

**PERAMALAN TREND PENJUALAN RETAIL
MENGUNAKAN *AUTOREGRESSIVE INTEGRATED
MOVING AVERAGE***



TESIS

LUKY FABRIANTO

14002269

**PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER**

NUSA MANDIRI

JAKARTA

2020

**PERAMALAN TREND PENJUALAN RETAIL
MENGUNAKAN *AUTOREGRESSIVE INTEGRATED
MOVING AVERAGE***



TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar
Magister Ilmu Komputer (M.Kom)

LUKY FABRIANTO

14002269

**PROGRAM PASCASARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2020**

SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISME

Yang bertanda tangan di bawah ini :

Nama : Luky Fabrianto
NIM : 14002269
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Peramalan Trend Penjualan Retail Menggunakan *Autoregressive Integrated Moving Average*” adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan

Jakarta, 19 Agustus 2020
Yang menyatakan,

Materai Rp 6 000,-

Luky Fabrianto

PERSETUJUAN TESIS

Tesis ini diajukan oleh :

Nama : Luky Fabrianto
NIM : 14002269
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Judul Tesis : “Peramalan Trend Penjualan Retail Menggunakan
Autoregressive Integrated Moving Average”

Untuk dipertahankan pada periode 2020-1 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 19 Agustus 2020

PEMBIMBING TESIS

Pembimbing : Dr. Yan Rianto, M. Eng.

D E W A N P E N G U J I

Penguji I : Dr. Windu Gata, M.Kom

Penguji II : Dr. Hilman F. Pardede, ST, MEICT

Penguji III /
Pembimbing : Dr. Yan Rianto, M. Eng

Catatan:

*Lembar Persetujuan & Pengesahan ini WAJIB ditanda tangani oleh Pembimbing & Asisten Pembimbing

*Penulisan tanggal, bulan, dan tahun adalah tanggal pengesahan tesis

HALAMAN PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh:

Nama	: Luky Fabrianto
NIM	: 14002269
Program Studi	: Ilmu Komputer
Jenjang	: Strata Dua (S2)
Konsentrasi	: <i>Data Mining</i>
Judul Tesis	: Peramalan Trend Penjualan Retail Menggunakan <i>Autoregressive Integrated Moving Average</i>

Telah dipertahankan pada periode 2020-1 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 19 Agustus 2020

PEMBIMBING TESIS

Pembimbing I : Dr. Yan Rianto, M.Eng



DEWAN PENGUJI

Penguji I : Dr. Windu Gata, M.Kom



Penguji II : Dr. Hilman Ferdinandus Pardede,
S.T, M.EICT



Penguji III /
Pembimbing I : Dr. Yan Rianto, M.Eng



	LEMBAR BIMBINGAN TESIS
	STMIK NUSA MANDIRI

NIM : 14002269
 Nama Lengkap : Luky Fabrianto
 Dosen Pembimbing : Dr. Yan Rianto, M. Eng.
 Judul Tesis : “Peramalan Trend Penjualan Retail Menggunakan *Autoregressive Integrated Moving Average*”

NO	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
1	5 Apr 2020	Proposal Tesis	
2	19 Apr 2020	Dataset Penelitian	
3	10 Mei 20	Hasil Pengujian	
4	7 Jun 2020	Revisi Hasil Pengujian (Outlier)	
5	24 Jun 2020	Revisi Hasil Pengujian (Smoothing)	
6	12 Jul 2020	Revisi Hasil Pengujian	
7	25 Jul 2020	Naskah Tesis	
8	2 Aug 2020	Revisi Naskah Tesis	

Catatan untuk dosen pembimbing Bimbingan Tesis

- Dimulai pada tanggal : 5 April 2020
- Diakhiri pada tanggal : 2 Agustus 2020
- Jumlah pertemuan bimbingan : 8 (Delapan) kali bimbingan

Disetujui oleh,

Dosen Pembimbing



(Dr. Yan Rianto, M. Eng.)

KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadiran Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan laporan tesis ini tepat pada waktunya. Dimana laporan tesis ini penulis sajikan dalam bentuk buku yang sederhana.

Adapun judul tesis, yang penulis ambil sebagai berikut: “Peramalan Trend Penjualan Retail Menggunakan *Autoregressive Integrated Moving Average*”.

Tujuan penulisan laporan tesis ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) STMIK Nusa Mandiri.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan laporan tesis ini, maka penulis tidak dapat menyelesaikan laporan tesis ini tepat pada waktunya. Untuk itu ijinkanlah penulis kesempatan ini untuk mengucapkan

1. Ibu Dr. Dwiza Riana, S.Si, MM, M.Kom sebagai Ketua STMIK Nusa Mandiri
2. Ibu Nita Merlina, M.Kom sebagai Wakil Ketua I Bidang Akademik STMIK Nusa Mandiri
3. Bapak Dr. Hilman F. Pardede, ST, MEICT sebagai Ketua Program Studi Ilmu Komputer
4. Bapak Dr. Yan Rianto, M. Eng. selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini.
5. Orang tua, istriku Novianti M. Faizah, S.Kom, M.MSi dan anakku Muhammad Fauzi Pramudya tercinta yang telah memberikan dukungan material dan moral kepada penulis.
6. Seluruh Dosen Program Studi Ilmu Komputer (S2) STMIK Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.
7. Seluruh staf di lingkungan STMIK Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah.

8. Mahasiswa Program Studi Ilmu Komputer (S2) STMIK Nusa Mandiri angkatan 2018 yang menjadi teman dalam berbagi ilmu.
9. Seluruh pengurus dan staf di Toko Teras Umat yang telah melayani penulis dengan baik selama penulisan berlangsung.

Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan laporan tesis ini. Penulis menyadari bahwa penulisan laporan tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah yang penulis hasilkan untuk yang akan datang.

Akhir kata semoga laporan tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 19 Agustus 2020

Luky Fabrianto

Penulis

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya :

Nama : Luky Fabrianto
NIM : 14002269
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalti-Free Right) atas karya ilmiah kami yang berjudul : “Peramalan Trend Penjualan Retail Menggunakan *Autoregressive Integrated Moving Average*”. beserta perangkat yang diperlukan (apabila ada).

Dengan Hak Bebas Royalti Non-Eksklusif ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau bentuk-kan, mengelolanya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini .

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 19 Agustus 2020
Yang menyatakan,

Materai Rp 6 000,-

Luky Fabrianto

ABSTRAK

Nama : Luky Fabrianto
NIM : 14002269
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Judul : “Peramalan Trend Penjualan Retail Menggunakan
Autoregressive Integrated Moving Average”

Dalam satu tahun sejak awal berdiri, pada sembilan bulan pertama penjualan berfluktuatif namun berada dalam nilai transaksi yang menguntungkan serta bisa menutupi biaya operasional, tetapi pada tiga bulan terakhir nilai penjualan cenderung flat dan menurun, hal tersebut diduga disebabkan oleh *supply* barang dari distributor yang tersendat sehingga kebutuhan konsumen banyak tidak tersedia di toko. Sudah saatnya pengurus toko untuk memberikan peluang bagi para anggota untuk berinvestasi menanamkan modal guna kelanjutan eksistensi toko.

Penelitian ini mengkaji metode yang tepat untuk meramalkan trend penjualan toko, data yang dipakai adalah data penjualan satu tahun sejak toko berdiri. Hasil yang didapat ada sekitar 5% pencilan pada transaksi dan fluktuasi transaksi antar hari cukup besar sehingga diperlukan penghalusan untuk meminimalisir gejolak pada data transaksi, tahap terakhir dari penelitian ini adalah peramalan trend penjualan menggunakan metode ARIMA, model yang didapat adalah ARIMA (3,1,2) dengan persentase kesalahan atau MAPE sebesar 2.5% dan nilai RMSE sebesar 112883.14236. Untuk 15 hari kedepan diramalkan nilai penjualan berada pada kisaran antara Rp. 3.187.116 sampai Rp. 3.612.883.

Kata kunci: Peramalan, Penjualan, Pencilan, Penghalusan, ARIMA, MAPE

ABSTRACT

Name : Luky Fabrianto
NIM : 14002269
Study of Program : Computer Science
Levels : Master (S2)
Concentration : Data Mining
Title : “Forecasting Retail Sales Trends Using Autoregressive Integrated Moving Average”.

One year since its inception, in the first nine months sales fluctuated but were in a profitable transaction value and could cover operational costs, but in the last three months the sales value tends to be flat and declining, this is thought to be caused by stagnant supply of goods from distributors so that many consumer needs are not available in stores. It is time for the store management to turn to other distributors and provide opportunities for members to invest in investing for the continued existence of the store.

This study examines the right method for predicting store sales trends, the data used is sales data one year since the store was established. The results obtained are about 5% outliers in transactions and large transaction fluctuations inter-day so that smoothing is needed to minimize shocks in transaction data, the last stage of this study is forecasting sales trends using the ARIMA method, the model obtained is ARIMA (3,1, 2) with a percentage of errors or MAPE of 2.5% and an RMSE value of 112883,14236. For the next 15 days the sales value is predicted to be in the range of Rp. 3,187,116 to Rp. 3,612,883.

Keyword: Forecasting, Outliers, Smoothing, ARIMA, MAPE

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISM	iii
PERSETUJUAN TESIS.....	iv
PERSETUJUAN DAN PENGESAHAN	v
LEMBAR BIMBINGAN TESIS.....	vi
KATA PENGANTAR	vii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Penulisan	1
1.2 Identifikasi Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Ruang Lingkup Penelitian	3
1.5 Sistematika Penulisan.....	4
BAB 2 LANDASAN PEMIKIRAN	6
2.1 Tinjauan Pustaka	6
2.2 Tinjauan Studi	13
2.3 Tinjauan Obyek Penelitian.....	17
2.4 Kerangka Pemikiran	18
BAB 3 METODE PENELITIAN	19
3.1 Jenis Penelitian.....	19
3.2 Sumber Data dan Variabel Penelitian.....	19
3.3 Metode Peramalan.....	19
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN	28
4.1 Karakteristik Data Penjualan	28
4.2 Penyisihan Pencilan.....	28
4.3 Penghalusan	31
4.4 Permodelan ARIMA	31
BAB 5 PENUTUP	43
5.1 Kesimpulan	43
5.2 Saran.....	43
DAFTAR PUSTAKA	44
DAFTAR RIWAYAT HIDUP	48

DAFTAR TABEL

	Halaman
Tabel 2.1. Pengamatan Plot ACF dan PACF.....	11
Tabel 2.2. Penelitian Terkait.....	15
Tabel 3.1. Tingkat Akurasi Peramalan Berdasarkan MAPE	26
Tabel 4.1. Kuartil	28
Tabel 4.2. Data Pencilan.....	30
Tabel 4.3. Hasil Percobaan	36
Tabel 4.4. Hasil Pengukuran Kesalahan Model.....	37
Tabel 4.5. Nilai Penjualan Aktual	42

DAFTAR GAMBAR

	Halaman
Gambar 3.1. Pencilan Dalam Dataset.....	20
Gambar 3.2. Pola Data Horizontal	21
Gambar 3.3. Pola Data Trend	21
Gambar 3.4. Pola Data Musiman (<i>seasonal</i>)	22
Gambar 3.5. Flowchart ARIMA	27
Gambar 4.1. Plot Data Penjualan Toko	28
Gambar 4.2. Skema Pencilan Dari Data Penjualan Toko.....	29
Gambar 4.3. Plot Data Penjualan Setelah dilakukan Penyisihan Pencilan.....	30
Gambar 4.4. Plot Data Penjualan dengan penghalusan.....	31
Gambar 4.5. Plot <i>differencing</i> 1 & 2	32
Gambar 4.6. Plot ACF	33
Gambar 4.7. Plot PACF	33
Gambar 4.8. Plot <i>Residual</i>	34
Gambar 4.9. Plot <i>Density</i>	34
Gambar 4.10. Plot Peramalan dan Aktual	35
Gambar 4.11. Plot Peramalan Model ARIMA (1,1,1)	36
Gambar 4.12. Plot Peramalan Model ARIMA (3,1,2)	38
Gambar 4.13. <i>Standardized residual</i>	39
Gambar 4.14. <i>Histogram plus estimated density</i>	40
Gambar 4.15. <i>Plot Q-Q</i>	40
Gambar 4.16. Correlogram.....	41
Gambar 4.17. Plot Peramalan 15 Hari Kedepan.....	42

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Penjualan	48
Lampiran 2. Data Pencilan	49
Lampiran 3. Grafik Standar Deviasi	50
Lampiran 4. <i>Summary</i> Hasil Estimasi Parameter	51
Lampiran 5. <i>Coding</i> Jupyter	56

BAB I

PENDAHULUAN

1.1. Latar Belakang Penelitian

Koperasi konsumen adalah salahsatu dari jenis koperasi yang ada di Indonesia seperti yang tertera pada UU RI No. 17 Tahun 2012 [1], dalam hal ini minimarket yang dimiliki oleh anggota koperasi menjadikan para anggota sebagai konsumen bagi minimarket tersebut dibanding khalayak umum , sehingga peran aktif anggota dalam berbelanja sangat diharapkan.

Minimarket dari prespektif teknologi minimalnya adalah pemanfaatan komputerisasi untuk perhitungan dan penyimpanan segala transaksi keuangan yang terjadi. Pemanfaatan teknologi informasi juga sangat membantu para pelaku usaha dari pelbagai kalangan dalam banyak hal salahsatunya adalah peramalan.

Tidak seperti penjualan konvensional masa lalu yang orientasi manajemen hanya supply and demand. Pemanfaatan teknologi informasi dapat menyimpan segala transaksi penjualan maupun pembelian, sehingga dapat memberikan kemudahan dalam transaksi, dapat mengukur margin keuntungan, mengetahui secara realtime kondisi dari berbagai aspek, hingga pemanfaatan basis data sebagai peramalan penjualan guna meningkatkan profit bagi para anggota.

Data yang dipakai dalam penelitian adalah data penjualan dari sebuah toko koperasi konsumen, data penjualan tersebut mencerminkan aktifitas transaksi yang terjadi, selain untuk memenuhi kebutuhan rumah tangga aktifitas belanja anggota koperasi juga diisi dengan pembelian item tertentu dalam jumlah besar untuk kegiatan seperti santunan, sumbangan, dan lain sebagainya sehingga pola data penjualan cukup bergejolak pada waktu-waktu tertentu. Dari data tersebut peneliti melakukan peramalan penjualan untuk beberapa masa ke depan dengan tujuan memberikan informasi kepada para anggota untuk meningkatkan aktifitas transaksi di minimarket tersebut, bisa juga peramalan dipakai untuk membuka peluang investasi baru dari para anggota.

Peramalan adalah upaya memperkirakan yang akan terjadi dimasa depan berdasarkan data sebelumnya dengan basis metode ilmiah dan kualitatif yang dilakukan dengan sistematis [2].

Metode peramalan dalam yang banyak dikembangkan oleh peneliti bidang ilmu ekonomi adalah metode peramalan kuantitatif dimana metode tersebut dibagi menjadi dua jenis yang pertama metode regresi dan yang kedua metode *time series*. Terdapat beberapa metode *time series*, antara lain: rata-rata bergerak, penghalusan eksponensial, ARIMA, regresi deret waktu, dan lain-lain. Sedangkan dalam ilmu komputer metode peramalan yang dikembangkan dengan menggunakan konsep kecerdasan buatan, seperti: Jaringan syaraf tiruan, algoritma genetika, Fuzzy time series dan lain-lain [3].

Terdapat beberapa penelitian terkait peramalan dalam berbagai bidang yang telah dilakukan sebelumnya, seperti : Perbandingan beberapa metode peramalan dari bidang ilmu statistic dan ilmu computer dalam peramalan penjualan, hasilnya ARIMA dari bidang ilmu statistik bisa mengungguli beberapa metode peramalan dari bidang ilmu komputer[4], penelitian yang membandingkan metode peramalan antara ARIMA dengan Support Vector Machine (SVM) didapat hasil MAPE ARIMA 7,07% dan SVM 9,59%[5], penggunaan ARIMA dipakai dalam peramalan nilai impor biji besi Indonesia 2008 – 2017 menerangkan kelebihan dan kekurangan ARIMA, salahsatu poinnya adalah metode ARIMA akurat untuk peramalan jangka pendek [6], dan perbandingan antara metode peramalan ARIMA dengan *Radial Basis Function* (RBF) didapatkan error ARIMA lebih kecil dibandingkan RBF dan RBF-ARIMA [7].

Dalam penelitian ini transaksi yang tercatat dalam penyimpanan basis data akan dijadikan penelitian untuk membuat suatu peramalan penjualan 15 hari atau dua pekan kedepan dengan metode peramalan ARIMA. Kondisi toko koperasi saat ini membutuhkan investasi untuk perbaikan dari sisi *supply* sehingga pengurus membuka peluang investasi bagi para anggota (perorangan atau kelompok) dengan durasi per 15 hari dikarenakan perkulakan atau belanja toko untuk persediaan dilakukan dua kali dalam satu bulan.

1.2. Identifikasi masalah

Untuk meramalkan penjualan kedepan maka dibutuhkan data penjualan dari waktu sebelumnya yang tersimpan didalam basis data minimarket, sehingga permasalahan peramalan penjualan dapat dirumuskan.

Perumusan masalah dapat dilakukan dalam dua tahapan. Pertama, permasalahan bisnis dengan mengumpulkan segala data yang berkaitan dengan penjualan yang akan digunakan untuk meramalkan penjualan. Kedua, desain dari model peramalan penjualan yang akan digunakan, dalam penelitian ini digunakan metode peramalan ARIMA (*Autoregressive Integrated Moving Average*) yang akan dicoba untuk menyelesaikan permasalahan peramalan penjualan.

Dibawah ini adalah beberapa permasalahan atau pertanyaan yang akan muncul ketika akan meramalkan trend penjualan:

- Bagaimana pola data penjualan toko dari waktu pertama berdiri hingga setahun?
- Bagaimana hasil peramalan penjualan beberapa hari atau pekan ke depan dengan pendekatan metode ARIMA?

1.3. Tujuan penelitian

1. Mengetahui pola penjualan dari minimarket Teras Umat dalam kondisi ideal selama setahun, yaitu dari mulai 1 Mei 2018 sampai dengan 30 April 2019.
2. Memprediksi penjualan dengan pendekatan ARIMA.
3. Mengetahui model terbaik untuk memprediksi penjualan kedepan.
4. Memberi gambaran peramalan penjualan toko kepada anggota agar aktifitas belanja para anggota meningkat atau membuka peluang investasi baru bagi anggota koperasi.

