# Note su Clusters (Sostenibilità)

1. Cluster Antonio G. ([Cluster\_3features\_2016](file:///Users/edmondo/Library/CloudStorage/OneDrive-ConsiglioNazionaledelleRicerche/WORK_IN_PROGRESS/Antonio-Gattone/res_cluster_Gd_2016.html); [Cluster\_3features\_2022](file:///Users/edmondo/Library/CloudStorage/OneDrive-ConsiglioNazionaledelleRicerche/WORK_IN_PROGRESS/Antonio-Gattone/res_cluster_Gd_2022.html))
2. Cluster Arianna ([*grano\_italia\_rica.pdf*](file:///Users/edmondo/Library/CloudStorage/OneDrive-ConsiglioNazionaledelleRicerche/WORK_IN_PROGRESS/Ecowheataly-R/Note/grano_italia_rica.pdf))
   1. Cluster Ceccarelli (derivato da Arianna) [GitHub](https://github.com/gfgprojects/ecowheataly/tree/main/clustering_pipeline)
3. Confronto SAU Rica vs ISTAT (a livello provinciale)

## Cluster Antonio G.

* Non visualizzare le caratteristiche dei cluster per provincia (essendo il dataset campionario non può essere utilizzato per estendere i risultati al di sotto della soglia regionale- *per la quale ci sono dei pesi da applicare*)
* Presenza di outliers che riducono la variabilità e la possibilità di ottenere gruppi diversificati

## Cluster Arianna

* QUE (Quantity Use efficient ratio):
* come lo Yield include tecniche e management così QUE include gli aspetti climatici (in zone con temperature + alte o eccessiva pioggia sono necessari maggiori trattamenti a prescindere dall’efficienza), più dettagliatamente
* “*nutrient use efficiency (NUE)… The ideal management applications are used to enhance the NUE i.e. application of nutrients at right place and time with right rate. The use of vital plant nutrients is significant for agricultural sustainability. However, the application of essential nutrient is under the influence of plant, climate, and economic conditions of farmers. In general agro-climatic conditions, the crop has fertilizer or NUE below 50%”* ([Ref](https://www.tandfonline.com/doi/full/10.1080/01904167.2019.1676907)).
* Questo spiegherebbe perché non c’è una relazione diretta tra resa e quantità di trattamenti
* Immagine che contiene testo, schermata, diagramma

  Descrizione generata automaticamente

**Se vogliamo avere un QUE (NUE in letteratura) efficace dovremmo scontare la quantità di trattamento per Prec e Evapotraspirazione (trovare una funzione e scalare la quantità di trattamento a seconda delle condizioni) e solo dopo fare la cluster.**

**Raccomandazione**:

**In questo quadro sarebbe utile utilizzare le regioni agrarie per avere le condizioni climatiche su una scala territoriale più piccola (quindi più vicine possibile dalle condizioni affrontate dall’agricoltore).**

## Cluster Ceccarelli

Utilizzando i ratio si assume che le condizioni delle farms (climatiche, disponibilità dell’azienda, stato della pianta) sono tutte uguali e che il maggiore uso di trattamenti è solo dovuto ad inefficienza.

**Raccomandazione**: Ritengo l’assunzione troppo forte.

* Una tabella con le caratteristiche dei cluster potrebbe essere caotica (ne abbiamo 15):
  + Potremmo procedere sicuramente con:
    - elbow plot (di seguito), per condividere come troviamo l'ottimo "k"
    - boxplot per ~ 2 variabili a scelta, come di seguito
  + Ricordo che, come input per *clustering*, usiamo solo le variabili (che avrebbe senso plottare):
    - * 'herbicide\_ratio\_over\_yield'
      * 'elements\_ratio\_over\_yield'
      * 'hours\_of\_machines\_ha\_over\_yield'

Immagine che contiene testo, linea, Diagramma, schermata

Descrizione generata automaticamente

Immagine che contiene testo, diagramma, schermata, linea

Descrizione generata automaticamente

Per contesto, nella cartella output sono già presenti tutte le statistiche che alimentano la definizione dei cluster/agenti della simulazione

* Lato outliers, escludiamo 1679 nel 2016 (unico anno qui analizzato):

## Confronto SAU Rica vs ISTAT (a livello provinciale)

Frumento duro

Immagine che contiene testo, mappa

Descrizione generata automaticamente