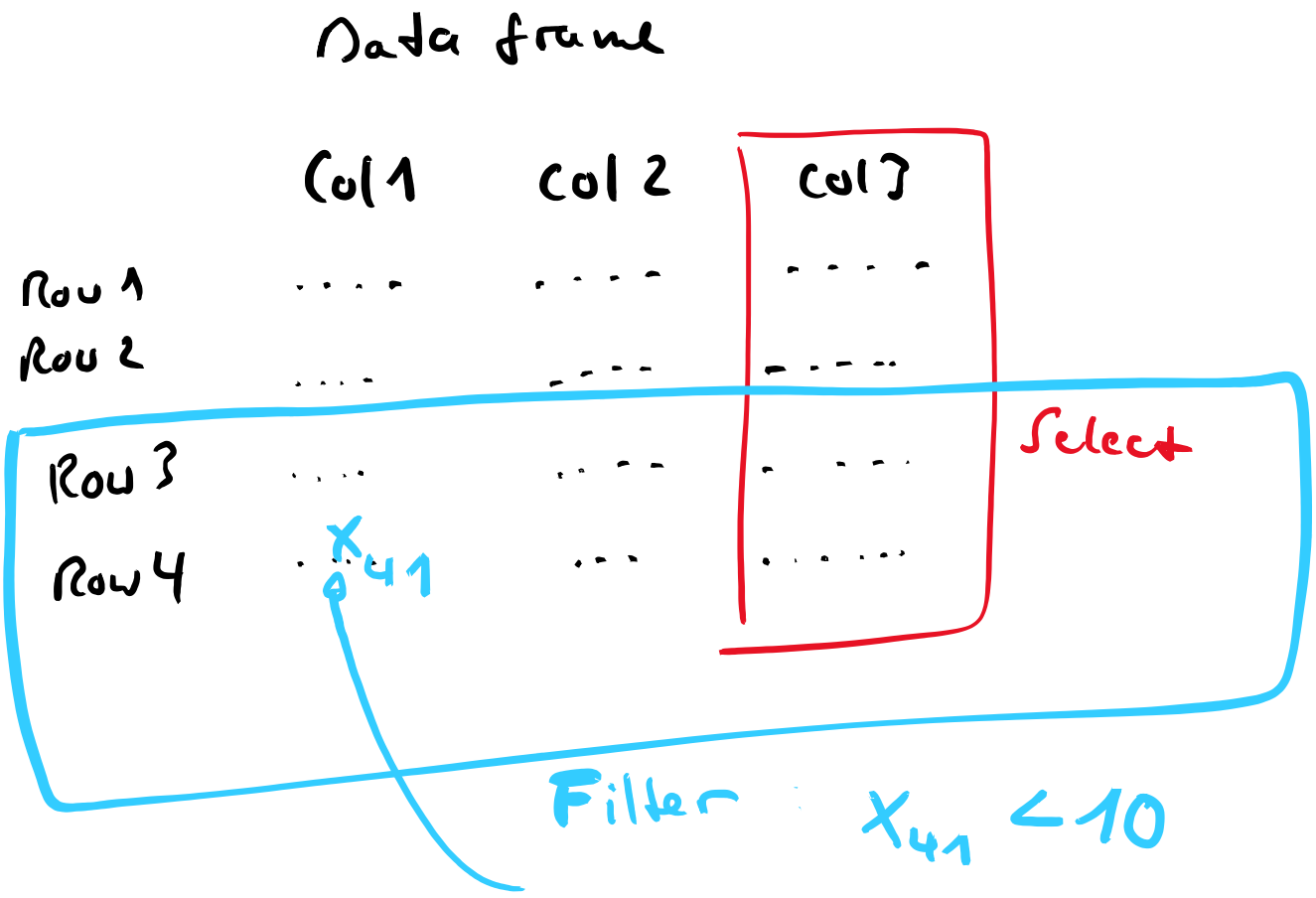
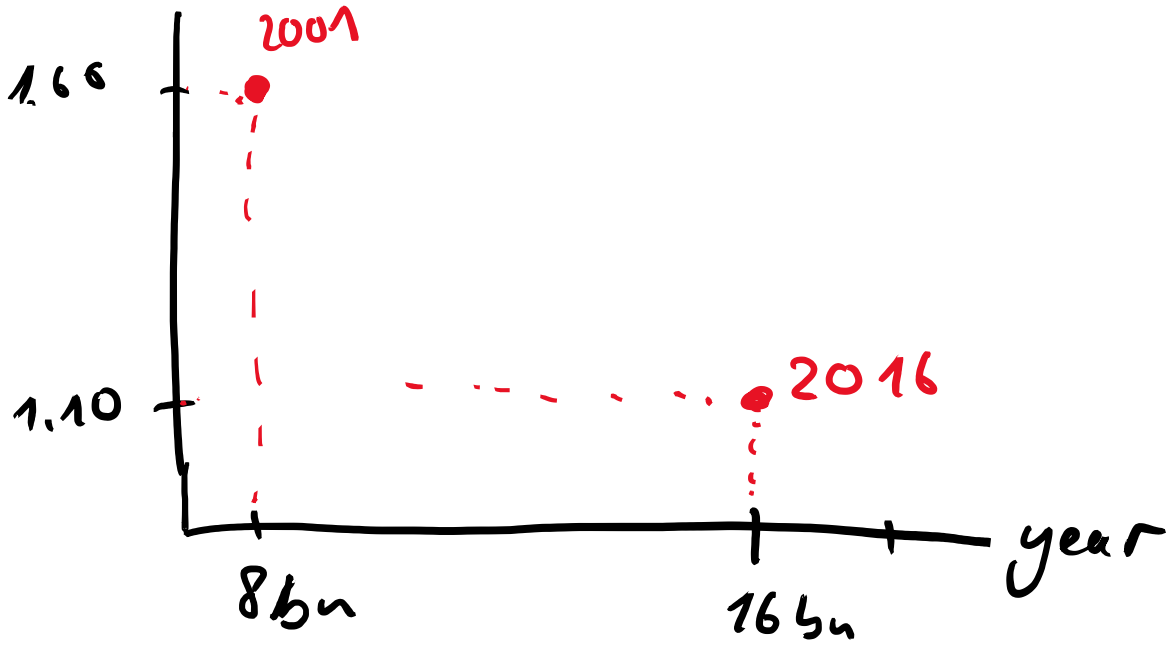
