



# Progetto Elaborazione Dati Scientifici

“Fat detection su immagini  
di carne”

---

Mattia Savoia, Gabriele Savoia

**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

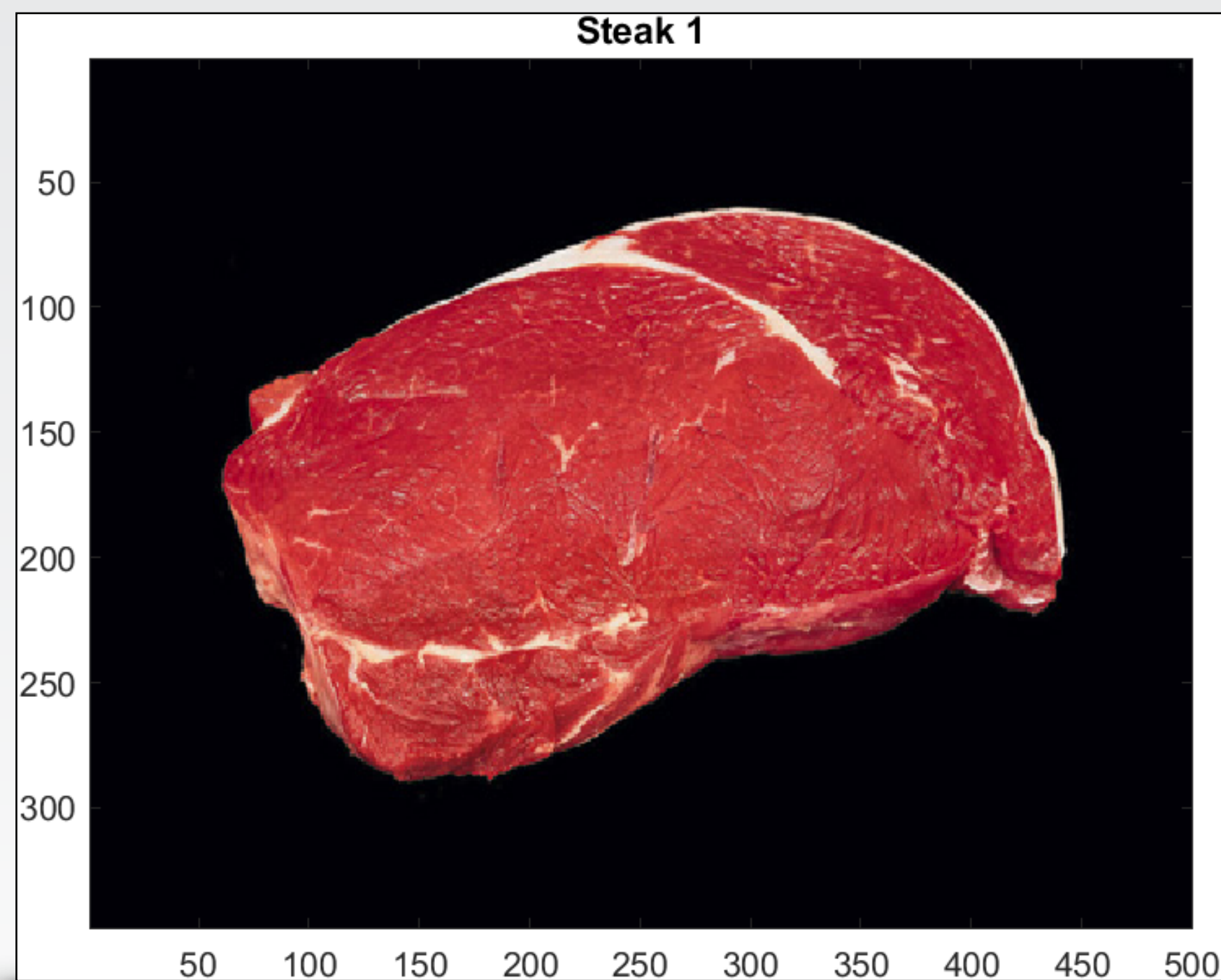
**5. Clustering**

**6. Previsione grasso**

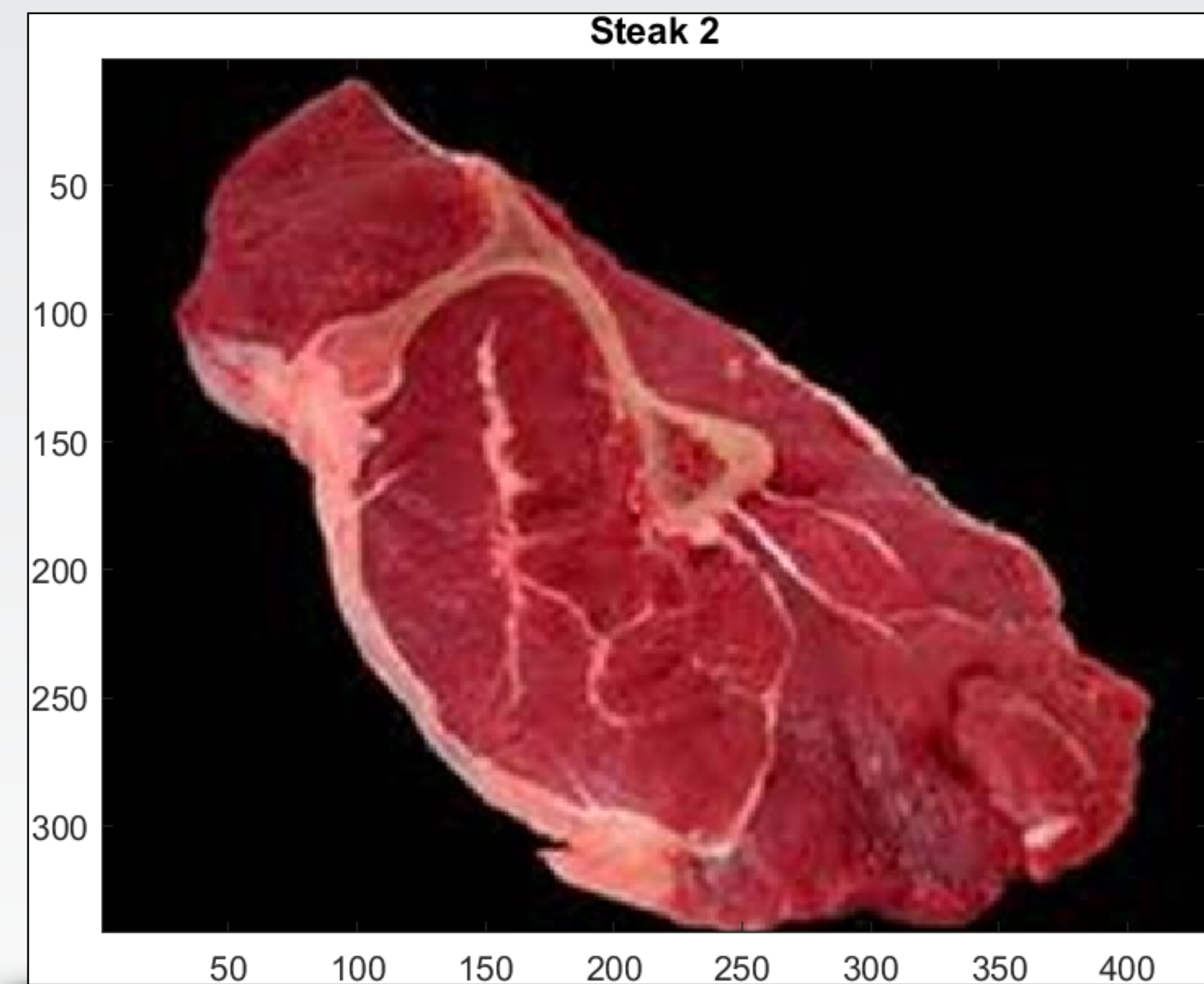
# 1. Descrizione dataset

# 1. Descrizione dataset

Il dataset è composto da due immagini R,G,B che rappresentano due pezzi di carne



500 x 348 px



431 x 341 px

**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

**5. Clustering**

**6. Previsione grasso**

## 2. Obiettivi del progetto

## 2. Obiettivi del progetto

- **Analisi Esplorativa:** estrapolazione di informazioni tramite l'utilizzo di tecniche come la scomposizione per canale e la conversione in scala di grigi (con i relativi istogrammi)

## 2. Obiettivi del progetto

- **Analisi Esplorativa:** estrapolazione di informazioni tramite l'utilizzo di tecniche come la scomposizione per canale e la conversione in scala di grigi (con i relativi istogrammi)
- **PCA Esplorativa:** usata per capire se proiettando i dati in uno spazio dimensionalmente più piccolo è possibile distinguere le due categorie carne e grasso



## 2. Obiettivi del progetto

- **Analisi Esplorativa:** estrapolazione di informazioni tramite l'utilizzo di tecniche come la scomposizione per canale e la conversione in scala di grigi (con i relativi istogrammi)
- **PCA Esplorativa:** usata per capire se proiettando i dati in uno spazio dimensionalmente più piccolo è possibile distinguere le due categorie carne e grasso
- **Clustering:** confronto tra metodi di clustering implementati con l'obiettivo di individuare “sfondo”, “carne” e “grasso” nell'immagine su cui sono applicati

## 2. Obiettivi del progetto

- **Analisi Esplorativa:** estrapolazione di informazioni tramite l'utilizzo di tecniche come la scomposizione per canale e la conversione in scala di grigi (con i relativi istogrammi)
- **PCA Esplorativa:** usata per capire se proiettando i dati in uno spazio dimensionalmente più piccolo è possibile distinguere le due categorie carne e grasso
- **Clustering:** confronto tra metodi di clustering implementati con l'obiettivo di individuare “sfondo”, “carne” e “grasso” nell'immagine su cui sono applicati
- **Previsione del grasso:** creazione di un modello PCA (basato solo sui pixel che definiscono il cluster del grasso) in grado di identificare il grasso su un'immagine di test mai vista

