computer-related-offense.R

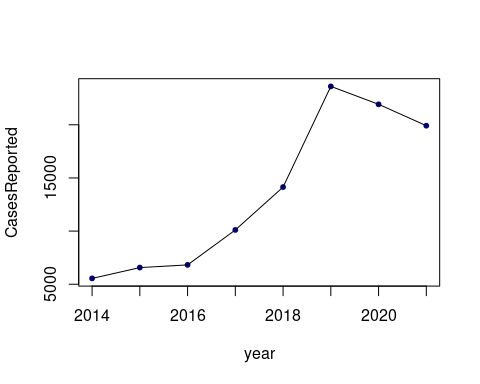
jasmeen

2023-03-26

library(readxl)  
data=read\_excel("computer related offense.xlsx")  
data

## # A tibble: 8 × 2  
## year CasesReported  
## <dbl> <dbl>  
## 1 2014 5548  
## 2 2015 6567  
## 3 2016 6818  
## 4 2017 10108  
## 5 2018 14141  
## 6 2019 23612  
## 7 2020 21926  
## 8 2021 19915

x=data$year  
y=data$CasesReported  
plot(data,pch=20,col="blue4")  
lines(data)



model1=lm(y~x)  
summary(model1)

##   
## Call:  
## lm(formula = y ~ x)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3321.8 -2224.4 -466.1 1492.9 5893.7   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5553255.5 1004317.1 -5.529 0.00147 \*\*  
## x 2759.3 497.8 5.543 0.00146 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3226 on 6 degrees of freedom  
## Multiple R-squared: 0.8366, Adjusted R-squared: 0.8094   
## F-statistic: 30.72 on 1 and 6 DF, p-value: 0.001456

model2=lm(y~poly(x,2,raw = TRUE))  
summary(model2)

##   
## Call:  
## lm(formula = y ~ poly(x, 2, raw = TRUE))  
##   
## Residuals:  
## 1 2 3 4 5 6 7 8   
## 1867.79 -79.66 -2726.05 -2264.39 -990.66 5790.13 1482.97 -3080.13   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -1.461e+08 1.108e+09 -0.132 0.900  
## poly(x, 2, raw = TRUE)1 1.421e+05 1.098e+06 0.129 0.902  
## poly(x, 2, raw = TRUE)2 -3.453e+01 2.722e+02 -0.127 0.904  
##   
## Residual standard error: 3528 on 5 degrees of freedom  
## Multiple R-squared: 0.8371, Adjusted R-squared: 0.772   
## F-statistic: 12.85 on 2 and 5 DF, p-value: 0.0107

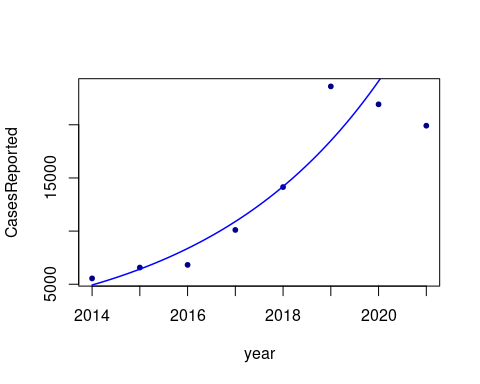
model3 = lm(log(y) ~ x)  
summary(model3)

##   
## Call:  
## lm(formula = log(y) ~ x)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.26652 -0.07994 0.01283 0.05997 0.35702   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -447.84694 63.83852 -7.015 0.000418 \*\*\*  
## x 0.22663 0.03164 7.162 0.000374 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2051 on 6 degrees of freedom  
## Multiple R-squared: 0.8953, Adjusted R-squared: 0.8778   
## F-statistic: 51.3 on 1 and 6 DF, p-value: 0.0003739

x\_axis=2014:2021  
data.frame(year=data$year,actual = data$CasesReported,linear=predict(model1,data.frame(x=x\_axis)),quad = predict(model2,data.frame(x=x\_axis)),loge = exp(predict(model3,data.frame(x=x\_axis))) )

## year actual linear quad loge  
## 1 2014 5548 3921.917 3680.208 5320.686  
## 2 2015 6567 6681.190 6646.661 6674.065  
## 3 2016 6818 9440.464 9544.054 8371.690  
## 4 2017 10108 12199.738 12372.387 10501.127  
## 5 2018 14141 14959.012 15131.661 13172.211  
## 6 2019 23612 17718.286 17821.875 16522.715  
## 7 2020 21926 20477.560 20443.030 20725.460  
## 8 2021 19915 23236.833 22995.125 25997.221

plot(data,pch=20,col="blue4")  
m3=function(x){  
 exp(-524.60367+0.26470\*x)  
}  
curve(m3,from = 2014 ,to = 2021,add = TRUE ,lwd = 1.5 , col = "blue")



##calculate percentage increase in x  
rm(list = ls())  
y1=2014:2021  
x1=c(5548,6567,6818,10108,14141,23612,21926,19915)  
df=data.frame(y1,x1)  
df

## y1 x1  
## 1 2014 5548  
## 2 2015 6567  
## 3 2016 6818  
## 4 2017 10108  
## 5 2018 14141  
## 6 2019 23612  
## 7 2020 21926  
## 8 2021 19915

pinc=function(x){  
 per\_inc = c()  
 for(i in 2:length(x)){  
 per\_inc[i] = x[i]/x[i-1] - 1  
 }  
 return(100\*per\_inc)  
}  
cbind(df,per\_inc = round(pinc(x1),4))

## y1 x1 per\_inc  
## 1 2014 5548 NA  
## 2 2015 6567 18.3670  
## 3 2016 6818 3.8221  
## 4 2017 10108 48.2546  
## 5 2018 14141 39.8991  
## 6 2019 23612 66.9755  
## 7 2020 21926 -7.1404  
## 8 2021 19915 -9.1718