

LAPORAN PRATIUM
BAGIAN 3 – PENANGANAN PELANGGARAN
ASUMSI PADA REGRESI LINEAR
MATA KULIAH STATISTIKA REGRESI
KELAS B



**“Hasil Analisis Penanganan
Pelanggaran Asumsi antara X dan Y”**

DISUSUN OLEH:
ANGELA LISANTHONI (21083010032)

DOSEN PENGAMPU:
TRIMONO, S.SI., M.SI
DR. ENG. IR. ANGGRAINI PUSPITA SARI., ST., MT.

PROGRAM STUDI SAINS DATA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAWA TIMUR
2022

1. BAB 1: PENDAHULUAN

1.1 Tujuan Pratikum

1.1.1 Tujuan Instruksional Umum (TIU)

Tujuan instruksional umum dari laporan ini adalah mahasiswa mampu melakukan pengolahan, analisis dan membuat model regresi dari data pengamatan berdasarkan model yang dibangun dengan menggunakan SPSS.

1.1.2 Tujuan Instruksional Khusus (TIK)

Tujuan instruksional khusus dari laporan ini adalah mahasiswa mampu menganalisis pelanggaran asumsi serta membuat transformasi variabel terhadap pelanggaran asumsi yang mungkin terjadi.

1.2 Permasalahan

Adapun permasalahan yang dibahas dalam laporan ini adalah:

- Bagaimana bentuk estimasi model regresi menggunakan model linier antara data uji X dan Y menggunakan SSPS?
- Bagaimana bentuk estimasi model regresi menggunakan model kuadratik antara data uji X dan Y menggunakan SSPS?
- Bagaimana cara penerapan metode curve estimation untuk menganalisis permasalahan pelanggaran asumsi dengan SPSS?
- Manakah model yang terbaik antara model linier dengan model kuadratik terhadap data uji X dan Y menggunakan perbandingan uji F, uji T, MSE, dan koefisien determinasi?

2. BAB II: TINJAUAN PUSTAKA

2.1 Model Linier

Berdasarkan (Ginting, Buulolo, & Siagian, 2019), Regresi linear sederhana adalah analisis regresi yang melibatkan ubungan antara satu variable terikat dan satu variable bebas yang memiliki model dengan rumus berikut:

$$Y = a + bx \quad (1)$$

dimana,

a = konstanta

b = koefisien regresi

Y = Variabel dependen (variable terikat)

X = variable independent (varianle bebas)

2.2 Model Kuadratik

Berdasarkan (Astutik, Wuryantini, & Nafidah, 2019),terkadang ada kemungkinan bahwa garis regresi yang dihasilkan tidak memiliki bentuk garis lurus, melainkan garis lengkung yang disebut regresi non linear. Bentuk persamaan model regresi polynomial kuadratik secara umum digambarkan sebagai:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 \quad (2)$$

dimana,

i : 1, 2, ..., n

Y_i : variabel respon pada pengamatan ke-i

β_0 : parameter intersep

β_1 dan β_2 : parameter slope
 X_i : variabel predictor pada pengamatan ke-i

2.3 Uji Hipotesis

2.3.1 Uji Koefisien Regresi Secara Bersama – sama (Uji F)

Uji F digunakan untuk melihat apakah model regresi yang ada layak atau tidak. Kelayakan yang dimaksud adalah apakah model regresi dapat digunakan untuk menjelaskan pengaruh variable independent pada dependen. Perhitungannya akan mengasumsikan H_0 : model tidak sesuai dan H_1 : model sesuai. H_0 ditolak apabila $|f_{hitung}| > f_{tabel}$ atau $sig < \alpha$ (Nanincova, 2019).

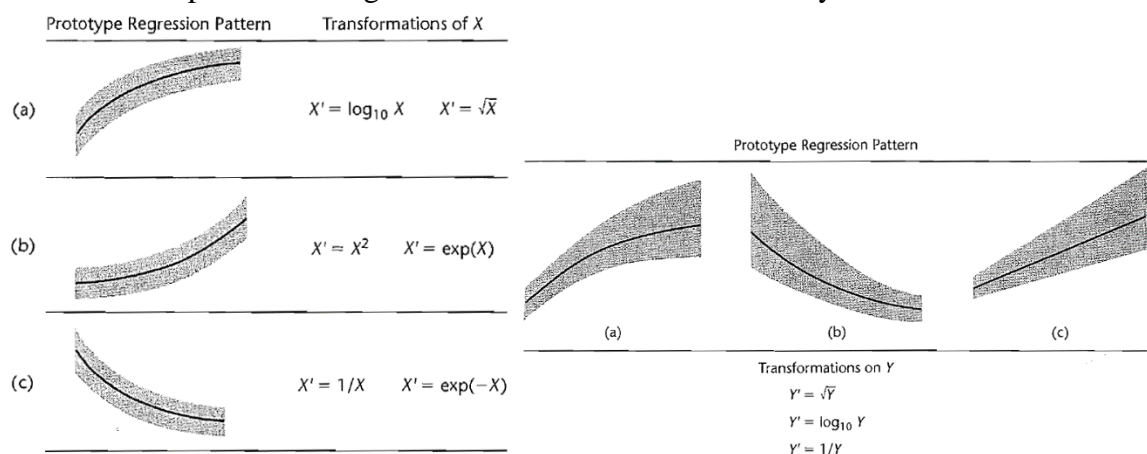
2.3.2 Uji Koefisien Regresi secara Individu (Uji T)

Uji t adalah suatu uji yang menunjukkan seberapa jauh pengaruh satu variable independent secara individual dalam menerangkan variable dependen. Perhitungannya akan mengasumsikan H_0 : variable independent tidak mempunyai pengaruh signifikan terhadap variable dependen dan H_1 : variable independent mempunyai pengaruh signifikan terhadap variable dependen. H_0 ditolak apabila $|t_{hitung}| > t_{tabel}$ atau $sig < \alpha$ (Magdalena & Krisanti, 2019).

2.4 Transformasi Data

2.4.1 Transformasi untuk kasus hubungan non linier

Berdasarkan (Kutner, 2004), ada model transformasi khusus untuk fungsi non-linear. Berikut adalah beberapa contoh fungsi non-linier beserta transformasinya:



2.4.2 Transformasi pada variabel Y dengan metode BOX-COX

Berdasarkan (Risma, Sahrman, & Thamrin, 2020), Transformasi Box Cox adalah transformasi pangkat pada respon (Y). Box Cox mempertimbangkan kelas transformasi berparameter tunggal yakni λ yang dipangkatkan pada variable respon Y sehingga transformasinya menjadi Y^λ . Perubahannya ditunjukkan pada tabel dibawah ini.

λ	Transformasi
2	$Y^2 = Y^2$
1	$Y^1 = Y^1$
0.5	$Y^{0.5} = \sqrt{Y}$

0	$Y^0 = \ln Y$
-0.5	$Y^{-0.5} = \frac{1}{\sqrt{Y}}$
-1	$Y^{-1} = \frac{1}{Y^1}$
-2	$Y^{-2} = \frac{1}{Y^2}$

2.4.3 Transformasi menstabilkan variansi

Berdasarkan (Sembiring, 2003), jika suatu kesamaan variansi tidak terpenuhi maka diperlukan transformasi untuk menstabilkannya. Berikut menunjukkan beberapa transformasi untuk menstabilkan variansi.

Hubungan var (y) dengan E(y)	Transformasi
Var(y) = k E(y) (Poisson)	$Y^* = \sqrt{Y}$ atau $\sqrt{Y} + \sqrt{Y + 1}$
Var(y) = k E(y) (1-E(y)) (binom)	$Y^* = \arcsin \sqrt{Y}$
Var(y) = k (E(y)) ²	$Y^* = \log y$
Var(y) = k (E(y)) ³	$Y^* = 1/\sqrt{Y}$
Var(y) = k (E(y)) ⁴	$Y^* = 1/y$

2.5 MSE (Mean Square Error)

Mean Squared Error adalah metrik evaluasi model yang sering digunakan dengan model regresi. MSE adalah rata-rata kesalahan prediksi kuadrat atas semua contoh dalam set pengujian. Kesalahan prediksi adalah perbedaan antara nilai sebenarnya dan nilai prediksi untuk suatu contoh (Webb, 2011).

2.6 Koefisien Determinasi

Koefisien determinasi, dilambangkan R^2 , adalah hasil bagi dari variasi yang dijelaskan (jumlah kuadrat residual) dengan variasi total (jumlah kuadrat total) dalam model regresi linier sederhana atau ganda. Biasanya sering dinyatakan sebagai persentase dan berada di antara 0 hingga 1. Rumus dari koefisien determinasi adalah (Dodge, 2008):

$$R^2 = \frac{\text{jumlah kuadrat residual}}{\text{jumlah kuadrat total}} \quad (3)$$

3. BAB III: ANALISIS DAN PEMBAHASAN

3.1 Model Linear

3.1.1 Model Awal

Coefficients					
	Unstandardized Coefficients	Std. Error	Standardized Coefficients	t	Sig.
	B		Beta		
X_032	-.218	.089	-.399	-2.463	.019
(Constant)	33.323	8.063		4.133	.000

Berdasarkan table Coefficients diatas, maka didapatkan:

$$\beta_0 = 33.323$$

$$\beta_1 = -0.218$$

Maka persamaan regresi awalnya adalah:

$$Y = 33.323 - 0.218X$$

3.1.2 Uji Kecocokan Model (Uji F)

- $H_0: \beta_0 = \beta_1 = 0$ (Model tidak sesuai)
 $H_1: \beta_1 \neq 0$ (Model sesuai)
- taraf signifikansi ($\alpha = 5\%$)
- Statistik uji

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	73.639	1	73.639	6.067	.019
Residual	388.379	32	12.137		
Total	462.018	33			

The independent variable is X_032.

Berdasarkan tabel ANOVA dapat diketahui:

$F = 6.607$

$\text{sig} = 0.019$

- Keputusan
Berdasarkan landasan teori diatas, H_0 ditolak apabila $\text{sig} < \alpha$. dan Diperhitungan ini, $\text{sig} < \alpha$ ($0.019 < 0.05$) sehingga H_0 ditolak yang artinya model yang digunakan cocok.

3.1.3 Uji Signifikansi Model (Uji T)

- $H_0: \beta_1 = 0$ (X tidak mempengaruhi Y)
- $H_1: \beta_1 \neq 0$ (X mempengaruhi Y)
- taraf signifikansi ($\alpha = 5\%$)
- Statistik uji

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
X_032	-.218	.089	-.399	-2.463	.019
(Constant)	33.323	8.063		4.133	.000

- Berdasarkan tabel Coefficients dapat diketahui:
 $t = -2.463$
 $\text{sig} = 0.019$
- Keputusan
Berdasarkan landasan teori diatas, H_0 ditolak apabila $\text{sig} < \alpha$. dan Diperhitungan ini, $\text{sig} < \alpha$ ($0.019 < 0.05$) sehingga H_0 ditolak yang artinya X berpengaruh signifikan terhadap Y.

3.1.4 Koefisien Determinasi

Model Summary			
R	R Square	Adjusted R Square	Std. Error of the Estimate
.399	.159	.133	3.484

The independent variable is X_032.

Berdasarkan table model summary, diperoleh nilai $R^2 = 0.159$ artinya sebesar 15.9% variabel Y dipengaruhi oleh variabel X dan 84.1% dipengaruhi oleh factor lain.

3.1.5 Mean Square Error (MSE)

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	73.639	1	73.639	6.067	.019
Residual	388.379	32	12.137		
Total	462.018	33			

The independent variable is X_032.

Berdasarkan table ANOVA, diperoleh nilai $MSE = 12.137$ artinya rata – rata kuadrat perbedaan nilai asli dan prediksi sebesar 12.137

3.1.6 Model Akhir

Karena berdasarkan Uji F, model linear cocok digunakan, sehingga dapat dilanjutkan dengan uji t. Berdasarkan uji t, disimpulkan bahwa variabel X berpengaruh signifikan terhadap variabel Y. Maka model akhir sama dengan model awal yakni: $Y = 33.323 - 0.218X$

3.2 Model Kuadratik

3.2.1 Model Awal

Coefficients					
	Unstandardized Coefficients B	Standard Error	Standardized Coefficients Beta	t	Sig.
X_032	3.725	1.020	.6810	3.651	.001
X_032 ** 2	-.022	.006	-.7228	-3.875	.001
(Constant)	-143.221	46.053		-3.110	.004

Berdasarkan table Coefficients diatas, maka didapatkan:

$$\beta_0 = -143.221$$

$$\beta_1 = 3.725$$

$$\beta_2 = -0.022$$

Maka persamaan regresi awalnya adalah:

$$Y = -143.221 + 0.218X_1 - 0.022X_1^2$$

3.2.2 Uji Kecocokan Model (Uji F)

- $H_0: \beta_0 = \beta_1 = \beta_2 = 0$ (Model tidak sesuai)
 $H_1: \beta_1 \neq \beta_2 \neq 0$ (Model sesuai)
- taraf signifikansi ($\alpha = 5\%$)
- Statistik uji

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	200.375	2	100.188	11.870	.000
Residual	261.643	31	8.440		
Total	462.018	33			

The independent variable is X_032.

Berdasarkan tabel ANOVA dapat diketahui:

$$F = 11.870$$

$$\text{sig} = 0.000$$

- Keputusan

Berdasarkan landasan teori diatas, H_0 ditolak apabila $\text{sig} < \alpha$. dan Diperhitungan ini, $\text{sig} < \alpha$ ($0.000 < 0.05$) sehingga H_0 ditolak yang artinya model yang digunakan cocok.

3.2.3 Uji Signifikansi Model (Uji T)

- $H_0: \beta_1 = 0$ (X tidak mempengaruhi Y)
- $H_1: \beta_1 \neq 0$ (X mempengaruhi Y)
- taraf signifikansi ($\alpha = 5\%$)
- Statistik uji

Coefficients					
Unstandardized Coefficients			Standardized Coefficients		
B	Std. Error		Beta	t	Sig.
X_032	3.725	1.020	6.810	3.651	.001
X_032 ** 2	-.022	.006	-.7.228	-3.875	.001
(Constant)	-143.221	46.053		-3.110	.004

- Berdasarkan tabel Coefficients dapat diketahui:
 $t = 3.651$
 $\text{sig} = 0.001$
- Keputusan
 Berdasarkan landasan teori diatas, H_0 ditolak apabila $\text{sig} < \alpha$. dan Diperhitungan ini, $\text{sig} < \alpha$ ($0.001 < 0.05$) sehingga H_0 ditolak yang artinya X berpengaruh signifikan terhadap Y.

3.2.4 Koefisien Determinasi

Model Summary			
R	R Square	Adjusted R Square	Std. Error of the Estimate
.659	.434	.397	2.905

The independent variable is X_032.

Berdasarkan table model summary, diperoleh nilai $R^2 = 0.434$ artinya sebesar 43.4% variabel Y dipengaruhi oleh variabel X dan 56.6% dipengaruhi oleh factor lain.

3.2.5 Mean Square Error (MSE)

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	200.375	2	100.188	11.870	.000
Residual	261.643	31	8.440		
Total	462.018	33			

The independent variable is X_032.

Berdasarkan table ANOVA, diperoleh nilai $MSE = 8.440$ artinya rata – rata kuadrat perbedaan nilai asli dan prediksi sebesar 8.440

3.2.6 Model Akhir

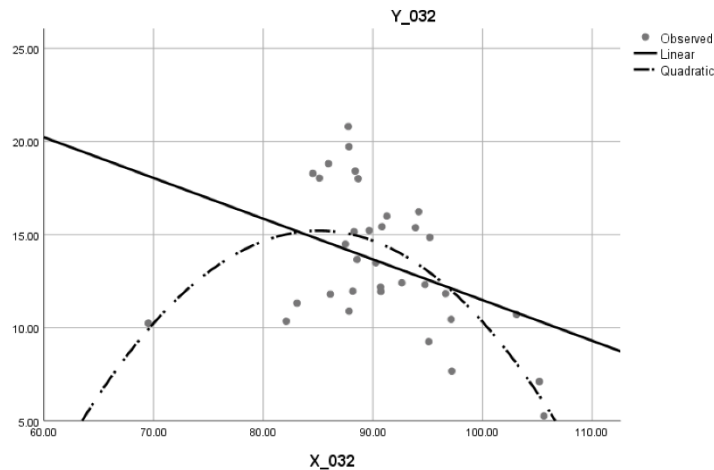
Karena berdasarkan Uji F, model linear cocok digunakan, sehingga dapat dilanjutkan dengan uji t. Berdasarkan uji t, disimpulkan bahwa variabel X berpengaruh signifikan terhadap variabel Y. Maka model akhir sama dengan model awal yakni: $Y = -143.221 + 0.218X_1 - 0.022X_1^2$

3.3 Pemilihan Model Terbaik Antara Model Linier dan Model Kuadratik

Model	Uji F	Uji t	MSE	R^2
Linier	Model cocok	β_1 signifikan	12.137	15.9 %
Kuadratik	Model cocok	β_1 signifikan	8.440	43.4%
		β_2 signifikan		

Berdasarkan tabel diatas, maka model terbaiknya adalah kuadratik jika dibandingkan dengan model linier karena model cocok dari uji F, signifikan dari uji t, nilai MSE lebih kecil dibandingkan model linier yaitu 8.440 dan nilai R^2 lebih besar dibandingkan model linier yaitu 43.4%

3.4 Curve Estimation



Secara visual data observasi tersebar acak dan mendekati garis kuadratik sehingga bisa dikatakan secara visual, model terbaiknya adalah model kuadratik.

4. BAB IV: KESIMPULAN

Model regresi yang paling terbaik adalah model kuadratik jika dibandingkan dengan model linier. Hal ini dibuktikan bahwa uji F model kuadratik menunjukkan model cocok, uji t model kuadratik menunjukkan signifikan berpengaruh, nilai MSE lebih kecil yaitu 8.440 dan nilai koefisien determinasi lebih besar yaitu 43.4%. Menurut bentuk visual curve estimation, didapatkan bahwa data observasi mendekati garis kuadratik dibanding garis linier yang memperkuat perhitungan ini. Sehingga model regresi yang digunakan untuk data uji coba adalah $Y = -143.221 + 0.218X_1 - 0.022X_1^2$.

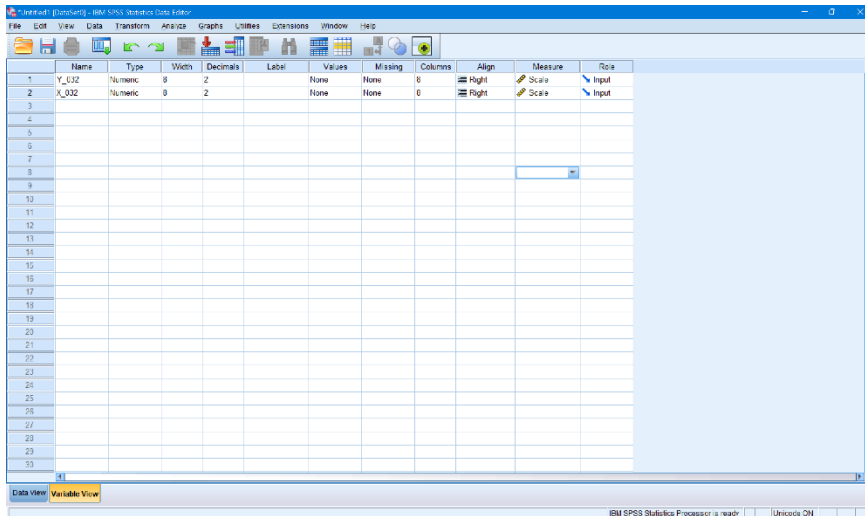
5. DAFTAR PUSTAKA

- Astutik, S., Wuryantini, S., & Nafidah, Q. A. (2019). Penerapan Analisis Regresi Nonlinear Kuadratik Terhadap Pengujian Toksisitas (LD50) Biopestisida Crude Extract Tembakau pada Kutu Daun Hijau (*Aphis Gossypii*). *Prosiding Seminar Nasional Integrasi Matematika dan Nilai Islami*, (pp. 430-436).
- Dodge, Y. (2008). *The Concise Encyclopedia of Statistics*. New York, NY: Springer.
- Ginting, F., Buulolo, E., & Siagian, E. R. (2019). IMPLEMENTASI ALGORITMA REGRESI LINEAR SEDERHANA DALAM MEMPREDIKSI BESARAN PENDAPATAN DAERAH (STUDI KASUS: DINAS PENDAPATAN KAB. DELI SERDANG). *KOMIK*, III(1), 274 - 279.
- Kutner, M. H. (2004). *Applied Linear Regression Models Fourth Edition*. New York: The McGraw-Hill.
- Magdalena, R., & Krisanti, M. A. (2019). Analisis Penyebab dan Solusi Rekonsiliasi Finished Goods Menggunakan Hipotesis Statistik dengan Metode Pengujian Independent Sample T-Test di PT. Merck, Tbk. *TEKNO*, 35 - 48.
- Nanincova, N. (2019). PENGARUH KUALITAS LAYANAN TERHADAP KEPUASAN PELANGGAN NOACH CAFE AND BISTRO. *AGORA*.
- Risma, Sahriman, S., & Thamrin, S. A. (2020). Perbandingan Estimasi Metode Kuadrat Terkecil Terboboti dan Metode Transformasi Box-Cox Pada Data Heteroskedastisitas. *Journal of Statistics and Its Application*, 83-93.
- Sembiring, R. (2003). *Analisi Regresi Edisi Kedua*. Bandung: ITB.
- Webb, G. I. (2011). *Encyclopedia of Machine Learning*. Springer, Boston, MA.

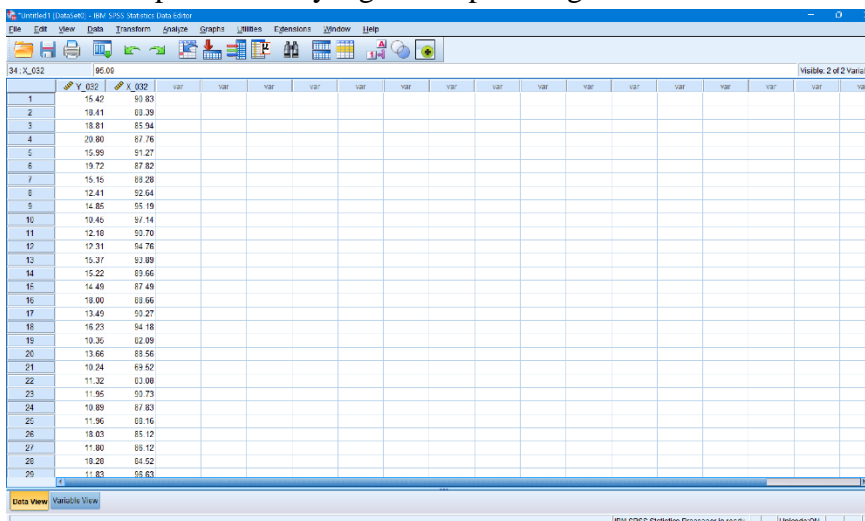
6. LAMPIRAN

6.1 Langkah Analisis

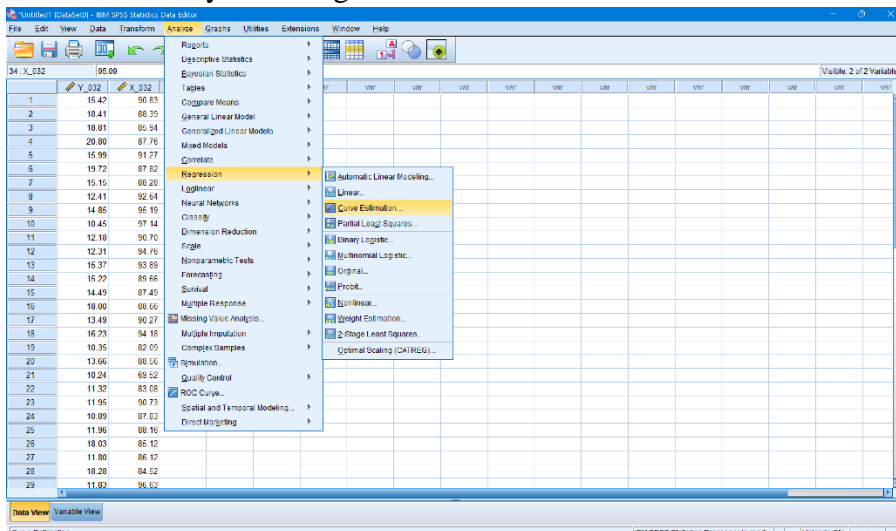
a. Definisikan variable Y dan variable X pada bagian Variabel View



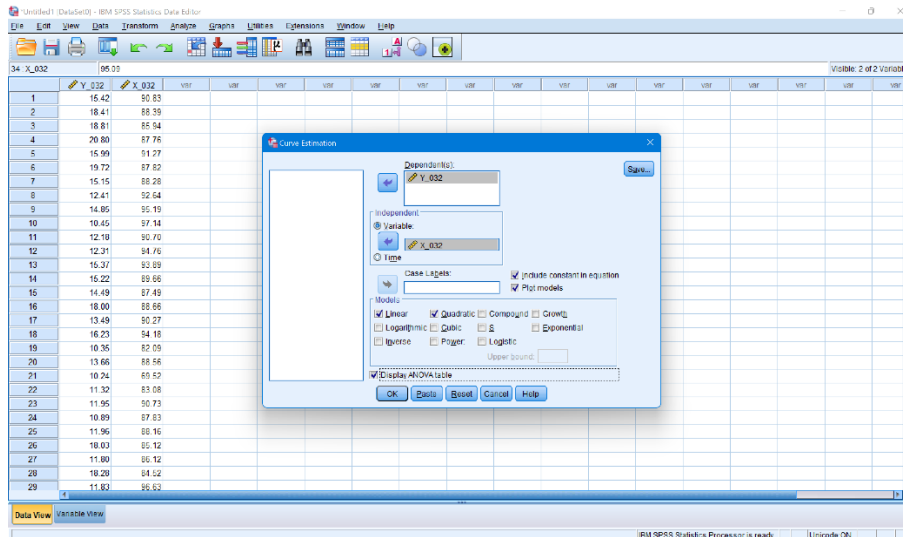
b. Isikan data pada kolom yang sesuai pada bagian Data View



c. klik menu Analyze → Regression → Curve Estimation



- d. Lalu akan muncul kotak seperti berikut. Masukkan bagian data Y ke Dependent dan Data X ke bagian independent. Karena ingin menguji untuk model linear dan kuadrat, maka hanya centang yang linear dan kuadrat, serta centang untuk menampilkan table Anova. Lalu klik ok



6.2 Output Analisis pada SPSS

Y_032

Linear

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.399	.159	.133	3.484

The independent variable is X_032.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	73.639	1	73.639	6.067	.019
Residual	388.379	32	12.137		
Total	462.018	33			

The independent variable is X_032.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		t	Sig.
	B	Std. Error	Beta			
X_032	-.218	.089	-.399		-2.463	.019
(Constant)	33.323	8.063			4.133	.000

Quadratic

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.659	.434	.397	2.905

The independent variable is X_032.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	200.375	2	100.188	11.870	.000
Residual	261.643	31	8.440		
Total	462.018	33			

The independent variable is X_032.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		t	Sig.
	B	Std. Error	Beta			
X_032	3.725	1.020	6.810		3.651	.001
X_032 ** 2	-.022	.006	-.7228		-3.875	.001
(Constant)	-143.221	46.053			-3.110	.004

