

Analisi_finale

Introduzione

Librerie

```
library(lme4)
library(tidyverse)
library(RColorBrewer)
library(lmerTest)
library(gamlss)
```

Descrizione del dataset

```
load("DataRegression2025_unical.RData")
data=data.frame(RiceFarms)
data$id = as.factor(data$id)
attach(data)
which(is.na(data)) #non ci sono missing value.
```

```
## integer(0)
```

```
head(data)
```

```
##      id  size status varieties bimas seed urea phosphate pesticide pseed purea
## 1 101001 3.000  owner      mixed mixed   90  900      80      6000      80    75
## 2 101001 2.000  owner      trad mixed   40  600       0      3000      70    75
## 3 101001 1.000  owner      high mixed  100  700     150      5000     140    70
## 4 101001 2.000  owner      high mixed   60  600     100      5000      90    70
## 5 101001 3.572 share      high   no  105  400     400     10200     350    80
## 6 101001 3.572 share      high   no  105  400     400     10200     250    80
```

```
##      pphosph hiredlabor famlabor totlabor   wage goutput noutput price
## 1      75      2875      40      2915  68.49      7980      6800      60
## 2      75      2110      45      2155  60.09      4083      3500      60
## 3      70       980      95      1075  51.99      2650      2242      65
## 4      70      2081      10      2091  56.98      4500      3750      70
## 5      80      3889       1      3889 152.03     16300     13584     120
## 6      80      3519       1      3519 154.49     17424     14520     140
```

```
##      region
```

```
## 1 wargabinangun
## 2 wargabinangun
## 3 wargabinangun
## 4 wargabinangun
## 5 wargabinangun
## 6 wargabinangun
```

```
str(data)
```

```
## 'data.frame':  1026 obs. of  20 variables:
```

```
## $ id      : Factor w/ 171 levels "101001","101017",...: 1 1 1 1 1 1 2 2 2 2 ...
## $ size    : num  3 2 1 2 3.57 ...
## $ status  : Factor w/ 3 levels "owner","share",...: 1 1 1 1 2 2 3 3 3 1 ...
## $ varieties : Factor w/ 3 levels "trad","high",...: 3 1 2 2 2 2 1 2 2 3 ...
## $ bimas    : Factor w/ 3 levels "no","yes","mixed": 3 3 3 3 1 1 3 3 3 1 ...
## $ seed     : int   90 40 100 60 105 105 50 20 15 7 ...
## $ urea     : int   900 600 700 600 400 400 120 100 150 50 ...
## $ phosphate : int    80 0 150 100 400 400 0 0 50 0 ...
## $ pesticide : int   6000 3000 5000 5000 10200 10200 0 0 900 0 ...
## $ pseed    : num   80 70 140 90 350 250 60 50 130 150 ...
## $ purea    : num   75 75 70 70 80 80 75 75 70 70 ...
## $ pphosph  : num   75 75 70 70 80 80 75 75 70 70 ...
## $ hiredlabor: int   2875 2110 980 2081 3889 3519 670 805 380 40 ...
## $ famlabor  : int    40 45 95 10 1 1 140 50 80 69 ...
## $ totlabor  : int   2915 2155 1075 2091 3889 3519 810 855 460 109 ...
## $ wage     : num   68.5 60.1 52 57 152 ...
## $ goutput   : int   7980 4083 2650 4500 16300 17424 3840 2800 950 240 ...
## $ noutput   : int   6800 3500 2242 3750 13584 14520 3200 2400 800 200 ...
## $ price    : num    60 60 65 70 120 140 60 50 62 60 ...
## $ region    : Factor w/ 6 levels "wargabinangun",...: 1 1 1 1 1 1 1 1 1 1 ...
```

```
names(data)
```

```
## [1] "id"      "size"      "status"      "varieties"  "bimas"
## [6] "seed"    "urea"      "phosphate"   "pesticide"  "pseed"
## [11] "purea"   "pphosph"   "hiredlabor"  "famlabor"   "totlabor"
## [16] "wage"    "goutput"   "noutput"     "price"      "region"
```

Descrizioni variabili

Vediamo una breve descrizione delle variabili presenti nel dataset: - id: identificativo univoco dell'azienda agricola.

- CARATTERISTICHE DEL TERRENO: SUDDIVISE IN ETTARI COLTIVATI E PROPRIETA' DEL TERRENO
 - size: area totale coltivata a riso (in ettari).
 - status: stato della terra coltivata, che può essere:
 - * owner: agricoltori proprietari o affittuari (non mezzadri).
 - * share: mezzadri.
 - * mixed: combinazione delle due categorie precedenti.
 - varieties: tipo di varietà di riso coltivate:
 - * trad: varietà tradizionali.
 - * high: varietà ad alta resa.
 - * mixed: combinazione delle due varietà.
 - bimas: partecipazione al programma di intensificazione BIMAS:
 - * no: non partecipante.
 - * yes: partecipante.
 - * mixed: solo una parte del terreno è registrata nel programma.
- FATTORI DI INPUT PRODUTTIVI: COSTO DELLE MATERIE PRIME E TIPOLOGIA
 - seed: quantità di semi utilizzati (kg).
 - urea: quantità di urea utilizzata (kg).
 - phosphate: quantità di fosfato utilizzata (kg).
 - pesticide: costo dei pesticidi (in Rupiah).
 - pseed: prezzo del seme (in Rupiah per kg).
 - purea: prezzo dell'urea (in Rupiah per kg).
 - pphosph: prezzo del fosfato (in Rupiah per kg).

- INPUT: COSTI DEL PERSONALE E ORE DI LAVORO
 - hiredlabor: ore di lavoro salariato.
 - famlabor: ore di lavoro familiare.
 - totlabor: totale ore di lavoro (escludendo il raccolto).
 - wage: salario della manodopera (in Rupiah per ora).
- PRODUZIONE LORDA E NETTA
 - goutput: produzione lorda di riso (kg).
 - noutput: produzione netta di riso, calcolata sottraendo il costo del raccolto dalla produzione lorda.
 - price: prezzo del riso grezzo (in Rupiah per kg).
- AREE GEOGRAFICHE IN CUI OPERANO LE AZIENDE:
 - region: area geografica di appartenenza dell'azienda agricola, tra:
 - * wargabinangun
 - * langan
 - * gunungwangi
 - * malausma
 - * sukaambit
 - * ciwangi.