LAPORAN PELAKSANAAN

MBKM Magang Mandiri Telkomsel Branch Surabaya

Semester: V TA Ganjil 2024/2025

PEMODELAN DERET WAKTU DATA KUNJUNGAN TELKOMSEL KE SEKOLAH MENGGUNAKAN ARIMA

Nama	**: Tsabita Rosyidah Putri**
NPM	**: 22083010012**
Dosen Pembimbing	**: Trimono, S.Si., M.Si.**
	(NIP. 199509082022031003)

Program Studi Sains Data

Fakultas Ilmu Komputer

Universitas Pembangunan Nasional "Veteran" Jawa Timur

2024

LAPORAN PELAKSANAAN

MBKM Magang Mandiri Telkomsel Branch Surabaya

Semester: V TA Ganjil 2024/2025

PEMODELAN DERET WAKTU DATA KUNJUNGAN TELKOMSEL KE SEKOLAH MENGGUNAKAN ARIMA

Nama	**: Tsabita Rosyidah Putri**
NPM	**: 22083010012**
Dosen Pembimbing	**: Trimono, S.Si., M.Si.**
	(NIP.199509082022031003)

Program Studi Sains Data

Fakultas Ilmu Komputer

Universitas Pembangunan Nasional "Veteran" Jawa Timur

2024

LEMBAR PENGESAHAN

Magang Mandiri – Telkomsel Branch Surabaya

Semester: V TA 2024/2025

"PEMODELAN DERET WAKTU DATA KUNJUNGAN TELKOMSEL KE SEKOLAH MENGGUNAKAN ARIMA"

Disetujui oleh,

Pembimbing Lapangan	Dosen Pembimbing
Moch Sutrio Febri Fitriyanto	Trimono, S.Si., M.Si. NIP. 199509082022031003

Mengetahui,

Koordinator Program Studi Sains Data

Dr. Eng. Ir. Dwi Arman Prasetya, ST., M.T., IPU., Asean. Eng

NIP. 19801205 2005011 002

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas limpahan rahmat, hidayah, serta kesempatan yang diberikan sehingga penulis dapat menyelesaikan kegiatan Magang Mandiri di Telkomsel Branch Surabaya. Magang ini telah memberikan banyak ilmu dan pengalaman berharga yang akan sangat bermanfaat bagi perkembangan penulis di dunia kerja.

Kegiatan ini tentunya tidak terwujud tanpa bantuan dan dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Orang tua tercinta yang senantiasa memberikan doa, dukungan, baik secara moral maupun materiil selama proses magang berlangsung.
2. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU, selaku Koordinator Program Studi Sains Data Universitas Pembangunan Nasional "Veteran" Jawa Timur,
3. Bapak Trimono, S.Si., M.Si., selaku Dosen Pembimbing Akademik, atas segala arahan dan dukungan selama masa studi.
4. Bapak Moch Sutrio Febri Fitriyanto, selaku pembimbing lapangan di Telkomsel Branch Surabaya, yang dengan sabar memberikan bimbingan dan wawasan praktis di dunia kerja.
5. Rekan-rekan magang serta seluruh pegawai Telkomsel Branch Surabaya yang telah bekerja sama dengan baik dan membantu selama pelaksanaan magang.
6. Serta semua pihak yang telah memberikan dukungan dan bantuan dalam bentuk apapun selama kegiatan ini berlangsung.

Penulis menyadari bahwa laporan ini masih memiliki kekurangan dan ruang untuk perbaikan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun guna peningkatan kualitas di masa mendatang. Semoga laporan ini dapat memberikan manfaat bagi semua pihak yang membaca.

Surabaya, 1 Januari 2025

DAFTAR ISI

LEMBAR PENGESAHAN II
KATA PENGANTAR III
DAFTAR ISI IV
DAFTAR GAMBAR VI
DAFTAR TABEL IX
BAB 1. PENDAHULUAN 1
1.1. Latar Belakang 1
1.2. Tujuan Magang Mandiri MBKM 2
1.3. Manfaat Magang Mandiri MBKM 2
1.4. Mahasiswa Menguraikann Tujuan Penulisan Topik Magang Mandiri MBKM 2
BAB 2. LOKASI PROYEK MAGANG MANDIRI 3
2.1. Lokasi & sejarah Mitra Magang Mandiri 3
2.2. Struktur Organisasi Mitra Magang Mandiri 5
2.3. Visi dan Misi Telkomsel 7
2.4. Kegiatan Produksi (barang/jasa) 8
BAB 3. PELAKSANAAN MAGANG MANDIRI 11
3.1. Posisi / Kedudukan Magang Mandiri 11
3.2. Metodologi Penyelesain Tugas 12
3.2.1. Tinjauan Pustaka 12
3.2.2. Diagram Alur 21
3.2.3. Timeline 22
3.2.4. Analisis dan Pembahasan 22
3.3. Pembelajaran Hal Baru 30
BAB 4. REKOGNISI MATA KULIAH 31
4.1. Rekognisi Mata Kuliah Praktik Kerja Lapangan 33
4.1.1. Praktik Lapangan Kerja [Indikator ke-1] 33
4.1.2. Menjalin Kerja Sama Dalam Praktik Lapangan Kerja[Indikator ke-2] 34
4.1.3. Menganalisis Data Serta Memvisualisasikan [Indikator ke-3] 35
4.1.4. Kesimpulan Data [Indikator ke-4] 36
4.2. Rekognisi Mata Kuliah Kerja Sama 37
4.2.1. Materi Softskill – Kerja Sama [Indikator ke-1] 37
4.2.2. Pemanfaatan Sumber Daya Manusia Dalam Tim Secara Optimal [Indikator ke-2] 39

4.2.3. Pola Kerja Sama [Indikator ke-3] 40
4.2.4. Mekanisme Monitoring Progress Pekerjaan Secara Berkala [Indikator ke-4]... 41
4.3. Rekognisi Mata Kuliah Analisis Segmentasi Pelanggan..... 42
4.3.1. Materi Analisis Segmentasi Pelanggan [Indikator ke-1] 42
4.3.2. Clustering Menggunakan Metode K-Means: Segmentasi Pelanggan Berdasarkan Area, Rasio Aktivitas dan Omzet untuk Strategi Pemasaran Menggunakan Data Event Telkomsel Menggunakan Bahasa Pemrograman Python [Indikator ke-2, Indikator ke-3, dan Indikator ke-4]..... 46
4.4. Rekognisi Mata Kuliah Instuisi dan Wawasan Data 58
4.4.1. Memahami Data [Indikator ke-1] 58
4.4.2. Menganalisis Data Events by Telkomsel Menggunakan Looker Studio[Indikator ke-2] 60
4.4.3. Mengambil keputusan & Menyelesaikan Permasalahan dari Data Events by Telkomsel [Indikator ke-3, Indikator ke-4] 62
4.5. Rekognisi Mata Kuliah Kecerdasan Bisnis..... 64
4.5.1. Materi Kecerdasan Bisnis [Indikator ke-1]..... 64
4.5.2. Metode Pengembangan Dashboard Kecerdasan Bisnis [Indikator ke-2, Indikator ke-3, dan Indikator ke-4] 69
4.6. Rekognisi Mata Kuliah Statistika Sains Data dan Analisis Kebijakan 80
4.6.1. Materi Statistika Sains Data dan Analisis Kebijakan [Indikator ke-1] 80
4.6.2. Analisis Strategi Sales Berdasarkan Area Menggunakan Bahasa Pemograman Python [Indikator ke-2, Indikator ke-3, dan Indikator ke-4] 85
4.7. Rekognisi Mata Kuliah Analisis Deret Waktu 90
4.7.1. Materi Analisis Deret Waktu [Indikator ke-1] 90
4.7.2. Analisis Deret Waktu Pada Data Kunjungan Telkomsel ke Sekolah di Indonesia Pada Bulan Januari – Oktober 2024 Dengan Metode Arima Menggunakan Bahasa Pemograman Python [Indikator ke-2, Indikator ke-3, Indikator ke-4]..... 92
4.8. Rekognisi Mata Kuliah Data Storytelling 111
4.8.1. Memahami sebuah data dan Menghasilkan Hasil Analisis dan Penyajian Data Menggunakan Looker Studio ke dalam sebuah narasi [Indikator ke-1, Indikator ke-2, Indikator ke-3, Indikator ke-4] 111
BAB 5. KESIMPULAN DAN SARAN..... 118
5.1. Kesimpulan..... 118
5.2. Saran 118
BAB 6. REFLEKSI DIRI 119
DAFTAR PUSTAKA 120
LAMPIRAN..... 121

DAFTAR GAMBAR

Gambar 2. 1 Lokasi Telkomsel Branch Surabaya 3
Gambar 2. 2 Logo Lama Telkomsel 3
Gambar 2. 3 Logo Telkomsel 4
Gambar 2. 4 Sejarah Telkomsel 5
Gambar 2. 5 Struktur Organisasi Telkomsel Branch Surabaya 6
Gambar 3. 1 Diagram Alur Pemodelan Analisis Deret Waktu Data Kunjungan Telkomsel Ke Sekolah Menggunakan Metode Arima 21
Gambar 3. 2 Data Kunjungan Telkomsel ke Sekolah 23
Gambar 3. 3 Plot Data Deret Waktu 23
Gambar 3. 4. Perbandingan Sebelum & Sesudah Transformasi Box-Cox 25
Gambar 3. 5 Plot PACF & ACF Sebelum Differencing 25
Gambar 3. 6. Plot PACF dan ACF 26
Gambar 3. 7 Estimasi Parameter 26
Gambar 3. 8. Distribusi Residual 27
Gambar 3. 9 Grafik Prediksi dan Data Aktual 28
Gambar 3. 10 Grafik Prediksi 20 Hari Kedepan 29
Gambar 4. 1 Berkolaborasi dan Berdiskusi 34
Gambar 4. 2 Koordinasi Dengan Mentor 34
Gambar 4. 3 Dashboard Populasi Penduduk Surabaya – Madura Tahun 2023 35
Gambar 4. 4 Dashboard Data 36
Gambar 4. 5 Membagi Tugas 37
Gambar 4. 6. Bukti Chat Saling Membantu 38
Gambar 4. 7 Bukti Chat Komunikasi Yang Terbuka 38
Gambar 4. 8 Gmeet Bersama 39
Gambar 4. 9 Kolaborasi Dengan Mentor 39
Gambar 4. 10 Bekerja Sama 40
Gambar 4. 11 Progress Melalui Gmeet 41
Gambar 4. 12 Workflow Analisis Segmentasi Pelanggan 49
Gambar 4. 13 Data Events 51
Gambar 4. 14 Output Jumlah Outlier 51
Gambar 4. 15 Output Setelah Menangani Outlier 52
Gambar 4. 16 Statistika Deskriptif 52
Gambar 4. 17 Jumlah Pada Regional 52
Gambar 4. 18 Grafik Regional 52
Gambar 4. 19 Grafik Branch 53
Gambar 4. 20 Jumlah Branch 53
Gambar 4. 21 Jumlah Cluster 53
Gambar 4. 22 Grafik Cluster 53
Gambar 4. 23 Plot Elbhow Method 54
Gambar 4. 24 Jumlah Pelanggan 55
Gambar 4. 25 Hasil Cluster Pelanggan 55
Gambar 4. 26 Visualisasi Cluster 55
Gambar 4. 27 Hasil Setiap Cluster 56
Gambar 4. 28 Hasil Pseudo F-Statistik 57
Gambar 4. 29. Hasil Strategi Setiap Cluster 57

Gambar 4. 30 Pemetaan Target Berdasarkan Cluster....57
Gambar 4. 31 MyNami Telkomsel....59
Gambar 4. 32 Data Events Daerah Surabaya - Madura....59
Gambar 4. 33 Data Events....60
Gambar 4. 34 Upload Excel Ke Looker Studio....61
Gambar 4. 35 Memfillter Data....61
Gambar 4. 36 Hasil Dashboard....62
Gambar 4. 37 Star Schema....66
Gambar 4. 38 Snowflake Schema....66
Gambar 4. 39 Model Multidimensional....67
Gambar 4. 40 Dashboard Visualisasi Data....69
Gambar 4. 41 Workflow Kecerdasan Bisnis....70
Gambar 4. 42 Diskusi....71
Gambar 4. 43 Studi Literatur....72
Gambar 4. 44 MyNami Telkomsel....72
Gambar 4. 45 Data Events....73
Gambar 4. 46 Data Events Setelah Merge....75
Gambar 4. 47 DataFrame....76
Gambar 4 48. Info Data....76
Gambar 4. 49 Rancangan Dashboard....77
Gambar 4. 50 Menambah Report di Looker Studio....77
Gambar 4. 51 Menghubungkan Excel Ke Looker Studio....77
Gambar 4. 52 Catatan Evaluasi....78
Gambar 4. 53 Hasil Dashboard Events....78
Gambar 4. 54 Prosedur Kebijakan....84
Gambar 4. 55 Workflow Statistika Sains Data dan Analisis Kebijakan....85
Gambar 4. 56 Data Events Surabaya....87
Gambar 4. 57 Sebelum & Setelah Mengangani Outlier....87
Gambar 4. 58 Hasil Analisis Deksriptif....88
Gambar 4. 59 Grafik Total Sales Tiap Kota....89
Gambar 4. 60 Diagram Alur Pemodelan Analisis Deret Waktu Data Kunjungan Telkomsel Ke Sekolah Menggunakan Metode Arima....92
Gambar 4. 61 Data School....94
Gambar 4. 62 Variabel Data School....94
Gambar 4. 63 Tipe Data School....95
Gambar 4. 64 DataFrame....95
Gambar 4. 65 Hasil Outlier....96
Gambar 4. 66 Info Data Schools....97
Gambar 4. 67 Hasil Statistika Deskripsi....98
Gambar 4. 68. Pivot Tabel....98
Gambar 4. 69 Jumlah Missing Value....99
Gambar 4. 70 Grafik Total Jumlah Kunjungan....100
Gambar 4. 71 Grafik Data....100
Gambar 4. 72 Hasil Uji Stasioner....103
Gambar 4. 73 Distribusi Sebelum dan Sesudah Transformasi Box-Cox....103
Gambar 4. 74 Plot acf dan pacf....104

Gambar 4. 75 Grafik Differencing 2 Kali 104
Gambar 4. 76 Plot PACF dan ACF Setelah di Differencing 2 kali 105
Gambar 4. 77 Hasil Estimasi Parameter 106
Gambar 4. 78 Distribusi dan Q-Q Plot Residual 106
Gambar 4. 79. Hasil Uji Ljung-Box dan Uji Normalitas 107
Gambar 4. 80 Grafik Prediksi dan Data Aktual 109
Gambar 4. 81 Grafik Hasil Prediksi 110
Gambar 4. 82 Data School 113
Gambar 4. 83 Menyampaikan Hasil Analisis 114
Gambar 4. 84 Dashboard Reporting School 114
Gambar 4. 85 Infografis 115

DAFTAR TABEL

Tabel 2. 1 Sejarah Telkomsel 5
Tabel 2. 2 Struktur Organisasi 6
Tabel 2. 3 Kegiatan Produksi 8
Tabel 3. 1 Posisi Magang 11
Tabel 3. 2 Identifikasi Masalah 19
Tabel 3. 3 Kriteria Nilai MAPE 20
Tabel 3. 4 Timeline 22
Tabel 3. 5 Evaluasi Model 28
Tabel 3. 6 Hasil Peramalan 20 Hari Kedepan 28
Tabel 4. 1 Rekognisi Mata Kuliah 31
Tabel 4. 2 Indikator Penilaian Konversi Mata Kuliah Praktik Lapangan Kerja 33
Tabel 4. 3 Indikator Konversi Kerja Sama 37
Tabel 4. 4 Indikator Penilaian Konversi Mata Kuliah Analisis Segmentasi Pelanggan 42
Tabel 4. 5 Indikator Penilaian Konversi Mata Kuliah Instuisi dan Wawasan Data 58
Tabel 4. 6 Indikator Penilaian Konversi Mata Kuliah Kecerdasan Bisnis 64
Tabel 4. 7 Indikator Penilaian Konversi Mata Kuliah Statistika Sains Data dan Analisis Kebijakan 80
Tabel 4. 8 Indikator Penilaian Konversi Mata Kuliah Analisis Deret Waktu 90
Tabel 4. 9 Evaluasi Model 109
Tabel 4. 10 Tabel Prediksi 20 Hari Kedepan 110
Tabel 4. 11 Indikator Penilaian Konversi Mata Kuliah Data Storytelling 111

BAB 1. PENDAHULUAN

1.1. Latar Belakang

Program Merdeka Belajar Kampus Merdeka (MBKM) yang diluncurkan oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia merupakan salah satu langkah strategis untuk menjembatani kesenjangan antara dunia pendidikan dan dunia industri. Program ini memberi kesempatan kepada mahasiswa untuk memperoleh pengalaman langsung di dunia kerja, sehingga mahasiswa dapat menerapkan ilmu yang telah dipelajari dalam konteks industri yang sesungguhnya (Kemendikbud, 2023). Salah satu bentuk pengalaman langsung ini adalah magang mandiri di perusahaan besar, yaitu Telkomsel. PT Telkomsel merupakan perusahaan operator telekomunikasi seluler di Indonesia. Pada tahun 1993, PT Telkom mulai mengembangkan teknologi nirkabel berbasis GSM. Setahun kemudian, pada tahun 1994, PT Satelit Palapa Indonesia, sebagai operator jaringan GSM pertama di Indonesia, meluncurkan kartu SIM untuk pertama kalinya (Fadlullah, 2018). Melalui program MBKM Magang Madiri ini, mahasiswa tidak hanya bisa mengasah keterampilan teknis di bidang teknologi informasi dan komunikasi, tetapi juga meningkatkan keterampilan non-teknis atau *soft skills*, seperti pemecahan masalah, analisis data, serta kemampuan berkolaborasi dalam tim multidisiplin yang diperlukan dalam dunia kerja yang serba cepat dan kompetitif saat ini.

Pada kesempatan magang mandiri ini, penulis berkesempatan untuk bekerja di Telkomsel Branch Surabaya, yang memberikan pengalaman yang sangat berharga. Selama magang, penulis terlibat dalam berbagai proyek penting yang berkaitan dengan analisis data tenaga kerja, segmentasi pelanggan, serta pengoptimalan strategi pemasaran yang disesuaikan dengan perilaku dan karakteristik pelanggan. Salah satu proyek utama yang dikerjakan adalah analisis deret waktu untuk memprediksi kunjungan Telkomsel ke sekolah. Kunjungan Telkomsel ke sekolah merupakan salah satu strategi penting dalam mendekatkan layanan dan produk kepada generasi muda sebagai target pasar potensial. Dengan memanfaatkan data historis kunjungan, analisis tren menggunakan model ARIMA dapat membantu Telkomsel dalam merancang strategi pemasaran yang lebih efektif dan efisien. Prediksi tren ini memberikan gambaran mengenai pola kunjungan di masa depan, memungkinkan perusahaan untuk mengoptimalkan alokasi sumber daya, seperti tenaga kerja, waktu, dan biaya operasional, serta menentukan waktu dan lokasi kunjungan yang strategis. Selain itu, pendekatan berbasis data ini mendukung transformasi digital Telkomsel, memberikan landasan ilmiah dalam pengambilan keputusan, serta menunjukkan pemanfaatan teknologi analisis data untuk meningkatkan efektivitas program edukasi digital di sekolah. Proyek ini juga memiliki kontribusi akademik dengan mengintegrasikan metode analisis deret waktu yang andal ke dalam konteks bisnis telekomunikasi, menjadikannya sebagai referensi untuk inisiatif serupa di masa mendatang.

Dalam industri telekomunikasi yang semakin kompetitif, memahami tren kunjungan ke sekolah menjadi isu mutakhir yang relevan untuk menjaga hubungan dengan segmen pelanggan muda. Sebagai salah satu program strategis Telkomsel, kegiatan ini menghadapi tantangan seperti fluktuasi kebutuhan pelanggan, jadwal pendidikan yang dinamis, serta pengelolaan sumber daya yang terbatas. Tanpa prediksi yang akurat, risiko inefisiensi dalam alokasi waktu dan tenaga kerja dapat meningkat, berpotensi mengurangi dampak kunjungan tersebut. Selain itu, keunikan proyek ini terletak pada penerapan analisis deret waktu, seperti model ARIMA, dalam konteks yang jarang diangkat di industri telekomunikasi. Hal ini menunjukkan upaya Telkomsel untuk

tidak hanya beradaptasi terhadap kebutuhan pasar tetapi juga mengantisipasi perubahan dengan pendekatan berbasis data.

1.2. Tujuan Magang Mandiri MBKM

Penulis mengikuti program magang mandiri MBKM dengan tujuan untuk memberikan platform bagi penulis sebagai mahasiswa magang agar dapat memahami dan menerapkan ilmu serta keterampilan yang telah dipelajari selama kuliah. Adapun tujuan penulis dalam mengikuti magang mandiri di Telkomsel Branch Surabaya adalah sebagai berikut:

1. Untuk mendapatkan pemahaman yang mendalam tentang divisi *data analyst* di Telkomsel dengan menganalisis data kunjungan sekolah untuk menghasilkan *insight* yang mendukung strategi bisnis Perusahaan.
2. Untuk menerapkan ilmu dan keterampilan yang telah dipelajari di Program Studi Sains Data, serta mengimplementasikan secara langsung ke dunia kerja di Telkomsel
3. Untuk melatih dan meningkatkan kemampuan soft skills seperti pemecahan masalah secara efektif, komunikasi, dan kerja tim dalam menangani tantangan nyata yang dihadapi oleh Perusahaan.
4. Untuk mengidentifikasi pola kunjungan ke sekolah berdasarkan tingkat pendidikan dan periode waktu menggunakan metode ARIMA

1.3. Manfaat Magang Mandiri MBKM

a. Manfaat bagi UPN “Veteran" Jawa timur
   1) Meningkatkan reputasi universitas di dunia industri telekomunikasi dan Masyarakat luas
   2) Membantu mengembangkan pengetahuan, pengalaman mahasiswa, dan meningkatkan kualitas Pendidikan di universitas
b. Manfaat bagi Telkomsel sebagai Mitra Penyelenggara MBKM
   1) Mendapatkan timbal balik dari mahasiswa dalam bentuk kontribusi pada tugas & proyek yang berfokus pada pengolahan data untuk mengoptimalkan distribusi kunjungan berdasarkan kebutuhan di tiap tingkat pendidikan.
   2) Terjalinnya hubungan & relasi yang kuat dengan universitas untuk potensi kolaborasi dalam penenlitian, pengembangan teknologi, dan inovasi Perusahaan.
   3) Memperoleh akses untuk mengenali dan merekrut bakat-bakat dari mahasiswa yang memiliki potensi menjadi calon karyawan di masa depan
c. Manfaat bagi Mahasiswa
   1) Mendapatkan pengalaman kerja nyata di lingkungan industri telekomunikasi, khususnya di Telkomsel.
   2) Memperolah wawasan yang luas dan keterampilan praktis sebagai *data analyst* dalam menganalisis data pelanggan dan menudukung strategi bisnis
   3) Memiliki kesempatana untuk menjalin relasi professional dengan pemangku kepentingan di Telkomsel, yang dapat bermanfaat untuk pengembangan karier di masa mendatang.

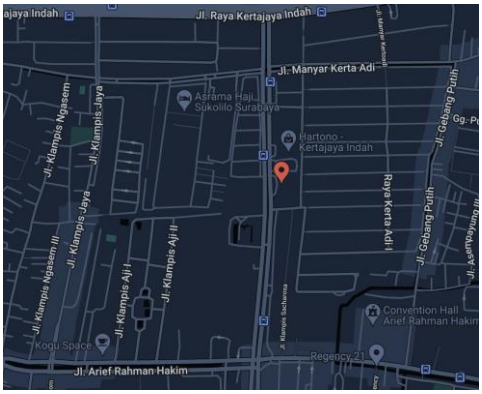
1.4. Mahasiswa Menguraikann Tujuan Penulisan Topik Magang Mandiri MBKM

Tujuan penulisan topik Magang Mandiri di Telkomsel yang berjudul **“ PEMODELAN DERET WAKTU DATA KUNJUNGAN TELKOMSEL KE SEKOLAH MENGGUNAKAN ARIMA”** adalah untuk mengembangkan kemampuan analisis data yang berfokus pada pemodelan analisis deret waktu menggunakan metode ARIMA melibatkan berbagai langkah, termasuk pemahaman teori dasar analisis deret waktu, pengumpulan dan pra-pemrosesan data historis terkait, eksplorasi pola musiman atau tren jangka panjang, penerapan model statistik yang relevan untuk menganalisis data, serta evaluasi kinerja model melalui metrik seperti MAPE, RMSE, dan MAE.

BAB 2. LOKASI PROYEK MAGANG MANDIRI

2.1. Lokasi & sejarah Mitra Magang Mandiri

Telkomsel merupakan Perusahaan telekomunikasi digital terdepan yang selama 28 tahun senantiasa melakukan transformasi digital dengan menghadirkan *Digital Connectivity*, *Digital Platfrom*, dan *Digital Services* yang dapat diandalkan dan mengutamakan kemanfaatan teknologi demi pertumbuhan setiap aspek kehidupan seluruh lapisan Masyarakat Indonesia. Pada 18 Juni 2021, Telkomsel menghadirkan identitas baru Perusahaan sebagai simbol perubahan yang semakin mengukuhkan komitmennya menjadi Perusahaan digital terdepan yang akan terus menjadi *enabler* dalam berkelanjutan untuk #BukaSemuaPeluang bagi Masyarakat dalam mendorong kemajuan bangsa. Kantor Telkomsel Branch Surabaya terletak di Jl. Dr. Ir. H. Soekarno No.175, Klampis Ngasem, Kecamatan Sukolilo, Surabaya, Jawa Timur 60116. Dengan titik koordinat -7.2845642520281055, 112.78127579853356



Gambar 2. 1 Lokasi Telkomsel Branch Surabaya

Dibawah ini adalah sejarah perusahaan Telkomsel

Gambar 2. 2 Logo Lama Telkomsel

Komponen warna logo lama tersebut, yaitu:

a. Bentuk segi enam melambangkan sistem seluler GSM.
b. Elips horizontal melambangkan jasa komunikasi darat yang ditawarkan Telkomsel, dan juga mewakili Telkom sebagai pemegang saham utama dan penyelenggara telekomunikasi domestik.
c. Elips vertikal menunjukkan bahwa komunikasi Telkomsel menggunakan gelombang udara dan juga mewakili Indosat sebagai pendiri dan pemegang saham utama lainnya (sebelum 2001) dan pelaksana telekomunikasi internasional terkemuka di Indonesia. Kedua elips tersebut berpotongan di atas segi enam berwarna merah dan abu-abu, serta membentuk huruf "T" berwarna putih yang merupakan huruf pertama dari Telkomsel.
d. Warna merah pada bagian atas segi enam melambangkan keteguhan, keyakinan dan kesiapan Telkomsel dalam menghadapi masa depan pertelekomunikasian.
e. Segi enam yang berwarna abu-abu, menyatakan komitmen Telkomsel terhadap pelanggan dan dukungan kuat dari perusahaan induknya.
f. Warna abu-abu yang merupakan warna logam, melambangkan teknologi dan kestabilan.

g. Warna putih, berarti keterbukaan dan etika yang tak perlu diragukan lagi.

Setelah menggunakan logo tersebut selama 25 tahun

Gambar 2. 3 Logo Telkomsel

Pada perayaan ulang tahunnya yang ke-26 tanggal 18 Juni 2021, Telkomsel resmi menyandang logo baru yang lebih sederhana. Peluncuran logo baru tersebut dimaknai sebagai simbol perubahan berupa transformasi digital dan nilai-nilai baru, yang diharapkan bisa berperan bagi kemajuan bangsa. Adapun logo baru Telkomsel yang merupakan hasil desain Deni Anggara dari Degarism Studio Bandung ini disebutkan terinspirasi dari batik, yang tercermin dari nama font yang digunakan yaitu Telkomsel Batik. Inspirasi dari batik juga terlihat dari adanya "portal" dengan guratan, yang bermakna hal-hal baru dan dunia baru yang bisa dijelajahi bersama Telkomsel. Portal itu juga terlihat dikelilingi oleh sinar, yaitu sinar optimisme. Adapun komponen warna logo baru tersebut, yaitu:

a. Warna merah dan putih sebagai warna dasar pada logo baru, mewakili keberanian, kekuatan, kemurnian, dan kejujuran.
b. Aksen warna kuning, melambangkan optimisme dan rasa senang.
c. Untuk komunikasi perusahaan juga dikenalkan tone warna-warna baru, termasuk biru tua yang melambangkan integritas dan kepercayaan (Kompas.com, 2021).

PT Telekomunikasi Selular atau biasa disingkat menjadi Telkomsel, adalah anak usaha Telkom Indonesia yang bergerak di bidang telekomunikasi bagi konsumen individu (hingga tahun 2023 hanya membidangi telekomunikasi seluler). Untuk mendukung kegiatan bisnisnya, hingga akhir tahun 2021, perusahaan ini memiliki 397 gerai GraPARI yang tersebar di seantero Indonesia.

Perusahaan ini adalah operator telekomunikasi seluler terbesar di Indonesia, dengan mengoperasikan 236.000 unit BTS untuk melayani lebih dari 170 juta pelanggan, sehingga perusahaan ini menguasai hampir 60% pangsa pasar telekomunikasi seluler di Indonesia pada tahun 2020. Dengan capaian tersebut, perusahaan ini juga menjadi operator seluler terbesar keenam di dunia. Bekerja sama dengan 575 mitra roaming, layanan perusahaan ini kini dapat digunakan di 200 negara di seluruh dunia.

Perusahaan ini beroperasi dengan menggunakan teknologi GSM, 4G LTE, dan 5G di frekuensi 900/1800 MHz, dan sebelumnya juga pernah mengoperasikan jaringan 3G hingga Mei 2023.

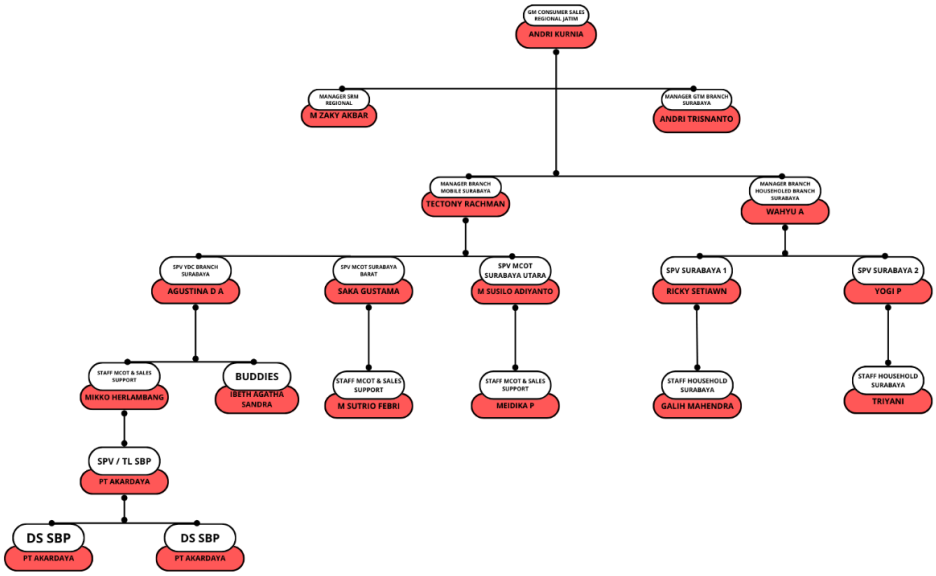
Gambar 2. 4 Sejarah Telkomsel

Tabel 2. 1 Sejarah Telkomsel

Tahun	Keterangan
1995	Resmi berdiri dengan layanan karuHalo
1997	Merilis simPati dan telah melayani seluruh provinsi di Indonesia
2003	Memperkenalkan layanan roaming prabayar pertama di Indonesia
2006	Menghadirkan jaringan 3G pertama di Indonesia
2011	Memimpin pasar Indonesia dengan melayani 100 juta pelanggan
2014	Menggelar 4G LTE pertama di Indonesia
2018	Uji coba 5G pertama di Indonesia melalui 5G *Experience Center Asian Games*
2019	Uji coba 5G pertama untuk kalangan industry di Indonesia dan video call 5G
2020	*Strategic investment* di sejumlah startup, termasuk GoJek
2021	Mengenalkan identitas baru diiringi penguatan ekosistem digital yang inklusif dan berkelanjutan di Indonesia
2022	Menghadirkan penerapan teknologi 5G *Underground Smart-Mining* pertama di Asia Tenggara
2023	Integrasi IndiHome ke Telkomsel melalui inisiatif FMC dan peluncuran produk Telkomsel One

2.2. Struktur Organisasi Mitra Magang Mandiri

Struktur organisasi adalah kerangka kerja formal yang menggambarkan pembagian tugas dan tanggung jawab di antara anggota perusahaan. Selain itu, struktur ini juga mencerminkan hubungan kerja sama antar pihak dalam organisasi yang bertujuan untuk mencapai sasaran bersama.



Gambar 2. 5 Struktur Organisasi Telkomsel Branch Surabaya

Berikut merupakan tugas dan tanggung jawab mitra magang mandiri Telkomsel Branch Surabaya :

Tabel 2. 2 Struktur Organisasi

Jabatan	Tugas dan Tanggung Jawab
GM Consumer Sales Regional	- Merencanakan strategi penjualan regional untuk mencapai target perusahaan. - Memimpin tim penjualan regional dan mengalokasikan sumber daya. - Mengidentifikasi peluang untuk memperluas pangsa pasar. - Berkolaborasi dengan distributor, mitra, dan stakeholders di regional Jatim.
Manager SRM Regional	- Merancang strategi untuk mengoptimalkan penjualan di berbagai kanal distribusi. - Mengelola hubungan dengan distributor dan mitra penjualan. - Menganalisis data untuk evaluasi kinerja penjualan. - Berkoordinasi dengan tim produk, pemasaran, distribusi, dan layanan pelanggan.
Manager GTM Branch Surabaya	- Menyusun dan menjalankan strategi peluncuran produk di branch Surabaya. - Melakukan analisis pasar dan kebutuhan pelanggan. - Mengelola kampanye pemasaran dan promosi lokal. - Memastikan aktivitas GTM selaras dengan target penjualan.
Manager Branch Mobile Surabaya	- Mengembangkan program penjualan untuk penetrasi layanan mobile. - Berkolaborasi dengan mitra dan distributor untuk meningkatkan penjualan. - Menganalisis data untuk menyusun strategi berbasis data. - Memastikan pencapaian target penjualan mobile.
Manager Household Branch Surabaya	- Merancang strategi dan program untuk segmen pelanggan household. - Meningkatkan penetrasi layanan rumah tangga (broadband, digital). - Mengelola loyalitas pelanggan dan mengawasi operasional penjualan.

	- Melakukan analisis dan evaluasi data pelanggan household.
SPV YDC Branch Surabaya	- Mengelola program untuk segmen Youth, Device, dan Community. - Merancang program pemasaran dan promosi sesuai karakteristik lokal. - Membangun kerja sama dengan mitra, komunitas, dan institusi lokal. - Memantau kinerja program dan memberikan edukasi ke komunitas.
SPV MCOT Surabaya Barat	- Mengelola aktivitas penjualan di kanal modern dan online. - Menyusun strategi kanal modern untuk wilayah Surabaya Barat. - Berkoordinasi dengan mitra dan distributor untuk distribusi produk. - Menganalisis data penjualan dan mengevaluasi program MCOT.
SPV MCOT Surabaya Utara	- Memimpin program penjualan di kanal modern dan online untuk wilayah Surabaya Utara. - Berkolaborasi dengan mitra lokal untuk mendukung program MCOT. - Menyusun laporan kinerja dan memberikan edukasi kepada mitra. - Meningkatkan daya saing produk di pasar kanal modern.
SPV Surabaya 1 & 2	- Mengawasi tim penjualan dan memberikan arahan untuk pencapaian target. - Melaksanakan strategi penjualan sesuai karakteristik wilayah. - Menyelesaikan kendala operasional dan memastikan distribusi produk berjalan optimal. - Membuat laporan kinerja wilayah masing-masing.
Staff MCOT & Sales Support	- Mendukung operasional penjualan di kanal modern dan transaksi online. - Menyusun laporan penjualan dan menganalisis performa kanal MCOT. - Berkoordinasi dengan tim internal dan mitra untuk distribusi produk. - Membantu pemantauan program promosi MCOT.
Staff Household Surabaya	- Mengelola data pelanggan household dan memantau program pemasaran. - Menyusun laporan kinerja pelanggan dan analisis kebutuhan pelanggan household. - Berkoordinasi dengan tim terkait untuk meningkatkan engagement. - Mendukung eksekusi program khusus untuk segmen household.
SPV / TL SBP	- Memimpin tim SBP untuk pencapaian target penjualan. - Mengelola hubungan dengan mitra bisnis SBP. - Memantau dan menganalisis kinerja tim penjualan. - Memberikan pelatihan dan pengembangan tim SBP.
DS SBP	- Mengelola data penjualan dan menyusun laporan kinerja tim SBP. - Melakukan analisis penjualan untuk mendukung strategi. - Berkoordinasi dengan tim penjualan dan marketing terkait kebutuhan data. - Memastikan kualitas dan akurasi data penjualan SBP.

2.3. Visi dan Misi Telkomsel

- Visi Telkomsel

  Menjadi penyedia layanan dan Solusi gaya hidup digital *mobile* kelas dunia yang terpecaya.
- Misi Telkomsel

Memberikan layanan dan Solusi digital *mobile* yang melebihi ekspetasi para pengguna, menciptakan nilai lebih bagi para pemegang saham, serta mendukung pertumbuhan ekonomi bangsa.

2.4. Kegiatan Produksi (barang/jasa)

Di Telkomsel, kegiatan produksi barang untuk membantu masyarakat dalam memenuhi kebutuhan terkait layanan telekomunikasi bertujuan untuk menghasilkan produk berkualitas yang dapat memenuhi kebutuhan pelanggan antara lain :

Tabel 2. 3 Kegiatan Produksi

No	Gambar Barang	Produksi Barang	Deskripsi Produk
1		Telkomsel One	Paket bundling yang menyediakan layanan internet rumah IndiHome, internet seluler Telkomsel, dan akses ke berbagai layanan digital. Paket ini dirancang untuk memberikan solusi konektivitas lengkap bagi pelanggan Telkomsel. Telkomsel One diluncurkan pada Juli 2023 sebagai wujud komitmen Telkomsel untuk mengimplementasikan Fixed Mobile Convergence (FMC). Layanan ini mempermudah pengguna mendapatkan layanan internet kabel dan seluler dalam satu paket.
2		Telkomsel Prabayar	Kartu perdana Telkomsel yang menggunakan sistem prabayar, yaitu sistem pembayaran di mana pengguna harus membayar terlebih dahulu sebelum menggunakan layanan. Telkomsel juga memiliki produk prabayar lainnya, seperti KARTU As, simPATI, LOOP, dan Kartu Facebook.
3		Telkomsel Lite	Kartu perdana Telkomsel yang menawarkan paket internet bulanan dengan harga terjangkau dan kuota internet hingga 60 GB. Kartu ini dirancang untuk segmen pekerja dan rumah tangga yang ingin berhemat dan menghemat pengeluaran internet.
4		Telkomsel Halo	Layanan pascabayar dari Telkomsel yang menawarkan berbagai keuntungan dan fasilitas kepada penggunanya. Berikut beberapa hal yang bisa didapatkan dengan menggunakan Telkomsel Halo: Kuota internet yang besar, Telepon dan SMS tanpa batas, Akses ke aplikasi hiburan, Layanan prioritas, Keuntungan roaming. Telkomsel Halo bisa digunakan kapan saja dan di semua jaringan Telkomsel (2G, 3G, 4G, dan 5G). Pengguna Telkomsel Halo

			akan menerima tagihan setiap bulannya pada akhir periode tagihan. Tagihan tersebut dapat dibayarkan melalui berbagai channel, seperti GraPARI, Digital Wallet, MyGraPARI, MyTelkomsel, dan ATM/Internet Banking yang bekerja sama dengan Telkomsel. Untuk membatasi pemakaian, pengguna Telkomsel Halo dapat mengaktifkan layanan pembatasan pemakaian (CLS/Credit Limit Service). Telkomsel Halo sebelumnya bernama kartu HALO hingga 18 Juni 2021.
5		by.U	Kartu internet serba digital milik Telkomsel yang menawarkan layanan seluler berbasis digital. by.U merupakan operator jaringan seluler prabayar digital pertama di Indonesia.
6		Telkomsel Orbit	Layanan internet rumah yang menggunakan modem WiFi dan jaringan seluler 4G LTE.
7		Eznet	Salah satu layanan broadband dari Telkomsel yang menghadirkan jaringan fixed, guna memberikan pengalaman Internet di rumah yang tejangkau bagi pelanggan. Saat ini layanan yang tersedia adalah layanan internet dengan kecepatan hingga 10 Mbps.
8		Eznet Wireless	Layanan internet rumah terjangkau dengan modem WiFi yang menawarkan pilihan non-fiber serta kuota tanpa batas dengan tagihan bulanan yang praktis.
9	MAX STREAM	MAXstream	Aplikasi video yang menampilkan ribuan film dan serial TV dari MAXstream Original, HBO GO, MyPlay, NOMO, Starvision, Genflix, Sushiroll, Vidio, Vision+, RCTI+, VIU, WeTV iFlix, Mola, dan lainnya.
10	DG DUNIA GAMES	Dunia Games	Layanan yang berisi pembahasan seputar games di mana pengguna dapat mengunduh game serta membeli kupon transaksi di dalam permainan online seluler maupun komputer.

11		LinkAja	Aplikasi layanan keuangan digital, sebelumnya bernama Telkomsel Cash atau TCASH. Didirikan oleh Telkomsel bersama anggota Badan Usaha Milik Negara yaitu Bank Mandiri, Bank Negara Indonesia, Bank Rakyat Indonesia, Bank Tabungan Negara, Pertamina, Asuransi Jiwasraya, dan Danareksa.
12		GraPARI	Gerai-gerai Telkomsel yang melayani pelanggan secara langsung dinamakan "GraPARI", Saat ini, terdapat ratusan gerai GraPARI yang tersebar di berbagai kabupaten/kota dan provinsi di Indonesia. Selain bertindak sebagai pusat layanan pelanggan yang memfokuskan diri untuk melayani para pelanggannya, GraPARI juga menyediakan penjualan produk sendiri seperti Telkomsel Halo dan Telkomsel PraBayar
13		Veronika	Telkomsel juga melayani penggunanya lewat "asisten virtual" bernama Veronika yang telah melayani pelanggan sejak 24 Agustus 2017. Veronika asisten virtual adalah sebuah sistem aplikasi komputer yang dirancang dengan algoritma jaringan syaraf tiruan dan artificial intelligence atau kecerdasan buatan untuk melaksanakan komunikasi dua arah kepada pelanggan dan memberikan informasi yang tepat dan akurat melalui platform perpesanan (messenger) dan mampu melaksanakan tugas untuk mengeksekusi sebuah perintah input seperti mengecek sisa kuota, menukar Telkomsel POIN, menemukan lokasi GraPARI terdekat, membeli paket, dan informasi lain seputar layanan dan produk dari telkomsel.

BAB 3. PELAKSANAAN MAGANG MANDIRI

3.1. Posisi / Kedudukan Magang Mandiri

Selama menjalani masa magang mandiri di Telkomsel Branch Surabaya, saya menempati posisi sebagai *Data Analyst* di bawah bimbingan mentor kami, Kak Febri. Peran saya sebagai seorang *Data Analyst* adalah untuk mengelola data perusahaan, terutama data yang berkaitan dengan operasional dan strategi bisnis di branch Surabaya Tugas utama saya mencakup pembuatan dashboard operasional menggunakan Looker Data Studio, mulai dari menghubungkan data, membersihkan dan mentransformasi data, hingga memilih visualisasi yang relevan. Dashboard menampilkan data performa jaringan dan analisis trafik dengan filter interaktif untuk mendukung pengambilan keputusan operasional. Dalam menjalankan peran ini, saya secara aktif berkolaborasi dengan tim untuk memahami kebutuhan bisnis dan mencari solusi yang efektif. Saya mengerjakan berbagai tugas yang diberikan oleh mentor, mulai dari tugas harian, mingguan, hingga tugas bulanan, serta berkontribusi dalam final project yang bertujuan memberikan nilai tambah nyata bagi perusahaan. Selama praktik ini, saya berfokus pada pengembangan dashboard interaktif yang menjadi salah satu proyek utama. Dashboard tersebut dirancang untuk menampilkan informasi penting terkait sales dan operation, sehingga dapat diakses dengan mudah oleh tim internal perusahaan. Informasi yang ditampilkan dalam dashboard mencakup data dalam bentuk grafik, diagram, atau tabel, yang dirancang agar mudah dipahami oleh semua pihak terkait. Proses pembuatan dashboard melibatkan pengolahan data yang mendalam, pembersihan data, hingga pemilihan desain visualisasi yang paling sesuai untuk menyampaikan informasi. Hal ini memastikan bahwa dashboard mampu memberikan wawasan yang jelas dan relevan, sekaligus mendukung efisiensi kerja.