1.4. Ruang Lingkup Penelitian

Data primer dari penjualan toko yang dipakai sebagai *dataset* dalam penelitian ini adalah didapat dari Toko Teras Umat yang terletak di wilayah Kota administrative Depok propinsi Jawa Barat. Data transaksi yang dipakai pada penelitian ini adalah data penjualan selama satu tahun sejak toko berdiri, yaitu dari

1 Mei 2018 sampai dengan 30 April 2019 dimana terdapat 364 hari dengan total nilai transaksi sebesar Rp. 1.579.862.842. Data penjualan satu tahun sejak toko berdiri adalah data pada saat penjualan dalam kondisi ideal, dimana *supply* dan *demand* serta partisipasi belanja anggota sangat baik, diharapkan dengan adanya investasi dari para anggota, kedepannya situasi tersebut akan kembali terulang.

Peramalan trend penjualan menggunakan metode ARIMA yang dikerjakan dengan bahasa pemrograman *Jupyter Notebook (Python)* menjadi batasan dalam penelitian ini dan persentase kesalahan terkecil (MAPE) serta pemilihan model berdasarkan nilai AIC terkecil sebagai pengukuran.

1.5. Sistematika Penulisan

Penelitian dibagi kedalam lima bab termasuk pendahuluan, yang disusun sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini pendahuluan yang menguraikan fakta-fakta berkaitan dengan permasalahan yang akan diteliti, mengenai deskripsi latar belakang penulisan, identifikasi masalah, tujuan masalah, ruang lingkup penelitian, hipotesis dan sistematika penulisan.

BAB II LANDASAN/KERANGKA PEMIKIRAN

Bab ini membahas tentang landasan-landasan teoritis yang digunakan untuk melaksanakan penelitian yang mencakup tinjauan pustaka, tinjauan studi dan tinjauan organisasi.

BAB III METODOLOGI PENELITIAN

Menguraikan proses pengumpulan data, konsep data mining, metodologi penelitian dan beberapa teori yang berkaitan dengan penelitian ini yang akan dijadikan solusi pada permasalahan yang diangkat.

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Bab ini menjelaskan tentang analisis penelitian dan pembahasan tahap-tahap yang dikerjakan, dari mulai pengambilan data, preprocessing data,

pengolahan data dengan metode ARIMA, selanjutnya menjelaskan hasil yang dapat dan implikasi dari penelitian tersebut.

BAB V PENUTUP

Bab kelima adalah kesimpulan dari temuan dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN / KERANGKA PENELITIAN

2.1. Tinjauan pustaka

Tinjauan pustaka dalam penelitian ini menggunakan beberapa referensi dari jurnal ilmiah, artikel dari internet, buku dan modul belajar. Berikut tinjauan pustaka yang mendukung teori dari penelitian, sebagai berikut:

2.1.1. Pencilan/*Outliers*

Pencilan adalah kondisi suatu data yang menyimpang dari kumpulan data yang lain [8], dapat juga diartikan pengamatan yang tidak mengikuti mayoritas pola dan letaknya jauh dari pusat data [9], dan mungkin berpengaruh besar terhadap koefisien regresi [10]. Jika dalam *dataset* terdapat pencilan maka dibutuhkan diagnostik yang dapat mengidentifikasi masalah pencilan tersebut, salah satunya dengan membuang pencilan dari kelompok data kemudian menganalisis data tanpa pencilan.

Metode umum untuk menyisihkan pencilan adalah dengan menggunakan nilai kuartil dan batasan. Kuartil-1 ($Q1$), kuartil-2 ($Q2$), dan kuartil-3 ($Q3$) membagi sebuah urutan data menjadi empat bagian. Batasan atau *IQR* (*Interquartile Range*) didefinisikan sebagai selisih $Q1$ terhadap $Q3$ atau $IQR = Q3 - Q1$. Pencilan dapat ditentukan dari nilai $(x) < 1.5 * IQR$ terhadap $Q1$ dan nilai $(x) > 1.5 * IQR$ terhadap $Q3$ [11].

2.1.2. Penghalusan

Prinsip dasar dari penghalusan adalah mengidentifikasi pola data dengan menghaluskan variasi lokal atau gejolak pendek pada suatu deret waktu dalam hal ini dilakukan penghitungan rata-rata bergerak atau *moving average* (*MA*).

Rerata bergerak (*MA*) secara sederhana adalah nilai pada suatu waktu dipengaruhi oleh beberapa nilai dari periode sebelumnya, rerata bergerak juga cocok digunakan untuk data berpola konstan atau stasioner.

Perhitungan rata-rata bergerak adalah[12] :

$$MA = \frac{\sum(n \text{ nilai data terbaru})}{n} \dots\dots\dots (2.1)$$

2.1.3. Peramalan Penjualan (*Forecasting*)

Analisis pada data *time series* ditujukan untuk melakukan suatu peramalan untuk masa yang akan datang. Data *time series* yang diperoleh yang kemudian akan dianalisis menjadi suatu acuan dalam meramalkan perubahan apa yang akan terjadi pada periode selanjutnya yang akan terjadi berdasarkan pola yang terjadi pada masa sekarang dan masa lampau. Perubahan yang diramalkan nantinya diharapkan dapat membantu pihak terkait untuk menentukan suatu keputusan untuk menghadapi kemungkinan yang akan terjadi di masa yang akan datang[13].

Peramalan data untuk masa depan yang dilakukan mengikuti langkah-langkah yang sistematis dan mengikuti model yang bersesuaian dengan sifat serta pola yang dimiliki oleh data asal. Sehingga hasil peramalan yang diperoleh dari model menjadi relevan dan memungkinkan untuk dilakukan.

2.1.4. Konsep Dasar Deret Waktu/*Time Series*

Data *time series* sering kali ditemukan dalam berbagai bidang disiplin ilmu seperti ekonomi, pertanian, meteorologi, biologi, serta disiplin ilmu lainnya. Data bentuk *time series* dapat dicatat berdasarkan periode waktu harian, pekanan, bulanan, tahunan, ataupun periode waktu tertentu lainnya dalam rentang waktu yang sama [14].

Data yang termasuk dalam jenis *time series* dapat diplot berdasarkan waktu. Hal ini dilakukan untuk mengamati pola dari data untuk selanjutnya menentukan langkah analisis yang akan dilakukan. Berdasarkan bentuk pola yang dibentuk, data *time series* dapat dibagi menjadi empat[13], sebagai berikut :

1. Data berpola horizontal, yaitu data yang relatif konstan dilihat dari plot data yang menyebar terhadap rata-rata (*mean*) dan ragam (*variance*).
2. Data berpola trend, plot data menyebar membentuk suatu kecenderungan menaik ataupun menurun secara konstan.

3. Data yang berpola musiman (seasonal), data ini merupakan data berpola trend yang dibentuk oleh faktor musiman dan cenderung berulang namun teratur.
4. Pola siklus, yaitu pola data berbentuk trend yang memuat musiman namun periode musimannya lebih panjang.

2.1.5. Stasioneritas Deret Waktu

Non stasioner dalam *time series* dibedakan menjadi 2, pertama non stasioner dalam *mean*/rata-rata, kedua non stasioner dalam *varians*/ragam. Non stasioner dalam *mean* diatasi dengan melakukan pembedaan/*differencing* dan untuk non stasioner dalam *varians* dilakukan transformasi [15].

Asumsi yang sangat penting dalam data runtun waktu adalah stasioneritas. Model runtun waktu yang sangat terkenal adalah model ARIMA untuk proses-proses non-stasioner dan untuk data yang *univariat* [16].

Stasioneritas dapat dideteksi melalui bentuk dari plot data runtun waktu atau melalui *correlogram*. Data stasioner adalah data yang memiliki nilai rata-rata (*mean*) dan *varians* yang tetap sepanjang waktu.

- Stasioner dalam *mean* (rata-rata)

Stasioner dalam hal rata-rata (*mean*) adalah suatu data *time series* yang memiliki rata-rata relatif konstan dari waktu ke waktu atau tidak memiliki trend dalam data.

- Stasioner dalam *variance* (ragam)

Stasioner dalam hal varian adalah suatu data *time series* yang memiliki varian konstan atau tidak berubah-ubah dari waktu ke waktu.

Untuk mendukung hasil dari identifikasi stasioneritas pada *correlogram* atau plot ACF dan PACF digunakan pengujian akar unit [17]. Akar unit pertama kali dikembangkan oleh *Dickey Fuller* sehingga disebut *Augment Dickey Fuller* (ADF). Uji akar unit digunakan untuk menguji adanya asumsi bahwa data *time series* tidak stasioner. Uji akar unit juga mengasumsikan bahwa residual bersifat independen dengan rata-rata nol, *varians* konstan, dan tidak saling berhubungan (non autokorelasi). Hipotesis yang digunakan dalam pengujian ini adalah pengujian akar unit, sebagai berikut:

H_0 : Data memuat akar unit (data tidak stasioner).

H_1 : Data tidak memuat akar unit (data stasioner).

Kriteria pengujian:

- Jika nilai mutlak *ADF Test Statistic* < nilai mutlak statistik *t* dengan nilai kritis 5% maka H_0 diterima dan H_1 ditolak.
- Jika nilai *p-value* > 0,05 maka H_1 diterima dan H_0 ditolak.

Bilamana *dataset* yang menjadi input model data *time series* tidak stasioner dalam hal rata-rata maka perlu dilakukan perubahan data dengan proses *differencing* atau pembeda supaya menghasilkan data yang stasioner dalam rata-rata. *Differencing* adalah sebuah perhitungan untuk mencari besar urutan perubahan nilai pada data *Time series*.

Metode *differencing* dilakukan dengan cara mengurangi nilai data pada suatu periode (Y_t) dengan nilai data periode sebelumnya (Y_{t-1}) dapat dinyatakan sebagai $Y_t - Y_{t-1}$. Jika hasil *differencing* pertama belum mendapat stasioneritas maka hasil *differencing* pertama ditransformasikan lagi, sehingga dihasilkan *differencing* kedua. Jika data tidak stasioner dalam hal varian maka perlu dilakukan transformasi data dengan tujuan untuk menstabilkan varian.

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

Fungsi autokorelasi atau *autocorrelation function* (ACF) merupakan hubungan (korelasi) antara suatu nilai variabel yang telah terjadi pada suatu periode dan yang terjadi pada periode berikutnya [18].

Grafik fungsi autokorelasi dan autokorelasi parsial digunakan untuk menentukan kestasioneran data *time series* yang digunakan, yaitu dengan melihat lag-lag nya yang turun secara eksponensial. Selanjutnya, pada proses *autoregressive* (AR)(*p*) grafik fungsi autokorelasi parsial digunakan untuk mendeteksi model awal dengan melihat fungsi *cut off* setelah lag- *k*, sedangkan pada proses *moving average* (MA)(*q*) grafik autokorelasi digunakan untuk mendeteksi model awal dengan melihat fungsi *cut off* setelah lag-*k* [19].

2.1.7. ARIMA (*Auto Regressive Integrated Moving Average*)

ARIMA sebenarnya merupakan usaha untuk mencari pola data yang paling cocok dari sekelompok data, sehingga metode ARIMA sepenuhnya memerlukan data historis dan data sekarang untuk menghasilkan ramalan jangka pendek [20].

Model linier time series secara umum ada dua jenis, yaitu model *time series* stasioner dan non stasioner. Model *time series* stasioner dilakukan tanpa proses *diffrencing* terlebih dahulu, model tersebut adalah:

1. Model *autoregressive* (AR) tingkat (p), model ini menunjukkan adanya keterkaitan antara suatu nilai pada waktu kini (Y_t) dengan nilai pada waktu sebelumnya Y_{t-k} , ditambah suatu nilai acak.
2. Model *moving average* (MA) tingkat (q), model ini menunjukkan adanya keterkaitan antara nilai pada waktu sekarang (Y_t) dengan residual pada waktu-waktu sebelumnya (α_{t-k} dengan $k = 1, 2, \dots$) [15].

Realitas dalam keseharian lebih banyak data yang non stasioner dibandingkan dengan data stasioner sehingga model time series *autoregressive integrated moving average* (ARIMA) tingkat (p, d, q) lebih populer dibanding model time series yang telah disebutkan sebelumnya [21]. Nilai d pada ARIMA adalah nilai *diffrencing* untuk menjadikan data non stasioner menjadi stasioner.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} \quad (2.2)$$

Dimana:

Y_{t-1} = Lag pertama dari series

β_1 = Koefisien lag pertama yang diperkirakan

α = Konstanta

ϵ = Nilai yang diharapkan untuk kesalahan random pada masa yang akan datang harus ditetapkan sama dengan nol

Secara sederhana,

Peramalan $Y_t = \text{Konstanta} + \text{kombinasi linier lag dari } Y \text{ [sampai lag } p] + \text{kombinasi linier dari lag error peramalan [sampai lag } q]$

ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik karena biasanya akan cenderung *flat* (mendatar/konstan) untuk periode yang cukup panjang, model ini secara penuh mengabaikan independen variabel dalam membuat peramalan model ini menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat [22].

Model ARIMA terdiri dari empat langkah dasar, yaitu tahap identifikasi, tahap estimasi parameter model, terakhir tahap pemeriksaan diagnostic dan peramalan. Berikut Tahapan dalam ARIMA.

Tahap Pertama – Identifikasi Model

Pada penelitian ini digunakan *ADF test* dengan hipotesa nol *time series* tidak stasioner yang mana jika *p-value* $< 0,05$ maka hipotesa nol ditolak dan data *time series* dianggap stasioner [21]. Langkah pengujiannya seperti berikut:

1. Hipotesis H_0 : Parameter model tidak signifikan
 H_1 : Parameter model signifikan
2. Area Penolakan *p-value* $< 0,05$ (level toleransi) tolak H_0

Identifikasi selanjutnya dengan melihat grafik fungsi autokorelasi dan autokorelasi parsial, yaitu dengan melihat lag-lag nya yang turun secara eksponensial. Selanjutnya, pada proses AR(*p*) grafik fungsi autokorelasi parsial digunakan untuk mendeteksi model awal dengan melihat fungsi *cut off* setelah lag-*k* , sedangkan pada proses MA(*q*) grafik autokorelasi digunakan untuk mendeteksi model awal dengan melihat fungsi cut off setelah lag-*k* [19].

Tabel 2.1. Pengamatan Plot ACF dan PACF [21]

Model	ACF	PACF
AR(<i>p</i>)	Turun secara eksponensial menuju nol dengan bertambahnya <i>k</i> (<i>dies down</i>)	Terpotong setelah lag <i>p</i>
MA(<i>q</i>)	Terpotong setelah lag <i>q</i> (<i>cut off</i> setelah lag <i>q</i>)	Turun secara eksponensial menuju nol dengan bertambahnya <i>k</i>

ARMA(p,q)	Turun secara eksponensial menuju nol dengan bertambahnya k	Turun secara eksponensial sejalan dengan bertambahnya k
---------------	--	---

Tahap Kedua - Estimasi Parameter Model

Tujuan Estimasi parameter dalam analisis *time series* untuk membentuk model yang baik dengan syarat parameter model harus signifikan atau *p-value* < 0,05. Metode estimasi parameter model yang dipakai adalah *conditional sum of squares* (css) dan *maximum likelihood estimation* (mle). Tahapan ini adalah untuk menguji kelayakan suatu model dengan pengujian signifikansi parameter dan kesesuaian asumsi, yang mana asumsi dasar yang harus terpenuhi adalah residual bersifat *white noise*.

Tahap Ketiga - Cek Diagnosa

Pemeriksaan ini dilakukan dengan menguji apakah data sudah *whitenoise* dan berdistribusi normal guna mendapatkan hasil peramalan yang baik [15]. Untuk mengetahui residual sudah memenuhi asumsi *whitenoise* dilakukan uji *Ljung-Box-Pierce* (LBQ). Cek diagnostik yang lain adalah menguji kenormalan residu menggunakan *Kolmogorov Smirnov*.

Tahap Keempat – Peramalan

Pada tahap terakhir ini data dibagi kedalam dua periode yaitu periode latih (*train*), periode uji (*test*) dan periode peramalan, pembentukan model dilakukan dengan menggunakan data latih.

Kriteria Model Terbaik

Setelah model diidentifikasi akan terbentuk lebih dari satu model yang dianggap sesuai, untuk itu digunakan *Mean Squared Error* (MSE) dan *Akaike Info Criterion* (AIC), dimana nilai terkecil menggambarkan nilai dari model terbaik [19].

Pemilihan Model Terbaik

Ada beberapa metode yang digunakan untuk menunjukkan kesalahan yang disebabkan oleh suatu teknik peramalan tertentu, hampir semua ukuran tersebut menggunakan beberapa fungsi dari perbedaan antara nilai sebenarnya dengan nilai peramalannya, perbedaan nilai sebenarnya dengan nilai peramalan ini biasanya disebut sebagai residual [22]. Dalam penelitian ini digunakan nilai RMSE, MAPE dan MPE untuk mengevaluasi model

2.2. Tinjauan Studi

Berikut dibawah ini adalah beberapa penelitian terkait dengan prediksi penjualan yang telah dilakukan oleh peneliti sebelumnya:

Penelitian tentang *time series* dalam penjualan yang mempertimbangkan pendekatan utama *machine learning*, *error* pada data *test* beberapa metode, sebagai berikut: Random Forrest 11.9%, Neural Network 11.3% dan ARIMA 11.4%[4].

Peramalan dengan ARIMA digunakan untuk dua buah objek data dilakukan [23] dalam kasus peramalan persediaan bahan baku untuk produksi, memberikan nilai error yang sangat rendah, yaitu 0.571%.

Analisis perbandingan antara metode ARIMA musiman dan *Support Vector Machine* (SVM) dalam peramalan dilakukan [5] terhadap *dataset* wisatawan mancanegara dengan membuat kategorisasi dalam tiga kuartil, didapatkan MAPE metode SARIMA 7,07% dan SVM 9,57%.

Metode ARIMA juga digunakan untuk memprediksi dan mengetahui karakteristik harga saham PT. Telkom oleh [24], didapatkan harga saham minimum dan maksimum saat itu adalah Rp. 7.099 Rp. 7.282, dengan nilai RMSE 91,16.

Penelitian yang dilakukan [25] dalam meramalkan kasus Demam Berdarah di Provinsi Jawa Timur, mendapatkan nilai MAPE sebesar 43,51%, cukup tinggi diantara penelitian yang lain.

Dalam penelitiannya [13] meramalkan data beban konsumsi listrik, peneliti membandingkan metode SARIMA dengan SAFRIMA yaitu metode ARIMA yang dengan nilai pecahan pada pembedanya (*differencing*), hasil yang diperoleh nilai MSE SARIMA = 5979,18 dan SARFIMA = 211833.

Perbandingan antara ARIMA dengan *Radial Based Function* (RBF) dilakukan oleh [7] dengan data set Indeks Harga Perdagangan Bebas Tahun 2006 – 2012. Didapatkan nilai error ARIMA terendah dibandingkan dengan RBF maupun ARIMA-RBF.

Perancangan system peramalan dibuat pada penelitian ini dalam kasus perusahaan kayu dengan menggunakan metode ARIMA, dimana hasil dari model terbaik memiliki nilai *error* kurang dari 5% [26]

Penilaian model terbaik pada SARIMA dengan parameter koefisien determinasi *adjusted* (R^2) 84,67% didapat dari kombinasi model ARIMA (1,1,0)(3,1,2)₆, dilakukan terhadap data Indeks Harga Konsumen Kota Palangka Raya pada tahun 2014 – 2017, penelitian dilakukan oleh [27].

Perbandingan keakuratan peramalan juga dilakukan, metode yang dibandingkan adalah ARIMA dengan *Exponential Smoothing* pada penjualan semen PT. Sinar Abadi, hasil yang didapat nilai MSE ARIMA lebih rendah dari *Exponential Smoothing*[28].

Penelitian dengan judul Peramalan Nilai Impor Besi Baja Indonesia 2008-2017 dengan Model Arima menerangkan kelebihan dan kekurangan metode ARIMA, dikatakan kelebihan dapat memodelkan berbagai macam pola data *time series* serta menghasilkan peramalan yang akurat untuk ramalan jangka pendek, sedangkan kelemahannya yaitu pada tahap identifikasi model terjadi subjektifitas, lalu kehandalan model juga tergantung dari keterampilan peneliti, dan pada peramalan jangka panjang kurang akurat [6].

Dibawah ini adalah penelitian terkait sebelumnya yang pernah dilakukan dan dijadikan sebagai referensi pada penelitian ini:

Tabel 2.2. Penelitian Terkait

NO	PENELITI	JUDUL	METODE	HASIL
1	Riza Rahmadayanti	Perbandingan Keakuratan Metode <i>Autoregressive Integrated Moving Average</i> (Arima) Dan <i>Exponential Smoothing</i> Pada Peramalan Penjualan Semen Di PT. Sinar Abadi[28]	ARIMA dan <i>Exponential Smoothing</i>	ARIMA lebih akurat karena memiliki nilai MSE lebih kecil dari metode <i>Exponential Smoothing</i>
2	Suhud Widiyanto	Peramalan Nilai Impor Besi Baja Indonesia 2008-2017 Dengan Model ARIMA[6]	ARIMA	Kelebihannya dapat memodelkan berbagai macam pola data deret waktu dan menghasilkan peramalan yang akurat untuk jangka pendek. Kelemahannya identifikasi model bersifat subjektif, kehandalan model tergantung dari keterampilan peneliti dan kurang baik peramalan jangka panjang
3	Molydah S	Analisis Perbandingan Implementasi Sarima Dan <i>Support Vector Machine</i> (SVM) Dalam Prediksi Jumlah Wisatawan Mancanegara[5]	ARIMA dan SVM	MAPE metode ARIMA = 7,07% dan SVM = 9,59%
4	Amira Herwindyani Hutasuhut, <i>et al.</i>	Pembuatan Aplikasi Pendukung Keputusan Untuk Peramalan Persediaan Bahan Baku Produksi Plastik <i>Blowing</i> dan <i>Inject</i> Menggunakan Metode ARIMA Di CV. Asia[23]	ARIMA	MAPE sebesar 0.52%

5	Bohdan, <i>et al</i>	<i>Machine-Learning Models For Sales Time Series Forecasting</i> [4]	ARIMA, NN, RF, Stacking	Error peramalan data test ARIMA lebih baik dari RF dan extra Tree
6	Djoni Hatidja	Penerapan Model Arima Untuk Memprediksi Harga Saham PT. Telkom Tbk. [24]	ARIMA	MSE 8310.124967 dan RMSE 91.15988683
7	Muhammad Bintang Pamungkas, <i>et al.</i>	Aplikasi Metode Arima Box-Jenkins Untuk Meramalkan Kasus Dbd Di Provinsi Jawa Timur [25]	ARIMA	MAPE 43,51%
8	Arina Dini Yuanti	Perbandingan Model Time Series Seasonal ARIMA dan Seasonal ARFIMA Pada Data Beban Konsumsi Listrik Jangka Pendek di Jawa Timur [13]	ARIMA dan ARFIMA	MSE ARIMA = 5979,18 dan ARFIMA = 211833
9	DT. Wiyanti, <i>et al</i>	Peramalan Deret Waktu Menggunakan Model RBF dan ARIMA[7]	ARIMA, RBF dan ARIMA-RBF	ARIMA memiliki nilai error paling kecil dibanding metode yang lain
10	Anjani Yuniarti	Perbandingan Metode Exponential Smoothing dengan ARIMA [29]	ARIMA dan Smoothing Ekspensial	MSE Smoothing Ekspensial = 562 dan ARIMA = 635
11	ANNISA UL UKHRA	Pemodelan Dan Peramalan Data Deret Waktu Dengan Metode Seasonal Arima [30]	ARIMA	MSE = 110,4
12	Ananto Wibowo	Model Peramalan Indeks Harga Konsumen Kota Palangka Raya Menggunakan Seasonal ARIMA (SARIMA) [27]	ARIMA	R adj = 84,67%
13	Puspa Linda, <i>et al.</i>	Peramalan Penjualan Produksi Teh Botol Sosro Pada PT. Sinar Sosro Sumatera Bagian Utara Tahun 2014 Dengan Metode ARIMA Box-Jenkins [31]	ARIMA	MAPE 1,37%

14	Mohamad As'ad, <i>et al.</i>	Peramalan Jumlah Mahasiswa Baru Dengan Model Autoregressive Integrated Moving Average (ARIMA)[32]	ARIMA	MSE=446.22 dan AIC = 144.02
15	Peter Sutanto, <i>et, al.</i>	Perancangan Sistem Forecasting di Perusahaan Kayu UD. 3G dengan Metode ARIMA[26]	ARIMA	Error pada model terbaik < 5%

Dari tinjauan studi tersebut diatas dapat disimpulkan bahwa metode ARIMA banyak digunakan untuk solusi peramalan.

Adapun *novelty* yang diperoleh dari penelitian ini adalah menggunakan data penjualan toko Teras Umat selama 1 tahun guna meramal penjualan agar meningkatkan minat belanja anggota dan membuka peluang investasi bagi anggota.

2.3. Tinjauan Objek Penelitian

Koperasi Teras Umat adalah sebuah minimarket atau toko modern yang didirikan oleh komunitas untuk memberdayakan Usaha Mikro Kecil dan Menengah (UMKM) yang berada di sekitaran Kota Administratif Depok khususnya kecamatan Cimanggis. Terdapat 303 orang anggota yang bergabung dengan Koperasi Teras Umat ini, dimana saat itu masing-masing anggota turut serta dalam memberikan modal awal yang nilai bervariasi.

Waktu operasional toko adalah buka pukul 07:00 WIB dan tutup pukul 22:00 WIB, toko beralamat di JL.Tugu Raya, Ruko Pondok Nurul Fikri, Kelurahan Tugu, Kecamatan Cimanggis, Kota Depok, Jawa Barat. Ada 4 orang karyawan yang dipekerjakan dalam operasional toko, sebagai berikut:

- Prima Tasbihi : Kepala Toko
- Noval Fabianto : Kasir/operator/penjaga toko
- Fahri Saputra : Kasir/operator/penjaga toko
- Nafiah : Kasir/operator/penjaga toko

Karyawan selain kepala toko bekerja secara shift, shift pertama pukul 07:00 WIB sampai 16:00 WIB, shift berikutnya pukul 14:00 WIB sampai 22:00 WIB,

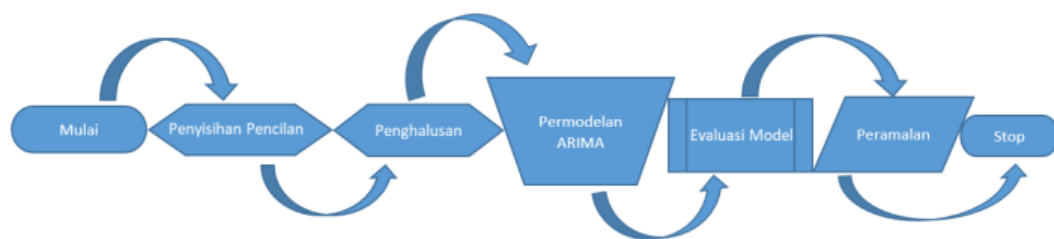
sedangkan kepala toko bekerja dari pukul 08:00 WIB sampai 17:00 WIB, berlaku pada semuanya 6 hari kerja dan 1 hari libur dalam 1 pekan.

2.4. Kerangka Pemikiran

Atas dasar tinjauan pustaka, tinjauan studi pustaka dan kondisi objek penelitian, penulis merancang kerangka pemikiran sebagai acuan yang akan dipakai pada penelitian ini, supaya penelitian dilakukan berjalan secara bertahap dan sesuai kerangka yang dibuat.

Terdiri dari beberapa bagian dan tahapan untuk mengetahui metode peramalan terbaik dalam memprediksi trend penjualan beberapa waktu kedepan. *Dataset* yang akan digunakan dalam penelitian ini adalah data penjualan periode bulan Mei 2018 – April 2019. Setelah data penjualan didapat maka selanjutnya adalah seleksi data, menyisihkan pencilan, melakukan penghalusan dan selanjutnya meramalkan trend penjualan dengan metode ARIMA.

Data yang sudah melewati tahap *pre-processing* maka akan dijadikan sebagai *dataset* yang siap digunakan untuk memprediksi penjualan menggunakan metode ARIMA. Setelah dilakukan *modelling* dengan ARIMA, maka dilakukan evaluasi menggunakan pengukuran nilai *error* yang didapat. Gambar 2.1 merupakan kerangka pemikiran yang akan diimplementasikan.



Gambar 2.1. Kerangka Pemikiran

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian ini adalah deskriptif, yaitu merupakan prosedur untuk membuat ikhtisar, menata, membuat grafik yang dilakukan dengan metode peramalan kuantitatif. Peramalan kuantitatif adalah metode peramalan yang melibatkan analisis statistik terhadap data masa lalu [33].

3.2 Sumber Data dan Variabel Penelitian

Dalam penelitian ini data yang diambil adalah data penjualan dari toko Teras Umat yang berlokasi di Ruko Nurul Fikri Jl. Tugu Raya, Kelurahan Tugu, Kecamatan Cimanggis, Kota Administratif Depok, Jawa Barat. Data penjualan yang digunakan adalah data penjualan toko dari mulai 1 Mei 2018 hingga 30 April 2019. Sebagai perbandingan antara data latih (*data train*) dan data uji (*data test*) digunakan 85% data *train* dan 15% data *test*.

3.3 Metode Peramalan

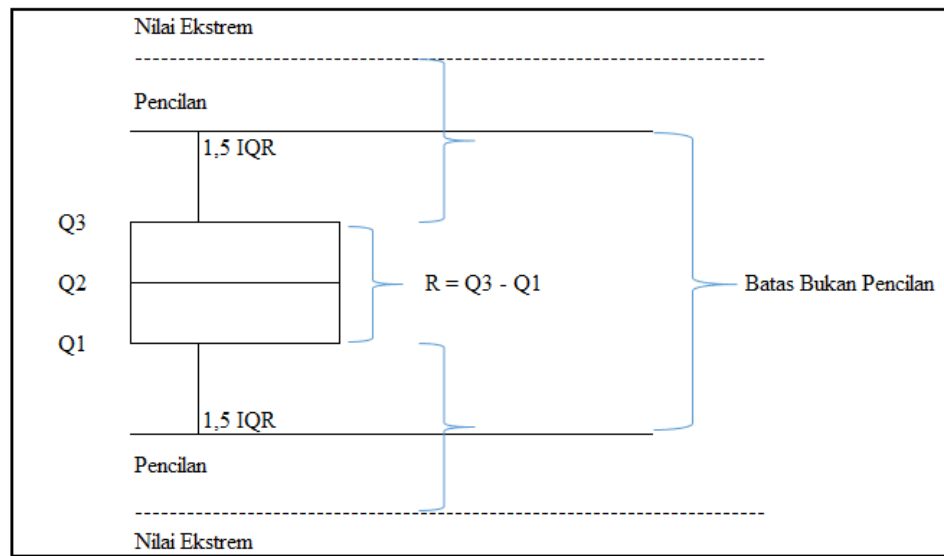
Peramalan dilakukan dengan menggunakan metode ARIMA (*Autoregressive Integrated Moving Average*). Sebelum dilakukan perhitungan dengan menggunakan metode ARIMA, terlebih dahulu dilakukan penyisihan pencilan dan penghalusan dengan menggunakan metode rata-rata bergerak (*moving average*).

3.3.1 Tahapan Penelitian

Langkah analisis yang akan dilakukan, sebagai berikut:

1. Menyisihkan pencilan yang ada pada data penjualan dengan menggunakan nilai kuartil dan batasan, secara singkat tahapannya sebagai berikut :
 - Membagi urutan data menjadi empat bagian kuartil: Kuartil-1 (Q_1), kuartil-2 (Q_2), dan kuartil-3 (Q_3).
 - Menentukan cakupan atau *IQR* (*Interquartile Range*) yang didefinisikan sebagai selisih Q_1 terhadap Q_3 atau $IQR = Q_3 - Q_1$.

- Pencilan dapat ditentukan dari nilai $(x) < 1.5 \cdot IQR$ terhadap $Q1$ sebagai batas bawah dan nilai $(x) > 1.5 \cdot IQR$ terhadap $Q3$ sebagai batas atas.
- Melakukan analisis terhadap data penjualan yang dinaggap pencilan. Skema pada gambar 3.1 merupakan ilustrasi pencilan pada sebuah data set.



Gambar 3.1. Pencilan Dalam Dataset

2. Melakukan penghalusan dengan rata-rata bergerak (*Moving Average*). Pada prinsipnya penghalusan adalah menghaluskan variasi lokal atau gejolak pendek pada suatu deret waktu dalam hal ini dilakukan penghitungan rata-rata bergerak (MA) [34].

Secara sederhana proses dalam melakukan *moving average* adalah sebagai berikut:

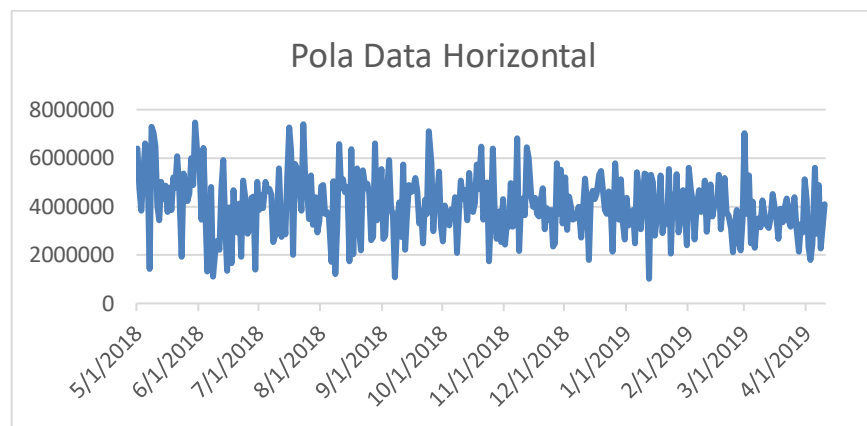
- Menggunakan n nilai data terbaru dalam suatu deret berkala untuk meramalkan periode yang akan datang.
- Rata-rata perubahan atau pergerakan sebagai observasi baru.
- Perhitungan rata-rata bergerak adalah :

$$MA = \frac{\sum(n \text{ nilai data terbaru})}{n} \dots\dots\dots (3.1)$$

Dimana n adalah jumlah data yang dijadikan rata-rata

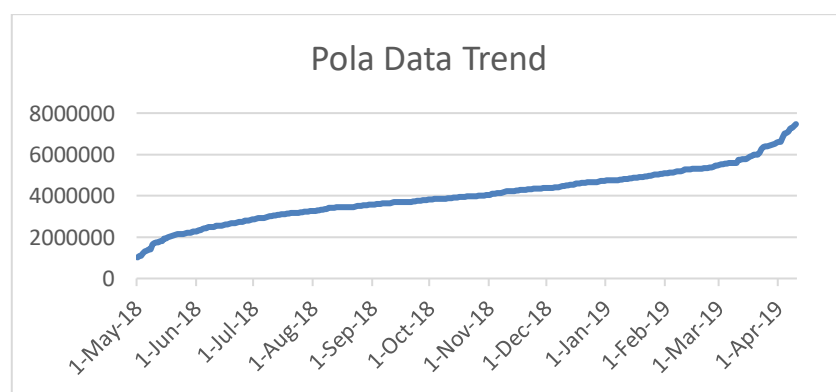
3. Melakukan identifikasi pola data dan karakteristik data penjualan toko, identifikasi dilakukan dengan mengamati bentuk pola data *time series*, pola data terbagi menjadi empat jenis[22], sebagai berikut :

- Data berpola horizontal, yaitu data yang relatif konstan dilihat dari plot data yang menyebar terhadap rata-rata (*mean*) dan ragam (*variance*), gambar 3.2 adalah contoh grafik pola data horizontal.



Gambar 3.2. Pola Data Horizontal

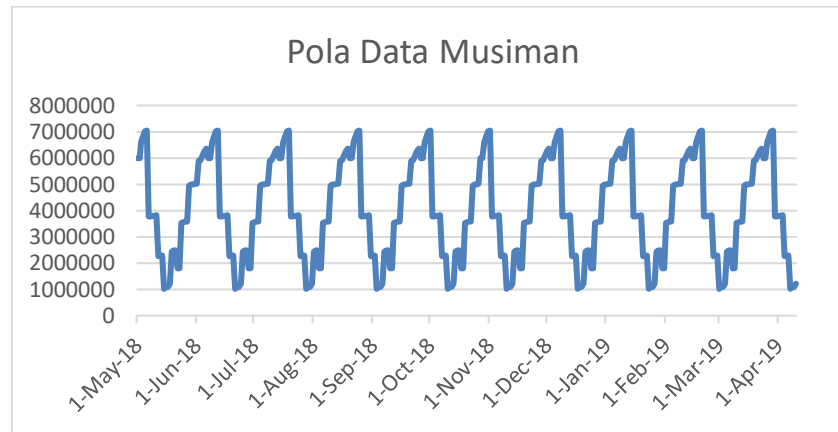
- Data berpola trend, plot data menyebar membentuk suatu kecenderungan menaik ataupun menurun secara konstan, gambar 3.3 merupakan contoh grafik pola data trend.



Gambar 3.3. Pola Data Trend

- Data yang berpola musiman (*seasonal*), data ini merupakan data berpola trend yang dibentuk oleh faktor musiman dan cenderung

berulang namun teratur, gambar 3.4 adalah contoh grafik pola data musiman (*seasonal*).



Gambar 3.4. Pola Data Musiman (*seasonal*)

- Pola siklus, yaitu pola data berbentuk trend yang memuat musiman namun periode musimannya lebih panjang.

4. Melakukan permodelan ARIMA

3.3.2 Permodelan ARIMA

Melakukan peramalan dengan metode ARIMA dengan langkah-langkah sebagai berikut:

1. Melakukan identifikasi kestasioneran data dengan melakukan (*Augmented Dickey Fuller*) ADF test, dimana hipotesis nol H_0 : Data tidak stasioner dan tingkat signifikansi yang dipakai 95% atau $p\text{-value} < 0,05$. Selanjutnya mengamati plot *time series* untuk memeriksa kestasioneran data penjualan, stasioner dalam *mean* maupun *variance*, bilamana nilai $p\text{-value} > 0,05$ dilakukan *differencing* untuk menentukan orde d , proses berikutnya memeriksa plot ACF dan PACF untuk menentukan orde p, q .
2. Jika data tidak stasioner maka harus dilakukan proses *differencing* sebanyak $I(d)$, d adalah jumlah *differencing* yang dilakukan hingga data mencapai nilai stasioner. Notasi yang dipakai dalam melakukan *differencing* adalah operator *shift* mundur atau *backward shift* (B), rumus nya sebagai berikut :

$$B X_t = X_{t-1} \dots\dots\dots (3.2)$$

Dimana: B = Pembeda (*Difference*)
 X_t = Nilai X pada ordo ke t
 X_{t-1} = Nilai X pada ordo ke $t-1$

Ketika deret stasioner dengan model umum ARIMA (0,d,0), *difference* menjadi :

$$(1 - B)^d X_t = e_t \dots\dots\dots (3.3)$$

Dimana: $(1 - B)^d X_t$ = diferensiasi orde ke d
 e_t = nilai *error*

Melakukan estimasi dan uji parameter dengan melihat signifikasi terhadap model. Proses *Autoregressive Integrated Moving Average* yang dilambangkan dengan ARIMA (p, d, q)

Dimana: p menunjukkan ordo/ derajat autoregressive (AR)
 d adalah tingkat proses differencing
 q menunjukkan ordo/derajat moving average (MA)

- Rumus autoregressive (AR) (p) adalah sebagai berikut :

$$X_t = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_p X_{t-p} + e_t \dots\dots\dots (3.4)$$

Dimana: X_t = data periode ke t
 β_p = parameter AR ke p
 $X_{t-2} + \dots + X_{t-p}$ = variable bebas (nilai masa lalu dari deret tersebut)
 e_t = nilai error pada t

Persamaan diatas dapat ditulis:

$$a(B)X_t = e_t \dots\dots\dots (3.5)$$

$$\text{Dimana } a(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p \dots\dots\dots (3.6)$$

- Rumus moving average (MA) (q) adalah sebagai berikut:

$$X_t = e_t - b_1 e_{t-1} - b_2 e_{t-2} - \dots - b_q e_{t-q} \dots\dots\dots (3.7)$$

Dimana: X_t = data periode ke t
 b_q = parameter MA ke q

$e_{t-1}, e_{t-2}, \dots, e_{t-q}$ = variable bebas (nilai masa lalu dari deret tersebut)

e_t = nilai error pada t

Persamaan diatas dapat ditulis:

$$b(B)X_t = e_t \dots\dots\dots (3.8)$$

$$\text{Dimana } b(B) = 1 + b_1B + b_2B^2 + \dots + b_qB^q \dots\dots\dots (3.9)$$

- Dari ketiga persamaan diatas didapat rumus autoregressive integrated moving average (ARIMA) (p,d,q) adalah sebagai berikut:

$$a_p(B)(1-B)^dX_t = b_0 + b_q(B)e_t \dots\dots\dots (3.10)$$

Dimana

$$a_p(B) = 1 - a_1B - a_2B^2 - \dots - a_pB^p \dots\dots\dots (3.11)$$

adalah operator AR

$$b_q(B)e_t = 1 - b_1B - b_2B^2 - \dots - b_pB^p \dots\dots\dots (3.12)$$

adalah operator MA

Dalam penelitian ini, pengujian parameter setiap model yang diuji harus diamati nilai AIC (*Akaike Information Criteria*) dan *p-value* nya, nilai AIC harus paling kecil dari semua parameter model yang uji dan nilai *p-value* harus tetap signifikan pada setiap parameter (AR dan MA).

3. Melakukan uji diagnostik untuk menguji kenormalan residual.
 - *Standardized residual* adalah residual yang distandardkan, keuntungan menggunakan *standardized residual* ini adalah tidak terpengaruh terhadap unit ukur, karena semua distandardkan. uji homoskedastisitas adalah uji keseragaman varian residual, jika varian bernilai konstan model tersebut dianggap baik.
 - *Plot Q-Q* adalah plot probabilitas, yaitu metode grafis yang dipakai untuk membandingkan dua distribusi probabilitas dengan memplot kuantil mereka satu sama lain[38].
 - *Correlogram* adalah plot antara *lag k* dengan *rk* dengan selang kepercayaan, dimana adalah autokorelasi sampel pada *lag k*. uji non-korelasi bertujuan mengetahui adanya korelasi antara data residual, suatu model yang baik nilai-nilai residualnya tidak saling berkorelasi.

- Histogram dapat digunakan untuk membandingkan distribusi data ke model teoretis, seperti distribusi normal, hal ini membutuhkan penggunaan skala kerapatan untuk sumbu vertikal. Plot kepadatan dapat dianggap sebagai plot histogram yang dihaluskan[38].
4. Melakukan peramalan atau *forecasting*.
 5. Mengukur kesalahan peramalan, ada beberapa metode yang dipakai untuk mengevaluasi peramalan, dalam penelitian ini digunakan data *test* sebanyak 15% terakhir, teknik yang dipakai untuk mengevaluasi peramalan yang dipakai antara lain :

- RMSE (*Root Mean Squared Error*) atau akar kesalahan rata-rata kuadrat. Nilai RMSE rendah menunjukkan bahwa variasi nilai yang dihasilkan oleh suatu model prakiraan mendekati variasi nilai obeservasinya. RMSE menghitung seberapa berbedaanya seperangkat nilai. Semakin kecil nilai RMSE, semakin dekat nilai yang diprediksi dan diamati[36].

$$RMSE = \sqrt{\frac{\sum y_1 - Y_t^A}{n}} \dots\dots\dots (3.13)$$

y adalah nilai hasil observasi

$\hat{y}$ adalah nilai hasil prediksi

n adalah jumlah data

- MAPE (*Mean Absolute Percentage Error*) atau nilai rata-rata persentase kesalahan peramalan[37].

$$MAPE = \frac{\sum_{t=1}^T \frac{|Y_t - Y_t^A|}{Y_t}}{T} \times 100\% \dots\dots\dots (3.14)$$

T adalah banyaknya periode peramalan

Y_t adalah nilai actual pada peroid ke t

Y_t^A adalah nilai ramalan pada peroid ke t

Tabel 3.1 Tingkat Akurasi Peramalan Berdasarkan MAPE [5].

MAPE	HASIL PERAMALAN
<10%	Sangat Baik
10-20%	Baik
20-50%	Layak Cukup
>50%	Buruk

Nilai MAPE merupakan acuan dalam menentukan akurasi suatu peramalan, nilai MAPE dikatakan sangat baik jika persentase kesalahan atau *error* dari hasil peramalan kurang dari 10%, namun suatu peramalan dianggap buruk jika persentase kesalahannya lebih dari 50% begitupun yang lainnya.

- MPE (*Mean Percentage Error*) atau rata-rata persentase kesalahan

$$MPE = \sum_{t=1}^n \frac{(Y_t - Y_t')}{n} \dots\dots\dots (3.15)$$

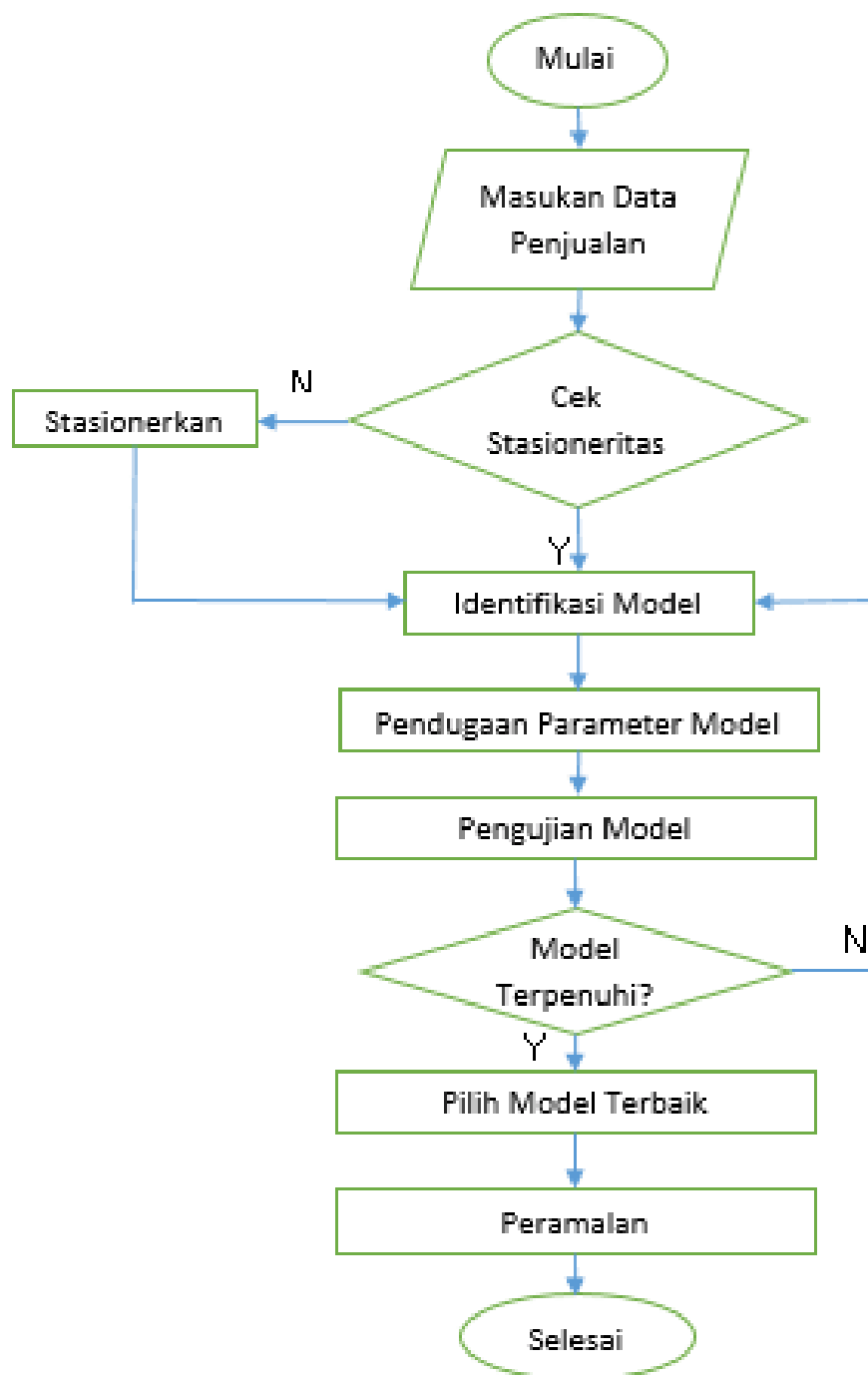
n adalah banyaknya peramalan

Y_t adalah nilai actual pada periode ke t

Y_t' adalah nilai ramalan pada periode ke t

MPE diperlukan untuk menentukan apakah suatu metode peramalan bias atau tidak. Jika pendekatan peramalan tersebut tidak bias, maka hasil perhitungan MPE akan menghasilkan persentase mendekati nol.

3.3.3. Flowchart ARIMA



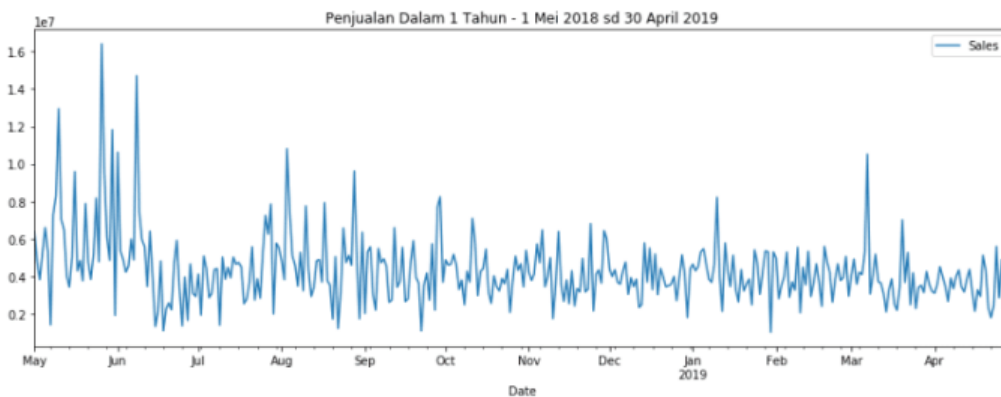
Gambar 3.5. Flowcart ARIMA

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Karakteristik Data Penjualan

Penelitian ini menganalisis data penjualan dari mulai tanggal 1 Mei 2018 hingga 30 April 2019, keseluruhan data berjumlah 364 hari, berkurang satu hari dikarenakan pada hari raya lebaran tahun 2018 atau tanggal 15 Juli 2018 toko diliburkan.



Gambar 4.1. Plot Data Penjualan Toko tanggal 1 Mei 2018 hingga 30 April 2019

4.2. Penyisihan Pencilan

Pada beberapa hari tertentu terjadi lonjakan penjualan yang mengakibatkan adanya pencilan, maka perlu diadakan penyisihan pencilan dengan menggunakan metode kuartil dan batasan.

Sebelum melakukan deteksi pencilan, dicari dahulu $Q1$ dan $Q3$ dari dataset, didapatkan nilai IQR yang berasal dari selisih antara $Q3$ dan $Q1$. Tabel 4.2 adalah hasil perhitungan kuartil.

Tabel 4.1 Kuartil

Q1	Q3	IQR
3,269,466	5,032,344	1,762,878.50

Setelah diketahui nilai IQR maka dapat bisa ditentukan batas atas dan batas bawah untuk menyisihkan pencilan dari dataset penjualan, adapun rumusan dalam menentukan batas bawah dan batas atas adalah sebagai berikut:

$$\text{Batas Bawah} = Q1 - 1.5 * IQR \dots\dots\dots (4.1)$$

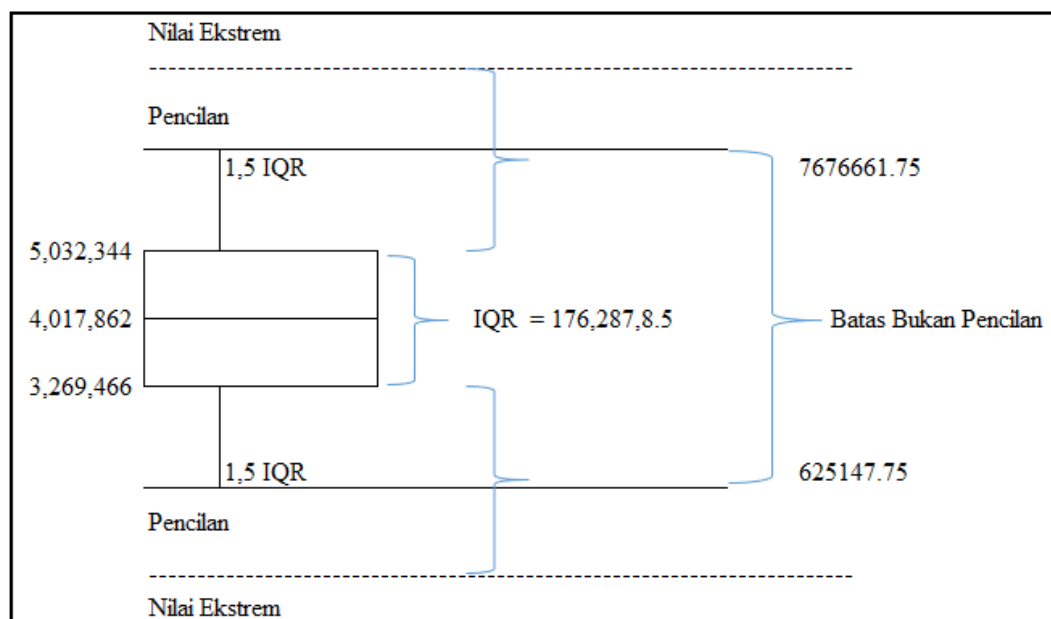
$$\text{Batas Atas} = Q3 + 1.5 * IQR \dots\dots\dots (4.2)$$

Diperoleh

Batas Bawah = 625,147.75 dan

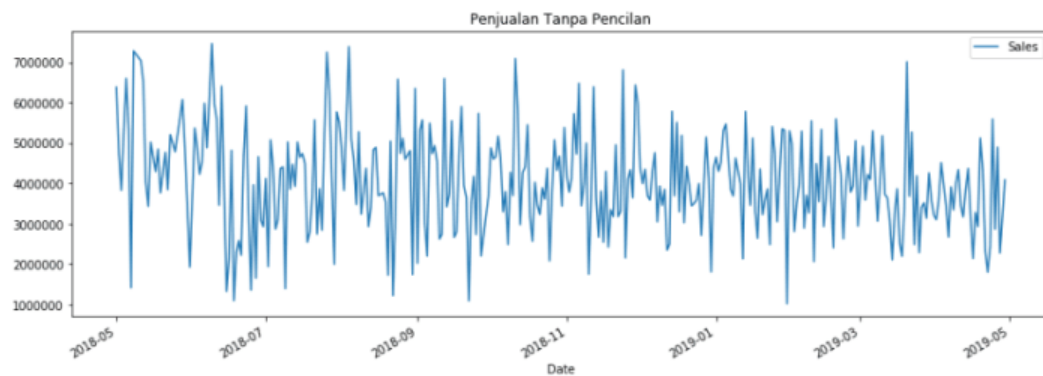
Batas Atas = 7,676,661.75.

Gambar 4.2 adalah skema batas atas dan bawah pencilan serta *IQR*, *Q1*, *Q2* dan *Q3* dari data penjualan selama satu tahun atau 365 hari mulai 1 Mei 2018 sampai 30 April 2019.



Gambar 4.2. Skema Pencilan Dari Data Penjualan Toko

Setelah dilakukan penyisihan pencilan terlihat data penjualan lebih rata atau tidak terdapat nilai penjualan harian yang terlalu besar atau terlalu kecil, seperti terlihat pada plot gambar 4.3.



Gambar 4.3. Plot Data Penjualan Setelah dilakukan Penyisihan Pencilan

Terdapat 20 hari atau 5% dari 365 hari yang akan disisihkan, setelah dilakukan penyisihan pencilan data yang tersisa sebanyak 345 hari penjualan, tabel 4.2 merupakan tabel yang berisi data pencilan dan keterangannya.