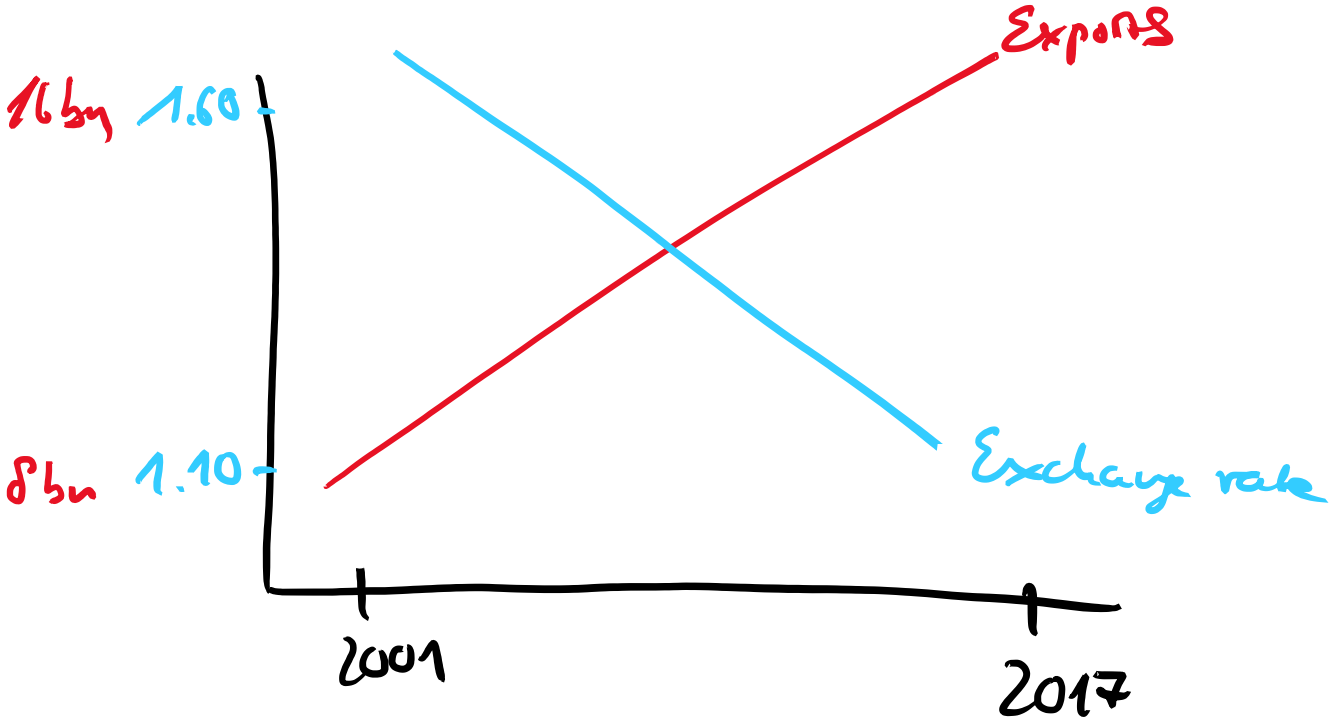

