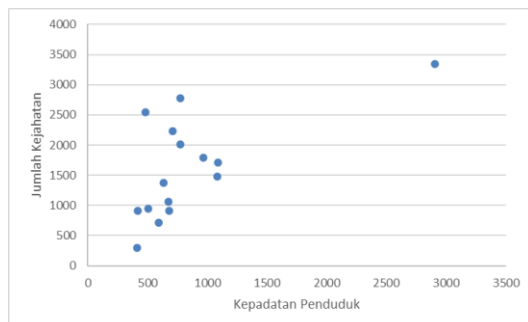


REGRESI LINEAR SEDERHANA

Data yang digunakan:

Kabupaten (Provinsi Jawa Timur)	Jumlah Kejahatan yang Dilaporkan (2022) (Y)	Kepadatan Penduduk per km ² (2022) (X)
Kabupaten Pacitan	300	414
Kabupaten Ponorogo	911	680
Kabupaten Trenggalek	709	592
Kabupaten Tulungagung	1787	966
Kabupaten Blitar	2238	711
Kabupaten Kediri	1710	1087
Kabupaten Malang	2010	773
Kabupaten Lumajang	1378	633
Kabupaten Jember	2273	775
Kabupaten Banyuwangi	2545	482
Kabupaten Bondowoso	947	503
Kabupaten Situbondo	911	418
Kabupaten Probolinggo	1057	673
Kabupaten Pasuruan	1480	1048
Kabupaten Sidoarjo	3344	2905



Model Regresi Linear Sederhana

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n$$

y_i merepresentasikan nilai ke-i dari peubah respon Y

x_i merepresentasikan nilai ke-i dari peubah respon X

ε_i merepresentasikan galat ke-i

i	x	y	$y_i - \bar{y}$	$x_i - \bar{x}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})(x_i - \bar{x})$
1	414	300	-1306,67	-432,4	1707377,8	186969,76	565002,67
2	680	911	-695,667	-166,4	483952,11	27688,96	115758,93
3	592	709	897,667	-254,4	805805,44	64719,36	228366,4
4	966	1787	180,3333	119,6	32520,111	14304,16	21567,867
5	711	2238	631,3333	-135,4	398581,78	18333,16	-85482,533
6	1087	1710	103,3333	240,6	10677,778	57888,36	24862
7	773	2010	403,3333	-73,4	162677,78	5387,56	-29604,67
8	633	1378	-228,667	-213,4	52288,444	45539,56	48797,467
9	775	2273	1166,333	-71,4	1360333,4	5097,96	-83276,2
10	482	2545	938,3333	-364,4	880469,44	132787,36	-341928,7
11	503	947	-659,667	-343,4	435160,11	117923,56	226529,53
12	418	911	-695,667	-428,4	483952,11	183526,56	298023,6
13	673	1057	-549,667	-173,4	302133,44	30067,56	95312,2
14	1048	1480	-126,667	237,6	16044,444	56453,76	-30096
15	2905	3344	1737,333	2058,6	3018327,1	4237833,96	3576474,4
Total	12696	24100			10150301	5184521,6	4630307
Rataan	846,4	1606,667					

REGRESI LINEAR SEDERHANA

A. Dugaan Parameter Regresi

- Penduga bagi slope β_1 :
$$\hat{\beta}_1 = \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{\sum(x_i - \bar{x})^2} = \frac{4630307}{5184521,6} = 0,8931$$
- Penduga bagi intersep β_0 :
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 1606,667 - (0,8931)(846,4)$$

$$\hat{\beta}_0 = 850,745$$

```
##{r}
(b1 <- model$coefficients[[2]])
[1] 0.8931021
```

```
##{r}
(b0 <- model$coefficients[[1]])
[1] 850.7451
```

Penduga bagi persamaan regresi: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

Berdasarkan hasil perhitungan, didapatkan dugaan persamaan regresi linear sederhana, yakni

$$\hat{y} = 850,745 + 0,8931x$$

- Interpretasi dugaan parameter regresi:
 - $\hat{\beta}_1$: Setiap penambahan 1 satuan kepadatan penduduk (1 penduduk per km²) maka akan meningkatkan dugaan rata-rata jumlah kejahatan sebesar 0,8931 kasus
 - $\hat{\beta}_0$: Nilai 850,745 merupakan nilai dugaan rata-rata jumlah kejahatan yang tidak dapat dijelaskan oleh kepadatan penduduk.

B. Uji Hipotesis bagi Parameter Regresi β_1 dan β_0

- Uji hipotesis bagi parameter β_1
 Dengan nilai slope ($\hat{\beta}_1$) sebesar 0,8931 akan diuji apakah ada hubungan linear antara kepadatan penduduk (X) dengan jumlah kejahatan (Y)

Hipotesis uji:

$H_0 : \beta_1 = 0$ (tidak ada hubungan linear antara kepadatan penduduk dengan jumlah kejahatan)

$H_1 : \beta_1 \neq 0$ (ada hubungan linear antara kepadatan penduduk dengan jumlah kejahatan)

i	x	y	$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
1	414	300	1220,489	-920,489	847300,605
2	680	911	1458,054	-547,054	299268,606
3	592	709	1379,461	-670,461	449518,621
4	966	1787	1713,482	73,51833	5404,94413
5	711	2238	1485,741	752,2594	565894,136
6	1087	1710	1821,547	-111,547	12442,739
7	773	2010	1541,113	468,887	219855,043
8	633	1378	1416,079	-38,0787	1449,98616
9	775	2273	1542,899	1230,101	1513148,03
10	482	2545	1281,22	1263,78	1597139,2
11	503	947	1299,975	-352,975	124591,643
12	418	911	1224,062	-313,062	98007,6516
13	673	1057	1451,803	-394,803	155869,225
14	1048	1480	1818,868	-338,868	114831,331
15	2905	3344	3445,207	-101,207	10242,7756
Total	12696	24100			6014964,54
Rataan	846,4	1606,667			

REGRESI LINEAR SEDERHANA

Statistik uji:

$$S_e = \sqrt{\frac{JKG}{n-2}} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{6014964,54}{15-2}} = 680,2129$$

$$S_{\hat{\beta}_1} = \sqrt{\frac{S_e^2}{\sum(x_i - \bar{x})^2}} = \sqrt{\frac{(680,2129)^2}{5184521,6}} = 0,2987$$

```
{r}
(se_b1 <- sqrt(KTG/sum((x-mean(x))^2)))
[1] 0.298738
```

$$t_h = \frac{\hat{\beta}_1 - \beta_1}{S_{\hat{\beta}_1}} = \frac{0,8931 - 0}{0,2987} = 2,9896$$

$$t_{(13;0,025)} = 2,1604$$

Kesimpulan: Karena nilai $t_h > t_{(13;0,025)}$ maka tolak H_0 yang menunjukkan adanya hubungan linear antara kepadatan penduduk dengan jumlah kasus kejahatan. Sehingga cukup bukti untuk menyatakan bahwa kepadatan penduduk mempengaruhi banyaknya jumlah kasus kejahatan.

- Uji hipotesis bagi parameter β_0

Dengan nilai intersep ($\hat{\beta}_0$) sebesar 850,745 akan diuji apakah ada jumlah kejahatan (Y) yang tidak dapat dijelaskan oleh kepadatan penduduk (X)

Hipotesis uji:

$H_0 : \beta_0 = 0$ (Semua jumlah kejahatan dapat dijelaskan oleh kepadatan penduduk)

$H_1 : \beta_0 \neq 0$ (ada jumlah kejahatan yang tidak dapat dijelaskan oleh kepadatan penduduk)

Statistik uji:

$$S_{\hat{\beta}_0} = \sqrt{S_e^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2} \right)}$$

$$S_{\hat{\beta}_0} = \sqrt{(680,2129)^2 \left(\frac{1}{15} + \frac{(846,4)^2}{5184521,6} \right)} = 307,864$$

$$t_h = \frac{\hat{\beta}_0 - \beta_0}{S_{\hat{\beta}_0}} = \frac{850,745 - 0}{307,864} = 2,76338$$

$$t_{(13;0,025)} = 2,1604$$

```
{r}
(se_b0 <- sqrt(KTG*((1/n)+mean(x)^2/sum((x-mean(x))^2))))
[1] 307.8637
```

Kesimpulan: Karena nilai $t_h > t_{(13;0,025)}$ maka tolak H_0 . Artinya terdapat cukup bukti untuk menyatakan bahwa ada jumlah kasus kejahatan yang tidak dapat dijelaskan oleh besarnya kepadatan penduduk pada taraf nyata 5%

C. Penduga Selang Kepercayaan bagi Parameter Regresi β_1 dan β_0

- Selang kepercayaan $(1 - \alpha)$ bagi parameter regresi β_1

$$\hat{\beta}_1 - t_{(n-2; \frac{\alpha}{2})} S_{\hat{\beta}_1} < \beta_1 < \hat{\beta}_1 + t_{(n-2; \frac{\alpha}{2})} S_{\hat{\beta}_1}$$

$$= 0,8931 - 2,1604 (0,2987) < \beta_1 < 0,8931 + 2,1604 (0,2987)$$

$$= 0,247717 < \beta_1 < 1,538486$$

```
{r}
(bb.b1 <- b1 - abs(qt(0.025, df=n-2))*se_b1)
[1] 0.2477178
```

```
{r}
(ba.b1 <- b1 + abs(qt(0.025, df=n-2))*se_b1)
[1] 1.538486
```

REGRESI LINEAR SEDERHANA

- Selang kepercayaan $(1 - \alpha)$ bagi parameter regresi β_0

$$\hat{\beta}_0 - t_{(n-2, \frac{\alpha}{2})} s_{\hat{\beta}_0} < \beta_0 < \hat{\beta}_0 + t_{(n-2, \frac{\alpha}{2})} s_{\hat{\beta}_0}$$

$$= 850,745 - 2,1604 (307,864) < \beta_0 < 850,745 + 2,1604 (307,864)$$

$$= 185,64604 < \beta_0 < 1515,8441$$

```
{r}
(bb.b0 <- b0 - abs(qt(0.025, df=n-2))*se_b0)
[1] 185.646
```

```
{r}
(ba.b0 <- b0 + abs(qt(0.025, df=n-2))*se_b0)
[1] 1515.844
```

D. Tabel Sidik Ragam

x	y	$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
414	300	1220,489	-386,17734	149132,936	-920,489	847300,605	-1306,67	1707377,8
680	911	1458,054	-148,61219	22085,5816	-547,054	299268,606	-695,667	483952,11
592	709	1379,461	-227,20517	51622,1883	-670,461	449518,621	897,667	805805,44
966	1787	1713,482	106,815008	11409,446	73,51833	5404,94413	180,3333	32520,111
711	2238	1485,741	-120,92602	14623,1026	752,2594	565894,136	631,3333	398581,78
1087	1710	1821,547	214,880359	46173,5688	-111,547	12442,739	103,3333	10677,778
773	2010	1541,113	-65,553692	4297,28658	468,887	219855,043	403,3333	162677,78
633	1378	1416,079	-190,58798	36323,7792	-38,0787	1449,98616	-228,667	52288,444
775	2273	1542,899	-63,767488	4066,29255	1230,101	1513148,03	1166,333	1360333,4
482	2545	1281,22	-325,4464	105915,357	1263,78	1597139,2	938,3333	880469,44
503	947	1299,975	-306,69125	94059,5244	-352,975	124591,643	-659,667	435160,11
418	911	1224,062	-382,60493	146386,532	-313,062	98007,6516	-695,667	483952,11
673	1057	1451,803	-154,8639	23982,8275	-394,803	155869,225	-549,667	302133,44
1048	1480	1818,868	212,201053	45029,2869	-338,868	114831,331	-126,667	16044,444
2905	3344	3445,207	1838,539931	3380229,08	-101,207	10242,7756	1737,333	3018327,1
Jumlah				4135336,79		6014964,54		10150301

$$JKT = \sum (y_i - \bar{y})^2 = 10150301$$

$$KTS = \frac{JKS}{n-2} = 462689,58$$

$$JKR = \sum (\hat{y}_i - \bar{y})^2 = 4135336,79$$

$$KTR = \frac{JKR}{1} = 4135336,79$$

$$JKS = JKT - JKR = 6014964,5$$

Sumber Keragaman	Derajat Bebas (db)	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)
Regresi	1	4135336,79	4135336,79
Sisaan	13	6014964,5	462689,58
Total	14	10150301	

E. Kebaikan Model Regresi

Kebaikan model diukur dengan seberapa besar keragaman data yang dapat diterangkan oleh model regresi
Diukur dengan koefisien Determinasi (R^2)

$$R^2 = \frac{JKR}{JKT} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = \frac{4135336,79}{10150301} = 0,40741$$

Multiple R-squared: 0.40741

Interpretasi:

- Nilai R^2 sebesar 0,40741 atau 40,741% menunjukkan bahwa 40,741% keragaman dalam jumlah kasus kejahatan (Y) mampu dijelaskan oleh kepadatan penduduk (X), sementara sisanya sebesar 59,259% dipengaruhi oleh faktor lain yang tidak termasuk dalam model (di luar model).

- Koefisien determinasi sebesar 40,741% menunjukkan bahwa kepadatan penduduk memiliki pengaruh yang cukup terhadap jumlah kasus kejahatan, tetapi masih ada faktor lain yang lebih dominan dalam menjelaskan keragaman jumlah kejahatan.

F. Selang Kepercayaan individu dan rata-rata dengan $x = 39$

Untuk suatu nilai $x = x_0$, nilai prediksi bagi rata-rata Y adalah

$$\hat{y}(x_0) = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

Prediksi rata-rata jumlah kasus kejahatan jika kepadatan penduduknya sebesar 39 penduduk per km² yakni:

$$\hat{y}(39) = 850,745 + 0,8931(39) = 885,5759 \text{ kasus kejahatan}$$

- Selang Kepercayaan $(1 - \alpha)$ bagi individu y untuk $x = 39$

$$\hat{y}(x_0) \pm t_{(n-2; \frac{\alpha}{2})} \times s_e \sqrt{\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]}$$

$$\hat{y}(39) \pm t_{(n-2; \frac{\alpha}{2})} \times s_e \sqrt{\left[1 + \frac{1}{n} + \frac{(39 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]}$$

$$= 885,5759 \pm 2,1604 \times 680,2129 \sqrt{\left[1 + \frac{1}{15} + \frac{(39 - 846,4)^2}{5184521,6}\right]}$$

$$= 885,5759 \pm 2,1604 \times 680,2129 (1,091973)$$

$$= 885,5759 \pm 2,1604 \times 742,7742$$

$$= 885,5759 \pm 1604,666$$

$$= (-719,09 ; 2490,242)$$

```
{r}
predict(model, amatan.diduga, interval = "prediction")

      fit      lwr      upr
1 885.5761 -719.0901 2490.242
```

Selang kepercayaan bagi individu untuk jumlah kejahatan pada kepadatan penduduk 39 penduduk per km² mencakup nilai 0 yang menunjukkan tingginya ketidakpastian dalam memprediksi jumlah kejahatan. Hal ini kemungkinan terjadi karena nilai $x = 39$ berada di luar selang pengamatan yang digunakan dalam model, sehingga hasil prediksi menjadi kurang akurat.

- Selang Kepercayaan $(1 - \alpha)$ bagi rata-rata y untuk $x = 39$

$$\hat{y}(x_0) \pm t_{(n-2; \frac{\alpha}{2})} \times s_e \sqrt{\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]}$$

$$\hat{y}(39) \pm t_{(n-2; \frac{\alpha}{2})} \times s_e \sqrt{\left[\frac{1}{n} + \frac{(39 - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]}$$

$$= 885,5759 \pm 2,1604 \times 680,2129 \sqrt{\left[\frac{1}{15} + \frac{(39 - 846,4)^2}{5184521,6}\right]}$$

$$= 885,5759 \pm 2,1604 \times 680,2129 (0,43864)$$

$$= 885,5759 \pm 2,1604 \times 298,3688$$

$$= 885,5759 \pm 644,5866$$

$$= (240,9895 ; 1530,163)$$

```
{r}
amatan.diduga <- data.frame(x=39)
predict(model, amatan.diduga, interval = "confidence")

      fit      lwr      upr
1 885.5761 240.9895 1530.163
```