

---

## IMPLEMENTASI DATA MINING DALAM MENGESTIMASI BIAYA PENGADAAN BUKU DAN ALAT TULIS DALAM DAFTAR KUANTITAS DAN HARGA (RAB) DENGAN MENGGUNAKAN REGRESI LINEAR BERGANDA PADACV.MARCO STAR

Rini Makdalena<sup>1</sup>, Puji Sari Ramadhan, S.Kom., M.Kom.<sup>#2</sup>, Milfa Yetri, S.Kom., M.Kom.<sup>#3</sup>

<sup>#1</sup> Program Studi Sistem Informasi, STMIK Triguna Dharma

<sup>#2,3</sup> Program Studi Sistem Informasi, STMIK Triguna Dharma

---

### Article Info

#### Article history:

Received xxxx xx<sup>th</sup>, 2020

Revised xxxx xx<sup>th</sup>, 2020

Accepted xxxx xx<sup>th</sup>, 2020

---

#### Keyword:

Estimasi Biaya

Regresi Linear

Data Mining

### ABSTRAK

*CV.Marco Star merupakan perusahaan yang bergerak dibidang penjualan*

*dalam menawarkan buku-buku pelajaran tingkat Paud hingga tingkat Sekolah Menengah Atas. Adapun masalah dalam mengestimasi biaya yang akan dibuat dalam bentuk proposal, agar perusahaan tidak terjadi mengalami kerugian. Berdasarkan hal tersebut penerapan dalam pengadaan buku dan alat tulis dalam daftar kuantitas dan harga (rab) konsep keilmuan Data Mining.*

*Penerapan Data Mining telah banyak digunakan dalam pengolahan data untuk menghasilkan pengetahuan, salah satunya. Data mining dapat mepredeksi dana yang akan dikeluarkan. Selanjutnya algoritma yang akan digunakan untuk pengelolaan Data Mining pada kasus mengestimasi biaya adalah Regresi Linear berganda. untuk penentuan dalam mepredeksi dana yang akan dikeluarkan. Selanjutnya algoritma yang akan digunakan untuk pengelolaan Data Mining pada kasus mengestimasi biaya adalah Regresi Linear berganda.*

*Hasil dari Algoritma Regresi Linear sangat sesuai dan dapat sebagai alat bantu untuk mengestimasi biaya pengadaan buku dan alat tulis dalam daftar kuantitas dan harga (rab) konsep keilmuan Data Mining.*

**Kata Kunci:** *Data Mining, Regresi Linear, Estimasi Biaya*

Copyright © 201x STMIK Triguna Dharma.  
All rights reserved.

---

Nama : Rini Makdalena Manihuruk  
Kator : STMIK Triguna Dharma  
Program Studi : Sistem Informasi  
Email : rinimanihuruk02@gmail.com

---



## 1. PENDAHULUAN

CV.Marco Star merupakan perusahaan yang bergerak dibidang penjualan dalam menawarkan buku-buku pelajaran tingkat Paud hingga tingkat Sekolah Menengah Atas. Pada perusahaan CV.Marco Star memiliki penjualan berbagai instansi pemerintah, seperti sekolah negeri maupun sekolah swasta. Selain itu perusahaan ini juga mengikuti tender yang di selenggarakan oleh Pemerintah. Adapun masalah dalam mengestimasi biaya yang akan dibuat dalam bentuk proposal, agar perusahaan tidak mengalami kerugian.

Berdasarkan hal tersebut penerapan dalam pengadaan buku dan alat tulis dalam daftar kuantitas dan harga (rab) konsep keilmuan *Data Mining*. Penerapan *Data Mining* telah banyak digunakan dalam pengolahan data untuk menghasilkan pengetahuan, salah satunya pada [1] *Data Mining* untuk penentuan dalam mepredeksi dana yang akan dikeluarkan. Selanjutnya algoritma yang akan digunakan untuk pengelolaan *Data Mining* pada kasus mengestimasi biaya adalah *Regresi Linear berganda*.

Penerapan *Regresi Linear* yang merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas. *Simple Regresi Linear* terdiri dari satu buah variabel bebas (x) dengan satu buah variabel terikat (y)[2]. Selain itu metode *Regresi Linear Berganda* dapat digunakan untuk meprediksi sesuatu di masa depan dengan menggunakan lebih dari 2 faktor atau variabel yang memiliki pengaruh terhadap apa yang akan diprediksi[3]. Sehingga Algoritma *Regresi Linear* sangat sesuai dan dapat sebagai alat bantu untuk mengestimasi biaya pengadaan buku dan alat tulis dalam daftar kuantitas dan harga (rab) konsep keilmuan *Data Mining*.

Berdasarkan deskripsi yang telah dikemukakan maka perlu diangkat dengan judul **“Implementasi *Data Mining* Dalam Mengestimasi Biaya Pengadaan Buku Dan Alat Tulis Dalam Daftar Kuantitas Dan Harga (Rab) Dengan Menggunakan *Regresi Linear Berganda* Pada CV.Marco Star”**.

## 2. Tinjauan Pustaka

### 2.1. *Data Mining*

*Data Mining* adalah sebuah proses menganalisa data untuk mengetahui suatu pola dari kelompok data yang tersembunyi. yang bertujuan untuk menemukan, menggali pengetahuan dari data atau informasi yang kita miliki. *Data Mining* berisi pencarian pola yang diinginkan pada *database* untuk membantu pengambil keputusan di waktu yang akan datang.

*Data Mining* merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar. yang tujuan utama *Data Mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki[4] .

*Data Mining* adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya. Secara teknis, *Data Mining* dapat disebut sebagai proses untuk menemukan korelasi atau pola dari ratusan atau ribuan *field* dari sebuah relasional *database* yang besar[5].

## 2.2. Regresi Linier Berganda

Metode regresi linear adalah alat statistik yang dipergunakan untuk mengetahui pengaruh antara satu atau beberapa variabel terhadap satu buah variabel. Manfaat dari regresi linear diantaranya analisis regresi lebih akurat dalam melakukan analisis korelasi, karena analisis itu kesulitan dalam menunjukkan tingkat perubahan suatu variabel terhadap variabel lainnya (slop) dapat ditentukan. Dengan analisis regresi peramalam atau perkiraan nilai variabel terikat pada nilai variabel bebas lebih akurat. Selain itu analisis ini untuk mengetahui arah hubungan antara variabel dependen apakah positif dan negatif dan untuk memprediksi nilai dari variabel dependen apabila nilai dari variabel independen mengalami kenaikan atau penurunan dan variabel independen. Data yang digunakan adalah data berskala interval atau rasio[8].

Teknik *Regresi Linier* berganda digunakan untuk mengetahui ada tidaknya pengaruh signifikan dua atau lebih variabel bebas ( $X_1, X_2, X_3, \dots, X_k$ ) terhadap variabel terikat atau tidak bebas ( $Y$ ). Model *Regresi Linier* berganda untuk populasi dapat ditunjukkan sebagai berikut :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Model *Regresi Linier* berganda untuk populasi diatas dapat ditaksir dengan model *Regresi Linier* berganda untuk sampel, yaitu :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$$

dengan :

$Y$  = nilai penduga bagi variabel  $Y$

$b_0$  = dugaan bagi parameter konstanta

$b_1, b_2, \dots, b_k$  = dugaan bagi parameter konstanta  $\beta_1, \beta_2, \dots, \beta_n$

$X$  = Variabel bebas.

### 3. Metodologi Penelitian

#### 3.1. Algoritma Sistem

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan *Data mining* dalam menyelesaikan permasalahan mengestimasi biaya pengadaan buku dan alat tulis pada CV.Marco Star berdasarkan variabel dengan menggunakan metode *Regresi Linear Bergandan*. Hal ini dilakukan untuk meningkatkan kinerja pada CV.Marco Star.

Berikut ini adalah data yang digunakan sebagai sampel dalam penelitian ini yaitu sebagai berikut:

X1, dan X2 merupakan variabel bebas, sedangkan Y merupakan terikat yang dipengaruhi oleh keberadaan variabel X1, dan X2. Karena estimasi biaya 2020 dan Realisasi 2021 dianggap mempengaruhi biaya alat dan tulis kantor, maka diinisialisasikan :

X1 = Vol

X2 = Satuan

Y = Jumlah Biaya

Tabel 3.3 Inisialisasi Variabel

No	URAIAN	X1	X2	Y
1	Pandai Berpidato	14000	1	14000
2	Rubah dan Anggur	24500	2	49000
3	Senin dan Teknik Mendengarkan	29300	3	87900
4	Pribahasa Indonesia 1	40500	1	40500
5	Pribahasa Indonesia 2	46750	1	46750
6	Pribahasa Indonesia 3	39000	2	78000
7	Profil dan Kiprah Sastrawan Indonesia 1	62000	2	124000
8	Profil dan Kiprah Sastrawan Indonesia 2	62000	2	124000
9	Profil dan Kiprah Sastrawan Indonesia 3	62000	2	124000
10	Profil dan Kiprah Sastrawan Indonesia 4	35800	1	35800
11	Puisi Misterius dan Jejak Kaki Piko	30500	1	30500
12	Batik Modern Karya Kasino	42800	1	42800
13	Santri Produktif Perintis Budidaya Lele	44600	1	44600
14	Ganggang Merah Membawa Berkah	49000	1	49000
15	Mahir Menggunakan Kamus Bhs.Indonesia	31200	1	31200
16	Always English	25500	1	25500
17	Baby Squirrel Learnt A Lesson	28500	1	28500
18	Bob's Prank	26400	2	52800
19	English Everywhere	31000	2	62000
20	Englsih is cool	32500	2	65000
21	English is fun	31200	1	31200
22	I Love English	26400	2	52800
23	I Speak English	26400	2	52800
24	Speaking English at Home and School	28500	1	28500
25	The Fish and The Tortoise	25500	1	25500
26	The Goat's Street	28500	1	28500
27	The Mousedeer and his magic flute	28500	2	57000
28	The Mousedeer cheats the farmer	26000	2	52000
29	wa all speak English	30400	2	60800
30	rhino's best friends and other stories	24500	3	73500
31	The Axe and the two traveler	24500	1	24500
32	The Bear and the fox	24500	1	24500
33	The crab and the fox	24500	2	49000
34	the crane and the fox	24500	2	49000
35	the ox and the flea	30600	1	24500

36	the snake and the man	29300	2	61200
37	the tadpoles look for their mother	98350	1	29300
38	the tears of a giant turtle and other stories	105000	2	196700
39	legend series : cinderelas	105000	1	105000
40	SETS - BUMI dan Ruang Angkasa	105000	1	105000
41	SETS - Dunia hewan dan Tumbuhan	83000	2	210000
42	SETS - Sains dan Teknologi	34000	1	83000
43	SETS - Tubuh kita dan Kesehatan	32000	1	34000
44	Aku makin tahu tentang astroid	35000	2	64000

Tabel 3.3 Inisialisasi Variabel (Lanjutan)

No	URAIAN	XI	X2	Y
45	Aku makin tahu tentang bumi dan iklimnya	28500	1	35000
46	Aku makin tahu tentang planet merkurius , venus dan mars	25000	2	57000
47	mengenal dunia serangga	30000	1	25000
48	mengenal keunikan hewan	32000	3	90000
49	mengenal moluska	30400	2	64000
50	menjelajah dunia kupu-kupu	24500	2	60800

Normalisasi data berfungsi untuk memudahkan proses perhitungan. Berikut adalah hasil normalisasi data dengan membagi variabel X1,X2, menjadi pengelompokkan data. Dari tabel diatas maka diketahui

3.4 Tabel Data Hasil Sigma

$\Sigma$	Nilai
Y	3009950
X1	78
X2	1959400
YX1	5278650
YX2	153405857500
X1^2	140
X2^2	101895595000
X1.X2	101895595000
Y2	3009950

Dari tabel 3.4 diatas maka dilakukan perhitungan proses Regresi Linier Berganda dengan menggunakan rumus persamaan sebagai berikut :

$$\sum Y = na + b_1 \sum X_1 + b_2 \sum X_2 \dots\dots\dots(1)$$

$$\sum YX_1 = a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 \dots\dots\dots(2)$$

$$\sum YX_2 = a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 \dots\dots\dots(3)$$

Dengan menggunakan rumus di atas maka didapatlah persamaan sebagai berikut :

$$3.009.950 = b_0 50 + b_1 78 + b_2 1.959.400 \quad [1]$$

$$5.278.650 = b_0 78 + b_1 140 + b_2 3.009.950 \quad [2]$$

$$153.405.857.500 = b_0 1.959.400 + b_1 3.009.950 + b_2 101.895.595.000 \quad [3]$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (2) adalah sebagai berikut :

$$3.009.950 = b_0 50 + b_1 78 + b_2 1.959.400 [1] \times [2]$$

$$5.278.650 = b_0 78 + b_1 140 + b_2 3.009.950 [2] \times [1]$$

$$-29156400 = -b_1 916 + 2335700 \quad [4]$$



Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (3) adalah sebagai berikut :

$$3.009.950 = b_0 + b_1 \cdot 78 + b_2 \cdot 1.959.400 \quad [1] \times [3]$$

$$153.405.857.500 = b_0 \cdot 1.959.400 + b_1 \cdot 3.009.950 + b_2 \cdot 101.895.595.000 \quad [3] \times [1]$$

---


$$-1772596845000 = b_1 \cdot 2335700 + b_2 \cdot 1255531390000 \quad [5]$$

Kemudian setelah melakukan proses eliminasi antara persamaan (2) hingga persamaan (3), maka diperoleh persamaan baru yaitu sebagai berikut.

$$5.278.650 = b_0 + b_1 \cdot 140 + b_2 \cdot 3.009.950 \quad [2] \times [3]$$

$$101895595000 = 1959400 \cdot b_0 + 1959400 \cdot b_1 + 101895595000 \cdot b_2 \quad [3] \times [2]$$

---


$$1622670075000 = (-b_1 \cdot 39539900) + 2050160380000,00 \quad [6]$$

Kemudian setelah melakukan proses eliminasi dapat persamaan dari 4,5,6, maka diperoleh persamaan baru yaitu sebagai berikut

$$-29156400 = -b_1 \cdot 916 + 2335700 \quad [4]$$

$$-1772596845000 = b_1 \cdot 2335700 + b_2 \cdot 1255531390000 \quad [5]$$

$$1622670075000 = (-b_1 \cdot 39539900) + 2050160380000,00 \quad [6]$$

Dari persamaan yang diperoleh maka dapat nilai  $b_0$ ,  $b_1$ ,  $b_2$  adalah sebagai berikut.

$$B_0 = -47378$$

$$B_1 = 31830$$

$$B_2 = 1,478$$

Seorang *staff* tender memesan peralatan persediaan alat tulis dan peralatan kantor Mengolah data pada bulan januari dengan jumlah volume 3 dan harga 50000 ingin mengetahui jumlah prediksi pada bulan januari berikutnya (2020), berikut adalah perhitungan prediksi

$$Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2$$

Keterangan :

Y = Biaya Estimasi

$b_0$  = Konstanta

$b_1$  = Koefisien Regresi  $X_1$

$b_2$  = Koefisien Regresi  $X_2$

$X_1$  = 3

$X_2$  = 50000

Maka :

$$Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2$$

$$Y = -47378 + 31830 \cdot X_1 + 1,478 \cdot X_2$$

$$Y = -47378 + 31830 \cdot 3 + 1,478 \cdot 50000$$

$$Y = 122,012$$

Jadi, menurut perhitungan diatas maka estimasi biaya pada tender yang akan di dapatkan Rp. 122,012.

