Cargamos las librerias necesarias

In [12]:
```
library(readr)
library(tidyr)
library(dplyr)
```

Cargamos los datos

In [3]:
```
datos <- read.csv("incumbents_subset.csv")
```

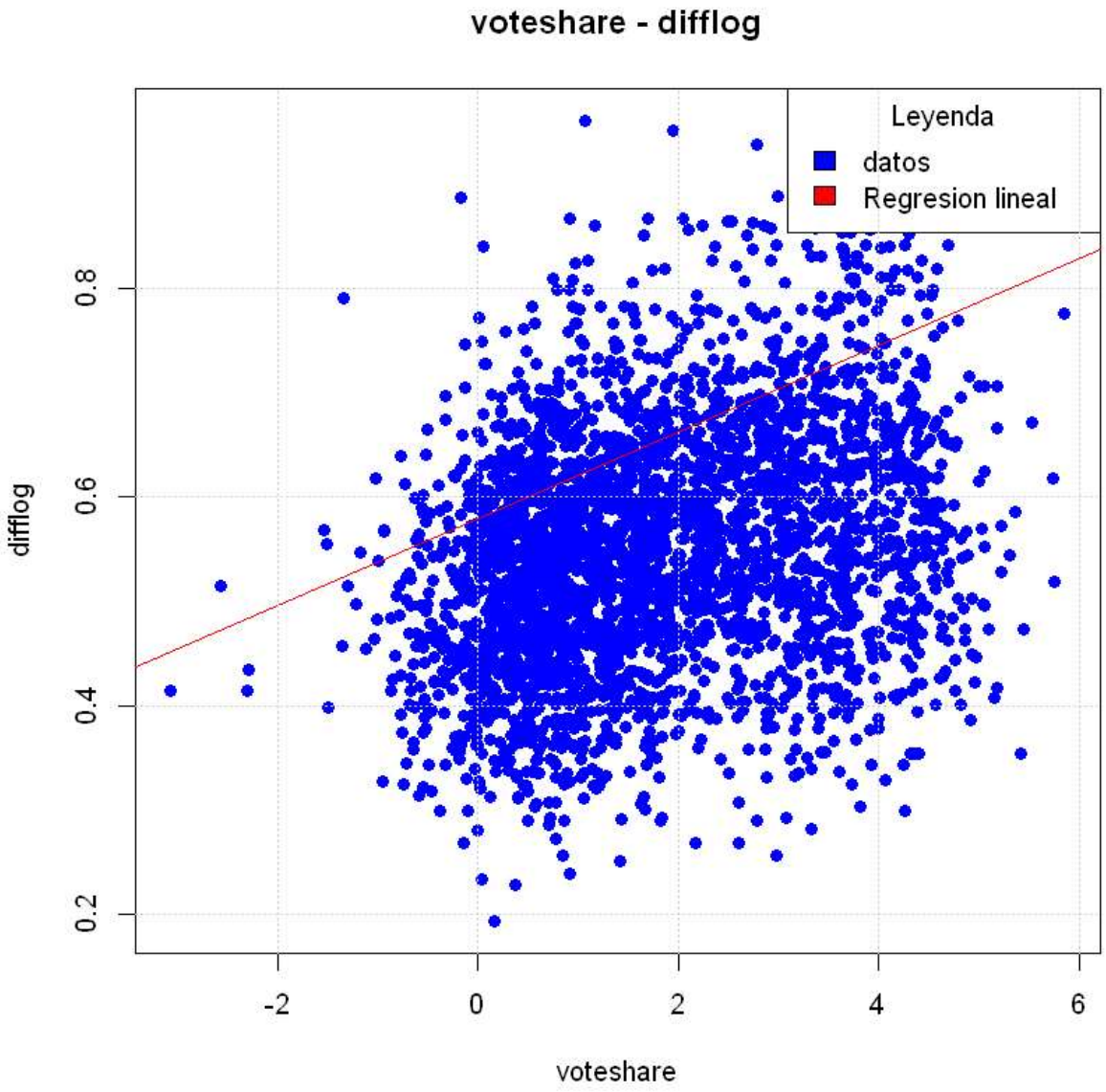
We are interested in knowing how the difference in campaign spending between incumbent and challenger affects the incumbent's vote share.