Tabel 3. 1 Posisi Magang

No	Nama	Asal Universitas	Jobdesk	Deskripsi Jobdesk
1	Erika Fatimatul Hidayanti	Universitas Pembangunan Nasional "Veteran" Jawa Timur	Membuat Dashboard Sales	Mengembangkan dashboard sales dengan Looker Data Studio, termasuk menghubungkan dan membersihkan data serta memilih visualisasi yang sesuai. Dashboard menampilkan pencapaian penjualan dan analisis pelanggan dengan filter interaktif berdasarkan waktu, lokasi, dan PIC.
2	Tsabita Rosyidah Putri	Universitas Pembangunan Nasional "Veteran" Jawa Timur	Membuat Dashboard Operation	Bertanggung jawab atas pembuatan dashboard operation menggunakan Looker Data Studio, mulai dari menghubungkan data, membersihkan dan mentransformasi data, hingga memilih visualisasi yang relevan. Dashboard menampilkan data performa jaringan dan analisis trafik dengan filter interaktif untuk mendukung pengambilan keputusan operasional.

3.2. Metodologi Penyelesain Tugas

3.2.1. Tinjauan Pustaka

- Analisis Deret Waktu

  Analisis deret waktu adalah suatu pengamatan yang dibangun berurutan dalam waktu. Analisis ini dilakukan untuk memperoleh pola data deret waktu, dengan menggunakan pengamatan sebelumnya untuk memprediksi suatu nilai pada masa yang akan datang. Data yang dikumpulkan secara periodik berdasarkan urutan waktu, baik dalam jam, hari, minggu, bulan, maupun dalam tahun. Lebih lanjut, akan dilakukan prediksi (*forecasting*) melibatkan pengamatan data masa lalu dan menempatkannya kemasa mendatang dalam suatu bentuk matematis (Atman Maulana, 2018).

- Bahasa Pemograman Python

  Python adalah bahasa pemrograman yang berbasis skrip dan berorientasi objek. Bahasa ini dirancang untuk pengembangan perangkat lunak dan memiliki fleksibilitas untuk dijalankan pada berbagai sistem operasi, menjadikannya sangat serbaguna. Saat ini, Python sangat populer di kalangan profesional dan praktisi di bidang data science dan analisis data. Kepopulerannya ini disebabkan oleh banyaknya pustaka (*library*) yang tersedia, yang menawarkan berbagai fungsi untuk analisis data, alat pemrosesan data (data preprocessing), serta model *machine learning*. Selain itu, Python juga memiliki berbagai pustaka visualisasi data yang memungkinkan pengguna untuk membuat grafik dan diagram untuk mempermudah pemahaman informasi yang terkandung dalam data (Edwardo, 2018).

- Sekolah

  Sekolah adalah suatu lembaga yang diselenggarakan baik oleh pemerintah maupun swasta, yang berfungsi sebagai wadah pendidikan dalam berbagai bentuk, baik formal, non-formal, maupun informal. Pendirian sekolah dapat dilakukan oleh negara ataupun pihak swasta dengan tujuan utama untuk memberikan pengajaran dan pendidikan kepada para murid. Melalui lembaga ini, pendidikan dikelola dan disampaikan oleh pendidik atau guru yang bertugas memberikan bimbingan serta arahan kepada murid agar dapat berkembang secara optimal baik dalam aspek akademis maupun karakter (R. Oktavian dan R. F. Aldya, 2020).

- Prediksi

  Prediksi merupakan suatu proses yang dilakukan untuk memperkirakan secara sistematis kemungkinan peristiwa atau kejadian yang akan terjadi di masa depan. Proses ini didasarkan pada analisis informasi yang tersedia dari masa lalu dan keadaan saat ini. Tujuan dari prediksi adalah untuk menghasilkan estimasi yang akurat dengan meminimalkan kesalahan, yaitu selisih antara hasil yang diperkirakan dan kenyataan yang terjadi di kemudian hari. Dengan kata lain, prediksi bertujuan untuk menghasilkan perkiraan yang lebih tepat agar keputusan yang diambil berdasarkan hasil perkiraan tersebut dapat lebih efektif dan optimal (Adiguno dkk., 2022)

- Pemeriksaan Stasioner Deret Waktu

- Uji *Augmented Dickey-Fuller* (ADF)
  Uji ADF digunakan untuk menguji stasioneritas data dalam mean. Menurut Juanda dan Junaidi (2012) model pada ADF test untuk pengujian korelasi serial antara residual dengan ΔZ_t dapat dinyatakan dalam bentuk umum proses *autoregresif* dinyatakan pada persamaan (1) sebagai berikut:
  $\Delta Z_t = \beta_1 + \beta_2 t + \delta Z_{t-1} + \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \cdots + \alpha_1 Z_{t-2} + \varepsilon_t$ (1),

  Dimana ε_t adalah proses yang white noise berdistribusi normal N(0,σ^2) dan $\Delta Z_{t-1} = (Z_{t-1} - Z_{t-2})$
  Hipotesis :
  $H_0 : \delta = 0$ (Data tidak stasioner terhadap mean (terdapat akar unit))
  $H_1 : \delta < 0$ (Data stasioner terhadap mean)

  Taraf signifikan : α = 0,05
  Statistik Uji t dinyatakan pada persamaan (2) sebagai berikut :
  $$t = \frac{\hat{\delta}}{SE\ (\hat{\delta})} \quad (2)$$
  Dimana :
  $\hat{\delta}$: estimasi koefisien $\hat{\delta}$ dari model (Z_{t-1})
  $SE\ (\hat{\delta})$: kesalahan standar dari estimasi $\hat{\delta}$
  Kriteria Uji :
  H_0 ditolak jika ADF < nilai kritis Augmented Dickey Fuller atau $p - value < \alpha$

- KPSS (*Kwiatkowski-Phillips-Schmidt-Shin*)
  Selain uji ADF dan uji PP untuk menentukan data stasioner atau tidak juga dapat dilakukan dengan uji KPSS yang kenalkan oleh *Kwiatkowski Phillips Schmidt Shin* (Zivot, E dan Wang, J. 2005). Persamaan uji stasioner dengan analisis KPSS pada persamaan (3) sebagai berikut :

  $$Z_t = \sigma_0 + \varepsilon_t \quad (3),$$
  Dimana :
  Z_t : Nilai variabel pada waktu ke-t

  Hipotesis :
  H_0 : Data stasioner
  H_1 : Data tidak stasioner

  Kriteria uji
  Jika $|t|$ < nilai mutlak nilai kritik Mackinnon, maka H_0 diterima yang berarti data tersebut stasioner.

- Transformasi Box-Cox
  Wei (2006) menjelaskan bahwa transformasi Box-Cox adalah transformasi pangkat pada respon. Box-Cox dengan mempertimbangkan kelas transformasi berpangkat tunggal, yaitu λ yang dapat dipangkatkan pada variabel respon Z_t, sehingga transformasinya menjadi Z_t^{λ}, dimana λ adalah parameter diduga, yang menentukan jenis transformasi yang diterapkan. Nilai λ dipilih untuk memaksimalkan normalitas dan kestabilan varians data. Dengan bentuk umum berikut yang dinyatakan pada persamaan (4) berikut :

$$Z_t\lambda = \frac{Z_t^{\lambda} - 1}{\lambda}, \lambda \neq 0 \quad (4),$$

Untuk $\lambda = 0$, rumus transformasi ini menjadi algoritma dari data

$$Z_t(\lambda = 0) = \log(Z_t)$$

Dimana :

Z_t : nilai asli dari variabel respon pada waktu t

- *Autocorrelation Function* (ACF)

  Menurut Makridakis et al. (1999) istilah autokorelasi digunakan untuk menjelaskan asosiasi atau ketergantungan bersama (*mutual dependence*) antara nilai - nilai suatu deret berkala yang sama pada periode berlainan. Autokorelasi pada lag k, didefinisikan pada persamaan (5) sebagai berikut:

$$\rho_{\mathrm{k}} = \frac{kov(Z_t, Z_{t-k})}{[Var(Z_t).Var(Z_{t-k})]^{1/2}} = \frac{\gamma_k}{\gamma_0} \quad (5),$$

  Dimana :

  ρ_{k} : Autokorelasi pada lag-k

  $kov(Z_t, Z_{t-k})$: $E[(Z_t - \mu)(Z_{t-k} - \mu)$

- *Partial Autocorrelation Function* (PACF)

  Menurut Makridakis et al. (1999) ukuran *partial autocorrelation* digunakan untuk menunjukkan besarnya hubungan antara nilai suatu variabel saat ini dengan nilai variabel sebelumnya dari variabel yang sama (nilai-nilai untuk berbagai kelambatan waktu) dengan menganggap pengaruh dari semua kelambatan waktu lainnya adalah konstan. *Partial autocorrelation* didefinisikan pada persamaan (6) sebagai berikut:

$$\phi_{k,k} = \frac{Cov(Z_t, Z_{t-k} | Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k+1})}{\sqrt{Var}(Z_t | Z_{t-1}, \ldots, Z_{t-k+1}).\sqrt{Var}(Z_{t-k} | Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k+1})} \quad (6),$$

  Dimana :

  $\phi_{k,k}$: Nilai PACF pada lag ke-k, menunjukkan korelasi parsial antara Z_t dan Z_{t-k} setelah mengeliminasi efek dari semua lags di antaranya. Nilainya berkisar antara -1 dan 1.

- *AutoRegressive* (AR)

  Autoregressive (AR) adalah sebuah model statistik yang digunakan untuk menganalisis deret waktu, yang mengasumsikan bahwa nilai data pada periode waktu saat ini dipengaruhi oleh nilai data pada periode sebelumnya. Dalam model ini, perubahan atau pola yang terjadi dalam data masa lalu digunakan untuk memprediksi nilai di masa depan. Dengan kata lain, model *autoregressive* menganggap bahwa hubungan antara pengamatan saat ini dan pengamatan sebelumnya sangat penting untuk memperkirakan nilai yang akan datang. Model AR sering kali digunakan dalam analisis deret waktu untuk memodelkan data yang menunjukkan pola atau ketergantungan temporal (Fitriani dkk., 2013). Model *Autoregressive* (AR) yang memiliki ordo p dinotasikan dengan AR(p) atau model ARIMA(p,0,0), ada beberapa nilai tertentu dari X_t berfungsi sebagai variabel dependen. Variabel-variabel ini memiliki nilai lag, yang berarti output

masa lalu atau saat ini akan mempengaruhi hasil di masa depan. Bentuk umum model *autoregresif* dinyatakan pada persamaan (7) :

$$X_t = C + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \cdots + \phi_p X_{t-p} + e_t \quad (7),$$

Dimana :

C : Nilai rata – rata (konstanta) dari data

X_{t-p} : Nilai data pada waktu $(t-p)$

X_t : Nilai data pada waktu t (yang diprediksi).

ϕ_p : Parameter *autoregresif* ke-p atau Koefisien yang menunjukkan seberapa besar pengaruh nilai sebelumnya (X_{t-p}) terhadap nilai saat ini (X_t)

e_t : Nilai *error* atau kesalahan pada waktu t

p : Jumlah lag

- *Moving Average* (MA)

  Moving Average (MA) adalah salah satu model dalam analisis deret waktu yang berfokus pada pola pergerakan residual atau kesalahan prediksi dari masa lalu. Model ini mengasumsikan bahwa nilai data pada periode saat ini dipengaruhi oleh rata-rata kesalahan prediksi yang terjadi dalam beberapa periode sebelumnya. Residual di sini mengacu pada selisih antara nilai yang sebenarnya dengan nilai yang diperkirakan oleh model. Model MA digunakan untuk menangkap pola fluktuasi jangka pendek dalam data deret waktu yang tidak dapat dijelaskan sepenuhnya oleh model tren atau komponen lainnya. Dengan menggabungkan informasi dari residual masa lalu, model ini membantu mengurangi pengaruh fluktuasi acak dan menghasilkan prediksi yang lebih stabil (Nur Komara & Dwi Agustin Nuriani Sirodj, 2023). Model MA yang memiliki ordo q dinotasikan dengan MA(q) atau ARIMA (0,0,q. Bentuk umum model *Moving Average* dinyatakan pada persamaan (8) sebagai berikut:

$$X_t = C + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \theta_3 e_{t-3} + \cdots + \theta_q e_{t-q} + e_t \quad (8),$$

  Dimana :

  X_t : Nilai data pada waktu t (yang diprediksi)

  C : Nilai rata – rata (konstanta) dari data

  θ_q : Parameter *Moving Average* atau koefesien yang menunjukkan pengaruh *error* masa lalu terhadap nilai X_t

  e_t : *Error* saat ini

  e_{t-q} : Nilai kesalahan pada $(t-q)$

- ARMA (*AutoRegressive Moving Average)*

  Autoregressive Moving Average (ARMA) adalah salah satu metode yang sering digunakan dalam analisis deret waktu untuk tujuan peramalan. Model ini menggabungkan komponen *Autoregressive* (AR), yang menggunakan nilai data masa lalu untuk memprediksi nilai saat ini, dengan *Moving Average* (MA) yang dinotasikan ARIMA $(p, 0, q)$, yang menggunakan rata-rata kesalahan prediksi dari periode sebelumnya (Melyani dkk., 2021). Berikut merupakan model ARMA dinyatakan pada persamaan (9) berikut :

$$X_t = C + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \quad (9),$$

  Dimana :

  X_t : Nilai data pada waktu t (yang diprediksi)

  C : Nilai rata – rata (konstanta) dari data

  $\phi_1, \ldots, \phi_p$: Parameter bagian AR

$\theta_1, \ldots., \theta_1$: Parameter bagian MA
e_t : Nilai *error* pada waktu t

- ARIMA (*AutoRegressive Integrated Moving Average)*
  Model *Autoregressive Integrated Moving Average* (ARIMA) adalah salah satu teknik dalam analisis deret waktu yang digunakan untuk membuat peramalan berdasarkan nilai masa lalu dan sekarang dari variabel dependen. ARIMA tidak melibatkan variabel independen dalam proses peramalan, tetapi sepenuhnya bergantung pada pola dan hubungan statistik yang ada dalam variabel dependen itu sendiri (Karnadi, 2023). Bentuk model ARIMA adalah kombinasi dari 3 komponen utama : *Autoregressive* (AR), *Integrated* (I), dan *Moving Average* (MA), yang dinotasikan ARIMA (p, d, q). Menurut Wei (2006) dinyatakan persamaan (10) sebagai berikut :
  $$t\phi_p(B)(1-B)^d X_t = \theta_q(B)\epsilon_t \quad (10),$$
  Dimana :
  X_t : Nilai data pada waktu t (yang diprediksi)
  $\phi_p(B)$: Koefesien AR dengan orde p, dimana $\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \cdots - \phi_p(B^p)$
  $\theta_q(B)$: Koefesien MA dengan orde q, dimana $\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \cdots - \theta_q(B^q)$
  $(1-B)^d$: Operator diferensiasi (lag), di mana d adalah orde diferensiasi untuk mencapai stasioneritas. B adalah operator lag dalam analisis deret waktu untuk menggambarkan pergeseran nilai data ke periode sebelumnya (*Backward Shift*), mendefinisikan $BX_t = X_{t-1}$
  ϵ_t : Nilai residual pada waktu t
  p: Orde *autoregresif* (AR).
  d: Orde diferensiasi (*differencing*) untuk membuat data menjadi stasioner.
  q: Orde *moving average* (MA)

- Stasioner Data Deret Waktu
  Data time series dikatakan stasioner jika rata-rata dan variansinya konstan, tidak ada unsur trend dalam data, dan tidak ada unsur musiman. Menurut Wei (2006) untuk proses stasioner Z_t memiliki 3 sifat berikut:
  1. $E(Z_t) = \mu$, konstan untuk semua t
  2. $Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2$, konstan untuk semua t
  3. $Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)] = y_k$, konstan untuk semua t dan $k \neq 0$ dan y_k adalah autokovarians pada lag ke-k

- Pemeriksaan Diagnostik
  Pemeriksaan diagnostik terdiri dari uji *white noise* dan uji normalitas. Uji *white noise* suatu model dikatakan baik jika nilai *error* bersifat acak yang menunjukkan tidak ada autokorelasi yang memiliki arti residual tidak berpola tertentu. Cara melihat proses white noise pada model yaitu dengan menggunakan uji statistik *Ljung-Box*. Soejoeti (1987) menjelaskan bahwa diagnosa atau verifikasi dimaksudkan untuk memeriksa apakah model yang diestimasi sudah cocok dengan data yang ada. Residual yang diperoleh harus memenuhi asumsi *white noise* dan mengikuti distribusi normal.
  - Uji *White Noise* Residual

Uji *White Noise* Residual pada model ARIMA bertujuan untuk memeriksa apakah residual dari model tersebut menunjukkan perilaku acak atau "*white noise*". *White noise* adalah serangkaian nilai yang tidak memiliki pola yang dapat diprediksi, dengan sifat-sifat utama seperti tidak adanya korelasi antar residual, memiliki mean nol, dan varian yang konstan sepanjang waktu. Uji ini dilakukan untuk memastikan bahwa residual yang dihasilkan oleh model ARIMA tidak mengandung informasi yang dapat digunakan untuk meningkatkan prediksi model. Jika residual bersifat *white noise*, maka model dianggap telah menangkap semua pola yang ada dalam data, dan tidak ada informasi yang tersisa untuk dipelajari lebih lanjut.

a. Uji Residual

Uji independensi *error* digunakan untuk mengetahui apakah *error* antar lag bersifat *independen*, artinya tidak terdapat korelasi *error* antar lag. Uji yang digunakan adalah uji Ljung-Box.

Hipotesis

H_0 : $\rho_1 = \rho_2 = \cdots = \rho_k = 0$ (autokorelasi residual ARIMA(p,d,q) tidak signifikan)

Hipotesis ini menyatakan bahwa semua autokorelasi residual hingga lag maksimum K adalah nol, yang berarti tidak ada pola autokorelasi dalam residual. Dengan kata lain, model sudah cukup menangkap pola data, dan residualnya bersifat acak (*white noise*)

H_1 : Minimal ada satu $\rho_i \neq 0$, $i = 1,2,\ldots.,k$ (autokorelasi residual ARIMA(p,d,q) signifikan)

Hipotesis ini menyatakan bahwa setidaknya ada satu autokorelasi residual pada lag tertentu yang tidak nol. Hal ini menunjukkan adanya pola yang masih tersisa di residual, yang berarti model belum sepenuhnya menangkap struktur data. Dinyatakan persamaan (11) *Ljung-Box* sebagai berikut :

Statistik Uji

$$Q = n(n+2)\ \sum_{k=1}^{K} \frac{\widehat{\rho_k^2}}{n-k}\ , n > k \qquad (11),$$

Daerah penolakan :

Jika $Q > X^2_{(\propto;K-p-q)}$: Hipotesis nol (H_0) ditolak

Jika $Q \leq X^2_{(\propto;K-p-q)}$: Gagal menolak Hipotesis nol (H_0)

Dimana

K : Lag maksimum

n : Jumlah data (observasi)

k : Lag ke-k

p & q : Orde dari ARMA (p,q)

ρ_k : *Autokorelasi* residual untuk lag ke-k

b. Uji Normalitas

Uji normalitas bertujuan untuk menguji apakah *error* dari model mengikuti distribusi normal. Menurut Kabasarang et al. (2012) salah satu cara untuk menguji asumsi normalitas dari suatu data yaitu dengan menggunakan uji *Jarque-Bera* (JB) dinyatakan persamaan (12) sebagai berikut :

Hipotesis

H_0 : *error* distribusi normal

H_1 : *error* tidak berdistribusi normal

Taraf signifikansi = α = 0,05

Statistik uji

$$JB = \frac{n}{6}\left[S^2 + \frac{(K-3)^2}{4}\right] \quad (12),$$

Dimana :

n : Jumlah data

S : Koefisien *skewness* (kemencengan)

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^{3/2}}$$

Jika S > 0 (distribusi memiliki ekor kanan lebih panjang (*positively skewed*))

Jika S < 0 (distribusi memiliki ekor kiri lebih panjang (*negatively skewed*))

K : Koefisien kurtosis

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2}$$

K > 3 (Distribusi memiliki puncak lebih tinggi (*leptokurtic*))

K < 3 (Distribusi lebih datar (*platykurtic*))

K = 3 (Distribusi normal (*mesokurtic*))

Kriteria uji

H_0 ditolak jika nilai $JB > X^2_{(\alpha;2)}$ atau $p - value < \alpha$

Tidak hanya menggunakan uji *Jarque-Bera* (JB), tetapi juga bisa menggunakan uji *Shapiro Wilk* adalah sebuah metode atau rumus perhitungan sebaran data yang dibuat oleh *shapiro* dan *wilk*. Metode *shapiro wilk* adalah metode uji normalitas yang efektif dan valid digunakan untuk sampel berjumlah kecil. Berikut rumus Uji *Shapiro Wilk* dinyatakan pada persamaan (13) sebagai berikut:

$$T_3 = \frac{1}{D}\left[\sum_{i=1}^{k} a_1(X_{n-i+1} - X_i)\right]^2 \quad (13),$$

Dimana :

X_{n-i+1} : Angka ke-n -I +1 pada data

$$D = \sum_{i=1}^{n}(X_i - \bar{X})^2$$

Dimana :

X_i : Angka ke i pada data

$\bar{X}$: Rata – rata data

- Identifikasi Model

Untuk menentukan orde p dan q dalam model ARIMA(p,d,q) digunakan grafik ACF dan PACF. Ada beberapa pedoman dalam pendugaan model ARIMA dari suatu data time series yang dijelaskan oleh Wei (2006) berkaitan dengan plot ACF dan PACF yang telah dirangkum

Tabel 3. 2 Identifikasi Masalah

Model	**ACF**	**PACF**
AR (p)	Turun secara eksponensial atau membentuk gelombang sinus.	*Cut off* setelah lag p
MA(q)	*Cut off* setelah lag q	Turun secara eksponensial atau membentuk gelombang sinus.
ARMA(p,q)	Turun cepat setelah lag (q-p)	Turun cepat setelah lag (p-q)
AR(p) or MA(q)	*Cut off* setelah lag q	*Cut off* setelah lag p
Tidak ada orde AR or MA (*White Noise or Random Process*)	Tidak ada lonjakan	Tidak ada lonjakan

- Estimasi Parameter
  Estimasi merupakan proses yang digunakan untuk menghasilkan suatu nilai tertentu terhadap suatu parameter. Data yang digunakan untuk melakukan estimasi parameter ini merupakan suatu sampel, yang pada perkembangannya akan digunakan oleh suatu estimator untuk menghasilkan suatu nilai parameter. Estimasi parameter pada mulanya akan digunakan untuk menduga suatu populasi dari sampel. Estimasi digolongkan menjadi dua yaitu estimasi titik dan estimasi parameter. Estimasi merupakan suatu tahapan yang terpenting dalam menentukan model peluang yang tepat dari sekumpulan data (Yendra & Noviadi, 2015). Estimasi parameter model ARIMA yang digunakan penulis adalah metode MLE (*Maximum Likelihood Estimation*)

- Evaluasi Model
  - MAE (*Mean Absolute Error*)
    Mean Absolute Error (MAE) adalah salah satu metrik evaluasi yang digunakan untuk mengukur rata-rata kesalahan antara nilai yang sebenarnya dengan nilai prediksi dalam sebuah model. MAE menghitung besarnya kesalahan tanpa memperhatikan arah deviasi (positif atau negatif) dengan mengambil rata-rata dari nilai absolut selisih antara nilai aktual dan nilai prediksi (Chai dkk., 2014). Dinyatakan pada persamaan (14) sebagai berikut :

$$MAE = \sum \frac{|Y\prime - Y|}{n} \quad (14),$$

Dimana :
Y' : Nilai Prediksi

Y : Nilai Aktual
n : Jumlah Data

- RMSE (*Root Mean Square Error*)

RMSE (*Root Mean Square Error*) adalah mengukur rata-rata kesalahan kuadrat antara nilai yang diprediksi oleh model dengan nilai yang sebenarnya (observasi nyata). Nilai RMSE yang lebih kecil menunjukkan bahwa model prediksi tersebut lebih akurat dalam menghasilkan estimasi yang mendekati nilai aktual (Hayati & Silfiani, 2022). Dinyatakan pada persamaa (15) sebagai berikut :

$$RMSE = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{N-P}} \quad (15),$$

y_i : Nilai aktual pada data ke-i
$\hat{y}_i$: Nilai prediksi pada data ke-i
P : Jumlah parameter yang diestimasi dalam model (p + d + q)
N : Jumlah data

- MAPE (*Mean Absolute Percentage Error*)

MAPE digunakan untuk mengetahui tingkat akurasi hasil peramalan. MAPE adalah menghitung nilai kesalahan mutlak pada setiap periode dibagi dengan nilai aktual untuk periode itu, kemudian menghitung rata-rata persentasi dari nilai mutlak tersebut (Sungkawa dkk.,2011). Nilai MAPE dapat dinyatakan pada persamaan (16) sebagai berikut :

$$MAPE = (\sum_{t=1}^{n} \frac{|A_t - F_t|}{A_t})(\frac{100\%}{n}) \quad (16),$$

A_t : Nilai aktual pada periode ke-t
F_t : Nilai hasil peramalan pada periode ke-t
n : Jumlah total periode

Hasil peramalan dikatakan baik jika nilai MAPE yang diperoleh semakin kecil. Adapun kriteria dari nilai MAPE adalah sebagai berikut :

Tabel 3. 3 Kriteria Nilai MAPE

Nilai MAPE	Kriteria
$x < 10\%$	Kemampuan peramalan sangat baik
$10\% \leq x < 20\%$	Kemapuan peramalan baik
$20\% \leq x < 50\%$	Kemampuan peramalan cukup baik
$x \geq 50\%$	Kemampuan peramalan buruk

- AIC (*Instantaneous Autocorrelation*)

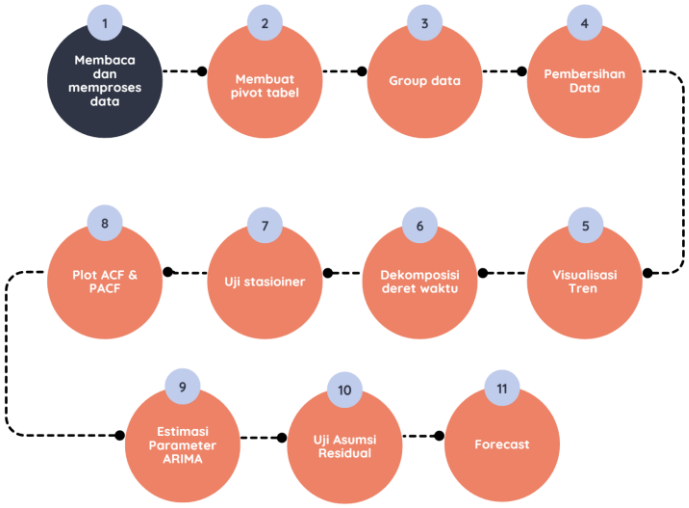
Metode AIC adalah metode yang dapat digunakan untuk untuk mengukur seberapa terkait data pada waktu yang berbeda dalam periode yang sangat singkat dan memilih model regresi terbaik yang ditemukan oleh Akaike (Grasa, 1989). Metode ini didasarkan pada metode *maximum likelihood estimation* (MLE). Menurut metode AIC, model regresi terbaik adalah model regresi yang mempunyai nilai AIC terkecil (Widarjono, 2007) dinyatakan pada persamaan (17) sebagai berikut :

$$AIC = \frac{2k}{n} + \ln(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}) \quad (17),$$

Dimana :
k : Jumlah parameter dalam model
n : Jumlah data dalalm model

$\hat{u}_i^2$: kuadrat dari residual atau *error* pada observasi i yaitu selisih antara nilai yang diprediksi oleh model dan nilai sebenarnya
$\sum_{i=1}^{n} \hat{u}_i^2$: Jumlah kuadrat residual (SSR)
$\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}$: Rata-rata kuadrat residual

3.2.2. Diagram Alur



Gambar 3. 1 Diagram Alur Pemodelan Analisis Deret Waktu Data Kunjungan Telkomsel Ke Sekolah Menggunakan Metode Arima

Untuk menyelesaikan hasil pemodelan analisis deret waktu data kunjungan telkomsel ke sekolah menggunakan metode arima penulis menggambarkan langkah-langkah dalam proses pemodelan analisis deret waktu data kunjungan telkomsel ke sekolah menggunakan metode arima

1. Membaca dan Memproses Data
   Tahap awal adalah membaca dataset yang akan dianalisis. Data perlu diproses agar memiliki format yang sesuai untuk analisis lebih lanjut, seperti memeriksa nilai yang hilang, duplikasi, atau format tanggal.
2. Membuat Pivot Tabel
   Langkah ini melibatkan penyusunan data ke dalam format pivot tabel untuk mempermudah analisis. Tujuan dari pivot tabel adalah untuk merangkum data dengan cara yang lebih terstruktur, misalnya agregasi berdasarkan kategori tertentu.
3. Group Data
   Data yang sudah dirapikan akan dikelompokkan berdasarkan variabel tertentu, seperti waktu atau kategori lainnya, agar lebih mudah dianalisis. Proses ini membantu memahami pola berdasarkan kelompok data.
4. Pembersihan Data
   Tahap ini mencakup penghapusan outlier, pengisian nilai yang hilang, dan koreksi data yang tidak sesuai. Tujuannya adalah memastikan data bersih dan bebas dari anomali sebelum digunakan untuk analisis.
5. Visualisasi Tren
   Setelah data bersih, tren pada data divisualisasikan untuk memberikan gambaran awal tentang pola yang terjadi dalam data dari waktu ke waktu.
6. Dekomposisi Deret Waktu

Pada langkah ini, data deret waktu diuraikan ke dalam komponen-komponennya, seperti tren, musiman (*seasonality*), dan residual. Ini dilakukan untuk memahami pola mendasar dalam data.

7. Uji Stasioner
   Uji stasioner dilakukan untuk memeriksa apakah data memiliki sifat stasioner (mean dan varians konstan dari waktu ke waktu). Jika tidak stasioner, transformasi data akan diperlukan, seperti diferensiasi.
8. Plot ACF & PACF
   Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) diplot untuk mengidentifikasi parameter dalam model ARIMA, seperti nilai p (*autoregressive*) dan q (*moving average*).
9. Estimasi Parameter ARIMA
   Berdasarkan analisis ACF dan PACF, parameter ARIMA dipilih (p, d, q), dan model diestimasi. ARIMA digunakan untuk menangkap pola dalam data deret waktu.
10. Uji Asumsi Residual
    Setelah model ARIMA dibuat, residual diuji untuk memastikan model sesuai. Residual yang baik harus berdistribusi normal dan tidak memiliki autocorrelation.
11. Forecast
    Model ARIMA digunakan untuk melakukan prediksi (*forecasting*) terhadap data masa depan. Prediksi ini didasarkan pada pola yang telah diidentifikasi dalam data.

3.2.3. Timeline

Proses pengerjaan Pemodelan Deret Waktu Data Kunjungan Telkomsel Ke Sekolah Menggunakan Arima, dengan waktu pengerjaan sebagai berikut :

Tabel 3. 4 Timeline

Tahapan Pengerjaan Project	Periode			
	1 – 15 November 2024	16 – 30 November 2024	1 – 8 Desember 2024	9 – 27 Desember 2024
Data *Collecting*				
Preprocessing Data				
Analisis Data Eksploratif				
Splitting Data				
Pemodelan ARIMA				
Prediksi				

3.2.3.1. Analisis dan Pembahasan

3.2.3.2. Data Penelitian

Data yang digunakan dalam analisis ini merupakan data tren kunjungan Telkomsel ke sekolah-sekolah dari tanggal 1 Januari 2024 hingga 9 Oktober 2024, dengan mencakup berbagai

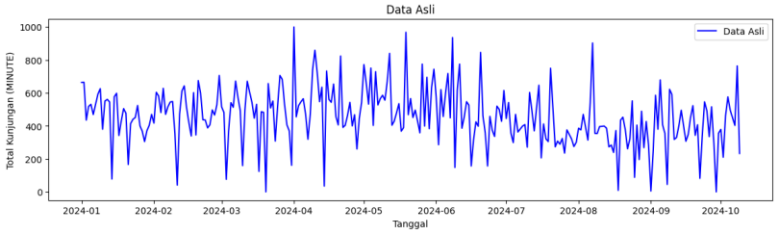
tingkat sekolah, yaitu KURSUS, PESANTREN, PKBM, SD, SKB, SLB, SMA, SMP, SPK, TK, dan UNIVERSITAS. Fokus analisis terletak pada durasi total kunjungan dalam satuan menit pada setiap tingkat sekolah dalam periode harian. Tujuan dari analisis ini adalah untuk memprediksi data kunjungan selama 20 hari ke depan, dengan variabel yang digunakan untuk analisis adalah 'TOTAL KUNJUNGAN (MINUTE)'.

LEVEL	PERIODE DATE	KURSUS	PESANTREN	PKBM	SD	SKB	SLB	SMA	SMP	SPK	TK	UNIVERSITAS	Total Kunjungan (MINUTE)
0	2024-01-01	0	0	0	0	0	0	67	179	0	0	227	664
1	2024-01-02	54	0	0	92	0	63	73	80	0	54	124	665
2	2024-01-03	0	0	0	81	0	71	88	64	0	39	92	435
3	2024-01-04	66	0	0	73	0	0	85	75	0	128	91	518
4	2024-01-05	0	0	0	73	0	33	83	91	0	145	106	531
...	...	...	...	...	...	...	...	...	...	...	...	...	...
269	2024-10-04	0	0	0	60	0	135	65	54	0	206	56	576
270	2024-10-05	0	0	0	61	0	0	60	69	0	170	135	495
271	2024-10-07	0	0	0	62	0	80	60	53	0	103	45	403
272	2024-10-08	0	0	0	52	0	60	57	54	0	98	74	764
273	2024-10-09	0	23	0	41	0	0	35	40	0	57	37	233

274 rows × 13 columns

Gambar 3. 2 Data Kunjungan Telkomsel ke Sekolah

3.2.3.3. Plot Data Deret Waktu



Gambar 3. 3 Plot Data Deret Waktu

Dapat disimpulkan bahwa menunjukkan fluktuasi yang signifikan dalam durasi total kunjungan (minute) dari waktu ke waktu, dengan rentang nilai berkisar antara 0 hingga sekitar 1000 menit atau sekitar 1,6 jam. Grafik ini memiliki variabilitas yang cukup tinggi, ditandai dengan lonjakan dan penurunan tajam di beberapa titik, namun tidak tampak adanya pola musiman atau tren yang konsisten.

3.2.3.4. Stasioner Data

Dalam hal ini penulis melakukan uji stasioner dalam mean menggunakan uji *Augmented Dickey-Fuller* (ADF). Hasil seperti dibawah ini :

Hasil *Augmented Dickey-Fuller Test* yaitu :

1. H_0 : Data tidak stasioner terhadap mean (ada akar unit)
   H_1 : Data stasioner terhadap mean.
2. Menentukan α = 0.05
3. Wilayah Kritis
   Jika ADF statistic < nilai kritis, maka H_0 ditolak (data stasioner)
   Jika ADF statistic > nilai kritis, maka H_0 gagal ditolak (data tidak stasioner)
4. Kesimpulan
   ADF Statistic: -3.812883434768411
   p-value: 0.0027782171924074766
   Critical Values:

1%: -3.4550813975770827
5%: -2.8724265892710914
10%: -2.5725712007462582
Kesimpulannya adalah Data stasioner terhadap mean (H0 ditolak).

Untuk melakukan uji stasioner dalam varians, penulis menggunakan uji KPSS dan didapatkan berikut :

Hasil Uji KPSS :

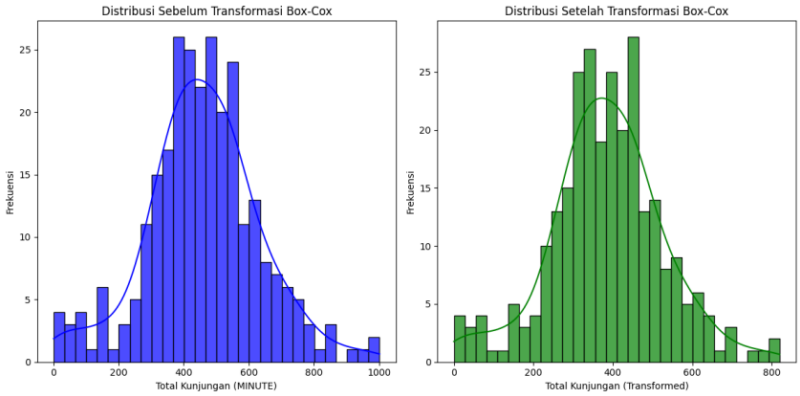
1. H_0: Data stasioner terhadap varians
   H_1 : Data tidak stasioner terhadap varians
2. Menentukan α = 0.05
3. Wilayah Kritis
   H_0 ditolak (Data tidak stasioner) jika P−value < α.
4. Kesimpulan
   Statistik Uji: 1.6002831966951836
   P-value: 0.01
   Nilai kritis :
   10%: 0.347
   5%: 0.463
   2.5%: 0.574
   1%: 0.739
   Kesimpulan p-value (0.01) lebih kecil dari 0.05, sehingga menolak hipotesis nol. Ini berarti Data tidak stasioner terhadap varians, dan terdapat perubahan signifikan dalam data.

Selanjutnya adalah melakukan transformasi Box-Cox pada data, alasan penulis melakukan transformasi box-cox dikarenakan transformasi ini dapat membantu mengatasi masalah distribusi data yang tidak normal dan meningkatkan kestabilan varians. Dengan menerapkan Box-Cox, dapat mengubah data ke bentuk yang lebih mendekati distribusi normal, yang penting untuk beberapa teknik analisis statistik seperti regresi atau model time series. Transformasi ini juga membantu dalam mengurangi pengaruh *outlier* dan meningkatkan kemampuan model untuk memberikan prediksi yang lebih akurat. Seperti pada hasil berikut :

Berikut ini hipotesis Uji *Box-Cox*:

1. H_0: $\sigma^2_{sebelum} = \sigma^2_{setelah}$ (Tidak ada perubahan signifikan pada varians atau distribusi data setelah transformasi Box-Cox)
2. H_1 : $\sigma^2_{sebelum} \neq \sigma^2_{setelah}$ (Ada perubahan signifikan pada varians atau distribusi data setelah transformasi Box-Cox dengan parameter λ)
3. Menentukan α = 0.05
4. Wilayah Kritis
   H_0 ditolak (perubahan yang signifikan dalam varians setelah transformasi) jika P−value < α.
5. Kesimpulan

Hasil Transformasi Box-Cox yaitu :
Lambda (Parameter Transformasi): 0.9663279782877713
Kesimpulan Stasioner Varians:
Standar Deviasi Sebelum Transformasi: 173.83
Standar Deviasi Setelah Transformasi: 141.80
Kesimpulannya adalah varians data lebih stabil dan data cenderung stasioner terhadap varians.

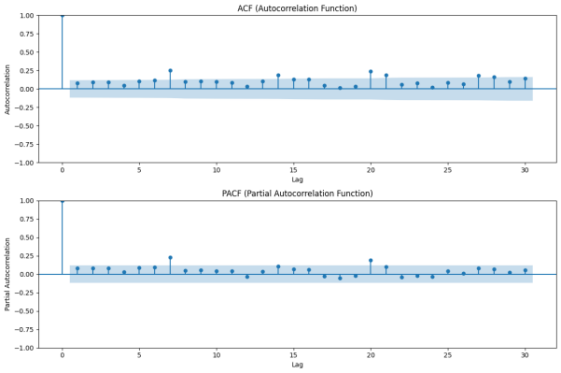


Gambar 3. 4. Perbandingan Sebelum & Sesudah Transformasi Box-Cox

Distribusi data terlihat lebih simetris dan mendekati distribusi normal. Transformasi Box-Cox telah berhasil mengurangi kemiringan data. Puncak distribusi masih berada di sekitar nilai yang sama, tetapi penyebaran data terlihat lebih terkonsentrasi di sekitar rata-rata. Rentang data juga sedikit menyempit, terutama di sisi kanan, karena transformasi telah mengurangi pengaruh nilai-nilai ekstrem.

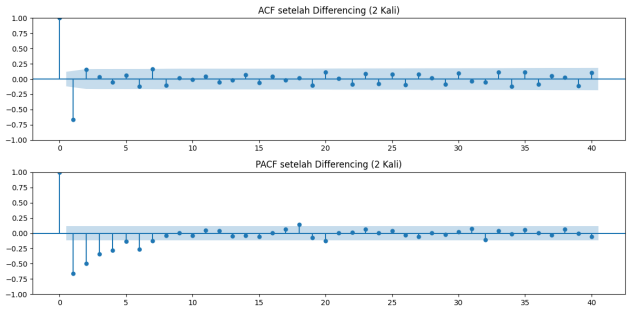
3.2.3.5. Identifikasi Model Dengan ACF dan PACF

Setelah mengetahui jika data yang dimiliki yaitu stasioner, didapatkan plot ACF dan PACF



Gambar 3. 5 Plot PACF & ACF Sebelum Differencing

Karena tidak terlihat pola *cut off* maka penulis ingin melakukan *differencing* sebanyak 2 kali. Tujuannya yaitu ingin melihat *cut off* pada lag yang lebih jelas.



Gambar 3. 6. Plot PACF dan ACF

Didapatkan hasil model ARIMA(p,d,q) adalah (6,2,1). Hal ini berarti model menggunakan enam lag *autoregressive* (p=6), dua kali *differencing* (d=2), dan satu lag moving average (q=1). Berikut adalah model persamaan ARIMA(6,2,1)

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4 - \phi_5 B^5 - \phi_6 B^6)(1 - B)^2 Y_t = \theta_1 B_{\epsilon_t}$$

Dimana :

$$B = Y_{t-1}\ (nilai\ data\ pada\ waktu\ t - 1)$$

$$B^2 = operator\ untuk\ mundur\ dua\ waktu\ (Y_{t-2})\ , dst$$

Penjelasan :

- Terdiri dari 6 lag.
- Menunjukkan dua kali *differencing* pada data, yang dapat menstabilkan variansi dan tren.
- Terdiri dari 1 lag *error*

3.2.3.6. Estimasi Parameter

```
Ringkasan Model Terbaik:
                                     SARIMAX Results
==========================================================================================
Dep. Variable:     Total Kunjungan (MINUTE)   No. Observations:                  274
Model:                        ARIMA(6, 2, 1)   Log Likelihood               -1788.451
Date:                       Tue, 31 Dec 2024   AIC                          3592.902
Time:                               20:56:52   BIC                          3621.749
Sample:                                    0   HQIC                         3604.483
                                       - 274
Covariance Type:                         opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8966      0.061    -14.769      0.000      -1.016      -0.778
ar.L2         -0.7797      0.083     -9.387      0.000      -0.942      -0.617
ar.L3         -0.6475      0.088     -7.394      0.000      -0.819      -0.476
ar.L4         -0.5759      0.087     -6.630      0.000      -0.746      -0.406
ar.L5         -0.4419      0.073     -6.046      0.000      -0.585      -0.299
ar.L6         -0.2963      0.048     -6.164      0.000      -0.390      -0.202
ma.L1         -1.0000      0.077    -12.934      0.000      -1.152      -0.848
sigma2      2.903e+04   2.66e-06   1.09e+10      0.000     2.9e+04     2.9e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.35   Jarque-Bera (JB):                 4.31
Prob(Q):                              0.55   Prob(JB):                         0.12
...
```

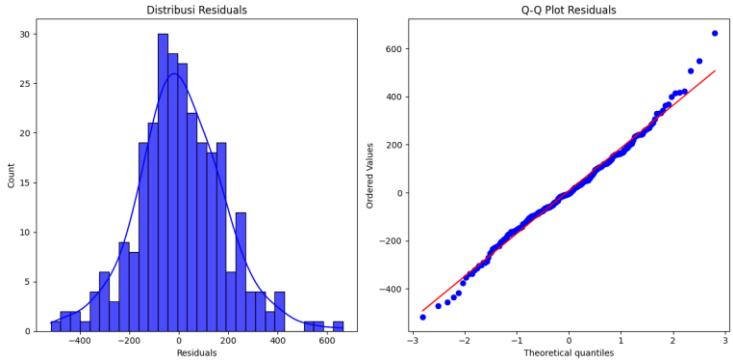
Gambar 3. 7 Estimasi Parameter

Model ARIMA(6,2,1) yang diestimasi menggunakan alat *Maximum Likelihood Estimation* (MLE) dengan jumlah observasi sebanyak 274 menunjukkan kinerja yang baik berdasarkan kriteria evaluasi model, yaitu Log *Likelihood* (-1788.451), AIC (3592.922), BIC (3621.749), dan HQIC (3604.483), di mana nilai AIC lebih kecil dibandingkan BIC dan HQIC, dengan semua parameter AR (L1 hingga L6) dan MA signifikan secara statistik (P-value < 0.05) yang mencerminkan

relevansi dalam menjelaskan pola data dengan hubungan langsung atau invers terhadap data masa lalu, sementara parameter sigma2 (*error variance*) menggambarkan variabilitas yang tidak dapat dijelaskan oleh model.

$$(1-(-0{,}8)B-(-0{,}7)B^2-(-0{,}6)B^3-(-0{,}5)B^4-(-0{,}4)B^5-(-0{,}2)B^6)(1-B)^2Y_t=(-1)B_{\epsilon_t}$$

3.2.3.7. Uji Asumsi Residual



Gambar 3. 8. Distribusi Residual

Untuk menguji asumsi residual penulis menggunakan Uji *Ljung-Box* dan untuk uji normalitas menggunakan uji *Shapiro-Wilk*.

