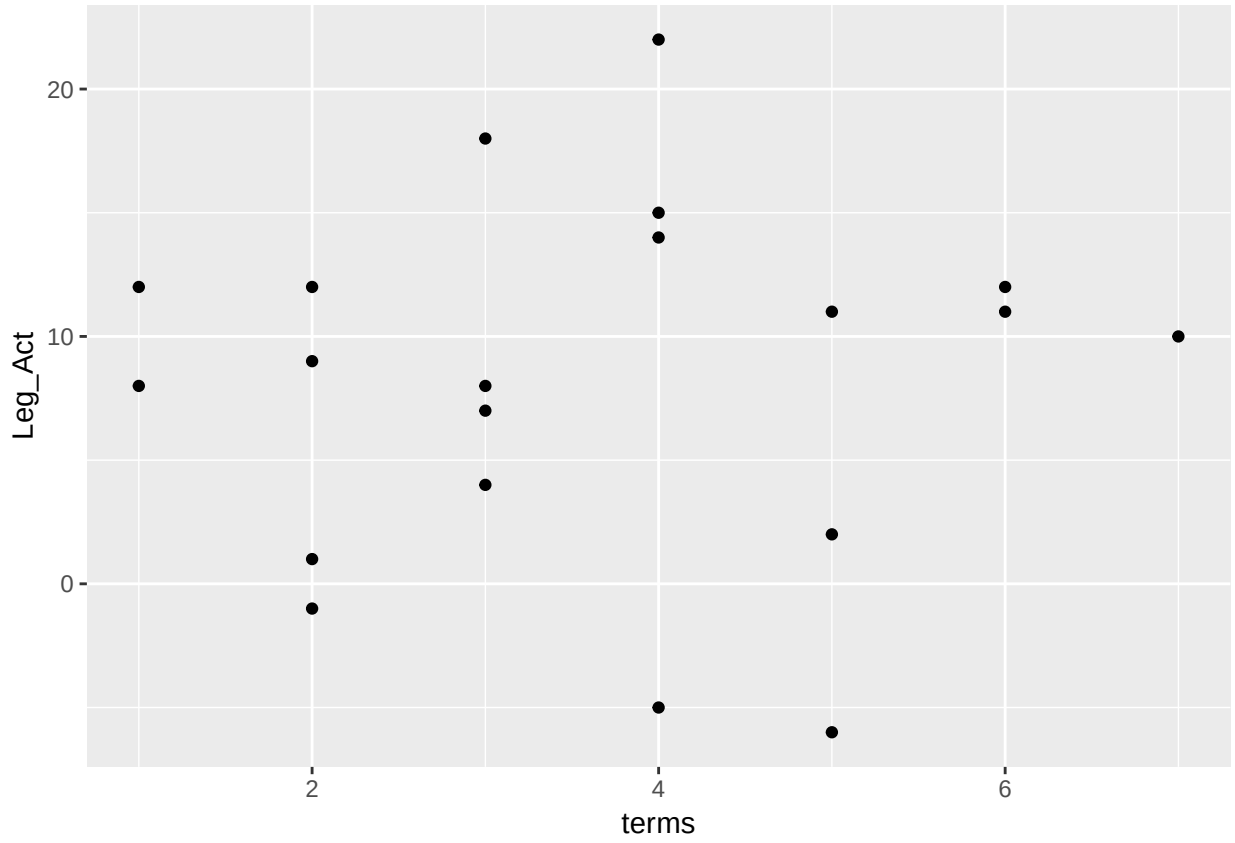
813 Data exercise 1

X. Zhang

January 25, 2019

```
## -- Attaching packages ------------------------------------------------------------------- tidyverse 1
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
## Warning: package 'tibble' was built under R version 3.5.2
## Warning: package 'readr' was built under R version 3.5.2
## -- Conflicts ---------------------------------------------------------------------- tidyverse_conflic
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
ggplot(EX1_813, aes(x = terms, y = Leg_Act)) + geom_point()
```



plotting legislative acts against terms, graph show no clear pattern.

```
leg_mod <- lm(Leg_Act ~ terms, data = EX1_813)
summary(leg_mod)
```