1. Run a regression where the outcome variable is voteshare and the explanatory variable is difflog.

In [44]:
```
# creamos las variables respectivas
x  = datos$difflog
y  = datos$voteshare
# aplicamos una regresion lineal
regresion1 <- lm(y ~ x)
```

1. Make a scatterplot of the two variables and add the regression line.

In [66]:
```
# graficamos para ver los resultados
plot(x,y,col= "blue",main = "difflog - voteshare", xlab = "difflog", ylab = "votesha
abline(regresion1,col="red")
legend(x = "topright", legend = c("datos", "Regresion lineal"), fill = c("blue", "re
       title = "Leyenda")
grid()
```



1. Save the residuals of the model in a separate object.

In [80]:
```
# guardamos los residuos
r1 = residuals( regresion1 )
```

1. Write the prediction equation

In [79]:
```
resultados = summary(regresion1)
coeficentes = resultados$coefficients
print(paste("Recta de regresion lineal: y = ",coeficentes[2],"x + ",coeficentes[1]))
```

```
[1] "Recta de regresion lineal: y =  0.0416663238227398 x +  0.579030710920673"
```

In [63]:
```
coeficentes[2]
```

0.0416663238227398

Question 2 We are interested in knowing how the difference between incumbent and challenger's

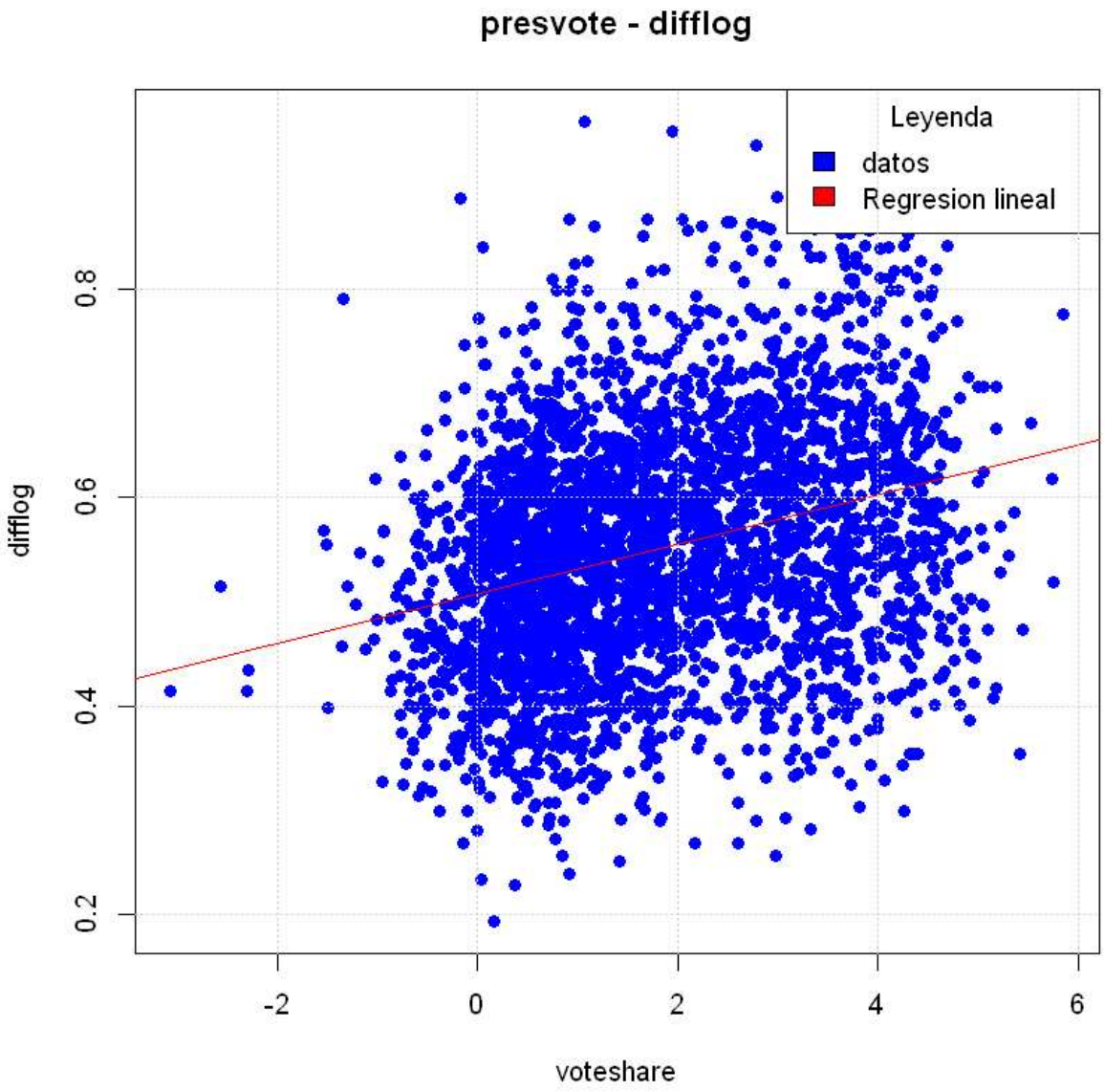
spending and the vote share of the presidential candidate of the incumbent's party are related.

