



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS 141501

**ANALISIS POLA ABNORMAL KONSUMSI AIR MINUM
PELANGGAN PDAM SURYA SEMBADA SURABAYA
MENGUNAKAN METODE *PEARSON'S CORRELATION*,
ABNORMALLY LOW CONSUMPTION, DAN *WINDOWED
ANALYSIS*.**

YOGI TRI CAHYONO
NRP 5212 100 045

Dosen Pembimbing I:
Prof. Ir. Arif Djunaidy, M.Sc., Ph.D.

Dosen Pembimbing I:
Faizal Mahananto, S.Kom., M.Eng., Ph.D.

JURUSAN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2016



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS 141501

***ABNORMAL PATTERNS DRINKING WATER
CONSUMPTION ANALYSIS OF PDAM SURYA SEMBADA
SURABAYA'S CUSTOMERS USING PEARSON'S
CORRELATION, ABNORMALLY LOW CONSUMPTION,
AND WINDOWED ANALYSIS METHOD.***

**YOGI TRI CAHYONO
NRP 5212 100 045**

**SUPERVISOR I:
Prof. Ir. Arif Djunaidy, M.Sc., Ph.D.**

**SUPERVISOR II:
Faizal Mahananto, S.Kom., M.Eng., Ph.D.**

**JURUSAN SISTEM INFORMASI
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Surabaya 2016**

LEMBAR PENGESAHAN

**ABNORMAL PATTERNS DRINKING WATER
CONSUMPTION ANALYSIS OF PDAM SURYA
SEMBADA SURABAYA'S CUSTOMERS USING
PEARSON'S CORRELATION, ABNORMALLY LOW
CONSUMPTION, AND WINDOWED ANALYSIS
METHOD.**

TUGAS AKHIR

Disusun Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Jurusan Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember

Oleh:

YOGI TRI CAHYONO

NRP. 5212 100 045

Surabaya, 25 Juli 2016

KETUA JURUSAN SISTEM INFORMASI



Dr. Ir. Aris Triahyanto, M.Kom

NIP.19650310 199102 1 001

LEMBAR PERSETUJUAN

**ANALISIS POLA ABNORMAL KONSUMSI AIR
MINUM PELANGGAN PDAM SURYA SEMBADA
SURABAYA MENGGUNAKAN METODE
PEARSON'S CORRELATION, ABNORMALLY LOW
CONSUMPTION, DAN WINDOWED ANALYSIS.**

TUGAS AKHIR

Disusun Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer
pada
Jurusan Sistem Informasi
Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember
Oleh :

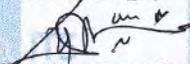
YOGI TRI CAHYONO
NRP. 5212 100 045

Disetujui Tim Penguji : Tanggal Ujian : 19 Juli 2016
Periode Wisuda : September 2016

Prof. Ir. Arif Djunaidy, M.Sc., Ph.D.


(Pembimbing I)

Faizal Mahananto, S.Kom., M.Eng., Ph.D.


(Pembimbing II)

Edwin Riksakomara, S.Kom., M.T.


(Penguji I)

Amalia Utamima, S.Kom., MBA.


(Penguji II)

**ANALISIS POLA ABNORMAL KONSUMSI AIR MINUM
PELANGGAN PDAM SURYA SEMBADA SURABAYA
MENGUNAKAN METODE PEARSON'S
CORRELATION, ABNORMALLY LOW CONSUMPTION,
DAN WINDOWED ANALYSIS.**

Nama Mahasiswa : Yogi Tri Cahyono
NRP : 5212 100 045
Jurusan : Sistem Informasi FTIf-ITS
Pembimbing 1 : Prof. Ir. Arif Djunaidy, M.Sc., Ph.D.
Pembimbing 2 : Faizal Mahananto, S.Kom., M.Eng., Ph.D.

ABSTRAK

Salah satu permasalahan utama yang dialami perusahaan air minum di negara berkembang adalah NRW. Non-revenue water (NRW) adalah perbedaan antara nilai air minum yang didistribusikan terhadap jumlah pendapatan yang diperoleh dari pelanggan. NRW disebabkan oleh tiga hal yaitu konsumsi resmi tidak ber-rekening, kehilangan akibat kesalahan teknis, dan kehilangan akibat kesalahan nonteknis. Dari masing-masing penyebab NRW tersebut, lebih dari 75 persen dari total jumlah NRW belum dapat di-identifikasi kehilangannya oleh PDAM Surya Sembada Surabaya. Penelitian yang dilakukan oleh Asian Development Bank (ADB) mengestimasi sekitar 50-65 persen NRW disebabkan oleh faktor nonteknis. Faktor nonteknis ini disebabkan oleh beberapa hal, yaitu: ketidakakuratan meter, kesalahan penanganan data, dan pencurian air yang dilakukan oleh pelanggan. Sebuah analisa pada pola konsumsi air minum pelanggan dapat membantu mengidentifikasi kehilangan air yang disebabkan oleh faktor nonteknis. Selama ini, analisa pada pola konsumsi air minum pelanggan yang dilakukan PDAM Surya Sembada Surabaya dilakukan melalui analisa penurunan konsumsi sebesar 30 persen. Cara identifikasi melalui pola

konsumsi seperti ini kurang memadai, karena hanya didasarkan pada pola konsumsi satu bulan terakhir.

Oleh karena uraian diatas, pada tugas akhir ini dibuat sebuah analisa pola konsumsi air minum pelanggan PDAM Surya Sembada Surabaya dengan menggunakan metode-metode statistik. Melalui analisa pola konsumsi ini diharapkan dapat menjadi rekomendasi tambahan serta dapat menentukan pelanggan yang telah mengkonsumsi air secara tidak sah (KTS) secara lebih optimal. Metode statistik yang digunakan terdiri dari korelasi produk momen (Pearson's correlation), metode konsumsi rendah abnormal (abnormally low consumption), dan algoritma analisis berjendela (windowed analysis). Kelompok pelanggan yang terklasifikasi terhadap masing-masing metode tersebut akan dianggap sebagai pelanggan yang berpotensi berperan dalam kehilangan air oleh faktor nonteknis.

Kinerja perhitungan yang diperoleh dari kalkulasi menggunakan metode-metode statistik dalam tugas ahir ini adalah 62% untuk akurasi, 68% untuk presisi, dan recall sebesar 44%. Nilai tersebut relatif rendah dikarenakan pelanggan yang telah teridentifikasi PDAM sebagai pelanggan mengkonsumsi air minum tidak sah tidak dapat dicakup oleh ketiga metode yang digunakan dan temuan kehilangan air nonteknis PDAM melalui analisis pola konsumsi air minum pelanggan juga masih terlampau rendah.

Kata kunci: *Metode Statistik, Pearson's Correlation, Abnormally Low Water Consumption, Konsumsi Air Tidak Sah, Pola Konsumsi Abnormal, PDAM Surya Sembada Surabaya, Windowed Analysis Algorithm.*

**ABNORMAL PATTERNS DRINKING WATER
CONSUMPTION ANALYSIS OF PDAM SURYA SEMBADA
SURABAYA'S CUSTOMERS USING PEARSON'S
CORRELATION, ABNORMALLY LOW CONSUMPTION,
AND WINDOWED ANALYSIS METHOD.**

Student Name : Yogi Tri Cahyono
NRP : 5212 100 045
Department : Sistem Informasi FTIf-ITS
Supervisor 1 : Prof. Ir. Arif Djunaidy, M.Sc., Ph.D.
Supervisor 2 : Faizal Mahananto, S.Kom., M.Eng., Ph.D.

ABSTRACT

One of the major issues affecting water utilities in the developing country is non-revenue water (NRW). NRW is difference between the amount of water put into the distribution system and the amount of water billed to customers. NRW comprises three component: unbilled authorized consumption, real losses, and apparent losses. More than 75% of total NRW suffered by PDAM Surya Sembada Surabaya couldn't be identified yet. A recent report by Asian Development Bank (ADB) estimated that in Asian cities around 50-65 % of NRW is due to apparent losses. Apparent losses is a result of unauthorized consumption and all types of inaccuracies associated with metering. Customer water consumption pattern analysis can be an approach to identify water losses which is caused by apparent losses. During this time, consumption analysis conducted by PDAM to identify apparent losses is through the 30% reduction of customer's consumption. The identification through consumption patterns like this is inadequate, because it's based only on consumption of the last month.

Therefore, in this final project is made an water consumption analysis of PDAM's customer using statistical methods. Through the analysis are made, hopefully can more optimal determine

customers who consuming water unauthorizedly and be an additional recommendation for PDAM. The statistical methods are used consist of Pearson's correlation, abnormally low consumption method, and windowed analysis algorithm. Consumer groups are classified to each of these methods will be considered as potential customer play a role in the loss of water by apparent losses.

The analysis performance which is obtained from calculation using statistical methods in this final project are 62% for the accuracy, 68% precision, and 44% recall. The result is relatively low because some of customer who consuming water unauthorizedly that have been identified by PDAM, cannot be covered by the three statistical methods and the findings of apparent loss through the analysis consumption pattern by PDAM is still low.

Keywords: *Statistical Methods, Pearson's Correlation, Abnormally Low Water Consumption, Windowed Analysis Algorithm, Unauthorized Water Consumption, Abnormal Water Consumption Pattern, PDAM Surya Sembada Surabaya.*

KATA PENGANTAR

Alhamdulillahirobbil'alamin. Atas karunia, berakah, dan jalan yang telah diberikan Allah SWT selama ini sehingga penulis dapat menyelesaikan tugas akhir dengan judul:

ANALISIS POLA ABNORMAL KONSUMSI AIR MINUM PELANGGAN PDAM SURYA SEMBADA SURABAYA MENGUNAKAN METODE PEARSON'S CORRELATION, ABNORMALLY LOW CONSUMPTION, DAN WINDOWED ANALYSIS

yang merupakan salah satu syarat kelulusan pada Jurusan Sistem Informasi, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya.

Terima kasih atas pihak-pihak yang telah mendukung, memberikan saran, motivasi, semangat, dan bantuan baik materi maupun moril demi tercapainya tujuan pembuatan tugas akhir ini. Secara khusus penulis akan menyampaikan ucapan terima kasih yang sedalam-dalamnya kepada:

- 1) Kedua orang tua penulis dan keluarga yang telah memberikan doa, kasih sayang, motivasi, dukungan materi dan moril sehingga penulis mampu menyelesaikan pendidikan S-1 ini dengan baik.
- 2) Bapak Prof. Ir. Arif Djunaidy, M.Sc., Ph.D. selaku dosen pembimbing pertama dan Bapak Faizal Mahananto, S.Kom., M.Eng., Ph.D. yang memberikan ilmu, pengalaman, petunjuk, dan motivasi sejalan dengan pengerjaan Tugas Akhir ini hingga selesai.
- 3) Bapak Arif Wibisono, S.Kom., M.Sc. dan Becti Cahyo Hidayanto, S.Si, M.Kom. sebagai dosen wali penulis selama menempuh pendidikan di Jurusan Sistem Informasi.
- 4) Karyawan PDAM Surya Sembada Surabaya yang telah memberi kesempatan dan bantuan, khususnya Mas Tatang, Mas Adit, Ibu Ayu, Bapak Nanang, Bapak Siri, dan Bapak

Nadir yang telah membantu mengumpulkan data dan informasi terkait keperluan tugas akhir ini.

- 5) Seluruh dosen pengajar beserta staf dan karyawan di Jurusan Sistem Informasi, FTIf ITS Surabaya yang telah memberikan ilmu dan bantuan.
- 6) Teman seperjuangan tugas akhir: Yukita, Adi, dan Yogha; Dagri Revolution: Ludy, Paung; Seluruh anggota MK56 Adit, Alwi, Andung, Benjo, Bowi, Danar, Dheni, Dody, Eca, Garis, Graha, Herman, Jack, Joshua, Kevin, Mayapada, Tino, Wildan; Geng Ampas Anggun, Abi, Amel, Faiz; Geng SE Aviva, Luqman, Ulvi; anggota lab RDIB Adit Darjo, Rifqi Bigpap; Teman-teman SOLA12IS Aga, Andhika, Andre, Ardhana, Denta, Eyang, Galent, Ican, Indi, Rowi; yang telah memberikan kesan, ilmu, inspirasi, semangat, dan pengalaman selama berkuliah dari awal hingga akhir pengerjaan tugas akhir ini.
- 7) Rekan-rekan mahasiswa Jurusan Sistem Informasi 8IOS, AE9IS, BASILISK dan seluruh angkatan SOLA12IS atas semua bantuan ketika penulis kuliah di Sistem Informasi ITS Surabaya.
- 8) Serta semua pihak yang telah membantu dalam pengerjaan Tugas Akhir yang belum mampu penulis sebutkan satu per satu diatas.

Semoga Allah SWT senantiasa melimpahkan rahmat hidayah serta membalas kebaikan yang telah diberikan kepada penulis.

Surabaya, 21 Juli 2016

Penulis

DAFTAR ISI

ABSTRAK.....	v
ABSTRACT.....	vii
KATA PENGANTAR	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xv
SEGMENT PROGRAM.....	xvii
DAFTAR TABEL.....	xix
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Perumusan Masalah.....	5
1.3. Batasan Tugas Akhir.....	5
1.4. Tujuan Tugas Akhir.....	6
1.5. Manfaat Tugas Akhir.....	6
1.6. Relevansi.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. PDAM Surya Sembada Surabaya.....	7
2.1.1. Gambaran Umum Hulu-hilir Air Minum PDAM.....	8
2.2. Rugi-rugi Distribusi Air Minum.....	10
2.2.1. Rugi Teknis.....	10
2.2.2. Rugi Nonteknis.....	11
2.3. Penelitian Analisis Pola Konsumsi Terdahulu.....	12
2.4. Dasar Teori.....	13
2.4.1. Normalisasi.....	13
2.4.2. Metode Statistik.....	14
2.4.3. Algoritma Statistik.....	18
2.5. Pengukuran Kinerja.....	25
BAB III METODOLOGI Pengerjaan Tugas AKHIR.....	27
3.1. Diagram Metodologi.....	27
3.2. Uraian Metodologi.....	27
3.2.1. Studi Literatur.....	27
3.2.1. Pengumpulan Data.....	29
3.2.2. Praproses Data.....	29

3.2.3.	Pembuatan Model metode Pearson's Correlation, Metode Abnormally Low Consumption, dan Algoritma Windowed Analysis.....	30
3.2.4.	Uji Coba.....	31
3.2.5.	Uji coba Deteksi Pelanggan KTS.....	31
3.2.6.	Buku Tugas Akhir.....	32
BAB IV	DESAIN DAN IMPLEMENTASI MODEL.....	33
4.1.	Pengumpulan Data.....	33
4.1.1.	Penggabungan Data.....	36
4.1.2.	Normalisasi.....	38
4.1.3.	Skenario Rentang Data Data.....	39
4.1.4.	Skenario Penentuan Nilai Ambang Metode.....	43
4.1.5.	Penentuan Konsumsi Minimal per Kapita per Hari.....	44
4.2.	Penentuan Pelanggan KTS.....	44
4.3.	Implementasi Aplikasi.....	45
4.3.1.	Implementasi Normalisasi Data.....	45
4.3.2.	Implementasi Memilih Rentang Data.....	46
4.3.3.	Implementasi Metode-metode Statistik.....	46
4.3.4.	Implementasi Masukan Parameter Metode.....	50
4.3.5.	Implementasi Penentuan Pelanggan KTS.....	51
BAB V	UJI COBA DAN ANALISIS HASIL.....	53
5.1.	Hasil Pembentukan Parameter Model.....	53
5.1.1.	Parameter Rentang Data.....	53
5.1.2.	Hasil Percobaan Parameter Nilai Ambang.....	63
5.2.	Hasil Uji Coba Model dan Aplikasi.....	67
5.2.1.	Uji Coba Parameter Model.....	68
5.2.2.	Uji Coba Aplikasi.....	72
5.3.	Hasil Penentuan Pelanggan KTS.....	77
5.4.	Analisis Hasil Uji Coba Aplikasi dan Penentuan Pelanggan KTS.....	79
5.4.1.	Analisis Hasil Uji Kebenaran Aplikasi.....	79
5.4.2.	Analisis Penentuan Pelanggan KTS.....	79
5.5.	Kesimpulan Hasil Eksperimen.....	83
BAB VI	KESIMPULAN DAN SARAN.....	85
6.1.	Kesimpulan.....	85
6.2.	Saran.....	86

DAFTAR PUSTAKA	87
BIODATA PENULIS	91
LAMPIRAN A	A-1
LAMPIRAN B	B-1
LAMPIRAN C	C-1

Halaman ini sengaja dikosongkan.

DAFTAR GAMBAR

Gambar 2.1 Hulu-Hilir Air Minum PDAM Surabaya.....	8
Gambar 2.2 Interpretasi Hasil Pearson's Correlation.....	17
Gambar 2.3 Ilustrasi Pola Data Penurunan Tiba-Tiba	19
Gambar 2.4 Ilustrasi Penggunaan Parameter Algoritma Jendela	21
Gambar 2.5 Ilustrasi Pergeseran Jendela.....	21
Gambar 2.6 Alur Metode Algoritma Jendela	24
Gambar 3.1 Metodologi Tugas Akhir	28
Gambar 4.1 Ilustrasi Penggabungan Data	36
Gambar 4.2 Ilustrasi kelipatan 3 bulanan.....	39
Gambar 4.3 Ilustrasi kelipatan 6 bulanan.....	40
Gambar 4.4 Ilustrasi kelipatan 12 bulanan.....	40
Gambar 4.5 Ilustrasi kelipatan 24 bulanan.....	41
Gambar 4.6 Ilustrasi 6 bulan setiap jendela	41
Gambar 4.7 Ilustrasi 5 bulan setiap jendela	42
Gambar 4.8 Ilustrasi 4 bulan setiap jendela	42
Gambar 4.9 Ilustrasi 3 bulan setiap jendela	43
Gambar 4.10 Ilustrasi 2 bulan setiap jendela	43
Gambar 4.11 Tampilan Antar Muka Aplikasi.....	50
Gambar 4.12 Tampilan Antar Muka Keluaran Aplikasi	51
Gambar 5.1 Grafik Percobaan Metode Pearson's Correlation.....	56
Gambar 5.2 Grafik Percobaan Metode Konsumsi Rendah Abnormal.....	59
Gambar 5.3 Grafik Percobaan Metode Windowed Analysis	63
Gambar 5.4 Nilai Korelasi Pelanggan Konsumsi Tidak Sah	65
Gambar 5.5 Nilai <i>Precision</i> dari Percobaan Nilai Ambang	65
Gambar 5.6 Nilai <i>Recall</i> dari Percobaan Nilai Ambang	66
Gambar 5.7 Nilai ambang optimal metode Windowed Analysis	67
Gambar 5.8 Hasil Recall Uji Coba Pearson's Correlation	69
Gambar 5.9 Hasil Recall Uji Coba Low Abnormal Consumption	70
Gambar 5.10 Hasil Recall Uji Coba Windowed Analysis	72
Gambar 5.11 Masukkan Parameter Dalam Aplikasi	75

Gambar 5.12 Hasil Deteksi Anomali Aplikasi	75
Gambar 5.13 Pola Konsumsi Penurunan 30 persen.....	81
Gambar 5.14 Pola Konsumsi Tidak Terdeteksi Penurunan 30 Persen Maupun Ketiga Metode Statistik	82
Gambar 5.15 Konsumsi Tidak Sah Pada Awal Tahun Analisis	83

SEGMENT PROGRAM

Segmen 4.1 Implementasi Normalisasi Data	45
Segmen 4.2 Implementasi Memilih Rentang Data.....	46
Segmen 4.3 Implementasi metode <i>Pearson's Correlation</i>	47
Segmen 4.4 Implementasi metode konsumsi rendah abnormal	48
Segmen 4.5 Implementasi Metode <i>Windowed Analysis</i>	49
Segmen 4.6 Implementasi Keluaran Aplikasi	52

Halaman ini sengaja dikosongkan.

DAFTAR TABEL

Tabel 4.1 Data-data Yang Digunakan	34
Tabel 4.2 Cuplikan Data Konsumsi Pelanggan Rumah Tangga	34
Tabel 4.3 Cuplikan Data Jumlah Kapita Pelanggan PDAM ..	35
Tabel 4.4 Cuplikan Data Pembayaran Pelanggan KTS.....	35
Tabel 4.5 Contoh Tindakan yang Diharuskan Membayar Denda	36
Tabel 4.6 Cuplikan data Konsumsi dan Kapita	38
Tabel 4.7 Data Konsumsi sebelum Normalisasi	38
Tabel 4.8 Data Konsumsi setelah Normalisasi.....	38
Tabel 5.1 Implementasi Pearson's Correlation 3 Bulanan	54
Tabel 5.2 Implementasi Pearson's Correlation 6 Bulanan	55
Tabel 5.3 Implementasi Pearson's Correlation 12 Bulanan	55
Tabel 5.4 Implementasi Pearson's Correlation 24 Bulanan	56
Tabel 5.5 Implementasi konsumsi rendah abnormal 3 bulan	57
Tabel 5.6 Implementasi konsumsi rendah abnormal 6 bulan	58
Tabel 5.7 Implementasi konsumsi rendah abnormal 12 bulan	58
Tabel 5.8 Implementasi konsumsi rendah abnormal 24 bulan	59
Tabel 5.9 Implementasi <i>windowed analysis</i> 6 bulan	60
Tabel 5.10 Implementasi <i>windowed analysis</i> 5 bulan	61
Tabel 5.11 Implementasi <i>windowed analysis</i> 4 bulan	61
Tabel 5.12 Implementasi <i>windowed analysis</i> 3 bulan	62
Tabel 5.13 Implementasi <i>windowed analysis</i> 2 bulan	62
Tabel 5.14 Hasil Uji Coba Metode Pearson's Correlation	68
Tabel 5.15 Hasil Uji Coba Metode Konsumsi Rendah Abnormal	70
Tabel 5.16 Hasil Uji Coba Metode Windowed Analysis	71
Tabel 5.17 Pelanggan KTS PDAM Januari – Desember 2015	73
Tabel 5.18 Hasil Deteksi Manual <i>Microsoft Excel</i>	76
Tabel 5.19 Hasil Penentuan Pelanggan KTS.....	78

Tabel 5.20 Detail Pelanggan yang Tidak Terdeteksi.....80

BAB I

PENDAHULUAN

Pada bagian pendahuluan ini dijelaskan mengenai latar belakang, masalah yang akan diselesaikan, batasan masalah, tujuan serta manfaat yang dihasilkan dari tugas akhir ini.

1.1. Latar Belakang Masalah

Air minum berasal dari sumber daya alam yang penting bagi manusia. Air minum sukar untuk dimurnikan, mahal untuk disalurkan, dan tidak mungkin untuk digantikan [1]. Air minum merupakan salah satu kebutuhan dasar bagi kualitas dan keberlanjutan hidup manusia oleh karenanya air minum mutlak harus tersedia [2].

Kondisi antara keberadaan air minum dan pertumbuhan konstan dari populasi di seluruh dunia serta perkembangan dari aktivitas industri menuntut kualitas juga kuantitas yang tinggi terhadap air [3]. Persediaan air minum umumnya dipenuhi oleh perusahaan pasokan air minum daerah. Perusahaan pasokan air daerah ini bertugas menyediakan air dalam jumlah yang cukup, pada tekanan yang tepat, kualitas yang dapat diterima, dan se-ekonomis mungkin yang dapat dibuat [4]. PDAM Surya Sembada Kota Surabaya sebagai perusahaan daerah pemasok air untuk daerah Pasuruan, Sidoarjo, Gresik, dan Surabaya meningkatkan produksi air minum dari tahun ke tahun untuk memenuhi kebutuhan penduduk yang terus bertambah. PDAM Kota Surabaya tercatat melakukan produksi air pada akhir tahun 2014 sekitar 294 juta meter kubik dan pada tahun 2015 sekitar 299 juta meter kubik air bersih [5]. Jumlah tersebut untuk melayani 92 persen penduduk Kota Surabaya dari total penduduk hampir mencapai 2.9 juta jiwa pada akhir tahun 2014 [6] [7] dan pendapatan yang didapatkan oleh PDAM Kota Surabaya dalam tahun yang sama adalah sebesar 674 juta rupiah [8].

Salah satu permasalahan utama yang mempengaruhi perusahaan air minum di negara berkembang adalah NRW. *Non-revenue water* (NRW) adalah nilai air yang didistribusikan terhadap jumlah pendapatan yang didapatkan berbeda. Di seluruh bagian dunia, sekitar 48 miliar meter kubik air hilang tiap tahunnya disebabkan oleh NRW [9]. NRW disebabkan oleh tiga hal, yaitu: konsumsi resmi tidak berekening, kehilangan akibat kesalahan teknis, dan nonteknis. Konsumsi resmi tidak berekening bukan mencerminkan operasional kerja yang tidak efektif melainkan sebuah kebijakan untuk menyediakan air tanpa menarik biaya [9]. Faktor kehilangan teknis adalah kehilangan air dari jalur distribusi air yang mencakup kebocoran pada saluran distribusi dan penyimpanan air yang meluap. Faktor kehilangan nonteknis dapat diakibatkan oleh meter pelanggan tidak akurat, kesalahan pencatatan tagihan pelanggan, dan penggunaan air yang ilegal. Kehilangan teknis dan nonteknis tersebut mengakibatkan berkurangnya pendapatan perusahaan dari air yang telah diproduksi serta berubahnya pola penggunaan air oleh pelanggan [10].

Dari masing-masing penyebab NRW tersebut, lebih dari 75 persen dari total jumlah NRW belum dapat di-identifikasi kehilangannya oleh PDAM Surya Sembada Surabaya. Persentase tersebut mencapai 61 juta kubik dari 77 juta kubik air minum belum dapat teridentifikasi pada akhir tahun 2014 dan 2015 [5]. Sebuah penelitian di wilayah Asia yang dilakukan oleh *Asian Development Bank* (ADB) pada tahun 2003 mengestimasi terdapat sekitar 50-65 persen kehilangan air disebabkan oleh kesalahan nonteknis [11]. Maka, jika diestimasi terdapat maksimal 50 juta meter kubik air minum disebabkan kehilangan nonteknis.

