini, dengan tulus, saya ingin menyampaikan rasa terima kasih kepada:

Mengucapkan terima kasih

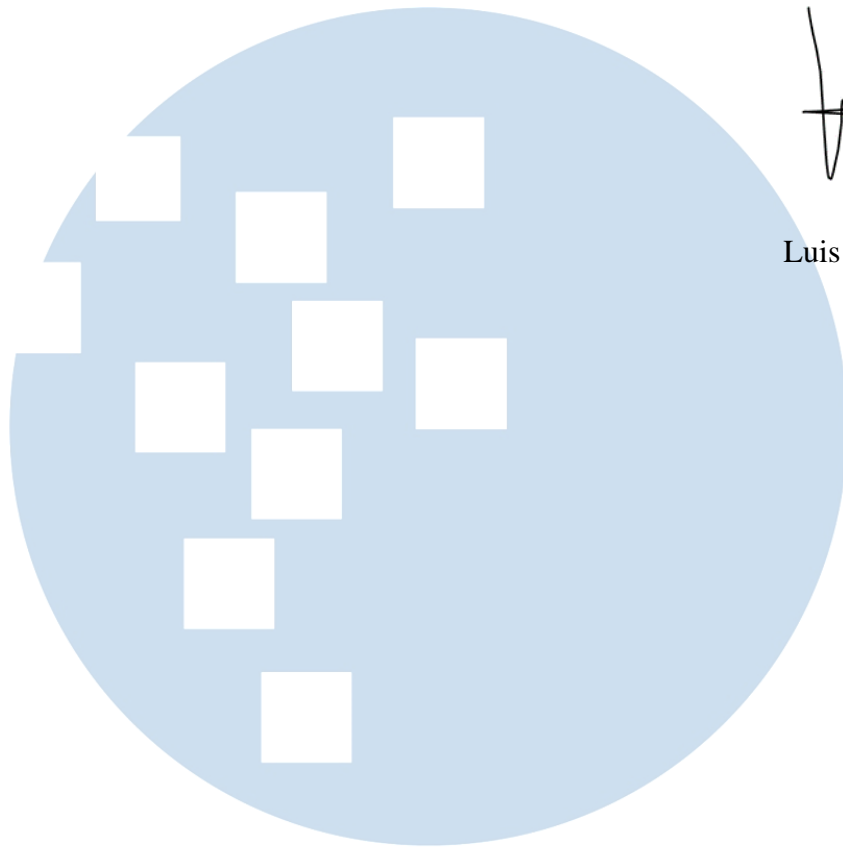
1. Bapak Dr. Ninok Leksono, M.A., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Melissa Indah Fianty, S.Kom., M.MSI., sebagai Pembimbing yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Semua teman-teman yang telah membantu dan memberikan dukungan pada saat pengerjaan laporan skripsi ini.

Harapan saya semoga skripsi ini dapat membantu bagi pembaca, dapat dijadikan referensi demi pengembangan yang lebih baik, dan berjalan dengan lancar.

Tangerang, 15 Mei 2024



Luis Fernando



UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA



# ANALISIS DATA PELANGGAN DARI PENJUALAN MOTOR

## PT XYZ MENGGUNAKAN ALGORITMA K-MEANS

Luis Fernando

### ABSTRAK

PT XYZ merupakan suatu perusahaan manufaktur yang beroperasi di industri kendaraan sepeda motor dan perlengkapan kendaraan lainnya. Sebagai perusahaan manufaktur aktif yang memerlukan manajemen operasional yang efektif untuk memastikan bahwa seluruh tahapan dan prosedur dalam aktivitas operasional dilakukan dengan efisien dan sesuai standar. Industri sepeda motor sebagai salah satu dari sektor otomotif di Indonesia mengalami pertumbuhan yang signifikan. Perusahaan ini menawarkan berbagai tipe sepeda motor kepada konsumen, Teknologi unggul yang dimiliki telah teruji dalam berbagai situasi, baik di jalanan maupun dalam pengembangan teknologi yang dapat memenuhi kebutuhan konsumen. Bahwa diperlukan untuk menganalisis dengan dataset perusahaan yang secara merinci data segmentasi pelanggan untuk memfasilitasi pembuatan laporan, dan menganalisis pelanggan dari penjualan motor. Hal ini bertujuan untuk memudahkan pengambilan keputusan yang lebih baik. Sistem informasi dirancang dengan menggunakan metode algoritma K-Means untuk mendukung pada penelitian ini. Metode pengumpulan data yang digunakan yakni pengambilan sampel pelanggan dan studi pustaka. Dari data yang dikumpulkan maka ditentukan tujuan perusahaan dalam hal ini, perlu melakukan analisa data menggunakan algoritma K-Means dengan metode Elbow dan Silhouette. Hasil evaluasi dalam mengelola data pelanggan dari penjualan motor. Pertama berdasarkan tipe motor yang penjualan didominasi oleh tipe motor yakni BEAT SPORTY CBS, PCX 160 CBS, dan REVO FIT FI, Kedua berdasarkan kreditor yang didominasi oleh PT. Federal International Finance. Ketiga berdasarkan provinsi yang didominasi oleh Jawa Barat, Jawa Tengah, dan Jawa Timur. Keempat berdasarkan usia yang didominasi oleh kelompok dewasa. Maka dari itu keempat pengelompokan segmentasi data pelanggan yang merupakan penjualan tertinggi dalam kebutuhan dan preferensi pelanggan.