Berikut ini hipotesis Uji Ljung-Box:

1. H_0: ρ1 = ρ2 = ⋯ = ρk = 0 (autokorelasi residual ARIMA(p,d,q) tidak signifikan)
2. H_1 : $\rho i \neq 0, i = 1,2,3,\ldots,k$ (autokorelasi residual ARIMA(p,d,q) signifikan)
3. Menentukan α = 0.05
4. Wilayah Kritis
   H_0 ditolak (autokorelasi residual signifikan) jika P−value < α.
5. Kesimpulan

Hasil uji *Ljung-Box* menunjukkan tidak adanya *autocorrelation* yang signifikan pada residual (p-value: 0.704). Ini menunjukkan bahwa model telah berhasil menangkap semua pola dalam data, dan residual tidak memiliki pola yang terdeteksi.

Berikut ini hipotesis Uji Shapiro-Wilk:

1. H_0 : Data residual terdistribusi normal.
   H_1 : Data residual tidak terdistribusi normal.
2. Menentukan α = 0.05
3. Wilayah Kritis

H_0 ditolak (residual tidak terdistribusi normal) jika p-value < α.
H_0 tidak ditolak (residual terdistribusi normal) jika p-value ≥ α.

4. Kesimpulan

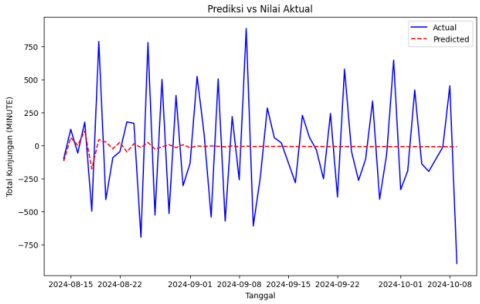
Hasil uji Shapiro-Wilk menunjukkan bahwa residual terdistribusi normal (p-value: 0.065), yang mengindikasikan bahwa asumsi normalitas terpenuhi. Ini adalah hasil yang baik karena banyak model *time series* yang mengasumsikan bahwa residual mengikuti distribusi normal.

3.2.3.8. Evaluasi Model

Langkah terakhir dalam analisis adalah melakukan evaluasi data dengan membagi data menjadi 80% pelatihan dan 20% pengujian terhadap model ARIMA(6,2,1) yang telah dihasilkan. Berdasarkan hasil evaluasi, diketahui bahwa nilai MAPE sebesar 143.60% mengindikasikan bahwa kesalahan prediksi rata-rata lebih besar dari nilai aktual itu sendiri. Hal ini menunjukkan bahwa model kurang akurat, terutama karena persentase kesalahan prediksi sangat tinggi. Nilai Mean Absolute Error (MAE) sebesar 315.504 dan Root Mean Squared Error (RMSE) sebesar 392.954 memberikan gambaran tentang kualitas model. Nilai RMSE yang lebih besar dibandingkan MAE menunjukkan adanya kesalahan besar pada beberapa prediksi (outlier). Menunjukkan bahwa MAE lebih baik untuk mengukur kesalahan rata-rata dan mudah dipahami dalam konteks prediksi total kunjungan. Nilai MAE yang rendah menunjukkan bahwa model memberikan prediksi yang cukup akurat secara keseluruhan.

Tabel 3. 5 Evaluasi Model

Model	MAE	RMSE	MAPE
ARIMA (6,2,1)	315.504	392.954	143.60%

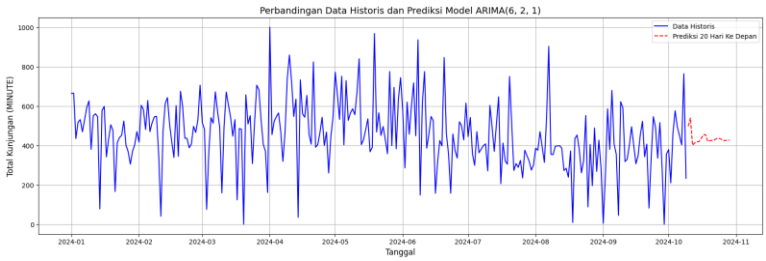


Gambar 3. 9 Grafik Prediksi dan Data Aktual

Bahwa prediksi ini hanya memberikan gambaran berdasarkan data historis dan tren yang ditangkap oleh model. Jika ada faktor eksternal atau kejadian mendadak (seperti promosi besar-besaran, kondisi cuaca, atau perubahan kebijakan) yang tidak tercermin dalam data historis, hasil prediksi mungkin tidak sepenuhnya akurat. Oleh karena itu, selain menggunakan hasil prediksi ini, penulis juga mencantumkan kondisi aktual dan mempertimbangkan faktor-faktor eksternal yang relevan untuk mendukung keputusan yang lebih tepat. Didapatkan grafik untuk data aktual memberikan yang menunjukkan variasi yang tajam, dengan puncak kunjungan yang tinggi pada beberapa tanggal tertentu, diselingi oleh penurunan yang cukup drastis. Berikut tabel perhitungan total kunjungan ke sekolah dalam prediksi selama 20 hari kedepan :

Tabel 3. 6 Hasil Peramalan 20 Hari Kedepan

Tanggal	Prediksi
2024-10-10	497.572319
2024-10-11	539.057898
2024-10-12	400.864929
2024-10-13	412.701807
2024-10-14	418.038390
2024-10-15	419.771158
2024-10-16	436.581762
2024-10-17	454.783585
2024-10-18	453.795988
2024-10-19	420.282617
2024-10-20	424.312649
2024-10-21	425.977893
2024-10-22	427.803285
2024-10-23	434.967007
2024-10-24	435.806465
2024-10-25	432.718216
2024-10-26	424.587975
2024-10-27	425.461612
2024-10-28	426.162324
2024-10-29	426.768019



Gambar 3. 10 Grafik Prediksi 20 Hari Kedepan

Grafik di atas menunjukkan prediksi jumlah kunjungan untuk 20 hari ke depan, namun, pada bagian peramalan (ditandai dengan garis merah), terlihat bahwa pola prediksi lebih teratur dan tidak menunjukkan fluktuasi ekstrem seperti data historis. Pola ini mencerminkan upaya model prediksi untuk menangkap tren jangka pendek atau rata-rata dari data historis, sementara variabilitas harian yang bersifat acak atau tidak terduga diabaikan. Dengan kata lain, model peramalan memberikan gambaran stabil yang dapat mempermudah dalam memahami pola umum jumlah kunjungan. Akan bermanfaat untuk menghubungkan wawasan ini dengan keputusan atau tindakan yang mungkin diambil Telkomsel berdasarkan prediksi ini, seperti penjadwalan ulang kunjungan, alokasi sumber daya, atau strategi promosi yang lebih efektif.

3.3. Pembelajaran Hal Baru

Selama mengikuti magang mandiri di Telkomsel Branch Surabaya, penulis sebagai data analyst mendapatkan banyak pengalaman baru, terutama dalam hal analisis deret waktu menggunakan metode ARIMA. Proses ini memberi wawasan tentang bagaimana memilih model yang tepat untuk memprediksi data deret waktu yang memiliki pola musiman dan tren jangka panjang. Saya juga belajar mengenai pentingnya mengevaluasi kinerja model menggunakan berbagai metrik, seperti MAPE, MAE, dan RMSE, yang membantu dalam menilai seberapa akurat model dalam memprediksi nilai yang akan datang berdasarkan data historis.

Melalui analisis ini, saya memahami betapa pentingnya melakukan visualisasi data secara efektif. Visualisasi ini bukan hanya berfungsi untuk menggambarkan pola dalam data, tetapi juga untuk mendeteksi tren dan musiman yang tersembunyi, yang bisa berperan besar dalam perencanaan bisnis. Dengan visualisasi yang jelas, saya dapat lebih mudah memantau perubahan pola dari waktu ke waktu dan memprediksi kemungkinan-kemungkinan yang dapat terjadi di masa depan. Teknik ini sangat bermanfaat dalam mempresentasikan data yang kompleks kepada stakeholder untuk pengambilan keputusan yang lebih baik dan terarah.

Selain itu, saya semakin menyadari bahwa pemrosesan data yang tepat sangat penting dalam mencapai hasil yang optimal. Proses seperti penanganan missing values dan pemilihan hyperparameter yang tepat terbukti mempengaruhi performa model secara signifikan. Saya belajar bahwa model yang efektif tidak hanya mampu menangkap pola yang ada, tetapi juga dapat menangani berbagai tantangan dalam data, seperti ketidaksempurnaan atau noise. Secara keseluruhan, pengalaman ini memberikan wawasan yang lebih mendalam tentang bagaimana menggunakan analisis deret waktu untuk mendukung perencanaan jangka pendek dan strategi pemasaran, serta bagaimana data dapat digunakan sebagai dasar pengambilan keputusan yang lebih cerdas dalam konteks bisnis.

BAB 4. REKOGNISI MATA KULIAH

Tabel 4. 1 Rekognisi Mata Kuliah

No.	Aktivitas Magang Mandiri	Konversi/Rekognisi Mata Kuliah Program Studi (Jumlah SKS)
1	Pengalaman langsung di dunia kerja yang relevan dengan bidang studi mahasiswa. Mahasiswa mengaplikasikan teori yang dipelajari, mengembangkan keterampilan teknis dan non-teknis, serta terlibat dalam tugas seperti analisis data, pembuatan visualisasi, dan pembuatan laporan. Di akhir magang, mahasiswa menyusun laporan dan melakukan presentasi evaluasi yang menjadi bahan penilaian. Magang ini mempersiapkan mahasiswa untuk dunia kerja dan meningkatkan keterampilan profesional mereka.	Praktik Kerja Lapangan (2 SKS)
2	1. Kolaborasi dengan Divisi Administrasi Bisnis Kami mengadakan agenda bersama dengan Divisi Administrasi Bisnis terkait pelaksanaan event yang akan diselenggarakan. Selain itu, kami juga mempelajari dan memahami proses kerja tim sales marketing dalam operasional lapangan. 2. Diskusi Tim Reguler Pertemuan rutin diadakan setiap minggu untuk membahas tugas-tugas yang akan dilakukan serta mengevaluasi hasil kerja sebelumnya. Diskusi ini bertujuan untuk memastikan kelancaran pelaksanaan tugas dan koordinasi tim. 3. Tugas Kolaboratif Pembagian tugas dilakukan secara kolaboratif dengan memperhatikan keahlian dan tanggung jawab masing-masing anggota tim. Hal ini bertujuan untuk meningkatkan efisiensi dan memastikan setiap pekerjaan terselesaikan dengan baik. 4. Presentasi Proyek Setelah menyelesaikan tugas mingguan, tim secara berkala melakukan presentasi proyek untuk melatih kemampuan berbicara di depan umum. Dalam presentasi ini, kami juga menyampaikan rekomendasi berdasarkan hasil analisis yang telah dilakukan.	Kerja Sama (2 SKS)
3	Clustering Menggunakan Metode K-Means: Segmentasi Pelanggan Berdasarkan Area, Rasio Aktivitas Pelanggan, dan Omzet untuk Strategi Pemasaran Berdasarkan Data Event Telkomsel Menggunakan Bahasa Pemrograman Python dengan memberikan sasaran pasar secara efektif dan efisien	Analisis Segmentasi Pelanggan (3 SKS)

4	Menganalisis data events Telkomsel menggunakan Looker Studio bertujuan mendukung strategi peningkatan layanan. Fokus utama adalah optimalisasi pemasaran di Surabaya sebagai wilayah dengan pengguna tertinggi, peningkatan aktivasi DS melalui pelatihan dan insentif, serta ekspansi bertahap di wilayah potensial seperti Bangkalan, Pamekasan, Sumenep, dan Sampang. Observasi lapangan direncanakan untuk memastikan strategi relevan dan efektif.	Instuisi dan Wawasan Data (2 SKS)
5	Mahasiswa membuat proyek "Pengembangan Dashboard Kecerdasan Bisnis dengan Looker dan Python untuk Mendukung Pengambilan Keputusan Strategis di Telkomsel Branch Surabaya" dengan tujuan untuk mengumpulkan dan membersihkan data operasional dari berbagai sumber, menganalisis data menggunakan Python untuk mendapatkan wawasan yang relevan, serta mengembangkan dashboard interaktif	Kecerdasan Bisnis (3 SKS)
6	Analisis strategi sales berdasarkan area menggunakan bahasa pemrograman Python, aktivitas magang yang dilakukan mencakup penyusunan laporan hasil analisis untuk dipresentasikan kepada tim terkait. Aktivitas ini melibatkan interpretasi data yang dihasilkan, penarikan kesimpulan, dan penyusunan rekomendasi kebijakan berbasis data untuk mendukung strategi pemasaran Telkomsel.	Statistika Sains Data dan Analisis Kebijakan (3 SKS)
7	Penulis menggunakan metode analisis deret waktu, termasuk ARIMA, untuk menganalisis data historis kunjungan Telkomsel ke sekolah, mengidentifikasi pola, tren, dan fluktuasi untuk mendukung pengambilan keputusan strategis dan perencanaan yang lebih akurat. Hasil analisis disampaikan melalui integrasi data analitik dengan elemen naratif, menggunakan visualisasi dan narasi untuk menjelaskan temuan secara efektif, sehingga mendukung pemahaman dan pengambilan keputusan, bahkan bagi pemangku kepentingan tanpa latar belakang analitis.	Analisis Deret Waktu (3 SKS) dan Data Storytelling (2 SKS)

4.1. Rekognisi Mata Kuliah Praktik Kerja Lapangan

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Praktik Kerja Lapangan. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Praktik Kerja Lapangan

Tabel 4. 2 Indikator Penilaian Konversi Mata Kuliah Praktik Lapangan Kerja

No	Kriteria/Indikator Penilaian
1	Kemampuan mengidentifikasi sejumlah solusi untuk masalah kompleks mengintegrasikan temuan dari beberapa disiplin ilmu yang berbeda, mengidentifikasi dan mengevaluasi berbagai opsi yang dikembangkan dan memilih solusi yang paling efektif
2	Kemampuan dalam menjalin kerjasama bersama tim kerja untuk menyelesaikan pekerjaan sesuai pembagian job desk, tenggat waktu yang telah ditentukan, dan memperhatikan kerapian dan ketepatan hasil pekerjaan
3	Kemampuan mengumpulkan, mengintegrasikan, dan menganalisis semua data dan informasi yang relevan dan mereduksi informasi tersebut menjadi komponen dan/atau bagan, diagram, atau grafik yang dapat diinterpretasikan dan dikomunikasikan kepada pemangku kepentingan
4	Kemampuan menarik kesimpulan logis dan objektif dari data dan memvalidasinya sebagai penyebab utama dan penyebab yang berkontribusi terhadap data yang akan dimodelkan

4.1.1. Praktik Lapangan Kerja [Indikator ke-1]

Praktik Kerja Lapangan (PKL) adalah mata kuliah yang dirancang untuk memberikan mahasiswa pengalaman langsung dalam menerapkan pengetahuan dan keterampilan yang telah mereka pelajari selama proses pendidikan di bangku kuliah. Mata kuliah ini bertujuan untuk menjembatani antara teori akademik dengan praktik nyata yang dihadapi di dunia kerja. Dalam konteks Program Studi Sains Data, PKL menjadi sarana penting bagi mahasiswa untuk memahami bagaimana ilmu sains data dapat digunakan dalam menyelesaikan berbagai masalah nyata yang ada di lingkungan kerja.

Dalam kegiatan magang yang dilaksanakan oleh penulis di PT Telkomsel Branch Surabaya, penulis menjalani program magang dengan rentang waktu mulai tanggal 2 September 2024 hingga 22 Januari 2025. Kegiatan magang ini dilaksanakan secara hybrid, yaitu kombinasi antara pelaksanaan secara online dan offline. Untuk pelaksanaan secara offline, kegiatan berlokasi di Telkom Landmark Tower Surabaya, yang beralamat di Jl. Dr. Ir. H. Soekarno No.175, Klampis Ngasem, Kec. Sukolilo, Surabaya, Jawa Timur.

Sebagai pemagang di perusahaan ini, penulis diberikan tanggung jawab utama dalam mengolah dataset yang dimiliki oleh Telkomsel Branch Surabaya. Proses pengolahan data ini mencakup berbagai tahapan, seperti membersihkan data untuk memastikan kualitasnya, menganalisis pola-pola yang relevan, hingga menginterpretasikan hasil analisis agar memberikan nilai tambah bagi perusahaan. Penulis juga ditugaskan untuk mengembangkan sebuah dashboard monitoring yang dirancang untuk membantu perusahaan dalam memantau kinerja operasional serta memberikan pandangan yang lebih mendalam terhadap berbagai indikator bisnis.

Selama kegiatan magang, penulis tidak hanya memperoleh pengalaman teknis terkait pengolahan data dan pembuatan dashboard, tetapi juga mendapatkan wawasan mengenai bagaimana data digunakan dalam proses pengambilan keputusan bisnis. Selain itu, penulis juga berkesempatan untuk berinteraksi dengan tim yang terdiri dari berbagai divisi, memperluas pengetahuan tentang budaya kerja

di PT Telkomsel, serta mengasah keterampilan komunikasi dan kolaborasi dalam lingkungan profesional.

PKL memberikan kesempatan kepada mahasiswa untuk mengeksplorasi dunia kerja secara langsung, mengenali dinamika organisasi, serta berkontribusi dalam pengelolaan dan analisis data. Dengan demikian, mahasiswa dapat mengasah keterampilan teknis yang mereka miliki, seperti kemampuan dalam pengolahan data, analisis statistik, pemrograman, dan pembuatan visualisasi data. Selain itu, mereka juga akan mendapatkan pemahaman yang lebih mendalam mengenai keterampilan nonteknis yang diperlukan dalam dunia kerja, seperti manajemen waktu, komunikasi profesional, kerja sama tim, serta etika kerja yang baik.

4.1.2. Menjalin Kerja Sama Dalam Praktik Lapangan Kerja[Indikator ke-2]

Gambar 4. 1 Berkolaborasi dan Berdiskusi

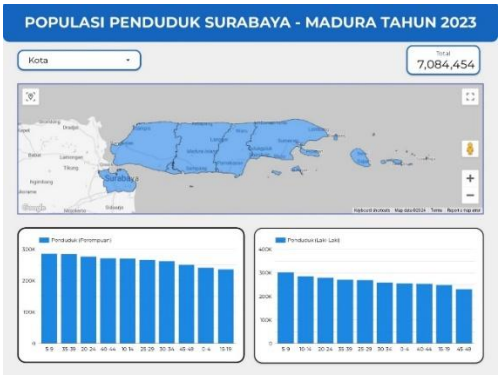
Selama menjalani program magang, penulis bersama tim menjaga kolaborasi yang erat dengan berbagai pihak, terutama yang berhubungan langsung dengan pengolahan data dan pengembangan dashboard monitoring di Telkomsel Branch Surabaya. Kerja sama dilakukan baik secara individu maupun tim dengan memanfaatkan platform komunikasi seperti WhatsApp dan *Google Meet* untuk mendukung diskusi, berbagi informasi, serta memperdalam pemahaman terkait kebutuhan bisnis perusahaan. Agar pekerjaan lebih terorganisir, tim melakukan pembagian tugas secara merata berdasarkan kompetensi masing-masing anggota, sehingga seluruh tanggung jawab dapat diselesaikan sesuai dengan target yang telah ditentukan dalam timeline proyek.

Gambar 4. 2 Koordinasi Dengan Mentor

Penulis secara aktif berkoordinasi untuk memastikan bahwa setiap anggota memahami tugas dan perannya dalam proyek. Selain itu, penulis dan tim secara konsisten mengikuti sesi mentoring dengan pembimbing lapangan di Telkomsel maupun dosen pembimbing dari kampus. Sesi ini menjadi momen penting untuk membahas kendala teknis maupun non-teknis yang ditemui selama pengerjaan proyek. Penulis juga menerima berbagai masukan yang berharga untuk

menyempurnakan proses pengolahan data dan pembuatan dashboard menggunakan Looker Data Studio.

4.1.3. Menganalisis Data Serta Memvisualisasikan [Indikator ke-3]



Gambar 4. 3 Dashboard Populasi Penduduk Surabaya – Madura Tahun 2023

Menganalisis data dengan menggunakan Looker Data Studio sebagai platform utama untuk membuat dashboard yang interaktif dan mudah dipahami. Proses ini melibatkan pengolahan dataset besar yang mencakup berbagai aspek operasional Telkomsel, seperti penjualan produk, penggunaan layanan, dan kepuasan pelanggan. Dashboard yang dibuat dirancang untuk memberikan wawasan yang berguna bagi berbagai divisi, termasuk tim manajemen dan marketing.

Selain itu, penulis memastikan tampilan dashboard memiliki antarmuka yang user-friendly dan profesional. Dengan fitur-fitur interaktif seperti filter data dan grafik yang dinamis, dashboard ini membantu tim Telkomsel dalam mengakses informasi yang dibutuhkan secara cepat dan efisien. Proses lebih lanjut beserta hasil yang lebih mendetail terkait dashboard ini dapat dilihat pada laporan magang atau dokumen pendukung lainnya. Dashboard ini bertujuan untuk menyediakan visualisasi data yang komprehensif dan *real-time*, yang mencakup berbagai metrik penting seperti jumlah pelanggan, pendapatan, tingkat churn, serta tren penggunaan layanan di wilayah Surabaya. Solusi ini dirancang untuk membantu manajemen dalam memantau performa operasional dengan lebih efisien dan mengambil keputusan berbasis data secara bijak.

Proses pengolahan dataset besar melibatkan berbagai tahapan, seperti membersihkan data untuk memastikan kualitasnya, menganalisis pola-pola yang relevan, hingga menginterpretasikan hasil analisis agar memberikan nilai tambah bagi perusahaan. Dashboard yang dibuat dirancang untuk memberikan wawasan yang berguna bagi berbagai divisi, termasuk tim manajemen dan marketing. Dengan fitur-fitur interaktif seperti filter data dan grafik yang dinamis, dashboard ini membantu tim Telkomsel dalam mengakses informasi yang dibutuhkan secara cepat dan efisien.

4.1.4. Kesimpulan Data [Indikator ke-4]

Gambar 4. 4 Dashboard Data

Setelah dilakukan analisis terhadap permasalahan yang ada, penulis melanjutkan dengan mengerjakan solusi hingga menghasilkan sebuah dashboard monitoring yang dapat membantu Telkomsel Branch Surabaya dalam mengelola data operasional dan mendukung pengambilan keputusan bisnis. Dari data dashboard penjualan Telkomsel Surabaya, tren harian menunjukkan puncak signifikan yang kemungkinan dipengaruhi promosi atau event tertentu. Penjualan tertinggi berasal dari Surabaya, diduga karena potensi pasar yang besar, sementara kontribusi kota lain perlu dianalisis lebih lanjut. Mayoritas penjualan berasal dari "*Event SO*" menunjukkan efektivitas strategi berbasis event, dengan peluang diversifikasi ke event lain. Aktivasi IMEI baru dan SIM dipengaruhi promosi bundling, sementara distribusi penjualan menunjukkan fokus pada beberapa POI dominan. Validasi diperlukan melalui analisis data historis, kampanye pemasaran, dan evaluasi strategi di area atau waktu tertentu.

4.2. Rekognisi Mata Kuliah Kerja Sama

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Kerja Sama. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Kerja Sama

Tabel 4. 3 Indikator Konversi Kerja Sama

No	Kriteria/Indikator Penilaian
1	Kemampuan untuk berkomitmen dalam tim melalui rasa saling percaya,keinginan untuk saling membantu, komunikasi terbuka, dan mencapai tujuan bersama
2	Kemampuan untuk dapat memanfaatkan potensi sumber daya manusia yang ada dalam tim secara optimal
3	Kemampuan menerapkan pola kerjasama untuk mempercepat pertumbuhan *teamwork*
4	Kemampuan mekanisme monitoring *progress* pekerjaan secara berkala dengan mengedepankan nilai kebersamaan

4.2.1. Materi Softskill – Kerja Sama [Indikator ke-1]

Secara Umum, kerjasama adalah salah satu soft skill yang sangat penting dalam dunia profesional. Ini melibatkan kemampuan untuk bekerja sama dengan orang lain secara efektif dan terciptanya rasa saling percaya dan keinginan untuk saling membantu di antara individu, tim, dan mitra kerja (Lilian Purba & Psikologi, 2013).

Konsep Dasar Kerja Sama

Kerja sama adalah suatu usaha bersama antara individu atau kelompok sosial untuk mencapai tujuan bersama. Kerja sama dilakukan sejak manusia berinteraksi dengan sesamanya. Kebiasaan dan sikap mau bekerja sama dimulai sejak kanak-kanak, mulai dalam kehidupan keluarga lalu meningkat dalam kelompok sosial yang lebih luas. Kerjasama juga dapat dilakukan antara Pemerintah dengan Organisasi masyarakat berdasarkan Peraturan Presiden. Kerja sama berawal dari kesamaan orientasi. Dalam kerja sama, tugas-tugas yang dibebankan kepada tiap individu dapat berbeda satu sama lain (wikipedia,2024)

Kami menerapkan Kemampuan untuk berkomitmen dalam tim diwujudkan melalui rasa saling percaya, keinginan untuk saling membantu, komunikasi yang terbuka, dan fokus untuk mencapai tujuan bersama. Dalam setiap aktivitas tim di Telkomsel, kami diajarkan untuk:

1. Membangun Rasa Saling Percaya

Gambar 4. 5 Membagi Tugas

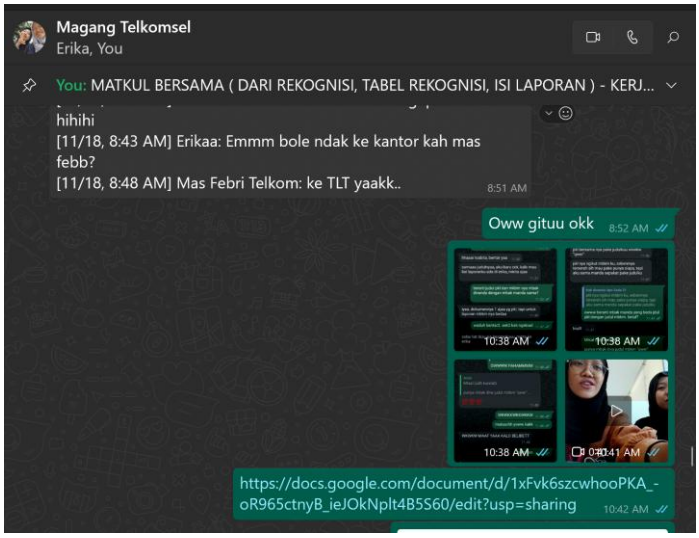
Memberikan kepercayaan penuh kepada rekan tim termasuk menghargai kemampuan dan tanggung jawab yang ia miliki dengan keyakinan bahwa kami akan menjalankan tugas masing-masing dengan baik dan sesuai ekspektasi.

2. Keinginan untuk Saling Membantu

Gambar 4. 6. Bukti Chat Saling Membantu

Bersikap proaktif dalam menawarkan bantuan mencerminkan inisiatif untuk mendukung rekan tim pada kami tanpa harus diminta, baik dalam menyelesaikan tugas yang menjadi tanggung jawab bersama maupun dalam membantu mereka menghadapi kendala atau hambatan yang muncul. Kami ada inisiatif saling membantu meskipun tidak bertatap muka. Sikap ini mencakup kesediaan untuk mendengarkan masalah, mencari solusi bersama, dan memberikan kontribusi secara nyata, sehingga dapat memperkuat hubungan kerja yang harmonis, meningkatkan efisiensi tim, serta memastikan setiap anggota merasa dihargai dan didukung dalam mencapai tujuan bersama.

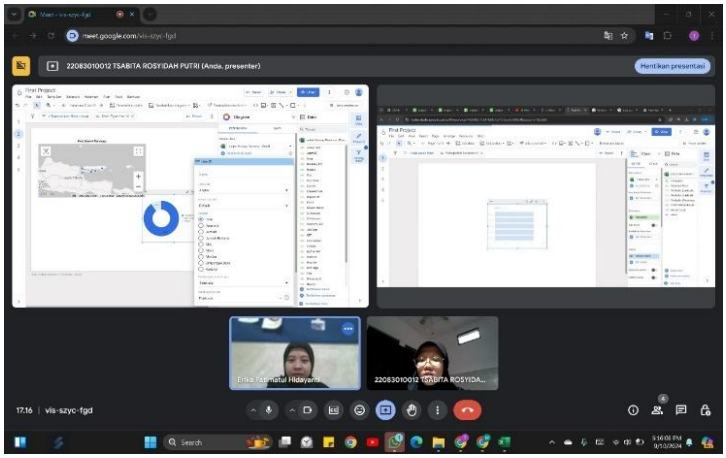
3. Komunikasi yang Terbuka



Gambar 4. 7 Bukti Chat Komunikasi Yang Terbuka

Mengedepankan transparansi dalam berdiskusi, berbagi informasi, dan memberikan umpan balik secara konstruktif adalah langkah penting untuk menciptakan komunikasi yang efektif dalam tim. Hal ini melibatkan keterbukaan dalam menyampaikan ide, pendapat, dan kendala yang dihadapi, sehingga anggota tim memiliki pemahaman yang sama dan dapat bekerja secara selaras. Kami membuat grup bersama untuk mengkomunikasi atau kabar terbaru dari mentor.

4. Mencapai Tujuan Bersama



Gambar 4. 8 Gmeet Bersama

Mengutamakan kepentingan bersama di atas kepentingan individu berarti setiap anggota tim menempatkan prioritas pada tujuan kolektif daripada kepentingan pribadi, sehingga tercipta sinergi yang kuat dalam bekerja. Sikap ini mencakup pengorbanan waktu, tenaga, atau ego demi keberhasilan tim, serta kesediaan untuk mendukung keputusan bersama meskipun mungkin berbeda dengan preferensi individu.

4.2.2. Pemanfaatan Sumber Daya Manusia Dalam Tim Secara Optimal [Indikator ke-2]

Kami di Telkomsel dilatih untuk memanfaatkan potensi sumber daya manusia secara optimal. Melalui interaksi dengan mentor, seperti :

Gambar 4. 9 Kolaborasi Dengan Mentor

1. Kolaborasi dengan Mentor atau Pemimpin Tim:
   Berinteraksi dengan mentor atau pemimpin tim yaitu Kak Febri untuk mendapatkan arahan, umpan balik, dan wawasan yang dapat meningkatkan kualitas pekerjaan tim.
2. Pengembangan Keterampilan Anggota Tim:
   Mendorong anggota tim untuk menggali potensi mereka melalui pelatihan, diskusi, dan proyek nyata, seperti yang dilakukan dalam diskusi dengan Kak Febri.
3. Pendelegasian Tugas yang Tepat:
   Mengalokasikan tugas berdasarkan keahlian dan minat individu agar setiap anggota tim dapat memberikan kontribusi terbaik mereka.
4. Membangun Lingkungan yang Mendukung:
   Menciptakan lingkungan kerja yang terbuka, mendukung inovasi, dan memberikan ruang bagi anggota tim untuk berkembang.
5. Pemanfaatan Proyek Nyata untuk Pengembangan:
   Menggunakan proyek nyata, seperti memanfaatkan data kunjungan sekolah dari Telkomsel yang kemudian disajikan melalui Looker Studio, sebagai media pembelajaran untuk mengembangkan keterampilan dan meningkatkan efisiensi kerja tim.

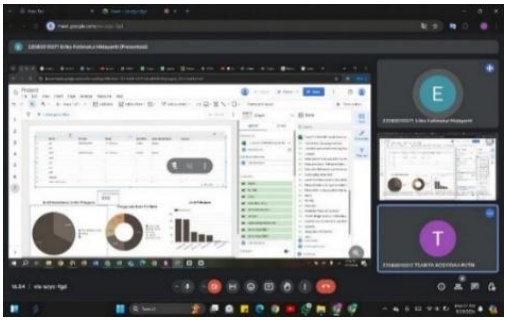
4.2.3. Pola Kerja Sama [Indikator ke-3]

Proyek-proyek nyata, seperti optimalisasi kampanye pemasaran dan analisis data pelanggan, menjadi sarana bagi mahasiswa untuk belajar menerapkan pola kerja sama yang efektif dalam lingkungan kerja. Pola kerja sama ini mencakup pembagian tugas yang jelas, komunikasi yang terbuka, koordinasi yang terstruktur, serta kemampuan untuk menyelesaikan konflik secara konstruktif.

Gambar 4. 10 Bekerja Sama

Melalui penerapan pola kerja sama tersebut, kami tidak hanya dilatih untuk bekerja secara individual, tetapi juga untuk berkolaborasi dalam tim dengan tujuan mencapai hasil yang maksimal. Selain itu, keterampilan kerja sama ini dirancang untuk meningkatkan efisiensi dalam penyelesaian proyek, memperkuat kemampuan analitis, dan membangun hubungan profesional yang mendukung pertumbuhan teamwork secara keseluruhan. Hal ini juga mempersiapkan kami untuk menghadapi tantangan di dunia kerja yang menuntut sinergi antar departemen dan pemahaman yang mendalam terhadap dinamika tim.

4.2.4. Mekanisme Monitoring Progress Pekerjaan Secara Berkala [Indikator ke-4]



Gambar 4. 11 Progress Melalui Gmeet

Mekanisme monitoring progress pekerjaan secara berkala adalah proses sistematis untuk memantau dan mengevaluasi kemajuan pekerjaan pada interval waktu tertentu. Proses ini dilakukan untuk memastikan bahwa semua tugas berjalan sesuai rencana, target tercapai, dan potensi masalah dapat diidentifikasi serta diatasi lebih awal. Monitoring dapat dilakukan melalui pertemuan tatap muka maupun secara daring menggunakan alat komunikasi dan kolaborasi digital.

Kami juga menerapkan mekanisme monitoring progress pekerjaan secara berkala, baik melalui pertemuan tatap muka di kantor maupun secara daring menggunakan Google Meet. Proses ini tidak hanya menjaga komunikasi tetap terbuka, tetapi juga memastikan kami memiliki pemahaman yang sama tentang perkembangan pekerjaan dan nilai kebersamaan tetap terjaga

4.3. Rekognisi Mata Kuliah Analisis Segmentasi Pelanggan

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Analisis Segmentasi Pelanggan. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Analisis Segmentasi Pelanggan

Tabel 4. 4 Indikator Penilaian Konversi Mata Kuliah Analisis Segmentasi Pelanggan

No	Kriteria/Indikator Penilaian
1	Kemampuan dalam menjelaskan konsep dasar analisis segmentasi pelanggan, identifikasi proses bisnis, tipe segmentasi pelanggan (demografi, geografis, perilaku, psikografis, dsb), tahapan-tahapan segmentasi pelanggan, dan pemanfaatan data segmentasi pelanggan untuk bisnis
2	Kemampuan memilih metode dan pendekatan analisis segmentasi pelanggan yang tepat yang dapat digunakan dalam strategi pemetaan target konsumen berdasarkan karakteristik tertentu
3	Kemampuan menggunakan bahasa pemrograman dan tools pendukung analisis segmentasi pelanggan dalam menyelesaikan studi kasus di bidang bisnis dan sasaran pasar secara efektif dan efisien
4	Kemampuan merancang dan memformulasikan, dan mendesain workflow analisis segmentasi pelanggan untuk menyelesaikan permasalahan pada suatu studi kasus tertentu

4.3.1. Materi Analisis Segmentasi Pelanggan [Indikator ke-1]

Konsep Dasar Analisis Segmentasi Pelanggan

Dalam merancang strategi pemasaran yang berorientasi pada pelanggan, segmentasi pelanggan memainkan peran penting dalam mengidentifikasi berbagai karakteristik yang dimiliki oleh pelanggan. Proses ini memungkinkan perusahaan untuk mengelompokkan pelanggan ke dalam kategori tertentu berdasarkan perilaku, kebiasaan, atau preferensi yang serupa. Dengan pendekatan ini, perusahaan dapat memahami kebutuhan setiap segmen pelanggan secara lebih mendalam dan menyusun strategi pemasaran yang lebih relevan dan efektif untuk menjangkau setiap kelompok tersebut. Hal ini tidak hanya meningkatkan kepuasan pelanggan, tetapi juga membantu perusahaan mencapai tujuan bisnis secara lebih optimal. Segmentasi pelanggan adalah proses membagi pasar yang luas menjadi kelompok-kelompok yang lebih kecil, di mana setiap kelompok atau segmen memiliki karakteristik, kebutuhan, atau perilaku yang serupa. Tujuan dari segmentasi pelanggan adalah untuk memahami perbedaan antara pelanggan dan menciptakan strategi pemasaran yang lebih terfokus dan efektif sesuai dengan kebutuhan masing-masing segmen (Brahmana dkk.,2020)

Segmentasi dalam konteks CRM

CRM merupakan sistem yang mengintegrasikan proses, kebijakan, dan teknologi untuk mengelola semua interaksi perusahaan dengan pelanggan. Dengan pendekatan ini, perusahaan dapat memahami kebutuhan dan preferensi pelanggan secara lebih baik, sehingga dapat meningkatkan kepuasan dan loyalitas mereka. Implementasi CRM sering kali melibatkan penggunaan perangkat lunak yang membantu mengumpulkan dan menganalisis data dari berbagai saluran komunikasi seperti email, telepon, media sosial, dan situs web (Solechan dkk.,2022). Dalam CRM, segmentasi pelanggan memainkan peran penting untuk:

- Memahami Pelanggan Secara Mendalam
  Memahami model Segmentasi Pelanggan Dengan Kernel K-Means Clustering, seperti membahas bagaimana segmentasi pelanggan dapat membantu perusahaan memahami perilaku dan preferensi pelanggan, termasuk identifikasi kelompok yang lebih sering menggunakan produk tertentu.Dengan segmentasi, perusahaan dapat mengetahui pola perilaku, kebutuhan, dan preferensi dari setiap kelompok pelanggan (Lubis & Ganesha, 2016). Dengan segmentasi, perusahaan dapat mengetahui pola perilaku, kebutuhan, dan preferensi dari setiap kelompok pelanggan.
- Meningkatkan Personalisasi
  Perancangan Analytical CRM untuk Mendukung Segmentasi Pelanggan di Institusi Pendidikan, seperti menjelaskan bagaimana data segmentasi memungkinkan perusahaan untuk memberikan penawaran yang lebih relevan dan personal, seperti diskon ulang tahun berdasarkan data pelanggan (Hananto dkk., 2017). Data segmentasi memungkinkan perusahaan memberikan penawaran atau layanan yang lebih relevan dan personal.
- Efisiensi dalam Pemasaran
  Memfokus pada segmen yang tepat, perusahaan dapat menghemat biaya pemasaran sekaligus meningkatkan efektivitas kampanye.
- Meningkatkan Loyalitas Pelanggan
  Memahami segmen pelanggan membantu perusahaan menciptakan program loyalitas yang sesuai dengan menunjukkan bagaimana memahami segmen pelanggan dapat membantu perusahaan menciptakan program loyalitas yang sesuai, seperti program *point rewards* untuk pelanggan premium.

Segmentasi dalam konteks CRM

Segmentasi pelanggan dapat dilakukan dengan pendekatan yang lebih spesifik tergantung pada kebutuhan bisnis. Berikut adalah jenis segmentasi beserta metode yang digunakan (equip, 2023):

a. Demografis
   Mengelompokkan pelanggan berdasarkan faktor-faktor demografis seperti usia, jenis kelamin, pendapatan, pendidikan, dan status pernikahan. Pendekatan ini membantu perusahaan memahami preferensi dan kebutuhan pelanggan berdasarkan karakteristik dasar mereka.
b. Geografis
   Membagi pelanggan berdasarkan lokasi geografis, termasuk negara, kota, dan iklim. Segmentasi ini memungkinkan perusahaan untuk menyesuaikan strategi pemasaran dengan preferensi lokal dan kondisi geografis yang berbeda
c. Perilaku
   Menganalisis perilaku pelanggan, termasuk kebiasaan membeli, frekuensi pembelian, dan respons terhadap promosi. Dengan memahami perilaku ini, perusahaan dapat merancang penawaran yang lebih relevan dan menarik bagi masing-masing segmen.
d. Nilai dan Manfaat
   Mengelompokkan pelanggan berdasarkan nilai yang mereka cari dari produk atau layanan. Segmentasi ini berfokus pada manfaat spesifik yang diinginkan pelanggan, memungkinkan perusahaan untuk menyesuaikan produk dan komunikasi pemasaran agar lebih sesuai dengan harapan pelanggan.

Ukuran Jarak & Ukuran Asosiasi

Perhitungan matriks ukuran jarak antar pelanggan

Perhitungan matriks ukuran jarak antar pelanggan adalah bagian penting dari manajemen logistik dan optimalisasi rute distribusi. Matriks ini dibuat dengan menghitung jarak antara setiap pasangan pelanggan. Ini dapat dilakukan dengan menggunakan berbagai rumus, tergantung pada jenis data lokasi yang tersedia, seperti rumus *Euclidean*. Dalam konteks distribusi, langkah pertama adalah menemukan jarak antara lokasi-lokasi yang relevan dengan menggunakan data koordinat atau aplikasi pemetaan seperti *Google Maps*. Setelah matriks jarak dibentuk, metode *saving matrix* dapat diterapkan untuk mengoptimalkan rute pengiriman dengan cara menggabungkan rute yang dapat menghemat total jarak tempuh. Dengan cara ini, perusahaan dapat merencanakan pengiriman yang lebih efisien, mengurangi biaya operasional, dan meningkatkan layanan pelanggan (Kasih & Maulidina, 2023). Berikut rumus *eculidean* jika dua lokasi memiliki koordinat (x_1, y_1) dan (x_2, y_2), maka jarak $J(1,2)$ adalah

$$J(1,2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Perhitungan ukuran Asosiasi

Perhitungan ukuran asosiasi adalah metode statistik yang digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel. Ukuran asosiasi yang umum digunakan termasuk koefisien Phi dan koefisien kontingensi. Koefisien *Phi* digunakan untuk tabel kontingensi 2x2 dan dihitung menggunakan rumus yang melibatkan nilai *Chi-Square*. Koefisien kontingensi adalah modifikasi dari koefisien Phi yang dapat digunakan untuk tabel yang lebih besar, memberikan nilai antara 0 dan 1 untuk menunjukkan kekuatan asosiasi; semakin mendekati 1, semakin kuat hubungan antara variabel (Intan Ratu dkk., 2013). Berikut adalah rumus koefesien Phi (Φ) dan Koefisien Kontingensi (C) :

- Koefesien Phi (Φ)

$$\Phi = \sqrt{\frac{x^2}{n}}$$

- Koefisien Kontingensi (C)

$$C = \sqrt{\frac{x^2}{x^2 + n}}$$

Dimana :
x^2 : Nilai *Chi-Square* yang dihitung dari tabel kontingensi.
n : Total jumlah data.

Deskripsi data customer secara numerik dan visual

Dalam analisis data, deskripsi data pelanggan secara numerik dan visual sangat penting untuk memahami karakteristik dan perilaku pelanggan. Data numerik dapat mencakup ukuran pusat seperti *mean*, *median*, dan modus, serta ukuran variabilitas seperti rentang, varians, dan deviasi standar. Misalnya, analisis deskriptif dapat

digunakan untuk menghitung rata-rata umur pelanggan atau total pengeluaran mereka dalam periode tertentu. Sementara itu, visualisasi data seperti grafik batang, diagram lingkaran, atau scatter plot dapat membantu menyajikan informasi ini dengan cara yang lebih mudah dipahami. Dengan menggunakan visualisasi, perusahaan dapat mengidentifikasi pola dan tren dalam data pelanggan, seperti segmentasi berdasarkan demografi atau perilaku pembelian. Ini tidak hanya memudahkan pengambilan keputusan tetapi juga membantu dalam merancang strategi pemasaran yang lebih efektif (Burnay dkk., 2019).

Analisis RFM

Analisis RFM (*Recency*, *Frequency*, *Monetary*) adalah metode yang digunakan untuk mengevaluasi nilai pelanggan berdasarkan tiga faktor utama: *Recency* (seberapa baru pelanggan melakukan pembelian), *Frequency* (seberapa sering pelanggan melakukan pembelian), dan *Monetary* (jumlah uang yang dibelanjakan oleh pelanggan). Metode ini membantu perusahaan dalam segmentasi pelanggan dengan memberikan skor kepada setiap pelanggan berdasarkan ketiga dimensi tersebut, sehingga memungkinkan penargetan yang lebih efektif dalam strategi pemasaran (algoritma, 2022). Keuntungan model RFM terletak pada relevansinya selamaberoperasi pada beberapa variabel yang dapat diamati dan bersifat obyektif. Variabel ini digolongkan menurut 3 kriteria, yaitu *recency*, *frequency*, dan *monetary*. *Recency* merupakan interval waktu antara pembelian terakhir dengan waktu sekarang, nilai yang lebih rendah sesuai dengan probabilitas yanglebih tinggi bahwa pelanggan akan melakukan pembelian ulang. *Frequency* adalah jumlah transaksi yang telah dilakukan pelanggan dalam jangka waktu tertentu dan *monetary* adalah jumlah uang yang dihabiskan dalam periode waktu yang ditentukan (Wang, C. H. ,2009). Berikut merupakan perhitungan skor RFM

$$skor\ RFM = ((nilai\ recency\ x\ 100) + (nilai\ frequency\ x\ 0) + nilai\ monetary$$

Tahapan Segmentasi Pelanggan

1. Survei
   Survei merupakan salah satu cara untuk mengeksplorasi target pelanggan untuk memahaminya secara mendalam. Misalnya dengan melalui wawancara kuesioner, dan lain sebagainya.
   Tahap survei ini memiliki peran yang sangat penting karena dapat membantu untuk mengumpulkan berbagai data informasi seputar sasaran pasar. Hal ini juga tentunya sudah sesuai dengan apa yang dibutuhkan oleh masyarakat.
2. Analisis
   Melakukan analisis informasi. Apabila data dari tahap survei telah didapatkan, selanjutnya yaitu menganalisis data informasi tersebut untuk mengetahui apa kesimpulannya. Hasil analisis tersebut digunakan sebagai acuan dasar untuk mengelompokkan pelanggan berdasarkan segmennya.
3. Identifikasi
   Melakukan identifikasi terhadap pelanggan. Setelah data konsumen didapatkan dan dianalisis, selanjutnya lakukan identifikasi pada setiap kelompok target pemasaran tersebut. Pada tahap ini, yang harus dilakukan adalah menemukan jenis pelanggan mana yang hendak menjadi target

pemasaran dalam bisnis yang dijalankan. Tapi identifikasi ini sangat penting karena menjadi tahap akhir untuk menentukan target konsumen.