1. Run a regression where the outcome variable is presvote and the explanatory variable is difflog.

In [68]:
```
# creamos las variables respectivas
x  = datos$difflog
y  = datos$presvote
# aplicamos una regresion lineal
regresion2 <- lm(y ~ x)
```

1. Make a scatterplot of the two variables and add the regression line.

In [69]:
```
# graficamos para ver los resultados
plot(x,y,col= "blue",main = "difflog - presvote", xlab = "difflog", ylab = "presvote
abline(regresion2,col="red")
legend(x = "topright", legend = c("datos", "Regresion lineal"), fill = c("blue", "re
       title = "Leyenda")
grid()
```



1. Save the residuals of the model in a separate object

In [81]:
```
# guardamos los residuos
r2 = residuals( regresion2 )
```

1. Write the prediction equation.

In [82]:
```
# guardamos las estadisticas de la regresion
resultados = summary(regresion2)
coeficentes = resultados$coefficients
print(paste("Recta de regresion lineal: y = ",coeficentes[2],"x + ",coeficentes[1]))
```

```
[1] "Recta de regresion lineal: y =  0.0238372338413339 x +  0.507583328405014"
```

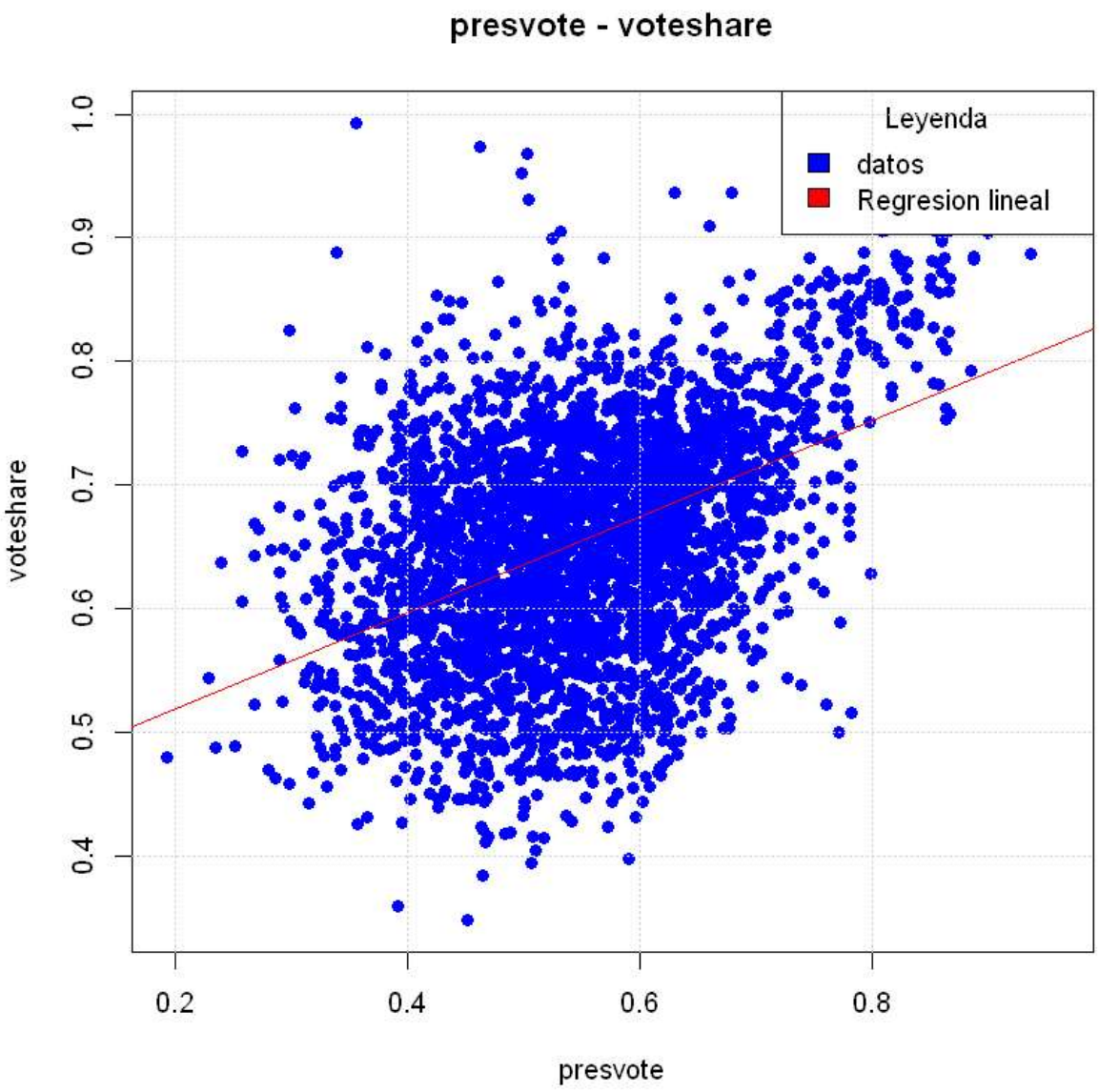
Question 3 We are interested in knowing how the vote share of the presidential candidate of the incumbent's party is associated with the incumbent's electoral success.