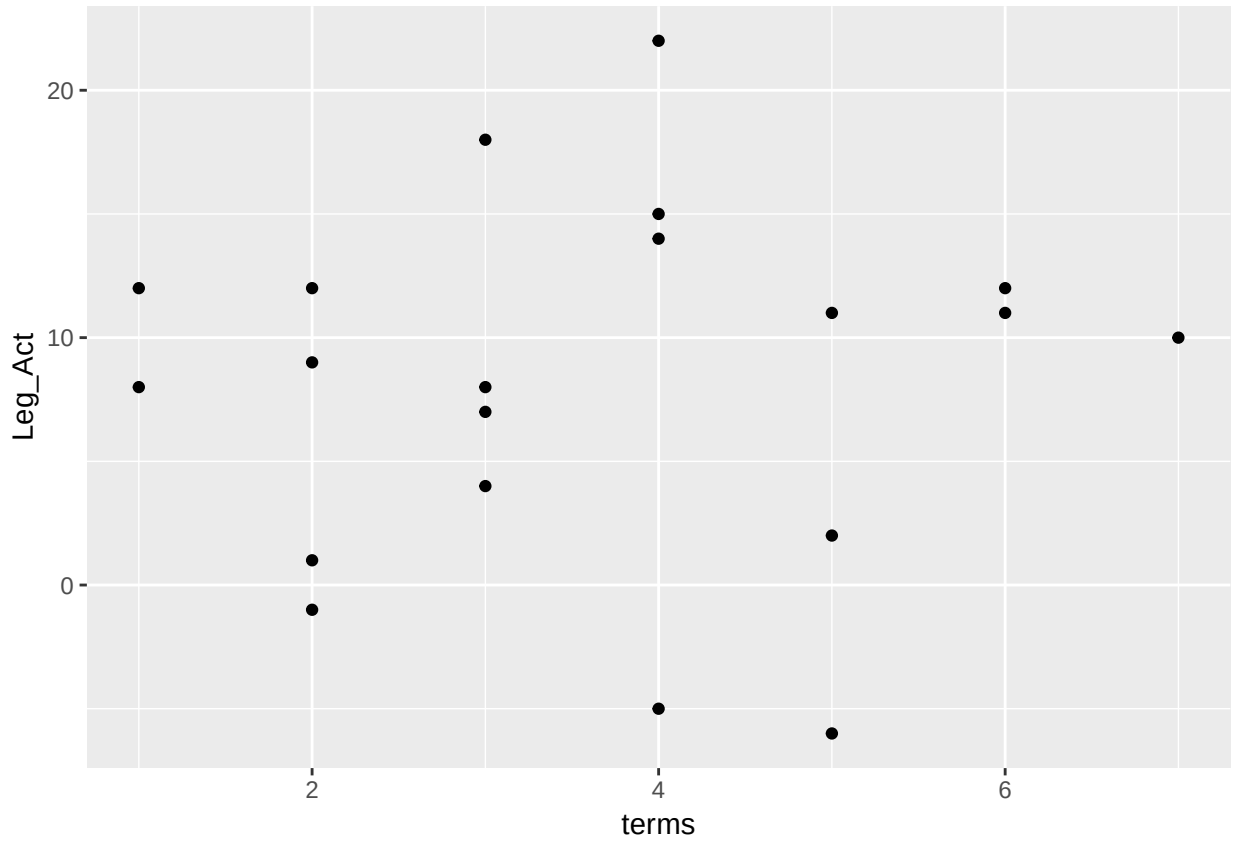
```
##
## Call:
```

```
## lm(formula = Leg_Act ~ terms, data = EX1_813)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -14.445  -4.683   1.142   4.124  13.730
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.5693     3.9846   1.900   0.0736 .
## terms         0.1752     1.0056   0.174   0.8636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.444 on 18 degrees of freedom
## Multiple R-squared:  0.001683,   Adjusted R-squared:  -0.05378
## F-statistic: 0.03035 on 1 and 18 DF,  p-value: 0.8636
```

intercept: 7.57; coefficient of terms variable: 0.17

```
EX1_813$residual <- resid(leg_mod)
EX1_813$pred <- predict(leg_mod)
```

```
ggplot(EX1_813, aes(x = terms, y = Leg_Act)) + geom_point()
```




```
correlation <- cor(EX1_813$Leg_Act, EX1_813$pred)
```

correlation of observed legislative act and predicted value is 0.041

```r
summary(leg_mod)$r.squared
```

```
## [1] 0.001683098
```

```r
c.sq <- correlation^2
```

correlation squared (c.sq) and r square of the model is the same.

```r
summary(leg_mod)
```

```
##
## Call:
## lm(formula = Leg_Act ~ terms, data = EX1_813)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -14.445  -4.683   1.142   4.124  13.730
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.5693     3.9846   1.900   0.0736 .
## terms         0.1752     1.0056   0.174   0.8636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.444 on 18 degrees of freedom
## Multiple R-squared:  0.001683,   Adjusted R-squared:  -0.05378
## F-statistic: 0.03035 on 1 and 18 DF,  p-value: 0.8636
```

Discussion:

Testing the hypothesis at p-value = 0.05.The model output suggests that although the terms variable have a coefficient 0.1725, which suggests more terms are associated with more legislative actions. However, the terms variable has a p-value of 0.86, which is larger than 0.05; Thus, the null hypothesis of b=0 cannot be rejected. The intercept 7.5693 suggesting that 0 term legislators have approximately 7.5 legislative acts. However, sine there is no 0 term legislators, it is difficult to assign substantive interpretation to the intercept.