Pemanfaatan Data Segmentasi Pelanggan untuk Bisnis

Pemanfaatan data segmentasi pelanggan memungkinkan bisnis untuk memahami kebutuhan, preferensi, dan perilaku pelanggan secara lebih spesifik, sehingga dapat merancang strategi yang lebih efektif. Dengan data ini, perusahaan dapat mengarahkan kampanye pemasaran secara terfokus, mengembangkan produk atau layanan yang relevan, menetapkan harga berdasarkan daya beli segmen, serta meningkatkan loyalitas pelanggan melalui program yang dipersonalisasi. Selain itu, segmentasi membantu mengoptimalkan distribusi produk ke lokasi yang paling potensial dan merencanakan aktivitas promosi atau event sesuai kebutuhan segmen target. Dengan pendekatan ini, perusahaan dapat meningkatkan efisiensi operasional, daya saing, dan potensi keuntungan secara signifikan (Monalisa & Nurahmah, 2024).

4.3.2. Clustering Menggunakan Metode K-Means: Segmentasi Pelanggan Berdasarkan Area, Rasio Aktivitas dan Omzet untuk Strategi Pemasaran Menggunakan Data Event Telkomsel Menggunakan Bahasa Pemrograman Python [Indikator ke-2, Indikator ke-3, dan Indikator ke-4]

Segmentasi pelanggan adalah proses pengelompokkan pelanggan berdasarkan karakteristik umum mereka. Segmentasi pelanggan bertujuan untuk membantu bisnis memahami kebutuhan dan perilaku pelanggannya agar dapat mengembangkan strategi pemasaran yang lebih efektif (Maretha, 2020). Dalam permasalahan ini telkomsel menghadapi tantangan untuk memahami perilaku pelanggan yang beragam terutama terkait dengan lokasi yang kurang strategis, rasio aktivitas pelanggan, dan kontribusi omzet mereka terhadap pendapatan perusahaan, maka penulis akan mengelompokkan pelanggan berdasarkan area, rasio aktivitas dan kontribusi omzet menggunakan metode K-Means *Clustering* dan juga merancang strategi pemasaran yang tepat sasaran untuk setiap segmen pelanggan berdasarkan karakteristik klaster yang ditemukan dengan melalui data *event* bisa menunjukkan bahwa sebagiaan besar pelanggan melalui *event* bisa mendeteksi adanya rasio aktivitas yang rendah ataupun yang tinggi.

- Metode K-Means Clustering

  Clusteringa dalah suatu teknik yang digunakan untuk mengelompokkan objek-objek data ke dalam kelompok-kelompok yang memiliki kesamaan tertentu (Witten dkk., 2016). Algoritma K-means mengelompokkan data ke dalam beberapa kelompok, di mana data dalam satu kelompok mempunyai karakteristik yang sama antara satu dengan yang lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Wanto dkk., 2020). Penulis menggunakan metode K-Means untuk mengelompokkan rasio aktivitas dan omzet sebagai dasar perancangan strategi pemasaran di berbagai area. Metode K-Means bekerja dengan meminimalkan nilai inertia, yaitu total jarak kuadrat antara titik data dan pusat klusternya. Nilai inertia dihitung dengan rumus berikut:

$$\sum_{i=1}^{n}\sum_{j=1}^{k} min\|x_i - c_j\|^2$$

Dimana :
x_i : Data ke-i
c_j : Pusat cluster ke-j
n : Jumlah data
k : Jumlah kluster

- *Silhouette Score*
  Silhouette Score adalah metrik evaluasi internal yang digunakan untuk mengevaluasi kualitas Clustering. Metrik ini mengukur seberapa baik setiap titik data ditempatkan dalam Cluster mereka sendiri dibandingkan dengan *Cluster* lain. *Silhouette Score* berkisar dari -1 hingga 1, di mana nilai yang lebih tinggi menunjukkan bahwa titik data lebih dekat ke *Centroid Cluster* mereka sendiri daripada ke *Centroid Cluster* lain

$$Silhouette\ Score = \frac{b - a}{\max(a, b)}$$

  a : Jarak rata-rata data ke data lain dalam cluster yang sama
  b : Jarak rata-rata data ke data di cluster terdekat

Selain metode K-Means, penulis juga memanfaatkan metode berbasis rumus statistik tertentu untuk mendukung analisis dan validasi hasil segmentasi, seperti :

- *Outlier Detection* (IQR Method)
  IQR (*Interquartile Range*) adalah ukuran statistik yang digunakan untuk mengukur sebaran data dengan cara menghitung selisih antara kuartil ketiga (Q3) dan kuartil pertama (Q1). Dalam analisis data, IQR berfungsi untuk mengidentifikasi *outlier*, yaitu nilai-nilai ekstrem yang berada jauh dari nilai lainnya dalam dataset (Wikipedia, 2023)
  Kuartil
  Kuartil Pertama (Q1)
  Nilai yang membagi 25% data terendah dari 75% data tertinggi.
  Kuartil Ketiga (Q3)
  Nilai yang membagi 75% data terendah dari 25% data tertinggi.

  Interquartile Range (IQR)
  IQR = Q3-Q1 (rentang antar kuartil)

  Batas untuk outlier
  Batas Bawah (*Lower Bound)* : Q1 – 1,5 × IQR
  Batas Atas (*Upper Bound)* : Q3 + 1,5 × IQR

  Oulier
  Data < *Lower Bound* < *Upper Bound*

- Normalisasi Data (*Standardization*)
  Standarisasi atau yang juga dikenal sebagai Z-Score *Normalization*, adalah teknik dalam statistika yang digunakan untuk mengubah data sehingga memiliki mean (rata-rata) nol dan deviasi standar satu. Tujuan dari standarisasi adalah untuk membuat distribusi data menjadi distribusi normal atau mendekati normal. Rumus standarisasi untuk mengubah setiap nilai dalam dataset menjadi nilai Z-score adalah sebagai berikut:

$$Z = \frac{x - \mu}{\sigma}$$

Dimana :
Z : Nilai standar (Z-score) dari x
x : Nilai dataset
μ : Rata-rata dari dataset,
σ : Deviasi standar dari dataset.

- Pseudo F-*Statistic*
  Pseudo F-*statistic* adalah rasio yang mengukur pembagian antara variansi dalam cluster (WCSS) dan variansi antara cluster (BCSS). Semakin besar nilai Pseudo F, semakin baik pemisahan antar-cluster.

$$P\ seudo\ F = \frac{BCSS\ /\ (k-1)}{WCSS\ /\ (N-K)}$$

Mengukur kualitas kluster
Within-Cluster Sum of Squares (WCSS) untuk mengukur sejauh mana data dalam satu cluster terdistribusi di sekitar pusat cluster (*centroid*). Semakin kecil nilai WCSS, semakin baik klaster yang dihasilkan.

$$WCSS = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \|x_j - c_i\|^2$$

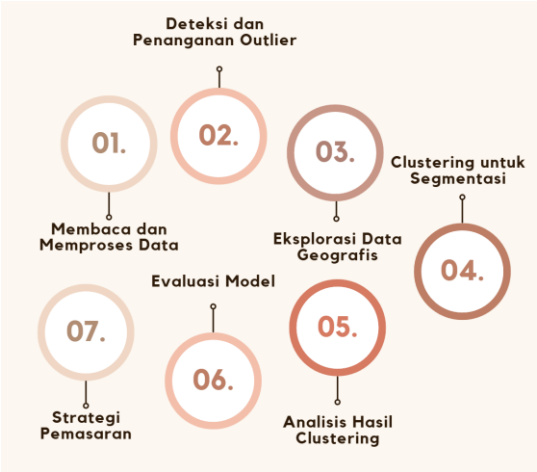
Dimana :
k : Jumlah cluster
n_i: Jumlah data dalam cluster ke- i
x_j : Titik data dalam cluster i
c_i : Centroid (pusat) dari cluster i

Between-Cluster Sum of Squares (BCSS) untuk mengukur sejauh mana centroid setiap cluster terdistribusi dibandingkan dengan rata-rata keseluruhan data (mean global). Semakin besar nilai BCSS, semakin baik pemisahan antara cluster.

$$BCSS = \sum_{i=1}^{k} n_i \|c_i - \mu\|^2$$

Dimana :
k : Jumlah cluster
n_i : Jumlah data dalam cluster ke-i
c_i: Centroid dari cluster i
μ : Rata-rata (mean) dari seluruh data

Berikut merupakan workflow untuk menganalisis segmentasi pelanggan menggunakan metode K-Means :



Gambar 4. 12 Workflow Analisis Segmentasi Pelanggan

1. Membaca dan Memproses Data
   Pertama, data CSV dibaca menggunakan pandas.read_csv untuk memuat dataset ke dalam sebuah DataFrame. Kemudian, kolom 'OMSET TOTAL.1' dihapus karena tidak memberikan kontribusi signifikan untuk analisis lebih lanjut. Selanjutnya, semua nilai kosong (NaN) dalam dataset diisi dengan nilai 0 untuk memastikan tidak ada nilai yang hilang, yang dapat mengganggu proses analisis dan pemodelan. Setelah itu, baris duplikat dalam dataset dihapus untuk memastikan bahwa hanya data yang unik yang tersisa, menghindari bias dan redundansi dalam analisis. Terakhir, kolom-kolom numerik seperti 'DS RATIO (%)' dan 'ACTUAL CALL (%)' yang sebelumnya mungkin berformat non-numerik dikonversi menjadi tipe data numerik untuk memudahkan perhitungan dan analisis lebih lanjut, dengan nilai-nilai tersebut dibulatkan ke bilangan bulat untuk keseragaman dan ketepatan analisis. Dan membuat kolom baru yaitu 'ACTIVE RATIO'
2. Deteksi dan Penanganan *Outlier*
   Proses deteksi dan penanganan outlier dimulai dengan menggunakan metode *Interquartile Range* (IQR) untuk mengidentifikasi data yang berada di luar batas wajar, yaitu nilai yang berada di bawah batas bawah atau di atas batas atas IQR. Setelah itu, untuk menangani outlier yang terdeteksi, metode *Winsorization* diterapkan dengan membatasi nilai-nilai ekstrem pada batas bawah dan atas yang wajar, sehingga outlier tersebut diubah menjadi nilai yang masih berada dalam rentang yang lebih masuk akal dan tidak mempengaruhi analisis secara berlebihan. Terakhir, setelah proses pembersihan data selesai, semua nilai numerik yang ada dalam dataset dikonversi menjadi tipe data int64 untuk memastikan konsistensi dan memungkinkan perhitungan yang lebih efisien dalam analisis selanjutnya.
3. Eksplorasi Data Geografis
   Menghitung jumlah pelanggan yang terdistribusi di setiap regional, branch, dan cluster dengan. Proses ini memberikan gambaran yang jelas tentang bagaimana pelanggan tersebar di berbagai lokasi dan kategori yang relevan. Setelah menghitung distribusi pelanggan, langkah berikutnya adalah membuat visualisasi dalam bentuk grafik batang untuk menampilkan 10 *regional*,*branch*, dan *cluster* dengan jumlah pelanggan terbanyak.
4. Clustering untuk Segmentasi
   Memilih kolom yang akan dianalisis yaitu 'OMSET TOTAL' dan 'ACTIVE RATIO', yang dapat memberikan wawasan terkait perilaku pelanggan atau area

bisnis. untuk menentukan jumlah cluster yang optimal, digunakan teknik yaitu *Elbow Method*. Terakhir adalah algoritma KMeans Clustering diterapkan untuk membagi data ke dalam beberapa cluster berdasarkan pola dan karakteristik yang ditemukan, dengan tujuan untuk mengelompokkan pelanggan.

5. Analisis Hasil *Clustering*
   Menentukan area dengan pelanggan terbanyak dan performa tinggi untuk difokuskan pada strategi promosi dan *event* dan menentukan area dengan performa rendah dan potensi pertumbuhan untuk strategi penetrasi pasar.
6. Evaluasi Model
   Evaluasi model k-means menggunakan sihoutte score
7. Strategi Pemasaran
   Melakukan strategi pemasaran dan memberikan rekomendasi pada produk yang akan dilakukan

Berikut merupakan langkah-langkah dari pengerjaan mata kuliah Analisis Segmentasi Pelanggan, yaitu :

a. Mengimpor library

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
from scipy.spatial.distance import cdist
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_score
```

Library-library yang akan digunakan yaitu memiliki fungsi utama untuk mendukung analisis data, klastering, visualisasi, dan evaluasi model. Pandas digunakan untuk manipulasi dan analisis data, seperti membaca, mengelola, dan membersihkan data dalam struktur DataFrame. NumPy mendukung komputasi numerik dengan array multidimensi, memungkinkan perhitungan statistik, algebra linear, dan transformasi data. Untuk visualisasi, Matplotlib menyediakan grafik dasar seperti garis, batang, atau scatter plot, sementara Seaborn memberikan visualisasi statistik yang lebih estetis dan informatif, seperti heatmap atau boxplot. Dalam klastering, KMeans dari sklearn.cluster digunakan untuk membagi data ke dalam kelompok berdasarkan kesamaan fitur, dengan data yang distandarisasi menggunakan ElbhowMethod. Selain itu, cdist dari scipy.spatial.distance menghitung jarak antar titik, berguna untuk metode seperti Elbow Method dalam menentukan jumlah klaster optimal. Kombinasi ini memungkinkan analisis data yang komprehensif dan mendalam. StandardScaler dari pustaka sklearn.preprocessing digunakan untuk standarisasi fitur dalam machine learning. Standarisasi mengubah data agar memiliki rata-rata (mean) 0 dan standar deviasi 1, sehingga data terdistribusi secara normal dalam skala yang seragam. from sklearn.metrics import silhouette_score untuk mengukur kualitas hasil clustering,

b. Membaca dan memproses data

	AREA	REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO (%)	OMSET TRX DIGITAL	OMSET TOTAL	...	ACTUAL POI	PLAN EVENT	ACTUAL EVENT	ACTUAL CALL (%)	SALES	OTA	SALES ORBIT	VF	TRX DIGITAL	ACTIVE RATIO
0	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	BULELENG	42	3	7	39	11043601	...	37	72	50	69	0	0	0	0	29	0.071429
1	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	JEMBRANA	23	3	13	39	14135322	...	43	75	56	75	0	0	0	0	30	0.130435
2	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	TABANAN	20	3	15	37	8446941	...	46	53	48	91	5	0	0	0	36	0.150000
3	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI TENGAH	BADUNG	42	6	14	92	24862947	...	77	156	100	64	59	0	0	125	89	0.142857
4	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI TENGAH	KOTA DENPASAR	107	6	6	84	32561574	...	29	152	105	69	6	0	0	38	80	0.056075
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1066	JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	13	7	18	9563982	...	43	51	49	96	24	0	0	0	18	0.065990
1067	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	36	4	11	0	3282292	...	1	4	4	100	0	0	0	0	0	0.111111
1068	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	32	4	12	0	3308591	...	1	4	4	100	11	0	0	0	0	0.125000
1069	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	4	17	0	1364900	...	1	5	4	80	37	0	0	0	0	0.166667
1070	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4	15	0	2593442	...	2	5	5	100	11	0	0	0	0	0.153846

1063 rows × 25 columns

Gambar 4. 13 Data Events

Data Events by Telkomsel memiliki 1.071 baris dan 24 kolom setelah membersihkan data. Dengan variabel ‘AREA’ wilayah operasional atau area cakupan yang lebih, ‘REGIONAL’ Pembagian wilayah operasional yang lebih spesifik dari Area, ‘BRANCH’ cabang atau kantor operasional Telkomsel di wilayah tertentu, ‘CLUSTER’ pembagian wilayah yang lebih spesifik, ‘CITY’ kota tempat aktivitas operasional atau pemasaran Telkomsel dilakukan, ‘DS TOTAL’ jumlah transaksi penjualan produk, ‘DS ACTIVE‘ jumlah Digital Sales (DS) yang aktif dalam periode tertentu, ‘DS RATIO (%)’ total Digital Sales dalam bentuk persentase, ‘OMSET TRX DIGITAL’ Total omzet yang dihasilkan dari transaksi digital, ‘OMSET TOTAL’ total omzet yang dihasilkan dari semua transaksi, ‘PLAN POI’ target jumlah Point of Interest (POI), ‘ACTUAL POI’ Jumlah Point of Interest, ‘PLAN EVENT’ target jumlah event atau kegiatan pemasaran yang direncanakan, ‘ACTUAL EVENT’ jumlah event atau kegiatan pemasaran, ‘ACTUAL CALL (%)’ persentase panggilan yang berhasil dilakukan dibandingkan target panggilan yang direncanakan, ‘SALES’ jumlah penjualan produk, ‘OTA’ penjualan produk Orbit Telkomsel, ‘SALES ORBIT’ jumlah unit Orbit yang berhasil terjual, ‘VF’ jumlah atau kontribusi dari penjualan voucher fisik, ‘TRX DIGITAL’ jumlah transaksi digital yang berhasil.

c. Menangani, Mendeteksi Oulier dan Hasil Statistika Deskriptif

```
Jumlah Outlier di kolom 'DS TOTAL': 72
Jumlah Outlier di kolom 'DS ACTIVE': 58
Jumlah Outlier di kolom 'DS RATIO (%)': 37
Jumlah Outlier di kolom 'OMSET TRX DIGITAL': 155
Jumlah Outlier di kolom 'OMSET TOTAL': 92
Jumlah Outlier di kolom 'OMSET DATA': 108
Jumlah Outlier di kolom 'OMSET DIGITAL': 120
Jumlah Outlier di kolom 'OMSET RECHARGE': 128
Jumlah Outlier di kolom 'OMSET OTHER': 127
Jumlah Outlier di kolom 'PLAN POI': 53
Jumlah Outlier di kolom 'ACTUAL POI': 59
Jumlah Outlier di kolom 'PLAN EVENT': 60
Jumlah Outlier di kolom 'ACTUAL EVENT': 70
Jumlah Outlier di kolom 'ACTUAL CALL (%)': 22
Jumlah Outlier di kolom 'SALES': 86
Jumlah Outlier di kolom 'OTA': 0
Jumlah Outlier di kolom 'SALES ORBIT': 0
Jumlah Outlier di kolom 'VF': 188
Jumlah Outlier di kolom 'TRX DIGITAL': 134
```

Gambar 4. 14 Output Jumlah Outlier

```
Jumlah Outlier di kolom 'DS TOTAL' setelah ditangani: 0
Jumlah Outlier di kolom 'DS ACTIVE' setelah ditangani: 0
Jumlah Outlier di kolom 'DS RATIO (%)' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET TRX DIGITAL' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET TOTAL' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET DATA' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET DIGITAL' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET RECHARGE' setelah ditangani: 0
Jumlah Outlier di kolom 'OMSET OTHER' setelah ditangani: 0
Jumlah Outlier di kolom 'PLAN POI' setelah ditangani: 0
Jumlah Outlier di kolom 'ACTUAL POI' setelah ditangani: 0
Jumlah Outlier di kolom 'PLAN EVENT' setelah ditangani: 0
Jumlah Outlier di kolom 'ACTUAL EVENT' setelah ditangani: 0
Jumlah Outlier di kolom 'ACTUAL CALL (%)' setelah ditangani: 0
Jumlah Outlier di kolom 'SALES' setelah ditangani: 0
Jumlah Outlier di kolom 'OTA' setelah ditangani: 0
Jumlah Outlier di kolom 'SALES ORBIT' setelah ditangani: 0
Jumlah Outlier di kolom 'VF' setelah ditangani: 0
Jumlah Outlier di kolom 'TRX DIGITAL' setelah ditangani: 0
```

Gambar 4. 15 Output Setelah Menangani Outlier

Statistik Deskriptif Data:

	DS TOTAL	DS ACTIVE	DS RATIO (%)	OMSET TRX DIGITAL	OMSET TOTAL	OMSET DATA	OMSET DIGITAL	OMSET RECHARGE	OMSET OTHER	PLAN POI	ACTUAL POI	PLAN EVENT	ACTUAL EVENT	ACTUAL CALL (%)	SALES	OTA	SALES ORBIT	VF	TRX DIGITAL	ACTIVE RATIO
count	1063.000000	1063.000000	1063.000000	1063.000000	1.063000e+03	1.063000e+03	1063.000000	1.063000e+03	1.063000e+03	1063.000000	1063.000000	1063.000000	1063.000000	1063.000000	1063.000000	1063.0	1063.0	1063.0	1063.000000	1063.000000
mean	23.034807	3.844779	18.038570	7.018815	1.490695e+07	4.534603e+06	161066.751646	3.119645e+06	4.194086e+06	45.971778	40.132643	71.228598	57.382879	83.566322	138.873942	0.0	0.0	0.0	9.193791	0.185090
std	11.424019	1.931352	8.146691	10.039452	1.583193e+07	5.042405e+06	214112.173169	3.642846e+06	4.878701e+06	37.622382	31.887017	46.502672	35.812721	13.866768	169.829967	0.0	0.0	0.0	12.709416	0.097071
min	3.000000	0.000000	0.000000	0.000000	0.000000e+00	0.000000e+00	0.000000	0.000000e+00	0.000000e+00	0.000000	0.000000	0.000000	0.000000	45.000000	0.000000	0.0	0.0	0.0	0.000000	0.000000
25%	14.000000	3.000000	12.000000	0.000000	2.892525e+06	7.163250e+05	0.000000	3.325000e+05	4.876870e+05	16.000000	15.000000	40.000000	34.000000	75.000000	2.000000	0.0	0.0	0.0	0.000000	0.125000
50%	21.000000	3.000000	17.000000	1.000000	8.425475e+06	2.294630e+06	50660.000000	1.495000e+06	1.902339e+06	37.000000	32.000000	63.000000	51.000000	86.000000	62.000000	0.0	0.0	0.0	2.000000	0.166667
75%	28.000000	5.000000	23.000000	11.000000	2.191753e+07	6.657713e+06	245024.500000	4.488000e+06	6.161750e+06	66.000000	56.500000	95.000000	75.000000	95.000000	215.000000	0.0	0.0	0.0	14.000000	0.230769
max	49.000000	8.000000	39.000000	27.000000	5.045504e+07	1.556980e+07	612561.000000	1.072125e+07	1.467284e+07	141.000000	118.000000	177.000000	136.000000	100.000000	534.000000	0.0	0.0	0.0	35.000000	0.750000

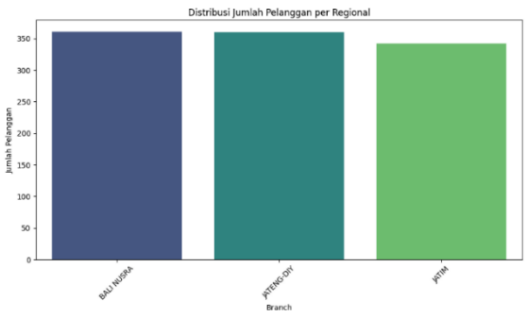
Gambar 4. 16 Statistika Deskriptif

Dapat disimpulkan bahwa rata-rata DS TOTAL adalah 23, sedangkan rata-rata DS ACTIVE hanya 3,8. Ini menunjukkan sebagian besar DS tidak aktif. Rasio DS ACTIVE terhadap total DS (DS RATIO) rata-rata adalah 18%. Untuk omset total memiliki rata-rata 7,02 juta, dengan omset digital rata-rata 161 ribu, yang merupakan kontribusi kecil terhadap total omset. Sebagian besar kontribusi berasal dari kategori omset data (rata-rata 4,53 juta) dan recharge (3,12 juta).

Omset total memiliki rata-rata 7,02 juta, dengan omset digital rata-rata 161.000, yang merupakan kontribusi kecil terhadap total omset. Sebagian besar kontribusi berasal dari kategori omset data (rata-rata 4,53 juta) dan recharge (3,12 juta). untuk perencanaan dan realisasi POI, event, dan call menunjukkan ada celah antara target dan pencapaian, seperti rata-rata plan event adalah 71 tetapi actual hanya 57.

Sales rata-rata adalah 138, dengan distribusi yang cukup besar (standar deviasi 169). Transaksi digital rata-rata rendah (9,2), menunjukkan adopsi digital masih kurang, dan active ratio rata-rata hanya 18,5%, yang menunjukkan hanya sebagian kecil individu yang aktif secara signifikan.

d. Eksplorasi data Geografis



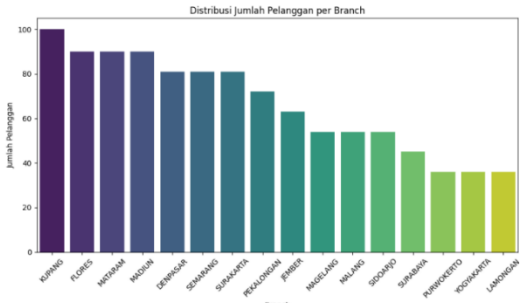
Gambar 4. 18 Grafik Regional

```
REGIONAL
BALI NUSRA    361
JATENG-DIY    360
JATIM         342
Name: count, dtype: int64
```

Gambar 4. 17 Jumlah Pada Regional

```
BRANCH
KUPANG        100
FLORES         90
MATARAM        90
MADIUN         90
DENPASAR       81
SEMARANG       81
SURAKARTA      81
PEKALONGAN     72
JEMBER         63
MAGELANG       54
MALANG         54
SIDOARJO       54
SURABAYA       45
PURWOKERTO     36
YOGYAKARTA     36
LAMONGAN       36
Name: count, dtype: int64
```

Gambar 4. 20 Jumlah Branch

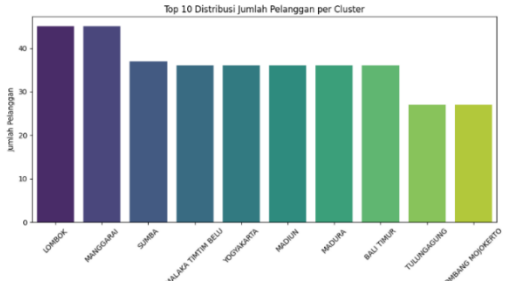


Gambar 4. 19 Grafik Branch

Dapat disimpulkan bahwa Branch tertinggi dengan jumlah pelanggan yaitu 100 orang pada daerah Kupang dan Branch terendah dengan jumlah pelanggan yaitu 36 orang pada daerah Lamongan

```
CLUSTER
LOMBOK                45
MANGGARAI             45
SUMBA                 37
MALAKA TIMTIM BELU    36
YOGYAKARTA            36
MADIUN                36
MADURA                36
BALI TIMUR            36
TULUNGAGUNG           27
JOMBANG MOJOKERTO     27
Name: count, dtype: int64
```

Gambar 4. 21 Jumlah Cluster



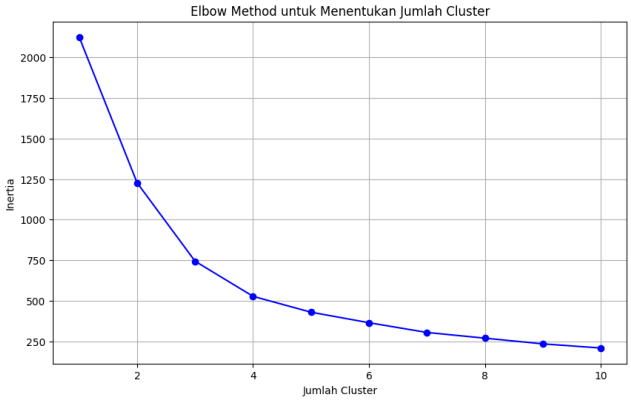
Gambar 4. 22 Grafik Cluster

Dapat disimpulkan bahwa Cluster tertinggi dengan jumlah pelanggan yaitu 45 orang di daerah Lombok dan Cluster terendah dengan jumlah pelanggan yaitu 27 orang di daerah Jombang Mojokerto.

e. Clustering untuk segmentasi menggunakan K-Means

```
# Persiapan Data
# Memilih kolom yang relevan untuk clustering
try:
    features = data[['OMSET TOTAL', 'ACTIVE RATIO']]
except KeyError as e:
     raise KeyError("Kolom yang dipilih tidak ditemukan dalam
dataset: 'OMSET TOTAL' atau 'ACTIVE RATIO'")

Normalisasi Data
# Standardisasi fitur agar skala variabel tidak mempengaruhi
hasil clustering
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features)

Menentukan Jumlah Cluster yang Optimal
# Elbow Method untuk menentukan jumlah cluster yang optimal
inertia = []
```

```
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=42, n_init=10)
    kmeans.fit(features_scaled)
    inertia.append(kmeans.inertia_)

# Visualisasi Elbow Method
plt.figure(figsize=(10, 6))
plt.plot(range(1, 11), inertia, marker='o', color='b',
label='Inertia')
plt.xlabel('Jumlah Cluster')
plt.ylabel('Inertia')
plt.title('Elbow Method untuk Menentukan Jumlah Cluster')
plt.grid(True)
plt.show()
```



Gambar 4. 23 Plot Elbhow Method

Dari grafik metode *elbow* tersebut, jumlah cluster yang terbaik terlihat ada perubahan yang signifikan pada nilai inertia jumlah cluster terbaik Dengan demikian, jumlah cluster terbaik adalah 3.

f. Analisis hasil clustering

```
# Menentukan jumlah cluster optimal
optimal_clusters = 3

# Clustering dengan K-Means
kmeans = KMeans(n_clusters=optimal_clusters, random_state=42,
n_init=10)
data['Cluster'] = kmeans.fit_predict(features_scaled)

# Menampilkan Hasil Clustering
print("\nHasil Clustering Pelanggan:")
display(data[['BRANCH', 'CLUSTER', 'CITY', 'Cluster']].head())

# Analisis Hasil Clustering
print("\nJumlah Pelanggan per Cluster:")
display(data['Cluster'].value_counts())
```

Kode di atas melakukan klasterisasi data menggunakan algoritma K-Means setelah menentukan jumlah klaster yang optimal, yaitu 3.

n_clusters=optimal_clusters digunakan untuk menentukan jumlah klaster yang diinginkan. random_state=42 mengatur nilai acak agar hasil klasterisasi dapat direproduksi. n_init=10 menetapkan jumlah pengulangan inisialisasi centroid untuk mencari hasil yang lebih stabil. fit_predict(features_scaled) menjalankan K-Means pada data yang telah distandarisasi dan mengembalikan hasil prediksi klaster untuk setiap baris data, yang kemudian disimpan dalam kolom baru 'Cluster' pada dataset.

Hasil Clustering Pelanggan:

	BRANCH	CLUSTER	CITY	Cluster
0	DENPASAR	BALI BARAT	BULELENG	0
1	DENPASAR	BALI BARAT	JEMBRANA	0
2	DENPASAR	BALI BARAT	TABANAN	0
3	DENPASAR	BALI TENGAH	BADUNG	0
4	DENPASAR	BALI TENGAH	KOTA DENPASAR	1

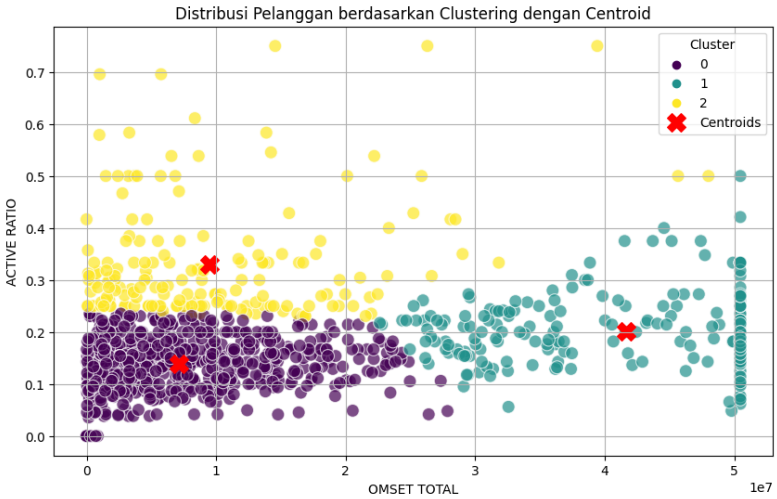
Gambar 4. 25 Hasil Cluster Pelanggan

```
Jumlah Pelanggan per Cluster:

Cluster
0    654
1    227
2    182
Name: count, dtype: int64
```

Gambar 4. 24 Jumlah Pelanggan

Dapat disimpulkan bahwa hasil pengelompokan pelanggan menunjukkan bahwa Cluster 0 memiliki jumlah pelanggan terbanyak dibandingkan dengan Cluster 1 dan Cluster 2. Dengan jumlah sebanyak 654 pelanggan, Cluster 0 mendominasi distribusi pelanggan, menunjukkan bahwa segmen ini mencakup mayoritas dari total pelanggan yang dianalisis. Hal ini bisa berarti bahwa pelanggan dalam cluster ini memiliki karakteristik, kebutuhan, atau perilaku yang lebih umum dan banyak ditemukan dibandingkan dengan dua cluster lainnya.

g. Visualisasi hasil clustering dengan centroid



Gambar 4. 26 Visualisasi Cluster

Dapat simpulkan bahwa **Cluster 0 (Ungu)** yaitu pelanggan yang memiliki kontribusi rendah terhadap total omset perusahaan, rasio aktivitas pelanggan sangat rendah (sekitar 0-0.2), dan kemungkinan besar ini adalah pelanggan yang kurang loyal atau hanya melakukan transaksi sekali atau dua kali. **Cluster 2 (Kuning)** yaitu pelanggan yang memiliki tingkat aktivitas yang cukup tinggi, namun omset totalnya bervariasi, mungkin kelompok pelanggan yang sering berinteraksi tetapi tidak semua dari mereka memberikan kontribusi besar dalam hal transaksi, dan pelanggan dalam kelompok ini bisa menjadi target untuk

program loyalitas atau promosi khusus untuk meningkatkan transaksi mereka. **Cluster 1 (Hijau)** yaitu pelanggan yang memberikan kontribusi omset yang tinggi, rasio aktivitas pelanggan cukup stabil (dalam rentang 0.2 hingga 0.5), dan ini adalah segmen pelanggan premium yang berkontribusi signifikan pada pendapatan perusahaan dan harus dipertahankan.

Centroid dalam gambar menunjukkan titik tengah dari masing-masing klaster, yang mewakili karakteristik rata-rata setiap kelompok pelanggan. Dengan memanfaatkan data centroid, perusahaan dapat merancang strategi pemasaran yang lebih tepat sasaran untuk masing-masing klaster. Misalnya **Cluster 0** Fokus pada reaktivasi pelanggan melalui promosi menarik. **Cluster 2** Program loyalitas untuk meningkatkan omset. **Cluster 1** Peningkatan layanan premium untuk mempertahankan loyalitas pelanggan utama.

```
Penjelasan untuk Setiap Cluster:

Cluster 0:
- Rata-rata OMSET TOTAL: 7126530.36 (Rendah)
- Rata-rata ACTIVE RATIO: 0.14
- Cluster ini memiliki potensi rendah berdasarkan OMSET TOTAL.

Cluster 1:
- Rata-rata OMSET TOTAL: 41639655.48 (Tinggi)
- Rata-rata ACTIVE RATIO: 0.20
- Cluster ini memiliki potensi tinggi berdasarkan OMSET TOTAL.

Cluster 2:
- Rata-rata OMSET TOTAL: 9522700.51 (Rendah)
- Rata-rata ACTIVE RATIO: 0.33
- Cluster ini memiliki potensi rendah berdasarkan OMSET TOTAL.
```

Gambar 4. 27 Hasil Setiap Cluster

h. Evaluasi Model menggunakan Silhoutte Score

```
# Menghitung Silhouette Score
sil_score = silhouette_score(features_scaled, data['Cluster'])

# Menampilkan hasil Silhouette Score
print(f"Silhouette Score: {sil_score:.4f}")

# Kesimpulan berdasarkan nilai Silhouette Score
if sil_score > 0.5:
    print("Model clustering bagus! Cluster terpisah dengan jelas
dan berbeda.")
elif sil_score > 0:
    print("Model clustering dapat diterima. Ada beberapa tumpang
tindih antar cluster, tetapi masih bisa dibedakan.")
else:
    print("Model clustering buruk. Cluster tidak terpisah dengan
baik, dan titik data mungkin masuk ke cluster yang salah.")
```

Didapatkan bahwa Silhouette Score: 0.5079, Model clustering bagus! Cluster terpisah dengan jelas dan berbeda.

i. Validasi Pseudo F-statistic

```
# Validasi Pseudo F-statistic
# Menghitung Within-Cluster Sum of Squares (WCSS) dan Between-
Cluster Sum of Squares (BCSS)
wcss = kmeans.inertia_  # WCSS sudah dihitung oleh K-Means
```

```
bcss = np.sum(cdist(features_scaled, centroids,
'euclidean')**2) - wcss

# Pseudo F-statistic
pseudo_f = (bcss / (optimal_clusters - 1)) / (wcss /
(features_scaled.shape[0] - optimal_clusters))
print(f"\nPseudo F-statistic: {pseudo_f:.4f}")

# Evaluasi Pseudo F-statistic
if pseudo_f > 1:
    print("Cluster dianggap sangat baik berdasarkan Pseudo F-
statistic.")
elif pseudo_f > 0.5:
    print("Cluster cukup baik, tetapi ada kemungkinan untuk
meningkatkan kualitas clustering.")
else:
    print("Cluster kurang baik berdasarkan Pseudo F-
statistic.")
```

```
Pseudo F-statistic: 8319.1676
Cluster dianggap sangat baik berdasarkan Pseudo F-statistic.
```

Gambar 4. 28 Hasil Pseudo F-Statistik

j. Strategi Pemetaan Target Konsumen

```
Strategi untuk Setiap Cluster:

Cluster 0: Fokus pada segmentasi ulang dengan memberikan tawaran produk yang lebih relevan atau peningkatan layanan.

Cluster 1: Fokus pada retensi pelanggan melalui program loyalitas, peningkatan produk, dan peningkatan kualitas layanan.

Cluster 2: Fokus pada peningkatan omzet melalui penawaran khusus dan promosi untuk mendorong aktivitas lebih tinggi.
```

Gambar 4. 29. Hasil Strategi Setiap Cluster

k. Implementasi Strategi dalam Pemetaan Target Konsumen Berdasarkan Area, Rasio Aktivitas, dan Omset Total

Pemetaan Target Konsumen Berdasarkan Cluster:

Cluster 0:

	BRANCH	CLUSTER	CITY	OMSET TOTAL	ACTIVE RATIO
0	DENPASAR	BALI BARAT	BULELENG	11043601	0.071429
1	DENPASAR	BALI BARAT	JEMBRANA	14135322	0.130435
2	DENPASAR	BALI BARAT	TABANAN	8446941	0.150000
3	DENPASAR	BALI TENGAH	BADUNG	24862947	0.142857
5	DENPASAR	BALI TIMUR	BANGLI	14542815	0.120000

Cluster 1:

	BRANCH	CLUSTER	CITY	OMSET TOTAL	ACTIVE RATIO
4	DENPASAR	BALI TENGAH	KOTA DENPASAR	32561574	0.056075
9	FLORES	ENDE SIKKA	ENDE	50455035	0.200000
10	FLORES	ENDE SIKKA	SIKKA	50455035	0.173913
11	FLORES	FLORES TIMUR	ALOR	49888637	0.333333
12	FLORES	FLORES TIMUR	FLORES TIMUR	35801991	0.230769

Cluster 2:

	BRANCH	CLUSTER	CITY	OMSET TOTAL	ACTIVE RATIO
13	FLORES	FLORES TIMUR	LEMBATA	47998003	0.50
16	FLORES	MANGGARAI	MANGGARAI TIMUR	9070265	0.25
21	KUPANG	KUPANG ROTE	ROTE NDAO	7533238	0.25
22	KUPANG	MALAKA TIMTIM BELU	BELU	19468159	0.25
23	KUPANG	MALAKA TIMTIM BELU	MALAKA	25857025	0.50

Gambar 4. 30 Pemetaan Target Berdasarkan Cluster

Didapatkan bahwa cluster 0 memiliki kontribusi yang rendah terhadap omset total perusahaan dan yang hanya melakukan transaksi sekali atau dua kali didapatkan pada daerah di Denpasar. Untuk cluster 1 memiliki kontribusi yang cukup stabil dengan omset yang tinggi pada daerah Denpasar dan Flores, sehingga segmen pelanggan pada cluster 1 adalah pelanggan premium dan harus dipertahankan. Yang terakhir adalah cluster 2 memiliki tingkat aktivitas yang cukup tinggi, namun untuk omset totalnya bervarias dalam hal ini pelanggan dalam cluster 2 bisa menjadi terget untuk program loyalitas atau promosi khusus untuk meningkatkan transaksi mereka dan didapatkan pada daerah Flores dan Kupang.

4.4. Rekognisi Mata Kuliah Instuisi dan Wawasan Data

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Instuisi dan Wawasan Data. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Instuisi dan Wawasan Data

Tabel 4. 5 Indikator Penilaian Konversi Mata Kuliah Instuisi dan Wawasan Data

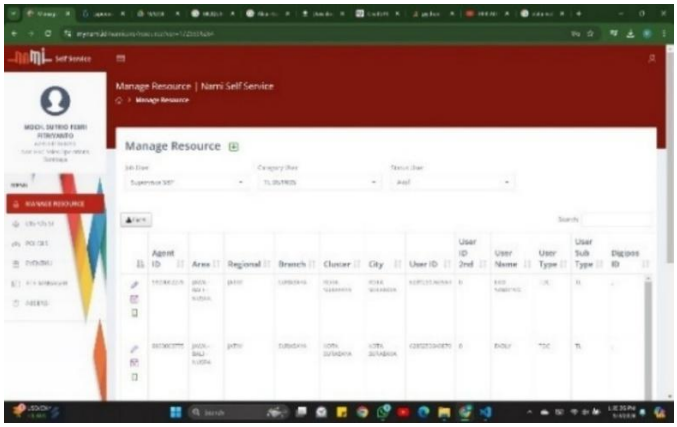
No	Kriteria/Indikator Penilaian
1	Kemampuan dalam memahami dan menjelaskan sebuah data yang ada dalam dunia riil dengan mengkombinasikan perasaan, penalaran dan pengalaman
2	Kemampuan dalam menggunakan data (data consumer) dan menyusun sebuah analisis sehingga tercipta sebuah intuisi dan wawasan data yang berguna
3	Kemampuan dalam mengambil keputusan dengan memadukan antara intuisi, wawasan, dan data
4	Kemampuan dalam menggunakan data untuk menyelesaikan permasalahan dari hasil observasi/lapangan sehingga dapat menghasilkan sebuah intuisi dan wawasan data yang aplikatif

4.4.1. Memahami Data [Indikator ke-1]

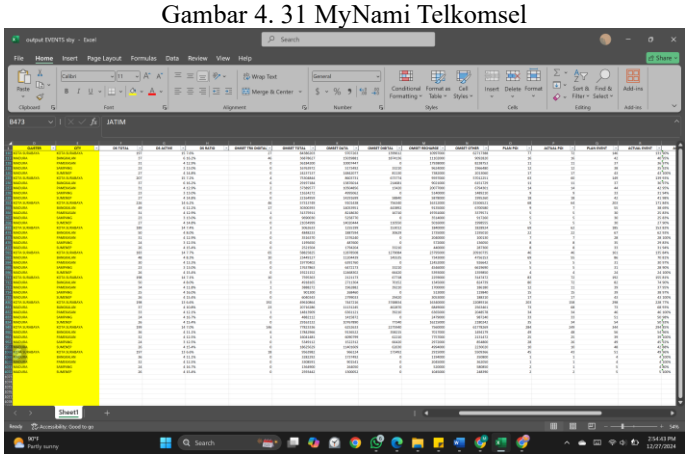
Secara umum, instuisi dan wawasan data merupakan dua aspek penting yang saling melengkapi dalam proses analisis data dan pengambilan keputusan. Intuisi dalam data adalah kemampuan untuk mengenali pola atau hubungan yang mungkin tidak langsung terlihat dalam data mentah. Ini lebih didasarkan pada perasaan atau pemahaman yang muncul tanpa analisis mendalam, sering kali dipengaruhi oleh pengalaman atau pengetahuan tentang domain yang dianalisis. Sementara, wawasan dalam data adalah pemahaman yang lebih mendalam dan terstruktur tentang data, yang diperoleh melalui penalaran logis dan analisis reflektif (HDL Anggrayani, 2021). Analisis data adalah proses sistematis yang melibatkan pengumpulan, pengolahan, dan interpretasi data untuk menarik kesimpulan yang bermanfaat. Proses ini penting dalam berbagai konteks, termasuk bisnis, penelitian, dan pengambilan keputusan sehari-hari. Tujuan utama dari analisis data adalah untuk mengekstraksi informasi yang berguna dari data agar dapat digunakan dalam pengambilan keputusan yang lebih baik (DQLab, 2021).

Untuk mengambil data, penulis menggunakan platform my.nami, Platform my.nami dirancang dengan sistem akses yang terbatas, di mana hanya satu pengguna yang dapat masuk ke dalam akun tersebut pada satu waktu. Akun pengguna yang terdaftar di platform ini tidak dapat diakses secara bersamaan melalui perangkat lain, sehingga memastikan keamanan dan eksklusivitas akses data. Setiap kali pengguna ingin mengambil data, platform ini akan menyediakan data dalam format Excel yang dapat diunduh, memudahkan proses analisis lebih lanjut di luar sistem.

Hal ini memungkinkan data yang diperoleh tetap terorganisir dan sesuai dengan kebutuhan.



Gambar 4. 31 MyNami Telkomsel



Gambar 4. 32 Data Events Daerah Surabaya - Madura

Salah satu data yang diperoleh penulis yaitu tentang data events. Data events ini adalah sampel data yang menunjukkan area Surabaya – Madura dengan jumlah pelanggan dan omset, dalam hal ini penulis akan menganalisis jumlah pengguna kartu Telkomsel melalui events khususnya di area Surabaya - madura, data events menjadi salah satu Indikator penting yang dapat memberikan gambaran mengenai intensitas dan keberhasilan kegiatan yang diadakan di berbagai area. Dengan menganalisis data events, diperlukan pemahaman yang lebih mendalam mengenai kegiatan events dalam menarik dan mempertahankan jumlah pengguna di berbagai area. Beberapa area mungkin belum mendapatkan perhatian yang cukup melalui strategi event yang tepat, sehingga potensi peningkatan jumlah pengguna di wilayah - wilayah tersebut belum optimal. Tujuannya yaitu untuk mengidentifikasi area dalam jumlah pengguna dengan potensi tinggi yang dapat diperoleh melalui strategi pengadaan events yang lebih terarah. Dengan memperbaiki dan menyesuaikan kegiatan events atau program di area tersebut dan juga meningkatkan jumlah pengguna baru serta mempertahankan jumlah pengguna yang sudah ada.

Tentunya dengan melibatkan kombinasi antara perasaan, penalaran, dan pengalaman, yang saling mendukung dalam memberikan pemahaman yang lebih baik dengan cara :

- Perasaan
  Melalui pengamatan awal, seperti data events memberikan kesan bahwa intensitas kegiatan di suatu area memiliki hubungan erat dengan peningkatan

jumlah pengguna kartu Telkomsel, menciptakan peluang untuk memperkuat strategi pemasaran.

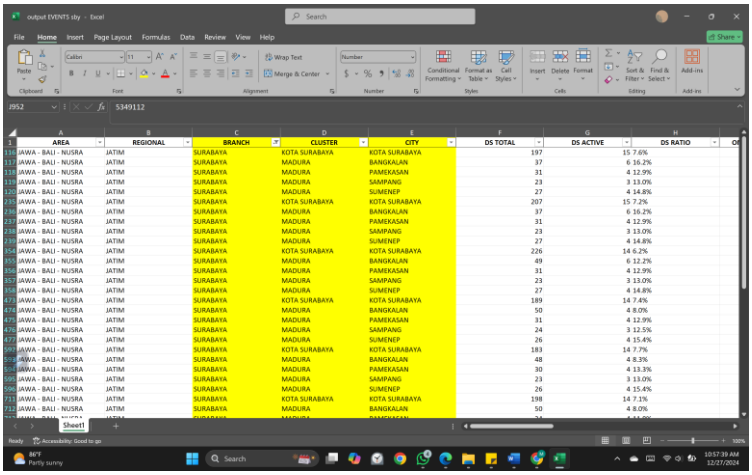
- Penalaran
  Seperti menganalisis pola dalam data events, hubungan antara intensitas kegiatan dan jumlah pengguna dapat dijelaskan secara logis. Penemuan ini mendukung kesimpulan bahwa area dengan tingkat kegiatan tinggi cenderung menunjukkan pertumbuhan pengguna yang lebih signifikan, memberikan dasar untuk merancang strategi events yang lebih efektif.
- Pengalaman
  Seperti menunjukkan bahwa area dengan pelaksanaan kegiatan yang terorganisasi dan relevan cenderung memiliki tingkat keberhasilan yang lebih tinggi dalam menarik pengguna baru. Insight ini menegaskan pentingnya memahami konteks lokal dalam merancang dan melaksanakan kegiatan pemasaran.

4.4.2. Menganalisis Data Events by Telkomsel Menggunakan Looker Studio[Indikator ke-2]

Data yang digunakan dalam analisis ini diperoleh dari MyNami Telkomsel, sebuah aplikasi atau layanan digital internal yang mendukung aktivitas manajemen dan operasional Telkomsel. Dalam hal ini penulis menganalisis dengan menggunakan Looker Studio yang bertujuan untuk menghasilkan intuisi dan wawasan data yang relevan dengan kondisi di lapangan, serta memberikan rekomendasi strategis untuk meningkatkan jumlah pengguna aktif di kota tertentu melalui pengadaan events.

Ada langkah – langkah untuk menganalisis data events menggunakan Looker studio, yaitu :

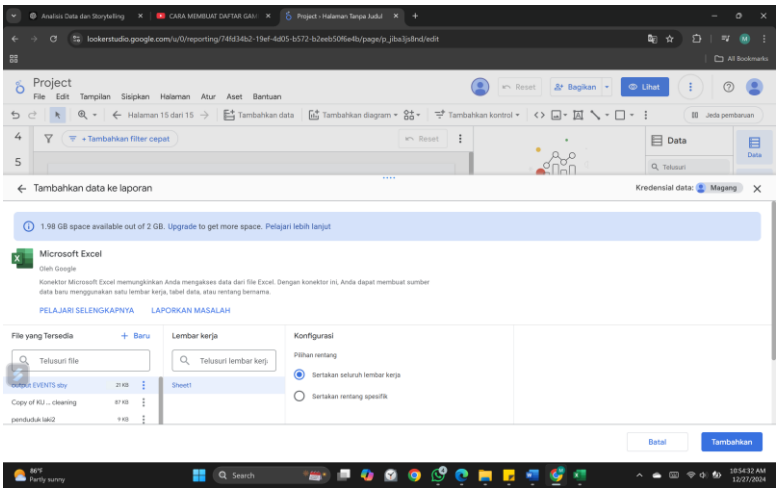
- Mengelola data events dengan mencleaning data mulai dari merubah type data, data events berupa file Excel yang mencakup periode Januari hingga Agustus 2024, dengan total 1.072 entri. Data mencakup wilayah Surabaya - Madura. Variabel-variabel yang digunakan dalam analisis meliputi AREA, REGIONAL, BRANCH, CLUSTER, CITY, dan DS ACTIVE (jumlah pengguna aktif) dan variabel lainnya dengan struktur data berikut :



AREA	REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	15	7.6%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	37	6	16.2%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4	12.9%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3	13.0%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4	14.8%
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	207	15	7.2%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	37	6	16.2%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4	12.9%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3	13.0%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4	14.8%
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	226	14	6.2%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	49	6	12.2%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4	12.9%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3	13.0%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4	14.8%
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	189	14	7.4%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4	8.0%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4	12.9%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	3	12.5%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4	15.4%
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	183	14	7.7%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	48	4	8.3%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	30	4	13.3%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3	13.0%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4	15.4%
JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	198	14	7.1%
JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4	8.0%

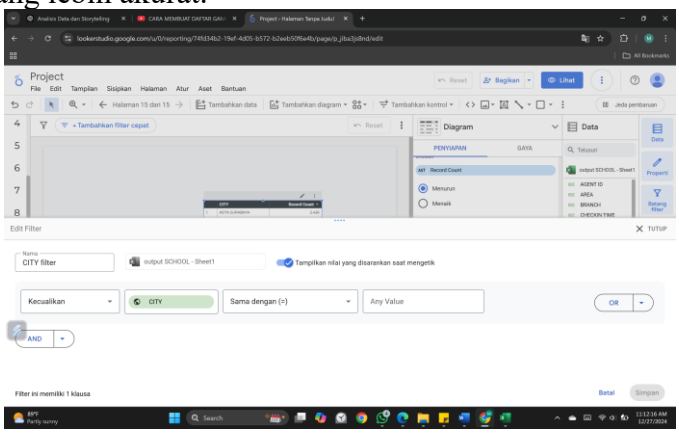
Gambar 4. 33 Data Events

- Setelah mendapatkan data events yang berformat excel penulis menghubungkan file excel ke Looker Studio



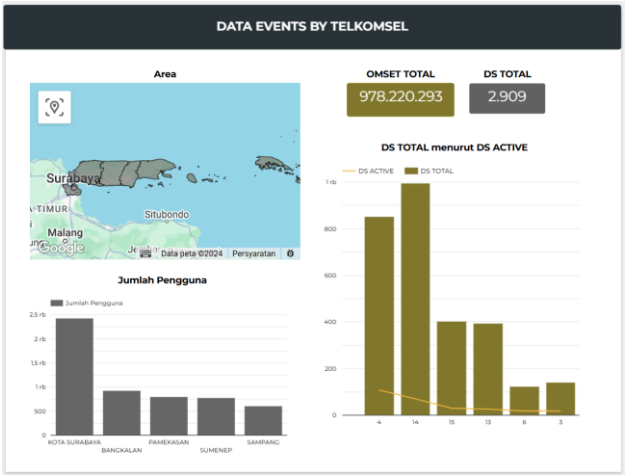
Gambar 4. 34 Upload Excel Ke Looker Studio

- Penulis melakukan proses pembersihan data yang akan dianalisis dengan mengidentifikasi dan menghapus bagian-bagian data yang bernilai null atau kosong. Proses ini dilakukan dengan cara mengatur filter secara langsung di Looker Studio untuk memastikan data yang digunakan dalam analisis dan visualisasi sudah bersih dan relevan, sehingga mendukung pengambilan keputusan yang lebih akurat.



Gambar 4. 35 Memfillter Data

- Penulis merancang dan membuat visualisasi data dan membuat dashboard yang relevan dan informatif, dengan mempertimbangkan tujuan analisis serta kebutuhan pengguna. Visualisasi ini dirancang untuk memudahkan pemahaman terhadap data, menyoroti area yang penting, dan mendukung penyampaian informasi secara jelas kepada para pemangku kepentingan.



Gambar 4. 36 Hasil Dashboard

Dapat disimpulkan bahwa

a. **Dominasi Kota Surabaya dalam Jumlah Pengguna**
   Grafik menunjukkan bahwa Kota Surabaya memiliki jumlah pengguna Telkomsel tertinggi dibandingkan dengan daerah lain, mencapai kurang dari 2.500 pengguna. Ini menunjukkan bahwa Surabaya adalah pasar utama yang perlu mendapat perhatian khusus dalam strategi pemasaran.
b. **Omset Total yang Signifikan**
   Dengan total omset mencapai **978.220.293**, kontribusi yang besar kemungkinan berasal dari wilayah dengan jumlah pengguna tinggi, terutama Surabaya. Hal ini menandakan pentingnya fokus pada daerah-daerah dengan potensi omset tinggi.
c. **Distribusi DS Total dan DS Active**
   Menunjukkan bahwa "DS Total" (jumlah agen penjualan) pada nilai tertentu, terutama pada angka 14, jauh lebih tinggi dibandingkan "DS Active". Ini mengindikasikan adanya potensi untuk meningkatkan aktivasi DS agar memaksimalkan kinerja.
d. **Konsistensi Jumlah Pengguna di Daerah Lain**
   Meskipun Surabaya mendominasi, jumlah pengguna di Bangkalan, Pamekasan, Sumenep, dan Sampang cukup merata pada kisaran 500-1000 pengguna. Ini dapat dijadikan dasar untuk strategi perluasan cakupan yang merata.
e. **Prioritas Aktivasi DS untuk Pertumbuhan**
   Perbandingan antara "DS Active" dan "DS Total" mengindikasikan perlunya optimalisasi aktivasi agar seluruh DS yang tersedia dapat lebih produktif dalam mendukung penjualan.

4.4.3. Mengambil keputusan & Menyelesaikan Permasalahan dari Data Events by Telkomsel [Indikator ke-3, Indikator ke-4]
Setelah mengananlisis dan menyimpulkan hasil dashboard data events by Telkomsel berikut adalah gabungan dari intuisi, wawasan, dan data untuk mendukung pengambilan keputusan:

- Pengoptimalan Daerah dengan Potensi Tinggi
  Berdasarkan data, Kota Surabaya memiliki jumlah pengguna tertinggi, sehingga prioritas strategi pemasaran dan peningkatan layanan perlu difokuskan pada wilayah ini untuk mengamankan kontribusi omset yang

signifikan. Observasi lapangan dapat dilakukan untuk memahami kebutuhan spesifik pengguna di Surabaya, sehingga strategi yang diterapkan lebih relevan dan efektif.

- Peningkatan Aktivasi DS untuk Efisiensi Operasional
  Perbedaan antara "DS Total" dan "DS Active" menunjukkan potensi yang belum dimanfaatkan. Dengan menganalisis kendala di lapangan yang menyebabkan rendahnya tingkat aktivasi DS, solusi seperti pelatihan tambahan, insentif kinerja, atau perbaikan proses dapat dirancang untuk meningkatkan produktivitas DS.
- Memanfaatkan Pasar Menengah untuk Ekspansi Stabil
  Jumlah pengguna yang cukup merata di Bangkalan, Pamekasan, Sumenep, dan Sampang menunjukkan peluang untuk mengembangkan pasar menengah. Observasi lokal diperlukan untuk memahami preferensi pengguna di daerah ini. Strategi ekspansi dengan pendekatan bertahap dapat meningkatkan penetrasi pasar dan memperluas cakupan omset.
- Mengintegrasikan Wawasan Lapangan untuk Validasi Data
  Data dashboard menunjukkan potensi yang besar, tetapi observasi lapangan harus dilakukan untuk mengidentifikasi hambatan spesifik, seperti kendala infrastruktur atau tingkat adopsi teknologi yang rendah. Wawasan dari lapangan dapat diintegrasikan untuk membuat strategi yang aplikatif dan berdampak langsung pada perbaikan kinerja.

Maka, keputusan yang dapat diambil adalah fokus pada penguatan layanan dan pemasaran di Surabaya untuk memaksimalkan omset. Menyusun program untuk meningkatkan aktivasi DS berdasarkan analisis kendala di lapangan

4.5. Rekognisi Mata Kuliah Kecerdasan Bisnis

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Kecerdasan Bisnis. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Kecerdasan Bisnis

Tabel 4. 6 Indikator Penilaian Konversi Mata Kuliah Kecerdasan Bisnis

No	Kriteria/Indikator Penilaian
1	Kemampuan dalam menjelaskan konsep dasar kecerdasan bisnis, konsep Extract-Transform-Load (ETL), skema data warehouse (star, snowflake), model multidimensional, data mining, reporting builder dan visualisasi data
2	Kemampuan memilih metode dan pendekatan business intelligence yang tepat yang dapat digunakan dalam menyelesaikan pengolahan data dan informasi guna pengambilan keputusan pada bisnis
3	Kemampuan menggunakan bahasa pemrograman dan tools pendukung business intelligence dalam menyelesaikan studi kasus guna pengambilan keputusan pada bisnis secara efektif dan efisien
4	Kemampuan merancang dan memformulasikan, dan mendesain business intelligence workflow untuk menyelesaikan permasalahan pada suatu studi kasus tertentu

4.5.1. Materi Kecerdasan Bisnis [Indikator ke-1]

Konsep Keilmuan Kecerdasan Bisnis

Kecerdasan Bisnis (BI) merupakan suatu disiplin ilmu yang berfokus pada pengumpulan, analisis, dan interpretasi data untuk mendukung pengambilan keputusan yang lebih baik dalam organisasi. Dalam era informasi yang semakin berkembang, BI telah menjadi kunci bagi perusahaan untuk memahami kondisi pasar, perilaku konsumen, dan tren industri. Dengan menggunakan alat dan teknik analisis yang canggih, BI memungkinkan organisasi untuk mengubah data mentah menjadi wawasan yang berharga, sehingga mereka dapat merespons perubahan dengan cepat dan efisien (Zein, 2023).

Konsep Extract- Transform-Load (ETL)

Extract-Transform-Load (ETL) adalah proses penting dalam pengolahan data yang digunakan untuk mengumpulkan, mengubah, dan memuat data dari berbagai sumber ke dalam sistem penyimpanan data yang terpusat, seperti data warehouse. Proses ETL berfungsi sebagai jembatan antara sumber data yang beragam dan sistem analisis yang membutuhkan data terintegrasi dan berkualitas tinggi. Dalam era digital saat ini, di mana volume data yang dihasilkan sangat besar dan bervariasi, ETL menjadi semakin penting bagi organisasi untuk mengelola dan memanfaatkan data mereka secara efektif (Putu dkk., 2021)

- Ekstraksi (Extract)
  Tahap pertama dalam proses ETL adalah ekstraksi, di mana data diambil dari berbagai sumber yang dapat mencakup database relasional, sistem ERP (Enterprise Resource Planning), file flat, API, dan sumber data lainnya. Proses ekstraksi harus dilakukan dengan hati-hati untuk memastikan bahwa semua data yang relevan diambil tanpa kehilangan informasi penting. Data yang diekstrak biasanya disimpan dalam area staging sebelum diproses lebih lanjut.

Pada tahap ini, penting untuk mempertimbangkan faktor-faktor seperti kecepatan ekstraksi dan dampaknya terhadap kinerja sistem sumber

- Transformasi (Transform)

  Setelah data diekstrak, langkah berikutnya adalah transformasi. Pada tahap ini, data akan dibersihkan dan diubah agar sesuai dengan kebutuhan analisis. Transformasi mencakup beberapa sub-proses yang krusial, seperti:

  - Pembersihan Data: Menghapus duplikat, memperbaiki kesalahan, dan menangani nilai hilang untuk meningkatkan kualitas data.
  - Standarisasi Data: Mengubah format data agar konsisten, misalnya mengubah format tanggal atau menyamakan satuan ukuran.
  - Penggabungan Data: Mengintegrasikan informasi dari berbagai sumber untuk menciptakan pandangan yang komprehensif tentang suatu entitas atau proses.
  - Penghitungan Nilai Tambahan: Menambahkan kolom baru berdasarkan perhitungan dari kolom yang ada, seperti menghitung total penjualan dari jumlah unit yang terjual dikalikan dengan harga per unit.

  Transformasi memastikan bahwa data siap untuk analisis dan memenuhi standar kualitas yang diperlukan. Proses ini juga memberikan kesempatan untuk mengoptimalkan struktur data agar lebih efisien dalam penyimpanan dan pemrosesan.

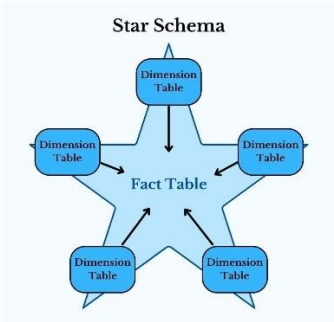
- Pemuatan (Load)
  Tahap terakhir dalam proses ETL adalah pemuatan, di mana data yang telah ditransformasi dimuat ke dalam sistem penyimpanan target, seperti data warehouse atau database analitik. Proses pemuatan harus dilakukan dengan cermat untuk memastikan integritas dan performa sistem penyimpanan. Pemuatan dapat dilakukan secara batch (pemuatan sejumlah besar data sekaligus pada waktu tertentu) atau real-time (pemuatan data secara terus-menerus saat data baru tersedia). Pemilihan metode pemuatan tergantung pada kebutuhan bisnis dan volume data yang dikelola.

Skema Data Warehouse

Data warehouse adalah sistem penyimpanan data yang dirancang untuk mengintegrasikan dan menyimpan data dari berbagai sumber dalam sebuah organisasi, sehingga memungkinkan analisis dan pengambilan keputusan yang lebih efektif. Salah satu aspek penting dari data warehouse adalah skema yang digunakan untuk mengorganisir data (Elektronik dkk., t.t.). Terdapat beberapa jenis skema dalam data warehouse, yang masing-masing memiliki karakteristik dan kegunaan tertentu. Berikut adalah penjelasan mengenai skema-skemanya:

Jenis-Jenis Skema pada Data Warehouse

1. Star Schema



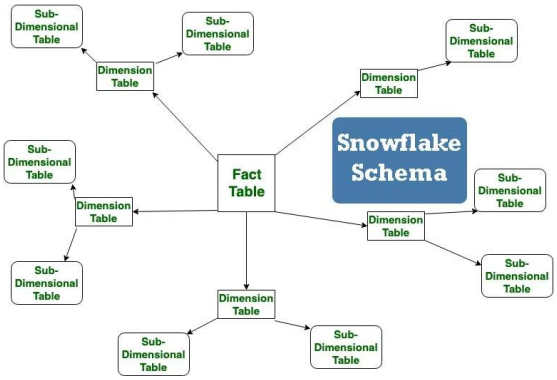
Gambar 4. 37 Star Schema

Star Schema adalah struktur dasar dalam pemodelan data warehouse yang terdiri dari satu tabel fakta di tengah dan beberapa tabel dimensi di sekitar tabel fakta tersebut. Tabel fakta berisi data kuantitatif yang relevan untuk analisis, sementara tabel dimensi memberikan konteks untuk data tersebut. Dengan Karakteristik:

- Tabel fakta terhubung langsung dengan tabel dimensi melalui kunci asing.
- Tabel dimensi tidak memiliki hubungan antar satu sama lain, sehingga membentuk pola seperti bintang.
- Memungkinkan kueri yang cepat dan efisien karena struktur yang sederhana.

Contoh penggunaan seperti dalam analisis penjualan, tabel fakta dapat mencakup total penjualan, sementara tabel dimensi dapat mencakup informasi tentang produk, waktu, dan lokasi.

2. Snowflake Schema



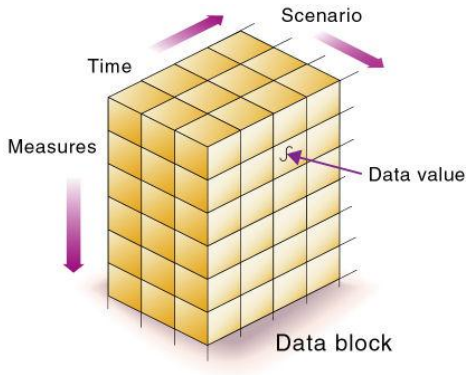
Gambar 4. 38 Snowflake Schema

Snowflake Schema adalah variasi dari Star Schema di mana tabel dimensi dipecah menjadi beberapa tabel lebih kecil untuk mengurangi redundansi. Dengan kata lain, tabel dimensi dapat memiliki subdimensi yang lebih terperinci. Dengan karakteristik:

- Memiliki struktur yang lebih kompleks dibandingkan Star Schema karena adanya normalisasi pada tabel dimensi.
- Mengurangi duplikasi data, tetapi dapat memperlambat kinerja kueri karena lebih banyak join yang diperlukan.

Contoh penggunaan seperti dalam analisis produk, informasi kategori produk dapat dipisahkan menjadi tabel tersendiri, sehingga memudahkan pengelolaan data.

Model Multidimensional



Gambar 4. 39 Model Multidimensional

Model multidimensional adalah pendekatan yang digunakan dalam analisis data untuk memahami dan mengelola data dalam beberapa dimensi. Pendekatan ini sangat penting dalam konteks data warehouse dan sistem analisis bisnis, di mana data kompleks perlu dianalisis dari berbagai sudut pandang untuk mendapatkan wawasan yang lebih dalam. Mengenai model multidimensional, termasuk konsep, komponen, dan manfaatnya. Model multidimensional merujuk pada struktur data yang memungkinkan pengguna untuk menganalisis informasi dari berbagai dimensi. Dalam model ini, data disusun sedemikian rupa sehingga pengguna dapat dengan mudah mengakses dan menganalisis data berdasarkan berbagai atribut sekaligus. Model ini sering digunakan dalam sistem OLAP (Online Analytical Processing) untuk mendukung analisis bisnis yang kompleks (Kuswandari dkk., 2021)

Data Mining

Data mining adalah proses penting dalam pengolahan data yang bertujuan untuk mengekstrak informasi dan pola yang berguna dari kumpulan data yang besar. Proses ini melibatkan penggunaan teknik statistik, algoritma, dan teknologi kecerdasan buatan untuk menemukan pola, tren, dan hubungan yang mungkin tidak terlihat secara langsung. Data mining sering kali digunakan dalam berbagai aplikasi, mulai dari analisis bisnis hingga penelitian ilmiah, untuk membantu organisasi membuat keputusan yang lebih baik berdasarkan data (Dibimbing.id, 2023)

Reporting Builder

Microsoft Report Builder adalah alat yang digunakan untuk membuat laporan paginasi dalam lingkungan SQL Server Reporting Services (SSRS) dan Power BI Report Server. Alat ini dirancang untuk pengguna bisnis yang lebih suka bekerja secara mandiri, tanpa harus menggunakan Perancang Laporan di Visual Studio atau SQL Server Data Tools (SSDT). Berikut adalah penjelasan lebih lanjut mengenai Microsoft Report Builder, fungsinya, dan fitur-fitur yang ditawarkannya. Microsoft Report Builder adalah aplikasi yang memungkinkan pengguna untuk merancang, mengedit, dan menerbitkan laporan dari berbagai sumber data. Laporan yang dihasilkan dapat berupa

laporan tabular, matriks, grafik, atau bentuk bebas lainnya. Dengan Report Builder, pengguna dapat membuat definisi laporan yang mencakup data yang akan diambil, sumber data, dan tata letak laporan (Microsoft Learn, 2024)

Fitur Utama Microsoft Report Builder

- Desain Laporan: Pengguna dapat membuat laporan dengan menggunakan antarmuka grafis yang intuitif. Laporan dapat disusun dengan elemen-elemen seperti tabel, grafik, dan kontrol lainnya untuk menyajikan data dengan cara yang menarik.
- Sumber Data Beragam: Report Builder mendukung berbagai jenis sumber data, termasuk SQL Server, Oracle, OLE DB, ODBC, dan XML. Pengguna dapat mengakses data dari berbagai sistem dan menggabungkannya dalam satu laporan.
- Definisi Laporan: Setiap laporan yang dibuat memiliki file definisi laporan (.rdl) yang menyimpan informasi tentang kueri yang digunakan untuk mengambil data, parameter laporan, dan tata letak visual. Definisi ini ditulis dalam format XML yang sesuai dengan Report Definition Language (RDL).
- Penerbitan Laporan: Setelah laporan selesai dirancang, pengguna dapat menerbitkannya ke server laporan SSRS atau Power BI Report Server. Ini memungkinkan laporan diakses oleh pengguna lain dalam organisasi.
- Pratinjau Laporan: Pengguna dapat mempratinjau laporan sebelum menerbitkannya untuk memastikan bahwa semua elemen tampil dengan benar dan data ditampilkan sesuai harapan.
- Kemampuan Mengubah Laporan: Report Builder memungkinkan pengguna untuk menyesuaikan dan memperbarui laporan yang telah ada sebelumnya. Ini memberi fleksibilitas dalam mengelola laporan seiring dengan perubahan kebutuhan bisnis.

Manfaat Menggunakan Microsoft Report Builder

- Meningkatkan Aksesibilitas Data: Dengan alat ini, pengguna non-teknis dapat membuat laporan tanpa perlu memahami SQL atau bahasa pemrograman lainnya.
- Pengambilan Keputusan Berbasis Data: Laporan yang dihasilkan membantu manajemen dalam membuat keputusan yang lebih baik berdasarkan analisis data.
- Efisiensi Waktu: Proses pembuatan laporan menjadi lebih cepat dan efisien berkat antarmuka drag-and-drop dan alat desain intuitif.

Visualisasi Data

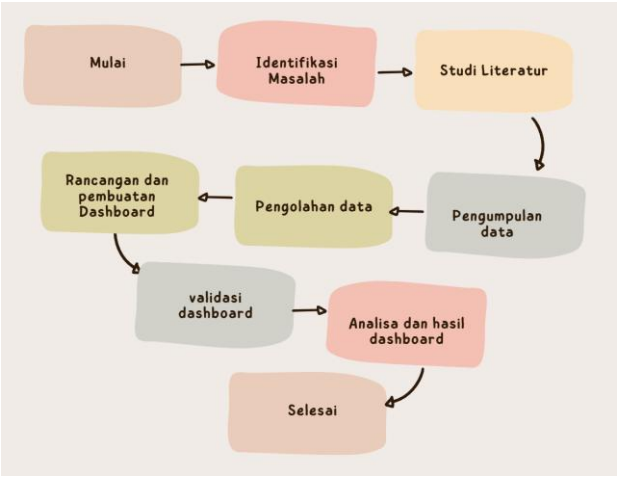
Gambar 4. 40 Dashboard Visualisasi Data

Visualisasi data adalah teknik yang digunakan untuk menyajikan data kuantitatif dalam bentuk yang mudah dipahami melalui elemen visual seperti grafik, diagram, dan peta. Teknik ini bertujuan untuk mengkomunikasikan informasi secara jelas dan efisien, sehingga audiens dapat dengan cepat memahami tren, pola, dan hubungan dalam data. Dengan menggunakan visualisasi, informasi yang kompleks dapat disederhanakan menjadi representasi visual yang lebih intuitif, memungkinkan pengguna untuk melihat dan menganalisis data dengan lebih efektif. Salah satu tujuan utama dari visualisasi data adalah meningkatkan pemahaman audiens terhadap data yang disajikan, membantu mereka memahami data kompleks dengan cara yang lebih intuitif. Selain itu, visualisasi juga memudahkan dalam mengidentifikasi pola, tren, dan outlier dalam dataset, yang pada gilirannya dapat memberikan wawasan berharga untuk pengambilan keputusan. Dalam konteks bisnis, visualisasi data sangat penting karena memberikan informasi yang diperlukan untuk membuat keputusan yang lebih baik berdasarkan analisis data. Dengan demikian, visualisasi data tidak hanya berfungsi sebagai alat komunikasi, tetapi juga sebagai sarana untuk mendemokratisasi akses terhadap informasi, memungkinkan individu dari berbagai latar belakang untuk memahami dan menggunakan data secara efektif dalam pengambilan keputusan (Revou, 2024)

4.5.2. Metode Pengembangan Dashboard Kecerdasan Bisnis [Indikator ke-2, Indikator ke-3, dan Indikator ke-4]

Dalam hal ini penulis mengembangkan dashboard menggunakan Looker dan Python untuk mendukung kecerdasan bisnis selama magang di Telkomsel Branch Surabaya. Di Telkomsel Branch Surabaya menghadapi tantangan dalam memonitor data penjualan yang berasal dari berbagai dataset dengan jumlah yang besar. Terutama untuk data penjualan, perusahaan kesulitan dalam mengelola dan menganalisis data yang tersebar, sehingga mempengaruhi efisiensi dalam pengambilan keputusan yang berbasis data. Proyek ini bertujuan membantu pengambilan keputusan strategis dengan memilih metode business intelligence yang tepat, seperti integrasi data dan visualisasi yang efektif. Dashboard yang dihasilkan digunakan untuk memantau tren bisnis, menganalisis perilaku pelanggan, dan memberikan wawasan guna meningkatkan kinerja operasional. Dalam proyek ini, penulis juga menerapkan konsep Extract-Transform-Load (ETL) dengan Python untuk mengolah data mentah dari berbagai sumber internal, memastikan data yang diolah sesuai untuk analisis dan kemudian dimuat ke dalam Looker untuk visualisasi yang dinamis dan informatif. Berikut merupakan

workflow untuk mata kuliah kecerdasan bisnis yang disesuaikan dengan kasus pembuatan dashboard untuk monitoring data penjualan di Telkomsel Branch Surabaya:



Gambar 4. 41 Workflow Kecerdasan Bisnis

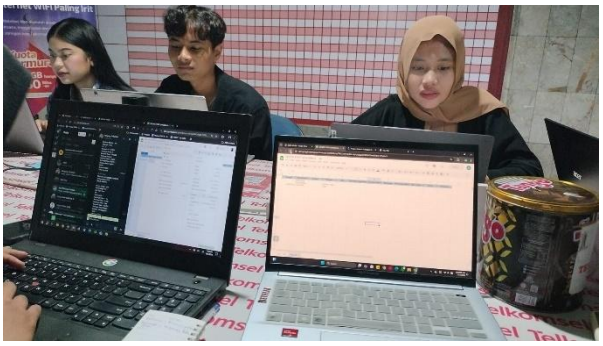
1. Mulai
   Proses dimulai dengan perencanaan untuk memecahkan masalah pemantauan data penjualan yang besar dan tersebar di berbagai dataset.
2. Identifikasi Masalah
   Mengidentifikasi masalah utama, yaitu kebutuhan sistem monitoring yang lebih efisien untuk menganalisis data penjualan cabang Surabaya. Tujuan dashboard adalah untuk menyajikan data penjualan secara visual dan real-time.
3. Studi Literatur
   Mengumpulkan informasi terkait metode terbaik untuk visualisasi data penjualan dan dashboard interaktif, serta alat yang dapat digunakan untuk pengembangan, seperti Tableau atau Power BI.
4. Pengumpulan Data
   Data penjualan yang terkait dengan cabang Surabaya dikumpulkan, menggabungkan dataset bulanan yang ada dari Januari hingga September.
5. Pengolahan Data
   Data yang telah dikumpulkan diproses dengan pembersihan data, penggabungan dataset, penyaringan untuk cabang Surabaya, dan pengecekan konsistensi format data. Python digunakan untuk efisiensi dalam memanipulasi data.
6. Rancangan Dashboard
   Desain awal dashboard dibuat, termasuk tata letak, elemen visualisasi yang tepat (seperti grafik penjualan, tren, dan metrik relevan), serta fitur interaktif untuk mempermudah analisis data.
7. Pembuatan Dashboard
   Dashboard dikembangkan berdasarkan desain yang sudah dibuat, menggunakan alat pengembangan seperti Tableau atau Looker Studio untuk memastikan tampilan dan fungsionalitas yang diinginkan.
8. Validasi Dashboard
   Dashboard diuji untuk memastikan data yang ditampilkan akurat, tata letaknya efektif, dan memenuhi kebutuhan analisis pengguna. Jika tidak lolos validasi, dilakukan revisi pada elemen visual atau fungsi dashboard.
9. Analisa Hasil Dashboard

Setelah dashboard selesai dan tervalidasi, hasil analisis data yang ditampilkan digunakan untuk menarik kesimpulan, seperti identifikasi tren penjualan dan pengambilan keputusan terkait strategi pemasaran.

10. Selesai
    Proses berakhir setelah dashboard selesai dibuat dan dapat digunakan untuk pemantauan dan analisis penjualan di Telkomsel Branch Surabaya.

Berikut langkah – langkah pengembangan dashboard kecerdasan bisnis dengan looker dan python untuk mendukung pengambilan keputusan strategis di Telkomsel Branch Surabaya :

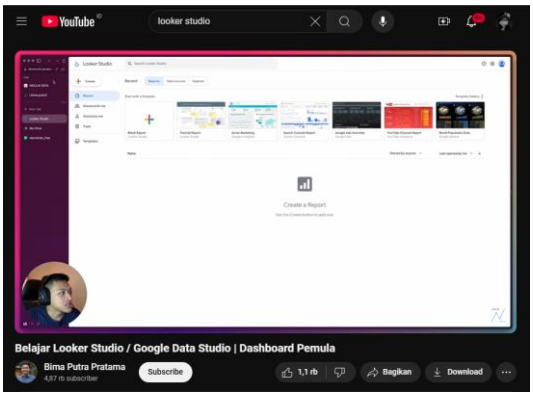
1. Identifikasi Masalah



Gambar 4. 42 Diskusi

Identifikasi masalah merupakan tahap awal yang penting untuk memahami isu yang akan diteliti. Langkah ini bertujuan untuk mengetahui jenis masalah yang ada serta faktor-faktor penyebabnya, sehingga fokus penelitian dapat diarahkan pada satu topik yang spesifik. Proses identifikasi dilakukan dengan cara mempelajari permasalahan yang terjadi di perusahaan secara mendalam. Berdasarkan hasil diskusi awal dengan mentor, terungkap bahwa PT. Telkomsel Branch Surabaya menghadapi tantangan dalam memonitor data penjualan yang berasal dari berbagai dataset dengan jumlah yang besar terutama Data Sales. Kebutuhan akan sistem monitoring yang lebih efisien ini diidentifikasi sebagai salah satu prioritas utama, terutama untuk mendukung analisis data penjualan secara lebih efektif. Dalam pertemuan pertama yang membahas tugas data analyst, mentor menyampaikan bahwa solusi berupa dashboard yang interaktif dan dapat diakses secara real-time diperlukan untuk mempermudah akses dan pemahaman terhadap data tersebut. Oleh karena itu, sesuai dengan arahan mentor penulis akan mengembangkan dashboard monitoring yang dirancang khusus untuk mengelola dan menampilkan data penjualan secara visual, sehingga dapat mendukung pengambilan keputusan yang lebih cepat dan berbasis data.

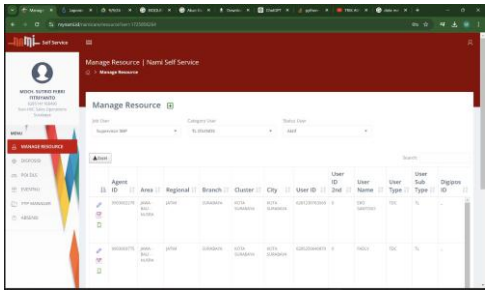
2. Studi Literatur



Gambar 4. 43 Studi Literatur

Penulis mulai mempelajari tentang looker data studio yang akan digunakan sebagai tools untuk membuat dashboard. Sumber pembelajaran berasal dari berbagai jenis, namun yang paling utama yaitu platform youtube. Melalui YouTube, penulis dapat mengakses berbagai tutorial dan demonstrasi yang memberikan panduan langkah-demi-langkah dalam menggunakan Looker Data Studio. Video tutorial ini memungkinkan penulis untuk melihat secara langsung cara-cara praktis dalam memanfaatkan fitur-fitur yang ada, seperti mengimpor data, melakukan pembersihan data, dan membuat visualisasi yang sesuai dengan kebutuhan.

3. Pengumpulan Data



Gambar 4. 44 MyNami Telkomsel

Dalam platform database Mynami, penulis mengakses data events yang tersimpan di dalam folder Manage Resource. Selain data events, platform ini juga menyediakan berbagai jenis data lainnya, seperti data sales, absensi, FTP manager, dan data operasional lainnya. Ketersediaan data yang beragam ini memungkinkan pengguna untuk mengelola informasi secara menyeluruh sesuai kebutuhan perusahaan. Penulis memanfaatkan data events secara spesifik karena relevan dengan tujuan penelitian, yaitu membuat dashboard interaktif untuk mendukung kebutuhan monitoring dan strategi bisnis perusahaan.

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO (	OMSET TR	OMSET TO	OMSET DA	OMSET DI	OMSET RE	OMSET OT	PLAN POI	ACTUAL P	PLAN EVE	ACTUAL E	ACTUAL C	SALES	OTA
BALI NUSR	DENPASAF	BALI BARA	BULELENG	42	3	7.1	39	11043601	6456732	626894	3004000	955975	49	37	72	50	69	0	
BALI NUSR	DENPASAF	BALI BARA	JEMBRANA	23	3	13.0	39	14135222	8340825	654798	2613000	2560198	48	43	75	56	75	0	
BALI NUSR	DENPASAF	BALI BARA	TABANAN	20	3	15.0	37	8446941	5681230	420320	1975000	370391	50	46	53	48	91	5	
BALI NUSR	DENPASAF	BALI TENG	BADUNG	42	6	14.3	92	24862947	19937451	1020468	3220000	685028	116	77	156	100	64	59	
BALI NUSR	DENPASAF	BALI TENG	KOTA DEN	107	6	5.6	84	32561574	14073039	966520	4313000	13347015	32	29	152	105	69	6	
BALI NUSR	DENPASAF	BALI TIMU	BANGLI	25	3	17.6	40	14542815	9085094	520052	3493000	1505668	32	30	48	40	83	0	
BALI NUSR	DENPASAF	BALI TIMU	GIANYAR	35	6	17.1	114	15545288	8141885	717153	2424000	4377250	21	18	105	80	76	0	
BALI NUSR	DENPASAF	BALI TIMU	KARANG A	28	3	10.7	66	13887805	6664495	892319	3396000	2934991	59	48	81	64	79	0	
BALI NUSR	DENPASAF	BALI TIMU	KLUNGKUI	19	3	15.8	36	6237676	4464847	454743	390000	943085	54	45	74	60	81	2	
BALI NUSR	FLORES	ENDE SIKK	ENDE	15	3	20.0	54	93223248	60437891	2172474	25019000	5897374	51	41	111	92	83	28	
BALI NUSR	FLORES	ENDE SIKK	SIKKA	23	4	17.4	90	2,81E+08	1,55E+08	3173675	96460000	26913263	53	37	82	65	79	46	
BALI NUSR	FLORES	FLORES TI	ALOR	9	3	33.3	49	49888637	20922081	1451451	24620000	2895105	30	29	156	127	81	112	
BALI NUSR	FLORES	FLORES TI	FLORES TI	13	3	23.1	30	35801991	19869056	1045255	10680000	4322026	36	35	131	112	85	8	
BALI NUSR	FLORES	FLORES TI	LEMBATA	6	3	50.0	37	47998003	26289336	1607029	17051000	3092136	9	8	42	38	90	206	
BALI NUSR	FLORES	MANGGAF	MANGGAR	15	4	26.7	39	84069798	30566154	1333375	46193000	5977269	32	30	141	103	73	177	
BALI NUSR	FLORES	MANGGAF	MANGGAR	11	2	18.2	21	27707826	15442746	842532	9279000	2143548	19	19	98	66	67	129	
BALI NUSR	FLORES	MANGGAF	MANGGAR	8	2	25.0	4	9070265	3969946	173496	4244000	682823	11	9	53	35	66	0	
BALI NUSR	FLORES	MANGGAF	NAGEKEO	9	2	22.2	15	16420982	8650405	630984	5346000	1857092	18	14	85	60	71	1	
BALI NUSR	FLORES	MANGGAF	NGADA	9	3	33.3	24	38851772	18868155	769987	15233000	3996129	28	22	93	53	57	61	
BALI NUSR	KUPANG	KUPANG R	KOTA KUP,	41	10	24.4	324	31143018	12855951	1667561	11645000	5161064	49	49	120	116	97	0	
BALI NUSR	KUPANG	KUPANG R	KUPANG	16	2	12.5	17	7664776	3448326	162145	2808000	1246305	15	15	65	49	75	75	
BALI NUSR	KUPANG	KUPANG R	ROTE NDA	8	2	25.0	61	7533238	2163268	326067	4458000	619402	20	18	62	42	68	60	
BALI NUSR	KUPANG	MALAKA T	BELU	12	3	25.0	103	19468159	14183268	705331	3464000	1230590	22	22	118	66	56	13	
BALI NUSR	KUPANG	MALAKA T	MALAKA	6	3	50.0	82	25857025	14130056	507234	8382000	2837735	15	12	81	52	64	0	

Gambar 4. 45 Data Events

Data events dalam platform Mynami terbagi berdasarkan bulan, dengan fokus proyek ini diarahkan pada data tahun 2024, mulai dari bulan Januari hingga September. Setiap file dataset memiliki struktur yang cukup kompleks dan ukuran yang besar, dengan jumlah baris yang melebihi 1000 dan terdiri dari 26 kolom. Kompleksitas ini mencerminkan volume data yang tinggi, sehingga memerlukan pengolahan data yang cermat untuk memastikan data siap digunakan dalam pembuatan dashboard interaktif yang dapat mendukung analisis dan pengambilan keputusan secara efektif.

4. Pengolahan Data

Pengolahan data ini dilakukan dengan menggunakan dua alat utama, yaitu Microsoft Excel dan pemrograman Python. Excel digunakan untuk manipulasi data yang bersifat manual atau memerlukan validasi cepat, sementara Python digunakan untuk otomatisasi proses pengolahan, analisis mendalam, dan penerapan teknik-teknik pemrosesan data yang lebih kompleks. Proses ini memastikan bahwa dataset yang dihasilkan bersih, terstruktur, dan siap digunakan untuk analisis lebih lanjut, yakni pembuatan dashboard menggunakan looker data studio

a. Data Merge

Langkah awal pengolahan data meliputi pembersihan data (data cleaning) untuk memastikan relevansi dan kualitas dataset. Fokusnya pada data branch Surabaya, dengan menggabungkan dataset bulanan Januari-September 2024. Proses ini mencakup penghapusan duplikat, pengecekan format, dan validasi data menggunakan Python dengan pustaka Pandas dan NumPy, guna mempersiapkan dataset untuk analisis dan visualisasi. Proses ini memakan waktu yang cukup lama sekitar kurang lebih 10 menit, yang membutuhkan waktu ekstra untuk memastikan setiap langkah pembersihan berjalan dengan baik dan akurat. Berikut adalah kode script proses tersebut.

