

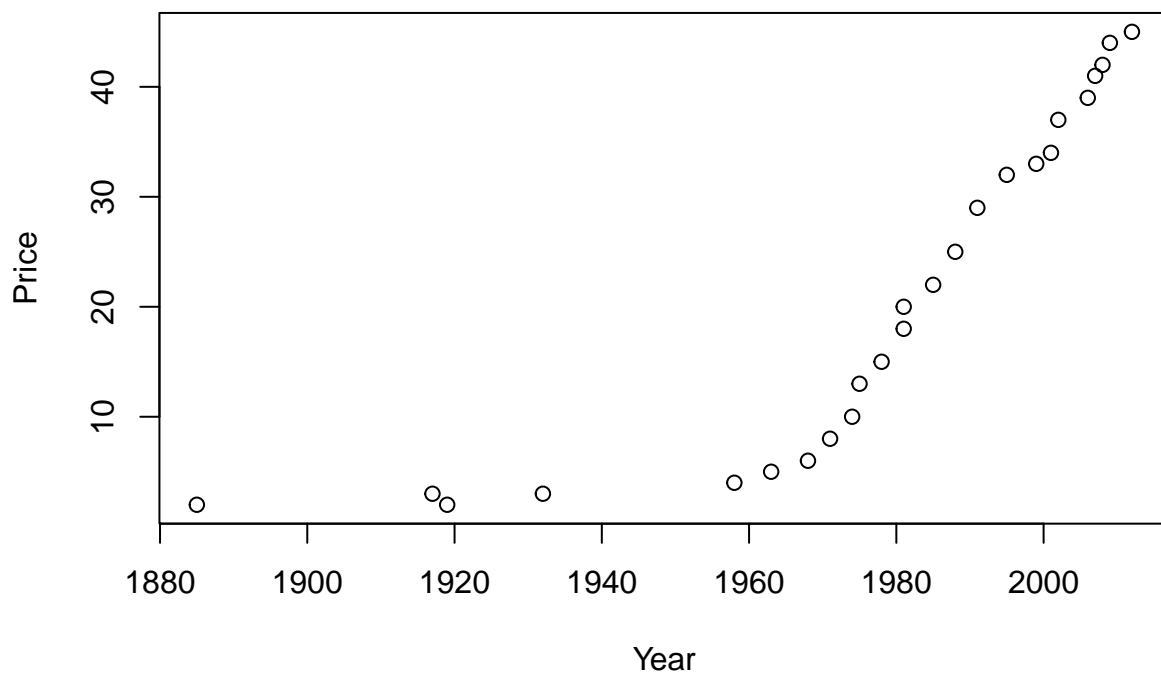
## Supplement for Lecture 5: Assessing Conditions

### Load Data from Textbook

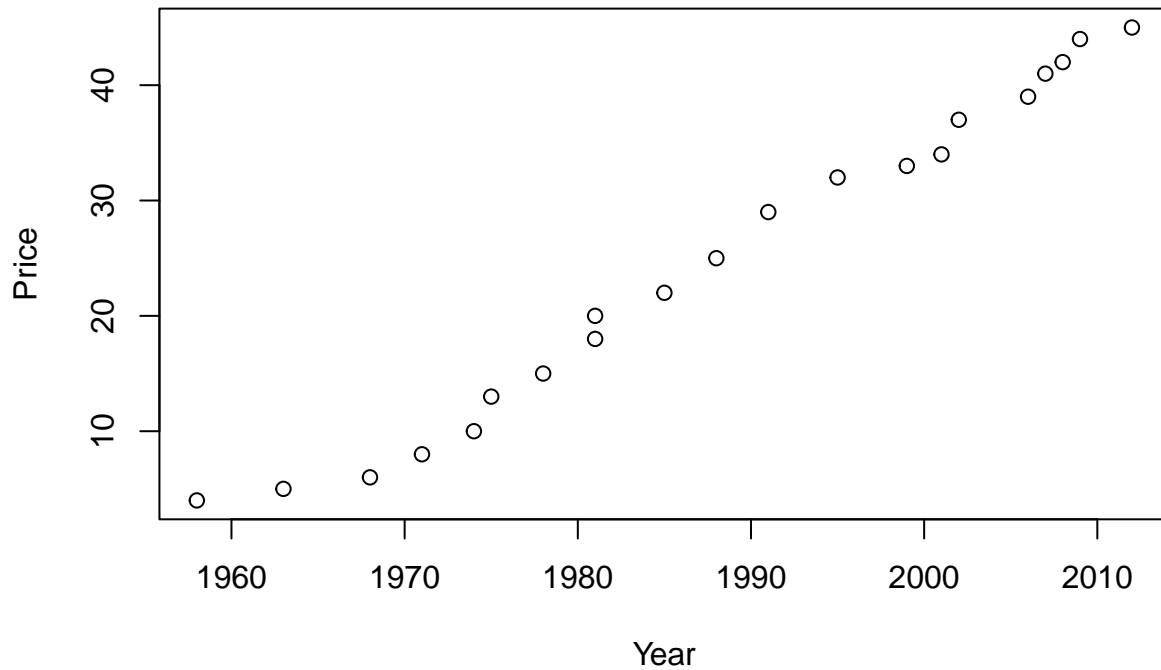
```
data("USstamps") # Load Data  
stamp = USstamps # Shorten Name  
rm(USstamps) # Removes Old Object from Environment
```