**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

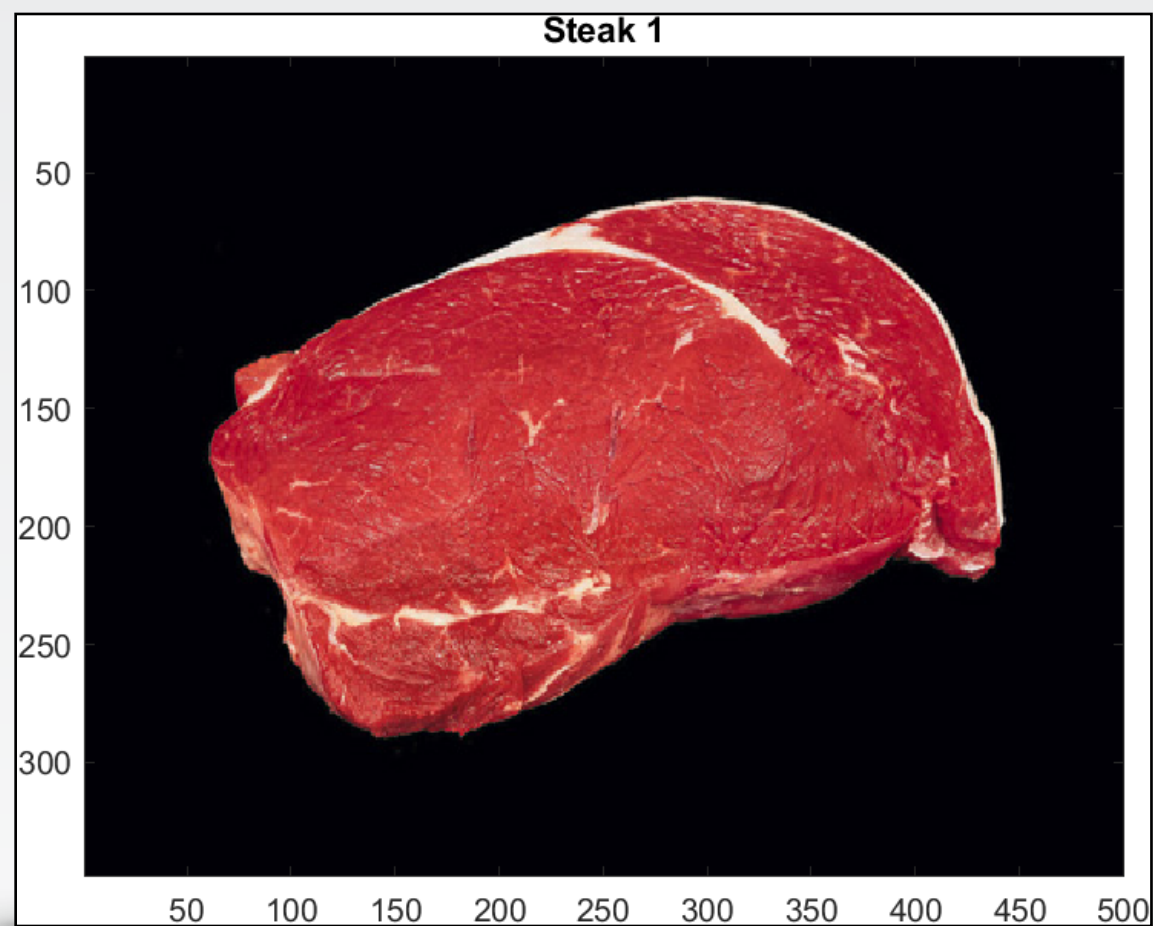
**5. Clustering**

**6. Previsione grasso**

### **3. Analisi esplorativa**

### 3. Analisi esplorativa

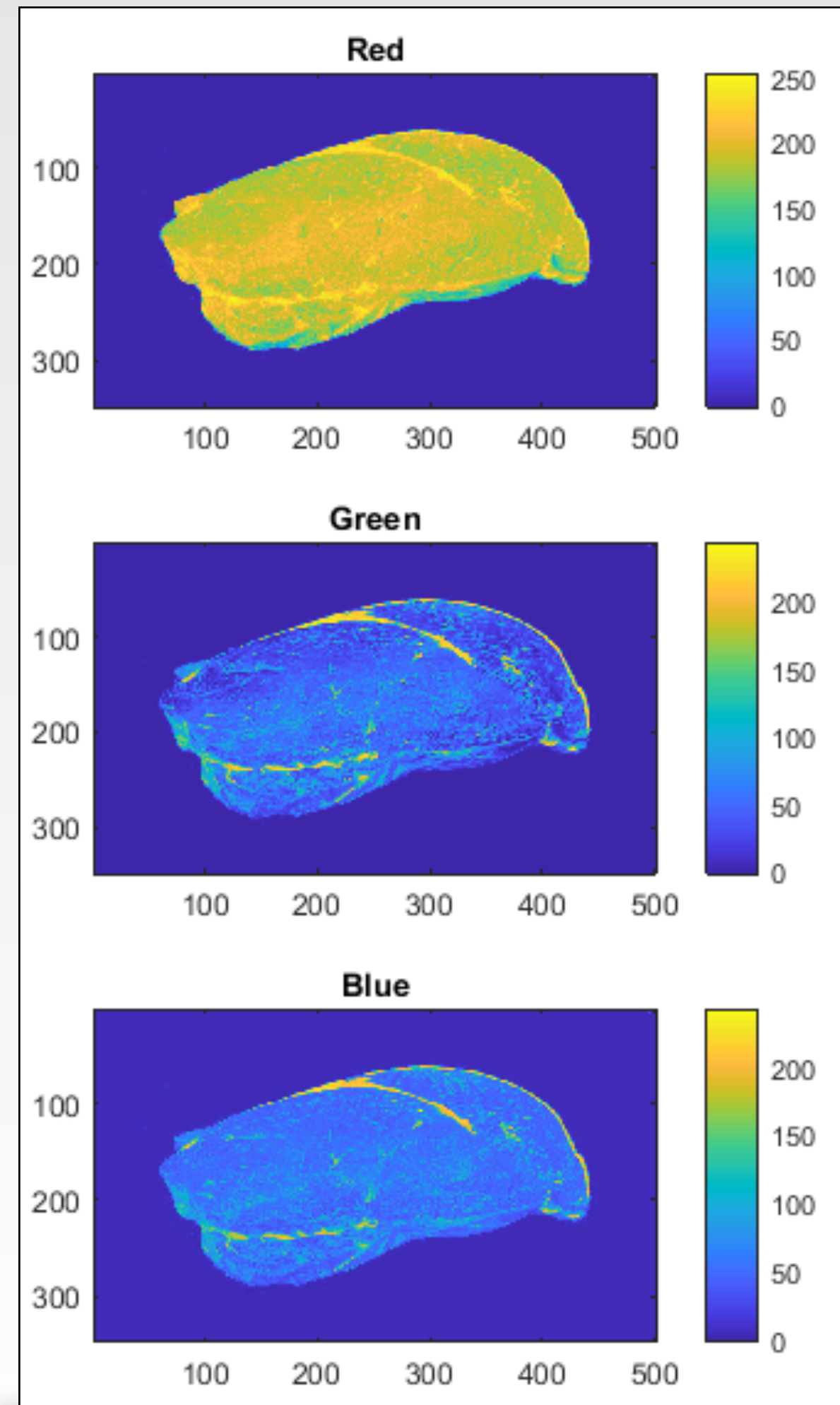
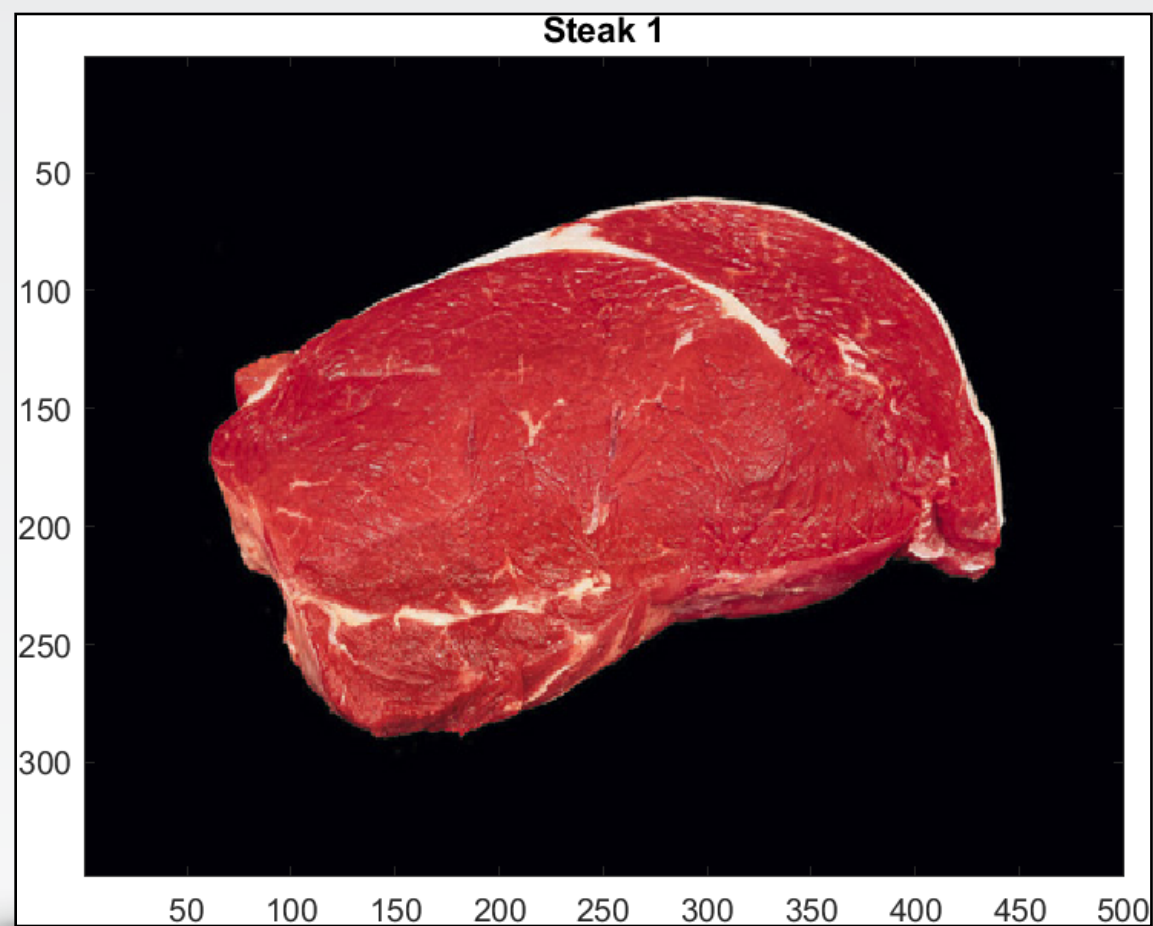
Analisi svolta solo su  
“steak1”



### 3. Analisi esplorativa

Canali R G B

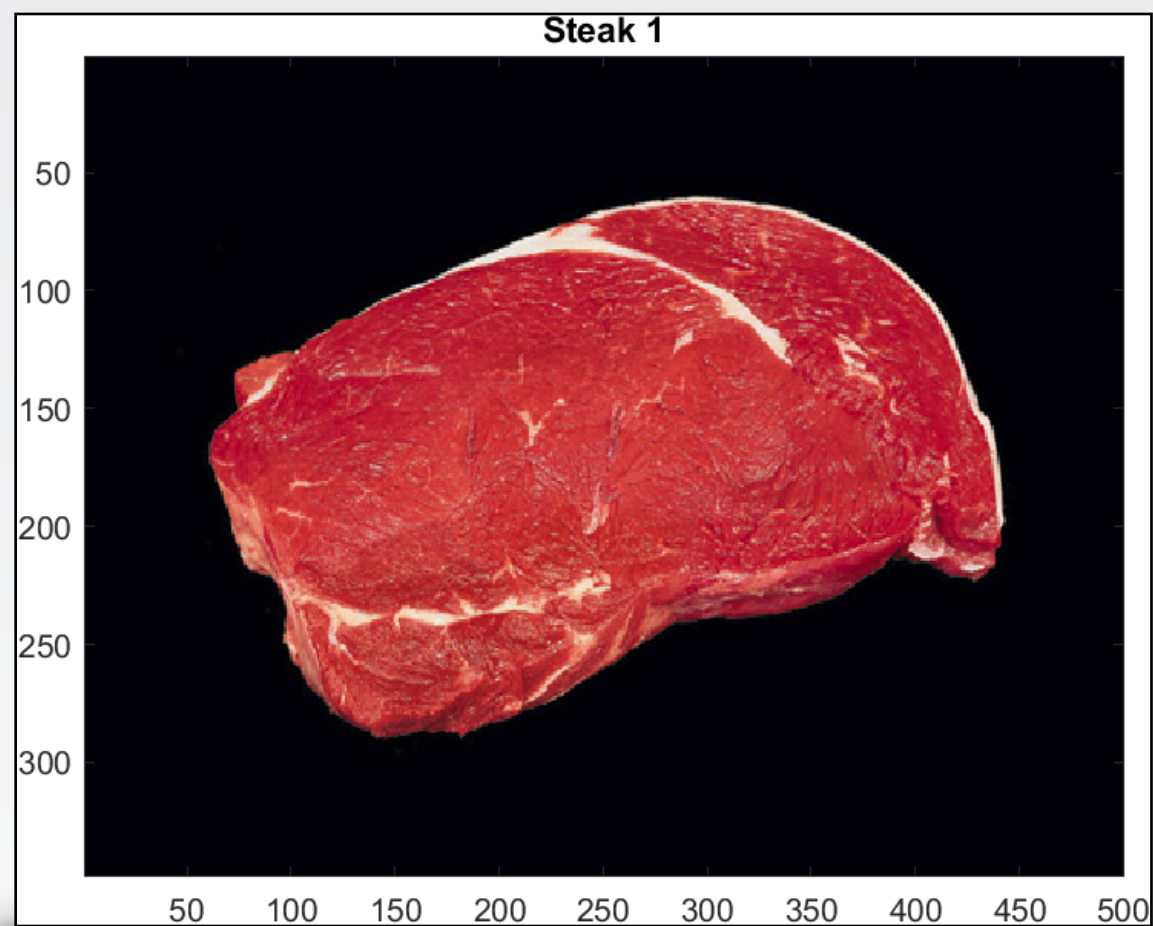
Analisi svolta solo su  
“steak1”



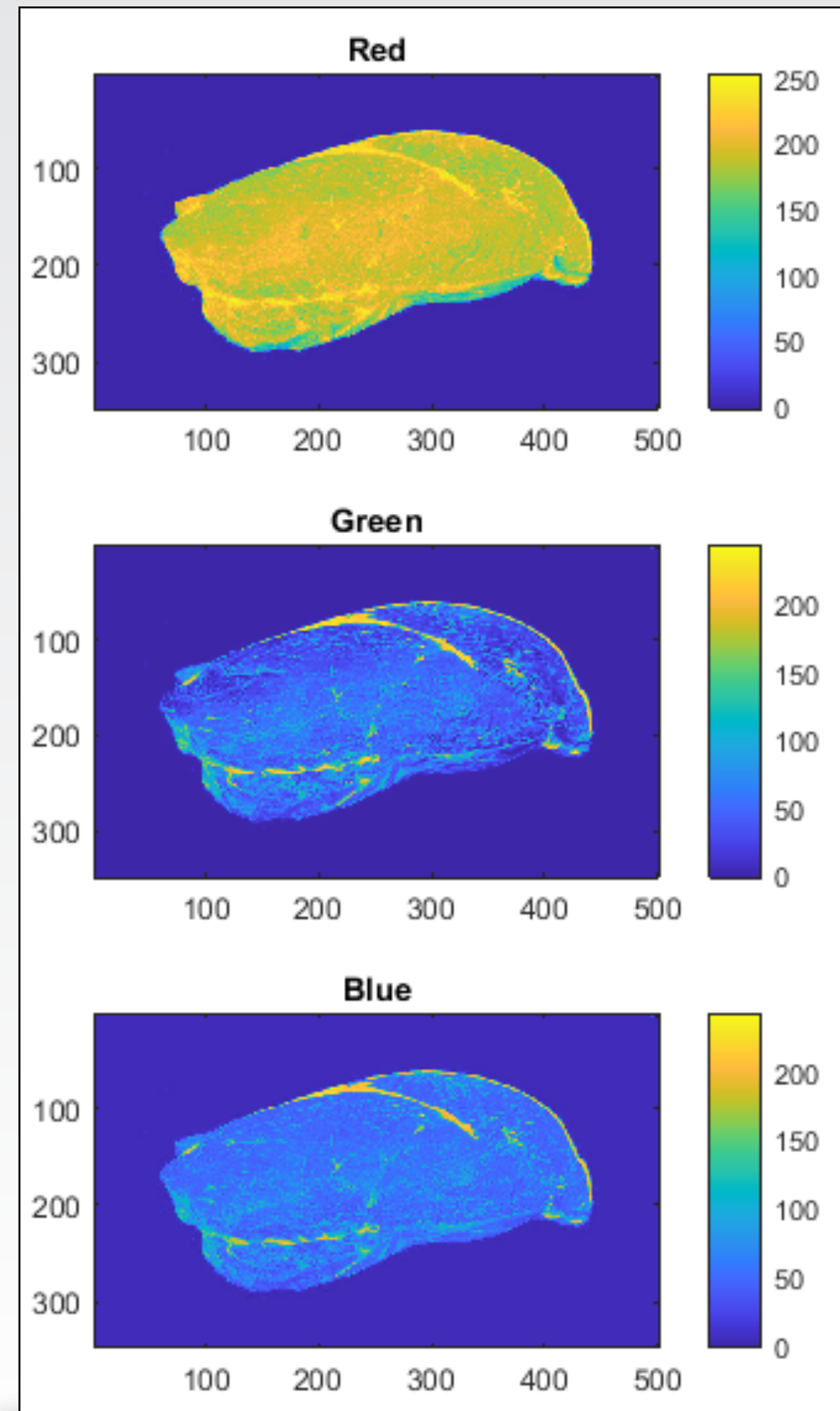


### 3. Analisi esplorativa

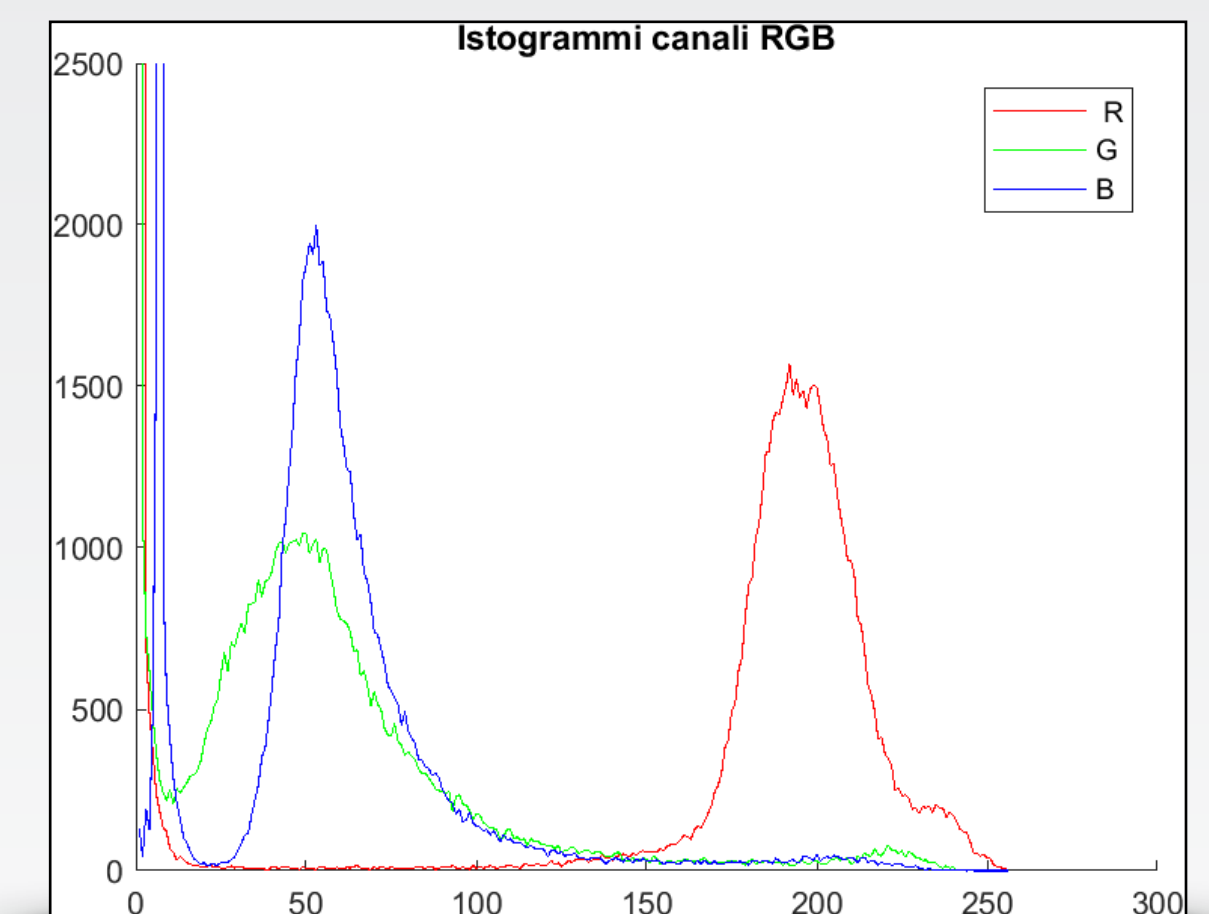
Analisi svolta solo su  
“steak1”



Canali R G B



- **ROSSO:** distingue carne da sfondo
- **VERDE / BLU:** risalta grasso dal resto

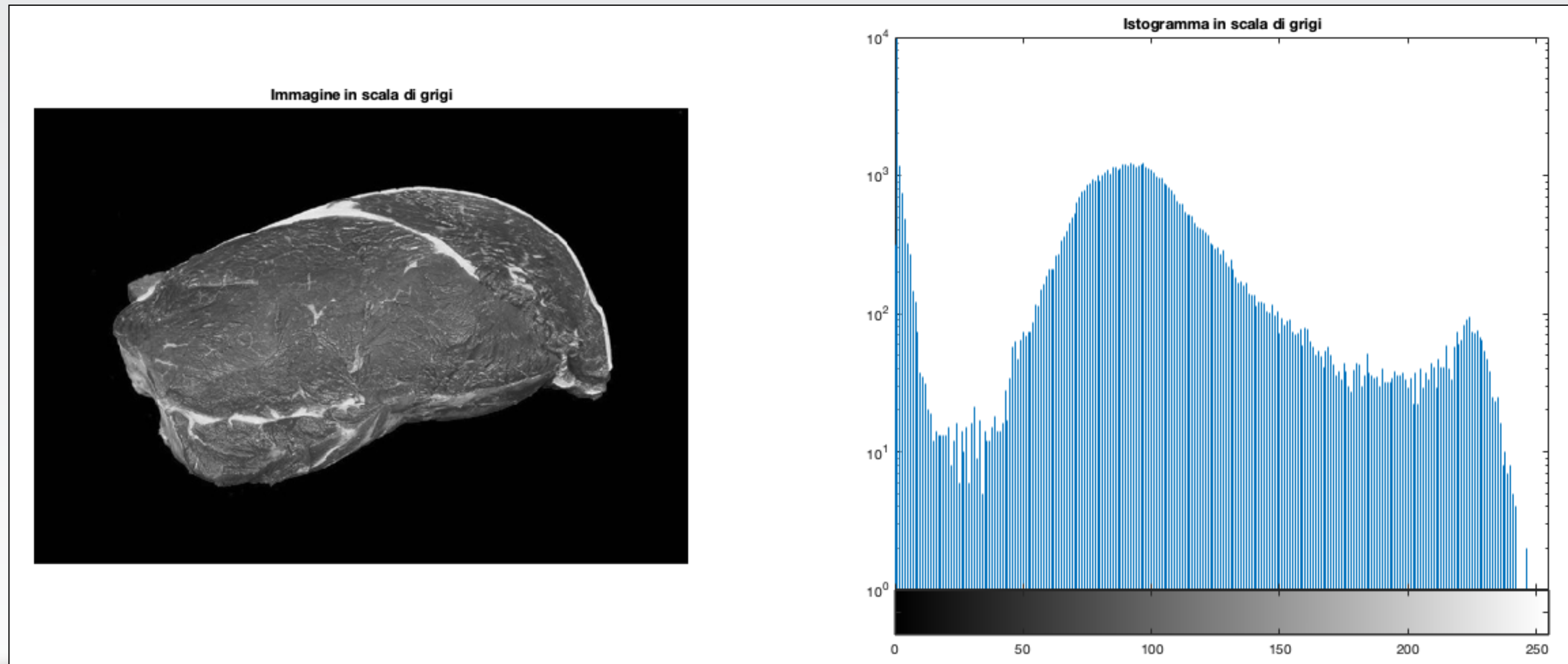


### 3. Analisi esplorativa



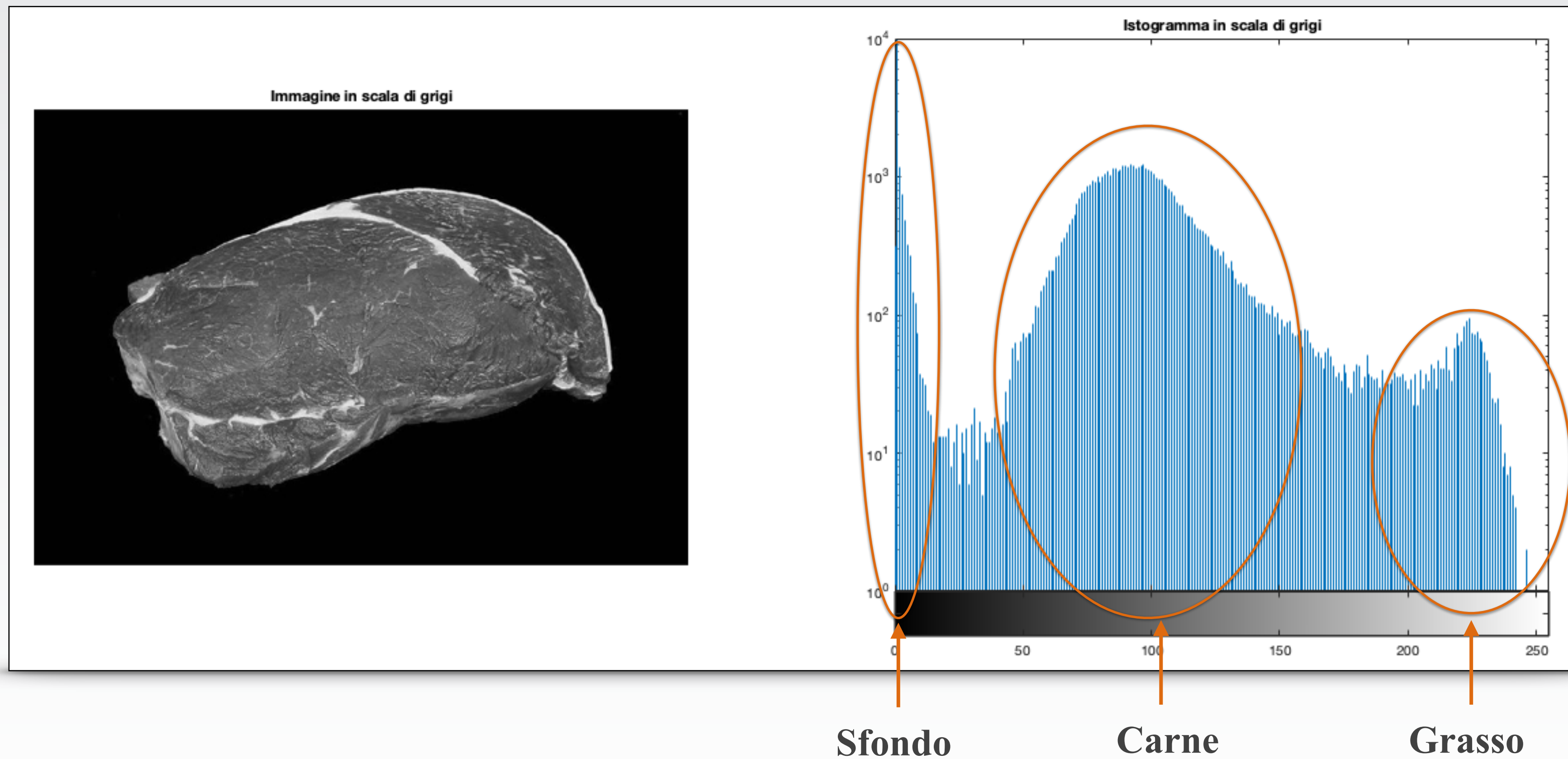
### 3. Analisi esplorativa

Rappresentazione immagine in **scala di grigi** con relativo **istogramma di frequenze**



### 3. Analisi esplorativa

Rappresentazione immagine in **scala di grigi** con relativo **istogramma di frequenze**



**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

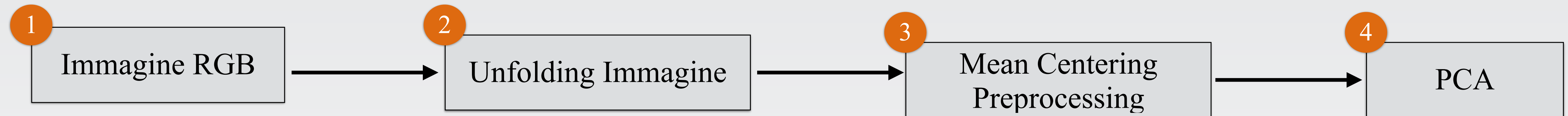
**5. Clustering**

**6. Previsione grasso**

## 4. PCA esplorativa

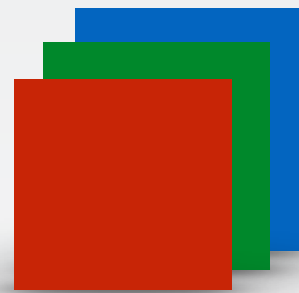
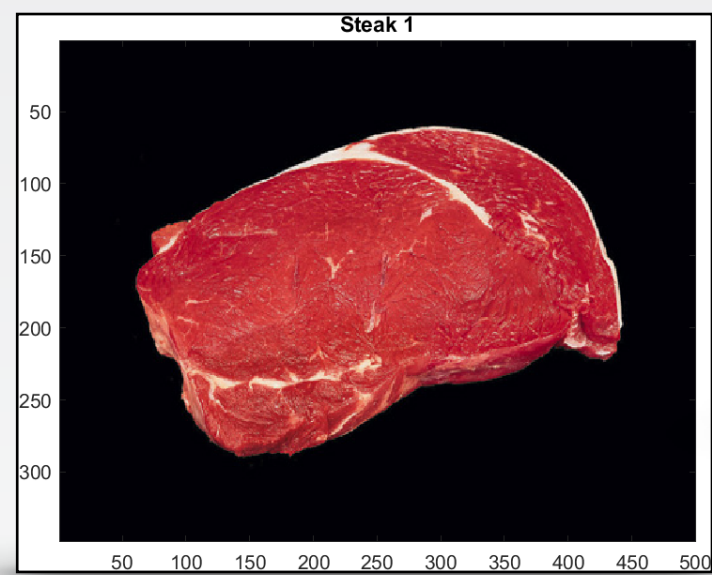
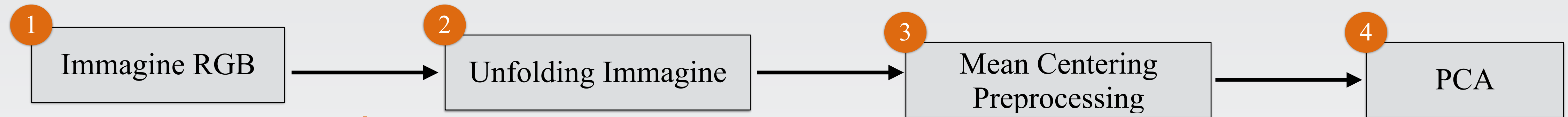
## 4. PCA esplorativa