```
import pandas as pd
import os

# Direktori di mana file Excel berada
folder_path = r'C:/Users/user/Downloads/Events

# Daftar file Excel dengan ekstensi .xlsx
excel_files = [
                                os.path.join(folder_path,
'Reporting_Events_202401.xlsx'),
```

```
        os.path.join(folder_path,    'Reporting_    Events
_202402.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202403.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202404.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202405.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202406.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202407.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202408.xlsx'),
        os.path.join(folder_path,    'Reporting_    Events
_202409.xlsx'),

    # Tambahkan file Excel lainnya jika ada
]

# List untuk menyimpan data dari setiap file Excel
data_frames = []
```

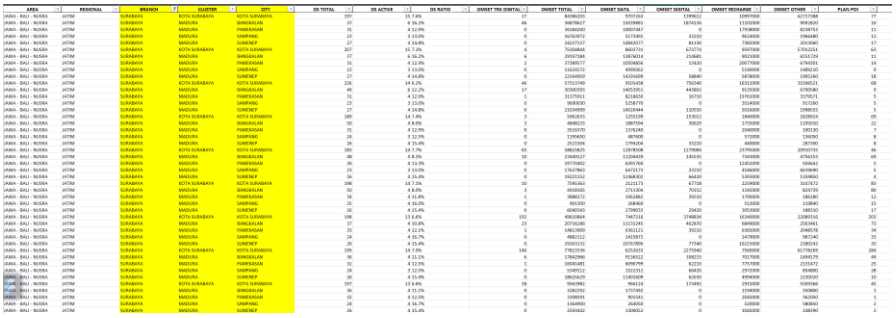
Kode tersebut bertujuan untuk menggabungkan data dari beberapa file Excel yang berada di folder tertentu. Langkah pertama adalah mengimpor dua library, yaitu ‘pandas’ dan ‘os’. Library ‘pandas’ digunakan untuk membaca dan mengolah data dari file Excel, sementara ‘os’ membantu dalam mengatur dan memanipulasi file serta direktori di komputer. Setelah itu, ditentukan lokasi folder tempat file Excel disimpan

```
# Loop untuk memeriksa apakah file ada dan membaca data
for file in excel_files:
    if os.path.exists(file):
        print(f"File ditemukan: {file}")
        # Membaca file Excel
        try:
            df = pd.read_excel(file)
            # Filter data dengan BRANCH = Surabaya
            if 'BRANCH' in df.columns:
               filtered_df = df[df['BRANCH'] == 'SURABAYA']
                data_frames.append(filtered_df)
            else:
                print(f"Kolom 'BRANCH' tidak ditemukan di
{file}")
        except Exception as e:
              print(f"Gagal membaca file: {file}. Error:
{e}")
    else:
        print(f"File tidak ditemukan: {file}")
```

Kode diatas memproses beberapa file Excel dengan memeriksa keberadaan file, membaca data, dan memfilter informasi berdasarkan cabang Surabaya. File yang ada diperiksa untuk kolom 'BRANCH', lalu data difilter untuk nilai 'SURABAYA'.

```
# Jika ada data yang berhasil dibaca, gabungkan semuanya
if data_frames:
            merged_data     =     pd.concat(data_frames,
ignore_index=True)

    # Simpan ke file Excel atau CSV
          output_excel     =     os.path.join(folder_path,
'merged_Events_data.xlsx')
           output_csv     =     os.path.join(folder_path,
'merged_Events_data.csv')

    try:
        # Simpan ke file Excel
        merged_data.to_excel(output_excel, index=False)
        print(f'Data berhasil disimpan ke {output_excel}')
    except Exception as e:
        print(f"Gagal menyimpan file Excel: {e}")

    try:
        # Simpan ke file CSV
        merged_data.to_csv(output_csv, index=False)
        print(f'Data berhasil disimpan ke {output_csv}')
    except Exception as e:
        print(f"Gagal menyimpan file CSV: {e}")

    # Tampilkan ringkasan hasil merge
    print("\nRingkasan Data yang Telah Digabungkan:")
    print(merged_data.info())  # Menampilkan informasi data
seperti jumlah kolom dan tipe data

     print("\nContoh Data yang Telah Digabungkan (5 baris
pertama):")
      print(merged_data.head())   #  Menampilkan  5  baris
pertama dari data yang digabungkan

else:
    print("Tidak ada file yang berhasil dibaca.")
```

Kode script di atas menghasilkan file Excel yang menggabungkan data events dari berbagai file Excel terpisah, di mana data tersebut telah difilter untuk hanya menyertakan baris dengan kolom 'BRANCH' yang bernilai 'SURABAYA'. File hasil akhir diberi nama "merged_ Events _data.xlsx". Setelah proses filter dan penggabungan selesai, dataset yang dihasilkan berisi total 45 baris data yang relevan dengan cabang Surabaya.



AREA	REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO	ONSET TRX DIGITAL	ONSET TOTAL	ONSET DATA	ONSET DIGITAL	ONSET RECHARGE	ONSET OTHER	PLAN POI
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	15 7.6%		27	8438635	5707263	1199012	1199000	8177188	77
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	37	6 16.2%		46	8487867	15019881	1874136	1131000	9093820	16
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4 12.9%		0	3604020	10007447	0	1797000	8230763	11
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3 13.0%		0	3676997	5173892	33250	962000	1964081	12
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4 14.8%		0	2423727	14842477	81130	1360000	2553600	17
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	207	15 7.2%		4	7559848	8602731	673776	934000	5701211	60
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	27	4 14.8%		4	2930258	10874014	214840	1007000	4154729	11
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	27	4 14.8%		0	3718577	10500853	1340	1007000	6180211	14
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3 13.0%		0	3168627	4966962	0	1360000	1980160	10
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4 14.8%		0	2154950	14250208	58840	1678000	1385260	18
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	226	14 6.2%		46	5733740	9104458	756340	1631000	3158521	68
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	49	6 12.2%		17	3030393	14053951	443882	3133000	6785780	9
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4 12.9%		1	3137011	8218630	16710	1976000	3370571	5
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3 13.0%		0	3008030	5258770	0	3514000	3127260	5
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4 14.8%		0	2104990	14024444	110550	5016000	1988555	5
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	189	14 7.4%		3	5062635	1255199	153012	1840000	1920914	69
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4 8.0%		3	4848235	1887594	30620	1710000	1195550	22
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4 12.9%		0	3510570	1379240	0	2660000	1931530	7
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	3 12.5%		0	1295650	487600	0	572000	136050	8
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	2517158	1794204	33250	440000	287500	8
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	180	14 7.7%		69	5865825	12878908	1279886	2379000	20930735	46
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	48	4 8.3%		10	2349527	11204439	145535	7403000	4766153	69
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	30	4 13.3%		0	1977002	6997760	0	1192000	926642	5
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3 13.0%		0	1761963	6473173	33250	4166000	6019690	5
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	1921152	12868302	46220	1350000	1328600	4
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	196	16 7.3%		10	7595262	2121174	67750	2250000	3147472	60
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4 8.0%		3	4950650	2713304	70152	1365000	824739	80
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	28	4 14.3%		5	3888172	1862882	39150	1700000	1861480	12
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	25	4 16.0%		0	2611302	358460	0	513000	133840	15
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	2423727	2604277	294420	1360000	1128000	17
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	196	13 6.6%		152	4960384	7497214	1748834	1230000	2108556	203
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4 8.0%		0	2073108	11352245	442870	1840000	2583441	25
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	33	4 12.1%		1	3482900	6301121	33250	4300000	2040578	34
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	4 16.7%		0	4882152	1415872	0	2479000	587240	35
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	2910152	10767890	77540	1623000	2180242	35
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	199	14 7.0%		146	7782150	6252633	2275040	750000	6177839	204
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	50	4 11.1%		0	1794296	5136522	556225	7017000	1494179	49
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	32	4 12.5%		5	2604581	6096799	82250	7757000	2155472	25
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	3 12.5%		0	3349512	1522312	86420	2972000	356800	28
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	1862562	11401009	63630	4994000	2230320	10
JABA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	13 6.6%		18	9563982	944124	173492	2915000	3109346	45
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	36	4 11.1%		0	3282292	1727492	0	1194000	260800	3
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	32	4 12.5%		0	3160391	915141	0	2660000	362050	5
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	4 16.7%		0	1366960	264050	0	520000	580800	2
JABA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4 15.4%		0	2501642	1600012	0	1660000	248360	2

Gambar 4. 46 Data Events Setelah Merge

b. Data Cleaning

```
import pandas as pd

# Membaca data dari file CSV
file_path    =    "C:/Users/user/Downloads/MY    MAGANG/DATA
TELKOMSEL/EVENTS/output EVENTS SBY.csv"

# Membaca file CSV dengan delimiter yang sesuai (ubah 'sep'
jika perlu)
data = pd.read_csv(file_path, sep=';')  # Ganti ';' dengan
',' jika file menggunakan koma sebagai pemisah

# Menghapus kolom 'OMSET TOTAL.1' jika kolom tersebut ada
if 'OMSET TOTAL.1' in data.columns:
    data.drop(columns=['OMSET TOTAL.1'], inplace=True)

# Memfilter data untuk BRANCH Surabaya
branch_surabaya = data.loc[data['BRANCH'] == 'SURABAYA']

# Menampilkan beberapa baris pertama untuk validasi
display(branch_surabaya.head())
```

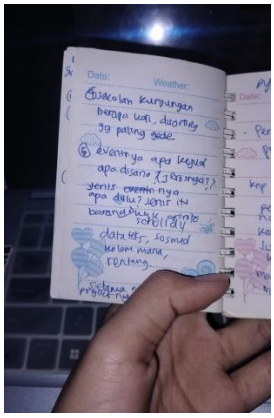
	AREA	REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO (%)	OMSET TRX DIGITAL	OMSET TOTAL	...	PLAN POI	ACTUAL POI	PLAN EVENT	ACTUAL EVENT	ACTUAL CALL (%)	SALES	OTA	SALES ORBIT	VF
114	JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	15	7.6	27	84386203	...	77	72	146	131	90.0	1	0	0	0
115	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	37	6	16.2	46	36878627	...	16	16	42	40	95.0	0	0	0	0
116	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	31	4	12.9	0	36184200	...	11	11	37	36	97.0	0	0	0	0
117	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	23	3	13.0	0	16763972	...	12	12	38	35	92.0	0	0	0	0
118	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	27	4	14.8	0	24237137	...	17	17	43	43	100.0	0	0	0	0

Gambar 4. 47 DataFrame

```
<class 'pandas.core.frame.DataFrame'>
Index: 45 entries, 114 to 1070
Data columns (total 24 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   AREA                45 non-null     object
 1   REGIONAL            45 non-null     object
 2   BRANCH              45 non-null     object
 3   CLUSTER             45 non-null     object
 4   CITY                45 non-null     object
 5   DS TOTAL            45 non-null     int64
 6   DS ACTIVE           45 non-null     int64
 7   DS RATIO (%)        45 non-null     int32
 8   OMSET TRX DIGITAL   45 non-null     int64
 9   OMSET TOTAL         45 non-null     int64
 10  OMSET DATA          45 non-null     int64
 11  OMSET DIGITAL       45 non-null     int64
 12  OMSET RECHARGE      45 non-null     int64
 13  OMSET OTHER         45 non-null     int64
 14  PLAN POI            45 non-null     int64
 15  ACTUAL POI          45 non-null     int64
 16  PLAN EVENT          45 non-null     int64
 17  ACTUAL EVENT        45 non-null     int64
 18  ACTUAL CALL (%)     45 non-null     int32
 19  SALES               45 non-null     int64
...
 22  VF                  45 non-null     int64
 23  TRX DIGITAL         45 non-null     int64
dtypes: int32(2), int64(17), object(5)
memory usage: 8.4+ KB
```

Gambar 4 48. Info Data

5. Rancangan Dashboard

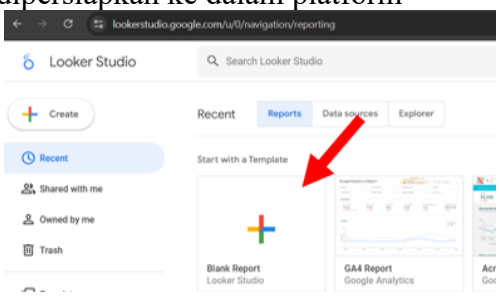


Gambar 4. 49 Rancangan Dashboard

Tahap perancangan dashboard secara konseptual bertujuan untuk menentukan informasi yang akan ditampilkan, struktur penyajiannya, dan jenis visualisasi yang sesuai. Proses ini melibatkan identifikasi data yang relevan untuk setiap variabel, memastikan kelengkapan dan kesesuaiannya dengan tujuan analisis. Pemilihan visualisasi difokuskan pada kemudahan interpretasi, kejelasan informasi, dan pengalaman pengguna, sehingga dashboard menjadi user-friendly dan mendukung pengambilan keputusan yang cepat dan tepat.

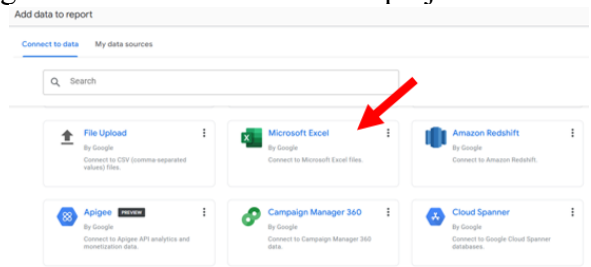
6. Pembuatan Dashboard

Pada tahapan ini dilakukan pembuatan dashboard yang terlah dirancang sebelumnya menggunakan software Looker Data Studio melalui URL: https://lookerstudio.google.com. Proses ini dimulai dengan menghubungkan dataset yang telah dipersiapkan ke dalam platform



Gambar 4. 50 Menambah Report di Looker Studio

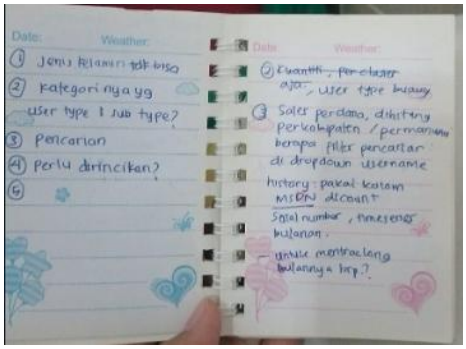
Hubungkan dataset ke dalam database looker data studio, karena penulis menghubungkan data dengan excel maka memilih ‘microsoft CSV seperti dan mengkonfigurasikan dataset ke halaman project



Gambar 4. 51 Menghubungkan Excel Ke Looker Studio

kemudian mengatur tata letak dan struktur dashboard agar sesuai dengan kebutuhan pengguna dengan membuat visualisasi data setiap variabel yang berelevan sampai menghasilkan dashboard untuk pengambilan dalam bisnis. Dan membuat desain layout dashboard sederhana yang nantinya akan mempermudah peletakan chart visualisasi. Proses ini hanya dilakukan langsung di looker data studio dengan memanfaatkan fitur shape.

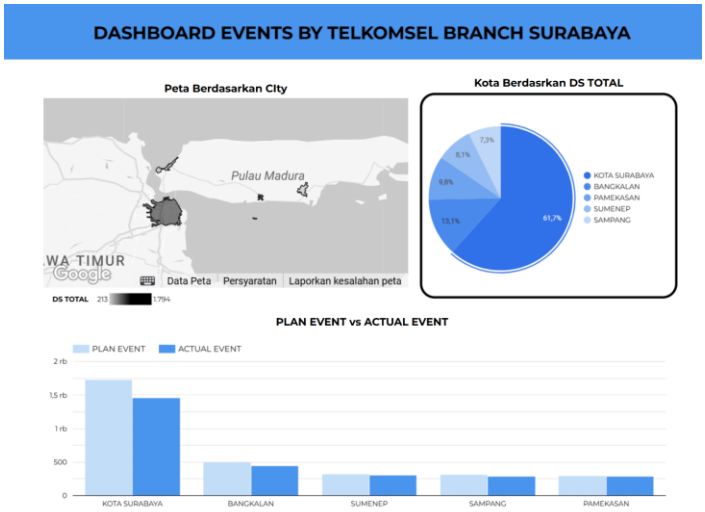
7. Validasi Dashboard



Gambar 4. 52 Catatan Evaluasi

Evaluasi dashboard yang dilakukan bersama mentor memberikan masukan penting untuk meningkatkan kualitas dan relevansi dashboard. Tujuan dari penambahan visualisasi ini adalah untuk memberikan konteks yang lebih komprehensif dalam analisis data. Informasi mengenai populasi penduduk dapat membantu pengguna dalam memahami Hubungan antara lokasi, khususnya di Surabaya dan Madura, dengan data dalam event dapat dilihat melalui pengaruh geografi terhadap kehadiran dan partisipasi dalam events. Dengan demikian, dashboard akan menjadi alat yang lebih strategis dalam mendukung pengambilan keputusan berbasis data. Penyesuaian ini juga diharapkan dapat meningkatkan daya tarik visual dashboard serta memberikan wawasan tambahan yang lebih informatif bagi pengguna, terutama dalam memahami potensi pasar di wilayah Surabaya dan Madura.

8. Hasil Dashboard



Gambar 4. 53 Hasil Dashboard Events

Dashboard ini menunjukkan distribusi dan performa pelaksanaan acara Telkomsel Cabang Surabaya di lima kota (Surabaya, Bangkalan, Sumenep,

Sampang, dan Pamekasan) dengan mengukur jumlah agen penjualan (DS Total) serta membandingkan jumlah acara yang direncanakan (Plan Event) dengan yang terealisasi (Actual Event). Kota Surabaya mendominasi jumlah agen penjualan dengan kontribusi 61.7%, sementara kota lain seperti Bangkalan (13.1%), Sumenep (8.5%), Sampang (8.1%), dan Pamekasan (7.2%) memiliki distribusi lebih kecil. Grafik Plan Event vs Actual Event menunjukkan bahwa Surabaya mengalami gap signifikan antara rencana dan realisasi acara, yang dapat memengaruhi indikator keberhasilan pelaksanaan event sesuai target dan, pada akhirnya, pendapatan bisnis. Sebaliknya, kota lain menunjukkan stabilitas dengan realisasi yang mendekati rencana. Dari sisi bisnis, hal ini menunjukkan bahwa Telkomsel dapat meningkatkan penetrasi pasar dan memperkuat brand awareness di kota-kota ini melalui pelaksanaan event yang konsisten. Namun, peluang terbesar ada di Surabaya, yang memiliki pasar besar tetapi masih belum optimal dari segi eksekusi acara. Optimalisasi eksekusi acara di Surabaya dapat berdampak langsung pada peningkatan akuisisi pelanggan, loyalitas pengguna, dan pendapatan. Untuk mencapainya, diperlukan monitoring rutin, alokasi sumber daya yang tepat, serta strategi berbasis data untuk memastikan pelaksanaan event mampu memberikan dampak bisnis yang maksimal di semua wilayah.

4.6. Rekognisi Mata Kuliah Statistika Sains Data dan Analisis Kebijakan

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Statistika Sains Data dan Analisis Kebijakan. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Statistika Sains Data dan Analisis Kebijakan

Tabel 4. 7 Indikator Penilaian Konversi Mata Kuliah Statistika Sains Data dan Analisis Kebijakan

No	Kriteria/Indikator Penilaian
1	Kemampuan dalam menjelaskan konsep statistika dan analisis data, dasar analisis kebijakan, pendekatan dalam proses pembuatan kebijakan publik, kerangka kebijakan publik, prosedur dalam analisis kebijakan publik, metode kajian isu kebijakan dan langkah analisis kebijakan publik, teknik penyusunan dan model dalam analisis kebijakan publik
2	Kemampuan memilih metode dan pendekatan statistika sains data yang sesuai untuk mengumpulkan, menganalisis, dan menginterpretasi data dalam konteks analisis kebijakan
3	Kemampuan menggunakan perangkat lunak pendukung (tools) seperti perangkat statistik, perangkat lunak analisis dan pemodelan data, serta perangkat lunak visualisasi data untuk menjalankan analisis statistika dan mengelola data dalam konteks praktik analisis kebijakan secara efektif dan efisien
4	Kemampuan merancang dan memformulasikan, dan mendesain workflow atau prosedur analisis data untuk mengevaluasi implikasi kebijakan berdasarkan hasil analisis statistik

4.6.1. Materi Statistika Sains Data dan Analisis Kebijakan [Indikator ke-1]

Konsep Statistika Dan Analisis Data

Statistika adalah ilmu yang mempelajari cara merencanakan, mengumpulkan, menganalisis, menginterpretasi, dan mempresentasikan data. Dalam konteks ini, analisis statistika terbagi menjadi dua kategori utama: statistika deskriptif dan statistika inferensial. Statistika deskriptif berfungsi untuk menyajikan data dalam bentuk yang lebih mudah dipahami, seperti tabel, grafik, dan ukuran pemusatan (mean, median, modus) yang memberikan gambaran umum tentang karakteristik variabel yang diteliti (Mardhotillah dkk., 2022).

Konsep dasar dari statistika mencakup berbagai teknik yang membantu dalam menganalisis data numerik dan membuat prediksi atau inferensi berdasarkan data yang ada. Di dalam statistika, terdapat dua cabang utama, yaitu:

1. Statistika deskriptif.
   Statistika deskriptif bertujuan untuk menggambarkan dan menyimpulkan karakteristik utama dari sebuah dataset, sedangkan statistika inferensial digunakan untuk membuat prediksi atau generalisasi tentang populasi berdasarkan sampel yang tersedia. Statistika deskriptif melibatkan penggunaan beberapa ukuran ringkasan seperti mean (rata-rata), median, modus, standar deviasi, serta varian untuk memberikan gambaran umum tentang data. Misalnya, untuk mengevaluasi kinerja sales, kita dapat menghitung rata-rata penjualan harian (mean) dan menghitung variansi atau standar deviasi untuk mengetahui seberapa besar fluktuasi penjualan di tiap hari. Beberapa rumus statistika deskriptif yang sering digunakan adalah sebagai berikut:
   - Mean (Rata-rata)

Mean (Rata-rata) adalah ukuran yang paling umum digunakan untuk menggambarkan pusat data. Rata-rata dihitung dengan menjumlahkan semua nilai data, lalu membaginya dengan jumlah data.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Dimana:
$\bar{x}$ adalah rata-rata,
x_i adalah nilai data ke-i,
n adalah jumlah data.

- Median (Nilai Tengah)
  Median adalah nilai tengah dari data yang sudah diurutkan. Jika jumlah data ganjil, median adalah nilai yang terletak tepat di tengah. Median sering kali digunakan ketika kita ingin menghindari pengaruh dari nilai ekstrim yang bisa mempengaruhi rata-rata.
  Rumus median untuk n ganjil

$$Me = x_{\frac{n+1}{2}}$$

  Rumus median untuk n genap

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}} + 1}{2}$$

- Modus
  Modus adalah nilai yang paling sering muncul dalam suatu kumpulan data. Berbeda dengan rata-rata dan median yang memerlukan semua data, modus hanya memerlukan pengidentifikasian nilai yang berulang.

Selain mean, median, dan modus, ada beberapa ukuran lain dalam statistika deskriptif, seperti varians yang mengukur sejauh mana data tersebar di sekitar rata-rata, dengan menghitung rata-rata kuadrat selisih antara setiap data dan rata-rata. Deviasi standar adalah akar kuadrat dari varians, memberikan gambaran yang lebih mudah dipahami tentang penyebaran data. Rentang mengukur selisih antara nilai maksimum dan minimum dalam data, memberikan gambaran sederhana mengenai sebaran data. Kurtosis mengukur bentuk puncak distribusi data, apakah lebih tajam atau datar dibandingkan dengan distribusi normal, sedangkan skewness menggambarkan sejauh mana distribusi data miring, apakah cenderung ke kiri (negatif) atau kanan (positif). Semua ukuran ini membantu dalam memahami distribusi dan karakteristik data yang lebih mendalam.

- Statistika Inferensia
  Statistika inferensia adalah cabang statistika yang digunakan untuk menarik kesimpulan atau membuat prediksi tentang populasi berdasarkan sampel data yang diambil dari populasi tersebut. Dalam statistika inferensia, kita menggunakan teknik-teknik untuk menguji hipotesis, memperkirakan parameter populasi, dan mengukur ketidakpastian hasil yang diperoleh. Berikut beberapa konsep utama dalam statistika inferensia:
  a. Estimasi Parameter
     Estimasi parameter digunakan untuk memperkirakan nilai suatu parameter populasi, seperti rata-rata atau proporsi. Estimasi ini dapat dilakukan melalui estimasi titik, di mana kita menghitung nilai rata-rata sampel untuk memperkirakan rata-rata populasi.

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

Selain itu, ada juga estimasi interval, yang memberikan rentang nilai dengan tingkat kepercayaan tertentu untuk parameter populasi.

$$\mu \pm Z_{\alpha/2} \times \frac{8}{\sqrt{n}}$$

Di mana $Z_{\alpha/2}$ adalah nilai Z untuk tingkat kepercayaan yang diinginkan, s adalah deviasi standar sampel, dan n adalah ukuran sampel.

b. Pengujian Hipotesis

Pengujian hipotesis adalah teknik yang digunakan untuk menguji klaim atau asumsi tentang populasi. Dalam pengujian hipotesis, kita membandingkan statistik sampel dengan nilai yang dihipotesiskan untuk populasi menggunakan uji statistik, seperti uji t. Uji t adalah salah satu teknik statistik yang digunakan untuk menguji apakah terdapat perbedaan yang signifikan antara dua kelompok atau lebih, terutama ketika data yang digunakan berdistribusi normal dan ukuran sampelnya relatif kecil. Uji ini biasanya digunakan untuk menguji hipotesis mengenai rata-rata dari satu atau dua kelompok data. Kegunaan uji t untuk menguji apakah terdapat perbedaan yang signifikan antara rata-rata dua kelompok atau lebih, terutama ketika ukuran sampel relatif kecil dan data berdistribusi normal. Berikut adalah rumus uji t:

$$t = \frac{\bar{x} - \mu_0}{\frac{8}{\sqrt{n}}}$$

Di mana $\bar{x}$ adalah rata-rata sampel, $\mu_0$adalah nilai yang dihipotesiskan untuk rata-rata populasi, s adalah deviasi standar sampel, dan n adalah ukuran sampel.

c. Analisis Regresi dan Korelasi

Analisis Regresi digunakan untuk memodelkan hubungan antara variabel dependen dan independen. Dalam regresi linier, kita berusaha untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen dengan menghitung koefisien regresi menggunakan data sampel. Berikut adalah rumus Regresi linear:

$$y = \beta_0 + \beta_1 x + \in$$

Di mana y adalah variabel dependen, x adalah variabel independen, β_0 adalah intercept, β_1 adalah koefisien regresi, dan $\epsilon 3$ adalah *error*. Untuk analisis korelasi digunakan untuk mengukur sejauh mana dua variabel saling berkaitan atau berhubungan satu sama lain. Korelasi tidak melibatkan prediksi atau pemodelan, tetapi lebih kepada seberapa kuat hubungan linear antara dua variabel tersebut. Berikut rumus analisi korelasi yang paling umu :

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

Konsep Dasar Analisis Kebijakan Publik

Analisis kebijakan adalah proses yang sistematis dan terstruktur yang bertujuan untuk mengumpulkan informasi secara menyeluruh dan menarik

kesimpulan logis yang dapat diandalkan guna mendukung pengambilan keputusan yang efektif dalam pembuatan kebijakan publik. Proses ini tidak hanya mencakup pengumpulan data, tetapi juga melibatkan sintesis informasi dari berbagai sumber yang beragam, termasuk hasil penelitian akademis, laporan pemerintah, serta masukan dari pemangku kepentingan dan masyarakat (Penelitian dkk., t.t.).

Proses pembuatan kebijakan publik adalah suatu rangkaian yang kompleks dan berlapis, yang melibatkan beberapa tahapan kunci yang saling terkait dan berkesinambungan. Tahapan-tahapan tersebut dimulai dengan identifikasi masalah, penentuan agenda kebijakan, perumusan kebijakan dengan rumus analisis biaya – manfaat sebagai berikut :

$$NPV = \sum_{t=0}^{T} \frac{B_t - C}{(1+r)^t}$$

Dimana :

NPV : Net Present Value

B_t : Manfaat pada tahun t

C_t : Biaya pada tahun t

r : Tingkat diskonto (suku bunga)

T : Total periode waktu

Pada tahap ini, penting untuk mempertimbangkan berbagai perspektif dan data yang relevan agar kebijakan yang diusulkan dapat memenuhi kebutuhan masyarakat secara holistik. Setelah perumusan selesai, proses berlanjut ke pengambilan keputusan, di mana pembuat kebijakan harus memilih opsi terbaik dari alternatif yang telah dirumuskan, sering kali dengan mempertimbangkan faktor politik, sosial, dan ekonomi yang ada. Tahap selanjutnya adalah implementasi, dan proses pembuatan kebijakan tidak lengkap tanpa tahap evaluasi, di mana dampak dari kebijakan yang telah diimplementasikan dianalisis untuk menentukan apakah tujuan awal telah tercapai dan bagaimana kebijakan tersebut dapat diperbaiki atau disesuaikan di masa depan. Setiap tahap dalam proses ini memiliki pertanyaan kunci yang harus dijawab untuk memastikan bahwa kebijakan yang dihasilkan tidak hanya efektif tetapi juga responsif terhadap kebutuhan masyarakat, sehingga mampu memberikan manfaat nyata bagi publik (Prabawati dkk, 2020)

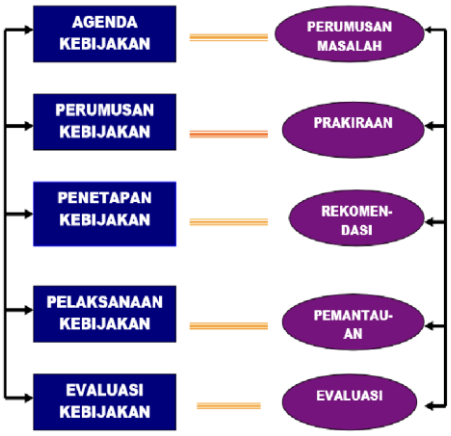
Kerangka kebijakan publik adalah suatu struktur yang komprehensif dan sistematis yang mendefinisikan proses bagaimana isu-isu publik diidentifikasi, dibahas, dan diatur dalam agenda kebijakan. Terdapat beberapa kerangka kerja analisis kebijakan, yaitu :

1. **Welfare Economics**:
   Fokus pada peningkatan kesejahteraan sosial tanpa mengorbankan orang lain (Pareto optimality). Kesejahteraan sosial tercapai jika ada peningkatan untuk sebagian orang tanpa merugikan orang lain. Pemerintah mengatasi trade-off antara efisiensi dan pemerataan melalui kebijakan redistribusi seperti pajak dan subsidi.
2. **Public Choice**:

Public choice menganalisis bagaimana pemerintah membuat keputusan berdasarkan kepentingan publik, dengan fokus pada perilaku aktor politik seperti pemilih, partai politik, dan birokrat. Pendekatan ini melihat agregasi preferensi individu untuk memaksimalkan kesejahteraan sosial dan menganggap pemerintah sebagai penyedia komoditas publik yang dipilih masyarakat.

3. **Social Structure**:
   Menganalisis kebijakan dari sudut pandang sosiologi, dengan fokus pada kekuasaan, struktur sosial, dan nilai budaya. Struktur sosial melibatkan pola interaksi yang stabil, didukung oleh peraturan dan norma. Pemerintah mempengaruhi kepercayaan, kohesi, dan partisipasi sosial melalui kebijakan dan kolaborasi sosial.

Prosedur kebijakan baru dimulai ketika para pelaku kebijakan mulai sadar bahwa adanya situasi permasalahan, yaitu situasi yang dirasakan adanya kesulitan atau kekecewaan dalam perumusan kebutuhan, nilai dan kesempatanmenurut Ackoff dalam Dunn (2000:121). Dunn (2000-21) berpendapat bahwa metodologi analisis kebijakan menggabungkan lima prosedur umum yang lazim dipakai dalam pemecahan masalah manusia: definisi, prediksi, preskripsi, deskripsi, dan evaluasi. Dalam analisis kebijakan prosedur-prosedur tersebut memperoleh nama-nama khusus, yakni:



Gambar 4. 54 Prosedur Kebijakan

1. Perumusan Masalah
   Perumusan masalah (definisi) menghasilkan informasi mengenai kondisi-kondisi yang menimbulkan masalah kebijakan.
2. Peramalan
   Peramalan (prediksi) menyediakan informasi mengenai konsekuensi di masa mendatang dari penerapan alternatif kebijakan.
3. Rekomendasi
   Rekomendasi (preskripsi) menyediakan informasi mengenai nilai atau kegunaan relatif dari konsekuensi di masa depan dari suatu pemecahan masalah.
4. Pemantauan
   Pemantauan (deskripsi), menghasilkan informasi tentang konsekuensi sekarang dan masa lalu dari diterapkannya alternatif kebijakan.
5. Evaluasi

Evaluasi, yang mempunyai nama sama dengan yang dipakai dalam bahasa sehari-hari, menyediakan informasi mengenai nilai atau kegunaan dari konsekuensi pemecahan atau pengatasan masalah

Metode kajian isu kebijakan yaitu dapat bersifat kuantitatif maupun kualitatif, untuk metode kuantitatif menurut Sugiyono berfokus pada pendekatan positivis dengan tujuan mengukur fenomena sosial secara objektif melalui analisis numerik. Penelitian ini bersifat behavioristik dan mekanistik, menggunakan desain yang spesifik, jelas, dan rinci. Tujuannya untuk menunjukkan hubungan antar variabel, menguji teori, dan mencari generalisasi prediktif. Selain itu untuk metode kualitatif Selain itu, untuk metode kualitatif, pendekatannya bersifat interpretatif dengan tujuan memahami fenomena sosial secara mendalam melalui analisis deskriptif. Penelitian ini menekankan pada makna, konteks, dan pengalaman individu, menggunakan desain yang fleksibel dan berkembang sesuai kebutuhan lapangan. Tujuannya adalah untuk mengeksplorasi dan menggali perspektif partisipan, memahami proses sosial, serta memberikan wawasan yang kaya dan holistik terhadap isu kebijakan yang dikaji.

Teknik Penyusunan Dan Model Dalam Analisis Kebijakan Publik

Abdul Wahab (2012:128) menjelaskan bahwa analisis kebijakan dapat menggunakan berbagai teknik pengambilan keputusan, di antaranya adalah:

1. Analisis SWOT
   Menilai kekuatan, kelemahan, peluang, dan ancaman untuk merumuskan alternatif kebijakan dan menentukan prioritas kebijakan.
2. AHP (Analytical Hierarchy Process)
   Teknik untuk memilih alternatif kebijakan berdasarkan faktor-faktor signifikan yang disaring secara berjenjang hingga muncul alternatif yang paling sesuai.
3. Dinamika Sistem (System Dynamics)
   Menganalisis interaksi antar faktor yang mempengaruhi masalah kebijakan, mengidentifikasi faktor pengungkit dan kendala, dan merumuskan alternatif kebijakan. Dengan model sederhana :

$$\frac{dX}{dt} = f(X, t)$$

di mana X adalah variabel sistem dan $f(X, t)$ adalah fungsi yang menggambarkan perubahan X berdasarkan waktu t.

4.6.2. Analisis Strategi Sales Berdasarkan Area Menggunakan Bahasa Pemograman Python [Indikator ke-2, Indikator ke-3, dan Indikator ke-4]

Gambar 4. 55 Workflow Statistika Sains Data dan Analisis Kebijakan

Berikut adalah workflow untuk analisis strategis sales berdasarkan area menggunakan bahasa pemograman python:

1. Mulai
   Proses dimulai dengan menentukan tujuan analisis, yaitu menilai kinerja sales pada daerah dengan konsep statistika dan data science.
2. Input Data
   Memasukkan data csv ke dalam bahasa pemograman python
3. Cleaning Data
   Menghapus kolom yang tidak relevan, mengisi nilai yang hilang, mengatasi outllier dan menghapus data duplikat, lalu memfilter data daerah surabaya, dan mengubah tipe data.
4. Analisis dan Visualisasi Data
   Proses analisis data mencakup menghitung hitung statistik deskriptif (sum, mean, max, dll) untuk kolom SALES
5. Kesimpulan dan analisis kebijakan
   Kesimpulan dari analisis implementasikan strategi pemasaran berbasis lokasi, dengan pendekatan yang lebih terfokus pada kota-kota yang memiliki potensi untuk berkembang serta menghasilkan kebijakan sales atau pemasaran berdasarkan area untuk Telkomsel Branch Surabaya
6. Selesai
   Hasil akhir dipresentasikan dalam laporan yang mencakup grafik dan insight yang didapat dari analisis.

Untuk menganalisis strategi sales berdasarkan area bahwa ada perbedaan yang signifikan dalam tingkat penjualan antara kota-kota yang berbeda. Kota dengan penjualan rendah menunjukkan potensi pasar yang belum tergali sepenuhnya, sementara kota dengan penjualan tinggi menunjukkan konsistensi dalam performa yang perlu dipertahankan. Tujuan penulis menganalisis hal ini untuk menerapkan dan mengetahui pendekatan pemasaran berbasis lokasi untuk memperkenalkan produk atau layanan Telkomsel lebih intensif di kota. Kebijakan dalam hal ini yaitu strategi pemasaran berbasis area atau lokasi, pendekatan ini berfokus pada strategi sales berdasarkan lokasi geografis untuk memaksimalkan potensi pasar yang ada di berbagai area, dan mengambil langkah-langkah yang lebih terfokus untuk meningkatkan penjualan di setiap lokasi. Berikut langkah – langkah untuk menganalisis strategi sales berdasarkan area menggunakan bahasa pemograman python

- Impor library

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy import stats
```

Kode di atas mengimpor berbagai pustaka yang berguna untuk analisis data dan visualisasi. Pandas digunakan untuk manipulasi dan analisis data, seperti mengimpor dan mengolah data dalam bentuk DataFrame. Seaborn adalah pustaka visualisasi yang dibangun di atas matplotlib, yang menyediakan cara mudah untuk membuat grafik statistik yang informatif dan menarik, seperti diagram kotak, heatmap, atau grafik distribusi. Matplotlib adalah pustaka utama untuk membuat visualisasi data secara umum, seperti grafik garis, bar,

atau scatter plot. SciPy adalah pustaka yang menyediakan berbagai fungsi statistik dan ilmiah, termasuk alat untuk melakukan uji hipotesis, analisis statistik, atau metode numerik lainnya, seperti uji ANOVA atau uji distribusi. Gabungan pustaka-pustaka ini memberikan kemampuan untuk menganalisis, memvisualisasikan, dan melakukan uji statistik terhadap data secara efisien.

- Input data

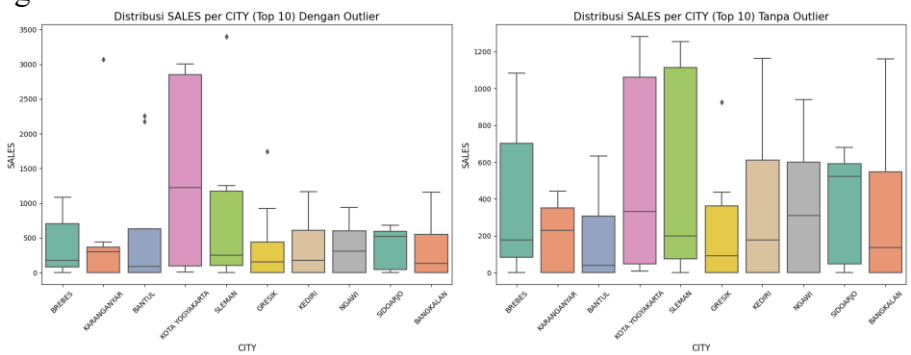
  Penulis menggunakan data events khususnya di daerah Surabaya - Madura untuk menganalisis jumlah sales berdasarkan kota-kota yang ada di wilayah tersebut. Variabel utama yang digunakan dalam analisis ini adalah kolom 'CITY', yang merepresentasikan nama-nama kota di daerah Surabaya-Madura, dan kolom 'SALES', yang mencatat jumlah penjualan untuk setiap kota.

	AREA	REGIONAL	BRANCH	CLUSTER	CITY	DS TOTAL	DS ACTIVE	DS RATIO (%)	OMSET TRX DIGITAL	OMSET TOTAL	...	PLAN POI	ACTUAL POI	PLAN EVENT	ACTUAL EVENT	ACTUAL CALL (%)	SALES	OTA	SALES ORBIT	VF	TRX DIGITAL
0	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	BULELENG	42	3	7	39	11043601	...	49	37	72	50	69	0	0	0	0	29
1	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	JEMBRANA	23	3	13	39	14135322	...	48	43	75	56	75	0	0	0	0	30
2	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI BARAT	TABANAN	20	3	15	37	8446941	...	50	46	53	48	91	5	0	0	0	36
3	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI TENGAH	BADUNG	42	6	14	92	24862947	...	116	77	156	100	64	59	0	0	125	89
4	JAWA - BALI - NUSRA	BALI NUSRA	DENPASAR	BALI TENGAH	KOTA DENPASAR	107	6	5	84	32561574	...	32	29	152	105	69	6	0	0	38	80
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1066	JAWA - BALI - NUSRA	JATIM	SURABAYA	KOTA SURABAYA	KOTA SURABAYA	197	13	6	18	9563982	...	45	43	51	49	96	24	0	0	0	18
1067	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	BANGKALAN	36	4	11	0	3282292	...	1	1	4	4	100	0	0	0	0	0
1068	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	PAMEKASAN	32	4	12	0	3308591	...	1	1	4	4	100	11	0	0	0	0
1069	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SAMPANG	24	4	16	0	1364900	...	2	1	5	4	80	37	0	0	0	0
1070	JAWA - BALI - NUSRA	JATIM	SURABAYA	MADURA	SUMENEP	26	4	15	0	2593442	...	2	2	5	5	100	11	0	0	0	0

1063 rows × 24 columns

Gambar 4. 56 Data Events Surabaya

- Mengatasi outlier



Gambar 4. 57 Sebelum & Setelah Mengangani Outlier

Berdasarkan visualisasi pada diagram boxplot, berikut kesimpulan :

Gambar Kiri (Dengan Outlier):

Terdapat beberapa nilai outlier yang signifikan di beberapa kota, terutama pada kota "KOTA YOGYAKARTA" yang memiliki nilai sales yang jauh lebih tinggi dibandingkan kota lain.

Variasi nilai sales di "KOTA YOGYAKARTA" sangat besar, terlihat dari rentang interkuartil (IQR) yang lebih lebar dibandingkan kota lainnya.

Sebagian besar kota lainnya memiliki distribusi sales yang lebih seragam dengan sedikit atau tanpa nilai outlier.

Gambar Kanan (Tanpa Outlier):

Setelah menghilangkan outlier, distribusi sales menjadi lebih proporsional di seluruh kota. Kota "KOTA YOGYAKARTA" tetap memiliki median sales yang relatif tinggi, tetapi perbedaannya dengan kota lain menjadi lebih kecil.

Kota-kota seperti "BIREUEN" dan "BANYUMAS" menunjukkan distribusi yang lebih homogen setelah outlier dihapus.
Rentang IQR dari setiap kota terlihat lebih seimbang dibandingkan gambar kiri.

- Analisis deskriptif

	sum	mean	max	min	std	count
CITY						
BANGKALAN	3273	363.666667	1161	0	484.508514	9
BANTUL	1282	183.142857	632	0	272.843929	7
BREBES	3488	387.555556	1084	0	395.816912	9
GRESIK	1882	235.250000	925	0	325.842534	8
KARANGANYAR	1617	202.125000	442	0	184.705509	8
KEDIRI	3307	367.444444	1165	0	441.121613	9
KOTA YOGYAKARTA	3213	535.500000	1284	9	594.009680	6
NGAWI	3200	355.555556	940	0	331.515125	9
SIDOARJO	3120	346.666667	681	0	308.399903	9
SLEMAN	4025	503.125000	1255	0	563.125826	8

Gambar 4. 58 Hasil Analisis Deksriptif

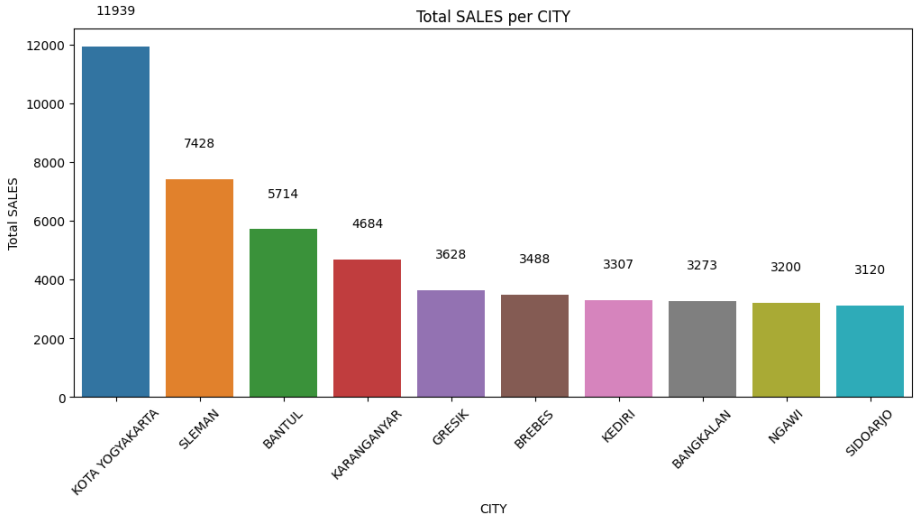
Statistika deskriptif memberikan gambaran distribusi data dari beberapa kota. Kota dengan total nilai tertinggi adalah SLEMAN (4.025), sedangkan BANTUL memiliki total terendah (1.282). Berdasarkan rata-rata, KOTA YOGYAKARTA memiliki nilai tertinggi (535,50), menunjukkan konsistensi nilai yang lebih besar dibandingkan kota lain, sementara BANTUL memiliki rata-rata terendah (183,14). Dari segi nilai maksimum, KOTA YOGYAKARTA mencatat nilai tertinggi (1.284), sedangkan KARANGANYAR memiliki nilai terendah (442). Semua kota memiliki nilai minimum 0, yang mungkin mencerminkan adanya kejadian atau data kosong. Untuk standar deviasi, KOTA YOGYAKARTA memiliki variabilitas data tertinggi (594,01), sedangkan KARANGANYAR memiliki variabilitas terendah (184,71), menunjukkan data yang lebih terpusat. Dari jumlah observasi, sebagian besar kota memiliki 9 observasi, namun BANTUL memiliki 7, dan KOTA YOGYAKARTA hanya memiliki 6, menunjukkan adanya variasi jumlah yang digunakan.

- Uji ANOVA

Hasil uji ANOVA dapat menunjukkan bahwa perbedaan SALES antar kota cukup signifikan. Berikut penjelasannya:

F-statistic: 2.9486 menunjukkan bahwa ada variasi yang cukup besar antara grup (kota) dibandingkan dengan variasi dalam grup tersebut.
p-value: 1.0966e-19: Nilai p-value yang sangat kecil (jauh di bawah 0.05) menunjukkan bahwa hasil uji ANOVA signifikan, yang berarti bahwa perbedaan SALES antar kota bukanlah kebetulan dan ada pengaruh nyata dari kota terhadap SALES. Jadi, hasilnya bagus, karena p-value yang sangat kecil menunjukkan bahwa perbedaan SALES antar kota itu signifikan dan layak untuk dipertimbangkan lebih lanjut dalam analisis atau keputusan.

- Visualisasi data



Gambar 4. 59 Grafik Total Sales Tiap Kota

Berdasarkan grafik "Total SALES per CITY," terlihat bahwa "KOTA YOGYAKARTA" mendominasi total penjualan dengan nilai tertinggi sebesar 11.939, untuk "SLEMAN" (7.428) dan "BANTUL" (5.714) yang juga berkontribusi signifikan terhadap total penjualan. Kota-kota seperti "KARANGANYAR" (4.684), "GRESIK," "BREBES," dan "KEDIRI" memiliki penjualan dalam kategori menengah (3.000–4.000), sementara "SIDOARJO" mencatat total penjualan terendah sebesar 3.120, tidak jauh berbeda dari "NGAWI" dan "BANGKALAN." Dengan adanya kesenjangan besar antara kota-kota teratas dan kota lainnya, strategi bisnis dapat difokuskan pada memperkuat dominasi di kota-kota dengan penjualan tinggi atau meningkatkan kontribusi kota-kota menengah untuk mendorong pertumbuhan pasar.

- Analisis kebijakan

  Untuk meningkatkan kebijakan pada kinerja penjualan dan mengoptimalkan strategi pemasaran di tiap kota, Telkomsel perlu mengimplementasikan strategi pemasaran berbasis lokasi dengan pendekatan yang berbeda untuk kota-kota dengan karakteristik pasar yang berbeda. Di kota dengan penjualan rendah berfokus pada kampanye pemasaran yang lebih terpersonalisasi, seperti iklan lokal, acara komunitas, atau promosi berbasis lokasi, guna meningkatkan penetrasi pasar dan kesadaran merek. Di kota dengan penjualan tinggi, tingkatkan kegiatan pemasaran digital dan loyalitas pelanggan untuk mempertahankan dan memperkuat posisi dominan Telkomsel di pasar. Pendekatan ini akan membantu Telkomsel dalam mengalokasikan sumber daya secara lebih efektif dan mencapai hasil penjualan yang lebih optimal di setiap lokasi.

4.7. Rekognisi Mata Kuliah Analisis Deret Waktu

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Analisis Deret Waktu. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Analisis Deret Waktu

Tabel 4. 8 Indikator Penilaian Konversi Mata Kuliah Analisis Deret Waktu

No	Kriteria/Indikator Penilaian
1	Kemampuan dalam menjelaskan konsep analisis deret waktu, eksplorasi data deret waktu (plot, transformasi, dekomposisi, dan karakteristiknya), proses stasioner dan nonstasioner serta pemodelannya, estimasi parameter model, diagnostik dan peramalan model, heteroskedastik pada model, model ARIMA dan Seasonal ARIMA, serta inferensi model peramalan
2	Kemampuan memilih metode dan pendekatan analisis deret waktu yang tepat yang dapat digunakan dalam menyelesaikan studi kasus model deret waktu stasioner, nonstasioner, musiman, dan berbagai model deret waktu lainnya
3	Kemampuan menggunakan perangkat lunak pendukung (tools) analisis deret waktu dalam menyelesaikan studi kasus deret waktu stasioner, nonstasioner, musiman secara efektif dan efisien
4	Kemampuan merancang dan memformulasikan, dan mendesain workflow analisis deret waktu untuk menyelesaikan permasalahan pada suatu studi kasus tertentu

4.7.1. Materi Analisis Deret Waktu [Indikator ke-1]

Analisis deret waktu merupakan teknik yang digunakan untuk mempelajari data yang tercatat secara berurutan dalam waktu tertentu, seperti data penjualan bulanan atau suhu harian. Tujuan utamanya adalah untuk menemukan pola dalam data dan menggunakannya untuk meramalkan nilai di masa depan. Fungsi momen dan stasioneritas deret waktu melibatkan statistik seperti rata-rata dan varians, yang seharusnya tetap stabil sepanjang waktu untuk memastikan data bersifat stasioner. Visualisasi data deret waktu biasanya dilakukan melalui plot untuk mengidentifikasi pola, tren, dan anomali.

Partial Autocorrelation Function (PACF) digunakan untuk mengukur hubungan antara nilai-nilai pada waktu yang berbeda dengan memperhitungkan pengaruh pengamatan sebelumnya. Partial autocorrelation didefinisikan sebagai berikut:

$$\phi_{k,k} = \frac{Cov(Z_t, Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1})}{\sqrt{Var}(Z_t | Z_{t-1}, \dots, Z_{t-k+1}) . \sqrt{Var}(Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1})}$$

Dimana :

$\phi_{k,k}$: Nilai PACF pada lag ke-k, menunjukkan korelasi parsial antara Z_t dan Z_{t-k} setelah mengeliminasi efek dari semua lags di antaranya. Nilainya berkisar antara -1 dan 1.

Uji White Noise digunakan untuk mengevaluasi apakah data tersebut mengandung variasi acak tanpa pola yang jelas. Simple Time Series Compositions memisahkan data menjadi komponen tren, musiman, dan residual, sementara Time Series Decomposition and Smoothing membagi data menjadi bagian-bagian utama dan menghaluskan data untuk analisis yang lebih mendalam. Backshift Operator dalam

pemodelan merujuk pada nilai sebelumnya dalam deret waktu. Teknik differencing digunakan untuk membuat data menjadi stasioner dengan menghilangkan tren. Uji stasioneritas, seperti uji Dickey-Fuller, dapat digunakan untuk memeriksa apakah data bersifat stasioner. Model Moving Average (MA) meramalkan nilai masa depan dengan menggunakan rata-rata kesalahan prediksi sebelumnya, sedangkan model Autoregressive (AR) memanfaatkan nilai sebelumnya untuk memprediksi nilai saat ini. Model gabungan antara AR dan MA disebut Autoregressive Moving Average (ARMA). Berikut merupakan model ARMA

$$X_t = \mathrm{C} + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q}$$

Dimana :

X_t : Nilai data pada waktu t (yang diprediksi)

C : Nilai rata – rata (konstanta) dari data

$\phi_1, \ldots., \phi_p$: Parameter bagian AR

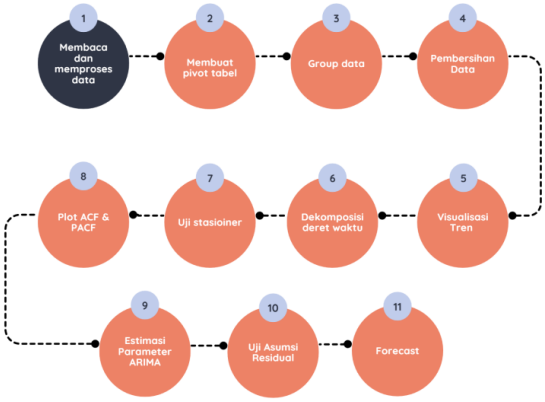
$\theta_1, \ldots., \theta_1$: Parameter bagian MA

e_t : Nilai *error* pada waktu t

Dalam membangun model, tantangan yang dihadapi meliputi pemilihan parameter yang tepat. Metode Estimasi digunakan untuk memperkirakan parameter model melalui teknik seperti kuadrat terkecil atau maksimum likelihood. Penentuan urutan model ARMA dan ARIMA merupakan langkah untuk menemukan parameter terbaik. Diagnosis model melibatkan pemeriksaan kecocokan model dengan menganalisis residual, sedangkan peramalan digunakan untuk memprediksi nilai masa depan dengan model yang telah dibangun.

SARIMA (Seasonal ARIMA) merupakan model yang digunakan untuk menganalisis data deret waktu dengan pola musiman. Model ini memperluas ARIMA dengan menambahkan komponen musiman untuk menangani variasi periodik dalam data. SARIMA cocok untuk menganalisis dan meramalkan data yang menunjukkan pola musiman yang jelas, seperti penjualan musiman atau suhu tahunan, dan memberikan prediksi yang lebih akurat dengan memperhitungkan faktor musiman yang tidak terdeteksi oleh model ARIMA standar.

4.7.2. Analisis Deret Waktu Pada Data Kunjungan Telkomsel ke Sekolah di Indonesia Pada Bulan Januari – Oktober 2024 Dengan Metode Arima Menggunakan Bahasa Pemograman Python [Indikator ke-2, Indikator ke-3, Indikator ke-4]



Gambar 4. 60 Diagram Alur Pemodelan Analisis Deret Waktu Data Kunjungan Telkomsel Ke Sekolah Menggunakan Metode Arima

Untuk menyelesaikan hasil pemodelan analisis deret waktu data kunjungan telkomsel ke sekolah menggunakan metode arima penulis menggambarkan langkah-langkah dalam proses pemodelan analisis deret waktu data kunjungan telkomsel ke sekolah menggunakan metode arima

1. Membaca dan Memproses Data
   Tahap awal adalah membaca dataset yang akan dianalisis. Data perlu diproses agar memiliki format yang sesuai untuk analisis lebih lanjut, seperti memeriksa nilai yang hilang, duplikasi, atau format tanggal.
2. Membuat Pivot Tabel
   Langkah ini melibatkan penyusunan data ke dalam format pivot tabel untuk mempermudah analisis. Tujuan dari pivot tabel adalah untuk merangkum data dengan cara yang lebih terstruktur, misalnya agregasi berdasarkan kategori tertentu.
3. Group Data
   Data yang sudah dirapikan akan dikelompokkan berdasarkan variabel tertentu, seperti waktu atau kategori lainnya, agar lebih mudah dianalisis. Proses ini membantu memahami pola berdasarkan kelompok data.
4. Pembersihan Data
   Tahap ini mencakup penghapusan outlier, pengisian nilai yang hilang, dan koreksi data yang tidak sesuai. Tujuannya adalah memastikan data bersih dan bebas dari anomali sebelum digunakan untuk analisis.
5. Visualisasi Tren
   Setelah data bersih, tren pada data divisualisasikan untuk memberikan gambaran awal tentang pola yang terjadi dalam data dari waktu ke waktu.
6. Dekomposisi Deret Waktu
   Pada langkah ini, data deret waktu diuraikan ke dalam komponen-komponennya, seperti tren, musiman (seasonality), dan residual. Ini dilakukan untuk memahami pola mendasar dalam data.
7. Uji Stasioner
   Uji stasioner dilakukan untuk memeriksa apakah data memiliki sifat stasioner (mean dan varians konstan dari waktu ke waktu). Jika tidak stasioner, transformasi data akan diperlukan, seperti diferensiasi.
8. Plot ACF & PACF

Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) diplot untuk mengidentifikasi parameter dalam model ARIMA, seperti nilai p (autoregressive) dan q (moving average).

9. Estimasi Parameter ARIMA
   Berdasarkan analisis ACF dan PACF, parameter ARIMA dipilih (p, d, q), dan model diestimasi. ARIMA digunakan untuk menangkap pola dalam data deret waktu.
10. Uji Asumsi Residual
    Setelah model ARIMA dibuat, residual diuji untuk memastikan model sesuai. Residual yang baik harus berdistribusi normal dan tidak memiliki autocorrelation.
11. Forecast
    Model ARIMA digunakan untuk melakukan prediksi (forecasting) terhadap data masa depan. Prediksi ini didasarkan pada pola yang telah diidentifikasi dalam data.

Penulis menggunakan bahasa pemograman Python untuk menerapkan metode ARIMA (Autoregressive Integrated Moving Average) pada data kunjungan sekolah Telkomsel dari bulan Januari hingga Oktober 2024 dengan model ARIMA Menurut Wei (2006):

$$t\phi_p(B)(1-B)^d X_t = \theta_q(B)\epsilon_t$$

Dimana :

X_t : Nilai data pada waktu t (yang diprediksi)

$\phi_p(B)$: Koefesien AR dengan orde p, dimana $\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \cdots - \phi_p(B^p)$

$\theta_q(B)$: Koefesien MA dengan orde q, dimana $\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \cdots - \theta_q(B^q)$

$(1-B)^d$: Operator diferensiasi (lag), di mana d adalah orde diferensiasi untuk mencapai stasioneritas. B adalah operator lag dalam analisis deret waktu untuk menggambarkan pergeseran nilai data ke periode sebelumnya (Backward Shift), mendefinisikan $BX_t = X_{t-1}$

ϵ_t : Nilai residual pada waktu t

p: Orde autoregresif (AR).

d: Orde diferensiasi (differencing) untuk membuat data menjadi stasioner.

q: Orde moving average (MA)

Untuk mengidentifikasi stasioner dalam mean penulis menggunakan Uji Augmented Dickey-Fuller (ADF) dengan model sebagai berikut :

$$\Delta Z_t = \beta_1 + \beta_2 t + \delta Z_{t-1} + \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \cdots + \alpha_1 Z_{t-2} + \varepsilon_t$$

Dan untuk stasioner dalam varians penulis menggunakan transformasi Box-Cox

Berikut proses untuk memprediksi data kunjungan Telkomsel ke sekolah di Indonesia pada bulan Januari – Oktober 2024 menggunakan metode ARIMA dan bahasa pemograman Python:

1. Import Library

```
import numpy as np
import pandas as pd
```

```
from scipy import stats
from scipy.stats import zscore
from scipy.stats import probplot
from scipy.stats import shapiro
from scipy.stats import boxcox
import matplotlib.pyplot as plt
import seaborn as sns
from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.stats.diagnostic import acorr_ljungbox
from statsmodels.tsa.stattools import adfuller
from statsmodels.tsa.stattools import kpss
from sklearn.metrics import mean_squared_error, mean_absolute_error
```

Library yang disebutkan di atas mencakup berbagai alat untuk analisis data dan pemodelan statistik. numpy dan pandas digunakan untuk manipulasi data, sementara scipy menyediakan berbagai fungsi statistik, seperti uji distribusi normal (shapiro), uji kebebasan dari outlier (zscore), serta transformasi Box-Cox untuk normalisasi data. matplotlib dan seaborn digunakan untuk visualisasi data, memungkinkan pembuatan grafik seperti distribusi atau scatter plot. statsmodels menawarkan berbagai fungsi untuk analisis deret waktu, termasuk dekomposisi musiman (seasonal_decompose), pemodelan ARIMA (ARIMA), serta uji stasioneritas seperti ADF (adfuller) dan KPSS (kpss). Selain itu, acorr_ljungbox digunakan untuk uji autocorrelation, dan plot_acf/plot_pacf untuk memeriksa autocorrelation dan partial autocorrelation. Terakhir, sklearn.metrics menawarkan metrik evaluasi model seperti mean_squared_*error* dan mean_absolute_*error* untuk mengukur akurasi model prediktif.

2. Impor Data dan Mengekplorasi Data

   Data yang digunakan dalam analisis ini merupakan data primer yang dikumpulkan melalui situs web mynami.id dalam format CSV selama magang di Telkomsel Cabang Surabaya. Data tersebut mencakup informasi mengenai kunjungan Telkomsel ke sekolah di Indonesia mulai dari tingkat TK hingga Universitas, dengan jumlah 129.719 baris dan 24 kolom.

	PERIODE	PERIODE DATE	DAY NAME	AGENT ID	USER ID	USER NAME	USER TYPE	AREA	REGIONAL	BRANCH	...	LEVEL	PJP FREQUENCY	DISPOSED DAY	DISPOSED STATUS	PJP CHECKIN	PJP EXTRA	CHECKIN TIME	CHECKOUT TIME	DURATION (MINUTE)	CHECKIN VALID
0	202401	1/1/2024	SENIN	9903006180	6,29E+12	ICAH	BUDDIES by.U	JABOTABEK & JABAR	WESTERN JABOTABEK	SERANG	...	SMP	F2	SENIN	UNDISPOSED	Y	N	1/1/2024 15:31	1/1/2024 20:14	282	Y
1	202401	1/1/2024	SENIN	9903000696	6,28E+12	ALEXANDER GAT ASER LALU	DSS	JAWA - BALI - NUSRA	BALI NUSRA	FLORES	...	UNIVERSITAS	F2	SENIN	DISPOSED	Y	N	1/1/2024 18:57	1/1/2024 21:31	153	Y
2	202401	1/1/2024	SENIN	9903000696	6,28E+12	ALEXANDER GAT ASER LALU	DSS	JAWA - BALI - NUSRA	BALI NUSRA	FLORES	...	SMA	F2	SENIN	DISPOSED	Y	N	1/1/2024 18:15	1/1/2024 18:57	41	Y
3	202401	1/1/2024	SENIN	9903006667	6,28E+12	SURYADITO MAULANA ROSIHAN	DSS	JAWA - BALI - NUSRA	BALI NUSRA	MATARAM	...	SMA	F2	SENIN	DISPOSED	Y	N	1/1/2024 16:18	1/1/2024 18:10	112	Y
4	202401	1/1/2024	SENIN	9903006667	6,28E+12	SURYADITO MAULANA ROSIHAN	DSS	JAWA - BALI - NUSRA	BALI NUSRA	MATARAM	...	SMA	F2	SENIN	DISPOSED	Y	N	1/1/2024 20:08	1/1/2024 20:42	34	Y
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
129714	202410	10/9/2024	RABU	9903006629	6,29E+12	NORA ASRIANI	DSS	SUMATERA	SUMBAGUT	MEDAN	...	SMP	F2	RABU	DISPOSED	Y	N	10/9/2024 11:09	10/9/2024 11:50	40	Y
129715	202410	10/9/2024	RABU	9903006636	6,28E+12	NIA INDAH LESTARI NASUTION	DSS	SUMATERA	SUMBAGUT	MEDAN	...	SMA	F2	JUMAT	DISPOSED	Y	Y	10/9/2024 11:40	10/9/2024 12:12	32	N
129716	202410	10/9/2024	RABU	9903010496	6,28E+12	AYU LISTIANA	DSS	SUMATERA	SUMBAGUT	MEDAN	...	SMA	F2	RABU	DISPOSED	Y	Y	10/9/2024 7:36	10/9/2024 8:45	69	N
129717	202410	10/9/2024	RABU	9903009079	6,28E+12	CANDIRA PRASETYO	DSS	SUMATERA	SUMBAGUT	PADANG SIDEMPUAN	...	SMA	F2	RABU	DISPOSED	Y	N	10/9/2024 10:01	10/9/2024 10:13	12	N
129718	202410	10/9/2024	RABU	9903010429	6,28E+12	INDAH SEKON LEDI SINAGA	DSS	SUMATERA	SUMBAGUT	PADANG SIDEMPUAN	...	SMP	F2	RABU	DISPOSED	Y	N	10/9/2024 9:19	NaN	0	N

129719 rows × 24 columns

Gambar 4. 61 Data School

```
Index(['PERIODE', 'PERIODE DATE', 'DAY NAME', 'AGENT ID', 'USER ID',
       'USER NAME', 'USER TYPE', 'AREA', 'REGIONAL', 'BRANCH', 'CLUSTER',
       'CITY', 'NPSN', 'INSTITUTION NAME', 'LEVEL', 'PJP FREQUENCY',
       'DISPOSED DAY', 'DISPOSED STATUS', 'PJP CHECKIN', 'PJP EXTRA',
       'CHECKIN TIME', 'CHECKOUT TIME', 'DURATION (MINUTE)', 'CHECKIN VALID'],
      dtype='object')
```

Gambar 4. 62 Variabel Data School

Data kunjungan Telkomsel ke sekolah di Indonesia mencakup 24 kolom, antara lain: tanggal harian, nama hari, nomor agen, ID pengguna, nama pengguna Telkomsel, tipe pengguna, lokasi (area, regional, branch, cluster, kota), NPSN, nama institusi atau nama sekolah, level sekolah, frekuensi PJP (Pelanggan Jasa Pemasaran) atau frekuensi kunjungan, status disposed, interaksi kunjungan Telkomsel ke sekolah, waktu check-in, waktu check-out, total durasi kunjungan dalam satuan menit, dan status validasi check-in agen.

```
PERIODE               int64
PERIODE DATE         object
DAY NAME             object
AGENT ID              int64
USER ID              object
USER NAME            object
USER TYPE            object
AREA                 object
REGIONAL             object
BRANCH               object
CLUSTER              object
CITY                 object
NPSN                 object
INSTITUTION NAME     object
LEVEL                object
PJP FREQUENCY        object
DISPOSED DAY         object
DISPOSED STATUS      object
PJP CHECKIN          object
PJP EXTRA            object
CHECKIN TIME         object
CHECKOUT TIME        object
DURATION (MINUTE)     int64
CHECKIN VALID        object
dtype: object
```

Gambar 4. 63 Tipe Data School

Data ini merupakan hasil dari tipe data kunjungan Telkomsel ke sekolah di Indonesia, mencakup informasi terkait waktu kunjungan, lokasi, interaksi agen, serta status validasi aktivitas agen selama kunjungan. Data ini memberikan gambaran komprehensif tentang frekuensi dan durasi kunjungan, tipe pengguna, serta detail terkait dengan sekolah yang dikunjungi, mulai dari tingkat sekolah hingga status kunjungan yang tercatat.

e. Pivot Table

LEVEL	PERIODE DATE	KURSUS	PESANTREN	PKBM	SD	SKB	SLB	SMA	SMP	SPK	TK	UNIVERSITAS	Total Kunjungan (MINUTE)
0	2024-01-01	0	0	0	0	0	0	67	179	0	0	227	664
1	2024-01-02	54	0	0	92	0	63	73	80	0	54	124	665
2	2024-01-03	0	0	0	81	0	71	88	64	0	39	92	435
3	2024-01-04	66	0	0	73	0	0	85	75	0	128	91	518
4	2024-01-05	0	0	0	73	0	33	83	91	0	145	106	531
...	...	...	...	...	...	...	...	...	...	...	...	...	...
269	2024-10-04	0	0	0	60	0	135	65	54	0	206	56	576
270	2024-10-05	0	0	0	61	0	0	60	69	0	170	135	495
271	2024-10-07	0	0	0	62	0	80	60	53	0	103	45	403
272	2024-10-08	0	0	0	52	0	60	57	54	0	98	74	764
273	2024-10-09	0	23	0	41	0	0	35	40	0	57	37	233

274 rows × 13 columns

Gambar 4. 64 DataFrame

f. Group Data

```
# Mengelompokkan data berdasarkan 'PERIODE DATE' dan menjumlahkan
nilai 'DURATION (MINUTE)'
grouped_dat = pivot_data_filled.groupby('PERIODE DATE').sum()

# Menambahkan kolom 'Total Kunjungan (MINUTE)' yang merupakan
jumlah dari seluruh interaksi untuk setiap tanggal
grouped_dat['Total Kunjungan (MINUTE)'] = grouped_dat.sum(axis=1)
```

```
# Mengubah tipe data 'Total Kunjungan (MINUTE)' menjadi Int64
grouped_dat['Total Kunjungan (MINUTE)'] = grouped_dat['Total
Kunjungan (MINUTE)'].astype('Int64')

# Mengubah 'PERIODE DATE' dari indeks menjadi kolom biasa
grouped_dat.reset_index(inplace=True)
```

Kode di atas bertujuan untuk mengelompokkan data berdasarkan tanggal pada kolom 'PERIODE DATE' dan menambahkan kolom baru, yaitu 'Total Kunjungan (MINUTE)', yang menghitung total durasi kunjungan dalam satuan menit untuk setiap tanggal atau hari. Selain itu, kode ini juga mengubah tipe data dan format DataFrame untuk memudahkan analisis lebih lanjut, sehingga data dapat diproses secara efisien dalam analisis berbasis waktu

g. Mengecek Outlier

```
# Mengecek Outlier
# Menghitung IQR untuk setiap kolom
Q1 = grouped_dat.quantile(0.25)
Q3 = grouped_dat.quantile(0.75)
IQR = Q3 - Q1

# Menentukan batas bawah dan atas untuk deteksi outlier
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

# Menandai outlier berdasarkan IQR
outliers_iqr = (grouped_dat < lower_bound) | (grouped_dat >
upper_bound)
print(outliers_iqr.sum())
```

Elemen-elemen yang berada di luar batas bawah atau batas atas ini akan ditandai sebagai outlier pada kolom 'outliers_iqr'. Dan juga kode ini juga menampilkan jumlah outlier yang ditemukan untuk setiap kolom dalam dataset

```
LEVEL
PERIODE DATE                0
KURSUS                     59
PESANTREN                  20
PKBM                        5
SD                         40
SKB                        23
SLB                        19
SMA                        24
SMP                        27
SPK                        20
TK                          4
UNIVERSITAS                 9
Total Kunjungan (MINUTE)   17
```

Gambar 4. 65 Hasil Outlier

h. Menangani Outlier

```
#  MENANGANI OUTLIER
grouped_data = grouped_dat.copy()

# Pilih hanya kolom numerik
numeric_columns =
grouped_data.select_dtypes(include=[np.number]).columns
```

```
# Hitung Z-score hanya untuk kolom numerik
z_scores = grouped_data[numeric_columns].apply(zscore)

# Membatasi Z-score lebih dari 3 atau kurang dari -3
grouped_data[numeric_columns] =
grouped_data[numeric_columns].where(abs(z_scores) <= 3, np.nan)

# Ganti NaN dengan 0 pada kolom numerik
grouped_data[numeric_columns] =
grouped_data[numeric_columns].fillna(0)

# Ubah tipe data menjadi int64 untuk kolom numerik
grouped_data[numeric_columns] =
grouped_data[numeric_columns].astype('int64')

# Hitung Z-score setelah penanganan
z_scores_after = grouped_data[numeric_columns].apply(zscore)

# Menandai outlier setelah penanganan
outliers_after = abs(z_scores_after) > 3

# Menghitung jumlah outlier setelah penanganan untuk setiap kolom
jumlah_outliers_after_per_column = outliers_after.sum()
```

Kode diatas bertujuan untuk menangani outlier menggunakan metode Z-score. Untuk menghitung dan memanipulasi Z-score dataset. Outlier yang memiliki Z-score lebih besar dari 3 atau kurang dari -3 diganti dengan NaN, yang menandakan bahwa nilai tersebut adalah outlier, kemudian diisi dengan nilai 0. Ini merupakan salah satu cara untuk menangani outlier, terutama jika data tersebut tidak terlalu besar atau signifikan secara statistik dan mengubah tipe data kolom numerik menjadi ‘int64’ untuk memastikan bahwa data tetap dalam format yang sesuai.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 274 entries, 0 to 273
Data columns (total 13 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   PERIODE DATE              274 non-null    datetime64[ns]
 1   KURSUS                    274 non-null    int64
 2   PESANTREN                 274 non-null    int64
 3   PKBM                      274 non-null    int64
 4   SD                        274 non-null    int64
 5   SKB                       274 non-null    int64
 6   SLB                       274 non-null    int64
 7   SMA                       274 non-null    int64
 8   SMP                       274 non-null    int64
 9   SPK                       274 non-null    int64
 10  TK                        274 non-null    int64
 11  UNIVERSITAS               274 non-null    int64
 12  Total Kunjungan (MINUTE)  274 non-null    int64
dtypes: datetime64[ns](1), int64(12)
memory usage: 28.0 KB
```

Gambar 4. 66 Info Data Schools

Informasi yang diperoleh dari variabel 'grouped_data', dapat disimpulkan bahwa tipe data yang dimiliki oleh kolom-kolom dalam dataset tersebut sesuai dengan yang diinginkan oleh penulis, yaitu tipe data datetime untuk kolom yang berkaitan dengan waktu, dan tipe data integer untuk kolom yang berisi nilai numerik seperti Total Kunjungan (MINUTE).

LEVEL	PERIODE DATE	KURSUS	PESANTREN	PKBM	SD	SKB	SLB	SMA	SMP	SPK	TK	UNIVERSITAS	Total Kunjungan (MINUTE)
count	274	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000	274.000000
mean	2024-05-19 03:35:28.467153408	10.281022	1.948905	0.368613	62.014599	2.281022	29.755474	74.394161	72.094891	2.554745	59.930657	93.441606	458.317518
min	2024-01-01 00:00:00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	2024-03-10 06:00:00	0.000000	0.000000	0.000000	55.000000	0.000000	0.000000	58.000000	54.000000	0.000000	0.000000	55.000000	359.250000
50%	2024-05-17 12:00:00	0.000000	0.000000	0.000000	66.500000	0.000000	0.000000	78.000000	75.000000	0.000000	53.000000	98.000000	456.500000
75%	2024-07-28 12:00:00	0.000000	0.000000	0.000000	78.000000	0.000000	44.000000	89.000000	87.000000	0.000000	95.000000	121.000000	551.000000
max	2024-10-09 00:00:00	134.000000	79.000000	67.000000	135.000000	103.000000	233.000000	197.000000	187.000000	108.000000	238.000000	311.000000	1000.000000
std	NaN	25.493759	9.383815	4.532249	28.144313	11.830203	43.021105	29.357680	32.323884	13.371513	57.186082	50.852209	173.832910

Gambar 4. 67 Hasil Statistika Deskripsi

Didapatkan kolom total kunjungan rata -rata adalah 458,32 menit, dengan nilai minimum 0 menit dan maksimum 1.000 menit. Hal ini menunjukkan adanya variasi besar dalam durasi kunjungan ke sekolah yang ada di Indonesia. Median kunjungan sekolah adalah 456,5 menit menunjukkan bahwa setengah dari kunjungan memiliki durasi dibawah angka tersebut.

Kategori dengan rata – rata jumlah kunjungan tertinggi adalah Universitas dengan nilai 93,44, lalu SMP dengan nilai 72,09 dan SMA dengan nilai 74,39. Menunjukkan bahwa institusi pendidikan menengah dan tinggi cenderung menjadi prioritas utama. Untuk kategori seperti Kursus dengan nilai 10,28 dan PKBM (Pusat Kegiatan Belajar Masyarakat) dengan nilai 0,37 memiliki rata – rata yang jauh lebih rendah, menujukkan bahwa tingkat fokus yang lebih rendah pada institusi nonformal. Banyak kategori yang memiliki nilai minimum 0, yang menunjukkan beberapa institusi tidak mendapatkan kunjungan sama sekali selama bulan januari

Untuk periode rentang waktu yang digunakan adalah bulan Januari hingga Oktober 2024. Nilai median menunjukkan bahwa sebagian besar kunjungan terjadi di sekitar bulan Mei hingga Juli 2024

```
display(grouped_data)
```

LEVEL	PERIODE DATE	KURSUS	PESANTREN	PKBM	SD	SKB	SLB	SMA	SMP	SPK	TK	UNIVERSITAS	Total Kunjungan (MINUTE)
0	2024-01-01	0	0	0	0	0	0	67	179	0	0	227	664
1	2024-01-02	54	0	0	92	0	63	73	80	0	54	124	665
2	2024-01-03	0	0	0	81	0	71	88	64	0	39	92	435
3	2024-01-04	66	0	0	73	0	0	85	75	0	128	91	518
4	2024-01-05	0	0	0	73	0	33	83	91	0	145	106	531
...	...	...	...	...	...	...	...	...	...	...	...	...	...
269	2024-10-04	0	0	0	60	0	135	65	54	0	206	56	576
270	2024-10-05	0	0	0	61	0	0	60	69	0	170	135	495
271	2024-10-07	0	0	0	62	0	80	60	53	0	103	45	403
272	2024-10-08	0	0	0	52	0	60	57	54	0	98	74	764
273	2024-10-09	0	23	0	41	0	0	35	40	0	57	37	233

274 rows × 13 columns

Gambar 4. 68. Pivot Tabel

i. Mengecek Nilai NaN

```
LEVEL
PERIODE DATE                0
KURSUS                      0
PESANTREN                   0
PKBM                        0
SD                          0
SKB                         0
SLB                         0
SMA                         0
SMP                         0
SPK                         0
TK                          0
UNIVERSITAS                 0
Total Kunjungan (MINUTE)    0
```

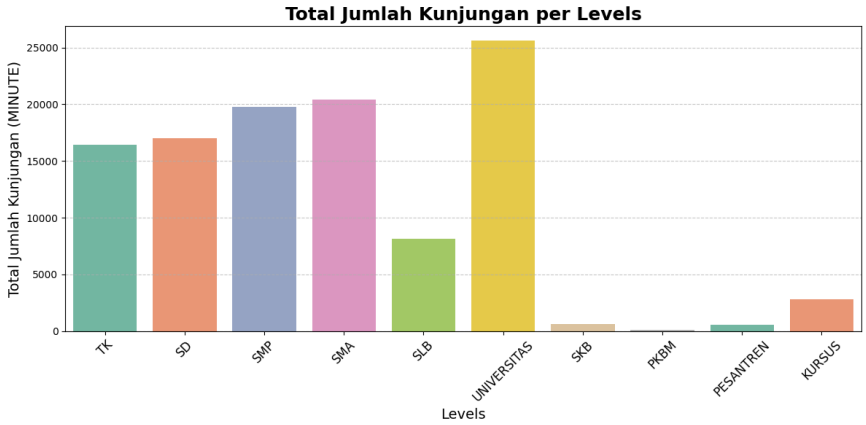
Dapat disimpulkan bahwa tidak ada nilai NaN pada setiap kolom

Gambar 4. 69 Jumlah Missing Value

j. Grafik Total Jumlah Interaksi Sekolah Pada Setiap Level Sekolah

```
# Daftar level yang tersedia dalam data
levels = ['TK', 'SD', 'SMP', 'SMA', 'SLB','UNIVERSITAS', 'SKB',
'PKBM', 'PESANTREN', 'KURSUS']

# Menghitung total jumlah interaksi per Levels
total_kunjungan = grouped_data[levels].sum()
print(total_kunjungan)

plt.figure(figsize=(12, 6))
sns.barplot(x=total_kunjungan.index, y=total_kunjungan.values,
palette='Set2')
plt.title('Total Jumlah Kunjungan per Levels', fontsize=18,
fontweight='bold')
plt.xlabel('Levels', fontsize=14)
plt.ylabel('Total Jumlah Kunjungan (MINUTE)', fontsize=14)
plt.xticks(rotation=45, fontsize=12)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

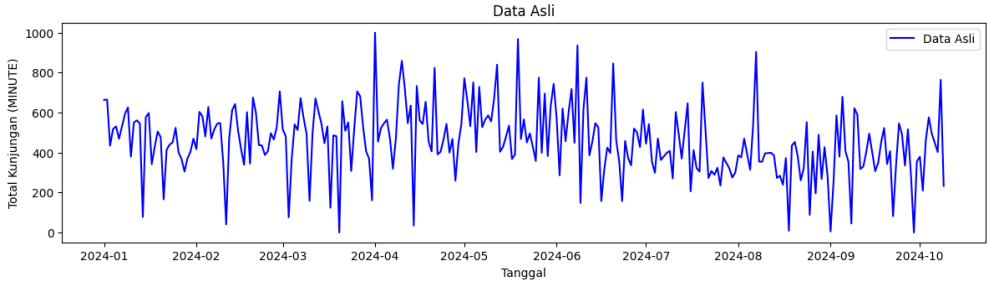
Kode di atas bertujuan untuk menganalisis dan memvisualisasikan tren interaksi Telkomsel di berbagai jenis institusi pendidikan berdasarkan levelnya, seperti TK, SD, SMP, SMA, hingga universitas dan level atau tingkat lainnya. Dengan menggunakan library matplotlib.pyplot untuk menyesuaikan detail visualisasi dan library seaborn untuk menghasilkan visualisasi statistik yang estetis. Kode ‘grouped_data[levels].sum()’ menghitung total jumlah interaksi untuk setiap level pendidikan dengan cara menjumlahkan seluruh nilai pada kolom yang sesuai dalam dataset ‘grouped data’. Hasil ini disimpan dalam variabel ‘total kunjungan’. Output seperti pada Gambar ……



Gambar 4. 70 Grafik Total Jumlah Kunjungan

Dapat disimpulkan bahwa total durasi kunjungan tertinggi tercatat pada tingkat Universitas, yaitu mencapai 25.603 menit, sedangkan total durasi kunjungan terendah terdapat pada tingkat PKBM (Pusat Kegiatan Belajar Masyarakat), dengan durasi sebesar 101 menit.

k. Visualisasi Grafik Untuk Data Asli



Gambar 4. 71 Grafik Data

Dapat disimpulkan bahwa menunjukkan fluktuasi yang signifikan dalam durasi total kunjungan (minute) dari waktu ke waktu, dengan rentang nilai berkisar antara 0 hingga sekitar 1000 menit. Grafik ini memiliki variabilitas yang cukup tinggi, ditandai dengan lonjakan dan penurunan tajam di beberapa titik, namun tidak tampak adanya pola musiman atau tren yang konsisten.

l. Uji Stasioner
Dalam Mean (Augmented Dickey-Fuller) dan varians (KPSS)

Didapatkan bahwa Hasil Augmented Dickey-Fuller Test:
ADF Statistic: -3.812883434768411
p-value: 0.0027782171924074766
Critical Values:
  1%: -3.4550813975770827
  5%: -2.8724265892710914
  10%: -2.5725712007462582

Data stasioner terhadap mean (H0 ditolak). Dan
Hasil KPSS Test:
KPSS Statistic: 1.4413282633577837
p-value: 0.01

Critical Values:
10%: 0.347
5%: 0.463
2.5%: 0.574
1%: 0.739

H0 ditolak: Data tidak stasioner terhadap varians.

m. Transformasi Box-Cox

```
# ==========================================
# MENANGANI DATA
# ==========================================
# Pastikan data memiliki indeks waktu
grouped_data['PERIODE DATE'] =
pd.to_datetime(grouped_data['PERIODE DATE'])
time_series_data = grouped_data.set_index('PERIODE DATE')['Total
Kunjungan (MINUTE)']

# ==========================================
# BOX-COX TRANSFORMATION
# ==========================================
# Fungsi untuk transformasi Box-Cox
def box_cox_transformation(timeseries):
    """
    Fungsi untuk melakukan transformasi Box-Cox pada deret waktu.

    Parameters:
    timeseries: pd.Series - Deret waktu yang akan
ditransformasikan

    Returns:
    transformed_data: pd.Series - Data yang sudah
ditransformasikan
    lambda_ : float - Nilai lambda yang digunakan dalam
transformasi Box-Cox
    """
    # Menghindari nilai negatif atau nol, karena Box-Cox hanya
bekerja pada data positif
    timeseries = timeseries + 1  # Menambahkan 1 untuk memastikan
data positif

    # Melakukan transformasi Box-Cox
    transformed_data, lambda_ = stats.boxcox(timeseries.dropna())

    # Kembalikan data yang sudah ditransformasi dan nilai lambda
    return pd.Series(transformed_data,
index=timeseries.dropna().index), lambda_

# Melakukan transformasi Box-Cox
transformed_data, lambda_ =
box_cox_transformation(time_series_data)

# Menampilkan hasil transformasi Box-Cox
print(f"Nilai Lambda untuk Transformasi Box-Cox: {lambda_}")
```

```
# ==========================================
# KESIMPULAN STASIONER VARIANS
# ==========================================
# Menghitung standar deviasi sebelum dan sesudah transformasi
std_before = time_series_data.std()
std_after = np.std(transformed_data)

# Membandingkan standar deviasi sebelum dan sesudah transformasi
print("\nKesimpulan Stasioner Varians:")
print(f"Standar Deviasi Sebelum Transformasi: {std_before:.2f}")
print(f"Standar Deviasi Setelah Transformasi: {std_after:.2f}")

if std_after < std_before:
    print("\nSetelah transformasi Box-Cox, varians data lebih
stabil. Data cenderung stasioner terhadap varians.")
else:
    print("\nTransformasi Box-Cox tidak berhasil menstabilkan
varians. Data mungkin tidak stasioner terhadap varians.")

# ==========================================
# PLOT DISTRIBUSI SEBELUM DAN SESUDAH BOX-COX
# ==========================================
plt.figure(figsize=(12, 6))

# Plot distribusi sebelum transformasi
plt.subplot(1, 2, 1)
sns.histplot(time_series_data + 1, kde=True, bins=30,
color='blue', alpha=0.7)  # Data asli dengan tambahan 1
plt.title("Distribusi Sebelum Transformasi Box-Cox")
plt.xlabel("Total Kunjungan (MINUTE)")
plt.ylabel("Frekuensi")

# Plot distribusi setelah transformasi Box-Cox
plt.subplot(1, 2, 2)
sns.histplot(transformed_data, kde=True, bins=30, color='green',
alpha=0.7)
plt.title("Distribusi Setelah Transformasi Box-Cox")
plt.xlabel("Total Kunjungan (Transformed)")
plt.ylabel("Frekuensi")

plt.tight_layout()
plt.show()

# ==========================================
# PLOT TIME SERIES SEBELUM DAN SESUDAH BOX-COX
# ==========================================
plt.figure(figsize=(12, 6))

# Plot data asli
plt.subplot(2, 1, 1)
plt.plot(time_series_data, label='Data Asli', color='blue')
plt.title('Data Asli')
plt.legend()

# Plot data setelah transformasi Box-Cox
plt.subplot(2, 1, 2)
```

```
plt.plot(transformed_data, label='Data Setelah Transformasi Box-Cox', color='orange')
plt.title('Data Setelah Transformasi Box-Cox')
plt.legend()

plt.tight_layout()
plt.show()

# Mengembalikan data hasil transformasi Box-Cox (tanpa periode tanggal)
transformed_data.reset_index(drop=True, inplace=True)
transformed_data
```

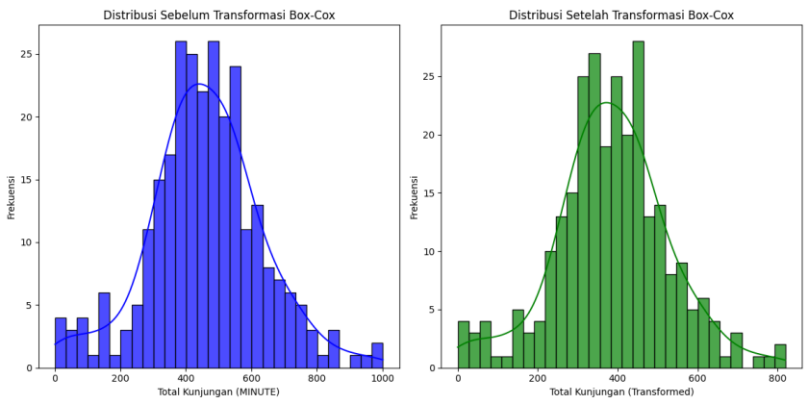
Output :

Didapatkan bahwa data stasioner dalam varians

```
Kesimpulan Stasioner Varians:
Standar Deviasi Sebelum Transformasi: 173.83
Standar Deviasi Setelah Transformasi: 141.80
```

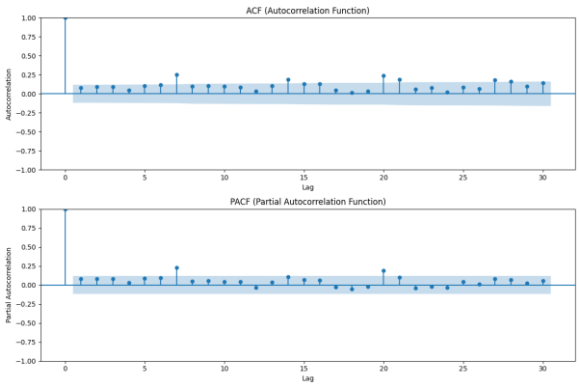
Gambar 4. 72 Hasil Uji Stasioner

Setelah transformasi Box-Cox, varians data lebih stabil. Data cenderung stasioner terhadap varians.



Gambar 4. 73 Distribusi Sebelum dan Sesudah Transformasi Box-Cox

n. Autokorelasi (ACF) dan Autokorelasi Parsial (PACF)



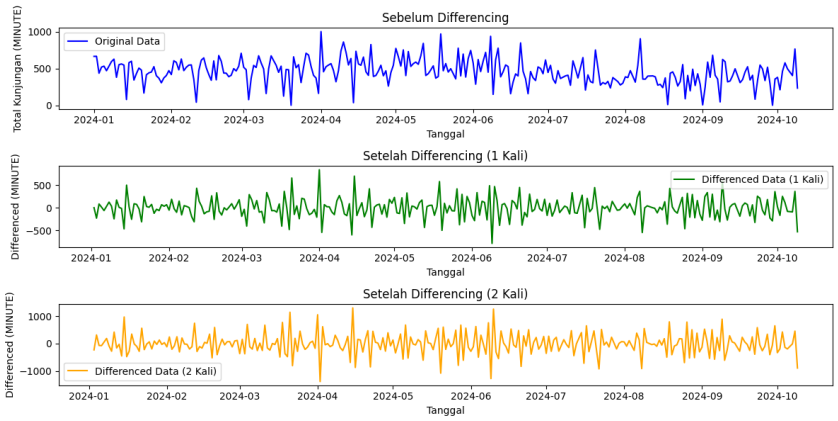
Gambar 4. 74 Plot acf dan pacf

Karena tidak terlihat pola cut off yang tidak berpola maka penulis ingin melakukan differencing sebanyak 2 kali. Tujuannya yaitu ingin melihat cut off pada lag yang lebih jelas.

