Rancang bangun data warehouse dan r studio serta pemanfaatannya dalam peramalan pola konsumsi masyarakat di Kabupaten Jember

by Lutfi Muharom

Submission date: 01-Aug-2019 10:03AM (UTC+0700)
Submission ID: 1156660014
File name: lutfi_jurnal_warehouse.pdf (604.9K)
Word count: 3943
Character count: 22622

Rancang Bangun Data *Warehouse* dan R Studio Serta Pemanfaatanya dalam Peramalan Pola Konsumsi Masyarakat di Kabupaten Jember

Lutfi Ali Muharom[1)]
[1)]*Prodi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember*
Email: [1)]lutfi.muharom@unmuhjember.ac.id

Alfian Futuhul Hadi[2)], Dian Anggraeni[3)]
[2,3)]*Jurusan Matematika, Fakultas MIPA, Universitas Jember*

Abstract

As we know that we have to process and store the data recording well. Data warehouse is one of data processing method that use to support the decission-making process. The data warehouse process started from colecting, selecting, designing and uploading data in to data warehouse. In this research, we use the data of SUSENAS from year of 1997 until 2012. We took the daily consumption data (household expendature) to be proceed in data warehouse. The implementation of web based R studio program can facilitate the users to acces R . R can be accessed by any kind of devices which have browser and internet acces by any kind of devices which have browse and internet acces. The connectivity of R studio to data warehouse can be simplify the users to access and process the data. As the result of consumption patterns (staple food) forecasting in jember, we conclude that the best forecasting method for forecasting method for forecasting using AR(1) model. The limited data collections caused the ensemble wouldn't become the best method , whereas, it should be the best method.

Keywords: *Data Warehouse, R program, ARIMA, ARIMA Ensemble.*

1. PENDAHULUAN

BPS melakukan pengumpulan data, menerbitkan publikasi statistik nasional maupun daerah ,serta melakukan analisis data statistik yang digunkan dalam pengambilan keputusan .sejak polapembangunan yang diarahkan pada otonomi daerah,dibutuhkan perhitungan statisik sampai dengan skala kecamatan bahkan desa untuk membantu pemerintah daerah dalam upaya mendongkrak pembangunan daerah.

Sejarah telah menunjukkan bahwa ada korelasi yang amat kuat antara kemampuan IPTEK suatu negara dengan tingkat pendapatannya per kapita.makin tinggi kemampuan IPTEK suatu negara makin tinggi pula tingkat pendapatan perkapitannya. Perlu dicatat bahwa negara-negara kaya hanya memiliki 20% dari penduduk bumi,tetapi sangat tinggi kemampuan IPTEK-nya,ternyata menguasai 80% dari pendapatan global (Kementan, 2012).

Perguruan tinggi sebagai salah satu pusat riset yang didalamnya terhimpun sarana dan prasarana IPTEK yang cukup mutakhir dan tersedia cukup sumber daya manusia ber kualitas tinggi.salah satu hambatan didalam penelitian / riset adalah masalah ketersediaan data, salah satu contoh adalah data sensus dan survei. Salah satu langkah yang perlu diambil adalah dengan membangun suatu data warehouse tersendiri yang berfungsi untuk menampung data dan angka yang dihasilkan oleh sensus dan survei. Kemudahan dalam mengakses data ini akan menjadikan penelitian terutama dalam bidang statistik akan semakin berkembang. Diharapkan hasil-hasil penelitian ini menjadi salah satu acuan pemerintah

maupun swasta didalam mengambil kebijakan sehingga memberikan dampak baik langsung maupun tidak langsung pada pertumbuhan ekonomi.

R adalah salah satu dari program sumber terbuka yang dapat diunduh gratis. R diluncurkan pertama kali tahun 1997, dan versi terakhir ketika tulisan diluncurkan ini dibuat tahun 1997, dan versi terakhir ketika tulisan ini dibuat adalah 3.1.0. R studio adalah salah satu GUI untuk R, salah satu keunggulan pada R studio ini adalah dapat dijalankan pada browser sehingga dijalankan diatas browser, maka pengguna tidak memerlukan lagi installasi R, kecuali paket (*package*) pemrograman sesuai dengan kebutuhan pengguna.

2. TINJAUAN PUSTAKA

2.1 SUSENAS

BPS melakukan pengumpulan data melalui kegiatan sensus dan survei (sampling). Survei Sosial Ekonomi Nasional (SUSENAS) adalah survei yang dirancang untuk mengumpulkan data sosial kependudukan yang relatif sangat luas dan dilaksanakan setiap tahun.data yang dikumpulkan antara lain menyangkut bidang bidang pendidikan, kesehatan/gizi, perumahan, sosial ekonomi lainnya, kegiatan sosial budaya, konsumsi/ pengeluaran dan pendapatan rumah tangga, perjalanan, dan pendapatan masyarakat mengenai kesejahteraan rumah tangganya.