### Scatterplot of Price vs Year

```
plot(Price ~ Year, data=stamp)
```



```
#Remove First Four Years in Data (1885,1917,1919,1932)  
#See Exercise 1.33 for Reason Why  
  
stamp2 = subset(stamp, Year > 1932)  
plot(Price ~ Year, data=stamp2)
```



## Fit Linear Regression Models

```
#Linear Regression on Original Data
```

```
mod1 <- lm(Price~Year,data=stamp)
```

```
mod1
```

```
##
```

```
## Call:
```

```
## lm(formula = Price ~ Year, data = stamp)
```

```
##
```

```
## Coefficients:
```

```
## (Intercept)      Year
```

```
##   -770.7811      0.4008
```

```
summary(mod1)
```

```
##
```

```
## Call:
```

```
## lm(formula = Price ~ Year, data = stamp)
```

```
##
```

```
## Residuals:
```

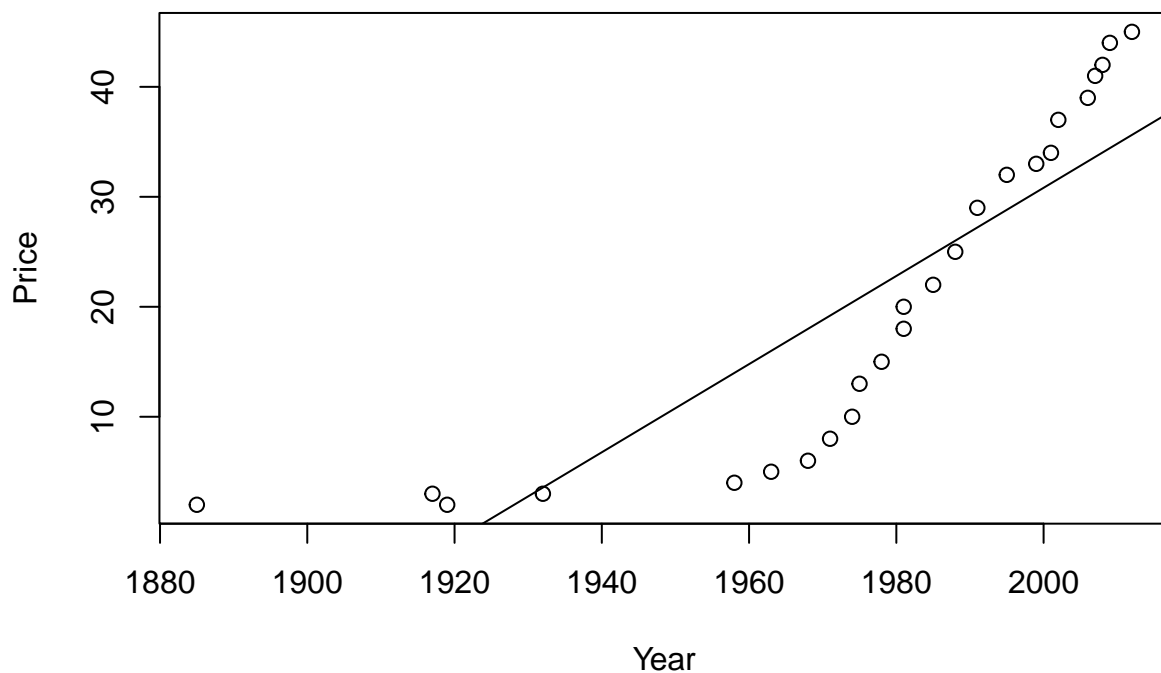
```
##      Min       1Q   Median       3Q      Max
```

```
## -11.993  -7.001   1.788   5.447  17.273
```

```
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) -770.78108   99.39045  -7.755 7.28e-08 ***
## Year         0.40080     0.05029   7.970 4.57e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.03 on 23 degrees of freedom
## Multiple R-squared:  0.7342, Adjusted R-squared:  0.7226
## F-statistic: 63.52 on 1 and 23 DF,  p-value: 4.572e-08

plot(Price ~ Year, data=stamp)
abline(mod1)
```

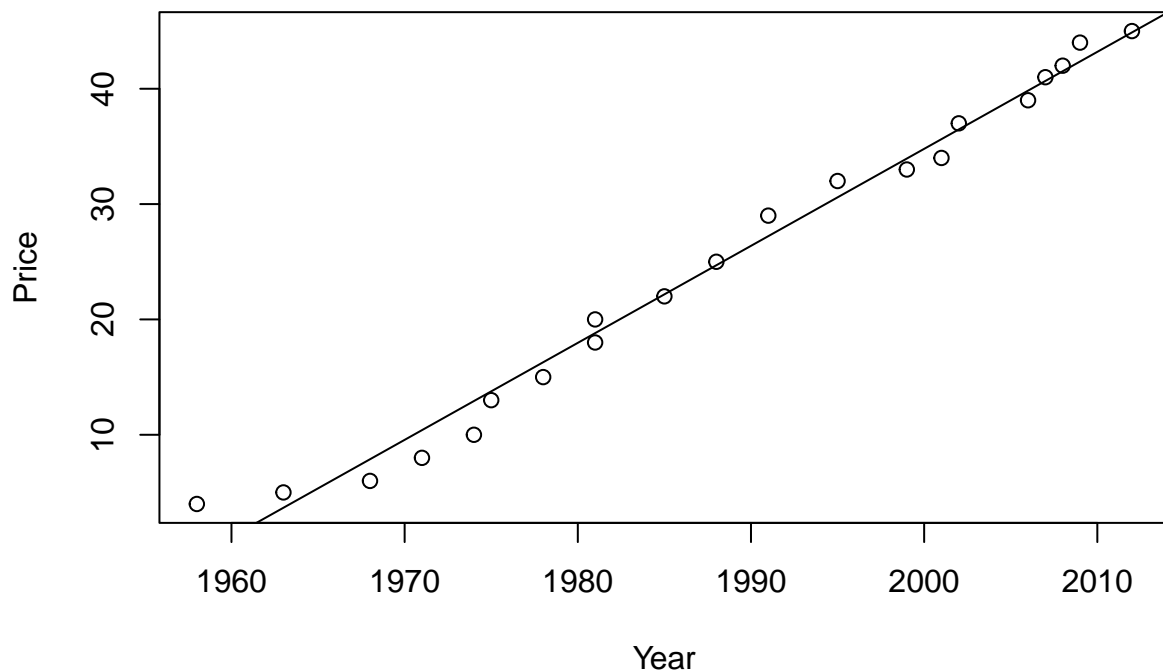


```
#Linear Regression on Subsetted Data
mod2 <-lm(Price~Year, data=stamp2)
mod2

##
## Call:
## lm(formula = Price ~ Year, data = stamp2)
##
## Coefficients:
## (Intercept)      Year
##   -1647.175      0.841
```

```
summary(mod2)
```

```
##
## Call:
## lm(formula = Price ~ Year, data = stamp2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9232 -0.9478  0.1195  1.1899  4.5325
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.647e+03  4.686e+01  -35.15  <2e-16 ***
## Year          8.410e-01  2.357e-02   35.68  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.737 on 19 degrees of freedom
## Multiple R-squared:  0.9853, Adjusted R-squared:  0.9845
## F-statistic: 1273 on 1 and 19 DF,  p-value: < 2.2e-16
plot(Price ~ Year, data=stamp2)
abline(mod2)
```



## Saving Fitted Values and Residuals

```
mod1$residuals
```

```
##          1          2          3          4          5          6
## 17.2729657  5.4473637  3.6457636 -0.5646372 -9.9854389 -10.9894392
##          7          8          9         10         11         12
## -11.9934395 -11.1958397 -10.3982399 -7.7990399 -7.0014401 -5.2038403
##         13         14         15         16         17         18
## -3.2038403 -2.8070406 -1.0094407  1.7881591  3.1849588  2.5817586
##         19         20         21         22         23         24
##  2.7801584  5.3793584  5.7761581  7.3753581  7.9745580  9.5737579
##         25
##  9.3713578
```

```
mod1$fitted.values
```

```
##          1          2          3          4          5          6          7
## -15.272966 -2.447364 -1.645764  3.564637 13.985439 15.989439 17.993439
##          8          9         10         11         12         13         14
## 19.195840 20.398240 20.799040 22.001440 23.203840 23.203840 24.807041
##         15         16         17         18         19         20         21
## 26.009441 27.211841 28.815041 30.418241 31.219842 31.620642 33.223842
##         22         23         24         25
## 33.624642 34.025442 34.426242 35.628642
```

```
stamp$fit=mod1$fitted.values
```

```
stamp$resid=mod1$residuals
```

```
mod2$residuals
```

```
##          5          6          7          8          9         10         11
##  4.5324928  1.3275847 -1.8773234 -2.4002683 -2.9232132 -0.7641948 -1.2871397
##         12         13         14         15         16         17         18
## -0.8100845  1.1899155 -0.1740110  0.3030441  1.7800992  1.4161728 -0.9477537
##         19         20         21         22         23         24         25
## -1.6297170  0.5293014 -0.8346251  0.3243933  0.4834117  1.6424300  0.1194852
```

```
mod2$fitted.values
```

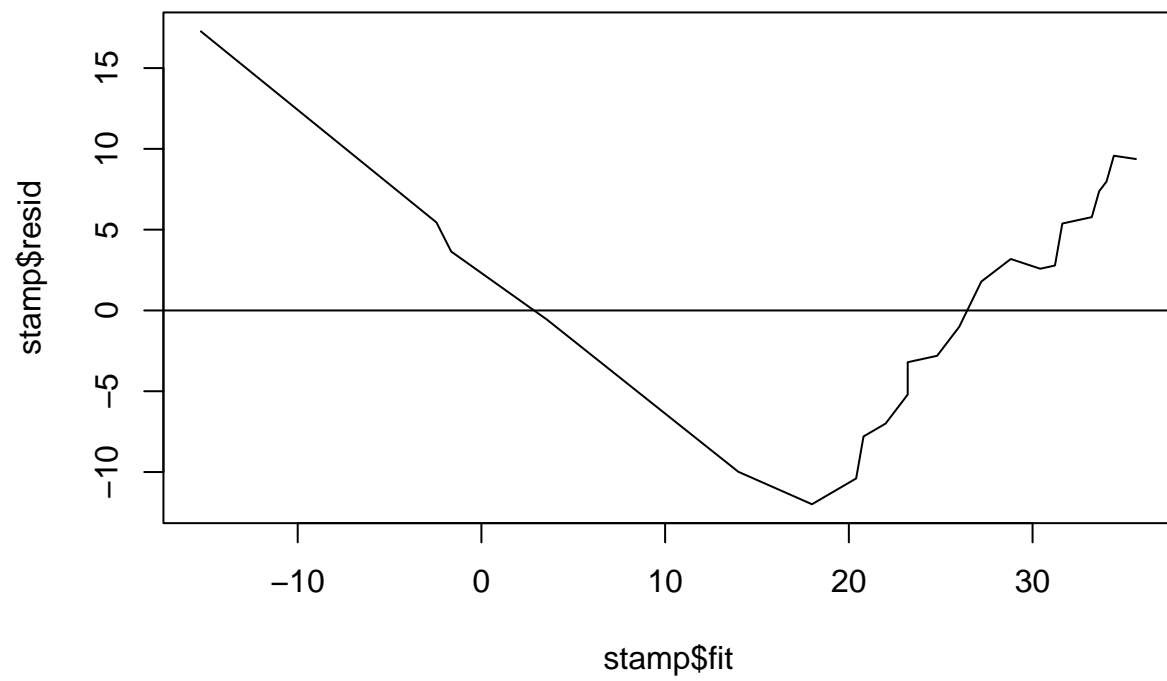
```
##          5          6          7          8          9         10         11
## -0.5324928  3.6724153  7.8773234 10.4002683 12.9232132 13.7641948 16.2871397
##         12         13         14         15         16         17         18
## 18.8100845 18.8100845 22.1740110 24.6969559 27.2199008 30.5838272 33.9477537
##         19         20         21         22         23         24         25
## 35.6297170 36.4706986 39.8346251 40.6756067 41.5165883 42.3575700 44.8805148
```

```
stamp2$fit=mod2$fitted.values
```

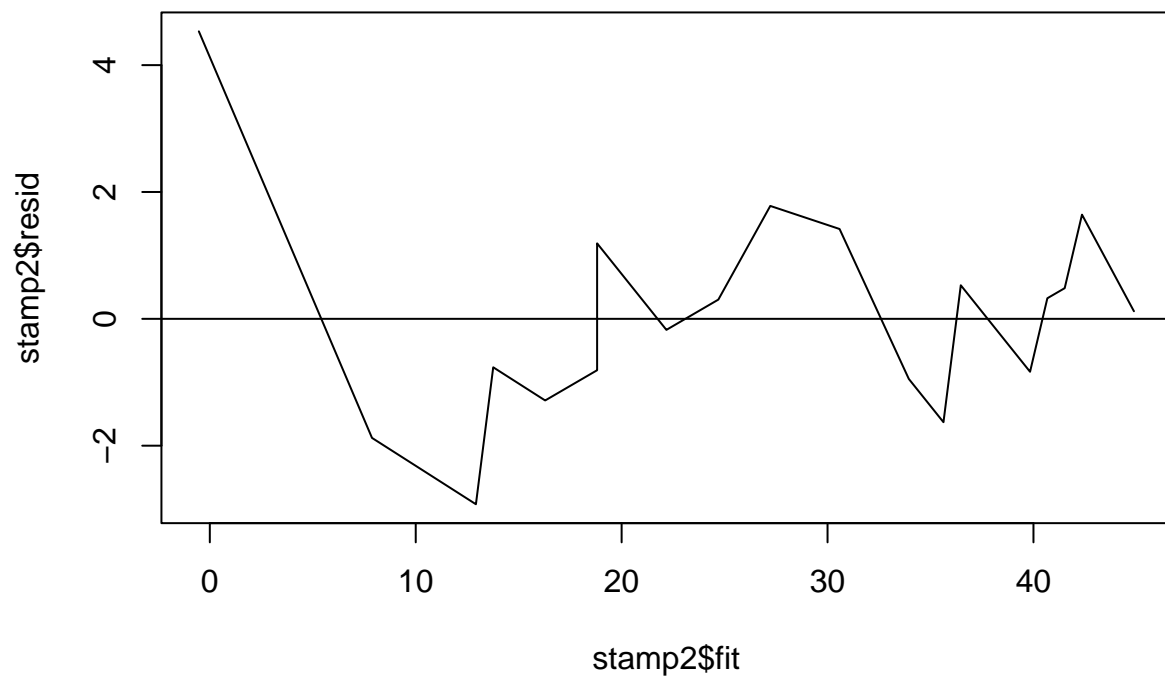
```
stamp2$resid=mod2$residuals
```

## Residuals vs Fit

```
plot(y=stamp$resid,x=stamp$fit,type="l")
abline(h=0) #h=location of horizontal line
```

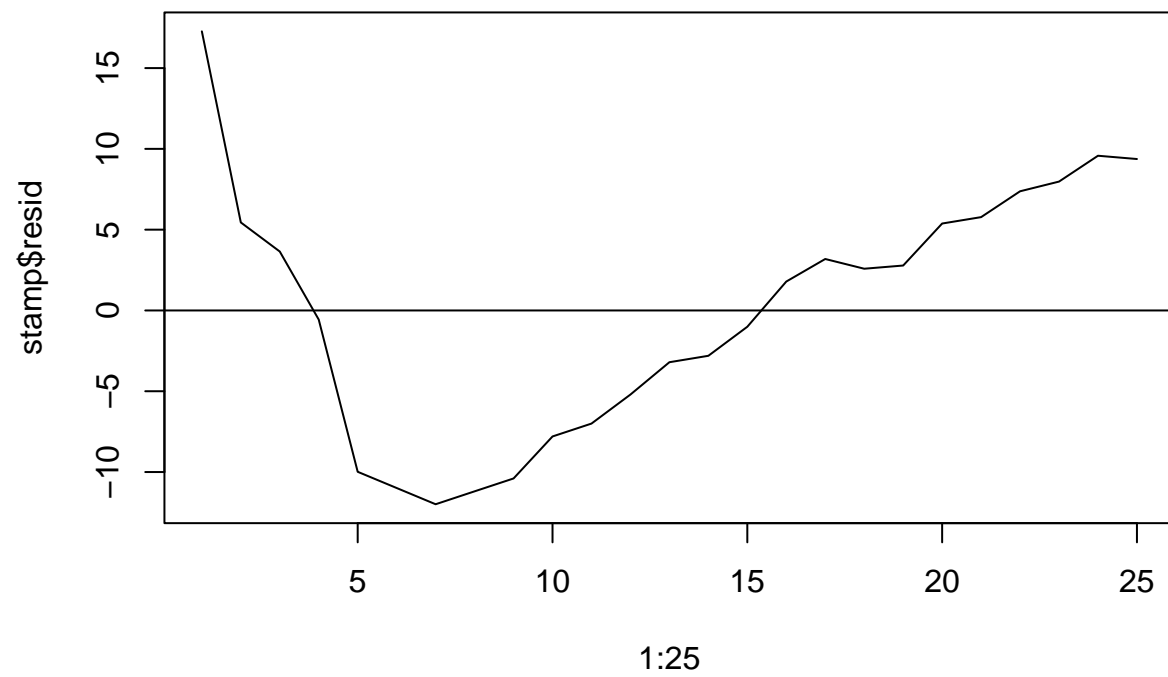


```
plot(y=stamp2$resid,x=stamp2$fit,type="l")  
abline(h=0)
```



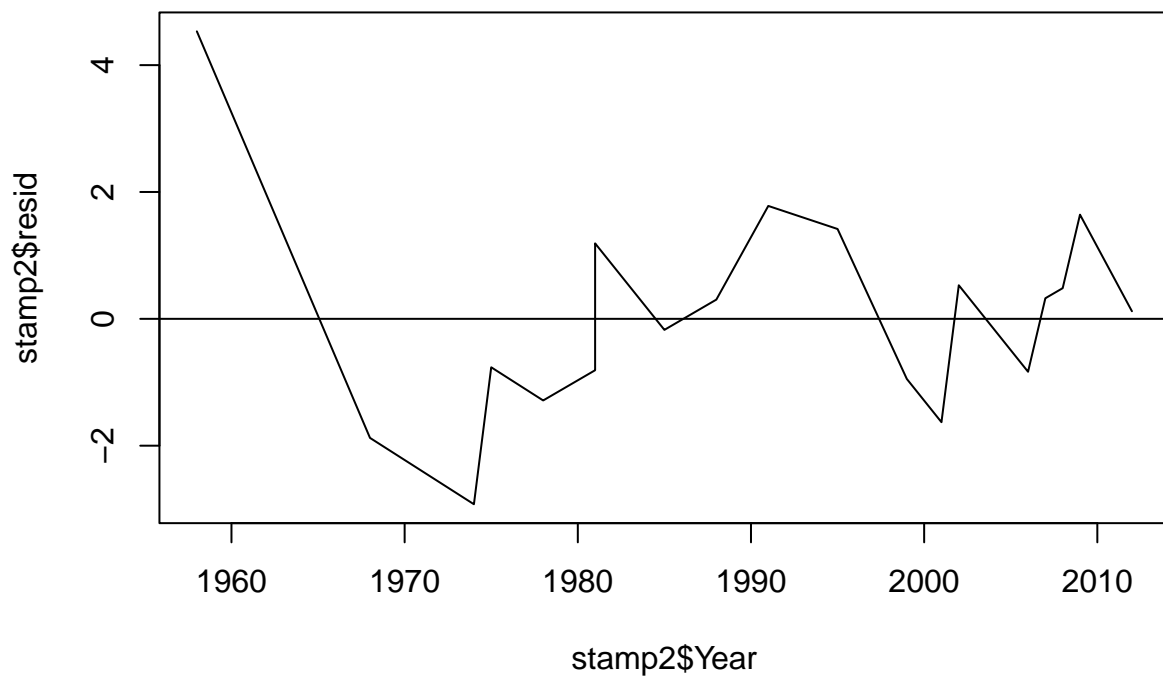
## Residuals vs Order (Time)

```
plot(y=stamp$resid,x=1:25,type="l")  
abline(h=0) #h=location of horizontal line
```



```
plot(y=stamp2$resid,x=stamp2$Year,type="l")  
abline(h=0)
```

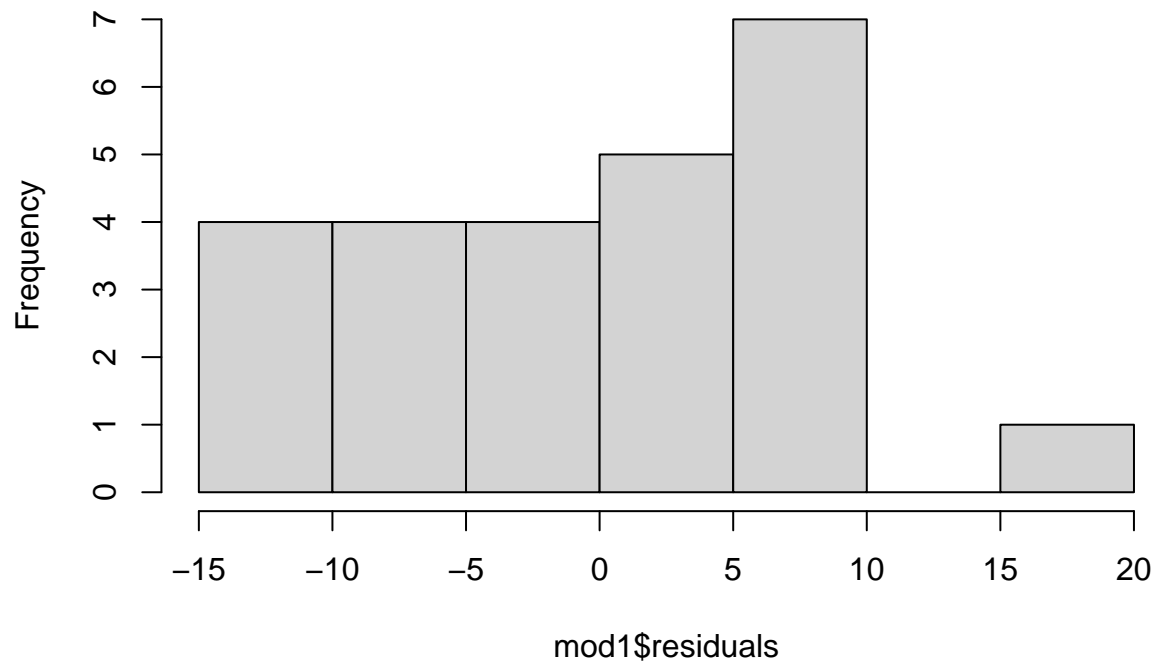




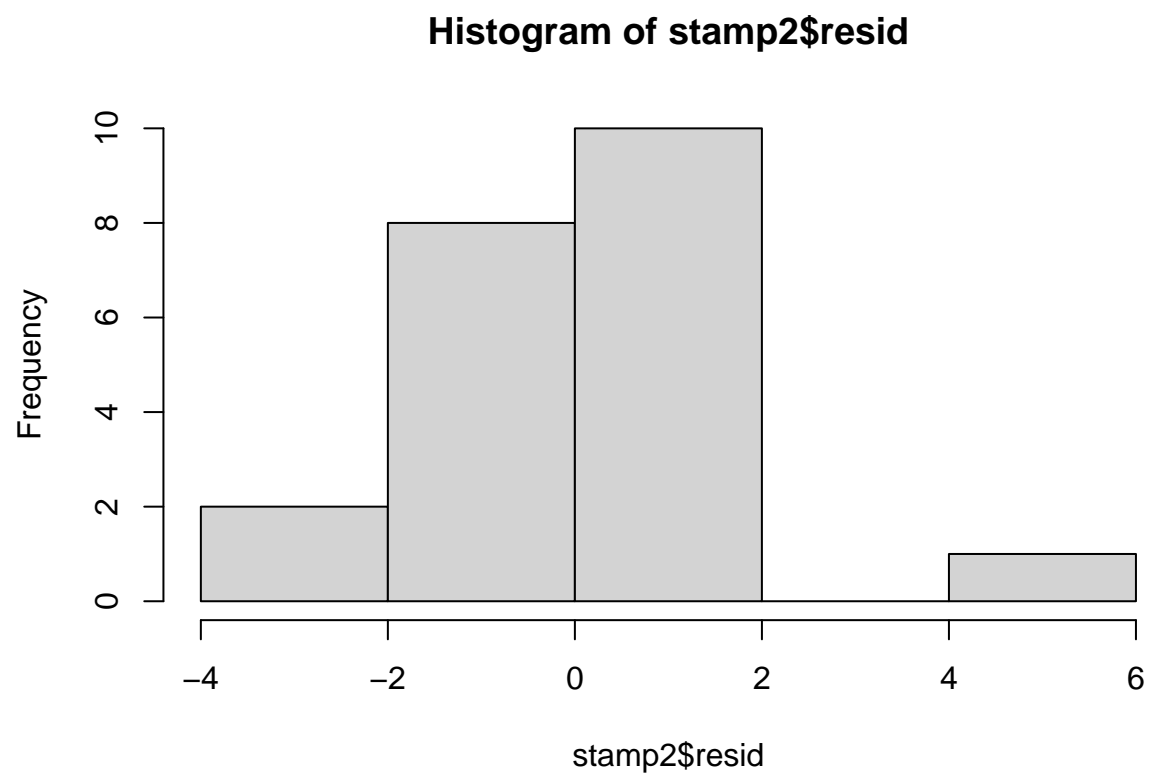
## Histogram/Boxplots of Resid

```
hist(mod1$residuals,breaks=5) #Remember: We can always plot residuals directly from model object
```

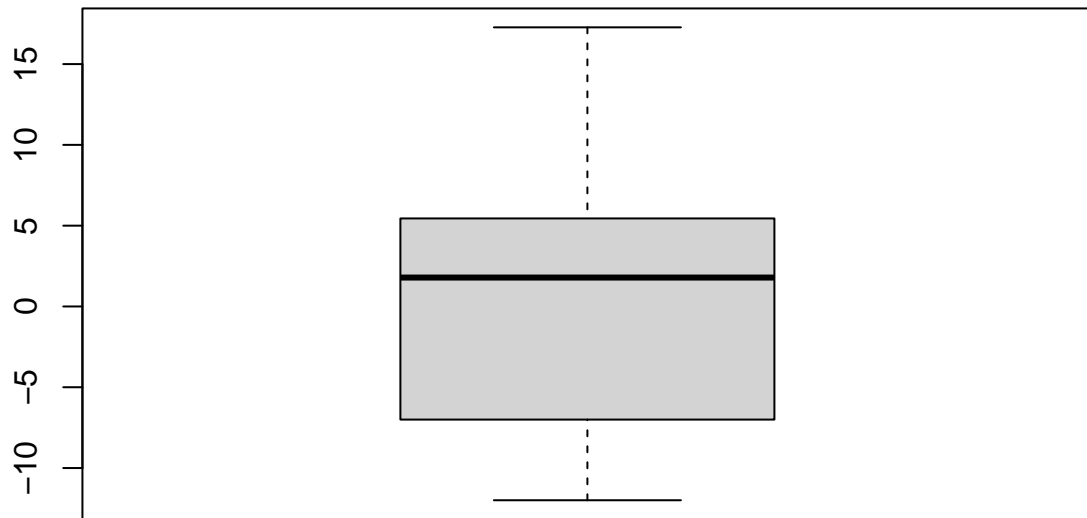
**Histogram of mod1\$residuals**



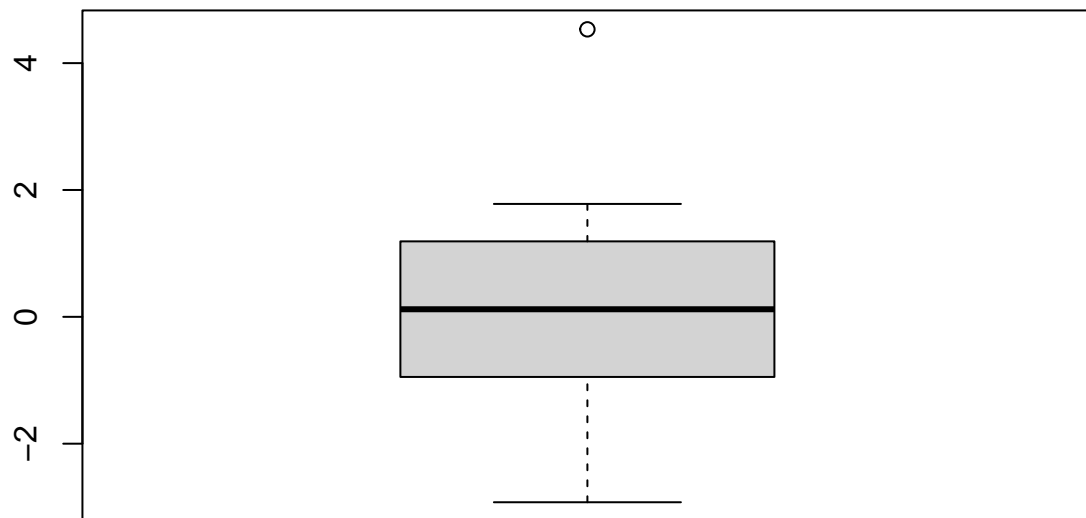
```
hist(stamp2$resid,breaks=5) #Remember: We can plot residuals that we saved into our dataset
```



```
boxplot(stamp$resid)
```



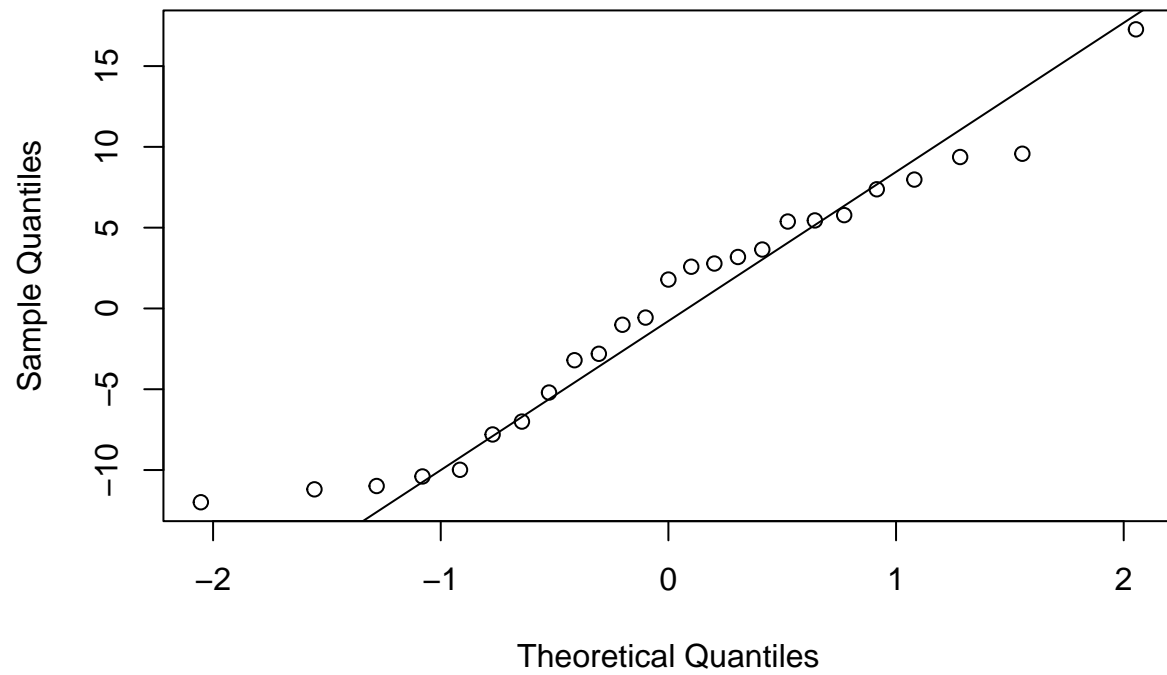
```
boxplot(mod2$residuals)
```



## Normal Plots

```
qqnorm(stamp$resid)  
qqline(stamp$resid)
```

**Normal Q-Q Plot**



```
qqnorm(stamp2$resid)  
qqline(stamp2$resid)
```

**Normal Q-Q Plot**

