Mini project

J WANG

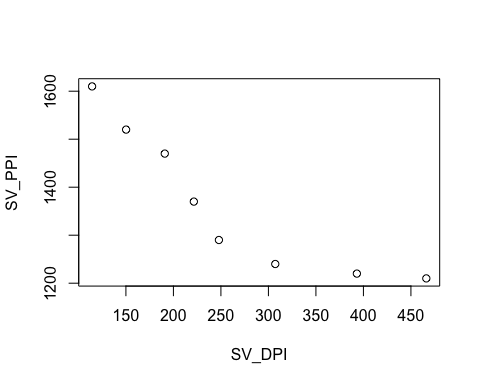
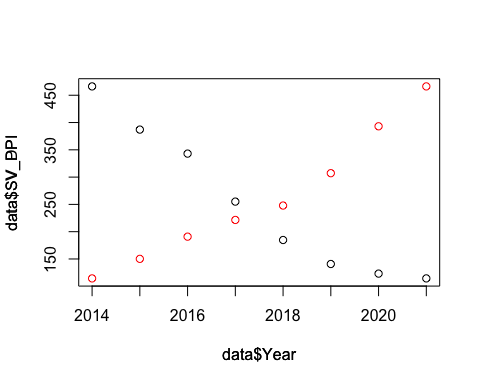
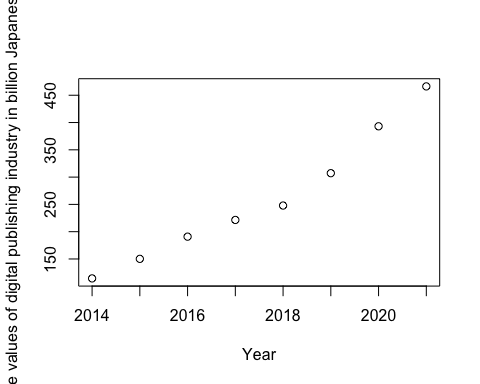
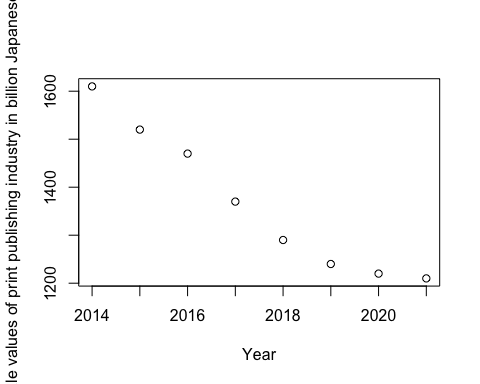
2023-02-15

#question Does digital publishing industry influence print publishing industry in Japan? Choose sale values as standard If the sale values of digital publishing industry increase while the sale values of print publishing industry decrease, then we can conclude that digital pulication industry has an effect on print publicaiton industry.

#import data and linear model Method All data was collected from the publishing industry in Japan in Statista (<https://www.statista.com/topics/9291/publishing-industry-in-japan/>). Sale values of print publishing industry and digital publishing industry in Japan from 2014 to 2022 were collected. In addition, salve values of printed and digital books, maganizes, and comic books in Japan fron 2013 to 2022 were collected. We selected the data in the period of 2014-2022 to analysis.

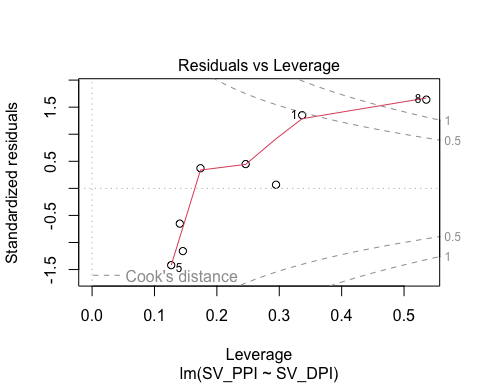
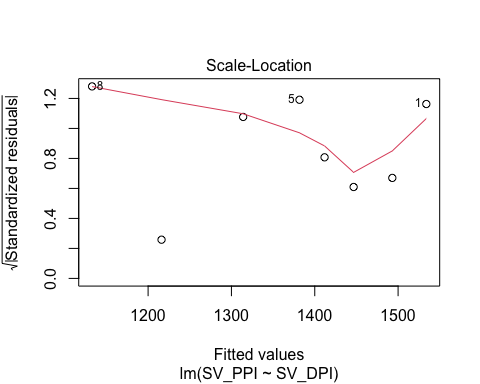
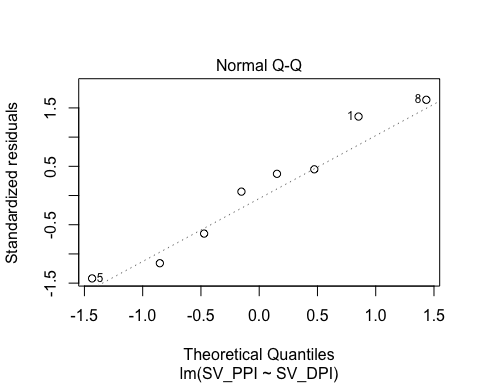
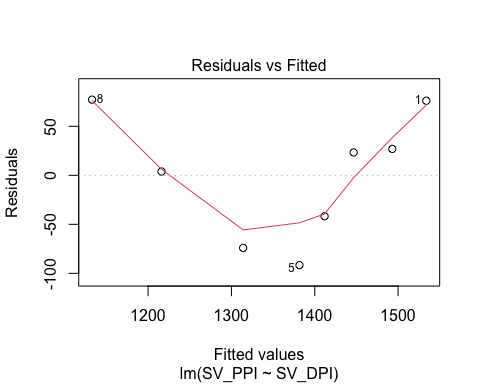
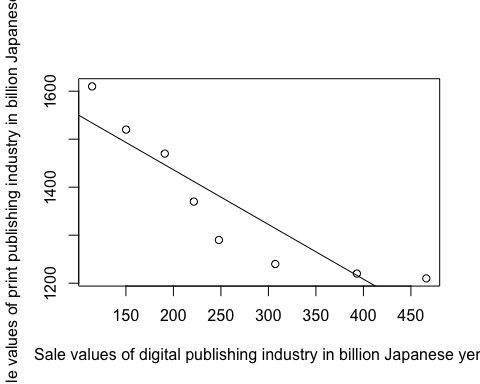
## Loading required package: ggplot2

## 'data.frame': 8 obs. of 12 variables:  
## $ Area : chr "Japan " "Japan " "Japan " "Japan " ...  
## $ Year : int 2014 2015 2016 2017 2018 2019 2020 2021  
## $ SV\_PPI: int 1610 1520 1470 1370 1290 1240 1220 1210  
## $ SV\_DPI: num 114 150 191 222 248 ...  
## $ SV\_PB : num 755 742 737 715 699 ...  
## $ SV\_DB : num 19.2 22.8 25.8 29 32.1 34.9 40.1 44.9  
## $ SV\_PM : num 852 780 734 655 593 ...  
## $ SV\_DM : num 6.5 10.5 16 17.8 15.6 13 11 9.9  
## $ SV\_PCB: num 226 210 195 167 159 ...  
## $ SV\_DCB: num 88.7 116.9 149.1 174.7 200.2 ...  
## $ GDP\_PC: num 38523 25006 39411 38903 39850 ...  
## $ AEPM : int 263 261 261 262 265 264 262 264



##   
## Call:  
## lm(formula = SV\_PPI ~ SV\_DPI, data = data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -91.67 -49.84 13.61 39.24 77.21   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1664.2927 61.6155 27.011 1.7e-07 \*\*\*  
## SV\_DPI -1.1401 0.2164 -5.269 0.00188 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 69.13 on 6 degrees of freedom  
## Multiple R-squared: 0.8223, Adjusted R-squared: 0.7927   
## F-statistic: 27.77 on 1 and 6 DF, p-value: 0.001884

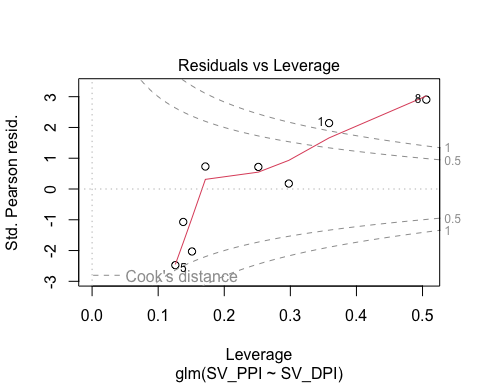
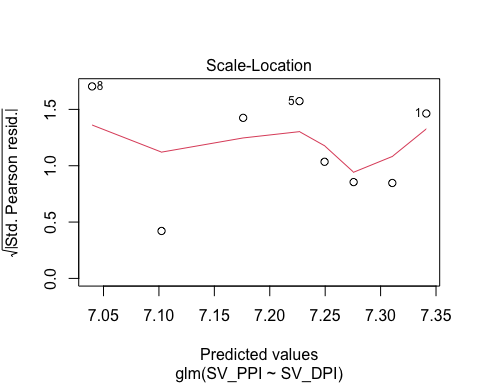
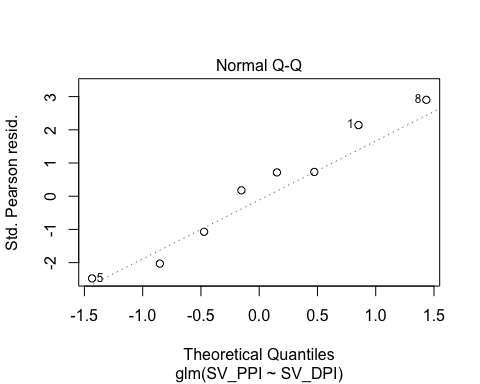
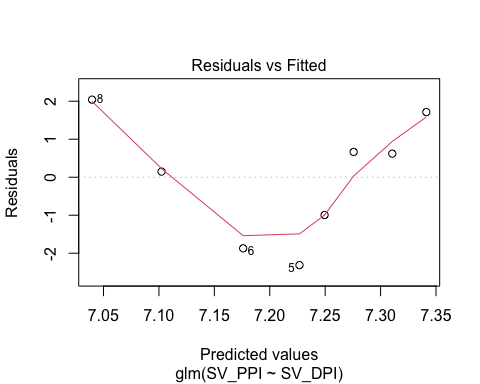
## (Intercept) SV\_DPI   
## 1664.29269 -1.14007

 linear model may not be a good choice.

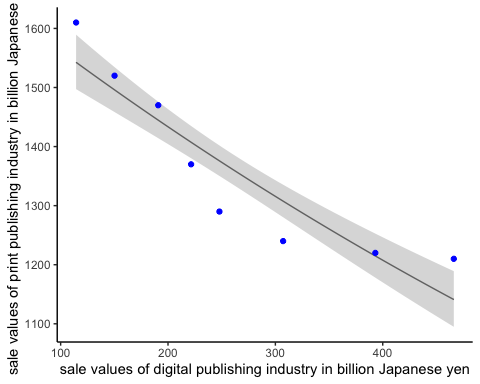
#poisson model

##   
## Call:  
## glm(formula = SV\_PPI ~ SV\_DPI, family = "poisson", data = data)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.3386 -1.2212 0.3836 0.9232 2.0214   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 7.439e+00 2.384e-02 312.114 <2e-16 \*\*\*  
## SV\_DPI -8.571e-04 8.706e-05 -9.846 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 116.71 on 7 degrees of freedom  
## Residual deviance: 17.86 on 6 degrees of freedom  
## AIC: 94.281  
##   
## Number of Fisher Scoring iterations: 3

## Analysis of Deviance Table  
##   
## Model: poisson, link: log  
##   
## Response: SV\_PPI  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev Pr(>Chi)   
## NULL 7 116.71   
## SV\_DPI 1 98.852 6 17.86 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

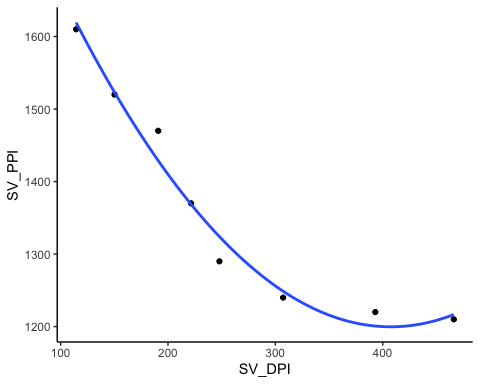


## [1] 114.4 466.2

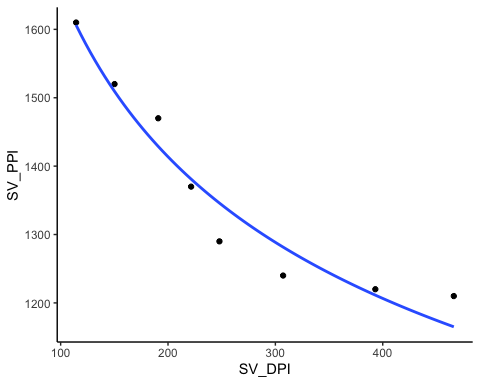


#non-linear least square

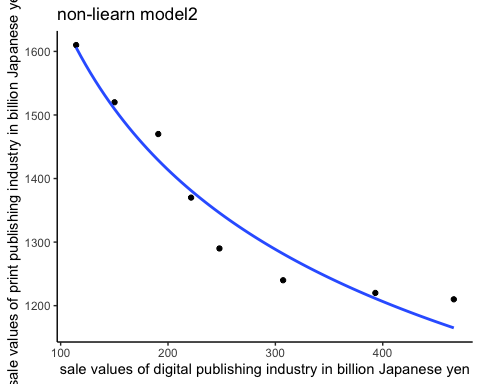
##   
## Formula: y ~ a \* x^2 + b \* x + c  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## a 4.901e-03 1.703e-06 2879 <2e-16 \*\*\*  
## b -3.983e+00 9.992e-04 -3987 <2e-16 \*\*\*  
## c 2.012e+03 1.342e-01 14989 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.295 on 349 degrees of freedom  
##   
## Number of iterations to convergence: 1   
## Achieved convergence tolerance: 1.49e-08



##   
## Formula: y ~ a \* x^b  
##   
## Parameters:  
## Estimate Std. Error t value Pr(>|t|)   
## a 4.612e+03 7.327e-01 6295 <2e-16 \*\*\*  
## b -2.230e-01 2.865e-05 -7783 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2871 on 350 degrees of freedom  
##   
## Number of iterations to convergence: 13   
## Achieved convergence tolerance: 1.19e-08



## [1] 2.409993

#final graph  #comic book 