Data frame
Col 1
Col 2
Col 3
Row 1
Row 2
Row 3
Row 4
Select
$x_{41}$
Filter
$x_{41} < 10$



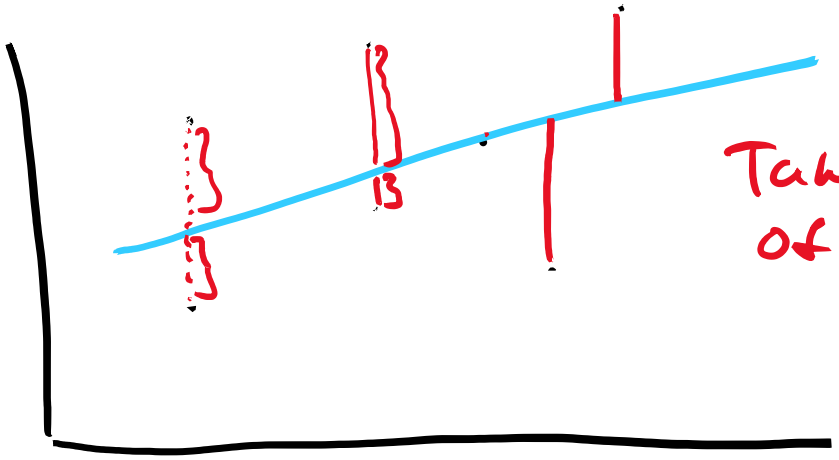

$y = a + bx$ : linear regression

$a \uparrow \quad b \uparrow$

$\alpha \quad \beta$

parametric statistical model

non-parametric

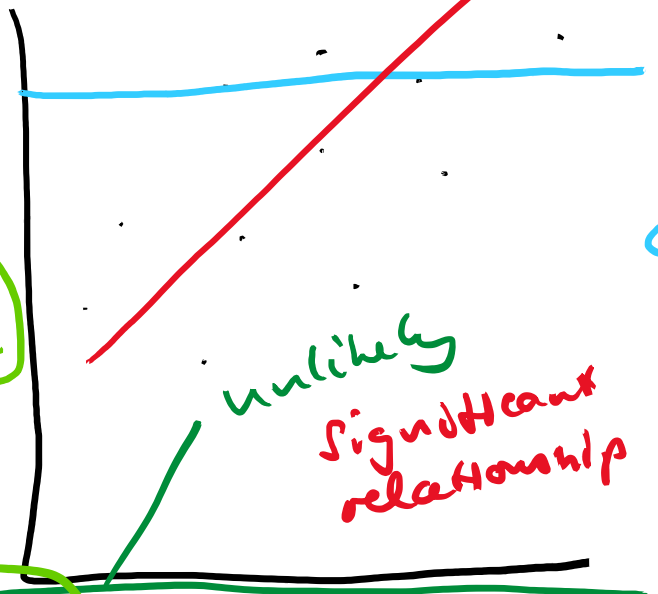


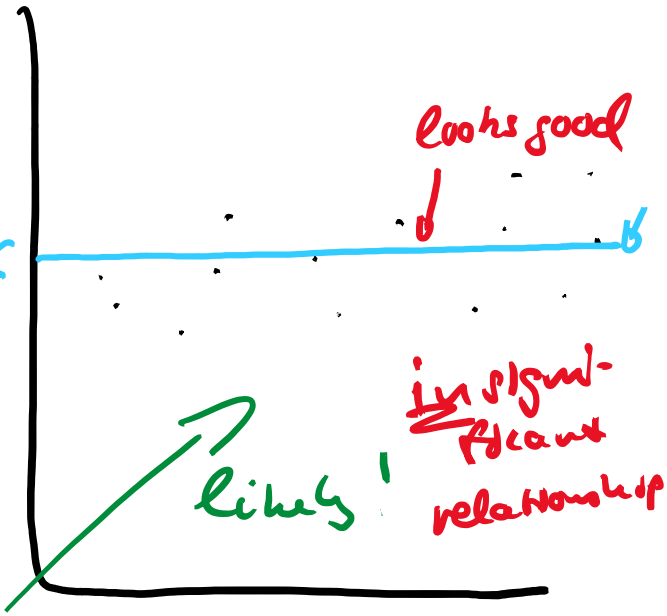
Takes squares of 1

$\Sigma 1$

p-value

much better






p value!

unlikely significant relationship

likely!

insignificant relationship

How likely is it, that the data as we observe it, is generated by a "flat line"?

