

## **PREDIKSI JUMLAH PRODUKSI KELAPA SAWIT DENGAN REGRESI LINEAR BERGANDA**

**Iat Solihat**

Universitas Dinamika Bangsa Jambi, iatsolihat389@gmail.com

**Jasmir Jasmir**

Universitas Dinamika Bangsa Jambi, ijay\_jasmir@yahoo.com

**Fachruddin Fachruddin**

Universitas Dinamika Bangsa Jambi, facruddin.stikom@gmail.com

### **Abstract**

PT. Plantation Nusantara 6 is a state-owned enterprise that has two core businesses with a composition of 95% of palm oil business and 5% of tea business. Bunut Business Unit is one of the business Unit and is a core garden of PT. Plantation Nusantara (PTPN) 6 located in Markanding Village, North Bahar District, Muaro Jambi Regency, Jambi province which has potential of palm oil plantation crops Large enough, in August 2019 the total production reached 41.414.550, this indicates that the amount of oil palm production is still not efficient because due to irregular rainfall, plant life, fresh fruit bunches, haze from fires and Harvester Power. It is therefore needed an analysis to know if the amount of production in the next month has increased or decreased and whether there are other production factors that affect it. Therefore, researchers conduct data mining analysis on the production of oil palm to become information that can be used by PTPN 6 business Unit Bunut Jambi. Researchers took production data in 2018 and in 2019. In conducting the analysis, researchers use multiple linear regression methods and use SPSS tools. After analysis with a partial T-test for fresh fruit bunches ( $X_1$ ) has a significant influence on the amount of production of oil palm ( $Y$ ), the list of present Harvester ( $X_2$ ) has an insignificant influence on the amount of production of oil palm ( $Y$ ). Simultaneously (together) fresh fruit bunches ( $X_1$ ) and Harvester ( $X_2$ ) greatly affects the number of palm oil production ( $Y$ ). Dearn also for MAD (Mean Absolute Deviation) 43.636.311,67, MSE (Mean Square Error) 1.916.114.651.658.140, RMSE (Root Mean Square Error) 43.773.446,88 and MAPE (Mean Square Percentage Error) of 78,59%.

### **Keywords:**

*Oil Palm, Multiple Linear Regression, SPSS*

### **Abstrak**

Pada PT. Perkebunan Nusantara 6 merupakan Perusahaan BUMN yang memiliki dua bisnis inti dengan komposisi 95% Bisnis Kelapa Sawit dan 5% Bisnis Teh. Unit Usaha Bunut adalah salah satu Unit Usaha dan merupakan kebun inti PT. Perkebunan Nusantara (PTPN) 6 yang lokasinya berada di Desa Markanding, Kecamatan Bahar Utara, Kabupaten Muaro Jambi, Provinsi Jambi yang memiliki potensi tanaman perkebunan kelapa sawit yang cukup besar, pada Agustus 2019 jumlah produksi mencapai 41.414.550, ini menunjukkan bahwa jumlah produksi kelapa sawit masih belum efisien karena disebabkan curah hujan yang tidak menentu, umur tanaman, tandan buah segar, kabut asap dari kebakaran dan tenaga pemanen. Oleh karena itulah dibutuhkan suatu analisis untuk mengetahui apakah jumlah produksi dalam bulan depannya mengalami peningkatan atau penurunan dan apakah ada faktor-faktor penentu produksi lain yang mempengaruhinya. Oleh karena itu, peneliti melakukan analisis data mining pada produksi kelapa sawit agar menjadi informasi yang dapat digunakan oleh PTPN 6 Unit Usaha Bunut Jambi. Peneliti mengambil data produksi tahun 2018 dan tahun 2019. Dalam melakukan analisis, peneliti menggunakan metode regresi linear berganda dan menggunakan tools SPSS. Setelah dilakukan analisis dengan uji t parsial untuk tandan buah segar ( $X_1$ ) memiliki pengaruh yang cukup signifikan terhadap jumlah produksi kelapa sawit ( $Y$ ), daftar hadir pemanen ( $X_2$ ) memiliki pengaruh tidak signifikan terhadap jumlah produksi kelapa sawit ( $Y$ ). Secara simultan (bersama-sama) tandan buah segar ( $X_1$ ) dan pemanen ( $X_2$ ) sangat berpengaruh terhadap Jumlah Produksi Kelapa Sawit ( $Y$ ). Diperoleh juga untuk MAD (Mean Absolute Deviation) 43.636.311,67, MSE (Mean Square Error) 1.916.114.651.658.140, RMSE (Root Mean Square Error) 43.773.446,88 dan MAPE (Mean Square Percentage Error) sebesar 78,59%.

### **Kata Kunci:**

*Kelapa Sawit, Regresi Linier Berganda, SPS*

## 1. PENDAHULUAN

Muaro Jambi memiliki 46.891 Unit Perkebunan Kelapa Sawit (BPS atau Badan Pusat Statistik, 2017)[1]. Unit Usaha Buntut adalah salah satu Unit Usaha dan merupakan kebun inti PT. Perkebunan Nusantara (PTPN) 6 yang lokasinya berada di Desa Markanding, Kecamatan Bahar Utara, Kabupaten Muaro Jambi, Provinsi Jambi yang memiliki potensi tanaman perkebunan kelapa sawit yang cukup besar, pada Agustus 2019 jumlah produksi mencapai 41.414.550, ini menunjukkan bahwa jumlah produksi kelapa sawit masih belum efisien karena disebabkan curah hujan yang tidak menentu, umur tanaman, tandan buah segar, kabut asap dari kebakaran dan tenaga pemanen. Oleh karena itulah dibutuhkan suatu analisis untuk mengetahui apakah jumlah produksi dalam bulan depannya mengalami peningkatan atau penurunan dan apakah ada faktor-faktor penentu produksi lain yang mempengaruhinya.

### a. Data Mining

*Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.[2] Tujuan utama *data mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki[3][4]

*Data Mining* adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting.[5][6]

### b. Algoritma Regresi Linear

Bidang *Regresi Linear* bertujuan untuk menjelaskan atau memodelkan hubungan antarvariabel. Dimana terdapat variabel  $y$  sebagai variabel respons, *output*, tak bebas, atau variabel yang dijelaskan dan variabel  $x$  sebagai variabel prediktor, masukan, bebas atau variabel penjelas.[7][8]

Dalam statistika, *regresi linear* adalah sebuah pendekatan untuk memodelkan hubungan antara variabel terikat  $y$  dan satu atau lebih variabel bebas yang disebut  $x$ . Salah satu kegunaan *regresi linear* adalah untuk melakukan prediksi berdasarkan data-data yang telah dimiliki sebelumnya. Hubungan di antara variabel-variabel tersebut disebut sebagai model *regresi linear*. Berdasarkan variabel bebas, maka *regresi linear* dapat dibagi menjadi dua, yaitu *regresi linear* sederhana dan *regresi linear* berganda.[9]

### c. Regresi Linear Berganda

Analisis *regresi linear* berganda digunakan oleh peneliti, bila peneliti meramalkan bagaimana keadaan (naik turunnya) variabel dependen (kriterium), bila dua atau lebih variabel independen sebagai faktor *predictor* dimanipulasi (dinaik/turunkan nilainya). Jadi analisis *regresi linear* berganda akan dilakukan bila jumlah variabel independennya minimal 2.[2]

Adapun persamaan *regresi linear* berganda dapat dirumuskan sebagai berikut:

$$Y = \alpha + b_1X_1 + b_2X_2 \dots + b_nX_n + \epsilon$$

Keterangan :

- Y = Variabel dependen
- a = Konstanta
- $b_1, b_2, b_3, \dots, b_n$  = Koefisien regresi variabel independen
- $X_1$  = Variabel independen pertama
- $X_2$  = Variabel independen kedua
- n = Banyak sampel
- $\epsilon$  = Standar error

### d. SPSS

SPSS merupakan salah satu program aplikasi statistik yang paling banyak dipakai oleh pengguna komputer. Program ini memiliki analisis cukup tinggi, memiliki *interface* pada lingkungan grafis dengan cara pengoperasiannya yang cukup sederhana sehingga mudah untuk dipahami pemakaiannya.[10]

SPSS adalah paket *software* yang digunakan untuk menganalisa statistik.[11]

Beberapa fitur statistik yang didukung oleh SPSS adalah:

1. Statistik Deskriptif: *Cross tabulation, frequencies, descriptive, explore, descriptive ratio statistics.*
2. Statistik bivariat: *Means, T-test, ANOVA, Correlation, dan nonparametric test.*
3. Predikat untuk hasil numerik: *Linear Regression.*
4. Prediksi untuk identifikasi kelompok: *Factor analysis, dan discriminant.*

Fitur-fitur diatas dapat diakses menggunakan menu *pull down* atau dapat diprogram dengan sintaksis khusus menggunakan bahasa 4GL proprietary. Pemrograman dengan sintaksis ini memiliki keuntungan mudah diulangisehingga dapat digunakan berulang kali. SPSS memungkinkan penanganan manipulasi data kompleks dan memudahkan analisisnya.

## 2. HASIL DAN PEMBAHASAN

### 2.1. Data

Data yang digunakan dalam penelitian ini adalah jenis data time series (runtun waktu) yakni jenis data yang terdiri dari variabel-variabel yang dikumpulkan menurut urutan waktu dalam suatu rentang waktu tertentu. Jika waktu dipandang bersifat diskrit (waktu dapat dimodelkan bersifat kontinu), frekuensi pengumpulan selalu sama (*equidistant*). Dalam kasus diskrit, frekuensi dapat berupa misalnya detik, menit, jam, hari, minggu, bulan atau tahun dan lain-lain.

Data runtun waktu dapat dituliskan sebagai :

$$y_t = y_1, y_2, y_3 \dots y_n ; t = 1, 2, 3, \dots n$$

Dengan  $y_t$  adalah data ke  $t$  dari runtun waktu tersebut. Jika  $t \in A$  dengan  $A$  himpunan bilangan asli maka runtun waktu ini dinamakan waktu diskrit. Sedangkan jika  $t \in R$  dengan  $R$  himpunan bilangan riil maka runtun waktu ini dinamakan runtun waktu kontinu.

### 2.2. Variabel Penelitian

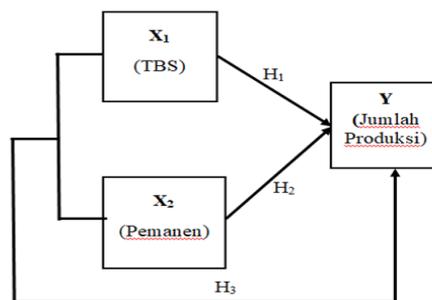
Berdasarkan pengertian diatas dan disesuaikan pada judul penelitian, maka penelitian menggunakan dua variabel, yaitu:

#### 1. Variabel Bebas atau Independen

Variabel bebas dalam pengertian ini adalah variabel yang dapat mempengaruhi variabel lain. Yang menjadi variabel bebas dalam penelitian ini adalah Tandan Buah Segar ( $X_1$ ) dan Daftar Hadir Pemanen ( $X_2$ ). Dalam penelitian ini dinamakan sebagai variabel ( $X$ ).

#### 2. Variabel Terikat atau Dependen

Yang dimaksud dengan variabel terikat adalah variabel yang dapat dipengaruhi oleh variabel lain. Dalam hal ini, yang menjadi variabel terikat adalah Jumlah Produksi Kelapa Sawit, yang kemudian dalam penelitian ini dinamakan sebagai variabel ( $Y$ ). Adapun model dalam penelitian ini bisa dilihat pada gambar 1:



Gambar 1. Pengaruh X terhadap Y

Gambar 1 menjelaskan tentang variabel X yang berpengaruh terhadap Y. Sebagaimana  $X_1$  merupakan Tandan Buah Segar.  $X_2$  adalah Daftar Hadir Pemanen dan Y Jumlah Produksi Kelapa Sawit.  $H_1$  menjelaskan apakah secara parsial berpengaruh terhadap Y,  $H_2$  menjelaskan apakah secara parsial berpengaruh terhadap Y, dan  $H_3$  menjelaskan apakah secara simultan atau bersama-sama ( $H_1, H_2$ ) berpengaruh terhadap Y.

### 2.3. Proses Pengolahan Data

#### a. Data Training

Pada proses ini sebelumnya jumlah data ada 22 data, kemudian dibagi menjadi 2 yaitu 17 data untuk data *training* dan 5 data untuk data *testing*. Adapun data yang digunakan untuk training dapat dilihat pada tabel 1:

**Tabel 1.** Data Jumlah Produksi Kelapa Sawit

Tahun 2018		Tahun 2019	
Bulan	Jumlah Produksi Kelapa Sawit (Kg)	Bulan	Jumlah Produksi Kelapa Sawit(Kg)
Januari	3.769.230	Januari	5.186.070
Februari	7.536.190	Februari	9.884.260
Maret	12.096.260	Maret	14.504.720
April	16.535.600	April	19.392.980
Mei	21.790.510	Mei	24.758.550
Juni	26.302.480	Juni	30.122.420
Juli	30.760.650		
Agustus	36.083.090		
September	41.207.160		
Oktober	46.885.860		
November	48.303.000		

Tabel 1 ini menampilkan jumlah produksi kelapa sawit tahun 2018 dan tahun 2019 pada PTPN 6 Unit Usaha Bunut Jambi. Data tersebut merupakan variabel dependen (Y). Adapun data tandan buah segar dapat dilihat pada tabel 2:

**Tabel 2.** Tandan Buah Segar

Tahun 2018		Tahun 2019	
Bulan	Jumlah Tandan Segar (Kg)	Bulan	Jumlah Tandan Segar (Kg)
Januari	443.842	Januari	531.357
Februari	868.348	Februari	1.004.878
Maret	1.365.873	Maret	1.466.676
April	1.840.094	April	1.961.877
Mei	2.385.323	Mei	2.486.039
Juni	2.852.110	Juni	3.002.836
Juli	3.327.575		
Agustus	3.886.331		
September	4.460.013		
Oktober	5.073.522		
November	5.637.239		

Tabel 2 ini menampilkan jumlah tandan buah segar tahun 2018 dan tahun 2019 pada PTPN 6 Unit Usaha Bunut Jambi. Data tersebut merupakan variabel independen ( $X_1$ ) atau variabel bebas yang mempengaruhi terhadap variabel dependen (Y). Adapun data absen pemanen dapat dilihat pada tabel 3:

**Tabel 3.** Data Absen Pemanen

Tahun 2018		Tahun 2019	
Bulan	Pemanen (KS. KNG, PKWT) (Hari Kerja)	Bulan	Pemanen (KS. KNG, PKWT) (Hari Kerja)
Januari	2.467	Januari	3.662
Februari	3.282	Februari	4.230
Maret	3.093	Maret	3.923
April	3.290	April	4.097
Mei	3.453	Mei	3.971
Juni	2.808	Juni	3.659
Juli	3.396		
Agustus	3.775		
September	4.319		
Oktober	4.262		
November	4.242		

Pada Tabel 3 menampilkan Data Absen Pemanen (Daftar Hadir) tahun 2018 dan tahun 2019 pada PTPN 6 Unit Usaha Bunut Jambi yang terdapat di 6 Afdeling. Pemanen PTPN 6 Unit Usaha Bunut Jambi terbagi menjadi 3 bagian yaitu Karyawan Sendiri / Karyawan Golongan (KS) yang berjumlah 15 Pemanen, 93 Karyawan /

Pemanen Non Golongan (KNG) dan 81 Karyawan / Pemanen Perjanjian Kerja Waktu Tertentu (PKWT). Data tersebut merupakan variabel independen ( $X_2$ ) atau variabel bebas yang mempengaruhi terhadap variabel dependen (Y).

Sistem melakukan penghitungan dengan menerapkan metode *Regresi Linear*. Proses yang pertama adalah menentukan komponen  $X_1$ ,  $X_2$ , Y dan data sampel yang digunakan yaitu jumlah produksi kelapa sawit, jumlah tandan buah segar, dan daftar hadir pemanen selama tahun 2018 dan tahun 2019. Adapun data tersebut bisa dilihat pada tabel 4:

**Tabel 4.** Perhitungan Data Training

Bulan	TBS ( $X_1$ )	Pemanen ( $X_2$ )	Jumlah Produksi (Y)
Januari 2018	443.842	2.467	3.769.230
Februari 2018	868.348	3.282	7.536.190
Maret 2018	1.365.873	3.093	12.096.260
April 2018	1.840.094	3.290	16.535.600
Mei 2018	2.385.323	3.453	21.790.510
Juni 2018	2.852.110	2.808	26.302.480
Juli 2018	3.327.575	3.396	30.760.650
Agustus 2018	3.886.331	3.775	36.083.090
September 2018	4.460.013	4.319	41.207.160
Oktober 2018	5.073.522	4.262	46.885.860
November 2018	5.637.239	4.242	48.303.000
Januari 2019	531.357	3.662	5.186.070
Februari 2019	1.004.878	4.230	9.884.260
Maret 2019	1.466.676	3.923	14.504.720
April 2019	1.961.877	4.097	19.392.980
Mei 2019	2.486.039	3.971	24.758.550
Juni 2019	3.002.836	3.659	30.122.420
<b>JUMLAH</b>	<b>42.593.933</b>	<b>61.929</b>	<b>395.119.030</b>

Tabel 4 menampilkan perhitungan data *training* pada tahun 2018 dan tahun 2019 yang akan di prediksi dan dilakukan perhitungan secara manual untuk mengetahui hasil persamaan *regresi linear* nya, kemudian  $X_1$ ,  $X_2$  dan Y di pangkatkan 2 atau dilakukan perkalian dengan variabel nya itu sendiri. Dilakukan juga perhitungan perkalian antara variabel independen yaitu tandan buah segar ( $X_1$ ) dengan variabel dependen yaitu jumlah produksi kelapa sawit (Y), variabel independen daftar hadir pemanen ( $X_2$ ) dengan variabel dependen yaitu jumlah produksi kelapa sawit (Y), dan perkalian antara variabel independen tandan buah segar ( $X_1$ ) dengan variabel independen daftar hadir pemanen ( $X_2$ ). Adapun hasil dari perhitungannya dapat dilihat pada tabel 5.

**Tabel 5.** Hasil Perkalian Variabel X dan Y

$X_1^2$	$X_2^2$	Y <sup>2</sup>
196.995.720.964	6.086.089	14.207.094.792.900
754.028.249.104	10.771.524	56.794.159.716.100
1.865.609.052.129	9.566.649	146.319.505.987.600
3.385.945.928.836	10.824.100	273.426.067.360.000
5.689.765.814.329	11.923.209	474.826.326.060.100
8.134.531.452.100	7.884.864	691.820.454.150.400
11.072.755.380.625	11.532.816	946.217.588.422.500
15.103.568.641.561	14.250.625	1.301.989.383.948.100
19.891.715.960.169	18.653.761	1.698.030.035.265.600
25.740.625.484.484	18.164.644	2.198.283.867.939.600
31.778.463.543.121	17.994.564	2.333.179.809.000.000
282.340.261.449	13.410.244	26.895.322.044.900
1.009.779.794.884	17.892.900	97.698.595.747.600
2.151.138.488.976	15.389.929	210.386.902.278.400
3.848.961.363.129	16.785.409	376.087.673.280.400
6.180.389.909.521	15.768.841	612.985.798.102.500
9.017.024.042.896	13.388.281	907.360.186.656.400
146.103.639.088.277	230.288.449	12.366.508.770.753.100
<b><math>X_1Y</math></b>	<b><math>X_2Y</math></b>	<b><math>X_1X_2</math></b>

1.672.942.581.660	9.298.690.410	1.094.958.214
6.544.035.514.120	24.733.775.580	2.849.918.136
16.521.954.934.980	37.413.732.180	4.224.645.189
30.427.058.346.400	54.402.124.000	6.053.909.260
51.977.404.684.730	75.242.631.030	8.236.520.319
75.017.566.232.800	73.857.363.840	8.008.724.880
102.358.369.923.750	104.463.167.400	11.300.444.700
140.230.831.242.790	136.213.664.750	14.670.899.525
183.784.469.293.080	177.973.724.040	19.262.796.147
237.876.442.198.920	199.827.535.320	21.623.350.764
272.295.555.417.000	204.901.326.000	23.913.167.838
2.755.654.596.990	18.991.388.340	1.945.829.334
9.932.475.420.280	41.810.419.800	4.250.633.940
21.273.724.710.720	56.902.016.560	5.753.769.948
38.046.641.423.460	79.453.039.060	8.037.810.069
61.550.720.883.450	98.316.202.050	9.872.060.869
90.452.687.183.120	110.217.934.780	10.987.376.924
1.342.718.534.588.250	1.504.018.735.140	162.086.816.056

Tabel 5 menampilkan hasil perhitungan  $X_1$ ,  $X_2$  dan  $Y$  yang di pangkatkan 2 atau dilakukan perkalian dengan variabel nya itu sendiri dan perkalian antara variabel independen yaitu jumlah tandan buah segar ( $X_1$ ) dengan variabel dependen yaitu jumlah produksi kelapa sawit ( $Y$ ), perkalian antara variabel independen yaitu daftar hadir pemanen ( $X_2$ ) dengan jumlah produksi kelapa sawit ( $X_2$ ) dan perkalian antara variabel independen tandan buah segar ( $X_1$ ) dengan variabel independen yaitu daftar hadir pemanen ( $X_2$ ). Adapun *summary output* dapat dilihat pada tabel 6:

**Tabel 6. Summary Output**

<i>Regression Statistics</i>	
Januari 2018	443.842
Februari 2018	868.348
Maret 2018	1.365.873
April 2018	1.840.094
Mei 2018	2.385.323

Tabel 6 menampilkan hasil *summary output* sebagaimana telah diuraikan perhitungannya. Dapat dilihat pada *Adjusted R Square* yaitu sebesar 0,99976472 atau 99,97% artinya variabel  $X_1$  (jumlah tandan segar) dan variabel  $X_2$  (daftar hadir pemanen) sangat mempengaruhi variabel  $Y$  (jumlah produksi kelapa sawit), sisanya  $Y$  dipengaruhi dari faktor luar selain  $X_1$  dan  $X_2$  dari penelitian ini. Adapun untuk mengetahui pengaruh secara bersama-sama atau tidaknya dapat dilihat pada tabel 7:

**Tabel 7. ANOVA (Analysis Of Varian)**

Anova					
Anova	df	SS	MS	F	Significance F
Regression	2	3161314602206010	1580657301103000	1018,804019	6,89026E-16
Residual	14	21720764538804	1551483181343		
Total	16	3183035366744810			

Tabel 7 menampilkan hasil *Analysis of Varian* yaitu untuk melihat pengaruh secara bersama-sama, dimana taraf signifikansinya 0,05 atau 5%. Dapat dilihat pada kolom *Significance F* yang nilainya 6,89026E-16 atau 0,0000000000000000689 ini menunjukkan nilai tersebut > 0,05. Yang artinya secara bersama-sama terdapat pengaruh terhadap variabel  $Y$  atau jumlah produksi kelapa sawit, tetapi mungkin tidak ada pengaruh secara *parsial* dan tdk signifikan terhadap  $Y$  atau jumlah produksi kelapa sawit. Adapun melihat pengaruh berdasarkan variabel nya atau koefisien *regresi* dapat dilihat pada tabel 8:

**Tabel 8.** Koefisien Regresi

Anova					
Anova	df	SS	MS	F	Significance F
Regression	2	3161314602206010	1580657301103000	1018,804019	6,89026E-16
Residual	14	21720764538804	1551483181343		
Total	16	3183035366744810			

Tabel 4.13 menampilkan koefisien *regresi*, dimana taraf signifikansinya 5% atau 0,05. Dilihat dari kolom *P-Value* yaitu variabel tandan buah segar ( $X_1$ ) dan daftar hadir pemanen ( $X_2$ ) tidak mempengaruhi terhadap jumlah produksi kelapa sawit secara signifikan. Tetapi besar pengaruhnya itu sebesar 99,92% karena ada faktor lain diluar dari dua variabel tandan buah segar ( $X_1$ ) dan daftar hadir pemanen ( $X_2$ ). Adapun untuk mengetahui *output* dari *regresi linear* berganda ini dapat dilihat pada tabel 9:

**Tabel 9.** Output Regresi Linear Berganda

<b>RESIDUAL OUTPUT</b>			
<i>Observation</i>	<i>Predicted Jumlah Produksi KS (Y)</i>	<i>Residuals</i>	<i>Standard Residuals</i>
Januari 2018	4.156.209,42	-386.979,1422	-0,332131445
Februari 2018	8.523.133,945	986.943,9449	-0,847061464
Maret 2018	12.768.314,81	-672.054,8086	-0,576802495
April 2018	17.102.376,04	-56.677,0382	-0,486445196
Mei 2018	22.036.960,31	-246.450,3055	-0,211520176
Juni 2018	25.663.077,57	639.402,4272	0,548778032
Juli 2018	30.306.406,1	454.243,9009	0,389862571
Agustus 2018	35.525.116,54	557.973,4579	0,478890231
September 2018	41.001.388,03	205.771,9727	0,176607303
Oktober 2018	46.370.529,2	515.330,7962	0,44229144
November 2018	51.328.612,22	-3.025.612,22	-2,596783262
Januari 2019	5.839.962,483	-653.892,4825	-0,561214369
Februari 2019	10.450.882,34	-566.622,335	-0,486313277
Maret 2019	14.290.843,36	213.876,6377	0,18356327
April 2019	18.792.452,09	600.527,906	0,515413312
Mei 2019	23.320.697,85	1.437.852,151	1,234061118
Juni 2019	27.642.067,97	2.480.352,028	2,128804407
<b>PROBABILITY OUTPUT</b>			
<i>Percentile</i>	<i>Jumlah Produksi (Y)</i>		
	2,941176471	3.769.230	
	8,823529412	5.186.070	
	14,70588235	7.536.190	
	20,58823529	9.884.260	
	26,47058824	12.096.260	
	32,35294118	14.504.720	
	38,23529412	16.535.600	
	44,11764706	19.392.980	
	50	21.790.510	
	55,88235294	24.758.550	
	61,76470588	26.302.480	
	67,64705882	30.122.420	
	73,52941176	30.760.650	
	79,41176471	36.083.090	
	85,29411765	41.207.160	
	91,17647059	46.885.860	
	97,05882353	48.303.000	

Tabel 9 menampilkan *output regresi linear* berganda (*Residual Output* dan *Probability Output*) dimana kolom pertama dari *Residual Output* ini adalah *Observation* atau jumlah data yaitu 17 Bulan. Kolom kedua *Predicted Jumlah Produksi Kelapa Sawit* adalah kolom yang memuat perkiraan atau prediksi variabel dependen untuk nilai-nilai dari variabel independen dari data asli penelitian. Kolom ketiga *Residuals* adalah selisih antara prediksi variabel dependen dengan nilai sebenarnya. Misalnya untuk observasi pertama (Januari 2018) nilai sebenarnya untuk jumlah produksi kelapa sawit adalah 3.769.230, sehingga selisihnya (*residual*) = 3.769.230 – 4.156.209,42 = -386.979,1422. Kolom keempat adalah *Standard Residuals* adalah residual yang

distandarisasikan, yang juga dikenal sebagai *Residuals Pearson*. Rata-rata dari standar residual = 0 dan standar deviasinya = 1. *Probability Output* menampilkan persentil dan nilai-nilai dari variabel dependen atau jumlah produksi kelapa sawit (Y).

#### b. Data Testing

Peneliti melakukan perhitungan data *testing* dengan jumlah data 5 (Bulan Juli, Agustus, September, November dan Desember 2019). Berdasarkan perhitungan data *training* sebelumnya diperoleh model *regresi linear* bergandanya sebagai berikut :

$$Y = (-1.641.609,438) + 8,822X_1 + 762,896X_2$$

Model *regresi linear* berganda tersebut kemudian digunakan pada perhitungan data *testing*. Adapun hasil perhitungan nya dapat dilihat pada tabel 10:

**Tabel 10.** Prediksi Jumlah Produksi Kelapa Sawit (*Data Testing*)

Bulan	Tandan Buah Segar (X <sub>1</sub> )	Pemanen (X <sub>2</sub> )	Jumlah Produksi(Y)	Prediksi (Y)	Error
Juli 2019	3.576.661	4.254	35.852.020	33.158.508,74	-2.693.511,257
Agus 2019	4.165.635	3.551	41.414.550	37.818.361,53	-3.596.188,47
Sept 2019	4.710.113	4.848	46.466.110	43.611.443,90	-2.854.666,10
Okt 2019	5.356.089	4.648	52.581.180	49.157.928,09	-3.423.251,911
Nov 2019	5.937.587	4.323	58.309.190	54.040.199,14	-4.268.990,862

Tabel 10 menampilkan prediksi jumlah produksi kelapa sawit dengan data *testing* yang menggunakan model *regresi linear* berganda. Adapun MAD (*Mean Absolute Deviation*), MSE (*Mean Square Error*), RMSE (*Root Mean Square Error*), dan MAPE (*Mean Absolute Percentage Error*), dapat dilihat pada tabel 11:

**Tabel 11.** Performa Model

Performa Model	
MAD	43.636.311,67
MSE	1.916.114.651.658.140
RMSE	43.773.446,88
MAPE	78,59%

Tabel 11 menampilkan performa model dari nilai MAD (*Mean Absolute Deviation*), MSE (*Mean Square Error*), RMSE (*Root Mean Square Error*), dan MAPE (*Mean Absolute Percentage Error*).

## 2.4. Praproses Data

Tahapan ini melakukan beberapa persiapan proses data dengan tujuan untuk mentransformasi data ke suatu format yang prosesnya lebih mudah dan efektif untuk dianalisis. Dalam tahap persiapan proses data penulis terlebih dahulu akan menyiapkan data jumlah produksi kelapa sawit (Y), data tandan buah segar (X<sub>1</sub>), dan data daftar hadir pemanen (X<sub>2</sub>) yang akan diolah. Peneliti juga menggunakan taraf signifikansi nya yaitu 5% atau 0,05.

## 2.5. Teknik Analisis Data

### Uji Validitas

Diketahui :  $r_{tabel} = N = 17 = 0,482$  (Dist nilai  $r_{tabel}$  5%). Diperoleh nilai  $r_{hitung}$  sebesar 0,996 > 0,482, maka Tandan Buah Segar atau variabel X<sub>1</sub> dinyatakan *valid* karena nilai  $r_{hitung} > r_{tabel}$ . Nilai  $r_{hitung}$  dari Daftar Hadir Pemanen yaitu 0,529 < 0,482, artinya Daftar Hadir Pemanen atau variabel X<sub>2</sub> tidak *valid* karena  $r_{hitung} < r_{tabel}$ .

### Uji Realibilitas

Uji *reliabilitas* atau *reliability* yang menggunakan rumus *Alpha* dengan signifikansinya dilakukan pada taraf  $\alpha = 0,05$  (5%). Diperoleh juga nilai  $Alpha < r_{tabel}$  yaitu 0,001 < 0,482, artinya kedua variabel antara X<sub>1</sub> dan X<sub>2</sub> tidak konsisten atau tidak reliabilitas.

## 2.6. Dasar Pengambilan Keputusan (Uji Asumsi Klasik)

- Uji *Multikolinearitas Tolerance* dan *VIF (Variance Inflation Factor)*  
Tidak ada gejala *multikolinearitas* karena nilai *tolerance* yaitu 0,740 > 0,1 dan nilai VIF yaitu 1,350 < 10,00.
- Uji *Autokorelasi Durbin Watson*  
Tidak ada gejala *autokorelasi*, karena nilai 2,4639. Nilai tersebut terletak antara dw (Dist  $dw_{tabel}$ ) s/d 4-dw.

- c. Uji *Kolmogorov Smirnov*  
Residual berdistribusi normal karena nilai signifikansinya adalah  $0,182 > 0,05$ .
- d. Uji *t parsial* pada kolom *coefficients* yaitu model 1 untuk  $X_1$  terdapat nilai sig 0,000, nilai sig < probabilitas 0,05 atau  $0,000 < 0,05$ . Variabel  $X_1$  mempunyai  $t_{hitung}$  yakni 38,250 dengan  $t_{tabel}$  2,144 atau  $38,250 > 2,144$ . Artinya Variabel  $X_1$  berpengaruh signifikan terhadap Y.
- e. Uji *t parsial* pada kolom *coefficients* yaitu model 1 untuk  $X_2$  terdapat nilai sig 0,273 > probabilitas 0,05 atau  $0,273 > 0,05$ , variabel  $X_2$  mempunyai  $t_{hitung}$  yakni 1,141 dengan  $t_{tabel}$  2,144 atau  $0,273 < 2,144$  yang berarti tidak berpengaruh terhadap Y.
- f. Uji *f simultan*  $X_1, X_2$  terhadap Y. Memiliki nilai  $f_{hitung}$  sebesar 1018,804 dengan nilai sig 0,000. Nilai  $f_{hitung}$  ( $1018,804$ ) >  $f_{tabel}$  (3,68) dan nilai sig < probabilitas 0,05, atau  $0,000 < 0,05$  artinya tandan buah segar ( $X_1$ ) dan daftar hadir pemanen ( $X_2$ ) yang secara *simultan* atau bersama-sama berpengaruh terhadap jumlah produksi kelapa sawit (Y).
- g. *Koefisien Determinasi ( $R^2$ )*  
Berdasarkan *model summary* dapat disimpulkan bahwa Tandan Buah Segar ( $X_1$ ) dan Daftar Hadir Pemanen ( $X_2$ ) berpengaruh sebesar 99,2% secara *simultan* (bersama-sama) terhadap Jumlah Produksi Kelapa Sawit (Y).

### 3. KESIMPULAN

Setelah melalui tahapan analisa dan pengujian pada perangkat lunak sistem prediksi jumlah produksi kelapa sawit pada PTPN 6 Unit Usaha Bunut Jambi menggunakan metode *regresi linier* berganda maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Persamaan *regresi* yang diperoleh dari hasil analisis *regresi linear* berganda adalah  
$$Y = (-1.641.609,438) + 8.822X_1 + 762,896X_2$$
2. Data produksi yang dianalisa dalam penelitian ini adalah data-data dalam tahun 2018 dan tahun 2019. Peneliti juga menggunakan taraf signifikansi sebesar 0,05 atau 5%.
3. Diperoleh nilai *Adjusted R Square* yaitu 0,992201239 atau 99,2%. Dan *P-Value* yaitu 0,000000000000001445 untuk variabel tandan buah segar ( $X_1$ ) berpengaruh signifikan, dan 0,272928485 untuk variabel daftar hadir pemanen ( $X_2$ ) yang artinya berpengaruh tetapi tidak signifikan terhadap jumlah produksi kelapa sawit.
4. Secara *parsial*, Tandan Buah Segar ( $X_1$ ) sangat berpengaruh terhadap Jumlah Produksi Kelapa Sawit (Y), tetapi tidak untuk Daftar Hadir Pemanen ( $X_2$ ) yang secara *parsial* tidak berpengaruh terhadap Jumlah Produksi Kelapa Sawit (Y).
5. Tandan Buah Segar ( $X_1$ ) dan Daftar Hadir Pemanen ( $X_2$ ) secara *simultan* atau bersama-sama memiliki pengaruh terhadap Jumlah Produksi Kelapa Sawit (Y).

### DAFTAR PUSTAKA

- [1] “Badan Pusat Statistik RI 2017.” 2017.
- [2] Amrin, “Data Mining Dengan Regresi Linier Berganda Untuk Peramalan Data Mining Dengan Regresi Linier Berganda Untuk,” vol. XIII, no. March 2016, pp. 74–79, 2018.
- [3] J. Jasmir, D. Zaenal A, P. Alam J, and E. Rasywir, “Prediksi Mahasiswa Drop Out dengan menggunakan Algoritma Klasifikasi Data Mining,” *Pros. Annu. Res. Semin. Comput. Sci. ICT*, vol. 4, no. 1, pp. 978–979, 2018.
- [4] F. Wulandari, P. A. Jusia, and J. Jasmir, “Klasifikasi Data Mining Untuk Mendiagnosa Penyakit ISPA Menggunakan Metode Naïve Bayes Pada Puskesmas Jambi Selatan,” vol. 2, no. 3, pp. 214–227, 2020.
- [5] T. Noviana, J. Jasmir, and Y. Novianto, “Penerapan Data Mining Menentukan Kelompok Prioritas Penerima Bantuan Beras Rastra Dengan Clustering K-Means,” *Progr. Stud. Tek. Inform. Stikom Din. Bangsa*, pp. 159–174, 2019.
- [6] J. Jasmir, D. Z. Abidin, S. Nurmaini, and R. F. Malik, “Penerapan Metode K-Nearest Neighbor dalam Memprediksi Masa Studi Mahasiswa ( Studi Kasus : Mahasiswa STIKOM Dinamika Bangsa ),” *Pros. Annu. Res. Semin.*, vol. 3, no. 1, pp. 133–138, 2017.
- [7] A. Kurniadi, Y. Novianto, and J. Jasmir, “Penerapan Metode Regresi Linier untuk Memprediksi Kebiasaan Pelanggan Studi Kasus : PT . Mensa Binasukses,” *J. Ilm. Mhs. Tek. Inform.*, vol. 2, no. 2, p. 107, 2020.

- [8] S. Sulistyono and W. Sulistiyowati, "Peramalan Produksi dengan Metode Regresi Linier Berganda," *PROZIMA (Productivity, Optim. Manuf. Syst. Eng.*, vol. 1, no. 2, p. 82, 2018.
- [9] D. N. A. JANIE, *STATISTIK DESKRIPTIF & REGRESI LINIER BERGANDA DENGAN SPSS*. 2012.
- [10] J. Pallant, "SPSS Survival Manual Survival Manual Pallant," *McGraw-Hill Educ.*, p. 361, 2010.
- [11] S. Yamin and H. Kurniawan, "Statistik SPSS complete: teknik analisis statistik terlengkap," *Anal. Koresp. Bab Anal. Diskrim.*, 2009.