$$\ln y_i = \beta_1 + \beta_2 \ln l_i + \beta_3 \ln k_i + \varepsilon_i$$

$$T = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n-k)}$$

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{\beta}_2^2 \mathbf{x}^T \mathbf{x}^{(2)}}{\mathbf{y}^T \mathbf{y}} = \frac{\mathbf{x}^T \mathbf{y}^{(2)}}{(\mathbf{x}^T \mathbf{x})(\mathbf{y}^T \mathbf{y})} = \mathrm{Corr}^2(X, Y)$$

```
library("dplyr")
data <- read.csv("data.csv", header=TRUE)
summary(data)
```

```
str_stand <- function(z) {
  z %>% tolower() %>% str_trim() %>%
    stri_trans_general(id = "Russian-Latin/BGN" ) %>%
    str_replace_all("[[:punct:]]"," ") %>%
    str_replace_all(" +"," ") %>% return()
}
```

```
ct.add <- left_join(d,ct) %>%
  mutate(dist=stringdist(str_stand(user_ans),in_cat)) %>%
  filter(dist<4) %>%
  select(in_cat=user_ans,out_cat) %>%
  unique()
```