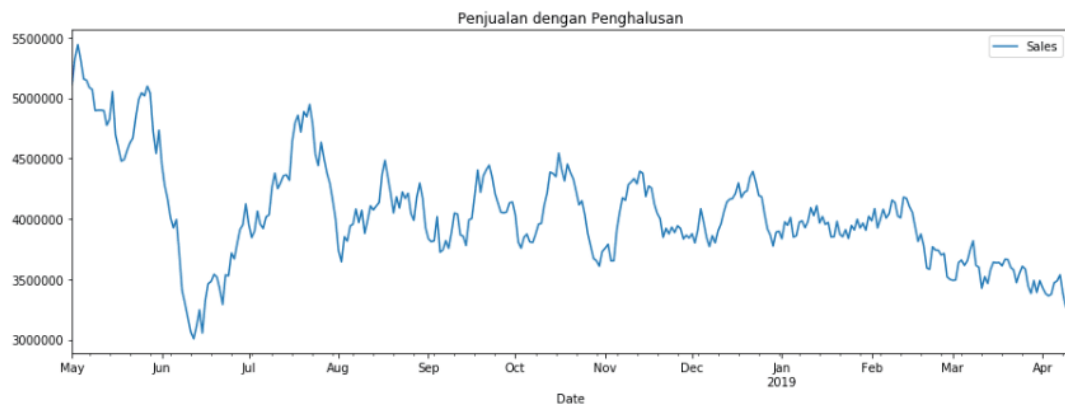
Tabel 4.2. Data Pencilan

Tgl	Nilai Transaksi (Rp) / Hari	Keterangan Pencilan
9 Mei 2018	8,283,700	Terdapat penjualan Beras @5 Kg sebanyak 85 karung
10 Mei 2018	12,947,000	Terdapat penjualan Beras @5 Kg sebanyak 80 karung
26 Mei 2018	16,406,520	Terdapat penjualan sembako dalam jumlah besar (Minyak goreng, Gula pasir, Beras, Tea, Terigu, Margarin, Kecap dan lain-lain)
8 Juni 2018	14,708,840	Terdapat penjualan Beras sebanyak 200 karung
4 Agustus 2018	10,812,598	Terdapat beberapa Pembayaran listrik dengan jumlah besar
11 Agustus 2018	7,757,995	Terdapat Pembayaran listrik dengan jumlah besar
29 Agustus 2018	9,624,547	Terdapat Pembayaran BPJS dengan jumlah besar
11 Januari 2019	8,219,110	Terdapat Pembayaran BPJS dengan jumlah besar
8 Maret 2019	10,521,704	Terdapat penjualan sembako dalam jumlah besar (Minyak goreng, Gula pasir, Beras, Mi instan dan lain-lain)
16, 20, 24, 27, 30 Mei, 1 Juni, 29 Juli, 18 Aug, 29, 30 Sep 2018	Rata-rata 8.957.397;-/hari	Terjadi lonjakan penjualan

4.3. Penghalusan

Berdasarkan gambar 4.3 diatas terlihat deret waktu terlihat data penjualan sangat fluktuatif, sehingga diperlukan adanya penghalusan atau *smoothing*, penghalusan dapat digunakan dalam dua hal, pertama sebagai peramalan dan kedua digunakan untuk mengurangi atau menghilangkan gejolak pendek pada suatu deret waktu [29].

Penghalusan dalam penelitian ini menggunakan rata-rata bergerak dengan jendela atau window setiap 15 hari, gambar 4.4 dibawah adalah hasil dari penghalusan deret waktu pada penelitian ini.



Gambar 4.4. Plot Data Penjualan dengan penghalusan

4.4. Permodelan ARIMA

Setelah tahapan *pre-processing* dilakukan, selanjutnya adalah melakukan permodelan ARIMA, adapun langkah-langkah yang dikerjakan sebagai berikut:

1. Identifikasi stasioneritas

Penelitian ini menggunakan *ADF test* dengan hipotesa nol (H_0) *time series* tidak stasioner yang mana jika *p-value* $< 0,05$ maka hipotesa nol ditolak dan *time series* dianggap stasioner [g].

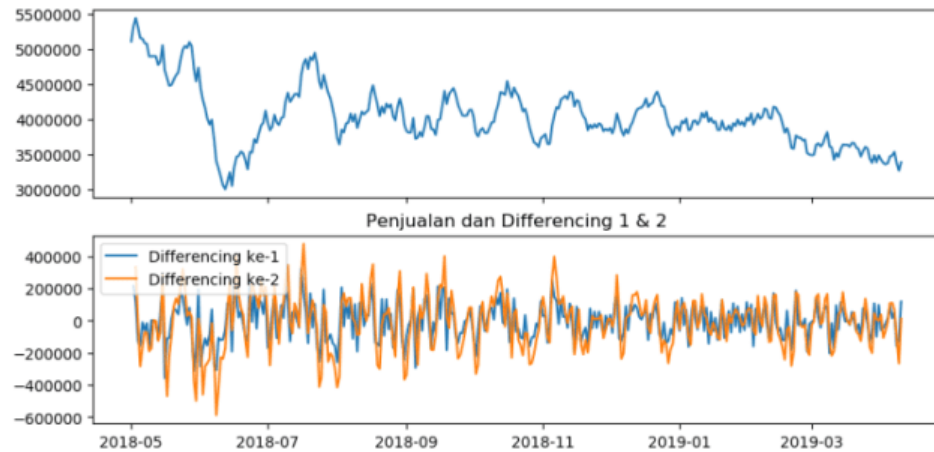
Hipotesis H_0 : Parameter model tidak signifikan

H_1 : Parameter model signifikan

Area Penolakan *p-value* $< 0,05$ (level toleransi) tolak H_0

Setelah dilakukan perhitungan terhadap data penjualan didapatkan nilai ***ADF Statistic*** = -2.56628 dan ***p-value*** = 0.100169 sehingga perlu

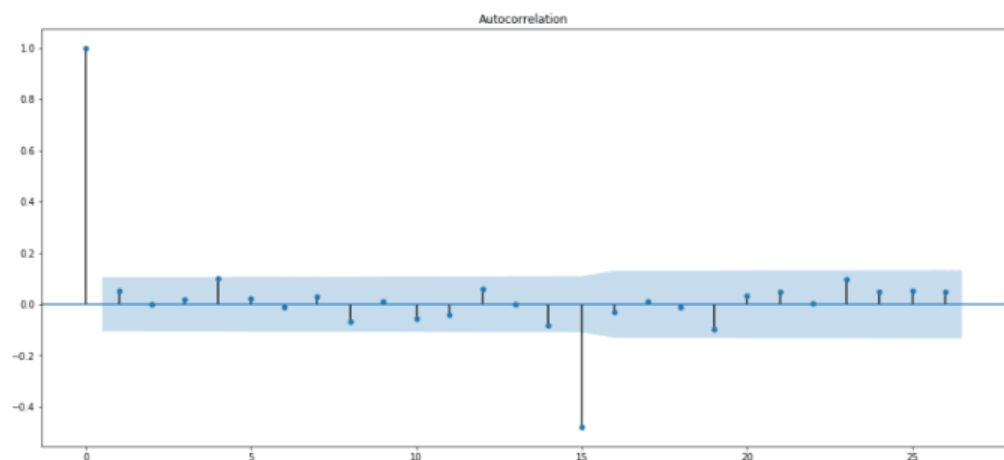
dilakukan *differencing*, gambar 4.5 merupakan visualisasi dari plot data penjualan original dan plot hasil dari *differencing* pertama dan kedua.



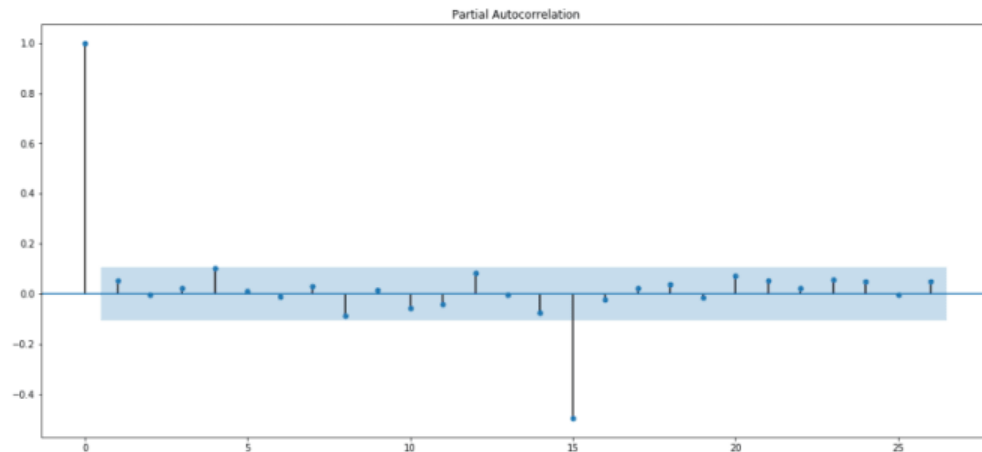
Gambar 4.5. Plot *differencing* 1 & 2

Dari plot tersebut terlihat *differencing* kedua lebih fluktuatif ke negative maupun positif, untuk menghindari *over differenced* yang bisa mempengaruhi parameter maka *differencing* yang digunakan satu saja atau $(d)(1)$.

Setelah dilakukan differencing dilanjutkan dengan menduga orde untuk AR dan MA dengan mengamati plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) yang berpatokan dari tabel 3.2, gambar 4.6 dan gambar 4.7 adalah plot ACF dan PACF.



Gambar 4.6. Plot ACF



Gambar 4.7. Plot PACF

Dari plot ACF dan PACF diatas didapat dugaan sementara orde 1 dipakai untuk nilai AR dan MA $(p,q)(1,1)$ karena lag pertama sudah memasuki wilayah signifikansi.

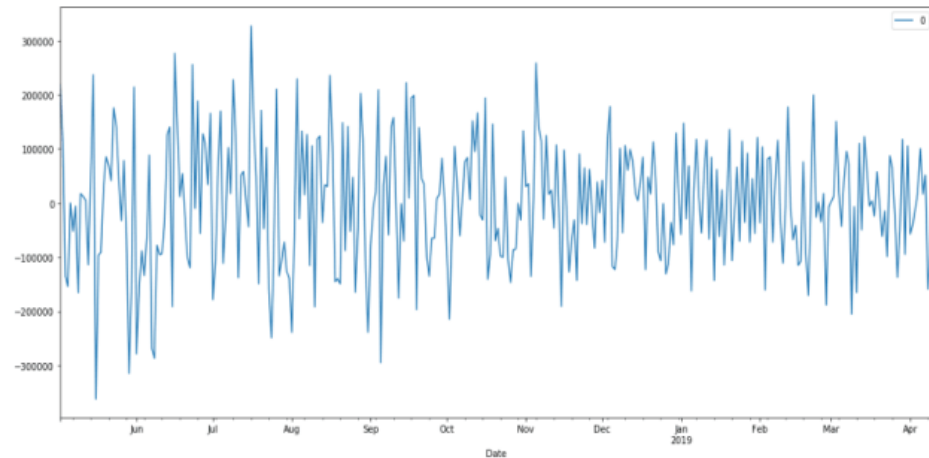
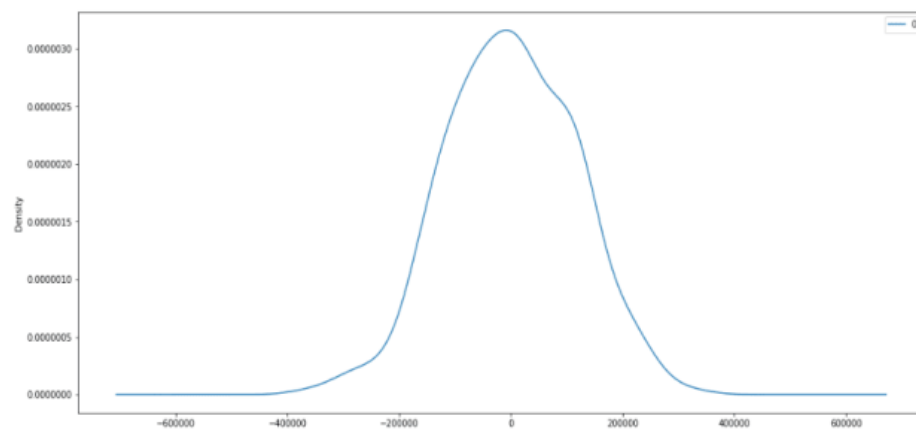
2. Membuat Model ARIMA

Dari langkah-langkah sebelumnya didapatkan dugaan sementara parameter untuk peramalan data penjualan dengan ARIMA (p,d,q) bernilai AR (p) (1), MA (q) (1) dan differencing (d) (1).

Nilai ARMA didapat dari analisis polt PACF dan ACF sedangkan nilai differencing q (1) didapat dari identifikasi p -value dengan *ADF test*. Dugaan sementara digunakan model ARIMA (p,d,q) (1,1,1).

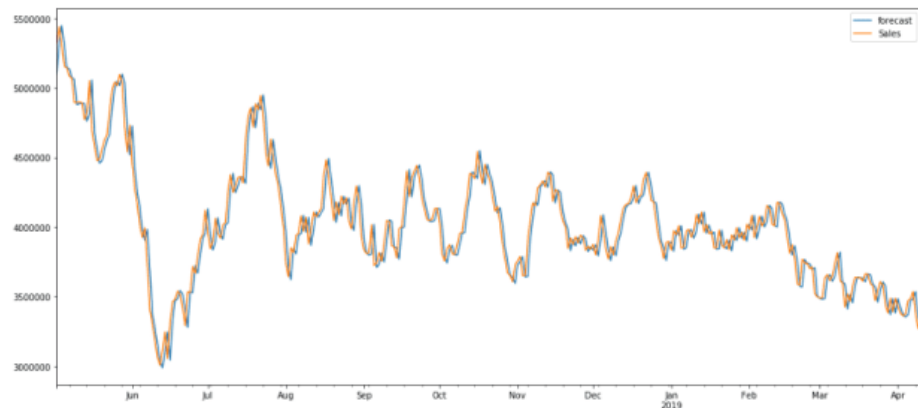
Hasil dari ARIMA (1,1,1) didapatkan nilai AIC untuk model tersebut sebesar 9002,171 dengan nilai p -value $< 0,05$ untuk AR (1) dan MA (1). Selanjutnya dilakukan lagi percobaan dengan parameter berbeda, untuk ARIMA (1,1,2) nilai AIC = 9004.540 dan p -value untuk semua AR dan MA = 0,9, untuk ARIMA (2,1,1) nilai AIC = 9004.137 dan p -value masih tinggi untuk MA (2).

Dari hasil diatas, ARIMA (1,1,1) memberikan hasil yang paling baik sejauh ini. Selanjutnya melihat plot residual untuk memastikan konsistensi *mean* dan *variance* pada gambar 4.8 dibawah.

Gambar 4.8. Plot *Residual*Gambar 4.9. Plot *Density*

Dari plot *residual* dan *density* diatas *mean* berada disekitar nol dan *variance* terlihat seragam.

Berikutnya membuat plot antara peramalan dengan penjualan aktual, seperti pada gambar 4.10 dibawah ini.



Gambar 4.10. Plot Peramalan dan Aktual

Plot pada gambar 4.10 adalah visualisasi model dilatih dari penjualan hari sebelumnya untuk meramalkan hari berikutnya. Namun untuk validitas, sebuah model membutuhkan kemampuan untuk peramalan kedepan serta membandingkan hasil peramalan dengan nilai sesungguhnya.

3. Uji Estimasi & Validasi Silang

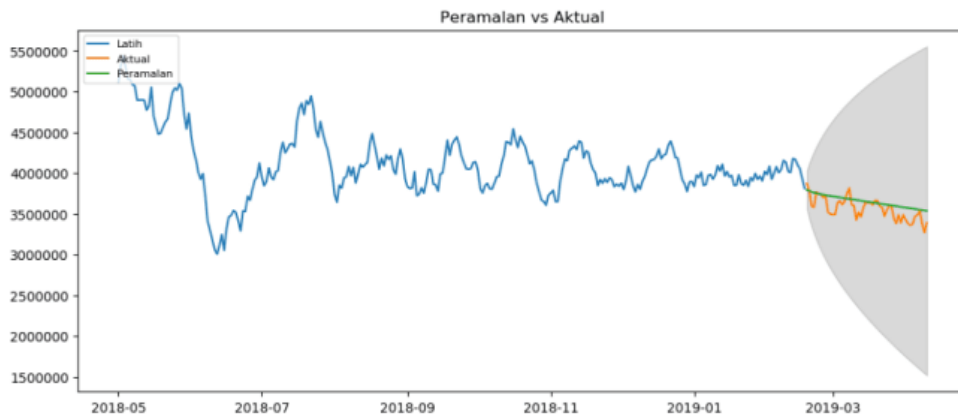
Diperlukan validasi silang untuk mengetahui kemampuan model dalam melakukan peramalan, validasi dalam deret waktu tidak diambil secara acak tetapi harus berurut, *series* data penjualan dibagi dua, sebagai berikut:

- Rentetan awal sebagai data latih sebanyak 85%
- Rentetan akhir sebagai data uji sebanyak 15%

Atau dalam penelitian ini terdapat:

- 293 hari sebagai data latih
- 52 hari sebagai data uji.

Plot dari ARIMA (1,1,1) dengan validasi silang terlihat seperti pada gambar 4.11.



Gambar 4.11. Plot Peramalan ARIMA (1,1,1).

Dari grafik plot pada gambar 4.11, ARIMA (1,1,1) terlihat arah dari peramalan cukup baik dan nilai hasil pengamatan berada dalam selang kepercayaan 95%, tetapi beberapa kombinasi parameter masih bisa dicoba agar mendapatkan nilai AIC dan nilai MAPE yang paling rendah, model paling baik adalah model dengan nilai persentase *error* yang paling rendah dan nilai p-value yang signifikan. Tabel 4.2.2 adalah hasil dari percobaan beberapa kombinasi parameter dari model ARIMA.

Tabel 4.3 Hasil Percobaan

Model	Nilai AIC	P-Value		Signifikansi
ARIMA(1, 1, 1)	7662.06	AR (1)	0.001	Signifikan
		MA (1)	0.004	
ARIMA(1, 1, 2)	7664.05	AR (1)	0.001	Ada tidak signifikan
		MA (1)	0.004	
		MA (2)	0.922	
ARIMA(1, 1, 3)	7664.87	AR (1)	0.005	Ada tidak signifikan
		MA (1)	0.011	
		MA (2)	0.653	
		MA (3)	0.266	
ARIMA(2, 1, 1)	7663.558	AR (1)	0.000	Ada tidak signifikan
		AR (2)	0.152	
		MA (1)	0.000	
ARIMA(2, 1, 2)	7653.283	AR (1)	0.020	Signifikan
		AR (2)	0.000	

		MA (1)	0.000	
		MA (2)	0.000	
ARIMA(2, 1, 3)	7659.592	AR (1)	0.000	Ada tidak signifikan
		AR (2)	0.000	
		MA (1)	0.000	
		MA (2)	0.000	
		MA (3)	0.151	
ARIMA(3, 1, 1)	7664.953	AR (1)	0.024	Ada tidak signifikan
		AR (2)	0.601	
		AR (3)	0.276	
		MA (1)	0.041	
ARIMA(3, 1, 2)	7639.772	AR (1)	0.000	Signifikan
		AR (2)	0.000	
		AR (3)	0.010	
		MA (1)	0.000	
		MA (2)	0.000	
ARIMA(3, 1, 3)	7641.695	AR (1)	0.000	Ada tidak signifikan
		AR (2)	0.004	
		AR (3)	0.387	
		MA (1)	0.000	
		MA (2)	0.000	
		MA (3)	0.778	

ARIMA (3, 1, 2) memberikan nilai AIC sangat baik, yaitu 7639.772 serta $p\text{-value} < 0,05$ atau signifikan untuk semua parameter, sehingga model ini layak untuk dipertimbangkan sebagai model peramalan.

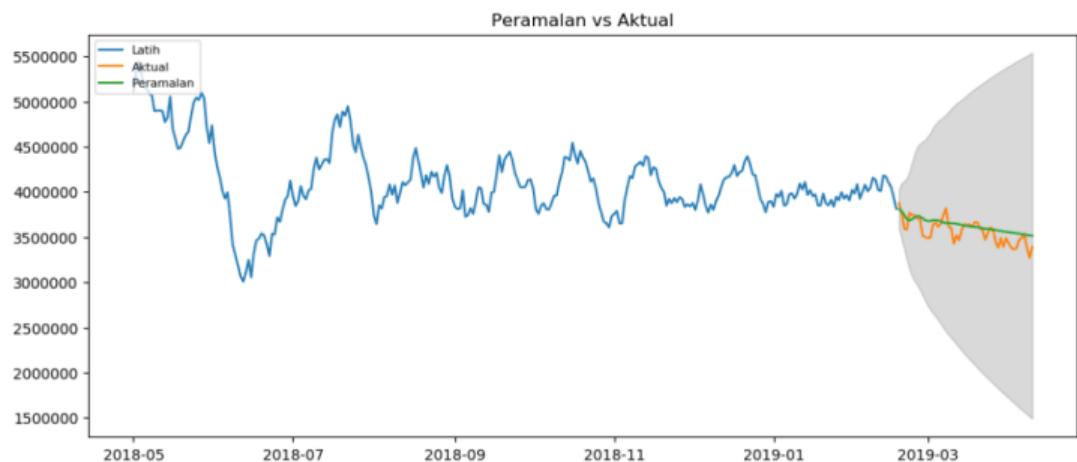
Setelah mendapatkan model yang paling layak berdasarkan nilai AIC dan signifikansi nilai $p\text{-value}$, dilakukan pengukuran kesalahan dengan mengamati nilai MAPE, MPE dan RMSE. Tabel 4.4 adalah hasil pengukuran dari semua model yang diuji.

Tabel 4.4. Hasil Pengukuran Kesalahan Model

Model	MAPE	MPE	RMSE
ARIMA(1, 1, 1)	0.02919	0.02567	128872.39000
ARIMA(1, 1, 2)	0.02959	0.02617	130147.88005
ARIMA(1, 1, 3)	0.02573	0.02019	115760.46519
ARIMA(2, 1, 1)	0.03779	0.03597	156921.37960
ARIMA(2, 1, 2)	0.03010	0.02521	130394.33900
ARIMA(2, 1, 3)	0.03270	0.02980	140582.02557

ARIMA(3, 1, 1)	0.02649	0.02172	119169.33710
ARIMA(3, 1, 2)	0.02561	0.01900	112883.14236
ARIMA(3, 1, 3)	0.02521	0.01813	111168.53618

Gambar 4.12 merupakan plot dari model ARIMA (3,1,2) dengan validasi silang 85% data latih dan 15% data uji.



Gambar 4.12. Plot Peramalan dari Model ARIMA (3,1,2)

Dalam pengukuran untuk membandingkan dua buah nilai, dalam hal ini peramalan penjualan dengan aktual penjualan, selain MAPE, MPE dan RMSE bisa juga digunakan korelasi dan *min-max error*, karena MAPE, MPE, korelasi dan *min-max error* mempunyai variasi penilaian antara 0 sampai 1 lalu dengan mempersentasikan MAPE, MPE, korelasi dan *min-max error* akan lebih mudah memberi penilaian.

Nilai korelasi dan *min-max error* dari Model ARIMA (3,1,2) adalah 0.7362432 atau sebesar 74% untuk korelasi dan 2.5% atau sebesar 0.02466715 untuk *min-max error*.

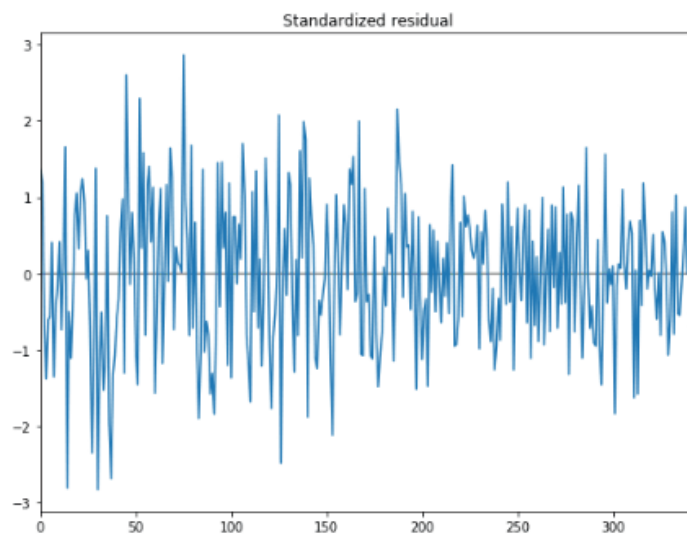
Selain pengukuran dengan range 0 – 1 dapat juga digunakan metric pengukuran dengan satuan berdasarkan data penjualan dalam hal ini Rupiah, berikut pengukuran dari:

- *Mean Absolute Error* (MAE) = 89503.4898
- *Mean Error* (ME) = 64722.1984
- *Root Mean Squared Error* (RMSE) = 112883.14236

4. Menguji diagnostik Model ARIMA

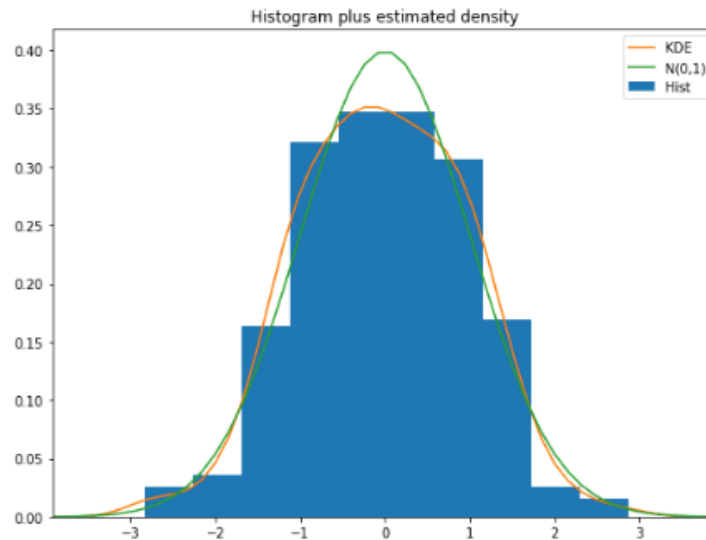
Uji diagnostik bisa dilakukan dengan menginterpretasikan plot residual yang ada pada beberapa gambar dibawah ini.

- *Standardized residual*: Menggambarkan *error* residual berfluktuatif disekitaran rata-rata angka nol dan varian yang cukup seragam, uji homoskedastisitas adalah uji keseragaman varian residual, jika varian bernilai konstan model tersebut dianggap baik.



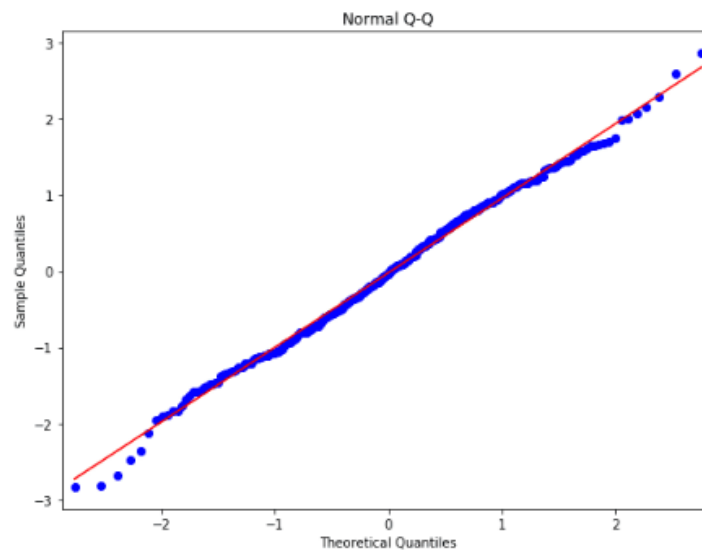
Gambar 4.13. *Standardized residual*

- *Histogram plus estimated density*: Plot kepadatan menunjukkan distribusi normal dengan rata-rata nol. Uji normalitas residual untuk melihat kenormalan residual, model dianggap baik jika residualnya berdistribusi normal. KDE (*Kernel Density Estimation*) adalah pendekatan statistik parametric guna mengestimasi fungsi distribusi variable acak.



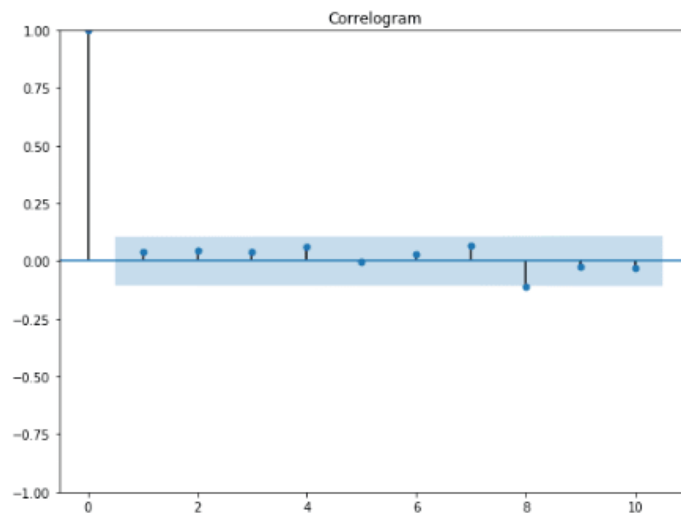
Gambar 4.14. *Histogram plus estimated density*

- Normal Q – Q: Letak titik-titik sesuai dengan garis merah, jika terjadi penyimpangan yang signifikan akan terlihat distribusi yang miring.



Gambar 4.15. *Plot Q-Q*

- Correlogram: Menunjukkan *error* residual tidak berkorelasi otomatis, uji non-korelasi bertujuan mengetahui adanya korelasi antara data residual, suatu model yang baik nilai-nilai residualnya tidak saling berkorelasi.



Gambar 4.16. Correlogram

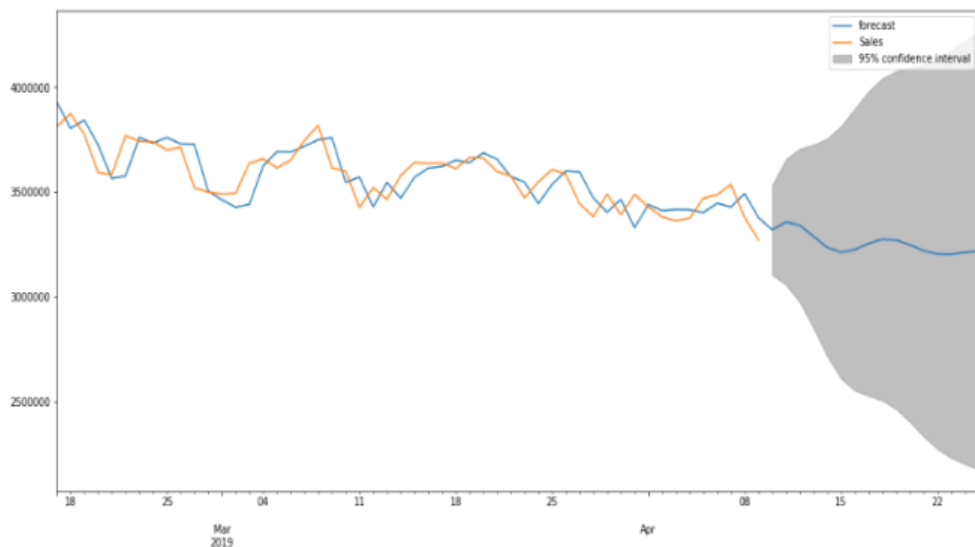
5. Pemilihan Model Terbaik ARIMA (p, d, q)

Pada penelitian ini model terbaik dilihat dari nilai AIC, signifikansi p -value dan MAPE yang terkecil, pada tabel 4.2.1 terdapat nilai AIC, signifikansi p -value dan MAPE dari beberapa model yang telah diujikan dan didapat model ARIMA (3,1,2) mendapatkan AIC sebesar 7639.772 dan nilai MAPE sebesar 0.0256123932 atau 2.5%, serta p -value yang signifikan. Persamaan model ARIMA (3,1,2) dari hasil estimasi adalah:

$$Y_t = -4437.2439 + 1.3177Y_{t-1} - 0.9924Y_{t-2} + 0.1514Y_{t-3} \\ - 1.3262 \epsilon_{t-1} + 0.9999 \epsilon_{t-2} + \epsilon_1$$

6. Peramalan

Setelah model terbaik telah didapatkan, yaitu model ARIMA (3,1,2), selanjutnya melakukan peramalan dengan melihat trend berdasarkan plot, gambar 4.17 adalah plot peramalan untuk 15 hari kedepan.



Gambar 4.17. Plot Peramalan 15 Hari Kedepan

Dari gambar 4.17 diatas dapat dilihat bahwa untuk 15 hari kedepan diperkirakan nilai penjualan berada pada *range* antara Rp. 3.300.000 sampai Rp. 3.500.000. Dengan nilai RMSE 112883.14236 diperkirakan nilai penjualan toko Teras Umat untuk 15 hari kedepan berada pada kisaran antara Rp. 3.187.116 sampai Rp. 3.612.883. Tabel 4.5 adalah nilai penjualan aktual toko dari tanggal 1 Mei 2019 sampai dengan 15 Mei 2019, rata-rata penjualan pada durasi 15 hari tersebut adalah Rp. 3.217.024 dan masuk dalam kisaran yang sesuai dengan peramalan pada penelitian ini.

Tabel 4.5. Nilai Penjualan Aktual

Date	Actual Sales (Rp)
1-May-19	3480933
2-May-19	3254640
3-May-19	2130120
4-May-19	1688450
5-May-19	3814800
6-May-19	2203450
7-May-19	3298876
8-May-19	3641600
9-May-19	3575150
10-May-19	3384800
11-May-19	3737050
12-May-19	2831100
13-May-19	3270500
14-May-19	4607000
15-May-19	3336900

BAB 5

PENUTUP

5.1. Kesimpulan

Kesimpulan yang didapatkan pada penelitian ini berdasarkan analisis dan pembahasan antara lain adalah:

1. Dari data penjualan toko Teras Umat terdapat lonjakan penjualan yang menyebabkan data penjualan sangat fluktuatif terutama pada awal bulan berdiri, beberapa hal disebabkan antara lain, penjualan untuk *charity*, pembayaran elektronik (listrik dan BPJS) dari beberapa instansi, euphoria anggota pada awal-awal toko berdiri dan lain-lain.
2. Pada delapan bulan pertama penjualan memiliki pola musiman pada lag 15, namun di empat bulan terakhir penjualan cenderung rata dan menurun.
3. Model ARIMA yang paling baik dalam meramalkan penjualan kedepan pada toko Teras Umat adalah ARIMA (3,1,2) dikarenakan model tersebut mendapatkan nilai AIC sebesar 7639.772 dan nilai MAPE sebesar 0.0256123932 atau persentase *error* sebesar 2.5%.
4. Diduga pada empat bulan terakhir pola penjualan cenderung flat dan menurun disebabkan *Supply* barang dari distributor ke toko mengalami masalah.

5.2. Saran

Dari hasil penelitian didapatkan saran untuk kemajuan toko Teras Umat, antara lain:

1. Menggunakan metode data mining dalam mengekstrak data penjualan untuk mengambil keputusan dan langkah kedepan, seperti aturan asosiasi, analisis belanja dan lain-lain.
2. Menganalisis dugaan hubungan antara *Supply* barang dan penjualan.
3. Melakukan peramalan dengan metode lain dan data penjualan yang lebih baru.
4. Mendorong agar anggota memenuhi kebutuhan hariannya dengan berbelanja di toko Teras Umat.

DAFTAR PUSTAKA

- [1] S. H. Wahid, “Ketentuan Pengembalian Setoran Pokok Dalam Undang-Undang No.17 Tahun 2012 Tentang Perkoperasian (Perspektif Undang-Undang Dasar 1945 Dan Hukum Islam),” JURISDICTION, 2017, doi: 10.18860/j.v5i2.4018.
- [2] Abdullah, A. G., & Mulyadi, Y. (2016). Peramalan Beban Listrik Jangka Pendek Melalui Pendekatan Statistik dan Soft Computing. KNTIA, 2.
- [3] Yusendra, M. A. E. (2015). Kajian Penerapan Metode Peramalan pada Ilmu Ekonomi dan Ilmu Komputer (Studi Kasus: Penerimaan Mahasiswa Baru Ibi Darmajaya). Prosiding Sembistek 2014, 1(01), 267-279. TIA, 2.
- [4] Pavlyshenko, B. M. (2019). Machine-learning models for sales time series forecasting. Data, 4(1), 15. Doi: 10.3390/data4010015.
- [5] Molydah, S. (2018). Analisis Perbandingan Implementasi Sarima Dan Support Vector Machine (Svm) Dalam Prediksi Jumlah Wisatawan Mancanegara.(Universitas Islam Indonesia, Yogyakarta)
- [6] Widiyanto, S. (2019). Peramalan nilai impor besi baja Indonesia 2008-2017 dengan model arima. JURNAL MANAJEMEN, 11(2), 217-225.
- [7] Wiyanti, D. T., & Pulungan, R. (2012). Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (RBF) dan Auto Regressive Integrated Moving Average (ARIMA). Jurnal Mipa, 35(2).
- [8] Ferguson, T. S. (1961, July). On the rejection of outliers. In Proceedings of the fourth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 1, pp. 253-287). Berkeley: University of California Press.
- [9] Barnett, V., & Lewis, T. (1984). Outliers in statistical data-Osd.
- [10] Sembiring, R. K. (1995). Analisis regresi. Bandung: ITB.
- [11] Paludi, S. (2009). Identifikasi dan Pengaruh Keberadaan Data Pencilan (Outlier). Majalah Panorama Nasional, 56-62.
- [12] Septiadi, M. N. K., Handayani, S., & Mulyadi, C. (2018). Penilaian Persediaan Barang Dagang Dengan Metode Rata-Rata Bergerak. EKA CIDA, 1(2).
- [13] Yuanti, A. D. (2016). Perbandingan Model Time Series Seasonal Arima (Sarima) Dan Seasonal Arfima (Sarfima) Pada Data Beban Konsumsi

- Listrik Jangka Pendek Di Jawa Timur (Studi Kasus di APD PT. PLN Distribusi Jawa Timur) (Doctoral dissertation, Airlangga University).
- [14] Cryer, J. D., & Chan, K. S. (2008). Time series analysis: with applications in R. Springer Science & Business Media.
 - [15] Wei, W. W. (2006). Time series analysis. In *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*.
 - [16] Azriati, K. F., Hoyyi, A., & Mukid, M. A. (2014). Verifikasi Model Arima Musiman Menggunakan Peta Kendali Moving Range (Studi Kasus: Kecepatan Rata-rata Angin di Badan Meteorologi Klimatologi dan Geofisika Stasiun Meteorologi Maritim Semarang). *Jurnal Gaussian*, 3(4), 701-710. ISSN: 2339-2541.
 - [17] Aziz, A. (2011). Analisis Critical Root Value pada Data Nonstasioner. *CAUCHY*, 2(1), 1-6. Doi: doi.org/10.18860/ca.v2i1.1794.
 - [18] Widodo, W. (2005). Metode Autoregresi dan Autokorelasi untuk meramalkan jumlah penjualan pakaian di toko Yuanita Purwodadi. Tugas Akhir Mahasiswa UNNES, Semarang.
 - [19] Krismiasari, S. (2012). Peramalan Produksi Padi Di Kabupaten Kampar Dengan Metode Box-Jenkins (Doctoral dissertation, Universitas Islam Negeri Sultan Syarif Kasim Riau).
 - [20] Sugiarto, S. (2017). Penduga Model Arima Untuk Peramalan Harga Tbs Kelapa Sawit Di Propinsi Riau. *Jurnal Sains dan Teknologi Industri*, 15(1), 35-40.
 - [21] Hanke, J. E., Reitsch, A. G., & Wichern, D. W. (2001). Business forecasting (Vol. 9). New Jersey: Prentice Hall.
 - [22] Makridakis, S. (1993). Dkk. Metode dan Aplikasi Peramalan, Jakarta: Airlangga.
 - [23] A. H. Hutasuhut, "Pembuatan Aplikasi Pendukung Keputusan untuk Peramalan Persediaan Bahan Baku Produksi Plastik Blowing dna Inject Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average) di CV. Asia," *J. Tek. Pomits*, vol. 3, no. 2, pp. 70–171, 2014, [Online]. Available: <http://ejurnal.its.ac.id/index.php/teknik/article/viewFile/8114/1846>.

- [24] Hatidja, D. (2011). Penerapan Model Arima Untuk Memprediksi Harga Saham PT. Telkom Tbk. *Jurnal Ilmiah Sains*, 11(1), 116-123. Doi: 10.35799/jis.11.1.2011.53.
- [25] Pamungkas, M. B., & Wibowo, A. (2019). Aplikasi Metode ARIMA Box-Jenkins Untuk Meramalkan Kasus Dbd Di Provinsi Jawa Timur. *The Indonesian Journal of Public Health*, 13(2), 183.doi: 10.20473/ijph.v13i2.2018.183-196.
- [26] Sutanto, P., Setiawan, A., & Setiabudi, D. H. (2017). Perancangan Sistem Forecasting di Perusahaan Kayu UD. 3G dengan Metode ARIMA. *Jurnal Infra*, 5(1), 325-330.
- [27] Wibowo, A. (2018). Model peramalan indeks harga konsumen kota Palangka Raya menggunakan Seasonal ARIMA (SARIMA). *Matematika*, 17(2).
- [28] Rahmadayanti, R., Susilo, B., & Puspitaningrum, D. (2015). Perbandingan Keakuratan Metode Autoregressive Integrated Moving Average (ARIMA) dan Exponential SMOOTHING Pada Peramalan Penjualan Semen di PT. Sinar Abadi. *Rekursif: Jurnal Informatika*, 3(1).
- [29] Yuniarti, A. (2010). Perbandingan metode peramalan eksponensial smoothing dan Box-Jenkins (ARIMA) musiman (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- [30] Ul Ukhra, A. (2014). Pemodelan dan peramalan data deret waktu dengan metode SEASONAL ARIMA. *Jurnal Matematika UNAND*, 3(3).
- [31] Linda, P., Situmorang, M., & Tarigan, G. (2014). Peramalan Penjualan Produksi Teh Botol Sosro Pada PT. Sinar Sosro Sumatera Bagian Utara Tahun 2014 Dengan Metode Arima Box-Jenkins. *Saintia Matematika*, 2(3), 253-266.
- [32] As'ad, M., Wibowo, S. S., & Sophia, E. (2017). Peramalan Jumlah Mahasiswa Baru Dengan Model Autoregressive Integrated Moving Average (Arima). *JIMP-Jurnal Informatika Merdeka Pasuruan*, 2(3), doi: 10.37438/jimp.v2i3.77.
- [33] Raihan, R., Effendi, M. S., & Hendrawan, A. (2016). Forecasting Model Exsponensial Smoothing Time Series Rata Rata Mechanical Availability

Unit Off Highway Truck CAT 777D Caterpillar. POROS TEKNIK, 8(1), 1-9.

- [34] Aprilia, D. (2016). Perbandingan Metode Peramalan Exponential Smoothing Dan Moving Average (Studi Prediksi Jumlah Penderita TB Paru di Provinsi Jawa Timur) (Doctoral dissertation, Universitas Airlangga).
- [35] Setiawan, E., Murfi, H., & Satria, Y. (2016). Analisis Penggunaan Metode Kernel Density Estimation pada Loss Distribution Approach untuk Risiko Operasional. Jurnal Matematika Integratif ISSN, 1412, 6184.
- [36] Putri, R. M., & Widodo, E. (2018). Application of Support Vector Machine Method For Rupiah Exchange Rate To Us Dollar Forecasting. In PROSIDING SEMINAR NASIONAL & INTERNASIONAL (Vol. 1, No. 1).
- [37] Naufal, M. F. (2017). Peramalan Jumlah Wisatawan Mancanegara Yang Datang Ke Indonesia Berdasarkan Pintu Masuk Menggunakan Metode Support Vector Machine (SVM) (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- [38] Samsiah, D. N. (2008). Analisis Data Runtun Waktu Menggunakan Model Arima (p, d, q). Program studi Matematika Fakultas Sains dan Teknologi. UIN Sunan Kalijaga. Yogyakarta.

DAFTAR RIWAYAT HIDUP

A. Biodata Mahasiswa

N I M : 14002269
 Nama Lengkap : Luky Fabrianto
 Tempat dan Tanggal Lahir : Jakarta, 5 Februari 1976
 Jenis Kelamin : Laki-laki
 Alamat : Jl. RTM Poncol No. 99 Kel. Tugu, Kec. Cimanggis,
 Kota Depok

B. Riwayat Pendidikan

1. SD Barunawati Jakarta, lulus Tahun 1987
2. SMPN 61 Jakarta, lulus Tahun 1990
3. SMAN 33 Jakarta, lulus Tahun 1993
4. D3 STMIK Gunadarma Depok, lulus Tahun 1996
5. S1 Universitas Gunadarma Depok, lulus Tahun 1999

C. Riwayat Pengalaman Pekerjaan

1. PT. Merbau Sakti (Tahun 1996 - 1997)
2. PT. Phintraco Ekasarana (Tahun 2000 - 2002)
3. PT. Jebsen and Jessen (Tahun 2002 - 2004)
4. PT. Moores Rowland (Tahun 2004 - 2005)
5. PT. Televisi Anak Spacetoon (Tahun 2005 - 2012)
6. PT. Televisi Anak Spacetoon – Indika Group (Tahun 2012 - 2013)
7. CV. Korabe Multi Teknika (Tahun 2013 - sekarang)



Jakarta, Agustus 2020

(Luky Fabrianto)

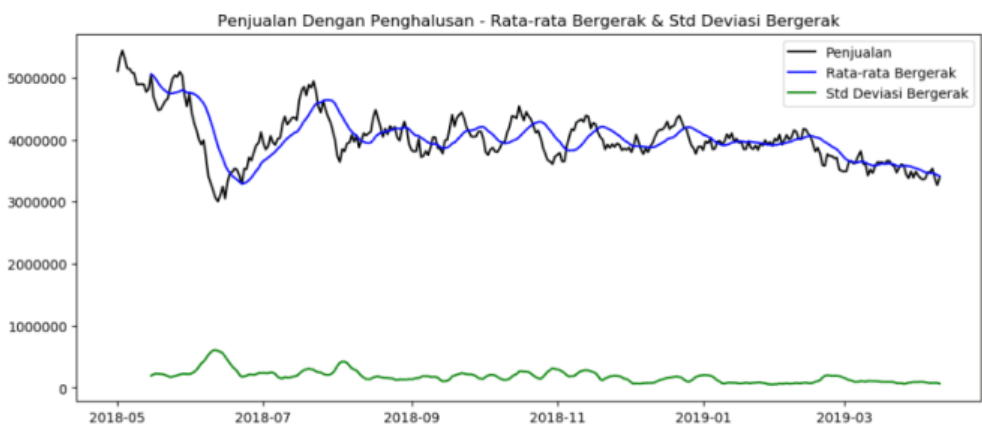
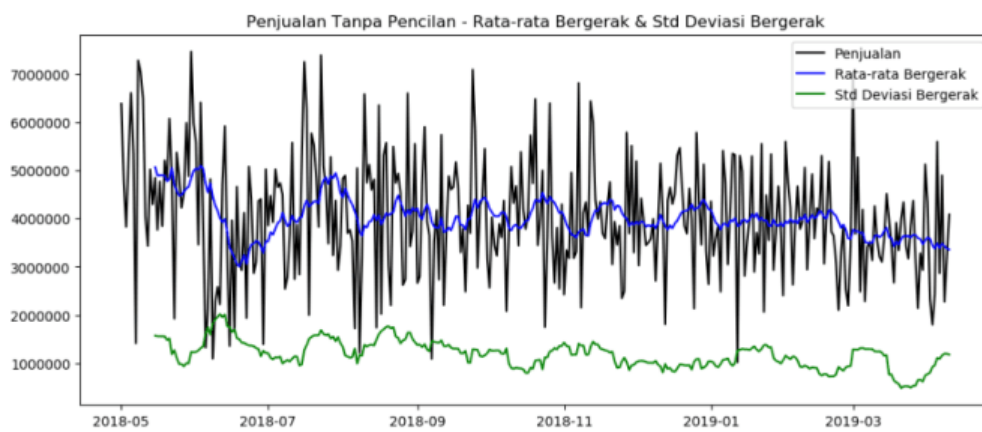
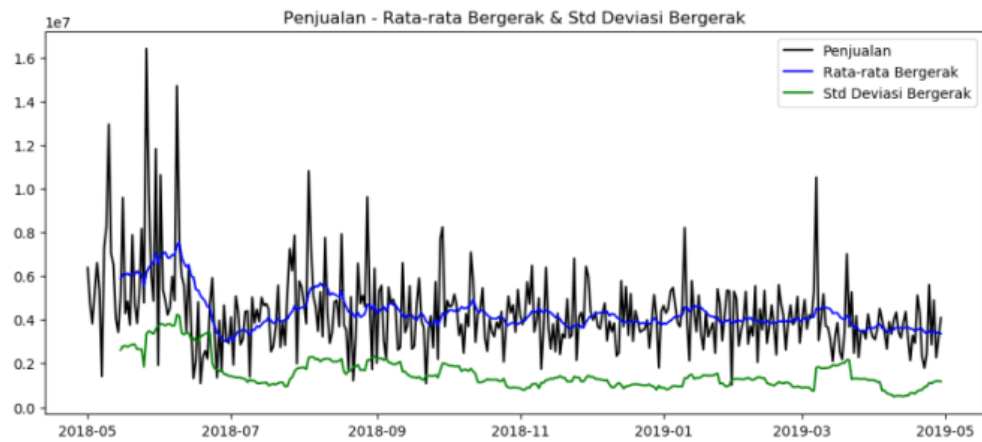
Lampiran 1. Data Penjualan

Sales	1325000	5525270	5550516	5384372	5512683.5	5302910	2194714
6384800	2139130	4825600	2663682	4209867.5	3298180	4982660.7	3437088
4737050	4819993	3830016	2801025	3787006	5192436	2810962	7015354
3831100	1096820	10812598	4718811.5	4134767	3028643	3509170	3684064
5270500	2283930	7391570	5904828	5729647.5	4421809	3991248	5272850
6607000	2581240	5098685	3945146.5	4733008	3981600	5295631	2486400
5336900	2221680	4601375	3649417.5	6482870	3447754	2895116	4187600
1413800	4713110	3481900	1090471.5	3447945	3506632	3705718	2285850
7284986	5918972	5281047	3532846.5	3981095	3587276	3269750	3413276
8283700	3278490	3238625	4174902.5	4998721.5	3994744	5551359	3518445
12947000	1356420	7757995	2734178.5	1749330	2710976	2064850	3137825
7045100	3964318	4375520	5736924	3626074	3935463	4487135	4260250
6502350	1652880	2931605	2203987	6396107.5	5146550	3552178.5	3574172
4040500	4662690.5	3418465	7698265.5	3634905	4117666	5338349.5	3216037
3433900	3096978	4832771	8245567	2667692.5	1807237	2932473.5	3103800
5021400	2930462.5	4892075	3695236.5	3811446	4382120	3696543	3593711
9592830	4121107	3698860	4880954.5	2544910	4648267	4669961.5	4514453
4295910	1937380	7928242	4606734	4298344	4303377.5	3698846.5	3977392
4849600	5080750	3766000	4659264	2424717	4519100	2406134	3448242
3768661	4426370	3540675	5171212	3346420	5310545	5599979	2670302
7883972	2868730	1724950	4614964	3179400	5471225	4746450	3914908
5/21/2018	3098273	5048715	3300500	4956339	4640610	4219617.5	3344772
3846158	4369250	1220205	3798779	3179547	3852015	2628205.5	3998733
5208940	4407375	3215315	2486977	3310924	3690440	3870928.5	4340994
8177134	1394296	6583395	4275700	6814364	4627150	4670696	3443283
4787890	5026887	4746525	3700000	2155950	8219110	3786707.5	3172202
16406520	3857727.5	5117510	7094127	4114923	3991330	3937714	3859692
9732180	4474161	4599856	5823990	4339500	2134914	5060139	4372326
6078500	3928857.5	9624547	2979890	3646500	5785404	2944571	3173820
4864430	5024817.5	4811750	4267170	6439000	4330933	4025138	2142328
11824950	4653522.5	1739800	4419150	6005140	3459455.5	4921870	3279378
1925870	4737167.5	6353850	5453340	4391390	5124843.5	3588334	2933322
10628570	4495769	2021850	3163500	3996880	3318084	4219608	5126425
5371910	2544817.5	5312528.5	2567300	4353864	2640100	4099273.5	4339305
4898310	2791272.5	5573179	4027699	3695600	4358416.5	5303075	2324450
4229780	3673400	3031782.5	3462600	3590436	3227620.5	10521703.5	1799400
4535490	5577030	2200660	3231500	4298700	3586950	3064250	2435594
5986020	2746307	5492602.5	3890500	4760866	3862250	4010586	5598400
4883948	3871570	4741831	3612300	3050310	2484965	5181804	2867650
14708840	2840970	4935794	4375114	3934476	5410839	3726350	4898600
7467696	5277030	4548612	2081300	3457797	4744532	3643831	2276150
5978740	7252465	2621316	3746000	3851553	3051354	3085542	3151115
5577809	6254334	2742743	5079600	2348987	4232741	2106033	4085363
3461700	7862260	6602716	4324816	2494129.5	5350831.5	3268612	
6411110	1998350	3423570	4674900	5787623.5	5316432.3	3866688	
3921570	5771281	3732557	3440400	3692425	1019750	2537850	

Lampiran 2. Data Pencilan

Tgl	Sales	No trans	Ket Barang	Qty	Harga	Tot Harga
5/9/2018	8,283,700	201801001922	HYDRO BERAS CINDE 5 KG	85	62,000	5,270,000
5/10/2018	12,947,000	201801001929	HYDRO BERAS CINDE 5 KG	80	62,000	4,960,000
5/16/2018	9,592,830	Hari Pertama bulan Ramadhan				
5/20/2018	7,883,972	Hari Minggu Pertama bulan Ramadhan				
5/24/2018	8,177,134	Lonjakan Penjualan				
5/26/2018	16,406,520	201801003211	GURIH MINYAK GORENG PCH 2L	70	23,900	1,673,000
		201801003211	SODAQO GULA PASIR LOKAL 1KG	70	11,700	819,000
		201801003211	GUNUNG MAS BERAS SAK 2.5KG	70	32,000	2,240,000
		201801003211	POCI TEH CELUP PAK 25X2G	70	5,800	406,000
		201801003211	STUPA HIJAU TEPUNG TERIGU 1KG	70	8,200	574,000
		201801003211	FORVITA MARGARINE 200G	70	4,900	343,000
		201801003211	ABC KECAP MANIS BTL 135ML	70	6,700	469,000
		201801003211	KARDUS PAKET KECIL	70	5,850	409,500
5/27/2018	9,732,180	Lonjakan Penjualan				
5/30/2018	11,824,950	Lonjakan Penjualan				
6/1/2018	10,628,570	Lonjakan Penjualan				
6/8/2018	14,708,840	1806010130	BERAS CAP BUNGA	200	59,000	11,800,000
7/29/2018	7,862,260	Lonjakan Penjualan				
8/4/2018	10,812,598	201801008188	EB P500 PLN_PREPAID	1	502,500	502,500
		201801008254	PLN_POSTPAID	1	2,978,792	2,978,792
		201801008277	PLN_POSTPAID	1	444,479	444,479
8/11/2018	7,757,995	201801008974	PLN_POSTPAID	1	1,540,395	1,540,395
8/18/2018	7,928,242	Lonjakan Penjualan				
8/29/2018	9,624,547	201801010508	BPJS_KESEHATAN	1	4,655,200	4,655,200
9/29/2018	7,698,266	Lonjakan Penjualan				
9/30/2018	8,245,567	Lonjakan Penjualan				
1/11/2019	8,219,110	201901000971	BPJS_KESEHATAN	1	4,242,300	4,242,300
3/8/2019	10,521,704	201901006083	PULOSARI BERAS SAK 5KG	20	63,900	1,278,000
		201901006083	SODAQO GULA PASIR LOKAL 1KG	30	11,900	357,000
		201901006083	SANIA MINYAK GRNG POUCH 2000ML	15	29,800	447,000
		201901006083	INDOMIE GORENG 85G	200	2,400	480,000
		201901006083	INDOMIE AYAM BAWANG 69G	200	2,100	420,000
		201901006090	EB P500 PLN_PREPAID	1	502,500	502,500

Lampiran 3. Grafik Standar Deviasi



Lampiran 4. *Summary* Hasil Estimasi Parameter

ARIMA Model Results						
=====						
Dep. Variable:	D.Sales		No. Observations:	292		
Model:	ARIMA(1, 1, 1)		Log Likelihood	-3827.030		
Method:	css-mle		S.D. of innovations	119050.889		
Date:	Tue, 14 Jul 2020		AIC	7662.060		
Time:	14:14:15		BIC	7676.767		
Sample:	05-02-2018		HQIC	7667.951		
	- 02-17-2019					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	-4437.2438	8423.193	-0.527	0.598	-2.09e+04	1.21e+04
ar.L1.D.Sales	0.7062	0.210	3.368	0.001	0.295	1.117
ma.L1.D.Sales	-0.6441	0.224	-2.878	0.004	-1.083	-0.205
Roots						
=====						
	Real	Imaginary	Modulus	Frequency		

AR.1	1.4161	+0.0000j	1.4161	0.0000		
MA.1	1.5526	+0.0000j	1.5526	0.0000		

ARIMA Model Results						
Dep. Variable:	D.Sales	No. Observations:	292			
Model:	ARIMA(1, 1, 2)	Log Likelihood	-3827.025			
Method:	css-mle	S.D. of innovations	119049.032			
Date:	Tue, 14 Jul 2020	AIC	7664.050			
Time:	14:16:18	BIC	7682.434			
Sample:	05-02-2018	HQIC	7671.414			
	- 02-17-2019					
	coef	std err	z	P> z	[0.025	0.975]
const	-4437.2438	8393.772	-0.529	0.597	-2.09e+04	1.2e+04
ar.L1.D.Sales	0.7114	0.215	3.304	0.001	0.289	1.133
ma.L1.D.Sales	-0.6452	0.223	-2.899	0.004	-1.081	-0.209
ma.L2.D.Sales	-0.0065	0.066	-0.098	0.922	-0.135	0.122
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	1.4057	+0.0000j	1.4057	0.0000		
MA.1	1.5266	+0.0000j	1.5266	0.0000		
MA.2	-101.4150	+0.0000j	101.4150	0.5000		

```

=====
ARIMA Model Results
=====
Dep. Variable:      D.Sales      No. Observations:      292
Model:              ARIMA(1, 1, 3)  Log Likelihood          -3826.435
Method:              css-mle       S.D. of innovations     118804.744
Date:                Tue, 14 Jul 2020  AIC                        7664.870
Time:                14:17:11       BIC                     7686.931
Sample:              05-02-2018     HQIC                    7673.707
                  - 02-17-2019
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const          -4437.2439   8784.057     -0.505     0.613    -2.17e+04    1.28e+04
ar.L1.D.Sales    0.6405     0.228       2.806     0.005      0.193      1.088
ma.L1.D.Sales   -0.5837     0.231     -2.529     0.011     -1.036     -0.131
ma.L2.D.Sales   -0.0289     0.064     -0.449     0.653     -0.155     0.097
ma.L3.D.Sales    0.0680     0.061      1.113     0.266     -0.052     0.188
=====
Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1           1.5612      +0.0000j       1.5612       0.0000
MA.1           1.9063      -0.8390j       2.0828      -0.0660
MA.2           1.9063      +0.8390j       2.0828       0.0660
MA.3          -3.3880      -0.0000j       3.3880      -0.5000
=====

```

```

=====
ARIMA Model Results
=====
Dep. Variable:      D.Sales      No. Observations:      292
Model:              ARIMA(2, 1, 1)  Log Likelihood          -3826.779
Method:              css-mle       S.D. of innovations     118932.307
Date:                Tue, 14 Jul 2020  AIC                        7663.558
Time:                14:20:21       BIC                     7681.942
Sample:              05-02-2018     HQIC                    7670.922
                  - 02-17-2019
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const          -4437.2438   7564.268     -0.587     0.557    -1.93e+04    1.04e+04
ar.L1.D.Sales   -0.8996     0.065    -13.801     0.000     -1.027     -0.772
ar.L2.D.Sales    0.0844     0.059      1.431     0.152     -0.031     0.200
ma.L1.D.Sales    0.9734     0.030     32.791     0.000      0.915     1.032
=====
Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1          -1.0150      +0.0000j       1.0150       0.5000
AR.2           11.6757      +0.0000j      11.6757       0.0000
MA.1          -1.0273      +0.0000j       1.0273       0.5000
=====

```

ARIMA Model Results						
=====						
Dep. Variable:	D.Sales	No. Observations:	292			
Model:	ARIMA(2, 1, 2)	Log Likelihood	-3820.642			
Method:	css-mle	S.D. of innovations	115507.903			
Date:	Tue, 14 Jul 2020	AIC	7653.283			
Time:	14:21:12	BIC	7675.344			
Sample:	05-02-2018	HQIC	7662.120			
	- 02-17-2019					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	-4437.2439	7294.296	-0.608	0.543	-1.87e+04	9859.314
ar.L1.D.Sales	-0.0683	0.029	-2.328	0.020	-0.126	-0.011
ar.L2.D.Sales	-0.9180	0.027	-34.030	0.000	-0.971	-0.865
ma.L1.D.Sales	0.1437	0.011	12.820	0.000	0.122	0.166
ma.L2.D.Sales	0.9998	0.019	53.239	0.000	0.963	1.037
Roots						
=====						
	Real	Imaginary	Modulus	Frequency		

AR.1	-0.0372	-1.0431j	1.0437	-0.2557		
AR.2	-0.0372	+1.0431j	1.0437	0.2557		
MA.1	-0.0719	-0.9975j	1.0001	-0.2614		
MA.2	-0.0719	+0.9975j	1.0001	0.2614		

ARIMA Model Results						
Dep. Variable:	D.Sales	No. Observations:	292			
Model:	ARIMA(2, 1, 3)	Log Likelihood	-3822.796			
Method:	css-mle	S.D. of innovations	117248.848			
Date:	Tue, 14 Jul 2020	AIC	7659.592			
Time:	14:18:14	BIC	7685.329			
Sample:	05-02-2018	HQIC	7669.902			
	- 02-17-2019					
	coef	std err	z	P> z	[0.025	0.975]
const	-4437.2440	7448.375	-0.596	0.551	-1.9e+04	1.02e+04
ar.L1.D.Sales	0.3072	0.046	6.687	0.000	0.217	0.397
ar.L2.D.Sales	-0.9194	0.214	-4.296	0.000	-1.339	-0.500
ma.L1.D.Sales	-0.2529	0.069	-3.677	0.000	-0.388	-0.118
ma.L2.D.Sales	0.8838	0.222	3.982	0.000	0.449	1.319
ma.L3.D.Sales	0.1199	0.084	1.435	0.151	-0.044	0.284
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	0.1670	-1.0295j	1.0429	-0.2244		
AR.2	0.1670	+1.0295j	1.0429	0.2244		
MA.1	0.2045	-1.0150j	1.0354	-0.2184		
MA.2	0.2045	+1.0150j	1.0354	0.2184		
MA.3	-7.7815	-0.0000j	7.7815	-0.5000		

```

=====
ARIMA Model Results
=====
Dep. Variable:      D.Sales      No. Observations:      292
Model:              ARIMA(3, 1, 1)  Log Likelihood          -3826.477
Method:              css-mle       S.D. of innovations     118822.597
Date:               Tue, 14 Jul 2020  AIC                        7664.953
Time:               14:22:31        BIC                     7687.014
Sample:             05-02-2018      HQIC                    7673.790
                   - 02-17-2019

=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const        -4437.2439   8708.737    -0.510    0.610   -2.15e+04   1.26e+04
ar.L1.D.Sales    0.6018     0.267     2.252    0.024     0.078     1.126
ar.L2.D.Sales   -0.0369     0.071    -0.523    0.601    -0.175     0.101
ar.L3.D.Sales    0.0661     0.061     1.088    0.276    -0.053     0.185
ma.L1.D.Sales   -0.5366     0.263    -2.044    0.041    -1.051    -0.022

Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1          1.4538      -0.0000j      1.4538      -0.0000
AR.2         -0.4479      -3.1943j      3.2255      -0.2722
AR.3         -0.4479      +3.1943j      3.2255      0.2722
MA.1          1.8636      +0.0000j      1.8636      0.0000
=====

```

```

=====
ARIMA Model Results
=====
Dep. Variable:      D.Sales      No. Observations:      292
Model:              ARIMA(3, 1, 2)  Log Likelihood          -3812.886
Method:              css-mle       S.D. of innovations     112262.639
Date:               Tue, 14 Jul 2020  AIC                        7639.772
Time:               14:23:13        BIC                     7665.509
Sample:             05-02-2018      HQIC                    7650.081
                   - 02-17-2019

=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const        -4437.2439   8441.924    -0.526    0.599   -2.1e+04   1.21e+04
ar.L1.D.Sales    1.3177     0.059    22.396    0.000     1.202     1.433
ar.L2.D.Sales   -0.9924     0.078   -12.652    0.000    -1.146    -0.839
ar.L3.D.Sales    0.1514     0.059     2.583    0.010     0.036     0.266
ma.L1.D.Sales   -1.3262     0.017   -78.694    0.000    -1.359    -1.293
ma.L2.D.Sales    0.9999     0.018    54.335    0.000     0.964     1.036

Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1          0.7260      -0.8759j      1.1377      -0.1399
AR.2          0.7260      +0.8759j      1.1377      0.1399
AR.3          5.1042      -0.0000j      5.1042      -0.0000
MA.1          0.6631      -0.7486j      1.0001      -0.1346
MA.2          0.6631      +0.7486j      1.0001      0.1346
=====

```

ARIMA Model Results						
=====						
Dep. Variable:	D.Sales	No. Observations:	292			
Model:	ARIMA(3, 1, 3)	Log Likelihood	-3812.847			
Method:	css-mle	S.D. of innovations	112246.480			
Date:	Tue, 14 Jul 2020	AIC	7641.695			
Time:	14:24:28	BIC	7671.109			
Sample:	05-02-2018	HQIC	7653.477			
	- 02-17-2019					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	-4437.2434	8566.374	-0.518	0.604	-2.12e+04	1.24e+04
ar.L1.D.Sales	1.4055	0.313	4.486	0.000	0.791	2.020
ar.L2.D.Sales	-1.0989	0.380	-2.889	0.004	-1.844	-0.353
ar.L3.D.Sales	0.2238	0.259	0.864	0.387	-0.284	0.731
ma.L1.D.Sales	-1.4156	0.318	-4.451	0.000	-2.039	-0.792
ma.L2.D.Sales	1.1194	0.425	2.634	0.008	0.287	1.952
ma.L3.D.Sales	-0.0902	0.320	-0.281	0.778	-0.718	0.538
Roots						
=====						
	Real	Imaginary	Modulus	Frequency		

AR.1	0.7191	-0.8774j	1.1344	-0.1407		
AR.2	0.7191	+0.8774j	1.1344	0.1407		
AR.3	3.4726	-0.0000j	3.4726	-0.0000		
MA.1	0.6627	-0.7489j	1.0001	-0.1347		
MA.2	0.6627	+0.7489j	1.0001	0.1347		
MA.3	11.0863	-0.0000j	11.0863	-0.0000		
=====						

Lampiran 5. Coding Jupyter

```
# Data Penjualan toko selama 1 tahun, 1 Mei 2018 sd 30 April 2019
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
data = pd.read_csv('E:\Kuliah Mkom\Thesis\Sales.csv',
parse_dates=True,index_col='Date')
data
# Plot Data Penjualan
data.plot(figsize=(15,5))
plt.title('Penjualan Dalam 1 Tahun - 1 Mei 2018 sd 30 April 2019')
# Deteksi Pencilan
Q1 = data.Sales.quantile(0.25)
Q2 = data.Sales.quantile(0.5)
Q3 = data.Sales.quantile(0.75)
Q1, Q2, Q3
IQR = Q3 - Q1
IQR
batas_bawah = Q1 - 1.5*IQR
batas_atas = Q3 + 1.5*IQR
batas_bawah, batas_atas
data[(data.Sales<batas_bawah)|(data.Sales>batas_atas)]
# Penyisihan Pencilan
df = data[(data.Sales>batas_bawah)&(data.Sales<batas_atas)]
df
# Plot Data Tanpa Pencilan
df.plot(figsize=(15,5))
plt.title('Penjualan Tanpa Pencilan')
# Penghalusan Rata-rata Bergerak
df1 = pd.read_csv('E:\Kuliah Mkom\Thesis\mlplus_sarima_out.csv',
parse_dates=True,index_col='Date')
df2 = df1.rolling(window=15, min_periods=5, center=True).mean()
df2.plot(figsize=(15,5))
plt.title('Penjualan dengan Penghalusan')
# Plot Rata2 Bergerak dan Rata2 Std Deviasi
roll_mean=df2.rolling(15).mean()
roll_std=df2.rolling(15).std()
plt.figure(figsize=(12,5), dpi=100)
plt.plot(df2, color='black',label='Penjualan')
plt.plot(roll_mean,color="blue",label='Rata-rata Bergerak')
plt.plot(roll_std,color='green',label='Std Deviasi Bergerak')
plt.legend(loc='best')
plt.title('Penjualan Dengan Penghalusan & Std Deviasi Bergerak')
plt.show(block=False)
# Cek Stasioneritas
from statsmodels.tsa.stattools import adfuller
from numpy import log
result = adfuller(df2.Sales.dropna())
```

```

print('ADF Statistic: %f' % result[0])
print('p-value: %f' % result[1])
# Differencing
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
fig, axes = plt.subplots(2, 1, figsize=(10,5), dpi=100, sharex=True)
axes[0].plot(df2[:,], label='Original')
axes[1].plot(df2[:,].diff(1), label='Differencing ke-1')
axes[1].plot(df2[:,].diff(2), label='Differencing ke-2')
axes[1].legend(loc='upper left', fontsize=10)
plt.title('Penjualan dan Differencing 1 & 2')
plot_acf(df2.Sales)
plot_pacf(df2.Sales)
plot_acf(df2.Sales.diff().dropna())
plot_pacf(df2.Sales.diff().dropna())
from statsmodels.tsa.arima_model import ARIMA
# p,d,q ARIMA Model
model = ARIMA(df2.Sales, order=(1,1,2))
model_fit = model.fit(dis=0)
print(model_fit.summary())
model = ARIMA(df2.Sales, order=(1,1,1))
model_fit = model.fit(dis=0)
print(model_fit.summary())
# plot residual errors
residuals = pd.DataFrame(model_fit.resid)
residuals.plot()
plt.figure(figsize=(12,5), dpi=100)
plt.show()
residuals.plot(kind='kde')
plt.show()
print(residuals.describe())
# Actual vs Fitted
plt.figure(figsize=(20,5), dpi=100)
model_fit.plot_predict(dynamic=False)
plt.show()
from statsmodels.tsa.stattools import acf
# Create Training and Test (293 = 85%)
train = df2.Sales[:293]
test = df2.Sales[293:]
# Build Model
model = ARIMA(train, order=(1, 1, 1))
fitted = model.fit(dis=-1)
# Forecast
fc, se, conf = fitted.forecast(52, alpha=0.05) # 95% conf
# Make as pandas series
fc_series = pd.Series(fc, index=test.index)
lower_series = pd.Series(conf[:, 0], index=test.index)
upper_series = pd.Series(conf[:, 1], index=test.index)
# Plot
plt.figure(figsize=(12,5), dpi=100)
plt.plot(train, label='Latih')

```

```

plt.plot(test, label='Aktual')
plt.plot(fc_series, label='Peramalan')
plt.fill_between(lower_series.index, lower_series, upper_series,
                 color='k', alpha=.15)
plt.title('Peramalan vs Aktual')
plt.legend(loc='upper left', fontsize=8)
plt.show()
# Build Model
model = ARIMA(train, order=(3, 1, 2))
fitted = model.fit(dispatch=-1)
print(fitted.summary())
# Forecast
fc, se, conf = fitted.forecast(52, alpha=0.05) # 95% conf
# Make as pandas series
fc_series = pd.Series(fc, index=test.index)
lower_series = pd.Series(conf[:, 0], index=test.index)
upper_series = pd.Series(conf[:, 1], index=test.index)
# Plot
plt.figure(figsize=(12,5), dpi=100)
plt.plot(train, label='Latih')
plt.plot(test, label='Aktual')
plt.plot(fc_series, label='Peramalan')
plt.fill_between(lower_series.index, lower_series, upper_series,
                 color='k', alpha=.15)
plt.title('Peramalan vs Aktual')
plt.legend(loc='upper left', fontsize=8)
plt.show()
# Accuracy metrics
def forecast_accuracy(forecast, actual):
    mape = np.mean(np.abs(forecast - actual)/np.abs(actual)) # MAPE
    me = np.mean(forecast - actual) # ME
    mae = np.mean(np.abs(forecast - actual)) # MAE
    mpe = np.mean((forecast - actual)/actual) # MPE
    rmse = np.mean((forecast - actual)**2)**.5 # RMSE
    corr = np.corrcoef(forecast, actual)[0,1] # corr
    mins = np.amin(np.hstack([forecast[:,None],
                              actual[:,None]]), axis=1)
    maxs = np.amax(np.hstack([forecast[:,None],
                              actual[:,None]]), axis=1)
    minmax = 1 - np.mean(mins/maxs) # minmax
    acf1 = acf(fc-test)[1] # ACF1
    return({'mape':mape, 'me':me, 'mae': mae,
           'mpe': mpe, 'rmse':rmse, 'acf1':acf1,
           'corr':corr, 'minmax':minmax})
forecast_accuracy(fc, test.values)
# Diagnostic
model.plot_diagnostics(figsize=(20,15))
plt.show()
# Final Forecast
from sklearn.metrics import mean_squared_error

```

```
X = df2.Sales
#size = int(len(X) * 0.85)
size = len(X) - 0
latih, uji = X[0:size], X[size:]
model = ARIMA(latih, order=(3,1,2))
model_fit = model.fit(dispatch=1)
model_fit.plot_predict(len(latih)-45, len(latih)+15)
plt.show()
```