noise

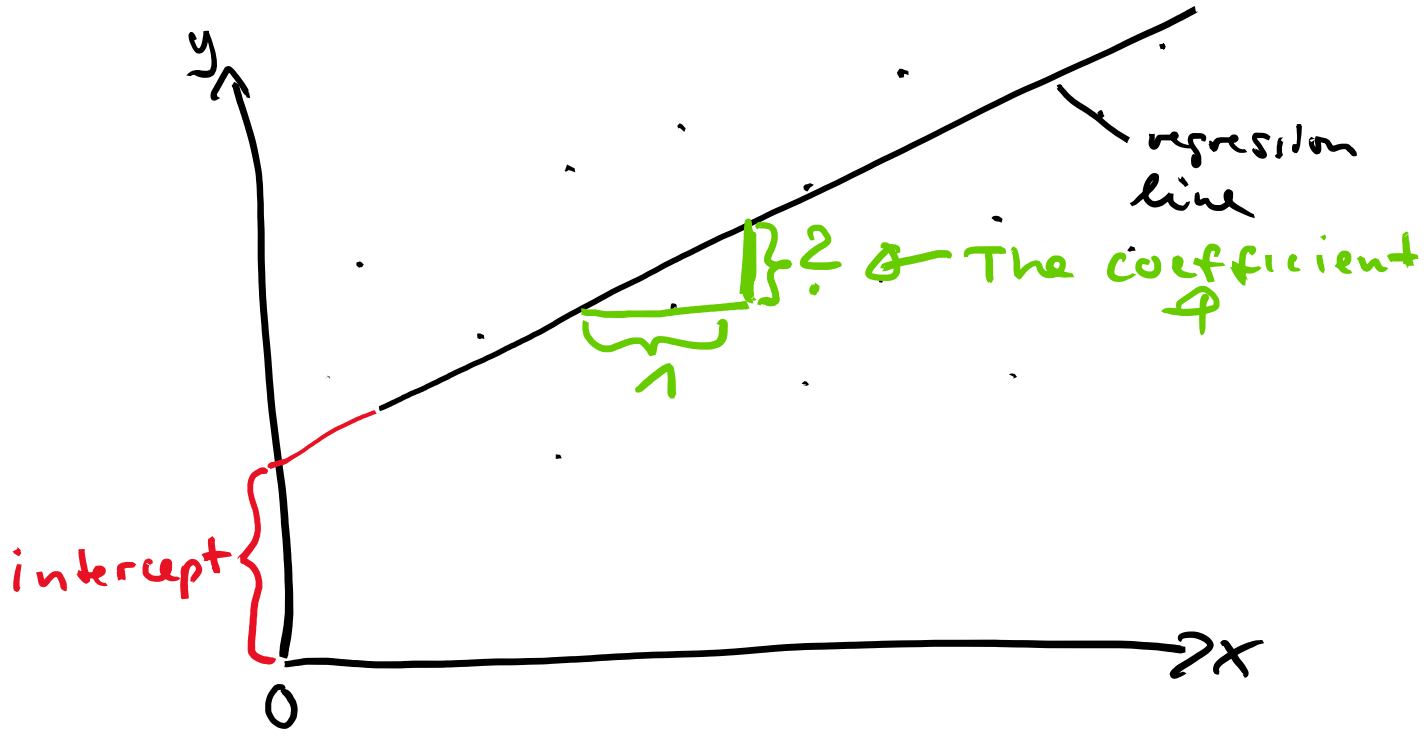
$c + \varepsilon$

* 0.10 (10%)

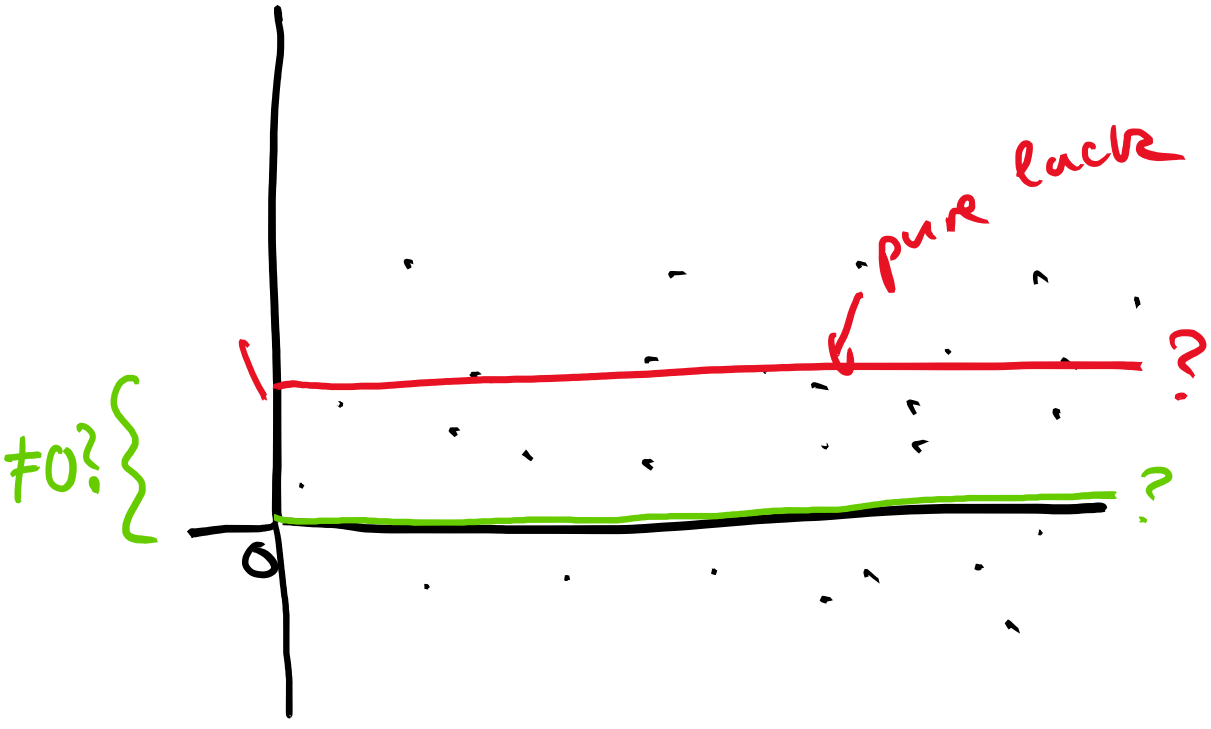
** 0.05 (5%)

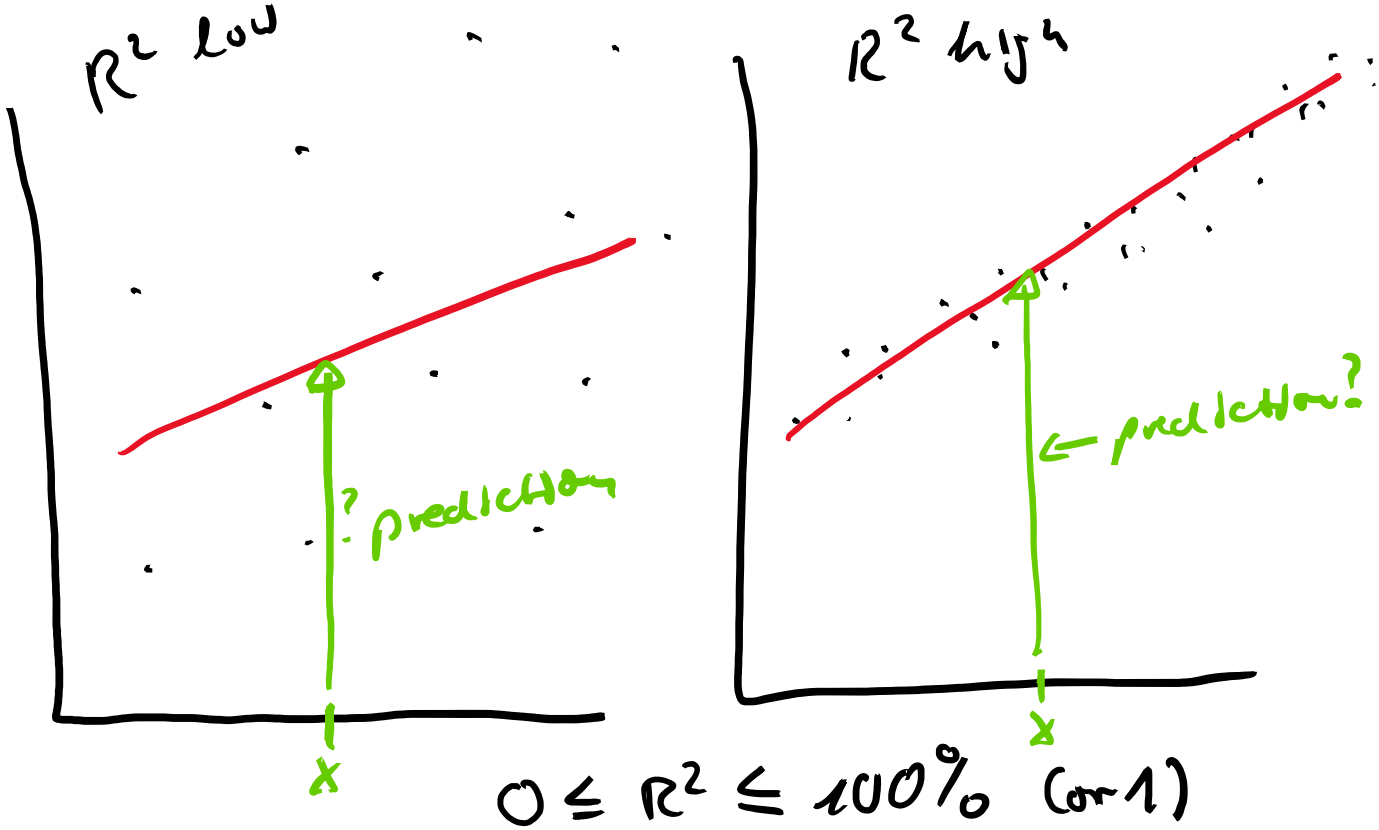
*** 0.01 (1%)

} standards of significance

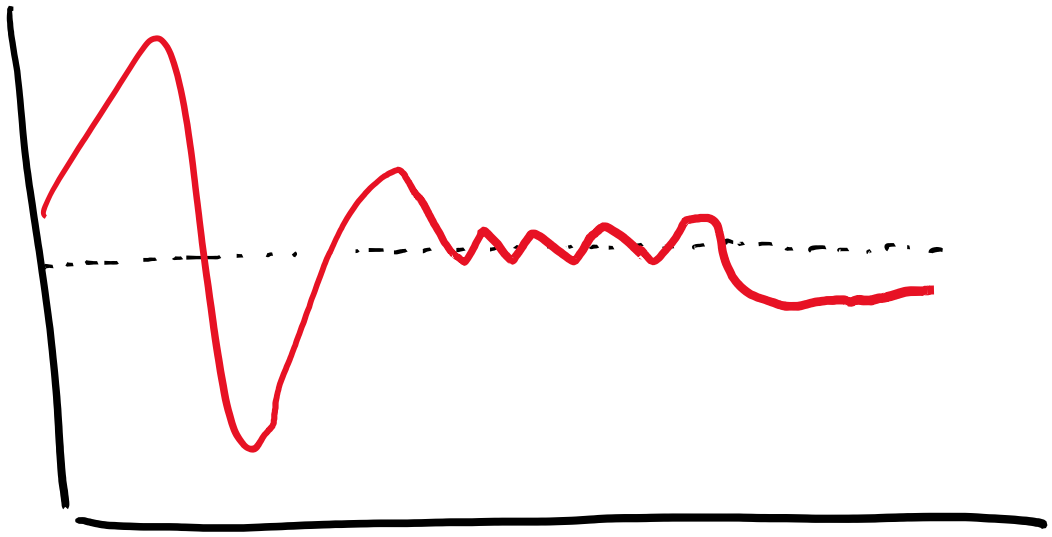

y
regression line
? The coefficient
intercept
0
x


coef = 0
intercept

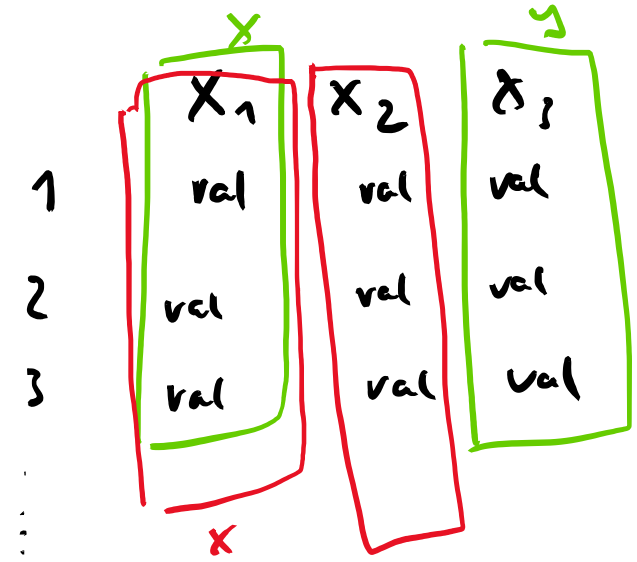

pure luck
?
?
≠0?
0

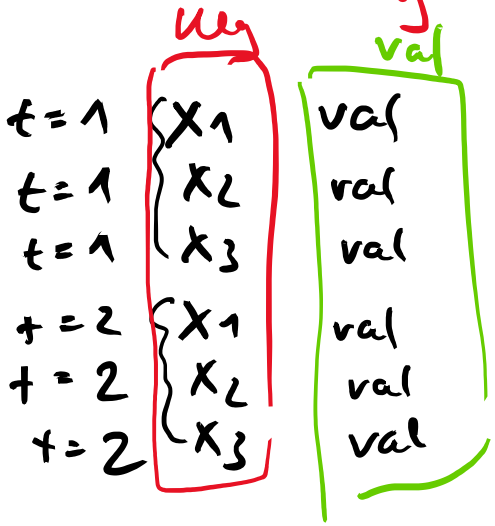



$0 \leq R^2 \leq 100\%$ (or 1)



Data frame




Long format

```
gsplot( ) +
  + geom_line( x1, x3 ) +
  + geom_line(x1, x2)
```

```
aes(x, y, color = Period)
```