1. Run a regression where the outcome variable is voteshare and the explanatory variable is presvote.

In [83]:
```
# creamos las variables respectivas
x  = datos$presvote
y  = datos$voteshare
# aplicamos una regresion lineal
regresion3 <- lm(y ~ x)
```

1. Make a scatterplot of the two variables and add the regression line

In [84]:
```
# graficamos para ver los resultados
plot(x,y,col= "blue",main = "presvote - voteshare", xlab = "presvote", ylab = "votes
abline(regresion3,col="red")
legend(x = "topright", legend = c("datos", "Regresion lineal"), fill = c("blue", "re
       title = "Leyenda")
grid()
```



1. Write the prediction equation.

In [85]:
```
# guardamos las estadisticas de la regresion
resultados = summary(regresion3)
coeficentes = resultados$coefficients
print(paste("Recta de regresion lineal: y = ",coeficentes[2],"x + ",coeficentes[1]))
```

```
[1] "Recta de regresion lineal: y =  0.388018443387439 x +  0.441329881204297"
```

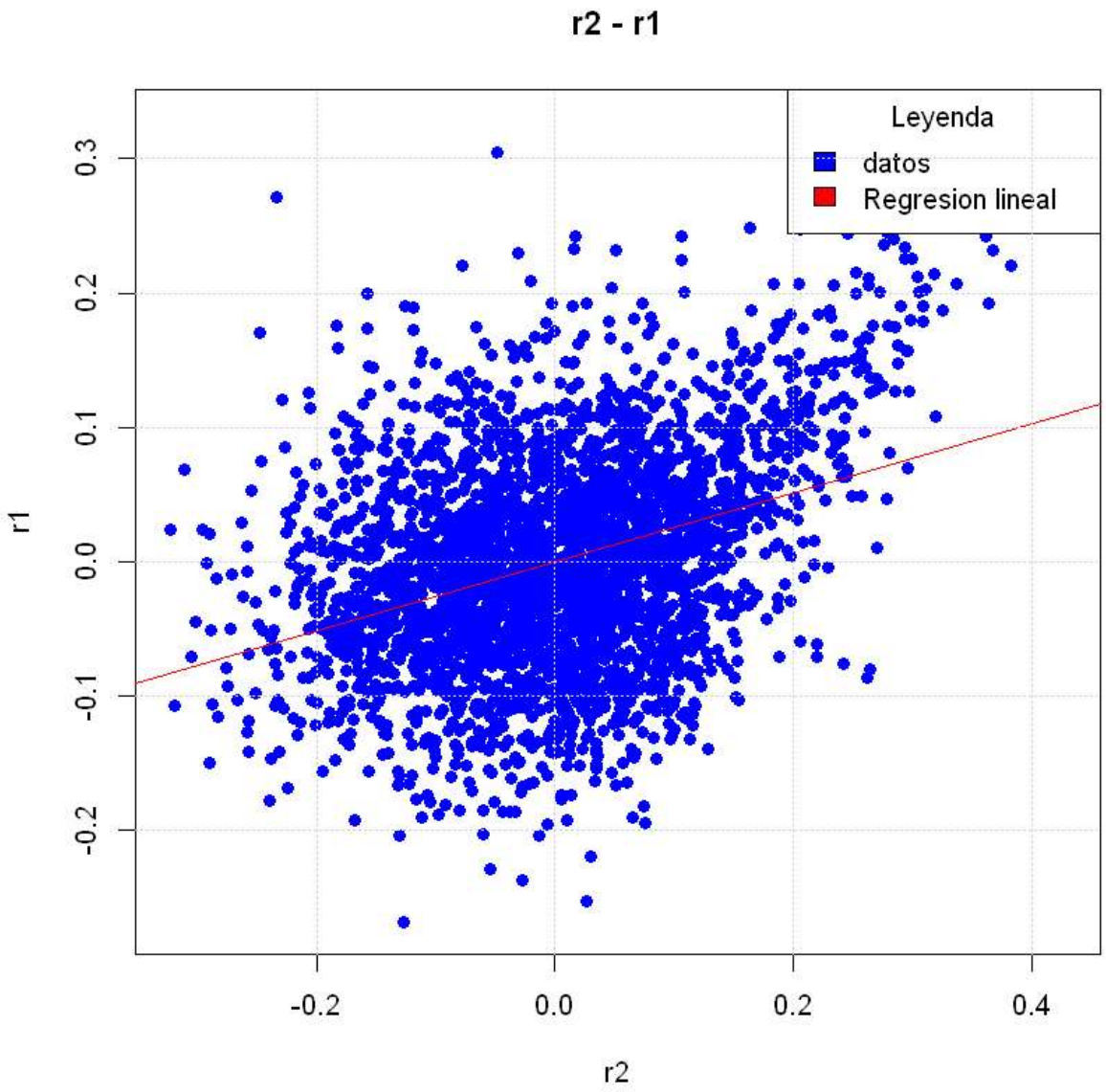
Question 4) The residuals from part (a) tell us how much of the variation in voteshare is not explained by the difference in spending between incumbent and challenger. The residuals in part (b) tell us how much of the variation in presvote is not explained by the difference in spending between incumbent and challenger in the district.

1. Run a regression where the outcome variable is the residuals from Question 1 and the explanatory variable is the residuals from Question 2.

In [90]:
```
# creamos las variables respectivas
x  = r2
y  = r1
# aplicamos una regresion lineal
regresion4 <- lm(y ~ x)
```