#### 4. Hasil Dan Pembahasan

Dalam halaman utama untuk menampilkan pada tampilan *form* pada awal sistem yaitu *form* login dan menu utama. Adapun *form* halaman utama sebagai berikut.

##### 1. Form Login

*Form Login* digunakan untuk mengamankan sistem dari *user-user* yang tidak bertanggung jawab sebelum masuk ke *Form* Utama. Berikut adalah tampilan *Form Login* :

Gambar 5.1 *Form Login*

## 2. *Form Menu Utama*

*Form Menu Utama* digunakan sebagai penghubung untuk *Form Rab* dan *ATK*. Berikut adalah tampilan *Form Menu Utama*:

Gambar 5.2 *Form Menu Utama*

### 4.1 **Halaman Adminstrator**

Dalam adminstrator untuk menampilkan *form* pengolahan data pada penyimpanan data kedalam *database* yaitu *Form Rab* dan *ATK*, dan *Form Proses Regresi Linear Berganda*. Adapun *form* halaman adminstrator utama sebagai berikut.

#### 1. *Form Rab* dan *ATK*

*Form Rab* dan *ATK* adalah *Form* pengolahan *Rab* dan *ATK* dalam penginputan data, ubah data dan penghapusan data. Adapun *Form* gejala adalah sebagai berikut.

No	Nama	x 1	x 2	y
12	Batik Modern Karya Kasino	40500	1	40500
13	Santri Produktif Pemrisa Buddidaya Lela	46750	1	46750
14	Ganggang Merah Membawa Berkah	39000	2	78000

Gambar 5.3 Form Rab dan ATK

## 2. Form Metode Regresi Linear Berganda

Form Metode Regresi Linear Berganda adalah proses perhitungan dalam memprediksi dalam mengestimasi biaya berdasarkan variabel yang sudah ditentukan. Adapun Form Metode Regresi Linear Berganda adalah sebagai berikut.

No	x 1	x 2	y	YX1	YX2	X1 <sup>2</sup>	X2 <sup>2</sup>	X1X2	Y2
1	14000	1	14000	196000	14000	196000	1	14000	19600
10	24500	2	49000	120050	98000	600250	4	49000	240100
11	29300	3	87900	257547	263700	858490	9	87900	772641

Gambar 5.4 Form Proses Regresi Linear Berganda

## 4.2 Pengujian

Pada bagian ini anda diminta untuk melakukan pengujian dengan sampling data baru atau adanya penambahan *record* data dari hasil pengolahan data sementara. Dan pada bagian ini anda diminta untuk dapat menguji keakuratan sistem yang anda rancang dengan *tools-tools* yang sudah teruji dan terkalibrasi sebelumnya. Adapun hasil proses program dalam mengestimasi *biaya* sebagai berikut.



The screenshot shows a software window titled "Metode Regresi". It contains a table with the following data:

No	x 1	x 2	y	YX1	YX2	X1 <sup>2</sup>	X2 <sup>2</sup>	X1X2	Y2
1	14000	1	14000	196000	14000	196000	1	14000	19600
10	24500	2	49000	120050	98000	600250	4	49000	24010
11	29300	3	87900	257547	263700	858490	9	87900	77264

Below the table, there are input fields for "KODE" (value: 1), "x 1" (value: 3), "x 2" (value: 50000), and "Tanggal" (value: Wednesday, December 2, 2020). There are also buttons for "Proses RLB", "SIMPAN HASIL", and "KELUAR".

Gambar 5.5 Hasil Memprediksi Regresi Linear Berganda

The screenshot shows a software window titled "Lp\_2" displaying a report for "CV. MARCO STAR". The report includes a logo and the following text: "PENGADAAN - KONTRAKTOR - LEVERANSIR - SUPPLIER" and "Jln. Surya No. 108 Medan - Sumatera Utara". Below this is a table with the following data:

No	Tanggal	X1	X2	Hasil
01	Wednesday, December 2, 2020	3	50000	122012
02	Friday, January 1, 2021	6	20000	173162
03	Tuesday, February 2, 2021	5	100000	259572

At the bottom of the report, it says "Diketahui Oleh".

Gambar 5.6 Laporan Hasil Estimasi Biaya

## 5. Kesimpulan dan Saran

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang di bahas tentang mengetahui prediksi biaya dengan menerapkan metode *Regresi* terhadap sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan sebagai berikut :

1. Regresi linear berganda berfungsi untuk melakukan untuk memprediksi sesuatu di masa depan dengan menggunakan lebih dari 2 faktor atau variabel yang memiliki pengaruh terhadap apa yang akan diprediksi dan di prediksi biaya pengadaan buku dan ATK..
2. Berdasarkan hasil analisa perancangan sistem ini, yang dimulai pertama kali dalam program ini adalah rancangan *database* rancangan program seperti tampilan *form login* tampilan menu utama tampilan *input data*

daftar kuantitas dan harga (rab) lalu tampilan *form* proses metode *regresi* dan yang terakhir adalah hasil dari data.

3. CV.Marco Star dapat menggunakan sistem berbasis desktop dalam memprediksi dengan menggunakan metode *regresi linear* berganda dengan menggunakan variabel yang sudah di tentukan dan dapat mengestimasi biaya.




## REFERENSI

- [1] D. Sunia, K. and A. P. Jusia, "Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *STIKOM Dinamika Bangsa*, pp. 121-134, 2019.
- [2] P. Katemba and R. K. Djoh, "Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear," *Jurnal Ilmiah FLASH*, vol. III, 2017.
- [3] E. Triyanto, H. Sismoro and D. A. Laksito, "Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul," *Jurnal Teknologi dan Sistem Informasi Univrab*, vol. IV, no. 2477-2062, pp. 73-86, 2019.
- [4] Alfannisa Annurullah Fajrin and Algifanri Maulana, " Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma FP-growth Pada Data Transaksi Penjualan Spare Part Motor," *Kumpulan jurnal Ilmu Komputer (KLIK)*, vol. Volume 05, 2018.
- [5] M. R. Julianti, M. I. Dzulhaq and A. Subroto, "Sistem Informasi Pendataan Alat Tulis Kantor Berbasis Web pada PT Astari Niagara Internasional," *JURNAL SISFOTEK GLOBAL* , vol. IX, no. 2088 – 1762, pp. 92-97, 2019.
- [6] Y. Darmi, A. Setiawan, J. Bali, K. Kampung Bali, K. Teluk Segara and K. Bengkulu, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," 2016.
- [7] R. Permana, D. N. Sulistyowati, A. O. Sar and T. A. Mutiara, "Rancang Bangun Sistem Informasi Penjualan Dan Pembelian Alat Tulis Kantor Pada CV. Putra Mandiri," *Jurnal Teknik Komputer AMIK BSI*, vol. VI, no. 2550-0120, pp. 141-148, 2020.
- [8] P.Katemba and R.Koro Djoh," Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear," *Jurnal Ilmiah FLASH*,2017.

### UCAPAN TERIMA KASIH

Terima kasih diucapkan kepada pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal selanjutnya.

### BIOGRAFI PENULIS

	<b>Rini Makdalena Manihuruk</b>
	<b>Puji Sari Ramadhan, S.Kom., M.Kom.</b>
	<b>Milfa Yetri, S.Kom., M.Kom.</b>