**Kata kunci:** K-Means, Pelanggan, Penjualan, Sistem Informasi, Teknologi

# **ANALYSIS OF CUSTOMER DATA FROM MOTORCYCLE SALES AT PT XYZ USING THE K-MEANS ALGORITHM**

Luis Fernando

## **ABSTRACT (English)**

*PT XYZ is a manufacturing company operating in the motorcycle and related vehicle equipment industry. As an active manufacturing company, it requires effective operational management to ensure that all stages and procedures in its operational activities are carried out efficiently and meet standards. The motorcycle industry, as a part of the automotive sector in Indonesia, has experienced significant growth. The company offers various types of motorcycles to consumers. Its superior technology has been tested in various situations, both on the road and in technological developments that meet consumer needs. It is necessary to analyze the company's dataset, which details customer segmentation data to facilitate report generation and to analyze customer behavior from motorcycle sales. This aims to simplify better decision-making. The information system is designed using the K-Means algorithm method to support this research. The data collection methods used include customer sampling and literature study. Based on the collected data, the company's objectives are determined. In this case, it is necessary to analyze the data using the K-Means algorithm with the Elbow and Silhouette methods. The evaluation results in managing customer data from motorcycle sales are as follows. First, based on the type of motorcycle, sales are dominated by BEAT SPORTY CBS, PCX 160 CBS, and REVO FIT FI. Second, based on creditors, sales are dominated by PT. Federal International Finance. Third, based on provinces, sales are dominated by West Java, Central Java, and East Java. Fourth, based on age, sales are dominated by the adult age group. Therefore, these four segments represent the highest sales in terms of customer needs and preferences.*

*Keywords: Customer, Information Systems, K-Means, Sales, Technology*

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

## DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT .....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA v	
KATA PENGANTAR.....	vi
ABSTRAK .....	viii
ABSTRACT (English) .....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR RUMUS.....	xiv
DAFTAR LAMPIRAN .....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	5
1.3 Batasan Masalah.....	6
1.4 Tujuan dan Manfaat Penelitian .....	6
1.4.1 Tujuan Penelitian .....	6
1.4.2 Manfaat Penelitian .....	7
1.5 Sistematika Penulisan .....	7
BAB II LANDASAN TEORI .....	9
2.1 Penelitian Terdahulu.....	9
2.2 Teori tentang Skripsi.....	15
2.2.1 Analisa Data.....	15
2.3 Teori tentang algoritma yang digunakan.....	18
2.3.1 Text Mining .....	18
2.3.2 K-Means .....	20
2.4 Teori tentang Tools yang digunakan .....	22
2.4.1 Python .....	22
BAB III METODOLOGI PENELITIAN .....	24
3.1 Gambaran Umum Objek Penelitian.....	24
3.1.1 Objek Penelitian.....	24
3.1.2 Struktur Organisasi .....	25
3.1.3 Visi dan Misi .....	27
3.2 Metode Penelitian .....	27
3.2.1 Alur Penelitian.....	27
3.3 Teknik Pengumpulan Data.....	30
3.3.1 Teknik Pengambilan Sampel .....	30
3.3.2 Studi Pustaka.....	31
3.4 Variabel Penelitian .....	31
3.5 Teknik Analisis Data .....	32
BAB IV ANALISIS DAN HASIL PENELITIAN .....	34
4.1 Identifikasi Permasalahan .....	34
4.2 Preprocessing Data.....	34