PDAM mengidentifikasi kehilangan air nonteknis melalui dua hal yaitu dengan analisis anomali terhadap pemakaian air pelanggan menggunakan aplikasi dan pelaporan masyarakat sekitar lokasi mengenai tidak dapat air (TDA). Analisis terhadap anomali yaitu melakukan identifikasi dan klarifikasi

terhadap naik turunnya penggunaan air secara drastis, yaitu 30 persen lebih tinggi maupun lebih rendah dari bulan lalu [12]. Sedangkan pelaporan masyarakat sekitar mengenai TDA salah satunya dapat mengindikasikan terdapat masyarakat yang melakukan sambungan ilegal terhadap pipa distribusi PDAM, sehingga tekanan air untuk menyalurkan air menjadi tidak normal. PDAM Surya Sembada Surabaya mencatat pada akhir tahun 2014 terdapat 1.613 kasus dan pada tahun 2015 tercatat terdapat hampir sebanyak 2.000 kasus konsumsi air tidak sah, dengan kehilangan air yang teridentifikasi pada akhir tahun 2014 sebesar 578 ribu meter kubik dan tahun 2015 sebanyak 688 ribu meter kubik [13] [14].

Konsumsi tidak resmi dapat diindikasikan melalui pola konsumsi abnormal dalam data konsumsi pelanggan. Konsumsi tidak resmi adalah konsumsi air minum yang seharusnya dibayarkan, namun tidak tercatat dalam tagihan perusahaan dikarenakan beberapa faktor nonteknis. Pola abnormal dalam data konsumsi pelanggan dapat dideteksi dengan berkurangnya konsumsi dan rendahnya konsumsi pelanggan dalam beberapa periode waktu. Tetapi, pola abnormal dalam data konsumsi pelanggan tidak hanya disebabkan oleh faktor-faktor nonteknis, dapat juga dikarenakan berubahnya jumlah penghuni dalam satu rumah dan berubahnya kebiasaan konsumsi [15]. Pelanggan dengan pola konsumsi abnormal seperti ini selanjutnya akan dimasukkan ke dalam daftar pelanggan yang harus di-investigasi lebih lanjut oleh PDAM Surya Sembada Surabaya. Investigasi tersebut dilakukan untuk dapat menekan kerugian dari konsumsi tidak resmi yang diakibatkan oleh faktor-faktor nonteknis. Kerugian dari kehilangan air ini dialami oleh PDAM daerah Surabaya menyentuh angka 177 miliar rupiah pada tahun 2014 dan 2015 [5] [14].

Dalam tugas akhir ini digunakan dua metode statistik dan satu algoritma statistik yang dapat mendeteksi pola anomali pada konsumsi air. Metode ini terdiri dari metode mendeteksi penurunan progresif dari pola konsumsi, metode mendeteksi

konsumsi rendah dan abnormal pelanggan dan algoritma untuk mendeteksi konsumsi turun secara tiba-tiba. Pola konsumsi anomali seperti itu merugikan PDAM dikarenakan hal tersebut merupakan indikasi adanya air yang dikonsumsi namun tidak dibayarkan. Berbeda halnya dengan pola konsumsi air minum yang naik secara drastis, hal tersebut memang hal yang dianggap anomali tetapi tidak dilakukan analisa karena tidak merugikan PDAM Surya Sembada Surabaya. Pola konsumsi air minum yang mengalami kenaikan secara drastis tidak dianggap sebagai konsumsi air minum tidak sah. Metode pertama yang digunakan adalah menggunakan korelasi produk momen (*Pearson's correlation*), kedua menggunakan algoritma konsumsi rendah abnormal, dan ketiga algoritma *windowed analysis* menggunakan kombinasi statistik deskriptif dengan pergeseran jendela. Metode korelasi produk momen mengidentifikasi hubungan antara konsumsi pada pelanggan PDAM dengan rentang data yang telah ditentukan. Metode konsumsi rendah abnormal (*algorithm based on abnormally low consumption*) adalah mengidentifikasi pelanggan mengkonsumsi air minum dalam satu tahun secara rendah yang tidak sesuai dengan jumlah kapita tiap pelanggan PDAM. Algoritma *windowed analysis* mengidentifikasi penurunan tiba-tiba dan kemudian konsumsi rendah juga stabil pada konsumsi air minum pelanggan yang pola konsumsi rentang data sebelumnya normal serta relatif stabil [15].

Dari dua metode dan satu algoritma tersebut diterapkan dengan maksud mendeteksi variasi dari masing-masing pola konsumsi yang memungkinkan terjadinya kehilangan air oleh faktor nonteknis. Metode pertama dapat mendeteksi peningkatan frekuensi penipuan yang dilakukan oleh pelanggan dengan contoh semakin lama menaruh alat manipulasi terhadap alat ukur dari bulan ke bulan ataupun alat ukur yang semakin lama semakin menurun perhitungannya akurasi. Metode kedua digunakan untuk mengidentifikasi pola abnormal dari kalkulasi penggunaan minimum air minum tiap orang dalam satu tahun terakhir. Algoritma *windowed analysis* dapat mendeteksi

manipulasi terhadap alat ukur meter secara permanen, sambungan ilegal tanpa melewati alat meter air, atau kerusakan alat ukur. Pola konsumsi pelanggan memenuhi kriteria dari metode-metode dan algoritma tersebut tidak serta merta pelanggan tersebut melakukan konsumsi air minum tidak resmi atau kehilangan akibat faktor nonteknis, namun mungkin saja disebabkan oleh berubahnya jumlah orang yang tinggal dalam satu rumah, rumah tertutup, berubahnya kebiasaan konsumsi [15].

Dengan hasil tugas akhir ini diharapkan dapat memperkirakan, mengevaluasi, dan mengawasi kelompok-kelompok pelanggan yang memiliki pola konsumsi air minum tidak normal yang terjadi di PDAM Surya Sembada Surabaya yang mana dari kelompok tersebut berpotensi diakibatkan oleh konsumsi tidak resmi atau kehilangan air minum oleh faktor nonteknis.

1.2. Perumusan Masalah

Perumusan masalah yang diangkat pada tugas akhir ini adalah bagaimana penentuan parameter model dari metode *Pearson's Correlation*, metode *Low Abnormally Consumption*, dan algoritma *windowed analysis* yang paling optimal digunakan dalam menentukan pola konsumsi abnormal pelanggan PDAM Surya Sembada Surabaya yang telah mengkonsumsi air minum secara tidak sah.

1.3. Batasan Tugas Akhir

Batasan masalah pada tugas akhir ini adalah :

- a. Cakupan subyek yang dianalisis adalah kelompok pelanggan rumah tangga PDAM Surya Sembada Surabaya.
- b. Data yang digunakan untuk diolah adalah data penggunaan air pelanggan dan jumlah jiwa tiap pelanggan dalam dua tahun terakhir Januari 2014 hingga Desember 2015
- c. Metode yang digunakan adalah *Pearson's Correlation*, *Low Abnormally Consumption*, dan algoritma *Windowed Analysis*.

1.4. Tujuan Tugas Akhir

Tugas akhir ini bertujuan untuk mendeteksi secara optimal pelanggan yang telah mengkonsumsi air minum secara tidak sah melalui analisa pola konsumsi pelanggan PDAM Surya Sembada Surabaya dengan menggunakan metode dan algoritma statistik.

1.5. Manfaat Tugas Akhir

Diharapkan dari penyelesaian tugas akhir ini dapat memberikan manfaat kepada beberapa subyek, yaitu:

- a. Bagi akademisi, dapat digunakan untuk mengetahui hasil analisis konsumsi air minum tidak sah oleh pelanggan di PDAM Surya Sembada Surabaya berdasarkan metode statistik.
- b. Bagi perusahaan, khususnya departemen analisis pemakaian air dapat membantu petugas pemantauan pemakaian konsumsi air untuk menganalisis pola abnormal dengan aplikasi berbasis metode-metode statistik sebagai rekomendasi untuk menekan jumlah dari NRW.

1.6. Relevansi

Relevansi tugas akhir ini terhadap jurusan Sistem Informasi adalah merupakan penerapan dari beberapa mata kuliah yang diantaranya adalah Statistika, Penggalan Data dan Analitika Bisnis (PDAB), Analisis dan Desain Perangkat Lunak (ADPL), dan Rancang Bangun Perangkat Lunak (RBPL).

BAB II

TINJAUAN PUSTAKA

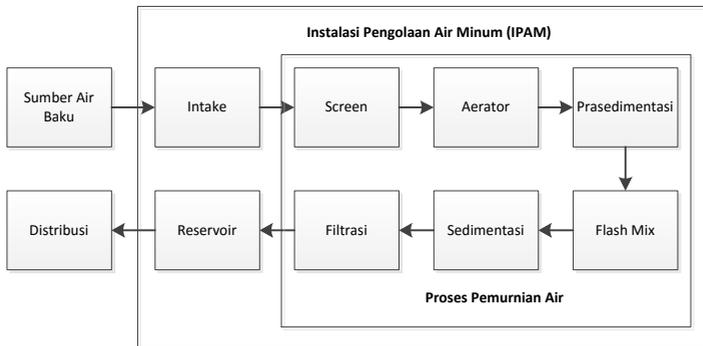
Pada bab ini tertulis dasar pengetahuan dan teori yang digunakan sebagai acuan dalam pengerjaan tugas akhir. Teori-teori yang dijelaskan pada bab ini meliputi teori mengenai deteksi anomali dan bagaimana cara melakukan deteksi anomali pada pola konsumsi air minum serta teori mengenai metode yang akan diimplementasikan pada aplikasi.

2.1. PDAM Surya Sembada Surabaya

Perusahaan Daerah Air Minum (PDAM) Kota Surabaya merupakan perusahaan umum milik daerah yang bertanggung jawab terhadap penyediaan air bersih bagi masyarakat kota Surabaya sejak tahun 1976. Saat ini, PDAM Kota Surabaya telah memiliki jumlah pelanggan sekitar 520 ribu pelanggan pada akhir tahun 2014, yang terdiri dari pelanggan domestik (rumah tangga), dan non domestik (pemerintah, komersial, dan industri). Dengan jumlah penduduk yang ada saat ini, PDAM Surabaya dituntut dapat memberikan pelayanan untuk memenuhi kebutuhan air bersih masyarakat di Kota Surabaya. Sumber air baku yang digunakan berasal dari sumber mata air tanah yang berada di daerah Umbulan, Pasuruan dan air dari sungai di Kota Surabaya. Pengembangan sistem air bersih Kota Surabaya sering dilakukan dengan berbagai kegiatan pengembangan sarana, seiring dengan peningkatan kualitas dan kuantitas air minum. Untuk memenuhi kebutuhan air minum tersebut, saat ini telah dioperasikan Instalasi Penjernihan Air Minum (IPAM) di dua lokasi Surabaya, yaitu di Ngagel dan di Karangpilang. IPAM Ngagel terbagi menjadi IPAM Ngagel I, Ngagel II, dan Ngagel III, sedangkan IPAM Karang Pilang terbagi menjadi PDAM Karang Pilang I, II, dan III. Sehingga saat ini terdapat 6 IPAM yang telah beroperasi untuk memenuhi kebutuhan air minum di wilayah Surabaya dan sekitarnya.

2.1.1. Gambaran Umum Hulu-hilir Air Minum PDAM

Proses distribusi air PDAM Surya Sembada Surabaya bermula dari air baku kemudian melalui proses produksi hingga menjadi air minum yang siap dikonsumsi oleh pelanggan [16]. Hulu hingga hilir air dapat dilihat pada gambar 2.1.



Gambar 2.1 Hulu-Hilir Air Minum PDAM Surabaya

Berikut adalah gambaran umum penjelasan dari proses distribusi air dari hulu hingga hilir pada PDAM Surya Sembada Surabaya.

a. Sumber Air Baku

Penyediaan air minum terdapat beberapa sumber yang kemudian dapat diolah menjadi air minum. Sumber air minum yang dapat digunakan adalah mata air, air permukaan, air tanah dangkal, dan Sumur Artesis. Air yang keluar dari mata air umumnya jernih dan memenuhi sebagai syarat air minum. Air dari mata air sebelum didistribusikan harus dikaptir atau dibungkus agar tidak ada pengotoran dari luar dan gangguan-gangguan lain. Air permukaan merupakan air yang terdapat dalam sungai, parit, saluran irigasi, atau danau. Tingkat kualitas air dari air permukaan juga berubah-ubah tergantung kondisi lingkungan sendiri. Air permukaan ini harus dilakukan pengolahan sebelum dapat dikonsumsi sebagai air minum.

Air tanah dangkal adalah air yang diperoleh dari dalam tanah kurang dari 15 meter dalamnya. Air tanah dangkal dimanfaatkan untuk memenuhi keperluan rumah tangga akan air minum. Air tanah dangkal tersebut sebagian besar dibuat menjadi sumur sebelum dikonsumsi. Sumur artesis merupakan sumber air yang lebih dalam dibandingkan dengan sumur yang dibuat dari air tanah dangkal. Kedalaman sumur artesis sendiri memiliki perbedaan disetiap tempat dan hanya dapat ditentukan dengan melakukan pengeboran percobaan.

b. Intake

Intake atau biasa disebut bangunan penangkap air adalah sebuah sumur yang berfungsi meratakan debit air yang masuk. Bangunan intake ini berguna sebagai penyeimbang beban pemompaan air ke instalasi pengolahan. Bangunan intake juga terdapat screen yang bertujuan untuk menyaring benda-benda terapung seperti sampah.

c. Proses Pemurnian air

1) Aerator

Aerator merupakan alat untuk aerasi yang digunakan untuk menyisihkan kandungan organik dan gas yang terlarut di permukaan atau untuk menambah oksigen ke air untuk mengubah substansi yang dipermukaan menjadi suatu oksida.

2) Pra-Sedimentasi

Bangunan pra-sedimentasi bertujuan untuk meminimalisasi kekeruhan air yang disebabkan oleh partikel makro seperti lumpur.

3) Flash Mix

Pada fase ini adalah penyuntikan bahan kimia yang dilakukan setelah proses pra-sedimentasi. Bahan kimia ini diinjeksi dengan tujuan

4) Sedimentasi dan filtrasi

Fase sedimentasi dan filtrasi merupakan sebuah fase pemurnian dalam proses air baku. Perbedaan antara keduanya adalah fase sedimentasi melakukan

pengendapan terhadap residu air yang tidak dibutuhkan, sedangkan pada fase filtrasi dilakukan penyaringan untuk menyaring partikel-partikel dalam air yang tidak dibutuhkan.

d. Reservoir

Reservoir merupakan tempat penyimpanan cadangan air yang telah melalui proses pemurnian setelah melewati media filter. Reservoir berfungsi untuk menampung air bersih yang akan disalurkan ke pelanggan melalui pipa-pipa distribusi

e. Distribusi

Pada bagian distribusi ini PDAM Surya Sembada Surabaya menggunakan pipa-pipa yang langsung menuju masing-masing sektor. Masing-masing IPAM mendistribusikan airnya ke beberapa sektor spesifik. Sebagai contoh

IPAM Karang Pilang III melayani industri dan warga di Surabaya Utara dan Timur, di antaranya daerah Made, Sambikerep, Pakal, Kenjeran, Wonorejo dan beberapa daerah di Surabaya Timur.

2.2. Rugi-rugi Distribusi Air Minum

Sistem distribusi air perkotaan sering kali dikubur dan dilupakan sampai sistem distribusi tersebut menunjukkan kebocoran dan semburan, hal tersebut menyebabkan kehilangan air secara signifikan dan kehilangan pendapatan yang besar. Menurut penelitian yang dilakukan *World Bank*, terdapat sekitar 48 miliar meter kubik air hilang tiap tahunnya dari saluran distribusi. Jumlah dari kehilangan air atau NRW adalah ukuran terhadap ke-efektifan dari operasional dari sistem distribusi air [17]. Kerugian air dapat disebabkan 2 tipe utama, yaitu kerugian oleh faktor teknis dan nonteknis.

2.2.1. Rugi Teknis

Kerugian air disebabkan oleh faktor teknis merupakan kehilangan air dari jalur distribusi air yang mencakup

kebocoran pada saluran distribusi, penyimpanan air (reservoir) yang meluap, dan seluruh kehilangan pada koneksi distribusi menuju pelanggan [10]. Kerugian oleh faktor teknis ini dapat dianalisa melalui pengamatan kebocoran, deteksi kebocoran, pembenahan kebocoran, analisa terhadap tekanan air, dan manajemen aset [17].

2.2.2. Rugi Nonteknis

Kerugian oleh faktor nonteknis adalah kehilangan air yang bukan disebabkan oleh faktor teknis dari sistem distribusi. Kehilangan nonteknis disebabkan oleh 4 hal utama, yaitu ketidak akuratan meter, pembacaan meter tidak tepat, penanganan data dan kesalahan penagihan, dan yang terakhir adalah konsumsi ilegal dari pelanggan. Kerugian oleh faktor nonteknis ini juga menyumbang kerugian berkali lipat dibandingkan oleh kehilangan oleh faktor teknis, dikarenakan air yang hilang oleh nonteknis adalah air yang telah melalui proses pemurnian terdapat variabel harga produksi dalam kerugian yang dialami perusahaan yang kemudian dikonsumsi oleh pelanggan perusahaan secara tidak resmi [17]. Kehilangan air oleh faktor nonteknis menyumbang sebanyak 50-65% dari seluruh kehilangan air yang terjadi pada negara-negara di wilayah Asia [11].

Banyak penelitian yang ditujukan untuk dapat menanggulangi kehilangan air oleh faktor nonteknis ini. Metodologi yang dapat mendeteksi lokasi dan ukuran cabang ilegal dalam pipa distribusi yaitu dengan *transient test* [18] dan pengukuran tekanan dengan *inverse calculation* untuk mendeteksi konsumsi tidak resmi [19]. Penelitian yang berfokus kepada penanganan ketidak-akuratan meter yaitu metodologi *Markov dynamic* yang dibuat dalam *Bayesian framework* untuk mengetahui hubungan antara akurasi meter dengan laju degradasi akurasi meter sehingga mengetahui waktu yang tepat untuk mengganti meter air [20]. Pendekatan sosial budaya juga menjadi penelitian di beberapa kota di Asia yang bekerja sama dengan beberapa komunitas lokal yang telah terbukti dapat menekan angka kehilangan air [21]. Sedangkan, jika melalui penggalan data

pelanggan yang dimiliki oleh perusahaan, deteksi kehilangan melalui faktor nonteknis dapat dianalisis melalui pola konsumsi pelanggan. Analisis pola konsumsi pelanggan dapat mengidentifikasi pelanggan yang memiliki potensi telah melakukan kecurangan atau memiliki meter pencatat yang tidak akurat. Pola konsumsi pelanggan dianalisis untuk dideteksi dalam beberapa kurun waktu apakah ada pola abnormal didalamnya yang memungkinkan pelanggan telah mengkonsumsi air secara tidak resmi. Pola abnormal pada konsumsi air minum memiliki kriteria pola konsumsi yang rendah atau mengalami penurunan dalam beberapa periode waktu [15].

2.3. Penelitian Analisis Pola Konsumsi Terdahulu

Penelitian terdahulu dilakukan oleh Inigo Monedero, Felix Biscarri, Juan I. Guerrero, Manuel Pena, Moises Roldan, dan Carlos Leon (2015) dengan judul “*Detection of Water Meter Under-Registration Using Statistical Algorithms*”. Dalam penelitian ini dijelaskan mengenai penggunaan ketiga algoritma statistik menggunakan IBM Modeler 15. Penelitian ini dilakukan untuk menganalisis pola konsumsi pelanggan perusahaan distribusi air minum di Seville, Spanyol. Penelitian di Seville ini dilakukan oleh Inigo Monedero dikarenakan sedang bertambahnya jumlah kasus penipuan sejalan dengan krisis finansial yang terjadi di Spanyol. Pada penelitian ini algoritma statistik yang digunakan adalah korelasi produk momen, atau *windowed analysis*, dan algoritma berdasarkan konsumsi rendah abnormal. Dalam tahap percobaan penelitian ini parameter atau hasil dari masing-masing algoritma statistik yang digunakan adalah -0.6 hingga -1.0 untuk korelasi produk momen, parameter dalam *windowed analysis* yaitu *desv* dan *desc* masing-masing 2 atau 2 dan 1.5 untuk *desv* serta *desc* secara berurutan, dan algoritma berdasar konsumsi rendah abnormal menggunakan nilai konsumsi air minum per kapita adalah sebesar $10m^2$. Hasil akhir dari penelitian ini yang dilakukan terhadap sekitar daerah Seville, Spanyol yang memiliki pelanggan berjumlah 1164. Setelah dilakukan proses

analisis menggunakan algoritma statistik terdapat sejumlah 334 pelanggan yang berpotensi melakukan tindak kecurangan dan kemudian dilakukan inspeksi terdapat 30 pelanggan yang terbukti melakukan tindak kecurangan atau memiliki tingkat keberhasilan sebesar 9 persen [15].

2.4. Dasar Teori

Dalam sub-bab tugas akhir ini dijelaskan mengenai teori dari teknis yang dilakukan untuk menganalisa pola abnormal dari konsumsi air minum PDAM Surya Sembada Surabaya. Masing-masing dasar teori ini dikutip dari beberapa publikasi ilmiah seperti teori normalisasi, metode statistik, dan algoritma statistik.

2.4.1. Normalisasi

Normalisasi merupakan bagian dari pra-proses data yang bertujuan untuk menyamakan skala atribut data kedalam sebuah range yang spesifik. Beberapa metode normalisasi yang biasa digunakan adalah :

1. Min-Max
Metode normalisasi dengan melakukan transformasi linier terhadap data asli.
2. Z-Score
Merupakan metode normalisasi yang berdasarkan mean (nilai rata-rata) dan standard deviation (deviasi standar) dari data.
3. Decimal Scalling
Metode ini melakukan normalisasi dengan menggerakkan nilai desimal dari data ke arah yang diinginkan
4. Sigmoidal
Melakukan normalisasi data secara nonlinier ke dalam range $-1 - 1$ dengan menggunakan fungsi sigmoid
5. Softmax
Metode ini merupakan pengembangan transformasi secara linier. Output range-nya adalah 0-1.

2.4.1.1. Metode *min-max*

Metode Min-Max merupakan salah satu metode normalisasi yang paling sederhana. Metode ini mentransformasikan data sehingga memiliki nilai antara 0 hingga 1 [22]. Nilai yang paling minimal adalah 0 dan nilai maksimal adalah 1. Rumus dari metode *min-max* dapat dilihat pada persamaan nomor (2.1).

$$MM(X_{ij}) = \frac{X_{ij} - X_{min}}{X_{max} - X_{min}} \quad (2.1)$$

Keuntungan dari metode ini adalah keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses normalisasi. Tidak ada data bias yang dihasilkan oleh metode ini. Kekurangannya adalah jika ada data baru, metode ini akan memungkinkan terjebak pada "out of bound" error.

2.4.2. Metode Statistik

Metode statistik adalah sebuah cara dengan berbasis statistik untuk dapat mencapai tujuan tertentu. Dalam tugas akhir ini metode statistik yang digunakan untuk mendeteksi pola abnormal konsumsi air minum pelanggan PDAM Surya Sembada Surabaya adalah korelasi produk momen (*Pearson's Correlation*) dan metode konsumsi rendah abnormal (*abnormally low consumption*). Metode korelasi produk momen digunakan untuk dapat mendeteksi penurunan progresif konsumsi air pelanggan. Penurunan progresif ini salah satu indikasi dari peningkatan progresif penipuan yang dilakukan oleh pelanggan [15]. Metode *abnormally low consumption* digunakan untuk mendeteksi konsumsi rendah dibawah estimasi rata-rata. Meskipun konsumsi rendah belum tentu sesuai dengan gangguan meter atau beberapa jenis kegagalan pengukuran, metode juga sangat relevan dalam pendeteksian konsumsi anomali [15].

2.4.2.1. Metode Berdasar Korelasi Produk Momen (*Pearson's Coefficient*)

Korelasi momen produk (*Pearson's coefficient*) adalah sebuah teknik statistik untuk menganalisis hubungan antara variabel. Sebuah perhitungan korelasi diantara kumpulan data biner atau kontinu yang memperlihatkan derajat hubungan linier antar satu dengan yang lain [23]. Persamaan korelasi momen produk dapat dilihat pada persamaan nomor (2.2) dan nomor (2.3).

$$\text{corr}(x,y) = \frac{\text{covariance}(x,y)}{\text{standard_deviation}(x) * \text{standard_deviation}(y)} = \frac{S_{xy}}{S_x * S_y} \quad (2.2)$$

Pada persamaan nomor (2.2) $\text{corr}(x,y)$ atau r adalah korelasi momen produk yang merupakan hasil dari korelasi antara variabel x dan y . Hasil dari korelasi produk momen (r) berjarak antara -1 (korelasi negatif sempurna atau berbanding terbalik) hingga $+1$ (korelasi positif sempurna atau berbanding lurus). Semakin tinggi nilai dari r semakin tinggi nilai dari variabel x [15]. Penjabaran persamaan nomor (2.2) dapat dilihat ke persamaan nomor (2.3).

$$-1 \leq \frac{S_{xy}}{S_x * S_y} = \frac{\frac{1}{N}(\sum_{t=1}^n (X_t - \bar{X}) x (Y_t - \bar{Y}))}{\sqrt{\sum_{t=1}^n (X_t - \bar{X})^2} x \sqrt{\sum_{t=1}^n (Y_t - \bar{Y})^2}} \leq 1 \quad (2.3)$$

Persamaan $\text{Corr}(x,y)$ merupakan pembagian dari $\text{covariance}(x,y)$ dengan standar deviasi (x,y) . $\text{Covariance}(x,y)$ adalah sebuah ukuran bagaimana kedua variabel x,y berhubungan [24] sedangkan $\text{standard_deviation}(x,y)$ merupakan ukuran dari sekumpulan persebaran data. Standar deviasi dibentuk dari persebaran(deviasi) dari setiap data dari rata-rata [25] Persamaan dari $\text{covariance}(x,y)$ dapat dilihat pada (2.4) dan persamaan dari $\text{standard_deviation}(x,y)$ dapat dilihat pada persamaan (2.5) dan (2.6).

$$\text{Covariance}(x,y) = \frac{1}{n} \left(\sum_{t=1}^n (X_t - \bar{X}) x (Y_t - \bar{Y}) \right) \quad (2.4)$$

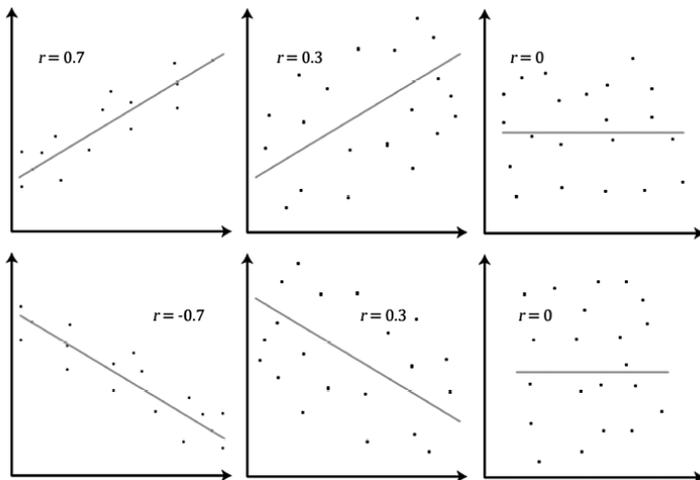
Pada persamaan nomor (2.4) adalah $covariance(x,y)$ di mana X_t, Y_t adalah nilai X atau Y saat t , \bar{X}, \bar{Y} adalah nilai rata-rata dari masing-masing X atau Y , n adalah jumlah populasi. Hasil dari nilai $covariance(x,y)$ merepresentasikan hubungan antara nilai dari kedua variabel, jika hasil dari $covariance(x,y)$ positif mengindikasikan kedua variabel tersebut berhubungan (positif $covariance$) lalu jika hasil dari $covariance(x,y)$ adalah negatif mengindikasikan variabel x,y berhubungan terbalik (negatif $covariance$) dan jika hasil dari $covariance(x,y)$ mendekati 0 baik positif maupun negatif mengindikasikan kedua variabel x,y tidak berhubungan [24].

$$Standard_deviation(x) = S_x = \sqrt{\sum_{k=1}^m (X_k - \bar{x})^2} \quad (2.5)$$

$$Standard_deviation(y) = S_y = \sqrt{\sum_{k=1}^n (y_k - \bar{y})^2} \quad (2.6)$$

Pada persamaan nomor (2.5) dan (2.6) adalah persamaan standar deviasi (σ) untuk masing-masing x dan y . Standar deviasi adalah akar pangkat dua dari $variance$ (σ^2) merupakan pengukuran yang digunakan untuk menghitung nilai dari persebaran dari nilai data. Semakin besar nilai standar deviasi semakin tersebar data dari data distribusi sedangkan semakin mendekati 0 nilai standar deviasi berarti persebaran data dekat dengan rata-rata seluruh data [26].

Intepretasi korelasi produk momen dari satu set data x dan y dapat dilihat pada Gambar 2.2.



Gambar 2.2 Interpretasi Hasil Pearson's Correlation

Nilai r positif menunjukkan hubungan kedua variabel positif, artinya kenaikan nilai satu variabel diikuti nilai variabel lain. Nilai r negatif menunjukkan hubungan kedua variabel negatif artinya menurunnya nilai salah satu variabel diikuti dengan meningkatnya yang lain. Nilai r sama dengan nol menunjukkan kedua variabel tidak menunjukkan hubungan. Nilai r semakin mendekati korelasi sempurna (-1 atau $+1$) memiliki arti semakin kuat hubungan antara kedua variabel (tiap-tiap data semakin mendekati garis) [27].

2.4.2.2. Metode berdasar Konsumsi Rendah Abnormal (*Abnormally Low Consumption*)

Metode berdasar konsumsi rendah abnormal adalah sebuah metode yang mengklasifikasikan pola konsumsi pelanggan dengan berdasarkan estimasi jumlah minimal konsumsi air per kapita atau tiap jiwa dalam kurun waktu tertentu.

Penggunaan dari air dapat menjadi tidak merata pada setiap pelanggan seperti rumah tidak digunakan lagi, berganti bisnis,

budaya penggunaan, peralatan yang digunakan, atau digunakan untuk irigasi. Konsumsi rendah tidak serta-merta selalu berarti berhubungan dalam pencurian air atau penggelapan meteran atau beberapa kesalahan dari meteran pencatat air [15]. Inigo Monedero dalam publikasinya memutuskan untuk mengklasifikasikan pelanggan yang memiliki konsumsi dibawah 10 meter kubik dalam satu tahun terakhir. Angka 10 meter kubik dihasilkan setelah diskusi dari perusahaan dan berdasarkan dari riset statistik dari deteksi kecurangan diwaktu lalu. Maka dari itu dalam tugas akhir ini kalkulasi konsumsi minimum dalam x bulan dapat dilihat pada persamaan (2.7).

$$MYC = P * \text{minimum/day/person} * 30 * x \quad (2.7)$$

Pada persamaan (2.7) MYC (*Minimum Year Consumption*) adalah konsumsi minimal dalam satu nomor rekening dalam satu tahun, P adalah jumlah kapita atau jumlah jiwa dalam satu nomor rekening. *Minimum/day/person* adalah konsumsi minimal jiwa dalam satu hari. Angka 30 adalah jumlah masing-masing hari dalam satu bulan, dan x adalah jumlah bulan sebagai panjang rentang data yang ingin dianalisis. Pelanggan yang memiliki konsumsi dibawah kalkulasi metode (2.7) dalam x bulan terakhir diklasifikasikan sebagai pelanggan yang memiliki konsumsi abnormal.

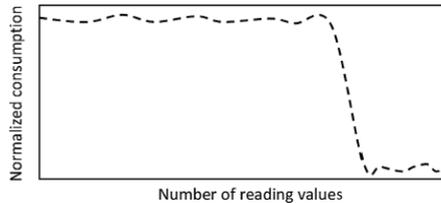
2.4.3. Algoritma Statistik

Algoritma statistik adalah analisis statistik yang digunakan untuk mendeteksi pola yang tidak biasa (abnormal) serta menjelaskan pola menggunakan model statistik [28]. Algoritma statistik yang digunakan, sekaligus metode ketiga dalam tugas akhir ini adalah algoritma *windowed analysis*. Algoritma *windowed analysis* digunakan untuk mendeteksi pola abnormal pada konsumsi pelanggan PDAM Surya Sembada Surabaya.

2.4.3.1. Algoritma Windowed Analysis

Algoritma *windowed analysis* adalah algoritma pencari pola data pelanggan, dari data konsumsi yang telah dinormalisasi.

Algoritma ini bertujuan untuk mendeteksi penggunaan air yang normal serta konstan didalam satu atau lebih jendela awal, kemudian diikuti dengan penurunan penggunaan air secara tiba-tiba kemudian stabil hingga akhir jendela yang digunakan [15]. Contoh pola data dapat dilihat pada Gambar 2.3.



Gambar 2.3 Ilustrasi Pola Data Penurunan Tiba-Tiba

Algoritma berdasar *windowed analysis* dibagi menjadi dua metode, yaitu metode pengidentifikasi penurunan dan metode pergeseran jendela.

2.4.3.1.1. Identifikasi Penurunan Pola Konsumsi

Metode ini dikembangkan dari kombinasi statistik deskriptif yaitu: max (maksimum konsumsi), min (minimal konsumsi), median (nilai tengah konsumsi), dan mean (rata-rata konsumsi) dari data penggunaan yang telah dinormalisasi. Komposisi data dalam jendela dapat ditentukan sesuai kebutuhan. Jendela berisi data penggunaan air secara berurutan dalam x bulan oleh pelanggan. Dari beberapa urutan nilai konsumsi pelanggan tersebut, kemudian dimasukkan kedalam empat jendela. Algoritma *windowed analysis* ini dapat dilihat pada persamaan (2.8), (2.9), (2.10), dan (2.11) yang merupakan satu kesatuan.

Dalam algoritma (2.8), (2.9), (2.10), dan (2.11) akhiran ($_x$) dari setiap variabel mengidentifikasi kalkulasi dari tiap nomor jendela. Kumpulan dari algoritma tersebut digunakan untuk mendeteksi konsumsi abnormal dari pelanggan berdasar pola penggunaan air tiap pelanggan untuk setiap jendelanya.

$$\{[max6 < (median6 \times desv) \text{ and} \quad (2.8)$$

$$min6 > (median6/desv) \text{ and}$$

$$max5 < (median5 \times desv) \text{ and}$$

$$min5 > (median5/desv) \text{ and}$$

$$mean6 < mean5/desc]\}$$

or

$$\{[max6 < (median6 \times desv) \text{ and} \quad (2.9)$$

$$min6 > (median6/desv) \text{ and}$$

$$max5 < (median5 \times desv) \text{ and}$$

$$min5 > (median5/desv) \text{ and}$$

$$max4 < (median4 \times desv) \text{ and}$$

$$min4 > median4/desv) \text{ and}$$

$$mean6 < mean4/desc \text{ and } mean5 < mean4/desc]\}$$

or

$$\{max6 < (median6 \times desv) \text{ and} \quad (2.10)$$

$$min6 > (median6/desv) \text{ and}$$

$$[max5 < (median5 \times desv) \text{ and}$$

$$min5 > (median5/desv)] \text{ and}$$

$$[max4 < (median4 \times desv) \text{ and}$$

$$min4 > (median4/desv)] \text{ and}$$

$$[max3 < (median3 \times desv \text{ and}$$

$$min3 > (median3/desv) \text{ and}$$

$$mean6 < (mean3/desc)] \text{ and}$$

$$[mean5 < (mean3/desc)] \text{ and}$$

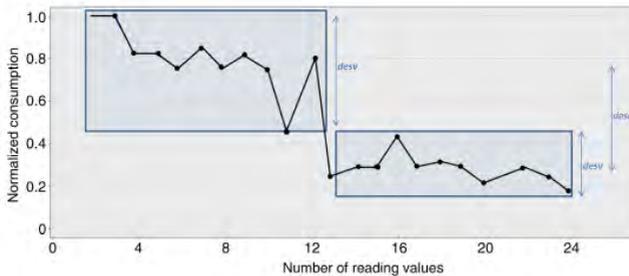
$$[mean4 < mean3/desc]\}$$

And

$$min3 > 0 \text{ and } min4 > 0 \text{ and } min5 > 0 \text{ and } min6 > 0 \quad (2.11)$$

Kalkulasi tiap jendela tersebut (*max*, *min*, *median*, dan *mean*) dikali dengan parameter *desv* atau dibagi dengan parameter *desc*. Dua parameter tersebut dapat dikonfigurasi untuk mendeteksi penurunan tiba-tiba diantara dua jendela. *Desc* merupakan selisih penurunan dari penggunaan yang telah dinormalisasi dari jendela kedua terhadap jendela pertama, sedangkan *desv* adalah lebar vertikal dari masing-masing jendela yang dianalisis atau batas pada penyimpangan dari nilai tengah penggunaan dari tiap jendela. *Desv* berguna untuk mengkategorikan pola konsumsi yang relatif stabil. Tujuan dari kedua parameter ini adalah untuk mengkategorikan pola konsumsi sesuai parameter penurunan drastis (*desc*) dan parameter konsumsi stabil (*desv*) yang telah ditentukan.

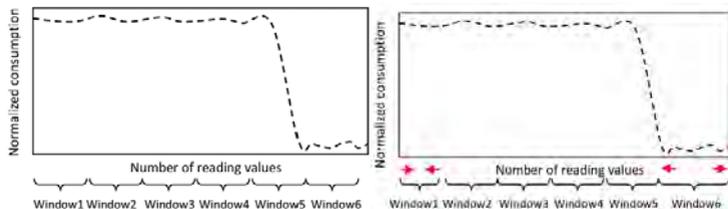
Semakin besar nilai *desv* maka batas fluktuasi dalam satu jendela konsumsi akan semakin lebar, jika sebaliknya nilai yang semakin kecil maka batas fluktuasi yang ditoleransi semakin kecil. Semakin besar nilai *desc* maka selisih antara rata-rata penurunan dua jendela haruslah semakin besar, sebaliknya jika *desc* semakin kecil selisih antara persentase penurunan dua jendela semakin ditoleransi tidak harus terlalu lebar. Sebagai ilustrasi dapat dilihat pada Gambar 2.4.



Gambar 2.4 Ilustrasi Penggunaan Parameter Algoritma Jendela

2.4.3.1.2. Metode Pergeseran Jendela

Pergeseran jendela yang dilakukan untuk pencari konsumsi yang secara tiba-tiba turun drastis. Pergeseran ini dilakukan karena algoritma (2.8), (2.9), (2.10), dan (2.11) hanya dapat mendeteksi penurunan drastis konsumsi pada masing-masing jendela yang berbeda. Metode ini dilakukan untuk memastikan konsumsi abnormal konstan pada jendela setelahnya dan memasukkan konsumsi normal konstan pada jendela terakhir sebelum penurunan. Ilustrasi pergeseran jendela dapat dilihat pada Gambar 2.5.



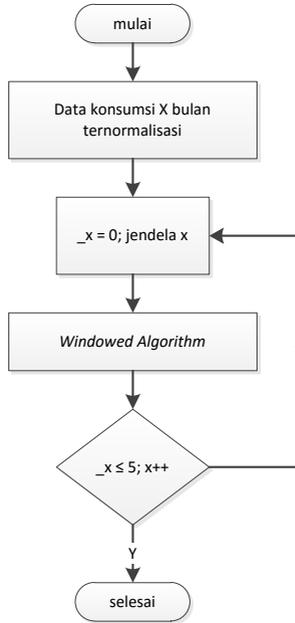
Gambar 2.5 Ilustrasi Pergeseran Jendela

Pada publikasi yang dilakukan Inigo Monedero masing-masing jumlah data dalam satu jendela berjumlah empat. Namun, jumlah data dalam satu jendela ini dapat disesuaikan dengan analisis yang dibutuhkan. Berikut ini adalah langkah-langkah pergeseran jendela dengan menggunakan 4 bulan konsumsi pada masing-masing jendela:

- a. Analisa dilakukan dengan enam jendela dari 18 nilai data historis atau 2 tahun konsumsi pada setiap pelanggan.
- b. Pada pergeseran jendela pertama, masukkan pada masing-masing jendela dengan empat tiap bulan nilai konsumsi pelanggan.
- c. Terapkan satu set jendela dari langkah *b* dengan algoritma jendela (7), (8), (9), dan (10).
- d. Pada pergeseran jendela kedua, masukkan pada jendela terakhir (jendela enam) dengan nilai konsumsi lima bulan terakhir pelanggan, masukkan pada jendela pertama (jendela satu) dengan nilai konsumsi tiga bulan pertama pelanggan dan jendela sisanya dengan masing-masing dengan nilai konsumsi empat bulan secara berurutan.
- e. Terapkan satu set jendela dari langkah *d* dengan algoritma jendela (7), (8), (9), dan (10).
- f. Pada pergeseran ketiga, masukkan pada jendela terakhir (jendela enam) dengan nilai konsumsi enam bulan terakhir pelanggan, masukkan pada jendela pertama (jendela satu) dengan nilai konsumsi dua bulan pertama pelanggan dan jendela sisanya dengan masing-masing dengan nilai konsumsi empat bulan secara berurutan.
- g. Terapkan satu set jendela dari langkah *f* dengan algoritma jendela (7), (8), (9), dan (10).
- h. Pada pergeseran keempat, masukkan pada jendela terakhir (jendela enam) dengan nilai konsumsi dua bulan terakhir pelanggan, masukkan pada jendela pertama (jendela satu) dengan nilai konsumsi enam bulan pertama pelanggan dan

- jendela sisanya dengan masing-masing dengan nilai konsumsi empat bulan secara berurutan.
- i. Terapkan satu set jendela dari langkah h dengan algoritma jendela (7), (8), (9), dan (10).
 - j. Pada pergeseran kelima, masukkan pada jendela terakhir (jendela enam) dengan nilai konsumsi tiga bulan terakhir pelanggan, masukkan pada jendela pertama (jendela satu) dengan nilai konsumsi lima bulan pertama pelanggan dan jendela sisanya dengan masing-masing dengan nilai konsumsi empat bulan secara berurutan.
 - k. Terapkan satu set jendela dari langkah j dengan algoritma jendela (7), (8), (9), dan (10).

Dengan demikian, alur proses algoritma *windowed analysis* dapat dilihat pada Gambar 2.6. Dari serangkaian metode pergeseran jendela dan metode identifikasi penurunan konsumsi tersebut, pergeseran jendela dilakukan sebanyak lima kali. Pergeseran jendela terdapat pada b , d , f , h , dan j . Dari masing-masing kelima pergeseran jendela tersebut didalamnya dijalankan keempat algoritma kombinasi statistik deskriptif yang telah dituliskan pada (7), (8), (9), dan (10). Algoritma dari kombinasi statistik deskriptif akan dijalankan untuk mendeteksi apakah ada pola konsumsi pelanggan yang mengalami penurunan secara tiba-tiba pada salah satu jendela dan dari algoritma tersebut harus terdapat dua sub-algoritma yang benar, yaitu (7) dan (10), atau (8) dan (10), atau (9) dan (10). Algoritma (10) selalu bernilai benar jika keseluruhan jendela analisis tidak terdapat konsumsi nol, karena konsumsi nol merupakan sebuah kejadian yang harus segera ditelusuri oleh perusahaan (kemungkinan alat ukur rusak, rumah kosong, dsb).



Gambar 2.6 Alur Metode Algoritma Jendela

Algoritma (9) bernilai benar jika penurunan drastis kemudian stabil berada pada jendela empat (bulan ke-13, ke-12, ke-11, ke-15, dan ke-14), algoritma (8) bernilai benar jika penurunan drastis kemudian stabil berada pada jendela lima (bulan ke-17, ke-16, ke-15, ke-19, dan ke-18), algoritma (7) bernilai benar jika penurunan drastis kemudian stabil berada pada jendela enam (bulan ke-21, ke-20, ke-19, ke-23, dan ke-22). Dengan kata lain jika terdapat penurunan dari bulan ke-11 kemudian pola konsumsi stabil hingga bulan ke-24 lalu bulan ke-23 kemudian pola konsumsi stabil hingga bulan ke-24 akan tercatat sebagai pola konsumsi abnormal berdasarkan algoritma *windowed analysis*.

2.5. Pengukuran Kinerja

Dalam proses pengerjaan tugas akhir ini diperlukan sebuah pengukuran Kinerja dari metode dan aplikasi yang telah dibuat. Pengukuran yang digunakan dalam tugas akhir ini, yaitu:

a. Akurasi

Akurasi merupakan persentase dari dari total sampel yang benar diidentifikasi. Persamaan dari akurasi dapat dilihat pada persamaan (2.12).

$$Akurasi = \frac{TP + TN}{P + N} \times 100\% \quad (2.12)$$

b. Recall

Recall merupakan perbandingan jumlah prediksi positif yang benar terhadap keseluruhan jumlah data sampel yang benar. Persamaan dari *recall* kurasi dapat dilihat pada persamaan (2.13).

$$Recall = \frac{TP}{P} \times 100\% \quad (2.13)$$

c. Precision

Precision merupakan perbandingan antara jumlah prediksi positif yang benar terhadap keseluruhan jumlah prediksi yang dianggap benar. Persamaan dari *precision* dapat dilihat pada persamaan (2.14).

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (2.14)$$

Halaman ini sengaja dikosongkan

BAB III

METODOLOGI Pengerjaan Tugas Akhir

Dalam bab ini dijelaskan mengenai langkah-langkah pengerjaan dari tugas akhir. Secara garis besar langkah-langkah pengerjaan tugas akhir dimulai dari pengumpulan data konsumsi air oleh pelanggan di wilayah Surabaya. Kemudian dilakukan pengolahan terhadap data untuk menghasilkan atribut yang merepresentasikan pola konsumsi air oleh pelanggan dalam kurun waktu 2 tahun terakhir. Pada akhirnya disusun laporan berdasarkan hasil pengerjaan tugas akhir.

3.1. Diagram Metodologi

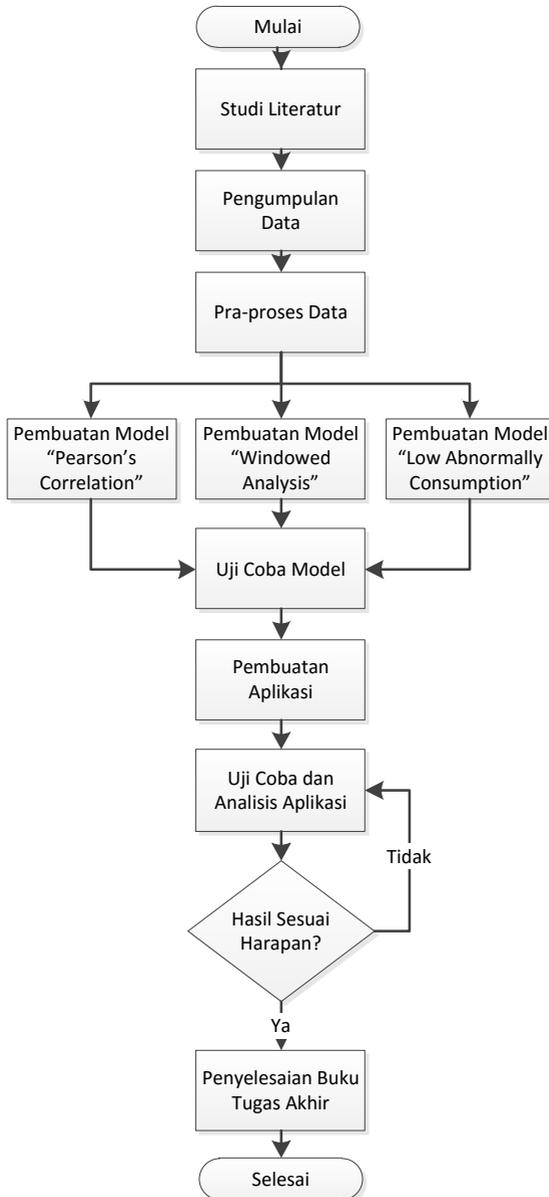
Dalam sub-bab ini diperlihatkan diagram pengerjaan dari tugas akhir yang dikerjakan. Metodologi pengerjaan tugas akhir ini berawal dari studi literatur hingga penyelesaian buku tugas akhir ini. Diagram metodologi dapat dilihat pada Gambar 3.1.

3.2. Uraian Metodologi

Pada uraian metodologi pada Gambar 3.1 dijelaskan mengenai masing-masing proses metodologi yang digunakan dalam tugas akhir ini.

3.2.1. Studi Literatur

Studi literatur dapat dilakukan dengan pengumpulan berbagai referensi, seperti wawancara narasumber, buku, penelitian sejenis, dan dokumen yang terkait. Studi literatur didasarkan pada topik yang telah didapatkan sebelumnya. Pada tugas akhir ini, diusulkan topik mengenai analisis pola konsumsi pelanggan yang memiliki potensi melakukan konsumsi tidak sah terhadap perusahaan PDAM Surya Sembada Kota Surabaya dengan metodologi klasifikasi tertentu. Melalui studi literatur ini pula diperoleh teknik analisis anomali yang sesuai dengan permasalahan studi kasus.



Gambar 3.1 Metodologi Tugas Akhir

Tahapan studi literatur digunakan untuk mendalami permasalahan yang diangkat serta mendalami penyelesaian masalah tersebut. Studi literatur dilakukan melalui berbagai referensi seperti publikasi ilmiah, paper, buku cetak, dan wawancara narasumber. Melalui studi literatur ini pula diperoleh teknik analisis anomali paling optimal dengan permasalahan studi kasus.

3.2.1. Pengumpulan Data

Pada tahapan ini, dilakukan pengambilan data pelanggan PDAM Surya Sembada Kota Surabaya. Terdapat tiga macam data yang diambil yaitu data penggunaan air bulanan oleh pelanggan yang akan diolah dengan algoritma deteksi anomali, data jumlah jiwa (kapita) dalam masing-masing nomor rekening, dan data pelanggan yang telah melakukan pembayaran atas pelanggaran pemakaian air yang digunakan untuk tahap validasi. Data yang digunakan adalah data pelanggan rumah tangga PDAM di area pelayanan jaringan Surabaya pada Januari 2014 hingga Desember 2015. Selain itu, dilakukan wawancara kepada narasumber yang terkait sebagai informasi pendukung terlaksananya tugas akhir ini. Pengambilan data dilakukan di kantor PDAM Surya Sembada Kota Surabaya, Jawa Timur.

3.2.2. Praproses Data

Pada tahap ini akan dilakukan pengolahan data yang telah didapatkan. Praproses ini terdiri dari beberapa tahapan sebagai berikut:

a. Seleksi dan Penggabungan Atribut

Seleksi atribut dan penggabungan data dilakukan pada proses praproses data. Penggabungan data dilakukan pada data pemakaian air oleh pelanggan dalam rentang data tertentu, data jumlah jiwa dalam masing-masing nomor rekening, dan yang terakhir adalah data pelanggan yang telah membayar atas pelanggaran pemakaian air. Data yang digunakan hanya dari

kelompok pelanggan rumah tangga PDAM Surya Sembada Surabaya di area Surabaya. Atribut yang dipilih merupakan atribut yang hanya mendukung tugas akhir yang dilakukan.

b. Eliminasi nilai nol

Proses eliminasi data konsumsi selama 24 bulan yang mengandung nilai nol dilakukan pada tugas akhir ini. Proses eliminasi ini dikarenakan ketiga metode yang digunakan adalah menganalisis pola penurunan pada konsumsi air minum pelanggan PDAM Surya Sembada Surabaya. Sedangkan konsumsi nol pada kenyataannya adalah pelanggan tidak menggunakan atau tidak mengkonsumsi air minum sama sekali. Hal tersebut salah-satunya bisa disebabkan karena rumah kosong.

c. Normalisasi

Proses Normalisasi data konsumsi selama rentang data yang ditentukan menggunakan metode *min-max* dengan rentang data normalisasi antara 0 hingga 1. Proses normalisasi data konsumsi ini digunakan sebagai pra-proses algoritma *windowed analysis*.

3.2.3. Pembuatan Model metode Pearson's Correlation, Metode Abnormally Low Consumption, dan Algoritma Windowed Analysis.

Ketiga metode model yang dibangun dari 2 skenario. Skenario pertama adalah penentuan jumlah rentang data yang digunakan dan skenario kedua adalah penentuan nilai ambang dari ketiga metode. Metode *Pearson's Correlation* dilakukan penentuan jumlah rentang data dari kelipatan per kuartal dan penentuan nilai ambang korelasi yang teroptimal. Metode *windowed analysis* dilakukan penentuan jumlah rentang data dari penentuan bulan tiap 'jendela' dan penentuan nilai ambang *desc* dan *desv*. Metode *low abnormally* dilakukan penentuan jumlah rentang data dari kelipatan per kuartal dan penentuan nilai konsumsi minimal per kapita per hari per rentang data yang telah ditentukan. Untuk mengetahui model yang terbaik dari

berbagai eksperimen tersebut dilakukan pengolahan dengan data pelatihan.

3.2.4. Uji Coba

Pada sub-bab ini dijelaskan mengenai uji coba dari model dan juga aplikasi yang telah dibangun. Tujuan dari sub-bab uji coba adalah untuk mengetahui kemampuan dari model yang telah dibangun jika diterapkan pada *dataset* berbeda dan juga untuk mengetahui kemampuan perhitungan aplikasi dengan dibandingkan dengan perhitungan yang dilakukan secara manual (*microsoft excel*).

3.2.4.1. Uji Coba Model

Bagian uji coba model dilakukan untuk mengetahui apakah model yang teroptimal pada eksperimen yang dilakukan pada pembentukan model, juga optimal pada *dataset* yang lain. Proses uji coba memiliki langkah-langkah yang sama dengan proses pembuatan model.

3.2.4.2. Uji Coba Aplikasi

Pada tahap ini dilakukan uji coba aplikasi yang telah dibangun dengan menggunakan bahasa pemrograman Java. Aplikasi yang dibuat menggunakan metode-metode statistik yang telah dijelaskan pada tiap sub-bab 2.4.2 dalam studi literatur. Data yang digunakan pada uji coba aplikasi adalah *dataset* yang digunakan pada uji coba model dan parameter metode yang digunakan adalah rentang terbaik dari pembuatan model. Proses ini dilakukan untuk memastikan bahwa setiap metode statistik yang telah di-implementasikan pada aplikasi telah memberikan hasil yang sesuai dengan perhitungan yang dilakukan secara manual.

3.2.5. Uji coba Deteksi Pelanggan KTS

Uji coba ini dilakukan untuk mengetahui kemampuan metode dan algoritma dalam mendeteksi pola abnormal konsumsi air minum dalam kumpulan pelanggan air minum PDAM

Surabaya. Hasil dari uji coba ini adalah kelompok pelanggan dengan pola konsumsi air minum yang dianggap abnormal melalui masing-masing metode statistik yang kemudian dibandingkan dengan data pelanggan konsumsi tidak sah (KTS) yang dimiliki oleh PDAM.

3.2.6. Buku Tugas Akhir

Tahapan terakhir dari pembuatan tugas akhir ini adalah penyelesaian buku tugas akhir.

BAB IV

DESAIN DAN IMPLEMENTASI MODEL

Pada bab ini dijelaskan bagaimana rancangan dari tugas akhir yang dilakukan meliputi objek yang diteliti, bagaimana tugas akhir dilaksanakan, penentuan parameter metode, dan penentuan pelanggan anomali.

4.1. Pengumpulan Data

Data yang digunakan adalah data konsumsi air minum pelanggan selama 2 tahun terakhir pelanggan rumah tangga, data jumlah kapita dalam satu nomor pelanggan, dan data pembayaran pelanggan terhadap pelanggaran pemakaian air. Data yang didapat adalah data konsumsi pelanggan berkode tarif 3A untuk pembuatan model berjumlah 182.821 pelanggan dan 113.196 pelanggan 4b2 untuk uji *testing* yang merupakan pelanggan rumah tangga dari Januari 2014 setiap bulannya hingga Desember 2015, jumlah kapita setiap nomor pelanggan dari Januari 2014 hingga Desember 2015, dan yang terakhir adalah data pembayaran pelanggan terhadap temuan pelanggaran pemakaian air dari Januari 2014 hingga Desember 2015. Tabulasi dari data yang digunakan dapat dilihat pada Tabel 4.1.

Cuplikan data pada data konsumsi pelanggan rumah tangga PDAM Surya Sembada Kota Surabaya sebelum dilakukan pengolahan para proses, tersaji pada Tabel 4.2.

Data jumlah kapita yang dimiliki oleh PDAM Surya Sembada Surabaya pada 2 tahun terakhir (Januari 2014 - Desember 2015) untuk seluruh pelanggan adalah 185.770, sedangkan seluruh pelanggan pada akhir tahun 2015 adalah 536.985. Hal tersebut berarti jumlah pelanggan yang tidak tercatat memiliki jumlah kapita adalah 351.215 atau sebesar 65%. Cuplikan data pada data jumlah kapita setiap nomor pelanggan di PDAM Surya

Sembada Kota Surabaya sebelum dilakukan pengolahan tersaji pada Tabel 4.3.

Tabel 4.1 Data-data Yang Digunakan

Data	Rentang Data	Sumber
Data konsumsi pelanggan rumah tangga.	Januari 2014 – Desember 2015	PDAM Surya Sembada Kota Surabaya.
Data jumlah kapita pelanggan PDAM Surya Sembada Surabaya	Januari 2014 – Desember 2015	PDAM Surya Sembada Kota Surabaya.
Data pembayaran pelanggaran pemakaian air oleh pelanggan (Salinan berbentuk cetak).	Januari 2014 – Desember 2015	PDAM Surya Sembada Kota Surabaya.

A	B	C	D	E	F	G	H	I	J	K	L
THBL	NO_PLG	NAMA	JALAN	GANG	NOMOR	NOTAMB	DA	KD_TARIF	PAKAI	MCODE	LCODE
201504	476	Sengaja Dikosongkan	Sengaja Dikosongkan		28			3A		65	
201504	1010128	Sengaja Dikosongkan	Sengaja Dikosongkan	1	3			3A		29	
201504	1010157	Sengaja Dikosongkan	Sengaja Dikosongkan	2	4			3A		69	
201504	1010158	Sengaja Dikosongkan	Sengaja Dikosongkan	3	1 A			3A		16	
201504	1010159	Sengaja Dikosongkan	Sengaja Dikosongkan	3	3			3A		10	
201504	1010160	Sengaja Dikosongkan	Sengaja Dikosongkan	3	2			3A		4	
201504	1010177	Sengaja Dikosongkan	Sengaja Dikosongkan		41			3A		42	
201504	1010181	Sengaja Dikosongkan	Sengaja Dikosongkan		6			3A		36	
201504	1010191	Sengaja Dikosongkan	Sengaja Dikosongkan		9 AFT		JETIS	3A		2	
201504	1010210	Sengaja Dikosongkan	Sengaja Dikosongkan	1	9 A			3A		29	
201504	1010214	Sengaja Dikosongkan	Sengaja Dikosongkan	1	17 A		TAP	3A		18	
201504	1010248	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 B			3A		10	
201504	1010250	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 D			3A		20	L9R
201504	1010251	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 E			3A		62	
201504	1010252	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 G			3A		14	
201504	1010253	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 M			3A		73	
201504	1010254	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 P			3A		28	
201504	1010255	Sengaja Dikosongkan	Sengaja Dikosongkan	1	2 R			3A		21	
201504	1010257	Sengaja Dikosongkan	Sengaja Dikosongkan	1	4 C		PERT	3A		38	
201504	1010260	Sengaja Dikosongkan	Sengaja Dikosongkan	1	6 A1			3A		35	
201504	1010261	Sengaja Dikosongkan	Sengaja Dikosongkan	1	6 BNT			3A		87	
201504	1010262	Sengaja Dikosongkan	Sengaja Dikosongkan	1	6 C			3A		24	

Tabel 4.2 Cuplikan Data Konsumsi Pelanggan Rumah Tangga

Cuplikan data pada data pembayaran pelanggan terhadap pelanggaran konsumsi air minum tidak sah (KTS) PDAM Surya

Sembada Kota Surabaya sebelum dilakukan pengolahan tersaji pada Tabel 4.4.

A	B	C
NO_PLG	TGL_KONTR	JML_PENHUNI
Sengaja dikosongkan	12/10/2005	6
Sengaja dikosongkan	20/11/2008	1
Sengaja dikosongkan	25/11/2008	1
Sengaja dikosongkan	27/05/2009	3
Sengaja dikosongkan	23/11/2005	9
Sengaja dikosongkan	21/07/2011	3
Sengaja dikosongkan	31/08/2015	6
Sengaja dikosongkan	4/8/2014	3
Sengaja dikosongkan	28/07/2008	4
Sengaja dikosongkan	17/10/2008	6
Sengaja dikosongkan	17/12/2010	1
Sengaja dikosongkan	15/09/2011	2
Sengaja dikosongkan	14/11/2011	2
Sengaja dikosongkan	15/09/2011	4
Sengaja dikosongkan	20/07/2012	3
Sengaja dikosongkan	26/11/2014	3
Sengaja dikosongkan	20/07/2007	250

Tabel 4.3 Cuplikan Data Jumlah Kapita Pelanggan PDAM

REKAPITULASI DISEWA PERAKHIRAN GAS TERBUKA PELAKSANAAN PEMERTIAN KUBUS/SWEEPIING			
NO. KUBUS TANGKAI	NO. PERAKHIRAN ASLI (MS)	JUMLAH	
STASION PERBETIK BENDULHERI PERMAI GOSIOW / 48.2	02647	43	150,000
ELCROS TERBUK BENTUL / 51.1	05500	30	90,000
BOYAN KOCHE MADR WONGKONO TENGAH 100317 / 3A	00184	308	708,700
KULTARITE PUSKINITE SANGILAS MCTOYO 86 104223 / 33.1	00243	667	1.282,500
MS. BAMBANG POEJOH MANG INDAH / 48.2	00087	2501	3.752,800
MS. KUNTOYO MUR ANDAYANI 11/3-0 110051 / 20.2	01039	2178	3.010,100
MSHENTY UNKEMITJOYO SUNGUT PERMAI 10/49 1112076 / 46	00574	100	382,800
FORBEC SWARTTER SUNGUT ASRI TENGAH 1120417 / 48.2	00642	82	213,500
SUPREPRET SUNGUT SIBU 40/22 1122858 / 48.2	00725	79	172,800
REDEKATRE SANGAJI LEMUR 7/11 114081 / 48.2	02588	101	282,800
RENY KENDANGING SARI BLOK 1145862 / 48.2	02582	4.791	11.020,200
RENY KEDANG AYU BLOK 1C/22 115057 / 48.2	01300	114	262,200
OLCRODORH SUNGUT 43.0A SK/1 1151287 / 33.1	00649	1.302	2.744,900
RE-TERPONGRE ARHITIT KEDANG AYU BLOK 1A/4 1151909 / 48.2	01692	88	202,500
ADITYA DONDONG KESAWA KENDANGING SARI 88 1162821 / 48.2	01298	241	553,000
MS. H. SURYABESTI KENDANGING SARI 88 1150836 / 48.2	01833	75	172,800
	JUMLAH	12,510	28,994,300

Tabel 4.4 Cuplikan Data Pembayaran Pelanggan KTS

Pelanggan yang tercatat pada data pembayaran pada Tabel 4.4 adalah pelanggan-pelanggan yang secara sengaja maupun tidak

sangaja melakukan konsumsi air secara tidak sah, contoh-contoh tindakannya dapat dilihat pada Tabel 4.5.

59	-	a. Tempat meter dipindah sendiri	26
1	-	b. Menggunakan pompa Langsung/paralel	-
-	-	c. Meter hilang / Tidak ada	-
14	-	d. Meter rusak	13
-	-	e. Tutupan air digunakan	-
-	-	f. Pasangan ilegal / Non Pelanggan	-
9	-	g. Kaca meter pecah	3
80	-	h. Segel meter & segel kopling putus	70
-	-	i. Air dijual / memperluas jaringan	-
9	-	j. Stop kran rusak / meter diperbaiki sendiri	-
-	-	k. Memasang T sebelum meter	-
39	-	l. Tambahan / kehilangan air	75

Tabel 4.5 Contoh Tindakan yang Diharuskan Membayar Denda

Pada sub-bab ini akan dijelaskan mengenai langkah-langkah pra-proses data dari data awal hingga dapat digunakan yaitu meliputi penggabungan data dan normalisasi data.

4.1.1. Penggabungan Data



Gambar 4.1 Ilustrasi Penggabungan Data

Berbagai salinan data yang didapatkan dari PDAM Surya Sembada Surabaya kemudian digabungkan menjadi satu berkas dalam format *comma delimited* (.csv). Ketiga dokumen tersebut digabungkan dengan menggunakan bantuan alat *PHPMYAdmin*. Atribut pertama yang dibutuhkan untuk

dianalisa adalah pola konsumsi pelanggan rumah tangga PDAM Surya Sembada Surabaya. Atribut kedua yang digunakan adalah konsumsi air minum tiap pelanggan. Konsumsi air minum yang digunakan pada tugas akhir ini adalah data pada rentang data Januari 2014 hingga Desember 2015. Pemilihan waktu tersebut didasarkan terhadap relevansi terhadap kondisi terkini, kapabilitas dalam masing-masing metode dalam menganalisa pola konsumsi, dan mempertimbangkan kesediaan data pelanggan konsumsi tidak sah (KTS). Atribut ketiga adalah atribut yang menjelaskan masing-masing pelanggan melakukan pelanggaran terhadap pemakaian air atau tidak yang diambil dari data pembayaran pelanggaran pemakaian air.

Masing-masing atribut tersebut dikelompokkan berdasarkan masing-masing nomor pelanggan. Atribut yang digunakan adalah nomor, nomor pelanggan, konsumsi selama masing-masing 24 bulan, jumlah kapita, dan yang terakhir adalah atribut yang menunjukkan apakah masing-masing pelanggan pernah melakukan konsumsi tidak sah dalam jangka waktu 2 tahun terakhir.

Dari keseluruhan pelanggan rumah tangga pada tahun 2015, terdapat 16% dari keseluruhan pelanggan tersebut yang tercatat tidak memiliki konsumsi lengkap selama 2 tahun terakhir. Sehingga data konsumsi rumah tangga yang dapat digunakan untuk analisis menggunakan metode pada tugas akhir ini memiliki total 60.406 pelanggan, yang digunakan untuk membangun model dan proses uji coba. Cuplikan data dapat dilihat pada Tabel 4.6.

	2014												2015												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1 1014434	79	75	74	62	68	47	75	85	103	94	83	93	91	95	96	98	102	102	107	95	116	107	109	118	7
2 1017944	64	52	58	51	65	64	63	58	71	72	81	76	80	90	73	61	75	77	76	78	105	86	103	101	10
3 1018346	35	30	34	32	37	35	41	33	39	47	45	52	52	46	45	47	44	53	61	47	58	68	115	109	17
4 1030043	159	140	155	116	152	150	163	119	151	190	184	191	20	53	66	68	66	70	70	73	76	91	61	74	9
5 1032724	38	51	80	48	63	57	87	50	68	67	80	77	44	30	64	71	82	51	80	36	50	55	89	129	7
6 1032736	90	92	107	90	117	129	166	128	116	161	131	114	129	98	121	241	281	255	253	232	255	259	292	303	10
7 1052788	9	13	12	18	15	18	21	13	17	16	27	38	20	18	35	48	53	47	45	52	61	46	40	43	9
8 1072760	73	34	45	10	56	46	61	68	87	98	117	59	85	45	34	34	66	73	96	89	100	106	100	131	10
9 1072897	36	25	22	26	62	54	56	88	91	41	71	113	118	72	33	36	70	95	105	128	179	111	114	181	9
10 1100689	46	51	54	42	66	54	55	49	47	59	54	80	71	65	62	63	77	69	81	75	88	71	74	80	9

Tabel 4.6 Cuplikan data Konsumsi dan Kapita

4.1.2. Normalisasi

Normalisasi dilakukan dengan tujuan untuk melakukan standarisasi data serta untuk menyamakan skala atribut data kedalam sebuah range yang spesifik. Normalisasi secara khusus akan digunakan untuk metode analisa jendela (*Windowed Analysis*). Metode normalisasi yang digunakan adalah metode min-max yang mentransformasikan data menjadi rentang dari 0 hingga 1. Data konsumsi air minum (dalam m^3) sebelum normalisasi dapat dilihat pada Tabel 4.7 dan setelah normalisasi dapat dilihat pada Tabel 4.8.

Nopel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5020339	6	6	64	13	17	3	7	5	16	24	18	27	20	20	10	7	9	6	9	20	25	24	27	21
4038422	8	10	4	10	11	13	6	7	11	10	5	4	3	7	5	2	11	60	4	4	6	8	4	42
3102264	66	63	69	60	17	17	20	21	20	25	21	22	19	18	16	16	25	23	22	20	22	20	23	20
2203326	5	4	2	1	1	1	10	2	2	3	3	4	2	1	4	2	1	2	3	2	6	10	11	12
4124073	79	74	75	70	77	10	66	42	62	60	52	58	50	47	49	50	47	43	59	37	65	49	39	49
4124561	14	18	24	14	20	21	18	14	24	17	30	14	22	20	22	18	18	16	15	13	23	13	11	17
4037878	8	10	16	41	59	98	104	138	138	192	138	158	143	120	19	161	15	100	100	100	17	2	11	46

Tabel 4.7 Data Konsumsi sebelum Normalisasi

Nopel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5020339	0.24	0.24	1.00	0.33	0.38	0.20	0.25	0.23	0.37	0.48	0.40	0.51	0.42	0.42	0.29	0.25	0.28	0.24	0.28	0.42	0.49	0.48	0.51	0.44
4038422	0.28	0.31	0.23	0.31	0.32	0.35	0.26	0.27	0.32	0.31	0.24	0.23	0.21	0.27	0.24	0.20	0.32	1.00	0.23	0.23	0.26	0.28	0.23	0.75
3102264	0.95	0.91	1.00	0.86	0.22	0.22	0.26	0.28	0.26	0.34	0.28	0.29	0.25	0.23	0.20	0.20	0.34	0.31	0.29	0.26	0.29	0.26	0.31	0.26
2203326	0.49	0.42	0.27	0.20	0.20	0.85	0.27	0.27	0.35	0.35	0.42	0.27	0.20	0.42	0.27	0.20	0.27	0.35	0.27	0.56	0.85	0.93	1.00	1.00
4124073	1.00	0.94	0.95	0.90	0.98	0.20	0.85	0.57	0.80	0.78	0.69	0.76	0.66	0.63	0.65	0.66	0.63	0.58	0.77	0.51	0.84	0.65	0.54	0.65
4124561	0.33	0.49	0.75	0.33	0.58	0.62	0.49	0.33	0.75	0.45	1.00	0.33	0.66	0.58	0.66	0.49	0.49	0.41	0.37	0.28	0.71	0.28	0.20	0.45
4037878	0.23	0.23	0.26	0.36	0.44	0.60	0.63	0.77	0.77	1.00	0.77	0.86	0.79	0.70	0.27	0.87	0.25	0.61	0.61	0.61	0.26	0.20	0.24	0.39

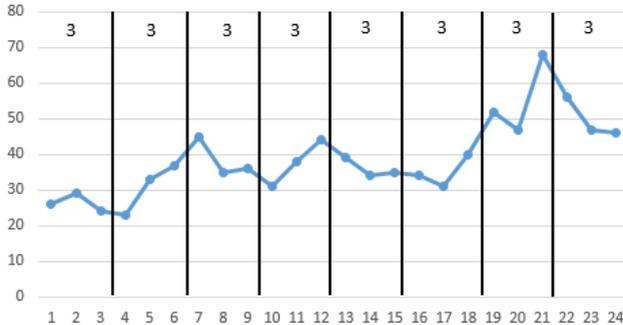
Tabel 4.8 Data Konsumsi setelah Normalisasi

Dalam tugas akhir ini, penentuan pelanggan yang melakukan konsumsi air minum tidak sah dilakukan dengan percobaan-

percobaan parameter. Parameter yang diteliti adalah parameter rentang data dan nilai ambang dari metode *pearson's*, *abnormally low consumption* dan metode jendela.

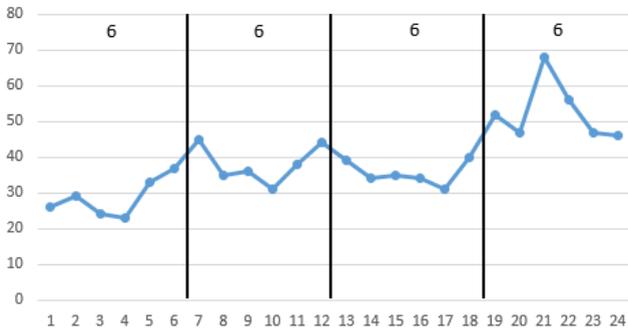
4.1.3. Skenario Rentang Data Data

- a. Panjang waktu analisis data. Panjang waktu yang digunakan akan berpengaruh tingkat akurasi dari masing-masing metode yang digunakan. Rentang panjang waktu dalam publikasi *Detection of Water Meter Under-Registration Using Statistical Algorithms* oleh Inigo Monedero adalah sepanjang 2 tahun pada metode *Pearson's Correlation*, 18 bulan pada metode *Windowed Analysis* dan 1 tahun terakhir untuk metode *Low Abnormal Consumption*. Pada tugas akhir kali ini akan dilakukan percobaan panjang waktu analisis untuk metode *Pearson's Correlation* dan *Low Abnormal Consumption* tiap:
 - 1) Kelipatan masing-masing kuartal dalam rentang data 2 tahun terakhir. Ilustrasi dapat dilihat pada Gambar 4.2.



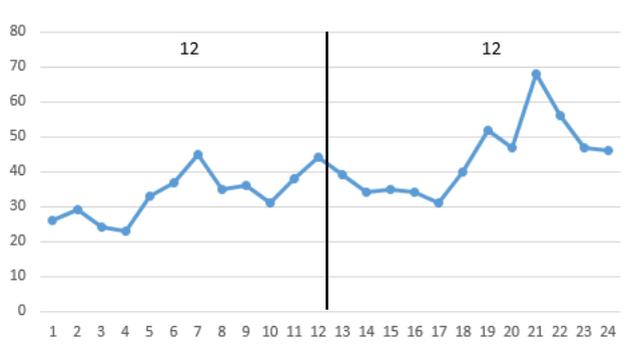
Gambar 4.2 Ilustrasi kelipatan 3 bulanan

- 2) Kelipatan masing-masing 6 bulan dalam rentang data 2 tahun terakhir. Ilustrasi dapat dilihat pada Gambar 4.3.



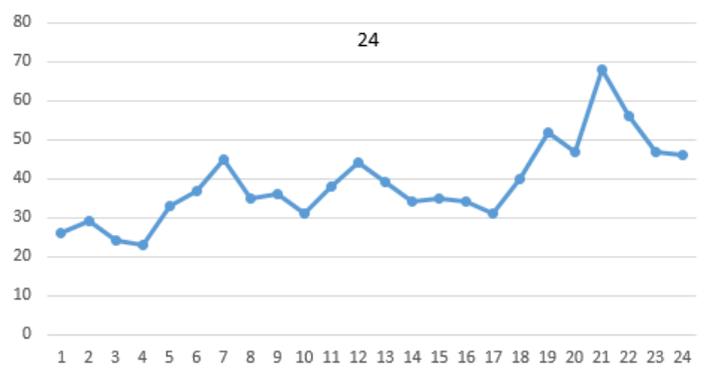
Gambar 4.3 Ilustrasi kelipatan 6 bulanan

- 3) 12 bulan dalam rentang data 2 tahun terakhir. Ilustrasi dapat dilihat pada Gambar 4.4.



Gambar 4.4 Ilustrasi kelipatan 12 bulanan

- 4) 2 tahun penuh dalam rentang data 2 tahun terakhir. Ilustrasi dapat dilihat pada Gambar 4.5.



Gambar 4.5 Ilustrasi kelipatan 24 bulanan

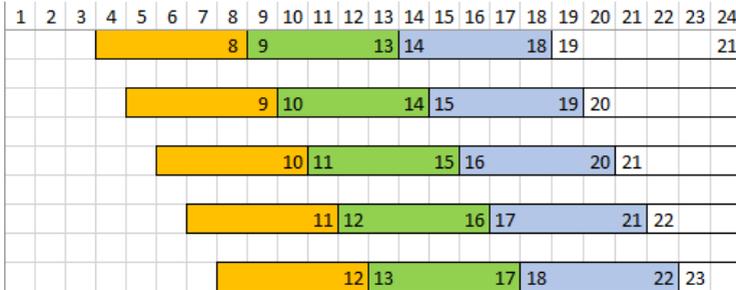
Sedangkan untuk metode *windowed analysis* tidak dapat menggunakan rentang data kelipatan dari tiap kuartal, maka dari itu percobaan rentang data pada metode ini berbeda dengan 2 metode yang lain. Percobaan panjang waktu analisis untuk metode *windowed analysis* adalah:

- 5) Jendela 3, jendela 4, dan jendela 5 berisi data konsumsi dalam 6 bulan. Dalam hal ini membutuhkan rentang data data konsumsi selama 24. Ilustrasi dapat dilihat pada Gambar 4.6.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
					6	7					12	13					18	19					24	
						7	8					13	14						19	20				
							8	9					14	15						21	22			
								9	10					15	16						22	23		
									10	11					16	17						23	24	

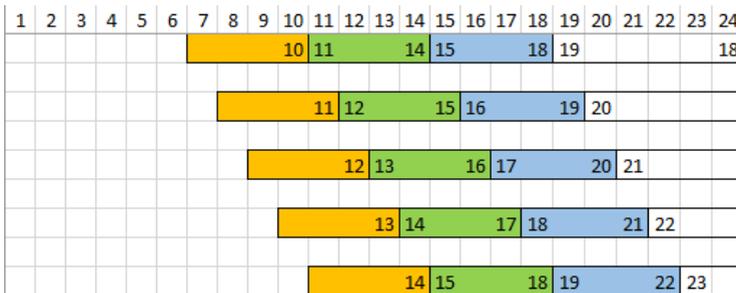
Gambar 4.6 Ilustrasi 6 bulan setiap jendela

- 6) Jendela 3, jendela 4, dan jendela 5 berisi data konsumsi dalam 5 bulan. Dalam hal ini membutuhkan rentang data data konsumsi selama 21 bulan. Ilustrasi dapat dilihat pada Gambar 4.7.



Gambar 4.7 Ilustrasi 5 bulan setiap jendela

- 7) Jendela 3, jendela 4, dan jendela 5 berisi data konsumsi dalam 4 bulan. Dalam hal ini membutuhkan rentang data data konsumsi selama 18 bulan. Ilustrasi dapat dilihat pada Gambar 4.8.



Gambar 4.8 Ilustrasi 4 bulan setiap jendela

- 8) Jendela 3, jendela 4, dan jendela 5 berisi data konsumsi dalam 3 bulan. Dalam hal ini membutuhkan rentang data data konsumsi selama 15 bulan. Ilustrasi dapat dilihat pada Gambar 4.9.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
									3	4	6	7	9	10									15	
									4	5	7	8	10	11										
									5	6	8	9	11	12										
									6	7	9	10	12	13										
									7	8	11	12	13	14										

Gambar 4.9 Ilustrasi 3 bulan setiap jendela

- 9) Jendela 3, jendela 4, dan jendela 5 berisi data konsumsi dalam 2 bulan. Dalam hal ini membutuhkan rentang data data konsumsi selama 12 bulan. Ilustrasi dapat dilihat pada Gambar 4.10.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
												2	3	4	5	6	7							12
												3	4	5	6	7	8							
												4	5	6	7	8	9							
												5	6	7	8	9	10							
												6	7	8	9	10	11							

Gambar 4.10 Ilustrasi 2 bulan setiap jendela

4.1.4. Skenario Penentuan Nilai Ambang Metode

- a. Nilai ambang digunakan dalam metode *Pearson's correlation* dan *windowed analysis*. Parameter ambang metode *Pearson's correlation* yang digunakan di dalam publikasi *Detection of Water Meter Under-Registration Using Statistical Algorithms* oleh Inigo Monedero adalah -0.6 atau -0.7. Sedangkan dalam tugas akhir ini akan dilakukan eksperimen untuk dapat mengetahui pengaruh

kinerja terhadap pengambilan masing-masing nilai ambang yang berada pada nilai korelasi negatif.

- b. Parameter yang digunakan dalam metode *windowed analysis* menggunakan ambang 2.0 dan 1.5 untuk *desv* dan *desc* secara berurutan atau 2.0 dan 2.0 untuk *desv* dan *desc* secara berurutan di dalam publikasi *Detection of Water Meter Under-Registration Using Statistical Algorithms* oleh Inigo Monedero. Pada tugas akhir ini akan dilakukan uji coba penentuan nilai ambang terbaik untuk metode *windowed analysis*. Eksperimen ini dilakukan antara nilai rekomendasi pada publikasi yang tersedia, sehingga eksperimen nilai *desv* akan dilakukan dari angka 1.5 hingga 2.5 dan eksperimen nilai *desc* akan dilakukan dari angka 1.5 hingga 2.5.

4.1.5. Penentuan Konsumsi Minimal per Kapita per Hari

Konsumsi minimal per kapita per hari adalah estimasi konsumsi minimal tiap orang tiap hari dalam satu nomor pelanggan/rumah tangga. Penentuan konsumsi minimal ini digunakan untuk kalkulasi metode *Low Abnormally Consumption*. Estimasi konsumsi minimal per kapita pada Kota Seville di Spanyol adalah sebanyak 114 liter -116 liter air minum dalam `satu hari [15] sedangkan konsumsi minimal tiap jiwa di perumahan pada tahun 2004 di wilayah Indonesia adalah 117 liter/jiwa/hari [29]. Sehingga dalam tugas akhir ini menggunakan angka 117 liter sebagai konsumsi minimal per kapita per hari yang digunakan untuk perhitungan metode.

4.2. Penentuan Pelanggan KTS

Penentuan pelanggan yang dianggap memiliki potensi melakukan konsumsi air minum secara tidak sah adalah jika terbukti minimal satu dari ketiga metode yang digunakan [15]. Dalam tugas akhir ini penentuan hasil akhir dari ketiga metode adalah dengan menggunakan konsep *OR* dari masing-masing metode. Yang berarti pelanggan yang teridentifikasi minimal 1 dari 3 metode akan dianggap pelanggan yang telah mengkonsumsi air minum secara tidak sah.

4.3. Implementasi Aplikasi

Pada bagian ini akan dijelaskan mengenai implementasi metode-metode statistik yang dibuat. Aplikasi dibuat dengan menggunakan bahasa pemrograman Java. Alur kerja dari sistem aplikasi terdiri dari implementasi beberapa pra-proses, implementasi memilih rentang data, implementasi metode *Pearson's Correlation*, *abnormally low consumption*, dan *windowed analysis*, dan implementasi hasil keluaran. Kode program selengkapnya dapat dilihat pada lampiran A.

4.3.1. Implementasi Normalisasi Data

Tahap normalisasi data ini digunakan sebagai pra-syarat penggunaan metode *windowed analysis*. Tahap normalisasi ini digunakan untuk merubah rentang data menjadi bernilai 0 hingga 1. Variabel konsumsi yang digunakan dalam tahap ini adalah variabel konsumsi pelanggan yang akan ditentukan sendiri oleh pengguna aplikasi. Cuplikan kode program implementasi normalisasi data dapat dilihat pada Segmen 4.1.

```
public void pickNormal(double[][] kons_new, int starts) {
    double norm_top = 1;
    double norm_bot = 0; //finding max
    for (int row = 0; row < lastfix; row++) {
        for (int col = 0; col < lastCol; col++) {
            if (kons_new[row][col+starts] > max[row]) {
                max[row] = kons_new[row][col+starts];
            } } //finding min;
    for (int row = 0; row < lastfix; row++) {
        for (int col = 0; col < lastCol; col++) {
            if (kons_new[row][col+starts] < min[row]) {
                min[row] = kons_new[row][col+starts];
            } } } //normalization
    for (int row = 0; row < lastfix; row++) {
        for (int col = 0; col < lastCol; col++) {
            kons_norm[row][col] = (norm_top-norm_bot)/(max[row]-
            min[row])*(kons_new[row][col+starts]-max[row])+norm_top;
        } this.kons_norm = kons_norm; }
}
```

Segmen 4.1 Implementasi Normalisasi Data

4.3.2. Implementasi Memilih Rentang Data

Pada tahap ini memilih rentang data digunakan dalam penentuan model yang digunakan pada masing-masing metode. Penentuan rentang data ini disesuaikan dengan syarat-syarat dasar yang harus dipenuhi oleh masing-masing metode. Maka dari itu rentang data metode *Pearson's Correlation* dan *low abnormally consumption* disesuaikan dengan metode *windowed analysis* yang memiliki rentang data yang telah ditentukan yaitu 12 bulan, 15 bulan, 18 bulan, 21 bulan hingga 24 bulan. Cuplikan dari implementasi memilih rentang data dapat dilihat pada Segmen 4.2.

4.3.3. Implementasi Metode-metode Statistik

Pada tahap ini adalah implementasi masing-masing metode kedalam aplikasi yang dibuat, yaitu metode *Pearson's Correlation*, metode *windowed analysis*, dan metode *low abnormally consumption*.

```
public double[][] splitRange(double[][] newArray, int
lastRow, int rangeStarts) {
//newCol adalah panjang data yang digunakan.
    int newCol = 24-rangeStarts;
    int starts = rangeStarts;
    double [][] rArray = new double[lastRow][newCol];
        for (int row = 0; row < lastRow; row++) {
            for (int col = 0; col < newCol; col++) {
rArray[row][col] = newArray[row][col+starts]; }
        } this.lastCol = newCol;
    return rArray; }
```

Segmen 4.2 Implementasi Memilih Rentang Data

4.3.3.1. Implementasi metode Pearson's Correlation

Pada tahap ini adalah implementasi dari metode Pearson's correlation. Metode ini mengambil konsumsi sesuai dengan inputan rentang data yang dipilih oleh pengguna. Korelasi ini adalah perbandingan antara konsumsi dengan banyaknya jumlah bulan yang dipilih. Didalam implementasi ini nilai ambang dapat dilihat secara manual oleh pengguna. Jika, terdapat salah satu pelanggan dengan nomor pelanggan tertentu memiliki konsumsi yang berkorelasi negatif dan berada dibawah nilai ambang yang ditentukan, maka akan aplikasi akan memberikan keluaran "Konsumsi turun progresif". Sebaliknya, jika pelanggan memiliki konsumsi diatas nilai ambang maka tidak akan tercatat dalam aplikasi yang dibuat. Cuplikan implementasi metode dapat dilihat pada Segmen 4.3.

```

public double correlation(double upper, double lower)
{
    SimpleRegression regression = new SimpleRegression();
    if (monthCorrel.length !=
    kons_range[0].length) { throw new
    DimensionMismatchException(monthCorrel.length,
    kons_range[lastfix-1].length);
    } else if (monthCorrel.length < 2) {
        throw new
        MathIllegalArgumentException(LocalizedFormats.INSUFFIC
        IENT_DIMENSION, monthCorrel.length, 2);
    } else {
        for(int row=0; row<lastfix;row++) {
            for(int col=0; col<lastCol; col++) {
                regression.addData(monthCorrel[col],
                kons_range[row][col]); }
            if(regression.getR() <= upper &&
            regression.getR() >= lower){ getR[row] = "Konsumsi
            turun progresif"; } else getR[row] = "";
            regression.clear(); }
        this.getR = getR;
        return regression.getR(); } }

```

Segmen 4.3 Implementasi metode *Pearson's Correlation*

4.3.3.2. Implementasi metode Abnormally Low Consumption

Pada tahap ini adalah implementasi dari metode konsumsi rendah abnormal (*low abnormally consumption*). Metode ini membandingkan antara jumlah dari konsumsi dengan konsumsi minimum rata-rata per kapita dalam rentang data yang telah ditentukan sendiri oleh pengguna. Jika konsumsi aktual pelanggan dalam rentang data tersebut lebih rendah dibandingkan dengan konsumsi minimum rata-rata maka pelanggan tersebut akan dianggap memiliki konsumsi rendah abnormal, lalu akan tercatat dalam aplikasi sebagai pelanggan “Konsumsi Rendah Abnormal”. Jika sebaliknya, maka pelanggan tersebut tidak akan tercatat dan tidak akan ditampilkan dalam aplikasi. Cuplikan dari metode dapat dilihat pada Segmen 4.4.

```
public void lowabn(int literConsump) {
//Menambahkan Jumlah Meter Kubik Konsumsi Selama Range yang
ditentukan
for(int row = 0; row<kapita.length;row++) {
DoubleStream jKons = Arrays.stream(kons_range[row]);
jStream[row] = jKons.sum();

//kalkulasi konsumsi minimum per kapita
for(int row = 0; row<kapita.length;row++) {
calc[row] = (kapita[row]*literConsump*30*lastCol)/1000; }
for(int row = 0; row<kapita.length;row++) {
if (jStream[row]<calc[row]) {
lowabn[row] = "Konsumsi Rendah Abnormal"; } else lowabn[row]
= "";} this.getA = lowabn; }
```

Segmen 4.4 Implementasi metode abnormally low consumption

4.3.3.3. Implementasi algoritma *Windowed Analysis*

Pada tahap ini adalah implementasi metode *windowed analysis*. Metode ini akan mendeteksi pelanggan yang memiliki konsumsi yang relatif stabil kemudian diikuti penurunan konsumsi yang drastis kemudian rendah stabil hingga akhir dari jangka waktu yang digunakan. Dalam metode ini membutuhkan parameter *desv* sebagai parameter kestabilan konsumsi dan *desc* adalah parameter penurunan drastis. Parameter tersebut dimasukkan sesuai dengan kehendak dari pengguna aplikasi. Cuplikan dari implementasi *windowed analysis* dapat dilihat pada Segmen 4.5.

```
public void windowed(double desv, double desc){
    //Membuat Jendela
    for(int b=0;b<kons_range.length*5;b++){
        sixth[index] =
java.util.Arrays.copyOfRange(newArray,x-a,x);
        fifth[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c, x-a);
        fourth[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c-c, x-a-c);
        third[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c-c-c, x-a-c-
c); index++; a++; if(a==7){ a=2;
        }
    }
    for (int itt = 0; itt <
5*kons_norm.length; itt++) {
        if( ((temp[itt] || temp_1[itt] ||temp_2[itt])
&& temp_3[itt]) == true){
            result[itt] = true;
        } else { result[itt] = false;
        }
    }
    //iteration for OR
    for(int i=0;i<kons_range.length;i++) {
        if((result[y-1] || result[y-2] || result[y-3]
|| result[y-4] || result[y-5])==true) {
r_wind[i] = "Konsumsi Turun Drastis";}
else{r_wind[i] = "";} } this.getW = r_wind; }
```

Segmen 4.5 Implementasi Metode *Windowed Analysis*

4.3.4. Implementasi Masukan Parameter Metode

Aplikasi yang dibuat memiliki tampilan antar-muka aplikasi yang dapat memudahkan pengguna untuk memasukkan parameter pada setiap metode.

Gambar 4.11 Tampilan Antar Muka Aplikasi

Pada Gambar 4.11 merupakan form masukan dari aplikasi deteksi konsumsi pelanggan tidak sah. Pada aplikasi ini masing-masing input inputan seperti nilai ambang metode *Pearson's Correlation*, nilai ambang metode *windowed analysis*

, dan konsumsi minimal rata-rata pada metode konsumsi rendah abnormal mempunyai form masukkan *textfield* karena parameter-parameter ini bersifat kontinu dan input yang diisikan berupa nilai angka. Sedangkan untuk masukkan panjang rentang data (*numbers month*) memiliki nilai diskrit dan mempunyai nilai yang sesuai dengan nama pada setiap *combo box*-nya.


```
public void showing() {
    for (int row = 0; row < lastfix; row++) {
        if (getR[row].equals("Konsumsi turun progresif") ||
            getW[row].equals("Konsumsi Turun Drastis") ||
            getA[row].equals("Konsumsi Rendah Abnormal")) { else if
            (getR[row].equals("Konsumsi turun progresif") &&
            getA[row].equals("Konsumsi Rendah Abnormal")) {
                jTableResult.setValueAt(getR[row]+", "+getA[row]+".", rows, 2); }
            else if (getW[row].equals("Konsumsi Turun Drastis") &&
            getA[row].equals("Konsumsi Rendah Abnormal")) {
                jTableResult.setValueAt(getW[row]+", "+getA[row]+".", rows, 2);
            } else { jTableResult.setValueAt(getR[row]+", "+getW[row]+",
            "+getA[row]+".", rows, 2); }
            rows++; } } this.lastshowing = rows; }
```

Segmen 4.6 Implementasi Keluaran Aplikasi

BAB V

UJI COBA DAN ANALISIS HASIL

Pada bab ini akan ditulis mengenai hasil dari perancangan dan implementasi dari setiap percobaan meliputi hasil eksperimen model, hasil penentuan parameter, uji coba aplikasi, dan kesimpulan hasil eksperimen.

5.1. Hasil Pembentukan Parameter Model

Pembentukan model dilakukan dengan menentukan jumlah rentang data yang digunakan dan nilai ambang pada masing-masing metode. Setelah dilakukan percobaan terhadap beberapa model, maka akan terlihat model yang memiliki parameter jangka waktu yang paling optimal berdasarkan nilai *recall* (hasil temuan metode dibandingkan dengan data aktual) dan nilai ambang berdasarkan penilaian kinerja menggunakan akurasi, *recall* dan presisi. Penggunaan nilai *recall* untuk mengukur tingkat keoptimalan parameter jangka waktu didasarkan terhadap perbedaan jumlah PKTS (Pelanggan Konsumsi Tidak Sah) dalam masing-masing rentang data yang digunakan dalam percobaan.

5.1.1. Parameter Rentang Data

Parameter rentang data adalah rentang data penggunaan data yang diubah untuk mengetahui dampak terhadap kinerja dari masing-masing metode. Rentang data yang digunakan adalah setiap kelipatan kuartal atau 3 bulanan, yaitu: 3 bulan, 6 bulan, 12 bulan, dan 24 untuk metode *Pearson's Correlation* dan *Abnormally Low Consumption*. Sedangkan untuk metode *Windowed Analysis* akan diubah komposisi jumlah bulan dalam masing-masing jendela ke-3, ke-4, dan ke-5 yang digunakan.

5.1.1.1. Hasil Model Metode Pearson's Correlation

Pada sub-bab diperlihatkan hasil pencarian model rentang data terbaik untuk metode *Pearson's correlation*. Percobaan awal

dibawah ini dilakukan dengan nilai ambang (*threshold*) Pearson's correlation sebesar -0.6 melalui rekomendasi publikasi [15].

5.1.1.1.1. Metode Pearson's Correlation 3 bulanan.

Pada tabel Tabel 5.1 memperlihatkan penerapan dari metode Pearson's correlation pada setiap kelipatan percobaan 3 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 23 pelanggan yang melakukan konsumsi tidak sah dari total 71 pelanggan yang terdata atau memiliki *recall* 0.323.

Tabel 5.1 Implementasi Pearson's Correlation 3 Bulanan.

Pearson's Correlation – 3 bulanan		
Rentang Data		
	Hasil benar	PKTS
1 – 3	1	4
4 – 6	6	17
7 – 9	3	5
10 – 12	0	4
13 – 15	3	3
16 – 18	4	12
18 – 21	3	11
22 – 24	3	15
Total	23	71

5.1.1.1.2. Metode Pearson's Correlation 6 bulanan.

Pada Tabel 5.2 memperlihatkan penerapan dari metode Pearson's correlation pada setiap kelipatan percobaan 6 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 21 pelanggan yang melakukan konsumsi tidak sah dari 71 pelanggan yang terdata atau memiliki *recall* 0.295.

Tabel 5.2 Implementasi Pearson's Correlation 6 Bulanan

Pearson's Correlation – 6 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 6	7
7 – 12	3	9
13 – 18	3	15
19 – 24	8	26
Total	21	71

5.1.1.1.3. Metode Pearson's Correlation 12 bulanan.

Pada Tabel 5.3 memperlihatkan penerapan dari metode Pearson's correlation pada setiap kelipatan percobaan 12 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 21 pelanggan yang melakukan konsumsi tidak sah dari 71 pelanggan yang terdata atau memiliki *recall* 0.295.

Tabel 5.3 Implementasi Pearson's Correlation 12 Bulanan

Pearson's Correlation – 12 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 12	11
13 – 24	9	40
Total	21	71

5.1.1.1.4. Metode Pearson's Correlation 24 bulanan.

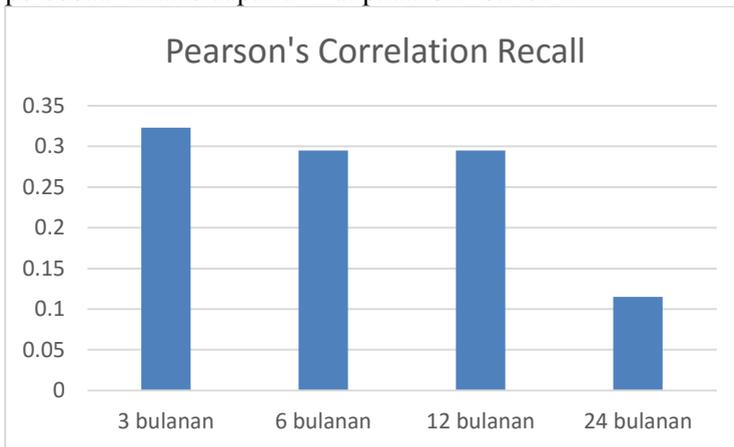
Pada Tabel 5.4 memperlihatkan penerapan dari metode Pearson's correlation pada setiap kelipatan percobaan 24 bulanan dalam 2 tahun. Terlihat dalam rentang 2 tahun tersebut berhasil mendeteksi 8 pelanggan yang melakukan konsumsi tidak sah dari 69 pelanggan yang terdata atau memiliki *recall* 0.115.

Tabel 5.4 Implementasi Pearson's Correlation 24 Bulanan

Pearson's Correlation – 24 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 24	8
Total	8	69

5.1.1.1.5. Hasil Optimal Pearson's Correlation

Prediksi benar terbaik didapatkan pada rentang data 3 bulanan yaitu sebanyak 23 pelanggan teridentifikasi dan prediksi rendah terendah didapatkan pada rentang data 2 tahunan yaitu sebanyak 8 pelanggan teridentifikasi. Visualisasi dari ke-empat percobaan diatas dapat dilihat pada Gambar 5.1.



Gambar 5.1 Grafik Percobaan Metode Pearson's Correlation

5.1.1.2. Hasil Model Metode Low Abnormal Consumption

Pada sub-bab diperlihatkan hasil pencarian model rentang data terbaik untuk metode konsumsi rendah abnormal (*low abnormal consumption*). Percobaan menggunakan konsumsi

rata-rata setiap orang dalam setiap harinya adalah sebanyak 117 liter [29].

5.1.1.2.1. Percobaan 3 bulanan

Pada Tabel 5.5 memperlihatkan penerapan dari metode konsumsi rendah abnormal (*low abnormally consumption*) pada setiap kelipatan percobaan 3 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 5 pelanggan yang melakukan konsumsi tidak sah dari total 71 pelanggan yang terdata atau memiliki *recall* 0.07.

Tabel 5.5 Implementasi konsumsi rendah abnormal 3 bulanan

Low Abnormal – 3 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 3	1
4 – 6	0	17
7 – 9	0	5
10 – 12	0	4
13 – 15	1	3
16 – 18	0	12
18 – 21	1	11
22 – 24	2	15
Total	5	71

5.1.1.2.2. Percobaan 6 Bulan

Pada Tabel 5.6 memperlihatkan penerapan dari metode konsumsi rendah abnormal (*low abnormally consumption*) pada setiap kelipatan percobaan 6 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 5 pelanggan yang melakukan konsumsi tidak sah dari total 71 pelanggan yang terdata atau memiliki *recall* 0.07.

Tabel 5.6 Implementasi konsumsi rendah abnormal 6 bulanan

Low Abnormal – 6 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 6	1
7 – 12	1	9
13 – 18	1	15
19 – 24	2	26
Total	5	71

5.1.1.2.3. Percobaan 12 bulan

Pada Tabel 5.7 memperlihatkan penerapan dari metode konsumsi rendah abnormal (*low abnormally consumption*) pada setiap kelipatan percobaan 3 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil mendeteksi 4 pelanggan yang melakukan konsumsi tidak sah dari total 71 pelanggan yang terdata atau memiliki *recall* 0.05.

Tabel 5.7 Implementasi konsumsi rendah abnormal 12 bulanan

Low Abnormal – 12 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 12	2
13 – 24	2	40
Total	4	71

5.1.1.2.4. Percobaan 24 bulanan

Pada Tabel 5.8 memperlihatkan penerapan dari metode konsumsi rendah abnormal (*low abnormally consumption*) pada setiap kelipatan percobaan 3 bulanan selama 2 tahun. Terlihat bahwa selama 2 tahun atau 24 bulan metode ini berhasil

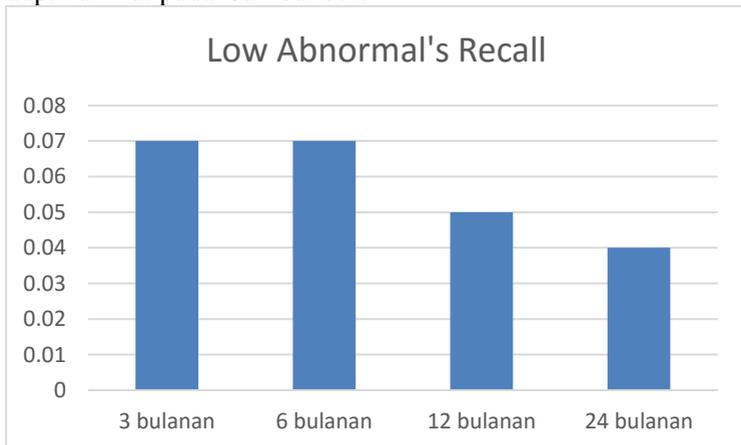
mendeteksi 3 pelanggan yang melakukan konsumsi tidak sah dari total 69 pelanggan yang terdata atau memiliki *recall* 0.04.

Tabel 5.8 Implementasi konsumsi rendah abnormal 24 bulanan

Low Abnormal – 24 bulanan		
Rentang Data	Hasil benar	PKTS
	1 – 24	3
Total	3	69

5.1.1.2.5. Hasil Optimal Metode Abnormally Low Consumption

Prediksi benar terbaik didapatkan pada rentang data 3 bulanan dan 6 bulanan yaitu sebanyak 5 pelanggan teridentifikasi dan prediksi rendah terendah didapatkan pada rentang data 2 tahunan yaitu sebanyak 3 pelanggan teridentifikasi. Visualisasi dari keempat percobaan metode *abnormally low consumption* dapat dilihat pada Gambar 5.2.



Gambar 5.2 Grafik Percobaan Metode Abnormally Low Consumption

5.1.1.3. Hasil Model Metode *Windowed Analysis*

Pada sub-bab diperlihatkan hasil pencarian model rentang data terbaik untuk metode *windowed analysis*. Percobaan dibawah menggunakan parameter *desv* sebesar 1.9 dan *desc* sebesar 1.9 melalui rekomendasi dari sub-bab 5.1.2.2.

5.1.1.3.1. Percobaan 6 bulan dalam satu jendela.

Pada Tabel 5.9 memperlihatkan hasil dari penerapan metode analisa jendela (*windowed analysis*) dengan hasil benar sebanyak 21 pelanggan berhasil teridentifikasi dari 69 pelanggan dalam rentang data 2 tahun atau memiliki *recall* 0.144. Metode ini menggunakan rentang data 6 bulanan pada jendela ke-3, ke-4, dan ke-5 sehingga membutuhkan rentang data keseluruhan sebanyak 24 bulan.

Tabel 5.9 Implementasi *windowed analysis* 6 bulanan

Windowed Analysis			
Jumlah bulan			
	Hasil benar	PKTS	Rentang data
6 bulan dalam satu jendela	10	69	24 bulan

5.1.1.3.2. Percobaan 5 bulan dalam satu jendela

Pada Tabel 5.10 memperlihatkan hasil dari penerapan metode analisa jendela (*windowed analysis*) dengan hasil benar sebanyak 15 pelanggan berhasil teridentifikasi dari 54 pelanggan dalam rentang data 21 bulan atau memiliki *recall* 0.203. Metode ini menggunakan rentang data 5 bulanan pada jendela ke-3, ke-4, dan ke-5 sehingga membutuhkan rentang data keseluruhan sebanyak 21 bulan.

Tabel 5.10 Implementasi *windowed analysis* 5 bulanan

Windowed Analysis			
Jumlah bulan			
	Hasil benar	PKTS	Rentang data
5 bulan dalam satu jendela	11	54	21 bulan

5.1.1.3.3. Percobaan 4 bulan dalam satu jendela

Pada Tabel 5.11 memperlihatkan hasil dari penerapan metode analisa jendela (*windowed analysis*) dengan hasil benar sebanyak 13 pelanggan berhasil teridentifikasi dari 44 pelanggan dalam rentang data 18 bulan atau memiliki *recall* 0.181. Metode ini menggunakan rentang data 4 bulanan pada jendela ke-3, ke-4, dan ke-5 sehingga membutuhkan rentang data keseluruhan sebanyak 18 bulan.

Tabel 5.11 Implementasi *windowed analysis* 4 bulanan

Windowed Analysis			
Jumlah bulan			
	Hasil benar	PKTS	Rentang data
4 bulan dalam satu jendela	8	44	18 bulan

5.1.1.3.4. Percobaan 3 bulan dalam satu jendela

Pada Tabel 5.12 memperlihatkan hasil dari penerapan metode analisa jendela (*windowed analysis*) dengan hasil benar sebanyak 15 pelanggan berhasil teridentifikasi dari 32 pelanggan dalam rentang data 15 bulan atau memiliki *recall* 0.34. Metode ini menggunakan rentang data 3 bulanan pada

jendela ke-3, ke-4, dan ke-5 sehingga membutuhkan rentang data keseluruhan sebanyak 15 bulan.

Tabel 5.12 Implementasi *windowed analysis* 3 bulanan

Windowed Analysis			
Jumlah bulan			
	Hasil benar	PKTS	Rentang waktu
3 bulan dalam satu jendela	11	32	15 bulan

5.1.1.3.5. Percobaan 2 bulan dalam satu jendela

Pada Tabel 5.13 memperlihatkan hasil dari penerapan metode analisa jendela (*windowed analysis*) dengan hasil benar sebanyak 31 pelanggan berhasil teridentifikasi dari 69 pelanggan dalam rentang data 2 tahun atau memiliki *recall* 0.374.

Tabel 5.13 Implementasi *windowed analysis* 2 bulanan

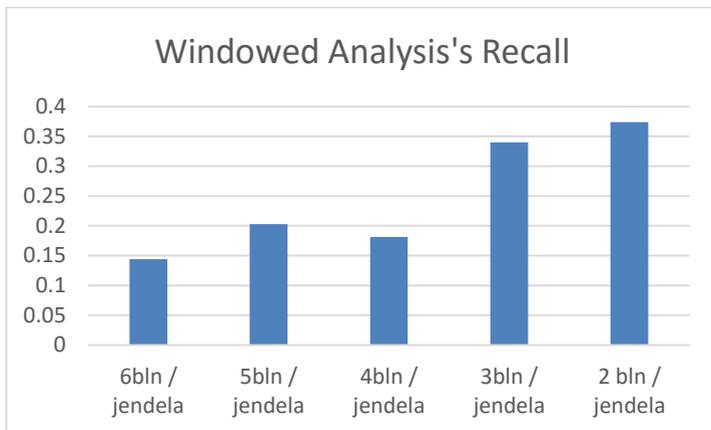
Windowed Analysis			
Jumlah bulan			
	Hasil benar	PKTS	Rentang waktu
2 bulan dalam satu jendela	12	40	12 bulan
2 bulan dalam satu jendela	13	29	12 bulan

Metode ini menggunakan rentang data 2 bulanan pada jendela ke-3, ke-4, dan ke-5 sehingga membutuhkan rentang data

keseluruhan sebanyak 12 bulan. Dalam penerapannya percobaan ini dilakukan 2 kali sehingga menganalisa konsumsi selama 2 tahun.

5.1.1.3.6. Hasil Optimal Metode *Windowed Analysis*

Prediksi benar terbaik didapatkan pada rentang data 12 bulanan atau 2 bulan dalam satu jendela yaitu sebanyak 25 pelanggan teridentifikasi dari 69 pelanggan KTS dalam 2 tahun atau memiliki *recall* 0.374. Visualisasi dari ke-empat percobaan metode *abnormally low consumption* dapat dilihat pada Gambar 5.3.



Gambar 5.3 Grafik Percobaan Metode *Windowed Analysis*

5.1.2. Hasil Percobaan Parameter Nilai Ambang

Parameter nilai ambang dimiliki oleh metode *Pearson's correlation* dan *windowed analysis*. Pada percobaan kali ini metode *Pearson's correlation* menggunakan model rentang data 12 bulan, dikarenakan mempertimbangkan jumlah PKTS yang mencukupi. Sedangkan pada percobaan metode analisis jendela menggunakan rentang data yang sama yaitu 12 bulan atau 2 bulan dalam satu jendela. Hasil percobaan ini diharapkan dapat memberikan gambaran mengenai pengaruh dari

perubahan nilai parameter ambang pada kedua metode terhadap kinerja dari sistem atau metode yang digunakan.

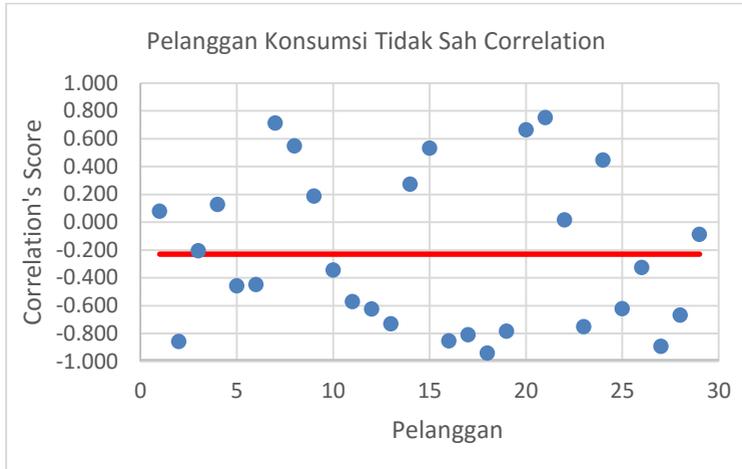
5.1.2.1. Percobaan Nilai Ambang Pearson's Correlation

Percobaan nilai ambang *Pearson's Correlation* ini menggunakan 58 data pelanggan, dengan komposisi 29 pelanggan yang telah terbukti KTS dan 29 pelanggan lainnya pelanggan yang belum atau tidak terbukti KTS. 29 pelanggan non-KTS dipilih secara acak dari kumpulan data. Nilai ambang pada metode *Pearson's correlation* bertujuan untuk mengklasifikasikan pelanggan yang memiliki pola konsumsi yang tiap bulannya mengalami penurunan. Semakin rendah nilai ambang yang ditentukan (semakin mendekati -1) maka pola konsumsi pelanggan yang terklasifikasi akan semakin menurun dengan curam. Sedangkan, semakin mendekati 0 nilai ambang *Pearson's correlation* maka pola konsumsi akan semakin menurun dengan landai. Sedangkan tidak semua pelanggan yang telah terbukti mengkonsumsi air secara tidak sah memiliki nilai korelasi yang negatif. Persebaran nilai korelasi pelanggan KTS dapat dilihat pada

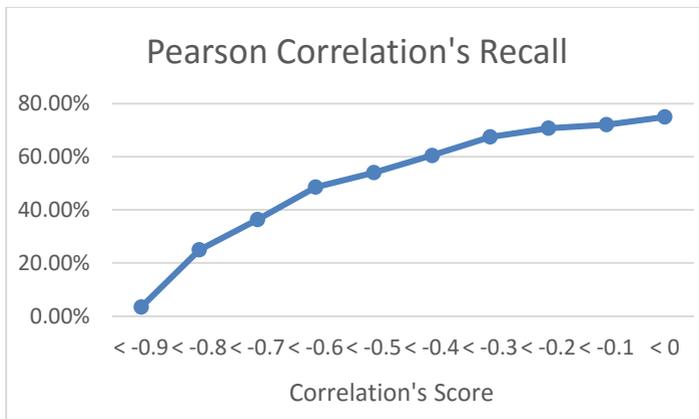
Gambar 5.4. Garis merah pada

Gambar 5.4 adalah nilai rata-rata dari nilai korelasi pelanggan KTS yang berada pada korelasi negatif, hal ini dimaksudkan pelanggan KTS pada data yang digunakan pada tugas akhir ini memiliki rata-rata korelasi berada pada nilai negatif.

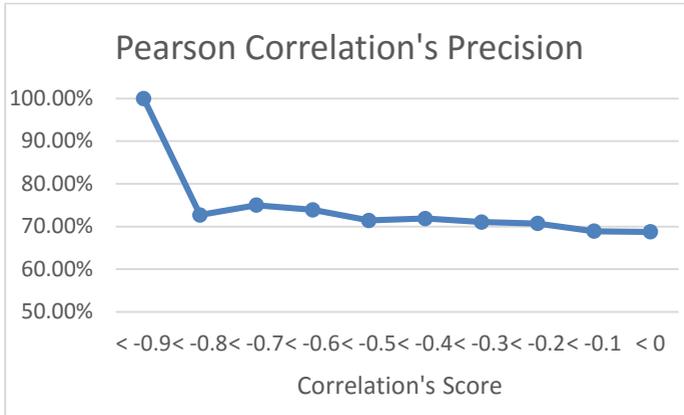
Percobaan perubahan nilai ambang dilakukan dari angka kurang dari -0.9 hingga kurang dari 0. Angka tersebut diambil dikarenakan hanya mengambil pola konsumsi menurun dalam pola konsumsi pelanggan. Dari percobaan didapatkan, semakin nilai ambang mendekati 0, maka nilai *recall* akan semakin baik dan *precision* akan semakin tidak baik, Begitu juga sebaliknya, semakin nilai ambang mendekati -1 maka *precision* akan semakin baik dan *recall* akan semakin tidak baik. Visualisasi dari nilai percobaan *recall* dan *precision* dapat dilihat pada Gambar 5.5 dan Gambar 5.6 secara berurutan.



Gambar 5.4 Nilai Korelasi Pelanggan Konsumsi Tidak Sah



Gambar 5.5 Nilai Precision dari Percobaan Nilai Ambang



Gambar 5.6 Nilai *Recall* dari Percobaan Nilai Ambang

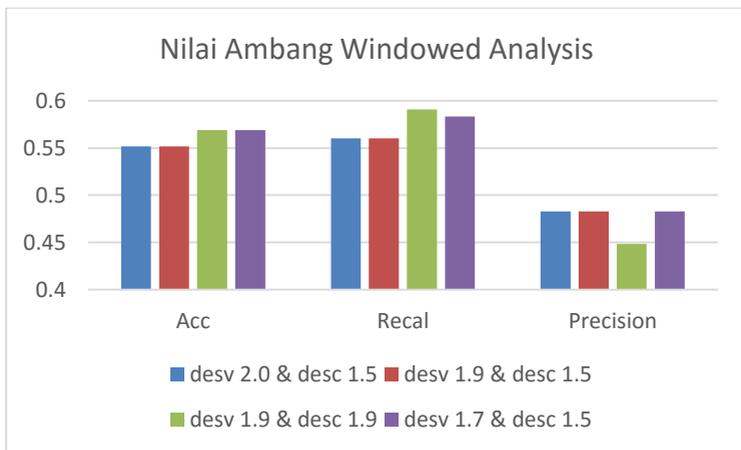
5.1.2.2. Percobaan Nilai Parameter *Windowed Analysis*

Percobaan nilai parameter *windowed analysis* ini menggunakan 58 data pelanggan, dengan komposisi 29 pelanggan yang telah terbukti KTS dan 29 pelanggan lainnya pelanggan yang belum atau tidak terbukti KTS. 29 pelanggan non-KTS dipilih secara acak dari kumpulan data. Percobaan nilai parameter dari metode *windowed analysis* ini dimulai pada nilai masing-masing 1.5 hingga 2.5. Nilai ini didasari oleh rekomendasi publikasi rujukan yang memberikan nilai 2.0 dan 1.5 untuk *desv* dan *desc* secara berurutan dan juga 2.0 dan 2.0 untuk *desv* dan *desc*.

Dalam percobaan yang telah didapatkan nilai kinerja *recall* dan *precision* terbaik dimiliki oleh nilai ambang 1.7 untuk *desv* dan 1.5 untuk *desc* dari ketiga kinerja metode yang digunakan. Terbaik kedua adalah nilai ambang 2.0 untuk *desv* dan 1.5 untuk *desc* dan juga 1.9 untuk *desv* dan 1.5 untuk *desc* yang memiliki kinerja akurasi sebesar 55%, 0.56 untuk *recall* dan 0.47 untuk *presisi*. Sedangkan untuk kombinasi nilai ambang 1.9 untuk *desc* dan 1.9 untuk *desv* memiliki nilai yang tinggi pada akurasi dan *recall*, namun lebih rendah pada nilai *presisi* yang diperoleh. Beberapa kombinasi *desv* dan *desc* teroptimal dapat

dilihat pada visualisasi Gambar 5.7. Detail dari hasil percobaan nilai parameter *windowed analysis* dapat dilihat pada lampiran B.

Dari percobaan nilai parameter metode *windowed analysis* yang telah dilakukan. Semakin besar nilai *desc* semakin sedikit pola yang terklasifikasi, hal ini dikarenakan *desc* merupakan sebuah nilai parameter untuk menentukan drastis atau landainya penurunan antara jendela dengan jendela lain. Semakin besar nilai *desc* maka pola penurunan dalam konsumsi akan semakin curam dan data yang terklasifikasi akan semakin sedikit. Sedangkan semakin besar nilai *desv*, kemampuan metode dalam mendeteksi anomali akan semakin berkurang. Dikarenakan *desv* adalah sebuah parameter yang menentukan kestabilan pola konsumsi dalam skala tertentu.



Gambar 5.7 Nilai ambang optimal metode Windowed Analysis

5.2. Hasil Uji Coba Model dan Aplikasi.

Data yang digunakan pada uji coba model dan aplikasi tugas akhir ini adalah data pelanggan rumah tangga berkode tarif pelanggan 4b2. Uji coba model dilakukan dengan rentang data yang sama dengan pembuatan model yaitu 24 bulan, yaitu

selama Januari 2014 hingga Desember 2015. Sedangkan uji coba aplikasi dilakukan dengan rentang data terbaik yang didapat setelah uji coba model dilakukan. Dari keseluruhan data ini hanya pelanggan yang memiliki data konsumsi minimal 24 bulan yang digunakan dalam proses uji coba.

5.2.1. Uji Coba Parameter Model

Uji coba dilakukan terhadap model yang telah dibangun dalam hasil pembuatan model. Tujuan uji coba model adalah untuk mengetahui hasil yang paling optimal dari proses pembuatan model pada proses pelatihan telah sesuai dengan hasil optimal pada proses uji coba dengan melihat tingkat prediksi benaran dibandingkan dengan data aktual (*recall*). Model terbaik jika memiliki nilai *recall* lebih tinggi dibanding model pada proses uji coba yang lainnya.

5.2.1.1. Hasil Uji Coba Metode *Pearson's Correlation*

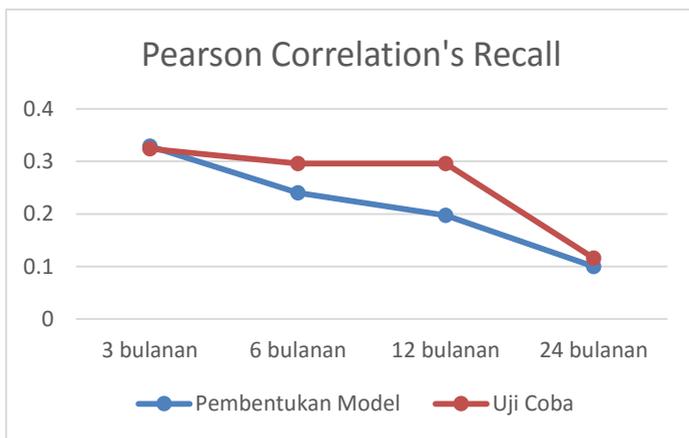
Model dari metode *Pearson's correlation* yang paling optimal pada proses pembuatan model selanjutnya dilakukan uji coba.

Tabel 5.14 Hasil Uji Coba Metode *Pearson's Correlation*

Pearson's Correlation – Uji Coba				
Rentang Data	Pembuatan Model		Uji Coba	
	Hasil benar	PKTS	Hasil benar	PKTS
3 bulanan	26	79	23	71
6 bulanan	18	75	21	71
12 bulanan	14	71	21	71
24 bulanan	7	70	8	69

Model yang paling optimal dari proses pembuatan model adalah dengan menggunakan rentang data 3 bulanan dengan

mendapatkan angka hasil benar terbanyak yaitu 26 sedangkan pada uji coba mendapatkan hasil benar terbanyak yaitu 23 yang masing-masing memiliki *recall* 0.329 dan 0.323. Nilai recall pada masing-masing percobaan dapat dilihat pada Gambar 5.8. Sehingga dapat dikatakan bahwa model Pearson's correlation dengan rentang data 3 bulanan baik untuk digunakan untuk memprediksi pelanggan dengan konsumsi tidak sah. Tabulasi dari percobaan uji coba dan pembuatan model metode *Pearson's Correlation* dapat dilihat pada Tabel 5.14.



Gambar 5.8 Hasil Recall Uji Coba Pearson's Correlation

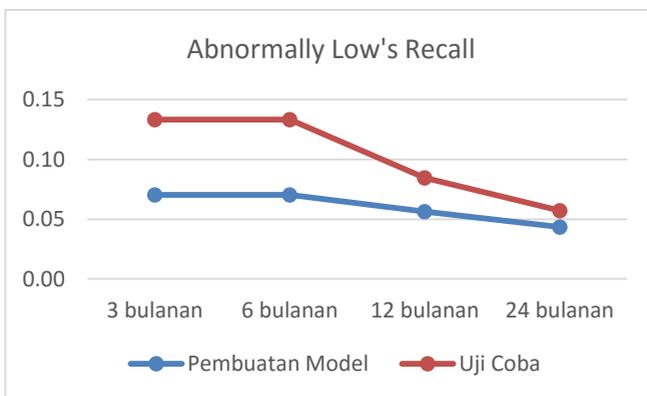
5.2.1.2. Hasil Uji Coba Metode *Abnormally Low Consumption*

Model dari metode *abnormally low consumption* yang paling optimal pada proses pembuatan model selanjutnya dilakukan uji coba. Model yang paling optimal dari proses pembuatan model adalah dengan menggunakan rentang data 3 bulanan atau 6 bulanan dengan mendapatkan angka hasil benar terbanyak yaitu masing-masing 5, sedangkan pada uji coba mendapatkan hasil benar terbanyak yaitu masing-masing 10 untuk rentang data 3 bulanan dan 6 bulanan.

Tabel 5.15 Hasil Uji Coba Metode Abnormally Low Consumption

Low Abnormal Consumption – 3 bulanan				
Rentang Data	Pembuatan Model		Uji Coba	
	Hasil benar	PKTS	Hasil benar	PKTS
3 bulanan	5	71	10	75
6 bulanan	5	71	10	75
12 bulanan	4	71	6	71
24 bulanan	3	69	4	70

Tabulasi dari uji coba dan pembuatan model dari metode *Abnormally Low Consumption* dapat dilihat pada Tabel 5.15. Sehingga dapat dikatakan bahwa metode *abnormally low consumption* dengan rentang data 3 bulanan atau 6 bulanan baik untuk digunakan untuk memprediksi pelanggan dengan konsumsi tidak sah. Perbandingan perolehan pembuatan model dan uji coba dari metode *Abnormally Low Consumption* dapat dilihat pada Gambar 5.9.



Gambar 5.9 Hasil Recall Uji Coba Low Abnormal Consumption

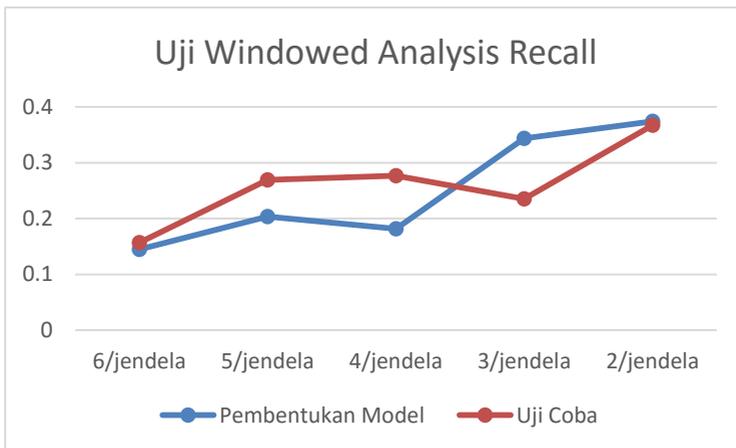
5.2.1.3. Hasil Uji Coba Metode *Windowed Analysis*

Model dari metode *windowed analysis* yang paling optimal pada proses pembuatan model selanjutnya dilakukan uji coba. Model yang paling optimal dari proses pembuatan model adalah dengan menggunakan rentang data 12 bulan atau 2 bulan dalam satu jendela dengan mendapatkan angka hasil benar terbanyak 31 dari total pelanggan KTS 69 dalam 24 bulan, sedangkan pada uji coba mendapatkan hasil benar terbanyak yaitu 24 dari total pelanggan KTS 79 pelanggan dalam 24 bulan yang masing-masing memiliki nilai *recall* 0.470 dan 0.30.

Tabel 5.16 Hasil Uji Coba Metode *Windowed Analysis*

Windowed Analysis				
Jumlah bulan	Pembuatan Model		Uji Coba	
	Hasil benar	PKTS	Hasil benar	PKTS
6 bulan dalam satu jendela	10	69	11	70
2 bulan dalam satu jendela	25	69	24	79
5 bulan dalam satu jendela	11	54	14	52
4 bulan dalam satu jendela	8	44	13	47
3 bulan dalam satu jendela	11	32	8	34

Tabulasi dari perolehan nilai benar dibandingkan dengan nilai aktual pada pembuatan model dan uji coba dari metode *windowed analysis* dapat dilihat pada ini dapat dilihat pada Tabel 5.16. Sehingga dapat dikatakan bahwa metode analisis jendela dengan rentang data 12 bulanan baik untuk digunakan untuk memprediksi pelanggan dengan konsumsi tidak sah. Pada Gambar 5.10 dapat dilihat perbandingan antara hasil *recall* pembuatan model dengan proses uji coba metode *windowed analysis*.



Gambar 5.10 Hasil Recall Uji Coba Windowed Analysis

5.2.2. Uji Coba Aplikasi

Setelah tahap implementasi sistem selesai dikerjakan, maka tahap berikutnya adalah tahap uji coba sistem. Yaitu melakukan pengujian terhadap aplikasi yang telah dikembangkan. Tujuan dari tahap uji coba ini adalah untuk menguji apakah aplikasi yang telah dikembangkan tersebut telah berjalan dengan benar dan sesuai dengan perhitungan manual dengan masing-masing metode yang telah dilakukan. Tahap uji coba ini terdiri dari empat sub-bab yaitu lingkungan uji coba, data uji coba, skenario uji coba, pelaksanaan dan hasil uji coba.

5.2.2.1. Lingkungan Uji Coba Aplikasi

Lingkungan yang digunakan untuk melakukan uji coba program adalah sistem operasi *Windows 10*. Perangkat lunak yang digunakan untuk melakukan komputasi adalah Netbeans 8.1 dan aplikasi pengolahan data *Microsoft Excel*. Sedangkan perangkat keras yang dimanfaatkan yaitu laptop Intel Core i5 dengan RAM 4 GB.

5.2.2.2. Data Uji Coba Aplikasi

Data yang digunakan pada uji coba aplikasi tugas akhir ini adalah data pelanggan rumah tangga berkode tarif pelanggan 4b2.

Tabel 5.17 Pelanggan KTS PDAM Januari – Desember 2015

No	Nopel	No	Nopel
1	1092329	16	4038426
2	1141425	17	4045039
3	1143419	18	4090746
4	1155766	19	4150671
5	1164716	20	4272668
6	2015204	21	4272735
7	2030746	22	4320267
8	2061737	23	5011164
9	2140977	24	5040481
10	2162806	25	5105163
11	2190744	26	5222638
12	2270768	27	5250953
13	2272333	28	5331023
14	3101333	29	5411138
15	4031604		

Data yang digunakan pada proses uji coba aplikasi ini terdiri dari 3 bagian, yaitu :

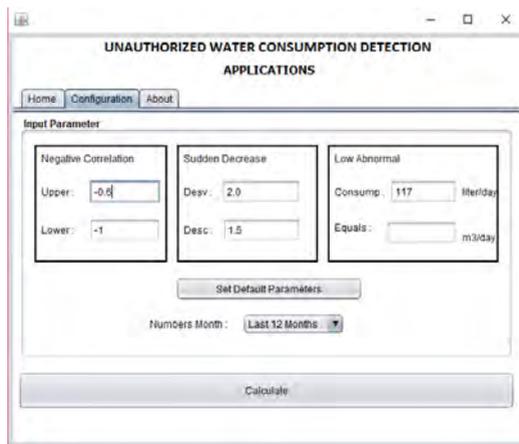
- a. Bagian pertama berupa data konsumsi air minum setiap bulan pelanggan bertipe 4b2 dari rentang data januari 2015 hingga desember 2015. Dari keseluruhan data ini hanya pelanggan yang memiliki data konsumsi minimal 24 bulan yang digunakan dalam proses uji coba.
- b. Bagian kedua adalah jumlah kapita pelanggan bertipe 4b2 selama 1 tahun terakhir. Dari keseluruhan data ini hanya pelanggan yang memiliki jumlah kapita yang tercatat yang dapat digunakan dalam proses uji coba.
- c. Bagian kedua terdiri dari 29 orang pelanggan air minum yang telah terbukti melakukan konsumsi air minum tidak sah menurut PDAM Surya Sembada Surabaya dalam satu tahun terakhir (januari 2015 hingga desember 2015) dengan melakukan pembayaran denda terhadap PDAM. Dapat dilihat pada Tabel 5.17.

5.2.2.3. Hasil Uji Kebenaran Aplikasi

Uji coba aplikasi yang dilakukan adalah uji kebenaran. Uji kebenaran adalah uji yang digunakan untuk memastikan bahwa implementasi algoritma pada aplikasi yang dibuat memberikan hasil yang sesuai. Dalam sub-bab ini dilakukan uji coba dilakukan dalam rentang data data dalam 1 tahun terakhir. Pemilihan rentang data ini berdasarkan uji coba pemilihan model yang paling optimal dari percobaan yang telah dilakukan pada sub-bab 5.2.1.

Proses uji kebenaran dilakukan dengan menggunakan data 58 nomor pelanggan dalam rentang data antara Januari 2015 hingga Desember 2015. Nilai ambang untuk Pearson's correlation adalah -0.6 dan nilai desv dan desc pada metode *windowed analysis* adalah 2 dan 1.5. Penentuan nilai ambang diambil dari percobaan yang dilakukan pada sub-bab 5.1.2. Ilustrasi masukan parameter masing-masing metode pada aplikasi dapat dilihat pada Gambar 5.11. Selanjutnya hasil deteksi anomali melalui aplikasi dibandingkan dengan hasil

deteksi anomali melalui perhitungan secara manual menggunakan *Microsoft Office Excel*.



Gambar 5.11 Masukkan Parameter Dalam Aplikasi

No	Nopol	Description
1	1092329	Konsumsi turun progresif
2	1143419	Konsumsi Turun Drastis
3	1155766	Konsumsi Turun Drastis
4	2015204	Konsumsi Turun Drastis
5	2030746	Konsumsi turun progresif, Konsumsi Turun Drastis.
6	2162806	Konsumsi turun progresif, Konsumsi Turun Drastis.
7	2270768	Konsumsi Turun Drastis
8	2272333	Konsumsi turun progresif, Konsumsi Turun Drastis.
9	4031604	Konsumsi Rendah Abnormal
10	4045039	Konsumsi Turun Drastis
11	4150671	Konsumsi Turun Drastis
12	5011164	Konsumsi Turun Drastis
13	5040481	Konsumsi turun progresif, Konsumsi Turun Drastis.
14	1114885	Konsumsi Rendah Abnormal
15	1145023	Konsumsi Turun Drastis
16	1180376	Konsumsi Rendah Abnormal
17	1181888	Konsumsi Rendah Abnormal
18	2011904	Konsumsi Rendah Abnormal
19	2014017	Konsumsi Rendah Abnormal

Gambar 5.12 Hasil Deteksi Anomali Aplikasi

Hasil deteksi anomali melalui aplikasi disajikan pada Gambar 5.12. Tabel berisi nomor pelanggan yang memiliki pola konsumsi anomali dan keterangan yang menjelaskan pola anomali tersebut. Proses deteksi anomali secara manual dilakukan dengan menggunakan nilai parameter yang sama

dilakukan pada *Microsoft Excel*. Proses deteksi anomali dilakukan pada masing-masing metode dengan masing-masing parameter. Hasil deteksi anomali secara manual dapat dilihat pada Tabel 5.18.

Tabel 5.18 Hasil Deteksi Manual *Microsoft Excel*

No	Pearson's Correlation	Windowed Analysis	Abnormally Low Consumption
1	1092329		
2		1143419	
3		1155766	
4		2015204	
5	2030746	2030746	
6	2162806	2162806	
7		2270768	
8	2272333	2272333	
9			4031604
10		4045039	
11		4150671	
12		5011164	
13	5040481	5040481	
14			1114885
15		1145023	
16			1180376
17			1181888
18			2011904
19			2014017

5.3. Hasil Penentuan Pelanggan KTS

Pada bagian penentuan pelanggan konsumsi tidak sah (KTS) ini dilakukan untuk mengetahui kemampuan metode dan juga aplikasi dalam mendeteksi pola penurunan konsumsi air minum yang diduga anomali. Pada bagian ini, pembuktian pelanggan KTS dilakukan dengan data pelanggan yang telah mengkonsumsi air minum secara tidak sah (KTS) yang dimiliki oleh PDAM Surya Sembada Surabaya. PDAM menemukan pelanggan KTS tersebut melalui berbagai cara, yaitu: dengan melakukan penertiban langsung, melakukan pengecekan terhadap meteran air pelanggan, dan juga menerima aduan / laporan dari masyarakat sekitar.

Dalam uji coba deteksi anomali ini menggunakan 3 metode secara bersama-sama. Dengan asumsi bahwa jika pelanggan terdeteksi dari salah satu metode maka pelanggan tersebut telah mengkonsumsi air minum secara tidak sah.

Jangka waktu yang digunakan dalam uji coba ini adalah konsumsi sepanjang 12 bulan atau satu tahun, dikarenakan pertimbangan pada sub-bab 5.1 yang dijelaskan dalam visualisasi Gambar 4.2 hingga Gambar 4.10, yang menerangkan semakin pendek jangka waktu yang digunakan dalam menganalisa pola konsumsi maka akan semakin akurat atau semakin optimal hasil yang diperoleh.

Nilai parameter ambang *Pearson's Correlation* yang digunakan pada tahap ini adalah -0.6, parameter *windowed analysis* adalah 2.0 dan 1.5 untuk *desv* dan *desc* secara berurutan, dan asumsi konsumsi minimal harian rata-rata rumah tangga per kapita adalah sebesar 117 liter air minum. Penentuan nilai-nilai parameter tersebut berdasarkan penjelasan pada sub-bab 5.1.2 yang telah ditunjukkan grafik pada Gambar 5.4 dan Gambar 5.5. Semakin besar nilai ambang dari *Pearson's Correlation* dan nilai *desv* pada *windowed analysis*

akan menyebabkan bertambahnya objek anomali yang terdeteksi dikarenakan nilai ambang dari kedua metode tersebut diperlebar. Sedangkan semakin tinggi nilai *desc* pada *windowed analysis* maka akan menyebabkan kemampuan metode dalam mendeteksi anomali semakin berkurang, dikarenakan metode hanya akan mengklasifikasikan pola penurunan yang semakin curam.

Tabel 5.19 Hasil Penentuan Pelanggan KTS

Periode waktu	Metode	Jumlah Pelanggan terdeteksi anomali	Nomor Pelanggan
1 tahun (Januari 2015 – Desember 2015)	<i>Pearson's Correlation</i>	5 Pelanggan	(1092329*), (2030746*), (2162806*), (2272333*), (5040481*).
	<i>Window-ed Analysis</i>	12 Pelanggan	(1143419*), (1155766*), (2015204*), (2030746*), (2162806*), (2270768*), (2272333*), (4045039*), (4150671*), (5011164*), (5040481*), (1145023).
	<i>Low Abnorm-ally Consumption</i>	6 Pelanggan	(4031604*), (1114885), (1180376), (1181888), (2011904), (2014017)

Proses uji coba deteksi anomali dilakukan pada 58 pelanggan rumah tangga dengan kode tarif 4b2. Komposisi data yang digunakan adalah 50:50 antara pelanggan KTS dengan pelanggan non KTS. Hasil uji coba deteksi anomali disajikan dalam Tabel 5.19. Dari 58 pelanggan dalam uji coba yang digunakan pelanggan dengan nomor pelanggan bertanda (*) merupakan pelanggan yang telah terbukti melakukan konsumsi air minum secara tidak sah oleh PDAM Surya Sembada Surabaya.

5.4. Analisis Hasil Uji Coba Aplikasi dan Penentuan Pelanggan KTS

Dari hasil uji coba untuk masing-masing model dan aplikasi yang telah dilakukan, maka dari hasil tersebut dapat dianalisis dan dirangkum. Analisis tersebut dapat dilihat pada sub-bab:

5.4.1. Analisis Hasil Uji Kebenaran Aplikasi

Dari hasil proses pengklasifikasian dengan menggunakan masing-masing metode. Baik itu yang dihasilkan oleh aplikasi maupun oleh perhitungan manual (dengan bantuan fasilitas *Microsoft Excel*) didapatkan hasil deteksi anomali yang sama dari ketiga metode yang dijalankan.

Hal tersebut dapat dilihat dari kolom "*Description*" pada tampilan aplikasi (dapat dilihat pada Gambar 5.12) yang menjelaskan pola anomali yang terdeteksi oleh aplikasi. Jika pada kolom tersebut bertuliskan "konsumsi turun progresif" maka nomor pelanggan tersebut terdeteksi dengan metode *Pearson's Correlation* masing-masing nomor pelanggan pada kolom masing-masing metode. Jika pada kolom tersebut bertuliskan "konsumsi turun drastis" maka nomor pelanggan tersebut terdeteksi dengan metode *Windowed Analysis* masing-masing nomor pelanggan pada kolom masing-masing metode. Jika pada kolom tersebut bertuliskan "konsumsi rendah abnormal" maka nomor pelanggan tersebut terdeteksi dengan metode *Low Abnormally Consumption* masing-masing nomor pelanggan pada kolom masing-masing metode.

5.4.2. Analisis Penentuan Pelanggan KTS

- a. Metode yang dibangun mampu mendeteksi anomali dari pola penurunan konsumsi pelanggan dan pola konsumsi rendah abnormal. Pada Tabel 5.19 adalah pelanggan-pelanggan yang terklasifikasikan oleh masing-masing metode yang digunakan.
- b. Hasil pengklasifikasian pelanggan sebagai anomali bukan hanya didasarkan pada selisih konsumsi air minum dari bulan sebelumnya, akan tetapi menganalisa

dari jangka waktu yang dapat ditentukan oleh pengguna aplikasi dengan ketiga metode analisis.

- c. Pelanggan yang telah terbukti melakukan konsumsi air minum tidak sah oleh PDAM Surya Sembada Surabaya namun tidak terdeteksi oleh masing-masing metode dalam perhitungan maupun aplikasi dikarenakan pelanggan dapat dilihat pada Tabel 5.20.

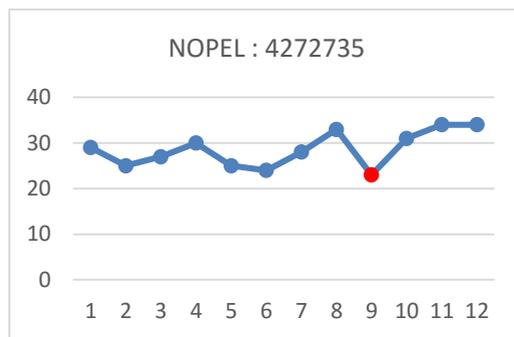
Tabel 5.20 Detail Pelanggan yang Tidak Terdeteksi

No	Nopel	Nilai Korelasi	Konsumsi Minimal Per Kapita Per Day (liter)
1	1141425	-0.2259926	325
2	1164716	0.54364483	181.25
3	2061737	0.36107653	361.1111111
4	2140977	0.88856147	320
5	2190744	-0.5427375	238.3333333
6	3101333	0.21561259	621.7592593
7	4038426	-0.4537423	253.8888889
8	4090746	0.62741945	240.625
9	4272668	-0.5206335	166.6666667
10	4272735	0.50810303	238.1944444
11	4320267	0.77053447	554.0123457
12	5105163	0.00440127	508.3333333
13	5222638	-0.0927357	191.6666667
14	5250953	0.37958267	341.6666667
15	5331023	0.30293869	357.2222222
16	5411138	0.39338237	245.3703704

Hal tersebut dikarenakan PKTS memiliki nilai korelasi lebih tinggi dari -0.6, tidak memenuhi kriteria penurunan drastis dari metode *windowed analysis*, dan konsumsinya diatas rata-rata batas minimum yang

dikalkulasi dari metode *abnormally low consumption* dalam satu tahunnya yaitu 117 liter. Seperti yang dapat dilihat pada Tabel 5.20.

- d. PDAM Surya Sembada Surabaya dalam mencari pola konsumsi anomali menggunakan analisa dari satu bulan sebelumnya, yaitu melihat penurunan konsumsi sebesar 30% dari bulan lalu. Salah satu contohnya adalah pelanggan bernomor 4272735 yang pola konsumsinya dapat dilihat pada Gambar 5.13.

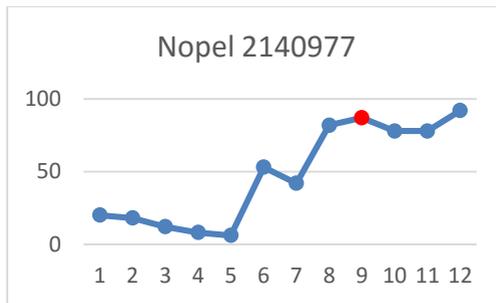


Gambar 5.13 Pola Konsumsi Penurunan 30 persen

Pelanggan yang memenuhi kriteria tersebut akan akan di-identifikasi lebih lanjut sebelum dapat ditentukan pelanggan tersebut melakukan KTS yang kemudian diharuskan membayar denda. Analisa penurunan konsumsi seperti hal tersebut tidak mampu untuk di-identifikasi oleh ketiga metode yang digunakan, tercatat dari data yang digunakan terdapat 9 dari 16 pelanggan yang teridentifikasi melalui analisis penurunan 30%. Detail dari informasi pelanggan ini dapat dilihat pada lampiran C. Pelanggan dengan nomor pelanggan ini melakukan pembayaran denda pada bulan ke-9 tepat terjadi penurunan.

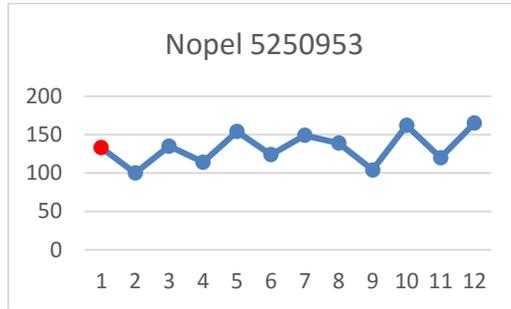
- e. Pola konsumsi air minum yang diperoleh juga terdapat yang tidak dapat diidentifikasi oleh analisa milik PDAM dan juga ketiga metode yang digunakan dalam

aplikasi ini. Seperti yang telah dijelaskan pada sub-bab 4.1 dan diperlihatkan pada Tabel 4.5 terdapat berbagai macam tindakan yang memicu terjadinya konsumsi tidak sah. Masing-masing dari tindakan tersebut perlu adanya pengamatan langsung dilapangan terhadap konsumsi ilegal yang sengaja atau tidak sengaja dilakukan. Pelanggan dengan pola konsumsi seperti ini terdeteksi terdapat sebanyak 5 dari 16 pelanggan yang tidak terdeteksi. Detail dari informasi pelanggan ini dapat dilihat pada lampiran C. Salah satu contoh pola data dari nomor pelanggan 2140977 dapat dilihat pada Gambar 5.14. Pelanggan tersebut melakukan pembayaran konsumsi tidak sah pada bulan 9 untuk konsumsi sebanyak 52m³.



Gambar 5.14 Pola Konsumsi Tidak Terdeteksi Penurunan 30 Persen Maupun Ketiga Metode Statistik

- f. Pelanggan yang terbukti melakukan KTS dan telah membayarkan denda terhadap konsumsinya pada awal-awal tahun akan tidak terdeteksi oleh ketiga metode. Tercatat terdapat 2 pelanggan dari 16 pelanggan yang tidak terdeteksi yang membayarkan denda konsumsinya pada awal tahun. Detail informasi pelanggan ini dapat dilihat pada lampiran C. Salah satu contoh pola konsumsi seperti diperlihatkan pada Gambar 5.15, pelanggan yang membayarkan denda pada Januari 2015.



Gambar 5.15 Konsumsi Tidak Sah Pada Awal Tahun Analisis

5.5. Kesimpulan Hasil Eksperimen

Dari eksperimen yang telah dilakukan maka didapatkan kesimpulan sebagai berikut :

1. Dari masing-masing metode yang diaplikasikan. Ketiga metode tersebut saling melengkapi untuk menganalisa pola konsumsi abnormal air minum.
2. Sebagai Pemilihan rentang data paling optimal untuk masing-masing metode adalah berikut:
 - a. Metode *Pearson's correlation* menggunakan rentang data 3 bulanan.
 - b. Metode konsumsi rendah abnormal menggunakan rentang data 3 bulanan atau 6 bulanan.
 - c. Metode *windowed analysis* menggunakan rentang data minimal 12 bulan atau 2 bulan dalam satu jendela.
3. Semakin banyak perhitungan yang dilakukan dalam rentang data tertentu (dalam tugas akhir ini, setiap 3 bulan dalam 2 tahun) akan meningkatkan hasil temuan.
4. Penentuan nilai ambang dan parameter pada masing-masing metode sangat berpengaruh terhadap jumlah anomali yang terdeteksi.
 - a. Semakin besar nilai ambang dari *Pearson's Correlation* dan nilai *desv* pada *windowed analysis* akan menyebabkan bertambahnya objek anomali yang terdeteksi (menyebabkan menurunnya presisi).

Dikarenakan nilai ambang dari kedua metode tersebut diperlebar.

- b. Semakin tinggi parameter *desc* maka akan menyebabkan kemampuan metode dalam mendeteksi anomali semakin berkurang, dikarenakan metode hanya akan mengklasifikasikan pola penurunan yang semakin curam.
 - c. Rekomendasi parameter pada metode *windowed analysis* untuk identifikasi anomali pada tugas akhir ini adalah 1.7 untuk *desv* dan 1.5 untuk *desc*.
 - d. Rekomendasi nilai ambang *Pearson's Correlation* adalah -0.7 atau 0.6.
5. Metode konsumsi rendah abnormal sangat bergantung kepada keakurasian jumlah kapita dalam satu nomor pelanggan yang dimiliki oleh perusahaan. Semakin mutakhir data yang dimiliki oleh perusahaan pada setiap nomor pelanggan, semakin akurat perhitungan terhadap anomali yang dapat dilakukan.
6. Kelemahan dari metode yang digunakan adalah antara lain:
- a. Tidak dapat mendeteksi penurunan secara tiba-tiba kemudian diikuti oleh konsumsi stabil seperti semula (dijelaskan pada poin d analisis hasil uji coba deteksi anomali).
 - b. Tidak dapat mendeteksi pelanggan yang mengkonsumsi air secara tidak sah namun tidak mengalami pola penurunan maupun pola dibawah konsumsi minimal rata-rata.
 - c. Tidak dapat mendeteksi pelanggan yang terbukti melakukan KTS pada awal bulan, yang mana dikarenakan menurunnya pola konsumsi pada tahun yang lalu (dijelaskan pada poin e analisis hasil uji coba deteksi anomali).
7. Kinerja dari uji coba pada aplikasi dan model yang dibangun memiliki akurasi sebesar 62%, presisi 68%, dan *recall* sebesar 44%.

LAMPIRAN A KODE PROGRAM

```
//NORMALISASI METHOD

public void pickNormal(double[][] kons_new, int
starts) {

double[] max = new double[lastfix];
Arrays.fill(max, Short.MIN_VALUE);
double[] min = new double[lastfix];
Arrays.fill(min, Short.MAX_VALUE);
double[][] kons_norm = new double [lastfix][lastCol];

double norm_top = 1;
    double norm_bot = 0;

        //finding max;
for (int row =0; row<lastfix;row++) {
for (int col = 0; col < lastCol; col++) {
    if (kons_new[row][col+starts] > max[row]) {
        max[row] = kons_new[row][col+starts];} } }

        //finding min;
for (int row =0; row<lastfix;row++) {
for (int col = 0; col < lastCol; col++) {
    if (kons_new[row][col+starts] < min[row]) {
        min[row] = kons_new[row][col+starts];}
} }

        //normalization
for (int row =0; row<lastfix;row++) {
for (int col = 0; col < lastCol; col++) {
    kons_norm[row][col] = (norm_top-
norm_bot)/(max[row]-
min[row])*(kons_new[row][col+starts]-
max[row])+norm_bot;} }
    this.kons_norm = kons_norm; }
```

```
//RENTANG WAKTU METHOD

public double[][]splitRange(double[][]newArray, int
lastRow, int rangeStarts) {
    int newCol = 24-rangeStarts;
    int starts = rangeStarts;

    //set value range untuk panjang array
    double [][]rArray = new
double[lastRow][newCol];
    for (int row = 0; row < lastRow; row++) {
        //range menunjukkan dia bakal iterasi
        col sebanyak berapa x
        for (int col = 0; col < newCol; col++)
        {
            rArray[row][col] =
newArray[row][col+starts];
        }
    } this.lastCol = newCol;
    return rArray;
}
```

```

// PEARSON'S CORRELATION METHOD

public double correlation(double upper, double lower)
{
    // Membuat Array Bulan
    double[] monthCorrel = new double[lastCol];
    for (int m = 0; m < lastCol; m++) { monthCorrel[m] =
m+1; }

        String[] getR = new String[lastfix];
        // correlation method
        SimpleRegression regression = new
SimpleRegression();
// throw exception
if (monthCorrel.length != kons_range[0].length) {
throw new
DimensionMismatchException(monthCorrel.length,
kons_range[lastfix-1].length); }

else if (monthCorrel.length < 2) {
throw new
MathIllegalArgumentException(LocalizedFormats.INSUFFIC
IENT_DIMENSION, monthCorrel.length, 2); }
else {
    //print result
    for(int row=0; row<lastfix;row++) {
        for(int col=0; col<lastCol; col++) {
            regression.addData(monthCorrel[col],
kons_range[row][col]); }
if(regression.getR() <= upper && regression.getR() >=
lower){
        getR[row] = "Konsumsi turun progresif"; }
else getR[row] = "";
        regression.clear(); }
    this.getR = getR;
}
}

```

```

// WINDOWED ANALYSIS METHOD

public void windowed(double desv, double desc){
    //inisiiasi array
    int constrain = 5*kons_range.length;
    double[][] sixth = new double[constrain][];
    double[][] fifth = new double[constrain][];
    double[][] fourth = new double [constrain][];
    double[][] third = new double [constrain][];

    double[] h_six = new double[sixth.length];
    double[] l_six = new double[sixth.length];
    double[] h_fif = new double[fifth.length];
    double[] l_fif = new double[fifth.length];
    double[] h_four = new double[fourth.length];
    double[] l_four = new double[fourth.length];
    double[] h_thir = new double[third.length];
    double[] l_thir = new double[third.length];

    double[] a_six = new double[sixth.length];
    double[] m_six = new double[sixth.length];
    double[] a_fif = new double[fifth.length];
    double[] m_fif = new double[fifth.length];
    double[] a_four = new double[fourth.length];
    double[] m_four = new double[fourth.length];
    double[] a_thir = new double[third.length];
    double[] m_thir = new double[third.length];

    boolean[] temp = new boolean[constrain];
    boolean[] temp_1= new boolean[constrain];
    boolean[] temp_2= new boolean[constrain];
    boolean[] temp_3= new boolean[constrain];
    boolean[] result = new boolean[constrain];
    String r_wind[] = new
String[kons_range.length];

    double newArray[] = new
double[kons_range.length*kons_range[0].length];

    for(int i=0; i<kons_range.length;i++) {
        double[] row = kons_norm[i];
        for(int j=0; j<row.length;j++) {
            double number = kons_norm[i][j];
            newArray[i*row.length+j] = number;
        }
    }

// WINDOWED ANALYSIS METHOD LANJUT HALAMAN SELANJUTNYA

```

```

// MEMBAGI JENDELA-JENDELA
    int c = 0;
if (lastCol == 24) { c = 6; }
else if (lastCol == 21 ) { c = 5;}
else if (lastCol == 18) { c = 4; }
else if (lastCol == 15) { c = 3; }
else if (lastCol == 12) { c = 2; }

    int next = 1;
    int x = kons_range[0].length;
    int a=2;
    int index = 0;
    for(int b=0;b<kons_range.length*5;b++){
        sixth[index] =
java.util.Arrays.copyOfRange(newArray,x-a,x);
        fifth[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c, x-a);
        fourth[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c-c, x-a-
c);
        third[index] =
java.util.Arrays.copyOfRange(newArray, x-a-c-c-c, x-a-
c-c);
        a++;
        index++;

        if(a==7){ next++; x =
kons_range[0].length; x = x*next; a=2; }}

//finding max values
    for(int row = 0;row<fifth.length;row++) {
        double max = Double.MIN_VALUE;
        double max_1 = Double.MIN_VALUE;
        double max_2 = Double.MIN_VALUE;
        for(int col = 0;col
<fifth[row].length;col++) {
if (fifth[row][col] > max){
max = fifth[row][col]; } h_fif[row] = max;

if (fourth[row][col] > max_1){
max_1 = fourth[row][col];
    } h_four[row] = max_1;

        if (third[row][col] > max_2){
max_2 = third[row][col];
    } h_thir[row] = max_2;
    }
}
// WINDOWED ANALYSIS METHOD LANJUT HALAMAN SELANJUTNYA

```

```

for(int row = 0;row<sixth.length;row++) {
    double max_3 = Double.MIN_VALUE;
    for(int col = 0;col
<sixth[row].length;col++) {
        /*
        Finding for highest element within window 6;
        */
if (sixth[row][col] > max_3){
max_3 = sixth[row][col]; } } }

    //finding min values
    for(int row =
0;row<fifth.length;row++) {
        double min = Double.MAX_VALUE;
        double min_1 = Double.MAX_VALUE;
        double min_2 = Double.MAX_VALUE;
        for(int col = 0;col
<fifth[row].length;col++) {
            if (fifth[row][col] < min){
                min = fifth[row][col];
            }
            l_fif[row] = min;
            if (fourth[row][col] < min_1){
                min_1 = fourth[row][col];
            }
            l_four[row] = min_1;
            if (third[row][col] < min_2){
                min_2 = third[row][col];
            }
            l_thir[row] = min_2;}
        }
for(int row = 0;row<sixth.length;row++) {
    double min_4 = Double.MAX_VALUE;
    for(int col = 0;col
<sixth[row].length;col++) {
        if (sixth[row][col] < min_4){
            min_4 = sixth[row][col];
        }
        l_six[row] = min_4; } } }
        //finding average
        for(int row = 0;row<fifth.length;row++) {
            double sum = 0;
            double sum_1 = 0;
            double sum_2 = 0;
            for(int col = 0;col
<fifth[row].length;col++) {
                sum += fifth[row][col];
                sum_1 += fourth[row][col];
                sum_2 += third[row][col];

            }
            a_fif[row] = sum/fifth[row].length;
            a_four[row] = sum_1/fourth[row].length;
            a_thir[row] = sum_2/third[row].length;
        }
}
// WINDOWED ANALYSIS METHOD LANJUT HALAMAN SELANJUTNYA

```

```

//finding average window #6
    for(int row = 0;row<sixth.length;row++) {
        double sum_3 = 0;
        for(int col = 0;col
<sixth[row].length;col++) {
            sum_3 +=sixth[row][col];
        } a_six[row] = sum_3/sixth[row].length;
    }
    //finding median
    for(int row = 0;row<fifth.length;row++) {
double[fifth[row].length];
        double[] list = new
double[fourth[row].length];
        double[] list_1 = new
double[third[row].length];
        double med = 0;
        double med_1 = 0;
        double med_2 = 0;

        for(int col = 0;col
<fifth[row].length;col++) {
            list[col] = fifth[row][col];
            list_1[col] = fourth[row][col];
            list_2[col] = third[row][col];
        } Arrays.sort(list);
        m_fif[row] = Process.median(list);
        Arrays.sort(list_1);
        m_four[row] = Process.median(list_1);
        Arrays.sort(list_2);
        m_thir[row] = Process.median(list_2);
    }
    /*
    finding median win#6
    */
    for(int row = 0;row<sixth.length;row++) {
double[sixth[row].length];
        for(int col = 0;col
<sixth[row].length;col++) {
            list_3[col] = sixth[row][col];

        } Arrays.sort(list_3);
        m_six[row] = Process.median(list_3);
    }
// WINDOWED ANALYSIS METHOD LANJUT HALAMAN SELANJUTNYA

```

```

//iterasi 5 kali di bawah ini rules of window
        for(int ittt =0;
        ittt<5*kons_range.length;ittt++){

if( (h_six[ittt]<(m_six[ittt]*desv) &&
l_six[ittt]>(m_six[ittt]/desv)
        && h_fif[ittt]<(m_fif[ittt]*desv) &&
l_fif[ittt]>(m_fif[ittt]/desv) &&
        a_six[ittt]<(a_fif[ittt]/desc)) == true
){
        temp[ittt] = true;
        }else temp[ittt]= false;

if( (h_six[ittt]<(m_six[ittt]*desv) &&
l_six[ittt]>(m_six[ittt]/desv)
        && h_fif[ittt]<(m_fif[ittt]*desv) &&
l_fif[ittt]>(m_fif[ittt]/desv)
        && h_four[ittt]<(m_four[ittt]*desv) &&
l_four[ittt]>(m_four[ittt]/desv)
        &&
a_six[ittt]<(a_four[ittt]/desc)&&a_fif[ittt]<(a_four[ittt]
/desc)) == true ){
        temp_1[ittt] = true;
        }else temp_1[ittt]= false;

if( (h_six[ittt]<(m_six[ittt]*desv) &&
l_six[ittt]>(m_six[ittt]/desv)
        && h_fif[ittt]<(m_fif[ittt]*desv) &&
l_fif[ittt]>(m_fif[ittt]/desv)
        && h_four[ittt]<(m_four[ittt]*desv) &&
l_four[ittt]>(m_four[ittt]/desv)
        && h_thir[ittt]<(m_thir[ittt]*desv) &&
l_thir[ittt]>(m_thir[ittt]/desv)
        && a_six[ittt]<(a_thir[ittt]/desc) &&
a_fif[ittt]<(a_thir[ittt]/desc)
        && a_four[ittt]<(a_thir[ittt]/desc)) ==
true ){
        temp_2[ittt] = true; } else temp_2[ittt]=
false;

if( (l_thir[ittt]>0 && l_four[ittt]>0 && l_fif[ittt]>0 &&
l_six[ittt]>0)== true) {
        temp_3[ittt] = true;
        } else temp_3[ittt] = false;

if( ((temp[ittt] || temp_1[ittt] ||temp_2[ittt]) &&
temp_3[ittt]) == true){
        result[ittt] = true;
        } else { result[ittt] = false;
        } }
// WINDOWED ANALYSIS METHOD LANJUT HALAMAN SELANJUTNYA

```

```

// FINAL RESULT WINDOWED ANALYSIS

int y = 5;
    int ab = 2;
    for(int i=0;i<kons_range.length;i++) {
        if((result[y-1] || result[y-2] ||
result[y-3] || result[y-4] || result[y-5])==true) {
            r_wind[i] = "Konsumsi Turun Drastis";
        } else{r_wind[i] = "";}

        y = 5;
        y = y*ab;
        ab++;
    } this.getW = r_wind;
}

```

```

// ABNORMALLY LOW CONSUMPTION METHOD

public void lowabn(int literConsump) {
    double [] jStream = new
double[kapita.length];
    double calc[] = new double[kapita.length];
    String lowabn[] = new String[kapita.length];
    double kons[][] = new
double[kapita.length][24];

    //ambil konsumsi aja
    for(int row = 0; row<kapita.length;row++) {
        DoubleStream jKons =
Arrays.stream(kons_range[row]);
        jStream[row] = jKons.sum();
    }

    //kalkulasi konsumsi minimum per kapita
    for(int row = 0;
row<kapita.length;row++) {
        calc[row] =
(kapita[row]*literConsump*30*lastCol)/1000;
    }

    for(int row = 0; row<kapita.length;row++) {

        if (jStream[row]<calc[row]) {
            lowabn[row] = "Konsumsi Rendah
Abnormal";
        } else lowabn[row] = "";
    } this.getA = lowabn; }
}

```

```

// METHOD MENENTUKAN PELANGGAN KTS

public void showing() {
    int rows = 0;
    String getR[] = new String[lastfix];
    String getW[] = new String[lastfix];
    String getA[] = new String[lastfix];
    getR=P.getR; getW=P.getW; getA=P.getA;
    for (int row = 0; row < lastfix; row++) {
        if (getR[row].equals("Konsumsi turun
progresif")
            || getW[row].equals("Konsumsi
Turun Drastis")
            || getA[row].equals("Konsumsi
Rendah Abnormal")) {
                jTableResult.setValueAt(rows+1, rows,
0);
                jTableResult.setValueAt(nopel[row],
rows, 1);

                if (getR[row].equals("Konsumsi turun
progresif")
                    && getW[row].equals("")
                    && getA[row].equals("")) {
                        jTableResult.setValueAt(getR[row],
rows, 2);
                    } else if (getR[row].equals("")
                        && getW[row].equals("Konsumsi
Turun Drastis")
                        && getA[row].equals("")) {
                            jTableResult.setValueAt(getW[row],rows,2);
                        }
                    }
            }
    }
// MENENTUKAN PELANGGAN KTS LANJUT HAL SELANJUTNYA

```

```

// METHOD MENENTUKAN PELANGGAN KTS
} else if(getR[row].equals(""))
    && getW[row].equals("")
    && getA[row].equals("Konsumsi
Rendah Abnormal")) {

jTableResult.setValueAt(getA[row],rows,2);
    } else if
(getR[row].equals("Konsumsi turun progresif")
    && getW[row].equals("Konsumsi
Turun Drastis")) {
    jTableResult.setValueAt(getR[row]+",
"+getW[row]+".", rows, 2);
    } else if (getR[row].equals("Konsumsi
turun progresif")
    && getA[row].equals("Konsumsi
Rendah Abnormal")) {
    jTableResult.setValueAt(getR[row]+",
"+getA[row]+".", rows, 2);
    } else if (getW[row].equals("Konsumsi
Turun Drastis")
    && getA[row].equals("Konsumsi
Rendah Abnormal")) {
    jTableResult.setValueAt(getW[row]+",
"+getA[row]+".", rows, 2);
    } else {
    jTableResult.setValueAt(getR[row]+",
"+getW[row]+", "+getA[row]+".", rows, 2);
    }
    rows++;
    }
} this.lastshowing = rows;
}

```

A-12

Halaman ini sengaja dikosongkan.

LAMPIRAN B

PERCOBAAN PARAMETER WINDOWED ANALYSIS

desv	2.5	Acc	0.5172414
desc	1.5	Recal	0.5185185
		Precision	0.4827586

desv	2.3	Acc	0.5344828
desc	1.5	Recal	0.5384615
		Precision	0.4827586

desv	2.5	Acc	0.5
desc	1.7	Recal	0.5
		Precision	0.4482759

desv	2.3	Acc	0.5172414
desc	1.7	Recal	0.52
		Precision	0.4482759

desv	2.5	Acc	0.5517241
desc	1.9	Recal	0.5652174
		Precision	0.4482759

desv	2.3	Acc	0.5517241
desc	1.9	Recal	0.5652174
		Precision	0.4482759

desv	2.5	Acc	0.5172414
desc	2.1	Recal	0.5294118
		Precision	0.3103448

desv	2.3	Acc	0.5172414
desc	2.1	Recal	0.5294118
		Precision	0.3103448

desv	2.5	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.2068966

desv	2.3	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.2068966

desv	2.5	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.137931

desv	2.3	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.137931

B-2

desv	2	Acc	0.5517241
desc	1.5	Recal	0.56
		Precision	0.4827586

desv	1.9	Acc	0.5517241
desc	1.5	Recal	0.56
		Precision	0.4827586

desv	2.1	Acc	0.5344828
desc	1.7	Recal	0.5416667
		Precision	0.4482759

desv	1.9	Acc	0.5517241
desc	1.7	Recal	0.5652174
		Precision	0.4482759

desv	2.1	Acc	0.5517241
desc	1.9	Recal	0.5652174
		Precision	0.4482759

desv	1.9	Acc	0.5689655
desc	1.9	Recal	0.5909091
		Precision	0.4482759

desv	2.1	Acc	0.5172414
desc	2.1	Recal	0.5333333
		Precision	0.2758621

desv	1.9	Acc	0.5344828
desc	2.1	Recal	0.5714286
		Precision	0.2758621

desv	2.1	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.2068966

desv	1.9	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.2068966

desv	2.1	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.137931

desv	1.9	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.137931

desv	1.7	Acc	0.5689655
desc	1.5	Recal	0.5833333
		Precision	0.4827586

desv	1.5	Acc	0.5344828
desc	1.5	Recal	0.5714286
		Precision	0.2758621

desv	1.7	Acc	0.5517241
desc	1.7	Recal	0.5714286
		Precision	0.4137931

desv	1.5	Acc	0.5172414
desc	1.7	Recal	0.5454545
		Precision	0.2068966

desv	1.7	Acc	0.5517241
desc	1.9	Recal	0.5714286
		Precision	0.4137931

desv	1.5	Acc	0.5172414
desc	1.9	Recal	0.5454545
		Precision	0.2068966

desv	1.7	Acc	0.5172414
desc	2.1	Recal	0.5384615
		Precision	0.2413793

desv	1.5	Acc	0.5
desc	2.1	Recal	0.5
		Precision	0.137931

desv	1.7	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.1724138

desv	1.5	Acc	0.5
desc	2.3	Recal	0.5
		Precision	0.0689655

desv	1.7	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.0689655

desv	1.5	Acc	0.5
desc	2.5	Recal	0.5
		Precision	0.0689655

B-4

Halaman ini sengaja dikosongkan.

LAMPIRAN C DETAIL NOMOR PELANGGAN KTS

NOPEL	Correlation Score	Cons/p/d	Month												month KTS	m3	Decrease (%)
			1	2	3	4	5	6	7	8	9	10	11	12			
1141425	-0.23	325.00	21	20	50	281	35	7	35	28	36	22	25	25	9	43	0.88
1164716	0.54	181.25	16	20	16	21	17	26	21	32	24	24	20	24	10	46	tidak teridentifikasi
2061737	0.36	361.11	39	41	48	46	42	43	40	41	40	42	48	50	12	52	tidak teridentifikasi
2140977	0.89	320.00	20	18	12	8	6	53	42	82	87	78	78	92	9	52	pola kenaikan
2190744	-0.54	238.33	44	33	45	42	45	33	27	40	22	29	38	31	11	52	0.45
3101333	0.22	621.76	22	21	20	20	20	27	22	914	25	30	192	30	12	48	0.84
4038426	-0.45	253.89	40	45	38	49	43	44	38	33	34	14	35	44	9	68	0.59
4090746	0.63	240.63	53	54	47	50	36	38	45	51	49	78	70	122	6	43	0.28
4272668	-0.52	166.67	18	29	19	16	15	15	3	15	7	13	16	14	12	43	0.53

C-2

Nopel	Correlation Score	Cons/p/d	Month												month KTS	m3	Decrease (%)
			1	2	3	4	5	6	7	8	9	10	11	12			
4272735	0.51	238.19	29	25	27	30	25	24	28	33	23	31	34	34	8	52	0.30
4320267	0.77	554.01	103	98	100	174	151	151	178	161	161	184	161	173	2	1236	awal bulan
5105163	0.00	508.33	26	28	32	32	32	38	28	31	25	36	39	19	11	24	0.51
5222638	-0.09	191.67	11	11	13	16	14	11	10	12	4	11	4	21	8	9	0.67
5250953	0.38	341.67	133	100	135	114	154	124	149	139	104	162	120	165	1	43	awal bulan
5331023	0.30	357.22	59	55	44	56	50	43	53	54	57	54	60	58	12	52	tidak teridentifikasi
5411138	0.39	245.37	18	21	19	19	24	21	21	30	29	18	23	22	6	43	tidak teridentifikasi

BAB VI

KESIMPULAN DAN SARAN

Pada bagian penutup ini dijelaskan kesimpulan yang diperoleh dari pengerjaan tugas akhir. Pada bab ini dijelaskan pula saran untuk pengembangan lebih lanjut dari tugas akhir ini.

6.1. Kesimpulan

Berdasarkan pada hasil uji coba dapat ditarik kesimpulan sebagai berikut:

1. Masing-masing metode yang digunakan mampu mendeteksi kondisi abnormal berupa pola penurunan dan konsumsi rendah dibawah rata-rata pada konsumsi air minum pelanggan PDAM Surya Sembada Surabaya. Metode *Pearson's Correlation* mampu mendeteksi pola penurunan progresif, metode *abnormally low consumption* mampu mendeteksi konsumsi rendah dibawah rata-rata kewajaran, dan metode *windowed analysis* mampu mendeteksi penurunan konsumsi secara drastis.
2. Dari hasil uji coba pada masing-masing metode, dapat disimpulkan bahwa:
 - a. Metode *Pearson's Correlation* dalam tugas akhir ini akan jauh baik jika menggunakan rentang data yang lebih singkat dalam analisis data yang dilakukan. Rentang data 3 bulanan yang digunakan dalam analisis tugas akhir ini akan lebih baik dibandingkan dengan rentang data 6 bulanan, 12 bulanan maupun 24 bulanan yaitu memperoleh nilai *recall* 0.32, 0.3, 0.3, 0.115 secara berurutan. Serta nilai ambang yang direkomendasikan dalam tugas akhir ini adalah -0.6 dan -0.7.
 - b. Metode *Abnormally Low Consumption* dalam tugas akhir ini akan jauh baik jika menggunakan rentang data yang lebih singkat dalam analisis data yang dilakukan. Rentang data 3 bulanan dan 6 bulanan

yang digunakan dalam analisis akan lebih baik dibandingkan dengan rentang data 12 bulanan maupun 24 bulanan yaitu memperoleh nilai *recall* 0.13, 0.13, 0.08, dan 0.06 secara berurutan.

- c. Metode *Windowed Analysis* dalam tugas akhir ini akan jauh baik jika menggunakan rentang data yang lebih singkat dalam analisis data yang dilakukan. Metode ini mampu mendeteksi lebih baik dibanding 2 metode lainnya. Rentang data 2 bulan dalam jendela 3,4, dan 5 yang digunakan dalam analisis akan lebih baik dibandingkan dengan rentang data 3 bulan dalam jendela 3,4, dan 5, rentang data 4 bulan dalam jendela 3,4, dan 5, maupun rentang data 6 bulan dalam jendela 3,4, dan 5 yaitu memperoleh nilai *recall* 0.30, 0.32, 0.28, 0.27, dan 0.16 secara berurutan. Serta parameter *desv* dan *desc* yang direkomendasikan dalam tugas akhir ini adalah 2.0 dan 1.5 atau 1.7 dan 1.5.

6.2. Saran

Saran yang dapat diberikan berdasarkan proses uji coba, penarikan kesimpulan, dan tujuan tugass akhir ini adalah: perlu dilakukan pemutakhiran secara periodik *dataset* yang dimiliki oleh PDAM Surya Sembada Surabaya mengenai jumlah kapita dalam satu nomor pelanggan. Mengingat terdapat lebih dari 65% nomor pelanggan pada tahun 2015 yang tidak tercatat memiliki nomor kapita. Hal ini penting untuk menunjang analisa dengan menggunakan metode *abnormally low consumption* atau metode lain yang memungkinkan menggunakan jumlah kapita.

DAFTAR PUSTAKA

- [1] B. Ruck, "Water," 2006. [Online]. Available: <https://www.worldvision.org.nz/media/72280/water.pdf>. [Accessed February 2016].
- [2] G. S. Putra, "Kinerja Perusahaan Daerah Air Minum (PDAM) Kabupaten Ponorogo Tahun 2007-2009," Universitas Indonesia, Jakarta, 2010.
- [3] P. B. Cheung, G. V. Girol, N. Abe and M. Propato, "Night Flow Analysis and Modeling for Leakage Estimation in a Water Distribution System," in *Taylor & Francis Group*, London, 2010.
- [4] P. W. Jowitt and C. Xu, "Optimal Valve Control in Water Distribution Network," in *American Society of Civil Engineer (ASCE)*, Edinburgh, U.K, 1990.
- [5] Departemen Pengendalian Kehilangan Air, "Water Balance Tahun 2014; Water Balance Tahun 2015," PDAM Surya Sembada Kota Surabaya, Surabaya, 2014; 2015.
- [6] Pers PDAM Surya Sembada Kota Surabaya, "Semarak Ulang Tahun PDAM Ke-38," 24 November 2014. [Online]. Available: http://www.pdam-sby.go.id/m/page.php?get=tampil_press_release&id=454001&bhs=1#.VrLj-Vn4bw0. [Accessed February 2016].
- [7] Badan Pusat Statistik Kota Surabaya, "Banyaknya Penduduk menurut Jenis Kelamin hasil Registrasi Tahun 2014," 2014. [Online]. [Accessed Februari 2016].
- [8] PDAM Surya Sembada Surabaya, "Laporan Keuangan," 2014. [Online]. Available: http://www.pdam-sby.go.id/page.php?get=tampil_keuangan&id=3001&bhs=. [Accessed February 2016].
- [9] B. Kingdom, R. Liemberger and P. Marin, "The Challenge of Reducing Non-Revenue Water (NRW) in Developing Countries," *Water Supply and Sanitation Sector Board*

Discussion Paper Series, vol. I, no. Non Revenue Water, p. No. 8, 2006.

- [10] A. Lambert, "Water Losses Management and Techniques," International Water Association (IWA), Berlin, 2001.
- [11] Arthur C. McIntosh, *Asian Water Supplies Reaching the Urban Poor*, Manila : Asian Development Bank, 2003.
- [12] T. D. TSI, Interviewee, *Anomaly Detection in PDAM Surya Sembada Kota Surabaya*. [Interview]. January 2016.
- [13] Jawa Pos, "Pencurian Air Meningkatkan Drastis," p. 39, 5 november 2015.
- [14] Departemen Pelaksanaan Penertiban Khusus , "Rekapitulasi Pembayaran Pelanggaran Pemakaian Air Pelaksanaan Penertiban Khusus / Sweeping," PDAM Surya Sembada Kota Surabaya, Surabaya, 2015.
- [15] I. Monedero, "Detection of Water Meter Under-Registration," *Water Resour. Plann. Manage.*, p. 4, 2015.
- [16] N. A. F. Vivin Sinita Indriani, "Evaluasi Unit Produksi Instalasi Pengelolaan Air Minum Karang Pilang I PDAM Surya Sembada Surabaya," Jurusan Teknik Lingkungan ITS, Surabaya, 2015.
- [17] H. E. Multikanga, S. K. Sharma and K. Vairavamoorthy, "Methods and Tools for Managing Losses in Water Distribution Systems," in *American Society of Civil Engineers*, Delft, 2013.
- [18] B. B. F. M. M. C. Meniconi S, "Transient tests for locating and sizing illegal branches in pipe systems.," in *Journal of Hydroinformatics*, Perugia, 2011.
- [19] L. J and C. L, "Inverse Transient Analysis in Pipe Networks," in *Journal of Hydraulic Engineering*, New York, 1994.
- [20] P. A and P. E, "Bayesian Modelling of Water Meters Ageing by Mixing Classes of Devices of Different States of Degradation," in *Revue de Statistique Appliquée*, 2004.
- [21] L. L.C and R. G, "Sustaining the NRW reduction strategy: The Manila Water Company territory management concept and

- monitoring tools," in *Water Loss Reduction Specialist*, Manila, 2012.
- [22] I. Mohamad and D. Usman, "Standardization and Its Effects on K-Means Clustering Algorithm," *Research Journal of Applied Sciences, Engineering and Technology*, 2013.
- [23] P.-N. Tan, "Introduction to data mining," in *Measures of Similarity and Dissimilarity*, Boston, Pearson Education, Inc., 2006, p. 77.
- [24] J. Church, in *Learning Haskell Data Analysis*, Birmingham, Packt Publishing Ltd., 2015, p. 138.
- [25] M. J. Katz, in *From Research to Manuscript A Guide to Scientific Writing*, Cleveland, Springer, 2006, p. 32.
- [26] "The idea of spread and standard deviation," Khan Academy, [Online]. Available: <https://www.khanacademy.org/math/probability/descriptive-statistics/variance-std-deviation/a/introduction-to-standard-deviation>. [Accessed February 2016].
- [27] D. A. Lund and M. Lund, "Pearson Product-Moment Correlation," Laerd, [Online]. Available: <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>. [Accessed February 2016].
- [28] "Encyclopedia of Database Technologies and Applications," in *Bioinformatics Data Management and Data Mining*, Hershey, Idea Group Reference, 2005, p. 31.
- [29] A. Danilenko, "The IBNET Water Supply and Sanitation Performance Blue Book," in *The International Benchmarking Network of Water and Sanitation Utilities Databook*, Washington, World Bank Publications, 2011, p. 82.
- [30] T. Fawcett, "Analysis, An Introduction to ROC," in *Science Direct*, Palo Alto, 2005.
- [31] Sugiyono, "Korelasi produk momen," in *Memahami Penelitian Kualitatif*, Bandung : ALPHABET, 2005, p. 212.

- [32] P. E. McKnight, K. M. McKnight and A. J. Figueredo, *Missing Data : A gentle Intriduction*, guilford Press, 2007.
- [33] G. Miligan and M. C. , "A study of standardization of variables in cluster analysis," *J. Classif*, 1988, pp. 181-204.
- [34] D. U. Ismail Mohamad, "Standardization and Its Effects on K-Means Clustering Algorithm," *Research Journal of Applied Sciences, Engineering and Technology* , 2013.

BIODATA PENULIS



Penulis lahir di Surabaya, 09 November 1994 yang merupakan anak bungsu dari 3 bersaudara. Penulis telah menempuh beberapa pendidikan formal, yaitu: SDN Medokan Ayu I Surabaya, SMP Negeri 12 Surabaya, SMA Negeri 16 Surabaya.

Penulis pada akhir tahun pelajaran SMA telah berhasil masuk Perguruan Tinggi Negeri Institut Teknologi Sepuluh Nopember Surabaya (ITS) pada jurusan Sistem Informasi melalui jalur SNMPTN tulis dengan NRP 5212100045. Selama berkuliah Penulis pernah aktif dalam beberapa kegiatan keorganisasian, seperti menjadi Ketua Angkatan 2012, staf Departemen Dalam Degeri HMSI ITS dan staf Kementerian Dalam Negeri BEM ITS pada tahun ke-2 perkuliahan, Kepala Departemen Dalam Negeri HMSI ITS pada tahun ke-3 perkuliahan. Disamping kegiatan keorganisasian penulis juga aktif dalam kegiatan pelatihan pengembangan karakter, seperti Latihan Keterampilan Manajemen Mahasiswa (LKMM) hingga tingkat dasar (LKMM TD), Future Leader Summit (FLS) BEM ITS dan Forum Indonesia Muda (FIM) 17. Serta penulis pernah aktif dalam kegiatan akademis, yaitu mengikuti perlombaan Manajemen Risiko tingkat Nasional yang berada di Universitas Katolik Parahyangan, Bandung hingga tahap finalis dan menjadi asisten praktikum Keamanan Aset Informasi (KAI) pada semester 7 perkuliahan.

Penulis mengambil laboratorium bidang minat Rekayasa Data dan Intelegensi Bisnis dikarenakan minatnya terhadap pengolahan dan analisis data. Penulis dapat dihubungi pada yogitricahyono@hotmail.com