

BAB II

KAJIAN PUSTAKA

2.1 *Business Intelligence*

Business Intelligence adalah seperangkat model matematika dan metodologi analisis yang mengeksploitasi data yang tersedia untuk menghasilkan informasi dan pengetahuan yang berguna untuk proses pengambilan keputusan yang kompleks. *Business Intelligence* menjelaskan tentang suatu konsep dan metode bagaimana untuk meningkatkan kualitas pengambilan keputusan bisnis berdasarkan sistem yang berbasis data. BI seringkali dipersamakan sebagaimana *briefing books, report and query tools*, dan sistem informasi eksekutif. BI merupakan sistem pendukung pengambilan keputusan yang berbasis data-data yang dilakukan dengan cara mengumpulkan, menyimpan, mengorganisasikan, membentuk ulang, meringkas data serta menyediakan informasi, baik berupa data aktifitas bisnis internal perusahaan, maupun data aktifitas bisnis eksternal perusahaan termasuk aktifitas bisnis para pesaing yang mudah diakses serta dianalisis untuk berbagai kegiatan manajemen *Business Intelligence* (BI) adalah rangkaian aplikasi dan teknologi untuk mengumpulkan, menyimpan, menganalisis, dan menyuguhkan akses data untuk membantu petinggi perusahaan dalam pengambilan keputusan.

2.1.1 Sejarah Perkembangan *Business Intelligence*

Pengambilan keputusan dalam setiap perusahaan membutuhkan arsitektur IT yang dapat memenuhi kebutuhan mereka. Menurut laporan dari *International Data Corporation* (IDC) di akhir musim gugur 2002, organisasi yang berhasil

mengimplementasikan dan menggunakan aplikasi analitik untuk bisnisnya memiliki peningkatan keuntungan, mulai dari 17% hingga 2000%. Sistem analitik bisnis atau umumnya dikenal sebagai *Business Intelligence* (BI) merupakan sebuah sistem yang mengevolusikan strategi, visi, dan arsitektur yang terus menerus mengarahkan operasional dan tujuan organisasi untuk mencapai tujuan bisnisnya. BI meliputi perolehan data dan informasi dari berbagai sumber yang bervariasi dan mengolahnya ke dalam pengambilan keputusan. BI dapat digunakan untuk mendukung perusahaan dalam mencapai berbagai kriteria keberhasilan [10].

2.1.2 Manfaat *Business Intelligence*

Beberapa manfaat yang bisa didapatkan bila suatu organisasi mengimplementasikan BI adalah sebagai berikut [10][11] seperti meningkatkan nilai data dan informasi organisasi dengan mengintegrasikan seluruh data sehingga menghasilkan pengambilan keputusan yang lengkap. Kemudian data dan informasi yang dihasilkan menjadi lebih mudah diakses dan lebih mudah untuk dimengerti (*user friendly*) lalu memudahkan pemantauan kinerja organisasi. Dalam mengukur kinerja suatu organisasi, seringkali dipergunakan ukuran yang disebut *Key Performance Indicator* (KPI). Kemudian meningkatkan nilai investasi teknologi informasi yang sudah ada tanpa harus mengubah atau menggantikan sistem informasi yang sudah digunakan sebelumnya. Kemudian menciptakan pegawai yang memiliki akses informasi yang baik (*wellinformed workers*) dan meningkatkan efisiensi biaya. BI dapat meningkatkan efisiensi karena mempermudah seseorang dalam melakukan pekerjaan, hemat waktu, dan mudah pemanfaatannya. Tabel. 2.1 dibawah ini

memaparkan perbedaan mendasar antara aplikasi BI dengan Sistem aplikasi umumnya.

Tabel 2.1 Perbedaan mendasar antara Sistem Aplikasi BI dan Sistem Aplikasi Umum lainnya

Sistem Aplikasi BI	Sistem Aplikasi Umum
1. Berpusat pada peluang bisnis.	1. Berpusat pada kebutuhan bisnis.
2. Diimplementasikan untuk seluruh departemen (<i>crossorganizational</i>).	2. Diimplementasikan untuk sebuah departemen tertentu.
3. <i>Requirement</i> bersifat informasi strategis.	3. <i>Requirement</i> bersifat fungsional
4. Membutuhkan analisis bisnis.	4. Membutuhkan analisis sistem.
5. Memerlukan pengembangan <i>iterative</i> dan evaluasi terus menerus	5. Pengembangan dapat dilakukan dengan teknik <i>waterfall</i>

2.2 Definisi Peramalan

Peramalan (*forecasting*) yaitu prediksi nilai-nilai sebuah peubah berdasarkan kepada nilai yang diketahui dari peubah tersebut atau peubah yang berhubungan. Meramal juga dapat didasarkan pada keahlian keputusan (*judgement*), yang pada gilirannya didasarkan pada data historis dan pengalaman [15]. Perpaduan antara seni dan ilmu dalam memperkirakan keadaan di masa yang akan datang, dengan cara memproyeksikan data-data masa lampau ke masa yang akan datang dengan menggunakan model matematika maupun perkiraan yang subjektif disebut sebagai peramalan [16].

2.2.1 Jenis-Jenis Metode Peramalan

Terdapat 2 pendekatan umum untuk jenis metode peramalan yaitu kualitatif dan kuantitatif. Metode peramalan kualitatif sangat penting dimana pada saat data historis tidak ada, tetapi metode ini bersifat sangat subjektif dan membutuhkan penilaian dari para ahli. Di lain pihak peramalan kuantitatif menggunakan data historis yang ada. Tujuan metode ini adalah mempelajari apa yang telah terjadi dimasa lalu untuk dapat meramalkan nilai- nilai dimasa yang akan datang [17].

a. Metode Peramalan Deret Berkala

Metode peramalan deret berkala, atau yang biasa disebut sebagai deret waktu (*time series*), merupakan salah satu metode yang termasuk dalam metode peramalan kuantitatif selain metode regresi atau kausal.

Metode peramalan deret berkala melibatkan proyeksi nilai yang akan datang dari sebuah variabel dengan berdasarkan seluruhnya pada pengamatan masa lalu dan sekarang dari variabel tersebut [17].

b. Metode Pemulusan

Metode pemulusan atau biasa disebut metode *Smoothing*, termasuk dalam metode peramalan deret berkala. Metode pemulusan memiliki dasar metode yaitu pembobotan sederhana atau pemulusan pengamatan masa lalu dalam suatu sumber deret berkala untuk memperoleh ramalan masa mendatang [18]. Dalam pemulusan nilai-nilai historis ini, galat acak dirata-ratakan untuk menghasilkan ramalan “halus”. Diantara keuntungannya yaitu biaya yang rendah, mudah digunakan dalam penerapannya, dan cepat dalam penyampaiannya. Karakteristik ini dapat membuatnya menarik terutama bilamana horison waktunya relatif pendek (kurang dari 1 tahun).

Metode pemulusan terdiri atas metode pemulusan perataan, dimana pada saat melakukan pembobotan yang sama terhadap nilai-nilai pengamayan sesuai dengan pengertian konvensional tentang nilai tengah, dan metode pemulusan eksponensial menggunakan bobot berbeda untuk data masa lalu, karena bobotnya berciri menurun seperti eksponensial dari titik data yang terakhir sampai terawal.

2.2.2 Jenis-Jenis Pola Data

Langkah penting dalam memilih suatu metode deret berkala (*time series*) yang tepat adalah dengan mempertimbangkan jenis pola data, sehingga metode yang paling tepat dengan pola tersebut dapat diuji.

Pola data dapat dibedakan menjadi empat jenis [15], yaitu:

1. Pola Horizontal atau *Horizontal Data Pattern*

Pola data ini terjadi bilamana data berfluktuasi di sekitar nilai rata-rata. Suatu produk yang penjualannya tidak meningkat atau menurun selama waktu tertentu termasuk jenis ini.

2. Pola Trend atau *Trend Data Pattern*

Pola data ini terjadi bilamana terdapat kenaikan atau penurunan sekuler jangka panjang dalam data. Contohnya penjualan perusahaan, GNP (*Gross National Product*) dan berbagai indikator bisnis atau ekonomi lainnya, selama perubahan sepanjang waktu.

3. Pola Musiman atau *Seasonal Data Pattern*

Pola data ini terjadi bilamana suatu deret dipengaruhi oleh faktor musiman (misalnya kuartal tahun tertentu, bulan atau hari-hari pada minggu tertentu).

Penjualan dari produk seperti minuman ringan, es krim dan bahan bakar pemanas ruang semuanya menunjukkan jenis pola ini.

4. Pola Siklis atau *Cyclied Data Pattern*

Pola data ini terjadi bilamana datanya dipengaruhi oleh fluktuasi ekonomi jangka panjang seperti yang berhubungan dengan siklus bisnis. Contohnya penjualan produk seperti mobil, baja.

2.2.3 Jangka Waktu Peramalan

Jangka waktu peramalan dapat dikelompokkan menjadi tiga kategori [16], yaitu:

1. Peramalan jangka pendek, peramalan untuk jangka waktu kurang dari tiga bulan.
2. Peramalan jangka menengah, peramalan untuk jangka waktu antara tiga bulan sampai tiga tahun.
3. Peramalan jangka panjang, peramalan untuk jangka waktu lebih dari tiga tahun.

2.2.4 Proses Peramalan

Didalam melakukan proses peramalan, apapun bentuk dan jenis peramalan yang akan dilakukan, terdapat lima langkah proses peramalan yang bisa dilakukan [20], yaitu:

1. Formulasi masalah dan pengumpulan data.

Jika metode peramalan kuantitatif yang dipakai maka data yang relevan harus tersedia dan benar. Jika data yang sesuai tidak tersedia maka mungkin perumusan

masalah perlu dikaji ulang atau memeriksa kembali metode peramalan kuantitatif yang dipakai.

2. Manipulasi dan pembersihan data

Ada kemungkinan kita memiliki terlalu banyak atau terlalu sedikit data yang dibutuhkan. Sebagian data mungkin tidak relevan pada masalah. Sebagian data mungkin memiliki nilai yang hilang yang harus diestimasi. Sebagian data mungkin harus dihitung dalam unit selain unit aslinya. Sebagian data mungkin harus diproses terlebih dahulu (misal, dijumlahkan dari berbagai sumber). Data yang lain kemungkinan sesuai tetapi hanya pada periode historis tertentu. Biasanya perlu usaha untuk mengambil data dalam suatu bentuk yang di butuhkan untuk menggunakan prosedur peramalan tertentu.

3. Pembentukan dan evaluasi model

Pembentukan dan evaluasi model menyangkut pengepasan data yang terkumpul pada suatu model peramalan yang sesuai dengan meminimalkan galat peramalan.

4. Implementasi model (peramalan sebenarnya)

Implementasi model terdiri dari model peramalan aktual yang dibuat ketika data yang sesuai telah terkumpul dan terpilihnya model peramalan yang sesuai. Peramalan untuk periode sekarang dengan nilai historis aktual diketahui sering kali digunakan untuk mengecek keakuratan dari proses.

5. Evaluasi peramalan

Implementasi model terdiri dari model peramalan aktual yang dibuat ketika data yang sesuai telah terkumpul dan terpilihnya model peramalan yang sesuai.

Peramalan untuk periode sekarang dengan nilai historis aktual diketahui sering kali digunakan untuk mengecek keakuratan dari proses.

2.2.5 Karakteristik Peramalan

Karakteristik dari peramalan yang baik harus memenuhi beberapa kriteria yaitu dari hal-hal sebagai berikut:

1. Ketelitian/ Keakuratan

Tujuan utama peramalan adalah menghasilkan prediksi yang akurat. Peramalan yang terlalu rendah mengakibatkan kekurangan persediaan (inventory). Peramalan yang terlalu tinggi akan menyebabkan inventory yang berlebihan dan biaya operasi tambahan.

2. Biaya

Biaya untuk mengembangkan model peramalan dan melakukan peramalan akan menjadi signifikan jika jumlah produk dan data lainnya semakin besar. Mengusahakan melakukan peramalan jangan sampai menimbulkan ongkos yang terlalu besar ataupun terlalu kecil. Keakuratan peramalan dapat ditingkatkan dengan mengembangkan model lebih kompleks dengan konsekuensi biaya menjadi lebih mahal. Jadi ada nilai tukar antara biaya dan keakuratan.

3. Responsif

Ramalan harus stabil dan tidak terpengaruhi oleh fluktuasi *demand*.

4. Sederhana

Keuntungan utama menggunakan peramalan yang sederhana yaitu kemudahan untuk melakukan peramalan. Jika kesulitan terjadi pada metode sederhana, diagnosa

dilakukan lebih mudah. Secara umum, lebih baik menggunakan metode paling sederhana yang sesuai dengan kebutuhan peramalan.

2.3 Exponential Smoothing Models

Terlepas dari kesederhanaannya, model pemulusan eksponensial adalah salah satu metode prediksi yang paling fleksibel dan akurat untuk analisis deret waktu. Metode ini telah terbukti sangat efektif dalam meramalkan fenomena ekonomi.

Meskipun model-model ini awalnya dirumuskan berdasarkan empiris dan intuitif, kemajuan terbaru dalam penelitian telah menempatkan mereka pada pijakan teoritis yang lebih baik. Pada bagian ini kami akan meninjau model *Smoothing* dari peningkatan kompleksitas, yang bertujuan untuk mengidentifikasi berbagai komponen dari rangkaian waktu, bahkan dalam bentuk *self-correction* dan adaptif [1].

2.3.1 Simple Exponential Smoothing

Model *simple Exponential Smoothing*, juga disebut model *Brown*, mengandalkan estimator s_t dari rata-rata pengamatan dari deret waktu hingga waktu t , disebut *smoothed mean* (rerata yang dihaluskan) dan didefinisikan melalui ekspresi rekursif

$$s_t = \alpha y_t + (1 - \alpha)s_{t-1}, \dots(2.1)$$

$$s_1 = y_1, \dots(2.2)$$

dimana $\alpha \in [0, 1]$ adalah parameter yang mengatur pentingnya nilai terbaru y_t sehubungan dengan rerata yang dihaluskan dari nilai sebelumnya. Prediksi untuk periode $t + 1$ diperoleh dengan menetapkan

$$f_{t+1} = s_t \dots(2.3)$$

Dapat dengan mudah diverifikasi bahwa persamaan berikut ini berlaku untuk metode pemulusan eksponensial:

$$\left. \begin{aligned} s_1 &= y_1, \\ s_2 &= \alpha y_2 + (1 - \alpha)s_1, \\ s_3 &= \alpha y_3 + (1 - \alpha)s_2, \\ &\dots, \\ s_t &= \alpha y_t + (1 - \alpha)s_{t-1}. \end{aligned} \right\} \dots(2.4)$$

Dengan menerapkan rekursif dari persamaan diatas, salah satunya dapat diperoleh hubungan antara peramalan untuk periode berikutnya dan pengamatan di masa lalu.

$$f_{t+1} = \alpha[y_t + (1 - \alpha)y_{t-1} + \dots + (1 - \alpha)^{t-2} y_2] + (1 - \alpha)^{t-1}y_1 \dots(2.5)$$

Persamaan diatas menunjukkan bahwa untuk model pemulusan eksponensial, fungsi F menyatakan prediksi sebagai kombinasi cembung linear dari pengamatan deret waktu, dengan bobot yang menurun secara eksponensial saat kita kembali ke masa. Persamaan ini juga memungkinkan kita untuk mendapatkan interpretasi untuk parameter α : jika α 0 model memiliki inersia yang lebih besar, dalam arti bahwa ia memberikan bobot yang hampir konstan untuk semua pengamatan sebelumnya; jika α 1 model lebih responsif, karena memberikan bobot yang lebih besar untuk pengamatan terbaru. Parameter α dipilih sedemikian rupa untuk meminimalkan kesalahan kuadrat rata-rata (*mean square error*).

2.3.2 Exponential Smoothing dengan penyesuaian tren

Menghadapi *time series* yang mencakup komponen tren, model pemulusan eksponensial sederhana pasti gagal menangkap tren. Akibatnya, ketika diterapkan pada *time series* non-stasioner dengan komponen tren, model sederhana selalu

tertinggal dari pengamatan aktual dan menghasilkan prediksi yang terdistorsi oleh kelebihan atau cacat. Namun, dimungkinkan untuk memperluas model sederhana untuk memasukkan komponen tren, yaitu melalui model *Holt Exponential Smoothing*.

Untuk tujuan ini, selain *smoothed mean* s_t , tren *smoothed* linier m_t juga didefinisikan, yang merupakan estimasi dari komponen tren aditif M_t . Dimungkinkan juga untuk menentukan penyesuaian tren yang lebih kompleks, misalnya dengan menggunakan fungsi kuadratik atau eksponensial. Ekspresi rekursif yang mendefinisikan *smoothed mean* harus disesuaikan untuk memperhitungkan *smoothed* tren.

$$s_t = \alpha y_t + (1 - \alpha) (s_{t-1} + m_{t-1}) \dots (2.6)$$

Ekspresi rekursif yang sesuai mengatur pembaruan *smoothed* tren

$$m_t = \beta (s_t - s_{t-1}) + (1 - \beta) m_{t-1} \dots (2.7)$$

dimana $\beta \in [0, 1]$ merupakan parameter yang memodulasi pentingnya nilai tren terbaru, dinyatakan sebagai perbedaan antara *smoothed mean* ($s_t - s_{t-1}$), berkenaan dengan *smoothed trend* dari periode sebelumnya, dinyatakan dengan nilai iterasi m_{t-1} . Jika $\beta \cong 0$ model memberikan bobot yang hampir sama untuk tren yang ditunjukkan di masa lalu, sedangkan jika $\beta \cong 1$ tren yang paling baru ditunjukkan adalah dominan. Perhatikan bahwa perbedaan antara *smoothed means* ($s_t - s_{t-1}$) adalah perkiraan tren terbaru yang ternyata jauh lebih kokoh dari perbedaan ($y_t - y_{t-1}$) antara nilai yang sesuai dari deret waktu, karena yang terakhir ini mungkin dipengaruhi oleh fluktuasi musiman dan acak.

Prediksi untuk periode $t + 1$ akhirnya diperoleh dengan menetapkan

$$f_{t+1} = s_t + m_t \dots (2.8)$$

Di sini juga pilihan parameter α dan β harus dibuat sedemikian rupa untuk meminimalkan jumlah kesalahan kuadrat (*sum of squared error*).

2.3.3 *Exponential Smoothing* dengan tren dan musiman

Jika *time series* mencakup komponen musiman, maka model *Exponential Smoothing* perlu dikembangkan. Selain *smoothed mean* dan tren, model *Winters* mendefinisikan *smoothed index* musiman q_t , untuk mendekati komponen musiman multiplikasi Q_t .

Dengan asumsi bahwa setiap siklus musiman terdiri dari periode L , pembaruan *smoothed mean* dan tren dimodifikasi, dan pembaruan rekursif juga diperkenalkan untuk *smoothed index* musiman q_t , berdasarkan

$$s_t = \alpha \frac{y_t}{q_{t-L}} + (1 - \alpha) (s_{t-1} + m_{t-1}), \dots (2.9)$$

$$m_t = \beta (s_t - s_{t-1}) + (1 - \beta) m_{t-1}, \dots (2.10)$$

$$q_t = \gamma \frac{y_t}{s_t} + (1 - \gamma) q_{t-L} \dots (2.11)$$

Di sini $\gamma \in [0, 1]$ adalah parameter yang memodulasi pentingnya nilai terkini dari musiman, dinyatakan sebagai rasio antara nilai terbaru y_t dari *time series* dan *smoothed mean* s_t , berkenaan dengan *smoothed index* musiman q_{t-L} dari periode sebelumnya homolog ke t . Adapun α dan β , jika $\gamma \cong 0$ model memberikan bobot yang hampir konstan untuk musiman yang terjadi di masa lalu, sedangkan jika $\gamma \cong 1$ musiman yang paling baru adalah yang dominan.

Akhirnya, prediksi untuk periode $t + 1$ didefinisikan sebagai

$$f_{t+1} = (s_t + m_t) q_{t-L+1}. \dots(2.12)$$

Pilihan parameter α , β dan γ harus dibuat sedemikian rupa untuk meminimalkan jumlah kesalahan kuadrat (*sum of squared error*).

2.3.4 Menghilangkan Tren dan Musiman

Dengan non-stasioner *time series* generik $\{Y_t\}$, beberapa teknik dapat diterapkan untuk mendapatkan *time series* stasioner yang diubah $\{B_t\}$, yaitu seri di mana distribusi $\{B_t, B_{t+1} \dots, B_{t+h}\}$ sama dengan $\{B_{t+k}, B_{t+k+1} \dots, B_{t+k+h}\}$ untuk setiap pilihan yang mungkin dari t , k dan h . Dalam hipotesis ini, *time series* $\{B_t\}$ memiliki komponen tren nol, diwakili oleh garis horizontal.

Moving Average (Rata-rata bergerak). Di satu sisi, dimungkinkan untuk membagi nilai *time series* $\{Y_t\}$ dengan nilai yang sesuai dari rata-rata bergerak untuk mendapatkan *time series* $\{B_t\}$ tanpa tren.

Trend (Kecenderungan). Di sisi lain, adalah mungkin untuk mengidentifikasi komponen tren M_t dari deret waktu melalui dekomposisi aditif atau multiplikasi, dan oleh karenanya untuk menghilangkannya masing-masing dengan pengurangan atau pembagian.

Differencing. Akhirnya, perbedaan antara nilai-nilai yang berdekatan dari *time series* dapat ditemukan, pengaturan

$$B_t = \nabla Y_t = Y_t - Y_{t-1},$$

$$\begin{aligned}
 B_t(2) &= \nabla^2 Y_t = Y_t - Y_{t-2}, \\
 &\dots, \\
 B_t(h) &= \nabla^h Y_t = Y_t - Y_{t-h} \dots (2.13)
 \end{aligned}$$

Differences (Perbedaan) urutan pertama atau paling banyak dari urutan kedua biasanya cukup untuk menghapus komponen tren.

Sebagai konsekuensi, diberi *time series* $\{Y_t\}$ dengan kedua tren dan komponen musiman, setidaknya ada tiga metode alternatif untuk mengembangkan model peramalan:

- menerapkan model *Winters* ke *time series* asli;
- menerapkan model *Holt* ke *time series* setelah data di de-seasonalisasi;
- menerapkan model sederhana pada *time series* yang diperoleh dengan mende-seasonalisasi dan menghapus komponen tren

2.3.5 *Autocorrelation Function (ACF)* dan *Partial Autocorrelation Function (PACF)*

Autocorrelation merupakan korelasi antar data di dalam *time series* tersebut yang dipisahkan dalam *lag*. Artinya, jika *lag* =2, maka nilai *t* akan dikorelasikan dengan *t+2* (artinya data pertama akan dikorelasikan dengan data ketiga), jika menggunakan *lag* = 4, maka kemudian *t* akan dikorelasikan dengan *t+4*, begitu seterusnya. Hal ini biasanya ditulis dengan rumus korelasi (Y_t, Y_{t+h}) dengan *k* adalah jumlah lag dan *t* adalah nilainya [40].

$$ACF_h = \text{corr}(Y_t, Y_{t-h}) \dots (2.14)$$

Oleh sebab itu, jika data akan mengandung tren, maka hal ini tidak akan berguna. Karena nilai ACF nya akan signifikan atau tinggi di setiap data atau *lag*nya. Itu lah sebabnya diperlukan *differencing* untuk menghilangkan tren pada data tersebut hingga menjadi data stationer.

Sama halnya *autocorrelation*, PACF juga menghitung korelasi antar data di *time series*. Perbedaannya adalah jika ACF mencari korelasi antara Y_t dan Y_{h+t} ($h =$ lag), maka PACF juga menghitung korelasi antara Y_t dan Y_{h+t} namun sebelumnya menghilangkan data antara Y_t dan Y_{h+t} terlebih dahulu. Berarti data yang dihilangkan adalah data Y_{t+1} sampai dengan Y_{t+h+1}

$$\text{PACF}_h = \text{corr}(Y_t, Y_{t-h}/Y_{t-1}, Y_{t-2}, \dots, Y_{t-h+1}) \dots (2.15)$$

PACF menghitung korelasi secara parsial antara nilai t dengan nilai $t+h$ tanpa dipengaruhi nilai nilai diantaranya. Sedangkan ACF menghitung yang sama (t dan $t+h$) namun masih memperhitungkan nilai nilai yang ada diantaranya secara keseluruhan data.

2.3.6 Hubungan Akurasi Peramalan

Menurut Jay Heinzer dan Barry Render (2009) ada beberapa perhitungan yang biasa dipergunakan untuk menghitung kesalahan peramalan (*forecast error*) total. Perhitungan ini dapat dipergunakan untuk membandingkan model peramalan yang berbeda, juga untuk mengawasi peramalan, untuk memastikan peramalan berjalan dengan baik. Model-model peramalan yang dilakukan kemudian divalidasi menggunakan sejumlah indikator. Indikator-indikator yang umum digunakan adalah

rata-rata penyimpangan absolut (*Mean Absolute Deviation*), dan rata-rata kuadrat terkecil (*Mean Square Error*). [11]

Mean Squared Error (MSE) adalah metode lain untuk mengevaluasi metode peramalan. Masing-masing kesalahan atau sisa dikuadratkan. Kemudian dijumlahkan dan dibagi dengan jumlah observasi. Pendekatan ini mengatur kesalahan peramalan yang besar karena kesalahan-kesalahan itu dikuadratkan. Metode itu menghasilkan kesalahan-kesalahan sedang yang kemungkinan lebih baik untuk kesalahan kecil, tetapi kadang menghasilkan perbedaan yang besar. *Mean squared error* adalah rata-rata dari kesalahan *forecast* dikuadratkan, atau jika dituliskan dalam bentuk rumus adalah: [11]

$$MSE = \frac{\sum_{t=1}^n (d_t - d'_t)^2}{n} \dots\dots\dots(2.16)$$

Untuk menghitung nilai persen error pada setiap observasi y_t yang berhubungan dengan hasil peramalan f_t dari periode $1, 2, \dots, k$ yaitu dengan menggunakan *Percentage Prediction Error* (PPE), didefinisikan sebagai berikut:

$$e_t^P = \frac{y_t - f_t}{y_t} \times 100. \dots\dots(2.17)$$

Dari error yang dihasilkan untuk masing-masing observasi, dilakukan perhitungan menggunakan *Mean Percentage Error* (MPE) untuk menghitung nilai rata-rata dari keseluruhan observasi, didefinisikan sebagai berikut:

$$MPE = \frac{\sum_{t=1}^k e_t^P}{k}. \dots\dots(2.17)$$

2.4 R Studio

R adalah suatu kesatuan *software* yang terintegrasi dengan beberapa fasilitas untuk manipulasi, perhitungan dan penampilan grafik yang handal. R berbasis pada bahasa pemrograman S, yang dikembangkan oleh AT&T Bell Laboratories sekarang Lucent Technologies pada akhir tahun '70 an. R merupakan versi gratis dari bahasa S dari *software* (berbayar) yang sejenis yakni S-PLUS yang banyak digunakan para peneliti dan akademisi dalam melakukan kegiatan ilmiahnya.

Pada awalnya, versi pertama R dibuat oleh Ross Ihaka and Robert Gentleman dari Universitas Auckland, namun selanjutnya R dikembangkan oleh tim yang disebut tim inti. Tim inti (*core team*) terdiri dari ahli statistik, ahli komputer & pemrograman, geografi, ekonomi dari institusi yang berbeda dari seluruh dunia yang mencoba membangun sebuah sistem (*software*) yang handal namun dengan biaya yang sangat murah. R dapat secara gratis didownload dan digunakan dengan berlisensi pada GNU (*General Public License*).

Menurut kutipan dari penghargaan Association for Computing Machinery Software bagi John Chamber 1998, menyatakan bahwa bahasa pemrograman dapat “memanipulasi, visualisasi dan menganalisis data”. R dibuat searah dengan ide yang ada pada bahasa pemrograman S.

2.4.1 R dan Program Statistik Lainnya

Seperti dijelaskan sebelumnya, R merupakan “kerabat” dekat dari S-PLUS dimana secara fungsi dan sintaks/tata bahasa sama-sama menggunakan bahasa S, namun tidak identik. R dapat berinteraksi dengan program statisik, manipulasi,

perhitungan dan penampilan grafik lainnya, seperti SPSS, yang cukup populer, Microsoft Excel dengan menyediakan fasilitas import dan ekspor data.

Selain software di atas, R dapat melakukan import file dari software lainnya seperti, Minitab, SAS, Stat, Systat dan EpInfo. R adalah bahasa fungsional¹, dimana terdapat inti bahasa yang menggunakan bentuk standar notasi aljabar, yang memungkinkan perhitungan numerik seperti $2+3$, atau 3^11 . Selain itu tersedia pula fasilitas perhitungan dengan menggunakan fungsi.

Dengan beberapa fitur tersebut, R menjadi alat yang tangguh bagi para statistikawan, ahli ekonomi, peneliti dalam membantu risetnya, dikarenakan R dibangun dan didukung dengan model dan teori statistik terdepan dan menggunakan standar tertinggi bagi analisis data. R hampir dapat digunakan untuk berbagai bidang, mulai dari kalkulasi biasa seperti kalkulator, statistik, ekonometri, geografi, hingga pemrograman komputer.

2.4.2 Kelebihan dan Fitur-fitur R

R mempunyai karakteristik tersendiri, dimana selalu dimulai dengan prompt “>” pada *console*-nya. R mempunyai beberapa kelebihan dan fitur-fitur yang canggih dan berguna, diantaranya:

- a) Efektif dalam pengelolaan data dan fasilitas penyimpanan. Ukuran file yang disimpan jauh lebih kecil dibanding software lainnya.
- b) Lengkap dalam operator perhitungan array

- c) Lengkap dan terdiri dari koleksi *tools* statistik yang terintegrasi untuk analisis data, diantaranya, mulai statistik deskriptif, fungsi probabilitas, berbagai macam uji statistik, hingga *time series*.
- d) Tampilan grafik yang menarik dan fleksibel ataupun *costumized*.

Dapat dikembangkan sesuai keperluan dan kebutuhan dan sifatnya yang terbuka, setiap orang dapat menambahkan fitur-fitur tambahan dalam bentuk paket ke dalam software R.

Selain kelebihan dan kelengkapan fitur-fiturnya, hal yang terpenting lainnya yakni, R bersifat multiplatform, yakni dapat diinstall dan digunakan baik pada sistem operasi Windows, UNIX/LINUX maupun pada Macintosh. Untuk dua sistem operasi disebutkan terakhir diperlukan sedikit penyesuaian. Selain kelebihan disebutkan di atas, R didukung oleh komunitas yang secara aktif saling berinteraksi satu sama lain melalui Internet dan didukung oleh manual atau R- *help* yang menyatu pada *software* R. Sebagai catatan, buku ini mengambil contoh pada penggunaan R pada sistem berbasis Windows.

2.5 Penelitian Sebelumnya

Gidion Aryo Nugraha Pongdatu melakukan penelitian peramalan dengan menggunakan metode SARIMA dan *HoltWinters Exponential Smoothing* untuk penjualan baju dan sepatu. Dalam bisnis pakaian, diketahui bahwa pilihan calon konsumen secara umum akan pakaian cenderung ditentukan oleh harga pakaian. Dalam kasus ini, kecenderungan pelanggan dalam pemilihan berdasarkan harga memicu pihak toko untuk melakukan penyesuaian harga dengan cara membeli stok

langsung dari pihak produsen dan mengurangi biaya penyimpanan. Pihak toko xyz sendiri saat ini mengalami masalah dalam penyimpanan barang dikarenakan barang yang disimpan di gudang mengalami penumpukan sehingga untuk mendatangkan barang baru dibutuhkan tempat penyimpanan baru sehingga dibutuhkan biaya tambahan. Untuk mengatasi hal ini perlu dilakukan perhitungan peramalan yang lebih akurat agar dapat diprediksi jumlah item yang perlu disediakan sehingga tidak terjadi penumpukan di gudang. Untuk kasus peramalan pelanggan di toko xyz ini akan digunakan metode peramalan deret waktu dengan pola data musiman sesuai dengan karakteristik data penjualan toko xyz. Untuk itu akan digunakan metode *Seasonal Autoregressive Integrated Moving Average (SARIMA)* dan *Holt-Winter's Exponential Smoothing*.

Fajar Sidqi melakukan penelitian peramalan dengan menggunakan metode *Single Exponential Smoothing* dan *Double Exponential Smoothing* untuk meramalkan penjualan produk. Hasil penelitian menunjukkan bahwa MAPE dari *Single Exponential Smoothing* sebesar 20% dan MAPE dari *Double Exponential Smoothing* sekitar 24%. Penggunaan metode Pemulusan Eksponensial Tunggal memiliki kesalahan yang lebih rendah. Hasil penelitian menunjukkan bahwa peramalan dengan pemulusan eksponensial tunggal direkomendasikan untuk digunakan oleh XYZ store dalam kasus ini. Oleh karena itu, data yang diolah dengan metode-metode tersebut akan sangat bermanfaat bagi toko XYZ dalam menentukan stok produknya di masa yang akan datang.

I A Zahra melakukan penelitian peramalan dengan menggunakan metode ARIMA dan *Single Exponential Smoothing* untuk meramalkan kebutuhan obat.

Tujuan dari penelitian ini adalah untuk memprediksi kebutuhan obat dengan menggunakan teknik peramalan dan menghitung nilai *Economic Order Quantities*. Fluktuasi penggunaan obat yang terjadi setiap tahun menjadi kendala bagi gudang obat dalam perencanaan pengadaan di rumah sakit. Hasil penelitian ini berupa nilai taksiran kebutuhan obat untuk satu periode mendatang ditunjukkan dengan nilai kesalahan peramalan terkecil yaitu ARIMA (1.0.0) dengan nilai kesalahan 13%, dan hasil perhitungan *Quantity of Economic Order* untuk kebutuhan obat masa depan. Hasil peramalan antara ARIMA dengan metode *Exponential Smoothing* menunjukkan bahwa peramalan memiliki nilai error yang paling kecil dengan menggunakan ARIMA (1.0.0).

Berikut ini beberapa Penelitian sebelumnya yang telah dilakukan:

Tabel 2.2 Daftar Penelitian Terdahulu

No.	Tahun Penelitian	Judul	Persamaan	Perbedaan
1.	2018	Peramalan Pola Transaksi Pelanggan Toko Xyz dengan Metode Sarima dan <i>Holt-Winters Exponential Smoothing</i>	Melakukan peramalan menggunakan metode <i>HoltWinters Exponential Smoothing</i>	<ol style="list-style-type: none"> 1. Objek penelitian berbeda 2. Hasil <i>output</i> yang dilakukan berbeda 3. <i>Smoothing parameter</i> yang digunakan berbeda. Pada penelitian ini parameter alpha, beta, dan gamma-nya menggunakan nilai yang sama. 4. Menggunakan 2 metode yang berbeda untuk proses peramalannya.

2.	2019	<i>Forecasting Product Selling using Single Exponential Smoothing and Double Exponential Smoothing Method</i>	Melakukan peramalan menggunakan metode <i>Exponential Smoothing</i>	<ol style="list-style-type: none"> 1. Objek penelitian berbeda 2. Hasil <i>output</i> yang dilakukan berbeda 3. Data variabel dalam melakukan peramalan berbeda
3.	2019	<i>Forecasting Methods Comparison Based on Seasonal Patterns for Predicting Medicine Needs with ARIMA Method, Single Exponential Smoothing</i>	Melakukan peramalan menggunakan metode <i>Exponential Smoothing</i>	<ol style="list-style-type: none"> 1. Objek penelitian berbeda 2. Hasil <i>output</i> yang dilakukan berbeda 3. Parameter alpha ditentukan dengan mencari nilai MAD terkecil yang dihitung secara manual