<b>4.3</b>	<b>Data Mining</b>	35
<b>4.4</b>	<b>Elbow</b>	36
4.4.1	Penentuan Jumlah Cluster Segmentasi Pelanggan Berdasarkan Jenis Motor Menggunakan Metode Elbow	37
4.4.2	Penentuan Jumlah Cluster Segmentasi Pelanggan Berdasarkan Kreditor Menggunakan Metode Elbow	38
4.4.3	Penentuan Jumlah Cluster Segmentasi Pelanggan Berdasarkan Provinsi Menggunakan Metode Elbow	39
4.4.4	Penentuan Jumlah Cluster Segmentasi Pelanggan Berdasarkan Kategori Usia Menggunakan Metode Elbow	40
<b>4.5</b>	<b>Silhouette Score</b>	40
4.5.1	Evaluasi Kualitas Cluster Segmentasi Pelanggan Berdasarkan Jenis Motor Menggunakan Metode Silhouette	40
4.5.2	Evaluasi Kualitas Cluster Segmentasi Pelanggan Berdasarkan Jenis Kreditor Menggunakan Metode Silhouette	42
4.5.3	Evaluasi Kualitas Cluster Segmentasi Pelanggan Berdasarkan Provinsi Menggunakan Metode Silhouette	43
4.5.4	Evaluasi Kualitas Cluster Segmentasi Pelanggan Berdasarkan Kategori Usia Menggunakan Metode Silhouette	44
<b>4.6</b>	<b>Modeling</b>	46
<b>4.7</b>	<b>K-Means</b>	46
4.7.1	Hasil Cluster Pelanggan Berdasarkan Tipe Motor	46
4.7.2	Hasil Cluster Pelanggan Berdasarkan Jenis Kreditor	48
4.7.3	Hasil Cluster Pelanggan Berdasarkan Provinsi	50
4.7.4	Hasil Cluster Pelanggan Berdasarkan Kategori Usia	52
<b>4.8</b>	<b>Hasil dan Pembahasan</b>	53
4.8.1	Segmentasi Pelanggan Berdasarkan Jenis Motor	53
4.8.2	Segmentasi Pelanggan Berdasarkan Jenis Kreditor	54
4.8.3	Segmentasi Pelanggan Berdasarkan Provinsi	55
4.8.4	Segmentasi Pelanggan Berdasarkan Kategori Usia	56
<b>BAB V</b>	<b>SIMPULAN DAN SARAN</b>	58
5.1	Simpulan	58
5.2	Saran	60
<b>DAFTAR PUSTAKA</b>		61
<b>LAMPIRAN</b>		64

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR TABEL

Tabel 1.1 Rekap Permasalahan Penjualan Motor .....	2
Tabel 2.1 Penelitian Terdahulu .....	9
Tabel 3.1 Perbandingan tools Python dengan Google Colab.....	32
Tabel 4.1 Nilai Silhouette Score Pelanggan Berdasarkan Tipe Motor .....	41
Tabel 4.2 Nilai Silhouette Score Pelanggan Berdasarkan Jenis Kreditor .....	42
Tabel 4.3 Nilai Silhouette Score Segmentasi Pelanggan Berdasarkan Provinsi ...	44
Tabel 4.4 Nilai Silhouette Score Segmentasi Pelanggan Berdasarkan Kategori Umur .....	45
Tabel 4.5 Hasil Cluster Pelanggan Berdasarkan Jenis Motor .....	46
Tabel 4.6 Hasil Cluster Pelanggan Berdasarkan Jenis Kreditor .....	48
Tabel 4.7 Hasil Cluster Pelanggan Berdasarkan Provinsi.....	50
Tabel 4.8 Hasil Cluster Pelanggan Berdasarkan Kategori Usia.....	52

UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR GAMBAR

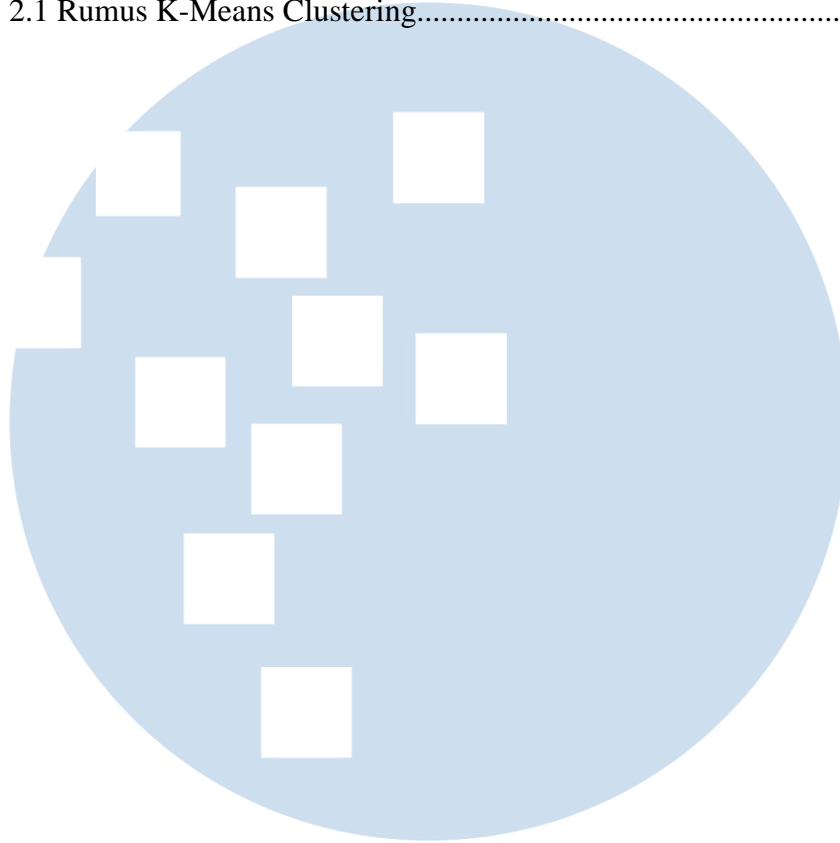
Gambar 1.1 Tipe Merek Sepeda Motor yang Terbanyak Tahun 2023 .....	4
Gambar 2.1 Komponen Analisis Data .....	17
Gambar 2.2 Tahapan Text Mining .....	19
Gambar 3.1 Struktur Organisasi PT XYZ.....	25
Gambar 3.2 Alur Penelitian.....	28
Gambar 4.1 Menampilkan Jumlah Data yang Hilang.....	35
Gambar 4.2 Dikonversi Data ke Datetime dalam Bentuk Numerik.....	36
Gambar 4.3 Jumlah Cluster Segmentasi Pelanggan Berdasarkan Jenis Motor.....	37
Gambar 4.4 Jumlah Cluster Segmentasi Pelanggan Berdasarkan Kreditor .....	38
Gambar 4.5 Jumlah Cluster Segmentasi Pelanggan Berdasarkan Provinsi .....	39
Gambar 4.6 Jumlah Cluster Segmentasi Pelanggan Berdasarkan Kategori Usia ..	40
Gambar 4.7 Segmentasi Pelanggan Berdasarkan Jenis Motor .....	41
Gambar 4.8 Segmentasi Pelanggan Berdasarkan Jenis Kreditor .....	42
Gambar 4.9 Segmentasi Pelanggan Berdasarkan Provinsi .....	43
Gambar 4.10 Segmentasi Pelanggan Berdasarkan Kategori Usia .....	45
Gambar 4.11 Scatter Plot Pelanggan Berdasarkan Tipe Motor .....	47
Gambar 4.12 Scatter Plot Pelanggan Berdasarkan Jenis Kreditor .....	49
Gambar 4.13 Scatter Plot Pelanggan Berdasarkan Provinsi .....	51
Gambar 4.14 Scatter Plot Pelanggan Berdasarkan Kategori Usia .....	53

UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

**DAFTAR RUMUS**

Rumus 2.1 Rumus K-Means Clustering.....21

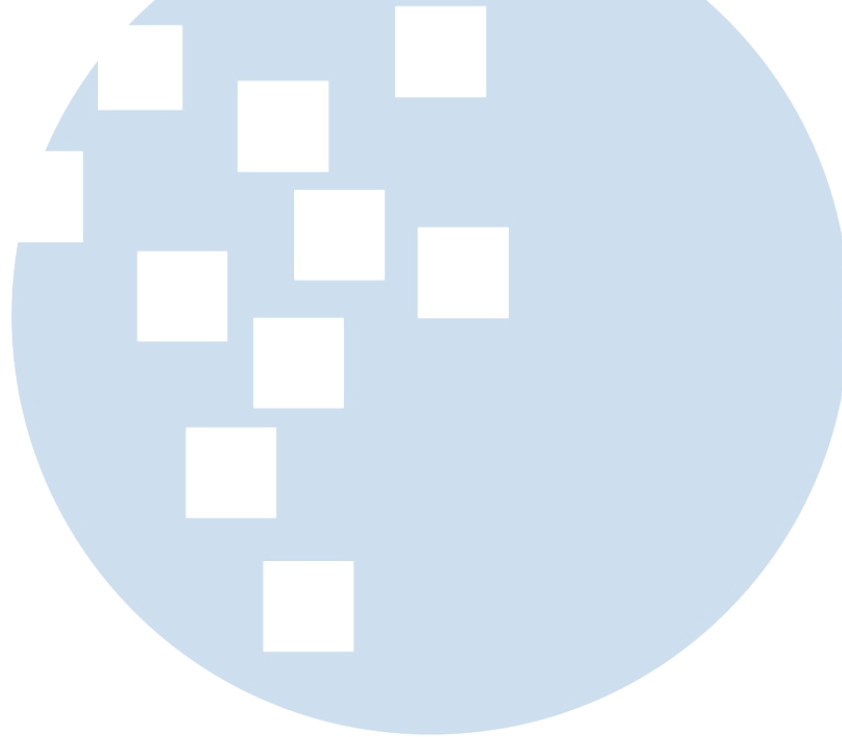


UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

## DAFTAR LAMPIRAN

Lampiran A Nama Form Bimbingan Skripsi.....	64
Lampiran B Hasil Surat Permohonan Izin.....	65
Lampiran C Hasil Turnitin.....	66



# UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA