

Prediksi Jumlah Mahasiswa yang Masuk ke Universitas Kristen Maranatha Menggunakan Regresi Linier

Stefanus Emmanuel Chandra Wijaya^{#1}, Andreas Widjaja^{*2}

[#]Program Studi Sistem Informasi, Universitas Kristen Maranatha
Jl. Prof. drg. Surya Sumantri No. 65 Bandung, Indonesia.

¹stefanusecw@gmail.com

²andreas.widjaja@maranatha.ac.id

Abstract — Regression analysis in statistics is a way to find out the causal relationship between one variable and another. The "cause" variable is referred to by various terms as explanatory variable, independent variable, or independent variable X (because it is often depicted as the horizontal axis or X axis on graphs). The variable that is affected is called the affected variable, the dependent variable, or the Y variable. Both variables can be random variables, but the affected variable must always be a random variable. Regression analysis is one of the most common and widely used analyzes. Regression analysis is usually used to make predictions and predictions. There are two types of regression: simple linear regression and multiple regression. Simple linear regression uses only one variable to make predictions. Multiple regression, on the other hand, predicts using two or more variables. Multiple regression is a regression or prediction model that contains several independent variables or predictors or variable X. The term multiple regression is also known as multiple regression. The word plural means plural or many variables.

Keywords— Multiple Regression Analysis, Regression Analysis, Types of Regression Analysis.

I. PENDAHULUAN

Istilah regresi pertamakali diperkenalkan oleh Francis Galton dalam artikelnya yang berjudul *Family Likeness in Stature* pada tahun 1886. Galton menyebutkan bahwa rata-rata tinggi badan anak yang lahir akan cenderung mundur atau 'regress' meskipun kecenderungan orangtua berbadan tinggi akan punya anak berbadan tinggi begitupun sebaliknya. Hukum regresi milik Galton ini dikonfirmasi oleh temannya Karl Pearson yang menyatakan bahwa bila ayahnya bertubuh tinggi maka anaknya akan bertumbuh pendek begitu pun sebaliknya. Dalam bahasa Galton hal ini disebut "*regression to mediocrity*" [1].

Tetapi sekarang makna regresi sudah berbeda jauh dengan teori yang pernah dikatakan oleh Galton. Secara umum sekarang regresi digunakan sebagai analisis sebagai ketergantungan suatu variabel terikat (Y) dengan variabel bebas (X) (bisa lebih dari satu). Dan kegunaan dari regresi linier saat ini lebih digunakan untuk melakukan prediksi menggunakan data-data yang telah dimiliki sebelumnya [2] [3].

Oleh sebab itu teknik regresi linier ini bisa digunakan untuk memprediksi jumlah mahasiswa yang masuk ke Universitas Kristen Maranatha. Untuk itu kita akan mengetahui bagaimana cara memprediksi jumlah mahasiswa yang hendak masuk ke Universitas Kristen Maranatha.

II. LANDASAN TEORI

Dalam proses melakukan analisis regresi ini ada beberapa teori yang dijadikan landasan dalam pembuatannya. Berikut adalah teori yang digunakan:

A. Regresi

Makalah Regresi adalah sebuah model analisis yang gunanya untuk melihat pengaruh antara dua variabel atau lebih dari dua variabel. Hubungan antar variabel itu bersifat fungsional dan diwujudkan dalam model matematika. Pada analisis regresi variabel dibedakan menjadi dua bagian, yaitu terikat (Y) dengan variabel bebas (X) [4].

B. Jenis Regresi

Terdapat dua jenis regresi, yaitu regresi linier sederhana dan regresi linier berganda. Regresi linier sederhana itu hanya menggunakan satu variabel saja untuk dapat memprediksinya. Sedangkan untuk regresi linier berganda menggunakan dua variabel atau lebih untuk dapat memprediksinya [5].

1) Regresi linear sederhana

Analisis Regresi Sederhana adalah sebuah metode pendekatan untuk pemodelan hubungan antara satu variabel dependen dan satu variabel independen. Dalam model regresi, variabel independen menerangkan variabel dependennya. Dalam analisis regresi sederhana, hubungan antara variabel bersifat linier, dimana perubahan pada variabel X akan diikuti oleh perubahan pada variabel Y secara tetap [15].

Sementara pada hubungan non-linier, perubahan variabel X tidak diikuti dengan perubahan variabel Y secara proporsional. seperti pada model kuadrat, perubahan X diikuti oleh kuadrat dari variabel X. Hubungan demikian tidak bersifat linier.

Analisis regresi sederhana dapat digunakan untuk mengetahui arah dari hubungan antara variabel bebas dengan variabel terikat, apakah memiliki hubungan positif atau negatif serta untuk memprediksi nilai dari variabel terikat apabila nilai variabel bebas mengalami kenaikan ataupun penurunan. Pada regresi sederhana biasanya data yang digunakan memiliki skala interval atau rasio [8].

Rumus regresi linier sederhana sebagai berikut.

$$Y = a + bX$$

Keterangan:

Y = Variabel dependen (variabel terikat)

X = Variabel independent (variabel bebas)

a = Konstanta (nilai dari Y apabila X = 0)

b = Koefisien regresi (pengaruh positif atau negatif)

Contoh regresi sederhana adalah jumlah uang saku yang diterima siswa dipengaruhi oleh jarak antara rumah dan sekolah. Jika erdasarkan penjelasan yang masuk akal semakin dekat ke rumah semakin rendah kantong siswa sebaliknya semakin jauh jarak dari rumah ke sekolah semakin esar kantong siswa. sangat penting sehingga jarak dari sekolah (variael X) akan erpengaruh positif terhadap nilai kantong siswa (variael Y) [5].

2) Regresi linear berganda

Regresi berganda adalah model regresi atau prediksi yang melibatkan lebih dari satu variabel bebas atau prediktor. Istilah regresi berganda dapat disebut juga dengan istilah *multiple regression*. Kata multiple berarti jamak atau lebih dari satu variabel.

Tidak sedikit orang yang salah kaprah dalam memahami istilah tersebut. Dimana tidak bisa membedakan antara multiple regression dengan multivariat regression. Perbedaannya adalah jika multiple regression atau regresi berganda adalah adanya lebih dari satu variabel prediktor (variabel bebas/*variabel independen*). Sedangkan *multivariat regression* atau regresi multivariat adalah analisis regresi dimana melibatkan lebih dari satu variabel response (variabel terikat/*variabel dependen*) [4].

Rumus regresi linier berganda sebagai berikut

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n$$

Keterangan :

Y = variabel terikat

X (1,2,3,...) = variabel bebas

b_0 = nilai konstanta

$b_1(2,3,\dots)$ = nilai koefisiensi regresi [6].

Sedangkan untuk kita dapat mencari nilai-nilai b_0 dan b_1, b_2 , dan seterusnya dapat dihitung dengan menggunakan Rumus dibawah ini :

Rumus mencari nilai b_0 dan b_1, b_2 , dan seterusnya

$$b = (X^T X)^{-1} X^T Y$$

Pertama kita harus mencari nilai dari X dan X^T setelah kita menemukan hasilnya maka kita sekarang tinggal mengkalikan X dan X^T setelah menemukan hasilnya maka sekarang tinggal di invers. Lalu setelah hasil invers keluar maka hasil di kalikan lagi dengan X^T dan setelah hasilnya keluar inilah yang di sebut nilai H . Lalu selanjutnya H tinggal dikali dengan Y dan hasilnya akan menghasilkan nilai b_0 dan b_1, b_2 , dan seterusnya [14].

C. Manfaat Regresi

Analisis regresi digunakan hampir pada semua bidang kehidupan, baik itu dalam bidang industri, pemerintahan, engineer, ekonomi dan lain sebagainya.

Dan ini adalah beberapa manfaat yang didapatkan dengan melakukan penerapan analisis regresi:

- Membuat estimasi rata-rata dan nilai variabel tergantung dengan didasarkan pada nilai variabel bebas.
- Meramalkan nilai rata-rata variabel bebas yang didasari nilai variabel bebas di luar jangkauan *sample* [5].

D. Kelebihan dan kekurangani

Berikut ini adalah bentuk dari kekurangan dan kelebihan dari regresi linier berganda

1) Kelebihan

Dengan menggunakan regresi berganda maka dapat menganalisis dengan menggunakan beberapa variable bebas (X) sehingga hasil prediksi yang didapatkan lebih akurat dibandingkan dengan yang regresi linier sederhana yang hanya menggunakan satu variable bebas (X) [13].

2) Kekurangan

Terkadang timbul kemungkinan kesalahan prediksi dengan kenyataan yang sebenarnya atau dengan kata lain semua data belum tentu berbentuk linear [12].

III. METODOLOGI

A. Analisis Sistem

Analisis ini akan menghasilkan suatu prediksi yang digunakan untuk mengetahui jumlah mahasiswa yang masuk ke Universitas Kristen Maranatha menggunakan regresi linier. Pada tahap ini kita akan mengetahui kebutuhan apa saja untuk dapat melakukan prediksi.

1) Analisis Kebutuhan

Agar kita dapat memprediksi jumlah mahasiswa yang masuk ke Universitas Kristen Maranatha menggunakan regresi linier maka dibutuhkan beberapa komponen penunjang, yaitu

- Input sistem
 - Input data Y (Jumlah Mahasiswa)
 - Input data X_1 (Rata-rata Potongan)

- Input data X_2 (Rata-rata Nilai)
- Input data X_3 (Rata-rata Biaya Masuk)
- Input nilai X .
- Hitung nilai X^T .
- Hitung nilai $X \times X^T$
- Hitung nilai $(X \times X^T)^{-1}$
- Hitung nilai $H = (X \times X^T)^{-1} \times X^T$
- Hitung nilai $H \times Y$
- Output sistem
 - Perhitungan nilai $X \times X^T$.
 - Perhitungan nilai $(X \times X^T)^{-1}$.
 - Perhitungan nilai $H = (X \times X^T)^{-1} \times X^T$.
 - Perhitungan nilai $H \times Y$.
 - Perhitungan $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$
 - Hasil prediksi atau ramalan.

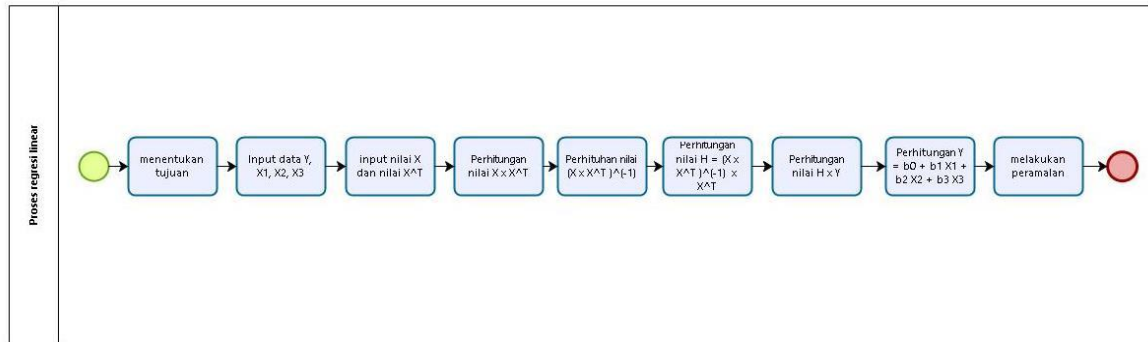
2) Alur kerja Sistem

Alur kerja dari memprediksi jumlah mahasiswa yang masuk ke Universitas Kristen Maranatha menggunakan regresi linier adalah pertama kita harus mempunyai data-data yaitu data Y (Jumlah Mahasiswa), data X_1 (Rata-rata Potongan), data X_2 (Rata-rata Nilai), dan data X_3 (Rata-rata Biaya Masuk). Lalu selanjutnya kita harus mencari nilai dari X dan X^T setelah kita menemukan hasilnya maka kita sekarang tinggal mengkalikan X dan X^T setelah menemukan hasilnya maka sekarang tinggal di invers $(X \times X^T)^{-1}$. Lalu setelah hasil invers keluar maka hasil di kalikan lagi dengan X^T yaitu menjadi $H = (X \times X^T)^{-1} \times X^T$ dan setelah hasilnya keluar inilah yang di sebut nilai H . Lalu selanjutnya H tinggal dikali dengan Y dan hasilnya akan menghasilkan nilai b_0 dan b_1 , b_2 , dan seterusnya. Lalu selanjutnya kita akan menghitung menggunakan rumus $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$. Lalu setelah kita menghitung menggunakan rumus tersebut maka hasilnya adalah hasil prediksi atau ramalan yang kita inginkan sebelumnya yaitu prediksi tentang penerimaan mahasiswa baru di Universitas Kristen Maranatha.

B. Analisis Sistem

Pada tahap ini kita akan mengetahui proses atau cara kerja untuk memprediksi jumlah mahasiswa yang masuk ke Universitas Kristen Maranatha menggunakan regresi linier.

1) Proses Bisnis



Powered by
bizagi
Modeler

Gambar 1. Proses Bisnis

Proses bisnis regresi linier

1. Menentukan tujuan : ini adalah proses dimana kita menentukan hal apa yang mau kita prediksi.
2. Input data Y, X_1, X_2, X_3 : Pada tahap ini kita diminta untuk melakukan pengumpulan data atau input data Y (Jumlah Mahasiswa), data X_1 (Rata-rata Potongan), data X_2 (Rata-rata Nilai), dan data X_3 (Rata-rata Biaya Masuk).
3. Input nilai X dan nilai X^T : pada tahap ini kita diminta untuk mendata dulu berapa nilai dari X dan X^T .
4. Perhitungan nilai $X \times X^T$: pada tahap ini setelah kita mengetahui data dari X dan X^T maka keduanya di kali kan.
5. Perhitungan nilai $(X \times X^T)^{-1}$: pada tahap ini setelah hasilnya keluar maka hasilnya di invers.
6. Perhitungan nilai $H = (X \times X^T)^{-1} \times X^T$: pada tahap ini kita akan menentukan nilai H yaitu dengan cara hasil dari invers pada proses sebelumnya akan di kali dengan X^T .
7. Perhitungan nilai $H \times Y$: pada tahap ini hasil dari H akan dikalikan dengan Y dan hasilnya akan menghasilkan nilai b_0 dan $b_1, b_2, dan b_3$.
8. Perhitungan $Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3$: ini merupakan rumus yang digunakan untuk melakukan peramalan.
9. Melakukan peramalan : Pada tahap ini kita sudah dapat melakukan peramalan atau memprediksi melalui data-data yang sudah kita dapatkan dan yang sudah kita hitung sebelumnya

IV. ANALISIS DAN HASIL

A. Analisis Sistem

Dikarenakan kondisi saat ini sedang pandemi covid19 maka penulis kesulitan mendapatkan data karna data hanya didapat secara online maka penulis hanya berhasil mengumpulkan data mahasiswa yang masuk ke Universitas Kristen Maranatha dari tahun 2008 sampai tahun 2016 saja.

Berikut adalah data mahasiswa yang masuk dari tahun 2008 sampai 2016 :

TABEL I
DATA MAHASISWA

No	Tahun	Jumlah (Y)	Rata-Rata Potongan (X1)	Rata-Rata Nilai (X2)	Biaya Masuk (X3)
1	2008	2639	5	71	20
2	2009	2425	5	80	23
3	2010	2206	4	58	26
4	2011	2199	9	78	34
5	2012	2167	4	60	37
6	2013	2180	5	67	39
7	2014	2070	5	77	43
8	2015	1825	7	78	47
9	2016	1724	9	78	50

B. Hasil Analisis

Selanjutnya dari data yang sudah kita punya maka kita akan memprediksi berapakah jumlah mahasiswa yang akan masuk ke Universitas Kristen Maranatha di tahun selanjutnya.

1) Input nilai X dan nilai X^T

Pada tahap ini nilai X akan diinput dan akan di transpose menjadi X^T

TABEL III
DATA NILAI X

X =			
1	5	71	20
1	5	80	23
1	4	58	26
1	9	78	34
1	4	60	37
1	5	67	39
1	5	77	43
1	7	78	47
1	9	78	50

TABEL IIIII
DATA NILAI X^T

X^T =								
1	1	1	1	1	1	1	1	1
5	5	4	9	4	5	5	7	9
71	80	58	78	60	67	77	78	78
20	23	26	34	37	39	43	47	50

Pada tabel II adalah data dari nilai X sedangkan pada tabel III adalah nilai X^T

2) Perhitungan nilai $X \times X^T$

Pada tahap ini nilai dari data X dan nilai dari data X^T akan dikalikan

TABEL IV
DATA NILAI $X \times X^T$

$X^T \times X =$			
9	53	647	319
53	343	3897	1962
647	3897	47075	23130
319	1962	23130	12209

3) Perhitungan nilai $(X \times X^T)^{-1}$

Pada tahap ini hasil perkalian dari nilai nilai $X \times X^T$ akan di invers.

TABEL V
DATA NILAI $(X \times X^T)^{-1}$

$(X^T \times X)^{-1} =$			
11.85059	0.415241527	-0.17824	-0.038683342
0.415242	0.070879322	-0.00936	-0.004507065
-0.17824	-0.009360161	0.00316	0.000174083
-0.03868	-0.004507065	0.000174	0.001487126

4) Perhitungan nilai $H = (X \times X^T)^{-1} \times X^T$

Pada tahap ini kita akan mencari nilai H yaitu dengan cara hasil dari nilai $(X \times X^T)^{-1}$ akan dikalikan dengan hasil dari data X^T

TABEL VI
DATA NILAI $H = (X \times X^T)^{-1} \times X^T$

$H = (X^T \times X)^{-1} \times X^T =$								
0.4978884	-1.222347803	2.167704633	0.369587359	1.385702	0.475877	-1.46128592	-0.96378	-0.249346116
0.0149254	-0.082837233	0.038685794	0.169822666	-0.02961	-0.03327	-0.144898041	-0.03053	0.097709632
0.0028223	0.031787682	-0.027857471	-0.010058742	-0.01962	-0.00651	0.025788301	0.010925	-0.007273421
-0.019116	-0.013088171	-0.007949545	-0.015106212	0.008757	0.008443	0.016132096	0.013241	0.0086878

5) *Perhitungan nilai H x Y*

Pada tahap ini hasil dari *H* akan dikalikan dengan *Y* dan hasilnya akan menghasilkan nilai b_0 , b_1 , b_2 , dan b_3 .

TABEL VII
 DATA NILAI *H X Y*

Beta = H*y =	
b0 =	2771.008794
b1 =	-26.60637089
b2 =	5.024825691
b3 =	-23.02508999

6) *Hasil nilai R²*

Pada tahap ini adalah hasil dari nilai R^2

TABEL VIVII
 DATA NILAI R^2

R ² =	0.891795388
R=	0.944349187

7) *Hasil aplikasi PSPP*

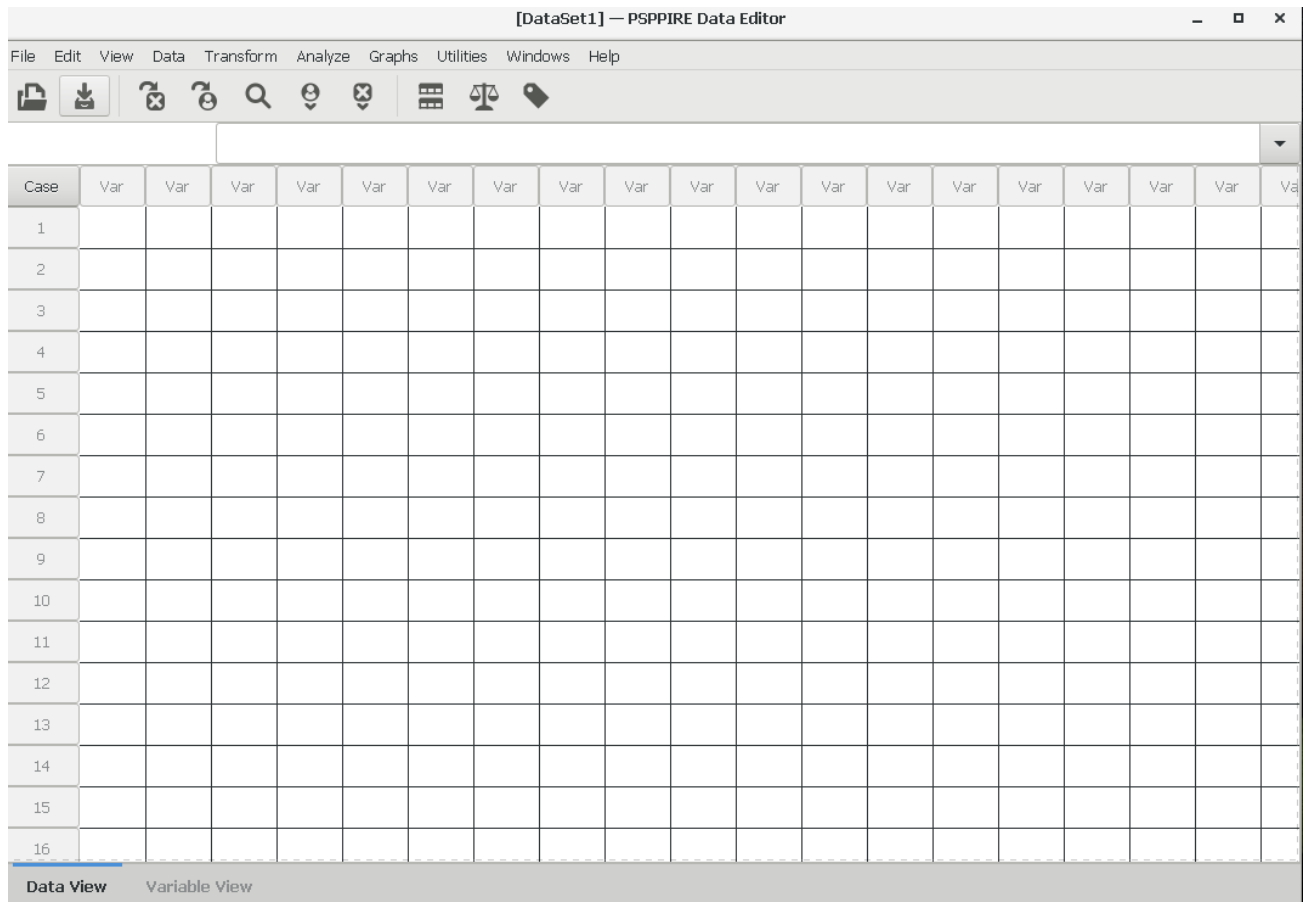
PSPP adalah aplikasi perangkat lunak analisis data sampel gratis yang ditujukan sebagai alternatif gratis untuk IBM SPSS Statistics. Ini memiliki antarmuka pengguna grafis dan antarmuka baris perintah tradisional [9].

Perangkat lunak ini menyediakan serangkaian kemampuan yang komprehensif mencakup frekuensi, perbandingan tabulasi silang rata-rata (uji nilai t dan ANOVA satu arah), regresi linier, regresi logistik, relokasi data, pengujian non-parametrik, faktor. Ini menyediakan rangkaian lengkap fitur seperti analisis. Analisis cluster, analisis komponen utama, analisis Chi-Square, dan lain-lain [10].

Hasil statistik dan grafik tersedia dalam format ASCII, PDF, PostScript, SVG, atau HTML. Dan juga dapat membuat berbagai grafik statistik. B. Histogram, diagram lingkaran, plot scree, dan diagram np.

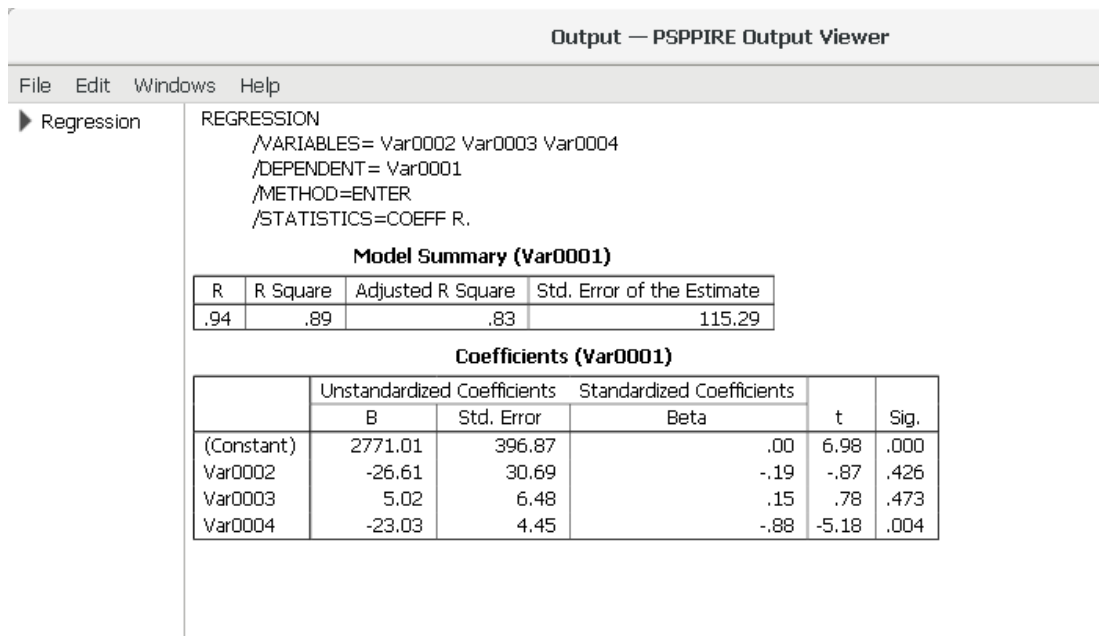
PSPP juga dapat mengimpor spreadsheet Gnumeric dan OpenDocument, database Postgres, nilai yang dipisahkan koma, dan file ASCII dan dapat mengekspor file dalam format file portabel dan sistem SPSS serta file ASCII. Beberapa library yang digunakan oleh PSPP dapat diakses secara terprogram [11].

Gambar 1 adalah contoh dari tampilan aplikasi PSPP



Gambar 2. Tampilan dari aplikasi PSPP

Gambar 2 adalah hasil perhitungan dari aplikasi pspp:



Gambar 3. Hasil output dari aplikasi PSPP

Keterangan : Constant (nilai b_0)
 Var0002 (nilai b_1)
 Var0003 (nilai b_2)
 Var0004 (nilai b_3)

Bisa kita lihat dimana hasil dari nilai b_0, b_1, b_2, b_3 , dan nilai R^2 menggunakan aplikasi PSPP sama dengan hasil yang dilakukan dengan menggunakan excel.

C. Hasil Prediksi

Pada tahap ini hasil dari nilai b_0, b_1, b_2 , dan b_3 akan dimasukkan kedalam rumus $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$ dimana datanya di hasilkan dari rata-rata data X_1 (Rata-rata Potongan), data X_2 (Rata-rata Nilai), dan data X_3 (Rata-rata Biaya Masuk).

Rata-rata $X_1 = 5.89$

Rata-rata $X_2 = 71.88888889$

Rata-rata $X_3 = 35$

Dan prediksisnya adalah $Y = 2159$

V. SIMPULAN DAN SARAN

A. Simpulan

Cara yang digunakan untuk memprediksi adalah dengan cara melakukan rata-rata dari data X_1 (Rata-rata Potongan), data X_2 (Rata-rata Nilai), dan data X_3 (Rata-rata Biaya Masuk) selanjutnya dimasukkan kedalam rumus sehingga dapat disimpulkan bahwa hasil prediksi akan berbeda-beda.tergantung dari data X_1, X_2, X_3 .

B. Saran

Metode ini juga dapat dilakukan untuk memprediksi jumlah mahasiswa baru pada tahun berikutnya berdasarkan data X_1 (Rata-rata Potongan), data X_2 (Rata-rata Nilai), dan data X_3 (Rata-rata Biaya Masuk) yang sebenarnya.

DAFTAR PUSTAKA

- [1] k. "konsultanstatistik," juli 2011. [Online]. Available: <https://www.konsultanstatistik.com/2011/07/sejarah-singkat-regresi.html>.
- [2] F. David, "Statistical Models: Theory and Practice," p. 26, 5 februari 2009.
- [3] S. Russell dan P. Norvig, Artificial Intelligence: A Modern Approach, USA: Prentice Hall Press, 2009.
- [4] R. Wardana, "lifepal," 30 oktober 2020. [Online]. Available: <https://lifepal.co.id/media/regresi/>.
- [5] U. Tanoto, "Jojo Nomic," 11 januari 2021. [Online]. Available: <https://www.jojonomic.com/blog/analisis-regresi/>.
- [6] A. Hidayat, "Stastistikian," 1 januari 2018. [Online]. Available: <https://www.statistikian.com/2018/01/penjelasan-tutorial-regresi-linear-berganda.html>.
- [7] Y. Liu, "Analytical Solution of Linear Regression," towards data science, 1 November 2018. [Online]. Available: <https://towardsdatascience.com/analytical-solution-of-linear-regression-a0e870b038d5>.
- [8] d. kho, "Analisis Regresi Linear Sederhana (Simple Linear Regression)," 2020. [Online]. Available: <https://teknikelektronika.com/analisis-regresi-linear-sederhana-simple-linear-regression/>.
- [9] b. pfaff, "pspp 1.4.1 released," 2020.
- [10] "pspp overview," 2013. [Online]. Available: <http://freestatisticalsoftware.com>.
- [11] "GNU Psp," 2020. [Online]. Available: <http://gnu.org>.
- [12] d. "Kenali Analisis Regresi Linear, Salah Satu Metode Pengolahan Data yang Sering Digunakan," 28 april 2021. [Online]. Available: <https://www.dqlab.id/kenali-analisis-regresi-linear-metode-pengolahan-data-yang-sering-digunakan>.
- [13] u. "Kelebihan Regresi Linier," 3 februari 2020. [Online]. Available: <https://lancanguning.com/post/14984/kelebihan-regresi-linier.html>.
- [14] "Contoh Soal Pembahasan Regresi Linier Berganda [Lengkap]," 16 mei 2020. [Online]. Available: <https://www.statmat.net/regresi-linier-berganda/>.
- [15] r. aulia, "Analisis Regresi Sederhana, Ini Penjasannya," [Online]. Available: <https://www.globalstatistik.com/analisis-regresi-sederhana-ini-penjasannya/>.