1. Make a scatterplot of the two residuals and add the regression line.

In [91]:
```
# graficamos para ver los resultados
plot(x,y,col= "blue",main = "r2 - r1", xlab = "r2", ylab = "r1",pch = 19)
abline(regresion4,col="red")
legend(x = "topright", legend = c("datos", "Regresion lineal"), fill = c("blue", "re
       title = "Leyenda")
grid()
```



1. Write the prediction equation.

In [92]:
```
# guardamos las estadisticas de la regresion
resultados = summary(regresion4)
coeficentes = resultados$coefficients
print(paste("Recta de regresion lineal: y = ",coeficentes[2],"x + ",coeficentes[1]))
```

```
[1] "Recta de regresion lineal: y =  0.256877012700098 x +  -3.14893081409062e-19"
```

Question 5 What if the incumbent's vote share is affected by both the president's popularity and the difference in spending between incumbent and challenger?

1. Run a regression where the outcome variable is the incumbent's voteshare and the explanatory variables are difflog and presvote.

In [106…

```
# creamos las variables respectivas
x  = datos$difflog
x2 = datos$presvote
y  = datos$voteshare
# aplicamos una regresion lineal
multiple = lm(formula = y ~ x + x2, data = datos)
```

1. Write the prediction equation. estadistica = summary(multiple) coeficientes = estadistica$Coefficients

In [121…

```
# guardamos las estadisticas de la regresion
estadisticas = summary(multiple)
c = estadisticas$coefficients
print(paste("Recta de regresion lineal: y = ",c[2],"x1 + ",c[3],"x2 + ",c[1]))
```

```
[1] "Recta de regresion lineal: y =  -2.29068334299495e-06 x1 +  0.388350170800448 x
2 +  0.44933185846982"
```

In [ ]: