

In [16]: linear <- read.csv(file="guru1.csv", header=TRUE, sep=",")  
linear

kabkota	kecamatan	negeri	swasta
Jakarta Pusat	Tanah Abang	419	142
Jakarta Pusat	Menteng	222	203
Jakarta Pusat	Senen	385	159
Jakarta Pusat	Johar Baru	384	60
Jakarta Pusat	Cempaka Putih	315	105
Jakarta Pusat	Kemayoran	856	319
Jakarta Pusat	Sawah Besar	263	266
Jakarta Pusat	Gambir	323	180
Jakarta Utara	Penjaringan	485	657
Jakarta Utara	Pademangan	339	157
Jakarta Utara	Tanjung Priok	765	721
Jakarta Utara	Koja	925	275
Jakarta Utara	Kelapa Gading	229	552
Jakarta Utara	Cilincing	875	436
Jakarta Barat	Kembangan	813	614
Jakarta Barat	Kebon Jeruk	852	498
Jakarta Barat	Palmerah	683	141
Jakarta Barat	Grogol Petamburan	510	386
Jakarta Barat	Tambora	408	422
Jakarta Barat	Taman Sari	288	211
Jakarta Barat	Cengkareng	992	614
Jakarta Barat	Kalideres	980	683
Jakarta Selatan	Jagakarsa	966	343
Jakarta Selatan	Pasar Minggu	732	358
Jakarta Selatan	Cilandak	652	527
Jakarta Selatan	Pesanggrahan	781	249
Jakarta Selatan	Kebayoran Lama	979	452
Jakarta Selatan	Kebayoran Baru	474	330
Jakarta Selatan	Mampang Prapatan	372	145
Jakarta Selatan	Pancoran	484	42
Jakarta Selatan	Tebet	750	332
Jakarta Selatan	Setia Budi	378	175
Jakarta Timur	Pasar Rebo	825	239
Jakarta Timur	Ciracas	914	223
Jakarta Timur	Cipayung	793	227
Jakarta Timur	Makasar	686	273
Jakarta Timur	Kramat Jati	881	326
Jakarta Timur	Jatinegara	870	306
Jakarta Timur	Duren Sawit	965	575
Jakarta Timur	Cakung	975	303
Jakarta Timur	Pulo Gadung	830	493
Jakarta Timur	Matraman	548	175
Kepulauan Seribu	Kepulauan Seribu Selatan	94	0
Kepulauan Seribu	Kepulauan Seribu Utara	113	0

In [17]: model <- lm(negeri ~ swasta, data=linear)  
summary(model)

Call:  
lm(formula = negeri ~ swasta, data = linear)

Residuals:

Min	1Q	Median	3Q	Max
-578.47	-161.26	52.58	201.33	364.68

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	374.3469	69.3907	5.395	2.93e-06 ***
swasta	0.7846	0.1894	4.142	0.000163 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

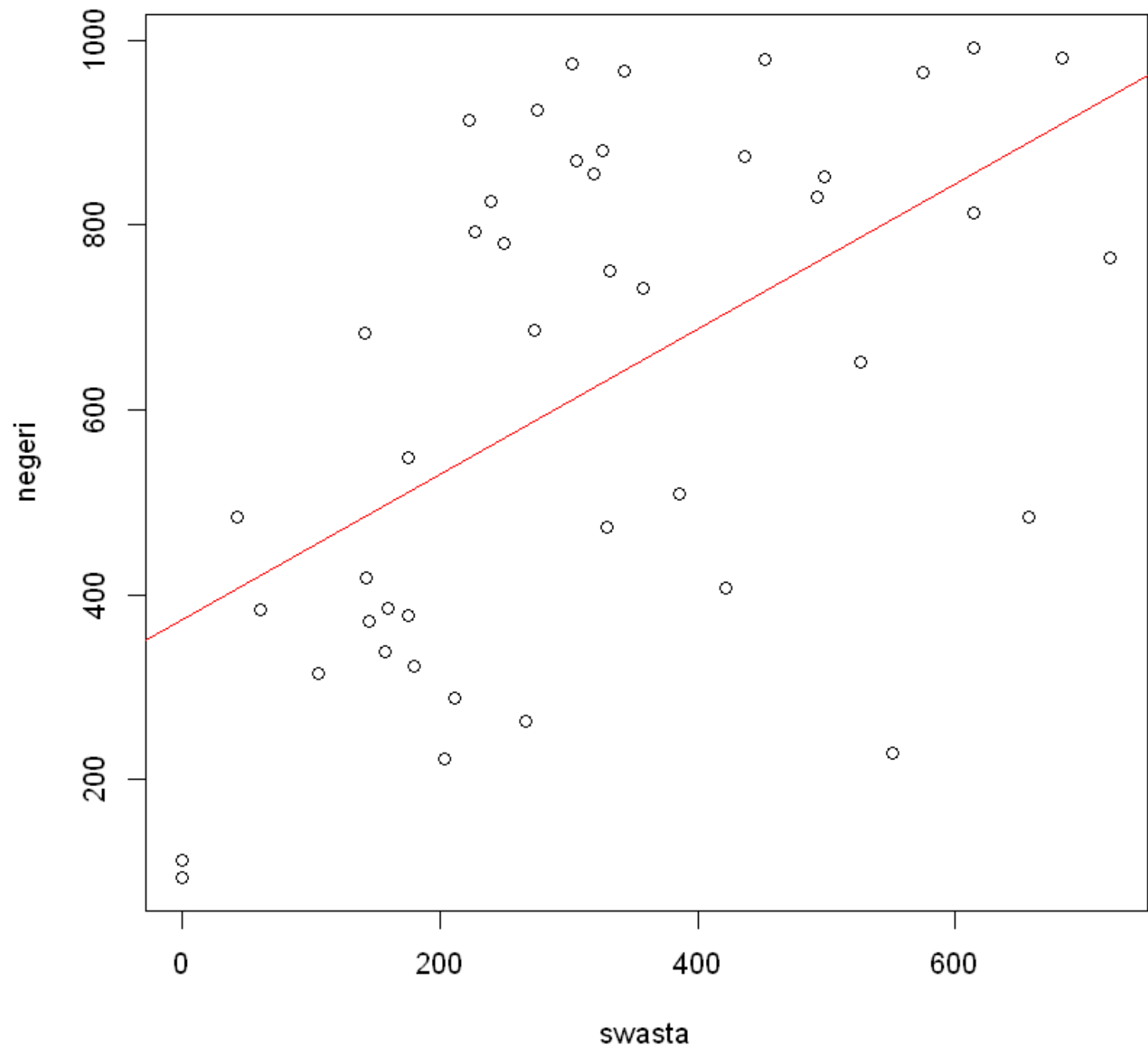
Residual standard error: 233.3 on 42 degrees of freedom  
Multiple R-squared: 0.29, Adjusted R-squared: 0.2731  
F-statistic: 17.16 on 1 and 42 DF, p-value: 0.0001626

Based on the output above, Let's understand the output. Values of coefficients( $\theta$ s) are **374.3469** as the *intercept* and **0.7846** as the *slope*, hence prediction equation for model is as below:

$$Negeri = 374.3469 + 0.7846 * swasta$$

Finally, we can add a best fit line (regression line) to our plot by adding the following text at the command line:

In [18]: plot(negeri ~ swasta, data=linear)  
abline(model, col = "red", lwd = 1)



In [21]: poly\_model <- lm(negeri ~ poly(swasta,degree=2), data = linear)  
summary(poly\_model)

Call:  
lm(formula = negeri ~ poly(swasta, degree = 2), data = linear)

Residuals:

Min	1Q	Median	3Q	Max
-572.33	-144.09	29.03	179.21	321.15

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	622.1	32.8	18.968	< 2e-16 ***
poly(swasta, degree = 2)1	966.3	217.6	4.442	6.62e-05 ***
poly(swasta, degree = 2)2	-587.5	217.6	-2.701	0.01 *

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 217.6 on 41 degrees of freedom  
Multiple R-squared: 0.3972, Adjusted R-squared: 0.3678  
F-statistic: 13.51 on 2 and 41 DF, p-value: 3.111e-05

Based on the output above, hence prediction equation for model is as below:

$$Negeri = 622.1 + 966.3 * swasta + (- 587.5 * swasta^2)$$

Finally, we can add a best fit line (regression line) to our plot by adding the following text at the command line:

In [19]: x <- with(linear, seq(min(swasta), max(swasta), length.out=2000))  
y <- predict(poly\_model, newdata = data.frame(swasta = x))  
  
plot(negeri ~ swasta, data = linear)  
lines(x, y, col = "red")

Call:  
lm(formula = negeri ~ poly(swasta, degree = 2), data = linear)

Coefficients:

(Intercept)	poly(swasta, degree = 2)1
622.1	966.3
poly(swasta, degree = 2)2	
-587.5	