SUSENAS sudah dilaksanakan sejak tahun 1963 dengan kuesioner yang dibagi dalam dua bagian besar yaitu Kor dan Modul. Kor adalah kumpulan pertanyaan-pertanyaan yang bersifat umum yang selalu ditanyakan setiap tahun, sedangkan Modul adalah pertanyaan-pertanyaan khusus sesuai dengan topiknya dan lebih rinci yang dilakukan tiga tahun sekali untuk setiap topik.

Pelaksanaan lapangan SUSENAS awalnya adalah pada bulan Januari Februari untuk setiap tahunnya. Namun sejak bergesernya tahun anggaran dari April- Maret menjadi Januari-Desember maka pelaksanaan SUSENAS pada bulan Januari menjadi terkendala, maka sejak 2004 pelaksanaan SUSENAS digeser menjadi bulan Juni-Juli untuk setiap tahunnya.

2.2 Konsep Data *Warehouse*

Data *warehouse* adalah database yang saling bereaksi yang dapat digunakan untuk *query* dan analisis, bersifat orientasi subjek, terintegrasi, time variant, tidak berubah yang digunakan untuk membuat para pengambil keputusan. Adapun karakteristik dari warehouse menurut Inmon yaitu:

1. Berorientasi subjek (*subject oriented*)

   Data *warehouse* berorientasi subjek artinya data *warehouse* didesain untuk analisa data berdasarkan subjek tertentu dalam organisasi, bukan pada proses atau fungsi aplikasi tertentu.

2. Terintegrasi (*integrated*)

   Data *warehouse* dapat menyimpan data-data yang berasal dari sumber-sumber yang terpisah kedalam suatu format yang konsisten dan saling terintegrasi satu dengan yang lainnya. Dengandemikian data tidak dapat dipecah-pecah karena data yang ada merupakan suatu kesatuan yang menunjang keseluruhan kensep data *warehouse* itu sendiri.

3. Rentang Waktu (*time-variant*)

   Seluruh data pada data warehouse dapat dikatakan akurat atau valid pda waktu tertentu. Untuk melihat interval waktu yang digunakan dalam mengukur keakuratan data *warehouse*.

4. *Non-Volatile*

   Karakteristik keempat dari data *warehouse* adalah non-volatile, maksudnya data pada data *warehouse* tidak diupdate secara *real time* tetapi di refresh dari sistem operasional secara reguler. Data yang baru selalu ditambahkan sebagai sebuah perubahan. Database tersebut secara kontinyu menyerap data baru ini, kemudian secara incremental disatukan dengan data sebelumnya.

2.3 Model ARIMA

Pada model ARIMA (p,d,q) nilai yang akan datang dari suatu variabel diasumsikan sebagai fungsi linier dari suatu variabel diasumsikan sebagai fungsi linier dari beberapa pengamatan

di masa lalu dan *random error* (Khashei et al, 2007). Wei (2009) menguraikan model ARIMA (p,d,q) secara umum dalam yaitu :

$$\emptyset_p(B)(1-)^d Z_t = \theta_p + \theta_q(B)a_t$$

dengan

$$\emptyset_p(B) = \left(1 - \emptyset_1 B - \cdots - \emptyset_y B^p\right), \theta_q(B) = (1 - \theta_1 B - \cdots - \theta_1 B^q)$$

dan $\emptyset_0$ adalah *intercept* pada model untuk *difference* ke-*d*.

Pembuatan model ARIMA dapat dilakukan dengan menggunakan tiga prosedur yaitu identifikasi modal, estimasi modal dan cek diagnosa (Jenkins 2006). Identifikasi modal merupakan metodologi untuk mengidentifikasi perlunya suatu tranformasi seperti transformasi untuk stasioner dalam varians transformasi differencing, keputusan untuk memasukan parameter $\emptyset_0$ ketika d>0 dan penentuan order p dan q pada ARIMA (Wei, 2009).

2.4 Model ARIMA *Ensemble*

ARIMA *ensemble* merupakan model ARIMA. Pembentukan ARIMA *ensemble* terdiri dari dua langkah. Pertama menciptakan anggota enseble dari beberapa modal ARIMA selanjutnya menggabungkan hasil ramalan anggota *ensemble* dari ARIMA yang terbentuk dengan menggunakan *averaging* dan *stacking* sehingga dipatkan hasil ramalan ARIMA *ensembel*.

Zaier et al. (2010) menjelaskan bahwa ada dua metode yang paling sering digunakan untuk mrngkombinasikan output yang berada dari anggota *ensemble*, yaitu:

a. Metode Rata-Rata (*Averaging*)

Dengan menggunakan metode rata-rata, output dari *ensemble* diperoleh dengan menghitung rata-rata dari output anggota *ensemble*. Misalkan N adalah jumlah dari anggota-anggota model individual dalam suatu ensemble, nilai peramalannya dapat ditulis sebagai berikut:

$$y_t = \frac{1}{N}\sum_k^N \hat{Z}_{k,t} + e_t$$

Dengan $\hat{Z}_{k,t}$ adalah suatu nilai prediksi dari metode ke-*k* pengamatan ke-*t*

Implementasi dari pendekatan rata-rata adalah relatif mudah dan ini telah ditunjukan sebagai pendekatan yang efektif untuk memperbaiki kinerja dari suatu model peramalan tunggal (Bishop, 1995).

b. Metode Penggabungan (*Stacking*)

Stacking adalah suatu metode umum yang menggunakan suatu kombinasi dari suatu model level lebih tinggi dan model level lebih rendah untuk mencapai suatu akurasi predeksi yang lebih tinggi.Secara umum,nilai peramalan *ensemble* dengan kombinasi yang digunakan ialah *stacking* dapat dihitung dengan rumus sebagai berikut:

$$y_t = \sum_k^N C_k \hat{Z}_{k,t}$$

Dengan $\hat{Z}_{k,t}$ adalah suatu nilai predeksi dari metode ke-k pengamatan ke-*t* dan C_k merupakan konstanta yang diestimasi dengan cara meminimumkan fungsi G. Fungsi G tersebut memiliki rumus sebagi berikut:

$$G = \sum_{t=1}^n \left[Z_t - \sum_{k=1}^N C_k \hat{Z}_{k,t}\right]^2$$

Dengan syarat $\sum_{k=1}^N C_k = 1$ dan $0 \le C_k \le 1$. Menurut Breiman (1996) meminimalisasi suatu fungsi G dapat memperbaiki kemampuan generalisasi dari suatu model.

3. METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan pada penelitian ini merupakan data yang didata oleh badan pusat statistic (BPS) yang diambil dari website http://ada.edu. Terdapat beberapa data sensus, untuk penelitian ini data yang diambil adalah data SUSENAS mulai tahun 1997-2012. Data tersebut diambil setiap tahun oleh BPS dan mempunyai variabel yang hampir sama setiap tahunnya.

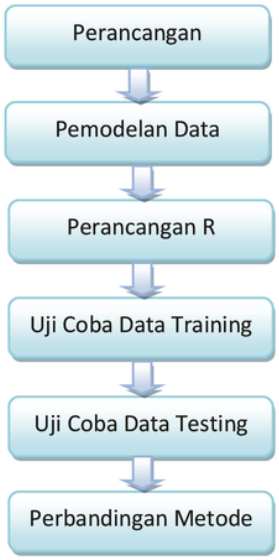
3.2 Variabel Penelitian

Variable penelitian yang akan digunakan pada penelitian ini adalah data belanja/ konsumsi kebutuhan pokok (beras, jagung)

masyarakat pada kabupaten Jember, provinsi Jawa Timur perminggu. Data sensus pertahun dimodelkan dengan rata-rata (*everaging*), diambil dari SUSENAS mulai tahun 1997-2012. Data rata-rata belanja kebutuhan pokok dibagi menjadi dua yaitu data *training* dan data *testing*. Data *training* digunakan untuk memodelkan yaitu mulai tahun 1997-2010, sedangkan data *testing* (dalam 2011-2012) digunakan untuk membagi data yang dihasilkan dari model ARIMA dan ARIMA *ensemble*.

3.3 Langkah Penelitian

Langkah-langkah dalam penelitian ini sesuai dengan tujuan yang ingin dicapai. Tujuan yang pertama diperoleh dengan mengaplikasikan perancangan arsitektur dan pemodelan data dalam membangun sebuah data *warehouse*. Tujuan kedua yaitu merancang aplikasi yang dapat memberikan akses ke data *warehouse* dan R berbasis web. Tujuan ke tiga dapat dicapai setelah tujuan pertama dan kedua dilakukan.



Gambar 1. Tahapan Penelitian

Tahapan penelitian dapat dijelaskan sebagai berikut ini.

1. Perancangan data *warehouse* ada 2 langkah yaitu perangcangan arsitektur dan pemodelan data. Data awal dalam proses perancangan ini adalah data SESUNAS Indonesia tahun 1997-2012. Data *warehouse* ini menggunakan database *engine mysql* versi 5.5.3x, dan berjalan pada sistem operasi linux U-buntu server 12.04 *precise*. Pada tahapan ini juga dilakukan proses konversi data dalam bentuk .sql, ini dilakukan untuk mempermudah konversi ke dalam bentuk data *warehouse*.
2. Perancangan R berbasis web dilakukan dengan melakukan instalasi pada server yang dijadikan uji coba.instalasi ini meliputi instalasi R server berikut dengan konfigurasinya serta instalasi web server sebagai DNS server, kemudian mengkoneksikan antar R server dan data *warehouse* sehingga aplikasi R berbasis web dapat terhubung dengan data *warehouse*.
3. Uji coba R berbasis web terhadap ARIMA dan ARIMA *ensemble*. Pada tahapan ini data dibagi 2 yaitu data *training* dan *testing*. Data *training* menggukan data periode tahun 1997-2010. Sedangkan data testing menggunakan periode tahun 2011-2012. Peramalan menggunakan metode ARIMA dan ARIMA *ensemble* dan hasilnya akan dibandingkan dengan data *testing*. Pada langkah ini diperoleh tujuan ketiga.

4. HASIL DAN PEMBAHASAN

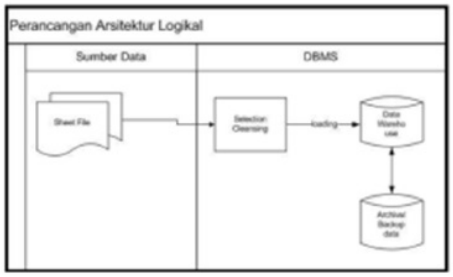
4.1 Perancangan Arsitektur *Logical*

Proses perancangan arsitektur data warehouse adalah proses yang rumit, untuk memulainya kita harus mendefinisikan kebutuhan dari pengguna.pada tahapan ini terdiri dari perancangan arsitektur logikal dan perencangan arsitektur fisik. Data yang akan digunakan sebagai data awal dalam proses perancangan ini adalah data SUSENAS BPS Indonesia (Data Susenas 1997-2012).

Pada perancangan arsitektur logikal pada data *warehouse*, sumber data yang akan digunakan adalah sumber yang berasal dari kegiatan SUSENAS BPS. Data yang digunakan atau di peroleh untuk data SUSENAS tahun 1997-2012 adalah berupa format *.csv, dimana didalam data tersebut terdapat data berupa kode administratif wilayah (mulai dari provinsi sampai dengan desa) yang dijadikan sampel

dan data berupa nilai angka dari survey yang dilakukan.

Dari sumber tersebut, kemudian dilakukan proses pemilihan data yang diperlukan proses seleksi yaitu proses pemilihan data yang diperlukan dari sumber tersebut, karena bisa jadi dalam pembuatan data *warehouse* tidak semua data akan digunakan. Proses yang akan dilakukan selanjutnya adalah proses *cleansing* yaitu proses pembersihan data dan proses transfomasi. Proses terakhir yang dilakukan adalah proses loading yaitu proses memasukan data hasil dari tahap sebelumnya ke dalam data *warehouse*.



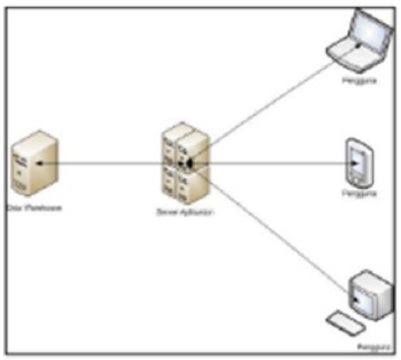
Gambar 2. Perancangan Arsitektur Logikal

Untuk mempermudah proses *import* data berikutnya,maka dibuat sebuah interface untuk import data secara otomatis. Pembuatan interface ini akan mempermudah pengguna untuk melakukan proses *query* data. Berikut adalah tampilan *interface* dalam proses *query* data pada data *warehouse*:

Gambar 3. *Interface* pada data *Warehouse*

4.2 Perancangan Arsitektur Fisik

Arsitektur Fisik data *warehouse* yang didapat data SUSENAS BPS akan terlihat sebagai berikut:



Gambar 4. Perancangan Arsitektur Fisik

4.3 Pemodelan Data

Skema yang digunakan untuk pemodelan data adalah dengan menggunakan *relationship schema*. Pengguna skema ini memungkinkan proses query yang lebih ringan dan memudahkan penjelajahan terhadap data dimensinya. Tabel tabel yang digunakan ditunjukkan dalam tabel 1 dibawah ini:

Tabel 1. Daftar Tabel

No	Tabel	Keterangan
1.	Propinsi	Tabel ini berisi data propinsi dan kode propinsi yang ada di indonesia
2.	Kabupaten	Tabel ini berisi data kabupaten pada setiap propinsi
3.	Kecamatan	Tabel ini berisi data kecamatan pada setiap kabupaten
4.	Desa	Tabel ini berisi data desa/ kelurahan pada setiap kecamatan
5.	SUSENAS	Tabel yang berisi catatan survey terhadap sampel yang diambil pada setiap desa, kecamatan kabupaten dan propinsi yang ada berdasarkan variabel yang menjadi variabel survey pada SUSENAS

4.4 Langkah-langkah Perancangan R Studio

Langkah-langkah installasi R Studio adalah sebagai berikut :

a. Installasi OS Linux Server
b. Ubah pada *source list* arahkan pada *mirror* CRAN. Pada uji coba ini CRAN diarahkan pada mirror unej.ac.id

```
## r installation package from mirror.unej.ac.id
$ debhttp://cran.unej.ac.id/bin/linux/ ubuntu/precise/
$ deb http://cran.unej.ac.id/raring-backports main restricted universe
```

c. Kemudian install paket R system menggunakan perintah berikut:

```
$ sudo apt-get install r-base-dev
```

d. Instalasi paket R pada Ubuntu membutuhkan kunci yang dibuat oleh "*micheal rutter*" gunakan perintah berikut:

```
$ sudo apt-key adv --keyserverKeyserver.ubuntu.com --recv-key E084DAB9
```

e. Langkah berikut adalah instalasi R studio server. R studio server ini adalah sebuah server yang berfungsi menjalankan program R lewat browser menggunakan perintah berikut (os 64 bit):

```
$ sudo apt-get install gdebi-core
$sudo apt-get install libappamor1
$wget http://download2.rstudio.org/rstudio-server-0.98.1091-amd64.deb
$sudo gdebi rstudio-server-0.98.1091-amd64.deb
```

f. Langkah berikut adalah pada konfigurasi server.hal ini diperlukan karena setting default dari R server menggunakan port <8787>. agar lebih mudah mengakses maka kita gunakan port <80> dan address dari 0.0.0.0. penggunaan alamat (address) 0.0.0.0 ini memungkinkan akses R dari berbagai IP menggunakan perintah berikut:

```
$/etc/rstudio/rserver/.conf
www-port=8787 -> ubah menjadi 80
www-address=127.0.0.1 -> ubah menjadi 0.0.0.0
```

4.5 Koneksi R Studio dan Data *Warehouse*

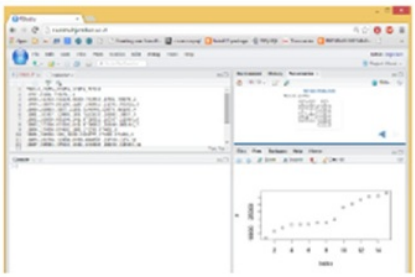
RMySQL merupakan sebuah paket *interface* yang menghubungkan R dan MySQL database. RMySQL dibuat oleh David A. James and Saikat DebRoy pada tahun 2012. Download code tgz dari RMySQL pada alamat `http://cran.r-project.org/web/package/RMYSQL/index.html` atau `http://cran-project.org/src/contrib/RMYSQL_0.9-3.tar.gz` kemudian install menggunakan perintah `$ sudo apt-get install libdbd-mysql libmysqlclient-dev`

Pada console linux server .setelah itu jalankan perintah berikut pada console

```
R CMD INSTALL RMYSQL_0.9-3.tar.gz.
Dbconnect (MYSQL),user='xxx',password='xxxx',dbname='SUSENAS',host='10.10.1xx.xx'
```

merupakan perintah yang harus dijalankan agar terbentuk koneksi antara R server dan database. Setelah perintah tersebut dijalankan, maka berikutnya adalah mencoba dengan membuka data tabel yang dimiliki oleh database SUSENAS. Jika berhasil maka akan tampak tabel-tabel yang dimiliki oleh database SUSENAS. Langkah berikutnya adalah menggali informasi data dengan menggunakan *query*. *Query* data R server menggunakan `dbGETQUERY(nm_koneksi,query)`.

Berikut ini adalah tampilan R studio yang berhasil di koneksikan dengan data warehouse.



Gambar 5. Aplikasi R Studio

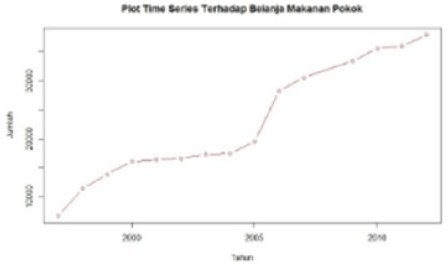
4.6 Uji Coba Program Dengan ARIMA

Ada beberapa langkah untuk mendapatkan nilai peramalan menggunakan metode ARIMA yang sesuai dengan metodologi BOX-Jenkins. Langkah-langkah tersebut meliputi identifikasi

estimasi parameter dan peramalan. Berikut ini merupakan penjelasan model ARIMA untuk data pola belanja makanan pokok di Kabupaten Jember.

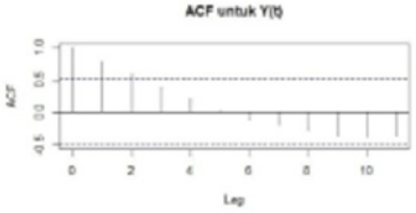
a. Identifikasi

Identifikasi pertama yaitu mengetahui perlu atau tidaknya tranformasi atau yang sering disebut dengan kestasioneran, dapat menggunakan *plot time series* atau uji yang lain. Pengujian untuk melihat kestasioneran dalam varian dapat menggunakan *box-cox*. Akan tetapi pengujian tersebut tidak dilakukan karena tidak merubah pola dari data. Sedangkan identifikasi kedua (uji kestasioneran dalam rata-rata) dapat menggunakan uji *Dickey-Fuller*. Karena hasil uji *Dickey-Fuller* dipengaruhi oleh adanya *outlier* (Suprapto, 2000), maka digunakan *plot time series* untuk melihat kestasioner.

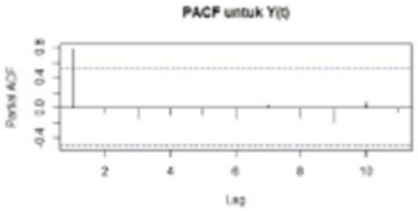


Gambar 6. Plot Time Series Terhadap Belanja Makanan Pokok

Pada uji *Dickey-Fuller* terlihat bahwa data uji memiliki nilai stasioner terlihat dari hasil hipotesa yang bernilai stasioner. Sehingga berdasarkan *plot time series* dan uji *Dickey-Fuller* data belanja makanan pokok di Kabupaten Jember stasioner. Kestasioneran ini berakibat pada nilai (I) pada ARIMA adalah nol), untuk nilai (AR) dan (MA) di lihat melalui plot ACF dan plot PACF



Gambar 7a. Plot ACF dan PACF data belanja makanan pokok



Gambar 7b. Plot ACF dan PACF data belanja makanan pokok

Pada gambar 7 terlihat plot ACF *cuts off* setelah lag 2. Plot ACF merepresentasikan nilai (MA),dari plot ACF tersebut maka nilai (MA) yang dapat diujikan adalah (1) dan (2) sedangkan plot PACF *cutsoff* setelah lag 1. Sehingga model ARMA yang memungkinkan adalah AR(1), ARMA(1,1), ARMA(1,2), tetapi juga akan menguji beberapa model ARMA yang lain yaitu : MA(1), MA(2), ARMA(2,2), dan ARMA(2,1) dan AR(2).

b. Estimasi Parameter

Berdasarkan hasil dari proses program terdapat nilai AIC pada masing masing model ARMA seperti pada tabel Sedangkan uji diagnosa yang dilakukan meliputi uji *white noise* dan uji distribusi normal seperti pada tabel 3.

Tabel 2. Model Arima dan Nilai AIC yang dihasilkan

Model	Nilai AIC
AR(1)	253,7
ARMA(1,1)	255,31
ARMA(1,2)	256,31
MA(1)	267,12
MA(0,2)	265,16
ARMA(2,2)	Error
ARMA(2,1)	254,62
AR(2,0)	253,3

Tabel 3. Hasil Pengujian Model ARMA

Model	Pengujian Asumsi White Noise			Pengujian Asumsi Normal	
	Lag	p-value	Ket	p-value	Ket
AR(1)	0	0,5134	Sig*	0,147	Sig*
	1	0,5134	Sig*		
	2	0,7151	Sig*		

Model	Pengujian Asumsi White Noise			Pengujian Asumsi Normal	
	Lag	p-value	Ket	p-value	Ket
AR(1,1)	0	0,1763	Sig*	0,145	Sig*
	1	0,1763	Sig*		
	2	0,1575	Sig*		
ARMA (1,2)	0	NA	Tidak Sig*	NA	Tidak Sig*
	1	<2e-16	Tidak Sig*		
	2	0,2514	Tidak Sig*		
MA(1)	0	NA	Tidak Sig*	1,05e-12	Tidak Sig*
	1	<2e-16	Sig*		
	2	0,01009	Sig*		
MA(2)	0	NA	Tidak Sig*	<2e-16	Tidak Sig*
	1	<2e-16	Tidak Sig*		
	2	0,08781	Sig*		
AR(2,2)	0	NA	Tidak Sig*	NA	Tidak Sig*
	1	NA	Tidak Sig*		
	2	NA	Tidak Sig*		
AR(2,1)	0	NA	Tidak Sig*	0,695	Tidak Sig*
	1	2,00e-16	Tidak Sig*		
	2	0,9178	Sig*		
AR(2)	0	NA	Tidak Sig*	0,552	Tidak Sig*
	1	<2e-16	Tidak Sig*		
	2	0,1753	Tidak Sig*		

Keterangan: Sig=Signifikan*

Hasil dari tabel 3 menunjukan bahwa terdapat dua model surat yang telah memenuhi asumsi baik *white noise* maupun asumsi berdistribusi normal ,yaitu model AR(1) dan model (1,1) selanjutnya akan digunakan model AR(1) dan model ARMA (1,1) untuk peramalan.

c. Peramalan

Berdasarkan data pada tabel 3 model ARMA yang memungkinkan AR(1,1) maka uji peramalan akan dilakukan dengan meramalkan data 2 tahun kedelapan, hasilnya sebagai berikut:

Tabel 4. Hasil Peramalan Model

Model	Tahun ke1	Tahun ke2
Ar(1)	35643,93	35643,89
ARMA(1,0,1)	35367,90	34572,54

Dari data tersebut kemudian dibandingkan dengan *testing* (tahun 2011 dan 2012), maka akan diperoleh hasil sebagai berikut:

Tabel 5. Uji Perbandingan Dengan Data Testing

Model	Tahun	Hasil	Testing	Selisih
AR(1)	2011	35643,93	36027,71	383,78
	2012	35643,89	37885,84	2241,95
ARMA (1,1)	2011	35367,90	36027,71	659,81
	2012	34572,54	37885,84	3313,30

d. Uji Proram Dengan ARIMA *Ensemble*

Pada pembahasan model ARIMA terdapat dua model ARIMA yang dijadikan sebagai model didalam *ensemble*, hasil dari semua data *training* yang dilakukan pada model ARMA kemudian akan diambil rata-rata (*average*).

Tabel 6. Hasil ARIMA *Ensemble*

Model	Tahun	Hasil	Teating	Selisih
ARMA(1)	2011	35643,93	36027,71	383,78
	2012	35643,89	37885,84	2241,95
ARMA (1,1)	2011	35367,90	36027,71	659,81
	2012	34572,54	37885,84	3313,30
ARIMA *ensemble*	2011	35781,87	36027,71	245,84
	2012	35370,81	37885,84	2515,03

Pengujian terhadap data *testing* memperlihatkan bahwa nilai model ARIMA *ensemble* pada tahun 2011 adalah sebesar 35505,92 dengan selisih terhadap data *testing* sebesar 521,79 pada tahun 2012, nilai pada model *ensemble* adalah sebesar 35108,22 dengan selisih terhadap data testing sebesar 2777,62. Nilai ini lebih besar dari nilai tahun 2012 pada model AR(1) dan ARMA(1,1). Hasil yang diperoleh pada model *ensemble* pada tahun 2011 dan 2012 menjadi tidak lebih baik dari model ARMA(1,1) dan AR (1), karena interval data yang terlalu lebar pada dua model ARMA yaitu AR(1), AR(1,1). Model terbaik terletak pada model AR(1).

5. KESIMPULAN

a. Pada rancang bangun data *warehouse*, penggunaan database *engine* MySQL sudah memadai dalam melakukan aktifitas *query* data.
b. Pada model *interface* data *warehouse* berbasis web memudahkan penggunaan dalam melakukan olah data yang berkaitan dengan *insert, update, import* dan *export*.
c. Penggunaan R studio mudah dilakukan sangat tepat apabila digunakan sebagai alat praktikum dalam perkuliahan.
d. *Session* pengguna aktif pada R studio hanya satu kali, jikapengguna mencoba login dari perangkat lain, maka koneksi dari perangkat lain terputus.
e. Pada uji coba peramalan pola konsumsi (makanan pokok) di Kabupaten Jember diperoleh model terbaik peramalan menggunakan model AR(1). Keterbatasan data bisa menjadi menyebabkan *ensemble* bukan yang terbaik. Seharusnya diharapkan *ensemble* adalah model terbaik.

DAFTAR PUSTAKA

Bishop, C.M. 1995. *Neural Network For Pattern Recognition*. New York: Oxford University Press Inc.

Box, G.E.P, Jenkins, G.M. 1976. *Time Series Analysis Forecasting And Control Revised Edition*. Oakland: Holden-Day, Inc.

Breiman, L. 1996. *Stacked Regression, Machine Learning*, Vol. 24, pp.59-64.

Data Susenas 1997-2012 ,<www.ada.edu> Diunduh Pada 25-28 Maret 2012

FAO. 2012. *The State Of Food Insecurity In The Word 2012*. Rome.

J. Supranto. 2000. *Statistic Teori Dan Aplikasi*. Jakarta: Erlangga.

Kementan. 2012. *Laporan Kinerja Kementrian Pertanian Tahun 2011*. Jakarta: Kementan.

Khashei, M., Bijari, M., Ardali,G.A.R. 2007. *Improvement Of Auto-Regressive Integrated Moving Average Models Using Fuzzy Logic And Artificial Neural Network*. Neurocomputing Vol.72:4

Mukhlis F., Bayu Adhi T. 2011. *Rancang Bangunan Data Warehouse Untuk Menunjang Evaluasi Akademik Di Fakultas*., Prosiding Konferensi Nasional Teknologi Informasi Dan Aplikasinya (KNTIA).

Prayoto. 2008. *Peranan Perguruan Tinggi Dalam Pengembangan IPTEK*. Seminar Nasional Dies Natalis 45 UGM

Pusdatin Kementaan. 2012. *Statistik Kosumsi Pangan Tahun 2012*. Jakarta, Kementan.

Radityo A,N., Johan T., Tony J, H. 2008. *Aplikasi Data Warehouse Untuk Analisa Penjual Mobil Berbasis Multidimensional Modeling (MDM) Dan Star Schema Design*. Jurnal Teknologi Informasi-AITI,Vol 5 No 2.

Shu,C., Burn,D.H. 2004. *Artificial Neurl Network Ensembles And Their Apliction In Pooled Flood Frequency Anlysis*. Water Resource Research, Vol 40,9.

Suryana, A. *Penganekaragaman Konsumsi Pangan Dan Gizi: Faktor Pendukung Peningkatan*.

Wei, W.W.S. 2009. *Time Series Analysis: Univariat And Multivariat Methods. Second Edition*. USA: Person Education,Inc.

Zaeir, I., Shu, C., Ouarda,T.B.M.J, Seidou, O., Chebana, F. 2010. *Estimation of Ice Thickness on Lakes using Artificial Neural Network Ensembles*. Journal Of Hidrology,383 ,330-340.

Rancang bangun data warehouse dan r studio serta pemanfaatannya dalam peramalan pola konsumsi masyarakat di Kabupaten Jember

ORIGINALITY REPORT

16%	13%	4%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.bps.go.id Internet Source	1%
2	pt.slideshare.net Internet Source	1%
3	blogozzy95.blogspot.com Internet Source	1%
4	journal.ugm.ac.id Internet Source	1%
5	faerlinaliliyana.blogspot.com Internet Source	1%
6	Submitted to National College of Ireland Student Paper	1%
7	documents.mx Internet Source	1%
8	Submitted to Universiti Kebangsaan Malaysia Student Paper	1%

9	rifkaanisa.blogspot.com Internet Source	1%
10	effendyimagination.blogspot.com Internet Source	1%
11	ejournal-s1.undip.ac.id Internet Source	<1%
12	text-id.123dok.com Internet Source	<1%
13	www.unmuhjember.ac.id Internet Source	<1%
14	www.pdf-archive.com Internet Source	<1%
15	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	<1%
16	Submitted to Universitas Andalas Student Paper	<1%
17	eprints.uny.ac.id Internet Source	<1%
18	ekanyanan.blogspot.com Internet Source	<1%
19	A Ohri. "R for Cloud Computing", Springer Nature, 2014 Publication	<1%

20	e-journal.unair.ac.id Internet Source	<1%
21	www.mcr.wi.tum.de Internet Source	<1%
22	users.ecs.soton.ac.uk Internet Source	<1%
23	iwaponline.com Internet Source	<1%
24	candra.unsri.ac.id Internet Source	<1%
25	www.ultimefile.it Internet Source	<1%
26	www.airqualitymodeling.org Internet Source	<1%
27	muhiklaten.blogspot.com Internet Source	<1%
28	sigitnugroho.id Internet Source	<1%
29	repositorio.unesp.br Internet Source	<1%
30	sevgiligiyim.com Internet Source	<1%
31	Submitted to State Islamic University of	<1%

	Alauddin Makassar Student Paper	
32	Malim Muhammad, Harjono Harjono, Lukmanul Akhsani. "Peramalan Mahasiswa Baru Ft dan Fkip Um Purwokerto dengan Model Arima", Techno (Jurnal Fakultas Teknik, Universitas Muhammadiyah Purwokerto), 2017 Publication	<1%
33	Zhang, Dan. "Database and NoSQL", R for Programmers, 2016. Publication	<1%
34	www.geek.sc Internet Source	<1%
35	withr.me Internet Source	<1%
36	Submitted to Universitas Muhammadiyah Surakarta Student Paper	<1%
37	eladobecapistrano.com Internet Source	<1%
38	zebradoc.tips Internet Source	<1%
39	ricestaff.weebly.com Internet Source	<1%
40	digilib.unmuhjember.ac.id Internet Source	

		<1%
41	ascelibrary.org Internet Source	<1%
42	Submitted to Politeknik Negeri Bandung Student Paper	<1%
43	Submitted to Forum Komunikasi Perpustakaan Perguruan Tinggi Kristen Indonesia (FKPPTKI) Student Paper	<1%
44	Submitted to Universitas Indonesia Student Paper	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		