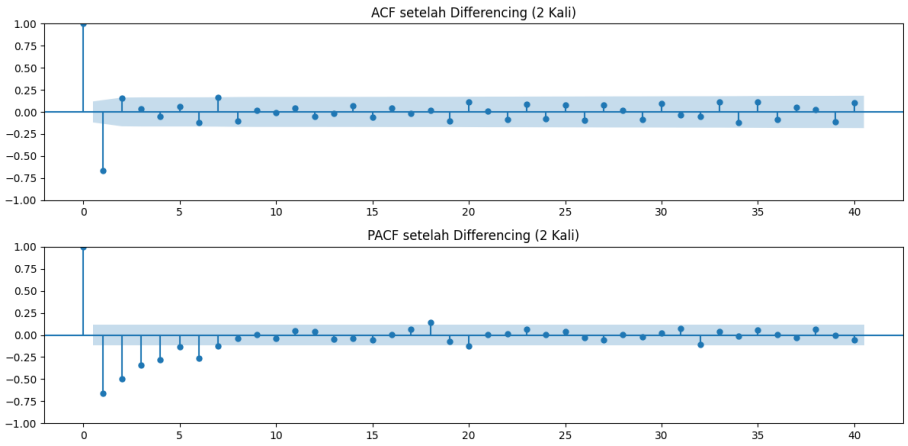
o. Melakukan Differencing

Alasan penulis melakukan differencing sebanyak dua kali meskipun data telah dinyatakan stasioner berdasarkan hasil uji Augmented Dickey-Fuller (ADF) adalah untuk memastikan identifikasi cut-off lag pada plot Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) menjadi lebih jelas. Dengan melakukan differencing tambahan, penulis dapat lebih mudah mengidentifikasi pola lag signifikan yang diperlukan untuk menentukan parameter p dan q dalam model ARIMA. Langkah ini bertujuan untuk memastikan pemilihan parameter model yang optimal, sehingga dapat meningkatkan akurasi dan interpretasi model time series yang dihasilkan.

Output :



Gambar 4. 75 Grafik Differencing 2 Kali



Gambar 4. 76 Plot PACF dan ACF Setelah di Differencing 2 kali

Didapatkan hasil model ARIMA(p,d,q) adalah (6,2,1). Hal ini berarti model menggunakan enam lag autoregressive (p=6), dua kali differencing (d=2), dan satu lag moving average (q=1). Berikut adalah model persamaan ARIMA(6,2,1)

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4 - \phi_5 B^5 - \phi_6 B^6)(1 - B)^2 Y_t = \theta_1 B_{\epsilon_t}$$

Penjelasan :

- Terdiri dari 6 lag.
- Menunjukkan dua kali differencing pada data, yang dapat menstabilkan variansi dan tren.
- Terdiri dari 1 lag *error*

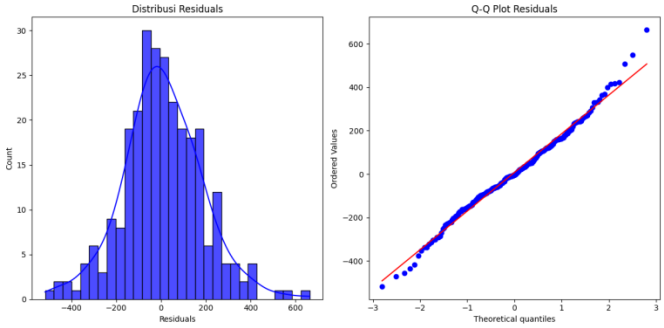
p. Estimasi Model ARIMA

Model ARIMA(6,2,1) yang diestimasi menggunakan metode Maximum Likelihood Estimation (MLE) dengan jumlah observasi sebanyak 274 menunjukkan kinerja yang baik berdasarkan kriteria evaluasi model, yaitu Log Likelihood (-1788.451), AIC (3592.922), BIC (3621.749), dan HQIC (3604.483), di mana nilai AIC lebih kecil dibandingkan BIC dan HQIC, dengan semua parameter AR (L1 hingga L6) dan MA signifikan secara statistik (P-value < 0.05) yang mencerminkan relevansi dalam menjelaskan pola data dengan hubungan langsung atau invers terhadap data masa lalu, sementara parameter sigma2 (*error* variance) menggambarkan variabilitas yang tidak dapat dijelaskan oleh model, didukung oleh hasil Ljung-Box Test (Prob(Q) = 0.55) yang menunjukkan tidak adanya autokorelasi signifikan pada residual dan Jarque-Bera Test (Prob(JB) = 0.12) yang menunjukkan bahwa residual mendekati distribusi normal, sehingga model ini dapat dianggap valid untuk analisis dan peramalan.

$$(1 - (-0{,}8)B - (-0{,}7)B^2 - (-0{,}6)B^3 - (-0{,}5)B^4 - (-0{,}4)B^5 - (-0{,}2)B^6)(1 - B)^2 Y_t = (-1)B_{\epsilon_t}$$

```
Ringkasan Model Terbaik:
                                     SARIMAX Results
==========================================================================================
Dep. Variable:     Total Kunjungan (MINUTE)   No. Observations:                  274
Model:                        ARIMA(6, 2, 1)   Log Likelihood               -1788.451
Date:                       Tue, 31 Dec 2024   AIC                           3592.902
Time:                               11:59:16   BIC                           3621.749
Sample:                                    0   HQIC                          3604.483
                                       - 274
Covariance Type:                         opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8966      0.061    -14.769      0.000      -1.016      -0.778
ar.L2         -0.7797      0.083     -9.387      0.000      -0.942      -0.617
ar.L3         -0.6475      0.088     -7.394      0.000      -0.819      -0.476
ar.L4         -0.5759      0.087     -6.630      0.000      -0.746      -0.406
ar.L5         -0.4419      0.073     -6.046      0.000      -0.585      -0.299
ar.L6         -0.2963      0.048     -6.164      0.000      -0.390      -0.202
ma.L1         -1.0000      0.077    -12.934      0.000      -1.152      -0.848
sigma2      2.903e+04   2.66e-06   1.09e+10      0.000     2.9e+04     2.9e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.35   Jarque-Bera (JB):                 4.31
Prob(Q):                              0.55   Prob(JB):                         0.12
...
```

Gambar 4. 77 Hasil Estimasi Parameter



Gambar 4. 78 Distribusi dan Q-Q Plot Residual

Distribusi data terlihat lebih simetris dan mendekati distribusi normal. Transformasi Box-Cox telah berhasil mengurangi kemiringan data. Puncak distribusi masih berada di sekitar nilai yang sama, tetapi penyebaran data terlihat lebih terkonsentrasi di sekitar rata-rata. Rentang data juga sedikit menyempit, terutama di sisi kanan, karena transformasi telah mengurangi pengaruh nilai-nilai ekstrem.

q. Uji Asumsi Residual

```
# ==========================================
# UJI ASUMSI RESIDUAL
# ==========================================

# Pastikan residuals diambil dari model terbaik yang sudah
dipelajari sebelumnya
residuals = final_results.resid

# 1. Uji Ljung-Box untuk menguji apakah residual memiliki
autocorrelation
ljung_box_test = acorr_ljungbox(residuals, lags=[10],
return_df=True)
print("\nHasil Uji Ljung-Box (Autocorrelation):")
print(ljung_box_test)

# Ambil p-value dari hasil uji Ljung-Box untuk lag 10
```

```
ljung_box_p_value = ljung_box_test['lb_pvalue'].iloc[0]  #
Mengakses p-value untuk lag 10

# 2. Uji Normalitas dengan Shapiro-Wilk Test
shapiro_stat, shapiro_p_value = shapiro(residuals)
print("\nHasil Uji Normalitas (Shapiro-Wilk Test):")
print(f"Statistik Uji: {shapiro_stat}, p-value:
{shapiro_p_value}")

# ==========================================
# KESIMPULAN UJI ASUMSI RESIDUAL
# ==========================================

# 1. Kesimpulan Uji Ljung-Box (Autocorrelation)
if ljung_box_p_value > 0.05:
    print("\nUji Ljung-Box: Tidak ada autocorrelation yang
signifikan pada residuals. P-value:", ljung_box_p_value)
else:
    print("\nUji Ljung-Box: Terdapat autocorrelation yang
signifikan pada residuals. P-value:", ljung_box_p_value)

# 2. Kesimpulan Uji Normalitas (Shapiro-Wilk Test)
if shapiro_p_value > 0.05:
    print("\nUji Normalitas (Shapiro-Wilk): Residuals
terdistribusi normal. P-value:", shapiro_p_value)
else:
    print("\nUji Normalitas (Shapiro-Wilk): Residuals tidak
terdistribusi normal. P-value:", shapiro_p_value)

# 3. Kesimpulan Visualisasi
# Histogram dan Q-Q plot digunakan untuk mengevaluasi visual
residuals
if shapiro_p_value > 0.05 and ljung_box_p_value > 0.05:
    print("\nKesimpulan: Model ARIMA memenuhi asumsi dasar yaitu
residuals tidak menunjukkan autocorrelation dan terdistribusi
normal.")
else:
    print("\nKesimpulan: Model ARIMA perlu dievaluasi lebih
lanjut karena asumsi dasar tidak sepenuhnya terpenuhi.")
```

Tujuan kode di atas adalah untuk menguji apakah residual dari model ARIMA terbaik memenuhi asumsi-asumsi dasar dalam analisis time series. Pengujian ini penting untuk memastikan validitas model dalam merepresentasikan data dan menghasilkan prediksi yang akurat.

Output :

```
Hasil Uji Ljung-Box (Autocorrelation):
     lb_stat  lb_pvalue
10  7.224382   0.704106

Hasil Uji Normalitas (Shapiro-Wilk Test):
Statistik Uji: 0.9903039932250977, p-value: 0.06577419489622116
```

Gambar 4. 79. Hasil Uji Ljung-Box dan Uji Normalitas

Uji Ljung-Box: Tidak ada autocorrelation yang signifikan pada residuals. P-value: 0.7041056914966717

Uji Normalitas (Shapiro-Wilk): Residuals terdistribusi normal. P-value: 0.06577419489622116

Kesimpulan: Model ARIMA memenuhi asumsi dasar yaitu residuals tidak menunjukkan autocorrelation dan terdistribusi normal.

r. Evaluasi Model

```
# ========================================
# Persiapan Data
# ========================================
# Menggunakan data setelah differencing 2 kali
data_diff_2 = grouped_data_diff_2[['PERIODE DATE',
'Differenced_2']]

# Menentukan 'PERIODE DATE' sebagai indeks
data_diff_2['PERIODE DATE'] = pd.to_datetime(data_diff_2['PERIODE
DATE'])  # Pastikan kolom 'PERIODE DATE' adalah datetime
data_diff_2.set_index('PERIODE DATE', inplace=True)

# Membagi data menjadi pelatihan dan pengujian (80% pelatihan,
20% pengujian)
train_size = int(len(data_diff_2) * 0.8)
y_train, y_test = data_diff_2['Differenced_2'][:train_size],
data_diff_2['Differenced_2'][train_size:]

# ========================================
# Melatih Model ARIMA(6, 2, 1)
# ========================================
model = ARIMA(y_train, order=(6, 2, 1))
model_fit = model.fit()

# ========================================
# Prediksi dan Evaluasi Model
# ========================================
# Prediksi menggunakan model yang sudah dilatih
pred = model_fit.predict(start=len(y_train), end=len(y_train) +
len(y_test) - 1, dynamic=False)

# Menghitung RMSE dan MAE
rmse = np.sqrt(mean_squared_error(y_test, pred))
mae = mean_absolute_error(y_test, pred)

# ========================================
# Menghitung MAPE (menghindari pembagian dengan 0)
# ========================================
# Menangani nilai Actual = 0 untuk menghindari MAPE = infinity
combined = pd.DataFrame({'Actual': y_test, 'Predicted': pred})

# Menghindari pembagian dengan 0 untuk MAPE
combined = combined[combined['Actual'] != 0]  # Menghapus baris
dengan nilai Actual = 0
```

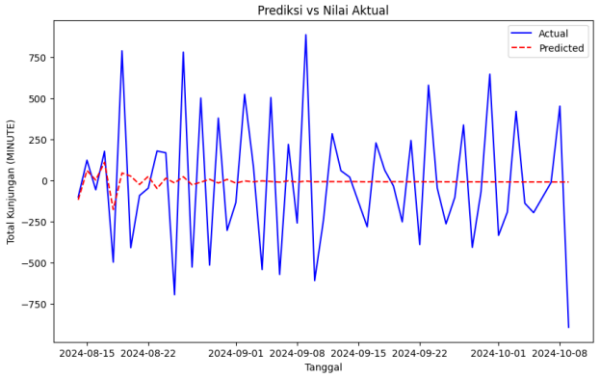
```
# Menghitung MAPE hanya jika data masih ada
if not combined.empty:
    mape = np.mean(np.abs((combined['Actual'] -
combined['Predicted']) / combined['Actual'])) * 100
    print(f"Model ARIMA(6, 2, 1):")
    print(f"  RMSE: {rmse}")
    print(f"  MAE: {mae}")
    print(f"  MAPE: {mape:.2f}%\n")
else:
    print("Tidak ada data yang valid untuk perhitungan MAPE
(nilai Actual = 0 ditemukan).")
```

Kode diatas bertujuan untuk membangun pipeline sederhana dalam analisis time series menggunakan model ARIMA terbaik yaitu (6,2,1). Dengan melakukan persiapan data ulang, pelatihan model, evaluasi model menggunakan MAPE, RMSE, dan MAE, dan visualisasi, kode ini memberikan Gambaran menyeluruh tentang seberapa baik model ARIMA dapat merepresentasikan pola data dan melakukan prediksi masa depan.

Output :

Tabel 4. 9 Evaluasi Model

Model	MAE	RMSE	MAPE
ARIMA (6,2,1)	138.603	178.875	143.60%



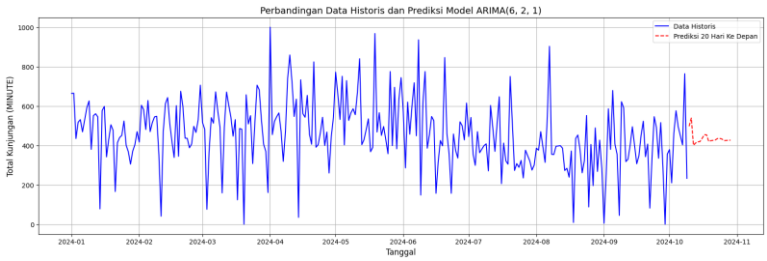
Gambar 4. 80 Grafik Prediksi dan Data Aktual

Bahwa prediksi ini hanya memberikan gambaran berdasarkan data historis dan tren yang ditangkap oleh model. Jika ada faktor eksternal atau kejadian mendadak (seperti promosi besar-besaran, kondisi cuaca, atau perubahan kebijakan) yang tidak tercermin dalam data historis, hasil prediksi mungkin tidak sepenuhnya akurat. Oleh karena itu, selain menggunakan hasil prediksi ini, penulis juga mencantumkan kondisi aktual dan mempertimbangkan faktor-faktor eksternal yang relevan untuk mendukung keputusan yang lebih tepat. Didapatkan grafik untuk data aktual memberikan yang menunjukkan variasi yang tajam, dengan puncak kunjungan yang tinggi pada beberapa tanggal tertentu, diselingi oleh penurunan yang cukup drastis.

s. Plot dan Nilai Forecast

Tabel 4. 10 Tabel Prediksi 20 Hari Kedepan

Tanggal	Prediksi
2024-10-10	497.572319
2024-10-11	539.057898
2024-10-12	400.864929
2024-10-13	412.701807
2024-10-14	418.038390
2024-10-15	419.771158
2024-10-16	436.581762
2024-10-17	454.783585
2024-10-18	453.795988
2024-10-19	420.282617
2024-10-20	424.312649
2024-10-21	425.977893
2024-10-22	427.803285
2024-10-23	434.967007
2024-10-24	435.806465
2024-10-25	432.718216
2024-10-26	424.587975
2024-10-27	425.461612
2024-10-28	426.162324
2024-10-29	426.768019



Gambar 4. 81 Grafik Hasil Prediksi

Grafik di atas menggambarkan prediksi jumlah kunjungan untuk 20 hari ke depan, namun, pada bagian peramalan (ditandai dengan garis merah), terlihat bahwa pola prediksi lebih teratur dan tidak menunjukkan fluktuasi ekstrem seperti data historis. Pola ini mencerminkan upaya model prediksi untuk menangkap tren jangka panjang atau rata-rata dari data historis, sementara variabilitas harian yang bersifat acak atau tidak terduga diabaikan. Dengan kata lain, model peramalan memberikan gambaran stabil yang dapat mempermudah dalam memahami pola umum jumlah kunjungan.

4.8. Rekognisi Mata Kuliah Data Storytelling

Pada Magang Mandiri di Telkomsel Branch Surabaya terdapat proyek yang relevan dengan konversi Mata Kuliah Data Storytelling. Berikut merupakan materi pembelajaran yang relevan dengan rekognisi mata kuliah Data Storytelling

Tabel 4. 11 Indikator Penilaian Konversi Mata Kuliah Data Storytelling

No	Kriteria/Indikator Penilaian
1	Kemampuan dalam memahami sebuah data dan mampu menyusunnya ke dalam sebuah narasi yang deskriptif, ringkas dan informatif
2	Kemampuan dalam menggunakan data dan menyajikannya dalam bentuk visualisasi tabel, grafik, dan infografik yang menarik serta mengkombinasikannya dengan narasi
3	Kemampuan dalam menyampaikan dan mengkomunikasikan secara verbal dan nonverbal dari sebuah hasil analisis data kepada pemangku kepentingan
4	Kemampuan dalam menggunakan data untuk menyelesaikan permasalahan dari hasil observasi/lapangan sehingga dapat menghasilkan sebuah penyajian data yang aplikatif

4.8.1. Memahami sebuah data dan Menghasilkan Hasil Analisis dan Penyajian Data Menggunakan Looker Studio ke dalam sebuah narasi [Indikator ke-1, Indikator ke-2, Indikator ke-3, Indikator ke-4]

Definisi & Konsep Dasar Storytelling

Storytelling adalah proses menyampaikan sebuah cerita melalui berbagai media, seperti kata-kata, Gambar, atau suara. Dalam bahasa Indonesia, istilah ini diartikan sebagai bercerita, yang terdiri dari dua kata: story (cerita) dan telling (penceritaan). Proses ini tidak hanya melibatkan penyampaian informasi, tetapi juga menciptakan pengalaman yang melibatkan interaksi antara pencerita (storyteller), cerita itu sendiri, dan pendengar (audience) (detikedu.com, 2022). Data storytelling adalah teknik yang menggabungkan data, visualisasi, dan narasi untuk menyampaikan informasi dengan cara yang menarik dan mudah dipahami. Ini adalah keterampilan penting bagi analis data untuk mengubah data mentah menjadi cerita yang informatif, sehingga memudahkan pengambilan keputusan dalam konteks bisnis. (RevoU, 2024)

Data storytelling juga bisa menjadikan data sebagai bagian dari narasi atau cerita, bukan hanya sekumpulan teks dan angka. Untuk membantu menyampaikan informasi melalui data yang sudah diolah, penulis menggunakan bantuan tools menggunakan Looker Studio dalam pembelajaran mata kuliah ini. Penulis menggunakan data dalam analisis ini diperoleh dari sistem pencatatan kunjungan Telkomsel ke institusi pendidikan, mulai dari jenjang taman kanak-kanak hingga perguruan tinggi, selama periode Januari hingga Oktober 2024 dan sumber data utama adalah dari platform mynami.id dengan format csv. Dataset ini terdiri dari 129.719 baris data dengan 24 kolom yang mencakup informasi yaitu

Atribut	Deskripsi
PERIODE	Menunjukkan bulan atau kuartal kunjungan.
PERIODE DATE	Tanggal spesifik kunjungan.
DAY NAME	Hari kunjungan berlangsung.
AGENT ID, USER ID	Identifikasi agen dan pengguna.
USER NAME	Nama petugas atau agen.
USER TYPE	Kategori pengguna, seperti staf lapangan
AREA, REGIONAL, BRANCH, CLUSTER, CITY	Mengindikasikan wilayah dan lokasi geografis kunjungan.
NPSN	Nomor identifikasi sekolah.
INSTITUTION NAME	Nama institusi atau sekolah.
LEVEL	Tingkat pendidikan (misalnya SD, SMP, SMA).
PJP FREQUENCY	Frekuensi kunjungan yang direncanakan.
DISPOSED DAY, DISPOSED STATUS	Hari dan status kunjungan (berhasil atau tidak).
PJP CHECKIN, PJP EXTRA	Aktivitas check-in utama dan tambahan selama kunjungan.
CHECKIN TIME, CHECKOUT TIME	Waktu mulai dan selesai kunjungan.
DURATION (MINUTE)	Durasi kunjungan dalam menit.
CHECKIN VALID	Validasi keberhasilan check-in.

Analisis ini bertujuan untuk memahami pola dan tren kunjungan Telkomsel ke sekolah berdasarkan wilayah dan tingkat pendidikan. Fokus utama meliputi distribusi kunjungan, frekuensi dan durasi aktivitas, serta efektivitas kunjungan dalam mendukung strategi pemasaran Telkomsel. Hasil analisis ini diharapkan dapat memberikan wawasan yang mendalam untuk pengambilan keputusan strategis, seperti alokasi sumber daya dan optimalisasi program kunjungan ke institusi pendidikan.

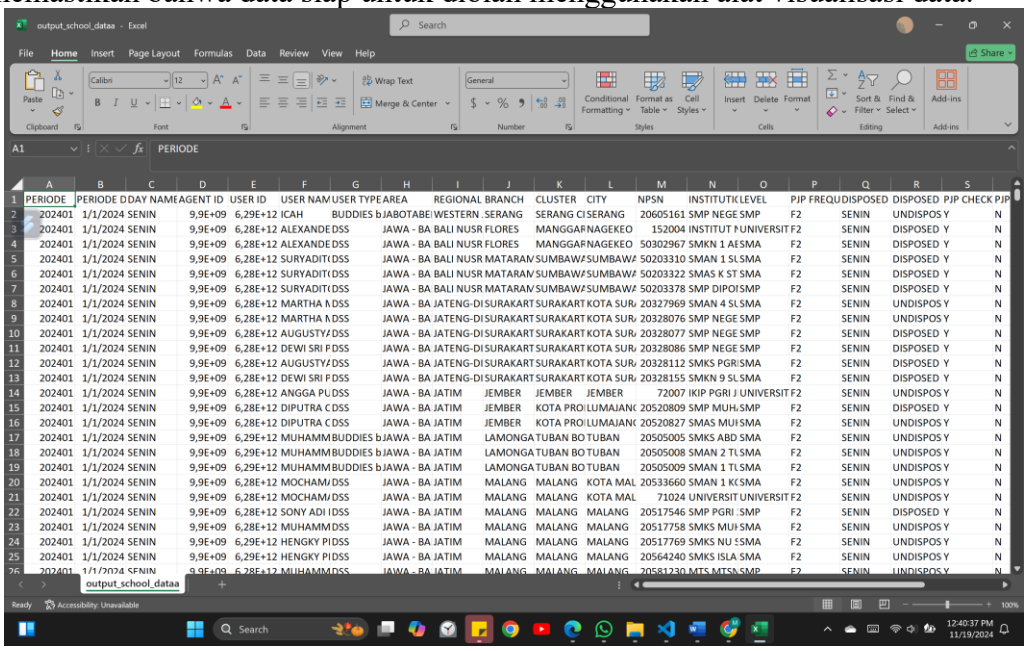
Proses pelaksanaan untuk menghasilkan visualisasi data kunjungan Telkomsel ke School memiliki beberapa tahap yang dilakukan, sebagai berikut :

a. Mengumpulkan Data

Pada tahap awal, dilakukan proses pengumpulan data dengan tujuan untuk memahami pola dan efektivitas kunjungan Telkomsel ke sekolah-sekolah selama periode Januari hingga Oktober 2024. Data yang dikumpulkan mencakup catatan rinci aktivitas kunjungan, dengan total 129.719 baris data dan 24 kolom

b. Analisis Data

Setelah data terkumpul, langkah selanjutnya adalah melakukan analisis data. Tahap awal dalam analisis ini melibatkan pengecekan dan penyesuaian tipe data untuk memastikan setiap variabel sesuai dengan format yang tepat, seperti angka, teks, atau tanggal. Hal ini penting untuk mendukung proses analisis lebih lanjut dan mencegah kesalahan saat data diolah. Proses selanjutnya adalah merapikan data menggunakan perangkat lunak Excel. Data yang tidak relevan dihapus, nilai yang hilang diimputasi jika diperlukan, dan tata letak data disesuaikan agar lebih terstruktur. Langkah ini bertujuan untuk meningkatkan keterbacaan dan memastikan bahwa data siap untuk diolah menggunakan alat visualisasi data.



Gambar 4. 82 Data School

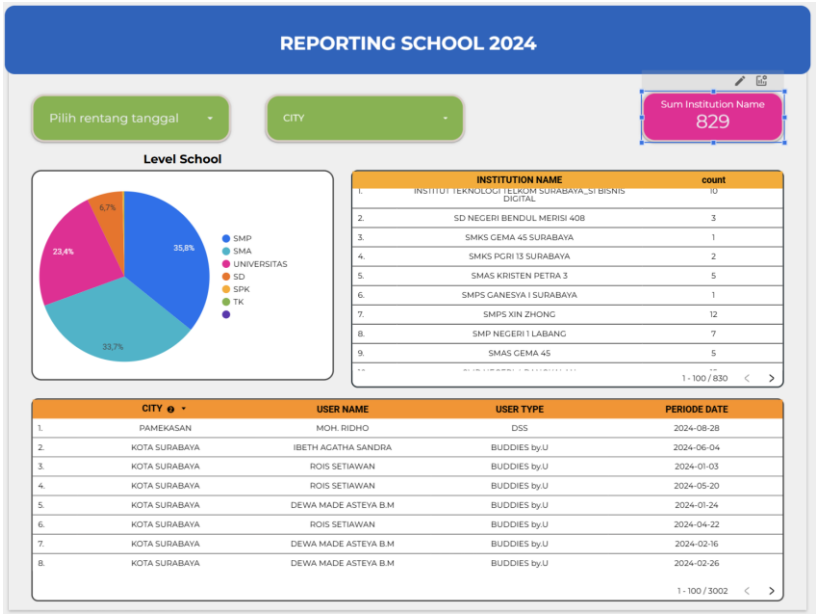
Data yang telah dirapikan kemudian ditransfer ke Google Spreadsheet sebagai platform penyimpanan yang terintegrasi. Hal ini dilakukan untuk mempermudah koneksi langsung antara dataset dan Looker Studio.

c. Visualisasi Data dan Infografis Dengan Menyampaikan Secara Verbal & Non-Verbal

Setelah data siap dan terhubung ke Looker Studio, langkah selanjutnya adalah membuat visualisasi data yang informatif dan interaktif. Visualisasi ini dirancang untuk menyajikan informasi secara efektif sehingga mempermudah pemahaman terhadap pola, tren, dan insight yang terkandung dalam data. Berbagai jenis data visualisasi seperti pie chart dan tabel dinamis. Visualisasi data di Looker Studio juga dilengkapi dengan fitur interaktif seperti filter dan kontrol waktu, sehingga pengguna dapat menyelami data lebih mendalam sesuai kebutuhan. Proses ini bertujuan untuk menghasilkan dashboard yang mudah dipahami oleh berbagai

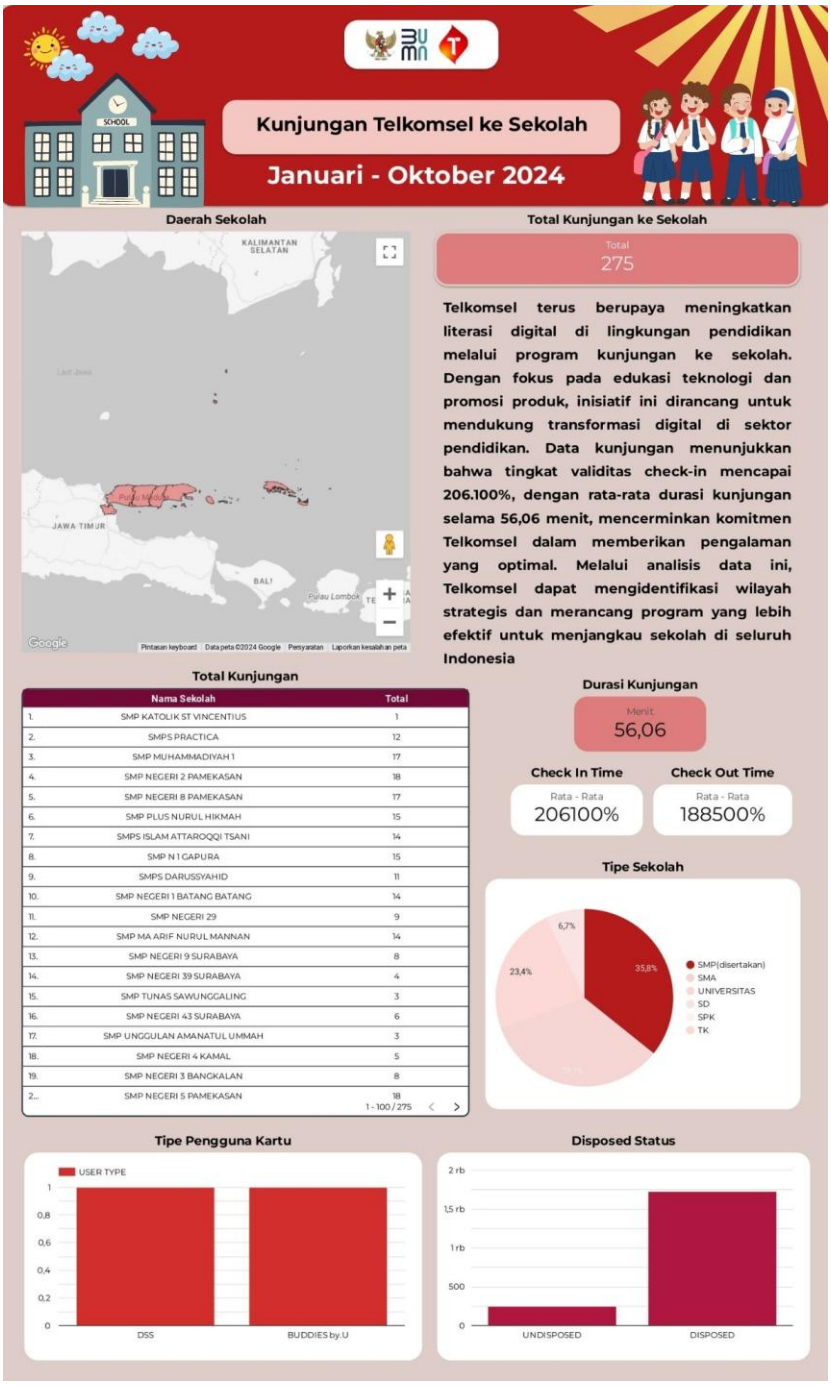
pemangku kepentingan dan dapat digunakan sebagai alat bantu pengambilan keputusan.

Gambar 4. 83 Menyampaikan Hasil Analisis



Gambar 4. 84 Dashboard Reporting School

Dapat disimpulkan bahwa dashboard diatas memberikan Gambaran awal mengenai aktivitas pelaporan sekolah pada tahun 2024 khususnya dari bulan Januari - Oktober. Fokus utama pelaporan saat ini adalah pada tingkat SMP, terutama di wilayah Kota Surabaya. Data ini dapat digunakan untuk berbagai keperluan, seperti: Mengevaluasi tingkat aktivitas pelaporan di berbagai institusi dan wilayah, Membantu dalam perencanaan program atau kegiatan terkait Pendidikan, Mengembangkan sistem pelaporan yang lebih baik di masa depan. Tidak hanya mencantumkan hasil dashboard, tetapi penulis juga menyampaikan ke dalam infografis. Seperti berikut :



Gambar 4. 85 Infografis

Narasi merupakan bentuk teks yang bertujuan untuk menceritakan suatu peristiwa atau kejadian secara kronologis, mendetail, dan runtut. Melalui narasi, pembaca diajak untuk memahami alur kejadian secara jelas dan terstruktur. Tujuan utamanya adalah memberikan wawasan, menggambarkan situasi, atau menyampaikan informasi sehingga pembaca dapat memperoleh Gambaran yang menyeluruh dan mendalam mengenai topik yang dibahas. (Wikipedia, 2024)

Narasi yang disusun dalam presentasi dashboard Looker Studio bertujuan untuk menjelaskan temuan utama terkait tren dan pola kunjungan Telkomsel ke sekolah selama periode Januari hingga Oktober 2024. Dashboard ini dirancang untuk

memberikan Gambaran menyeluruh mengenai distribusi kunjungan, frekuensi, durasi, serta efektivitas kunjungan berdasarkan wilayah dan kategori sekolah. Dan disampaikan secara terstruktur untuk membantu pemangku kepentingan memahami data dengan lebih mudah dan memberikan dasar yang kuat untuk pengambilan keputusan strategis guna meningkatkan efisiensi kunjungan Telkomsel ke sekolah.

Berikut adalah narasi yang telah dibuat untuk menyampaikan dan mengkomunikasikan secara verbal dan nonverbal dari sebuah hasil analisis data kepada pemangku kepentingan:

Selamat siang, para hadirin yang terhormat. Perkenalkan, saya Tsabita Rosyidah Putri, mahasiswa dari Universitas Pembangunan Nasional "Veteran" Jawa Timur. Pada kesempatan ini, saya akan menyampaikan hasil analisis yang telah saya lakukan mengenai tren dan pola kunjungan Telkomsel ke sekolah, serta memberikan rekomendasi untuk meningkatkan efisiensi kunjungan sekolah yang perlu ditingkatkan atau dipertahankan, berdasarkan data kunjungan Telkomsel ke sekolah dari bulan Januari hingga Oktober 2024. Data yang digunakan telah saya olah dan sajikan dalam bentuk dashboard visualisasi menggunakan tools Looker Studio. Adapun data yang disajikan yaitu total institusi yang dikunjungi, distribusi tingkat pendidikan, top institusi berdasarkan jumlah kunjungan, dan data kunjungan berdasarkan pengguna.

Berikut adalah hasil analisis yang telah saya lakukan :

1. Total Institusi yang Dikunjungi:

Dashboard ini menunjukkan bahwa sebanyak 829 institusi pendidikan telah dikunjungi, tersebar di berbagai wilayah dan tingkat pendidikan. Jumlah ini mencerminkan cakupan yang luas dari kegiatan kunjungan yang dilakukan.

2. Distribusi Tingkat Pendidikan:

Distribusi tingkat pendidikan yang dikunjungi dapat dilihat melalui pie chart "Level School". Dari data yang ada, tingkat SMP mendominasi dengan persentase 35,8%, diikuti oleh tingkat SMA yang mencatatkan angka 33,7%. Sedangkan tingkat SD menyumbang 23,4%, dan tingkat Universitas, SPK, serta TK masing-masing memiliki proporsi yang lebih kecil, yaitu sekitar 6,7%.

3. Top Institusi Berdasarkan Jumlah Kunjungan:

Tabel "Institution Name" menunjukkan institusi dengan jumlah kunjungan terbanyak. Institut Teknologi Telkom Surabaya menempati urutan teratas dengan 10 kunjungan. Beberapa institusi lainnya juga tercatat melakukan kunjungan, namun dengan jumlah yang lebih rendah.

4. Data Kunjungan Berdasarkan Pengguna:

Tabel bagian bawah memberikan rincian lebih lanjut mengenai kunjungan berdasarkan kota, nama pengguna (User Name), tipe pengguna (User Type), dan tanggal kunjungan (Period Date). Beberapa pengguna seperti "MOH. RIDHO" dan "IBETH AGATHA SANDRA" tercatat melakukan kunjungan ke berbagai institusi pada periode tertentu, menunjukkan variasi dalam distribusi kunjungan berdasarkan pengguna.

Berdasarkan analisis yang telah disimpulkan, ada beberapa rekomendasi untuk meningkatkan efisiensi kunjungan Telkomsel ke sekolah, yaitu :

1. Optimalisasi Kunjungan di Tingkat Pendidikan

Proporsi kunjungan ke tingkat Universitas masih rendah. Disarankan untuk meningkatkan jumlah kunjungan ke institusi pada tingkat ini, terutama jika potensi pengembangan layanan Telkomsel di sektor tersebut masih besar.

2. Evaluasi Strategi Kunjungan

Wilayah dengan kunjungan rendah, seperti yang terlihat di tabel bawah, perlu dievaluasi lebih lanjut. Fokuskan kunjungan ke kota-kota dengan tingkat kebutuhan layanan digital yang tinggi.

3. Peningkatan Keterlibatan di Institusi Populer

Institusi dengan jumlah kunjungan tinggi, seperti Institut Teknologi Telkom Surabaya, dapat menjadi mitra strategis untuk program khusus yang melibatkan teknologi atau layanan Telkomsel.

BAB 5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan analisis yang dilakukan menggunakan model ARIMA untuk memprediksi total kunjungan (*minute*), dapat disimpulkan bahwa model ARIMA(6, 2, 1) yang telah dipilih berdasarkan nilai AIC memberikan hasil yang baik dalam menggambarkan pola data historis. Proses pemilihan parameter p, d, dan q dilakukan dengan memanfaatkan visualisasi ACF dan PACF serta evaluasi menggunakan nilai AIC. Uji asumsi residual dengan uji *Ljung-Box* dan *Shapiro-Wilk* menunjukkan bahwa model ini memenuhi asumsi dasar, yaitu tidak ada autocorrelation yang signifikan dan residual terdistribusi normal, yang mengindikasikan bahwa model ARIMA ini dapat diandalkan untuk melakukan prediksi. Evaluasi model menggunakan metrik MAPE (143.60%), RMSE (392.954), dan MAE (315.5047) menunjukkan bahwa MAE lebih baik untuk mengukur kesalahan yang rata-rata dan mudah dipahami dalam konteks prediksi total kunjungan. Nilai MAE yang rendah menunjukkan bahwa model memberikan prediksi yang cukup akurat secara keseluruhan.

5.2. Saran

1. Peningkatan Model yaitu meskipun model ARIMA(6, 2, 1) memberikan hasil yang baik, masih mungkin untuk mengoptimalkan model lebih lanjut dengan mencoba kombinasi parameter yang lebih beragam, atau mempertimbangkan model lain seperti SARIMA atau model berbasis *machine learning* jika ada faktor musiman yang lebih kompleks yang perlu diperhitungkan.
2. Pemeliharaan Model yaitu model ini perlu terus diperbarui dengan data terbaru untuk menjaga akurasi prediksi, mengingat perubahan pola perilaku pengguna dapat terjadi seiring waktu. Implementasi pembelajaran berkelanjutan (*online learning*) dapat dipertimbangkan untuk meningkatkan responsivitas model terhadap perubahan data.
3. Evaluasi Lebih Lanjut yaitu untuk memperoleh pemahaman yang lebih mendalam tentang kinerja model, disarankan untuk melakukan uji validasi silang (*cross-validation*) dan menguji model di berbagai segmen data untuk memastikan stabilitas dan keandalannya di berbagai kondisi.

BAB 6. REFLEKSI DIRI

Selama magang Telkomsel, saya menemukan bahwa banyak pengetahuan yang saya peroleh dari kuliah sangat relevan dengan pekerjaan saya saat ini. Mata kuliah seperti Business Intelligence, Analisis Segmentasi Pelanggan, dan Data Storytelling menjadi dasar kuat untuk menganalisis data pelanggan dan membuat visualisasi yang menarik dan informatif. Pengalaman ini meningkatkan pemahaman teknis saya tentang cara data dapat membantu keputusan strategis bisnis, terutama dalam hal pemasaran dan retensi pelanggan.

Pengembangan soft skill, terutama dalam hal komunikasi, kerja sama, dan adaptasi, sangat dipengaruhi oleh program MBKM ini. Meskipun saya terbiasa bekerja dalam tim dan berkomunikasi dengan berbagai rekan dari berbagai latar belakang di Telkomsel, menyampaikan insight secara jelas dan efektif menjadi tantangan. Kemampuan untuk berkolaborasi dengan orang dari berbagai divisi dan beradaptasi dengan lingkungan kerja yang dinamis juga sangat penting untuk meningkatkan fleksibilitas dan keterampilan interpersonal.

Saya belajar dari pengalaman ini untuk tidak menyerah, mencari solusi, meminta bantuan, dan belajar dari kesalahan saya. Melalui pemikiran ini, saya menyadari bahwa program magang mandiri adalah proses pembelajaran diri yang berkelanjutan selain mendapatkan pengalaman kerja. Saya berkomitmen untuk terus berkembang secara pribadi dan profesional dengan memanfaatkan setiap kesempatan yang tersedia untuk meningkatkan diri saya. Dan ke depannya saya berencana untuk memperdalam keterampilan dalam analisis prediktif dan data science, serta fokus mengembangkan karir di bidang Business Intelligence. Di sisi pendidikan, saya juga akan mengambil kursus tambahan dan aktif dalam kegiatan yang dapat mendukung pemahaman teknis saya lebih lanjut, sehingga siap untuk menghadapi tantangan di masa depan.

DAFTAR PUSTAKA

Zulkarnaina, M. F., Ardiyantia, N. P. N., Sandia, I. W. W. K., Dewa, I., Hendrawana, N. T., & Mahendra, I. B. M. (2021). Perancangan dan Implementasi Data Warehouse Penjualan (Studi Kasus: Northwind Sample Database). Jurnal Elektronik Ilmu Komputer Udayana p-ISSN, 2301, 5373.

Rahmawati, S. N. E., Hasanah, M., Rohmah, A., Pratama, R. A. P., & Anshori, M. I. (2023). Privasi dan etika dalam manajemen sumber daya manusia digital. Lokawati: Jurnal Penelitian Manajemen Dan Inovasi Riset, 1(6), 01-23.

Ariani, A. F., & Aulia, K. (2024). Pengembangan Dashboard Interaktif Menggunakan Looker Studio untuk Visualisasi dan Prediksi Harga Komoditas Cabe di Jawa Timur. JATI (Jurnal Mahasiswa Teknik Informatika), 8(4), 8067-8074.

Maulana, H. A. (2018). Pemodelan deret waktu dan peramalan curah hujan pada dua belas stasiun di Bogor. Jurnal Matematika, Statistika dan Komputasi, 15(1), 50-63.

Melyani, C. A., Nurtsabita, A., Shafa, G. Z., & Widodo, E. (2021). Peramalan inflasi di Indonesia menggunakan metode Autoregressive Moving Average (ARMA). Journal of Mathematics Education and Science, 4(2), 67-74.

Adiguno, S., Syahra, Y., & Yetri, M. (2022). Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda. Jurnal Sistem Informasi Triguna Dharma (JURSI TGD), 1(4), 275-281.

Komara, L. N., & Sirodj, D. A. N. (2023, August). Metode Autoregressive Integrated Moving Average (ARIMA) untuk Meramalkan Produksi Padi di Provinsi Jawa Tengah. In Bandung Conf. Ser. Stat (Vol. 3, No. 2, pp. 496-504).

Karnadi, E. V. (2023). Peramalan Harga Bitcoin, Ethereum dan Binance Menggunakan Metode Arima (Autoregressive Integrated Moving Average) dengan Variabel Eksogen (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).

Fadlullah, N. M. Laporan Praktik Kerja Lapangan (Pkl) Pada Pt. Telekomunikasi Indonesia Witel Telkom Jakarta Timur Divisi Digital Service And Wifi (Dsw).

Lubis, A. H., & Ganesha, P. (2016). Model Segmentasi Pelanggan Dengan Kernel K-Means Clustering Berbasis Customer Relationship Management. Jurnal & Penelitian Teknik Informatika, 1(1).

Herlinda, V., Darwis, D., & Dartono, D. (2021). Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means. Jurnal Teknologi Dan Sistem Informasi, 2(2), 94-99.

Yendra, R., & Noviadi, E. T. (2015). Perbandingan estimasi parameter pada distribusi eksponensial dengan menggunakan metode maksimum likelihood dan metode bayesian. Jurnal Sains Matematika dan Statistika, 1(2), 62-72.

LAMPIRAN

1. Logbook Harian
2. Surat Rekomendasi
3. Surat Pengajuan Konversi MBKM
4. Surat Persetujuan Konversi MBKM
5. Kartu Konsultasi Mahasiswa
6. RPS Skema MBKM
7. Daftar Hadir Dosen
8. Berita Acara
9. Lembar Revisi Dosen Pembimbing
10. Lembar Revisi Dosen Penguji
11. Kode Script Rekognisi Mata Kuliah Kecerdasan Bisnis
    https://bit.ly/CodeKecerdasanBisnis
12. Kode Script Rekognisi Mata Kuliah Analisis Segmentasi Pelanggan
    https://bit.ly/CodeAnalisisSegmentasiPelanggan
13. Kode Script Analisis Deret Waktu
    https://bit.ly/CodeAnalisisDeretWaktu