### Costruzione del modello PCA



## 4. PCA esplorativa

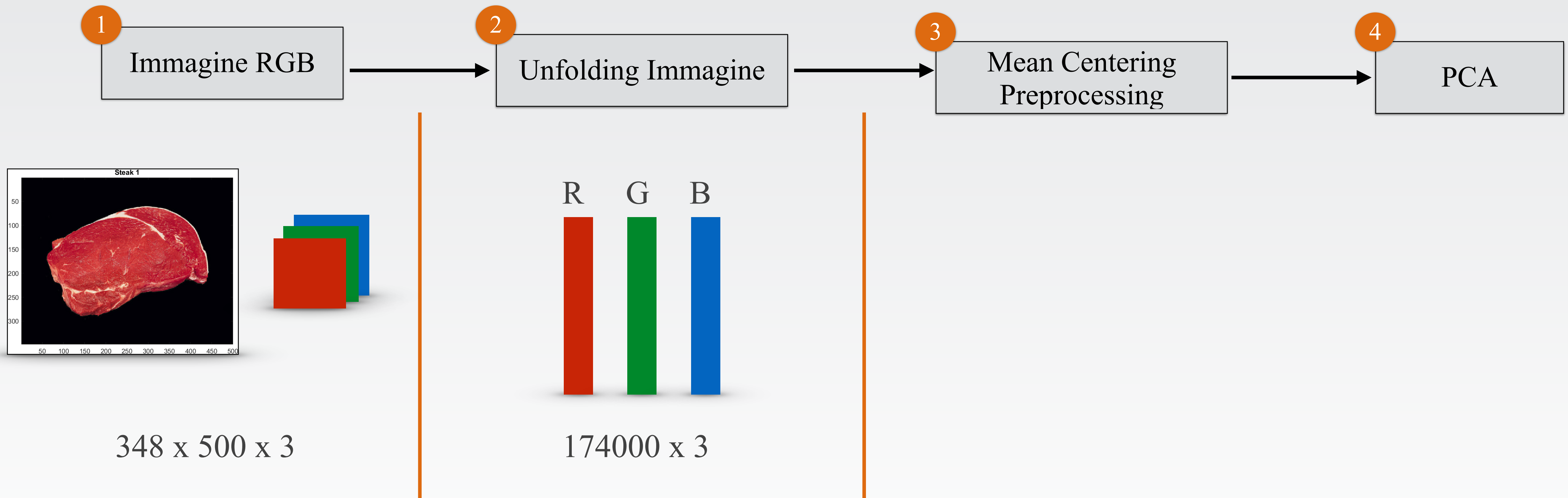
### Costruzione del modello PCA



348 x 500 x 3

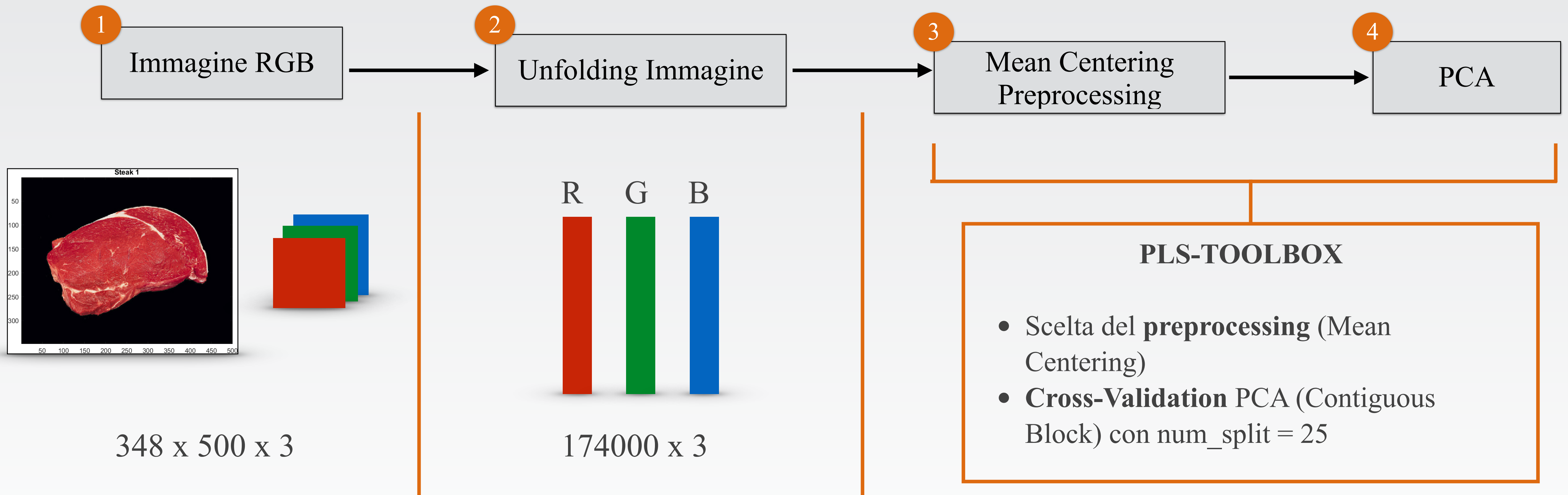
## 4. PCA esplorativa

### Costruzione del modello PCA



## 4. PCA esplorativa

### Costruzione del modello PCA





## 4. PCA esplorativa

## **4. PCA esplorativa**

**Scelta numero delle componenti  
modello PCA**

## 4. PCA esplorativa

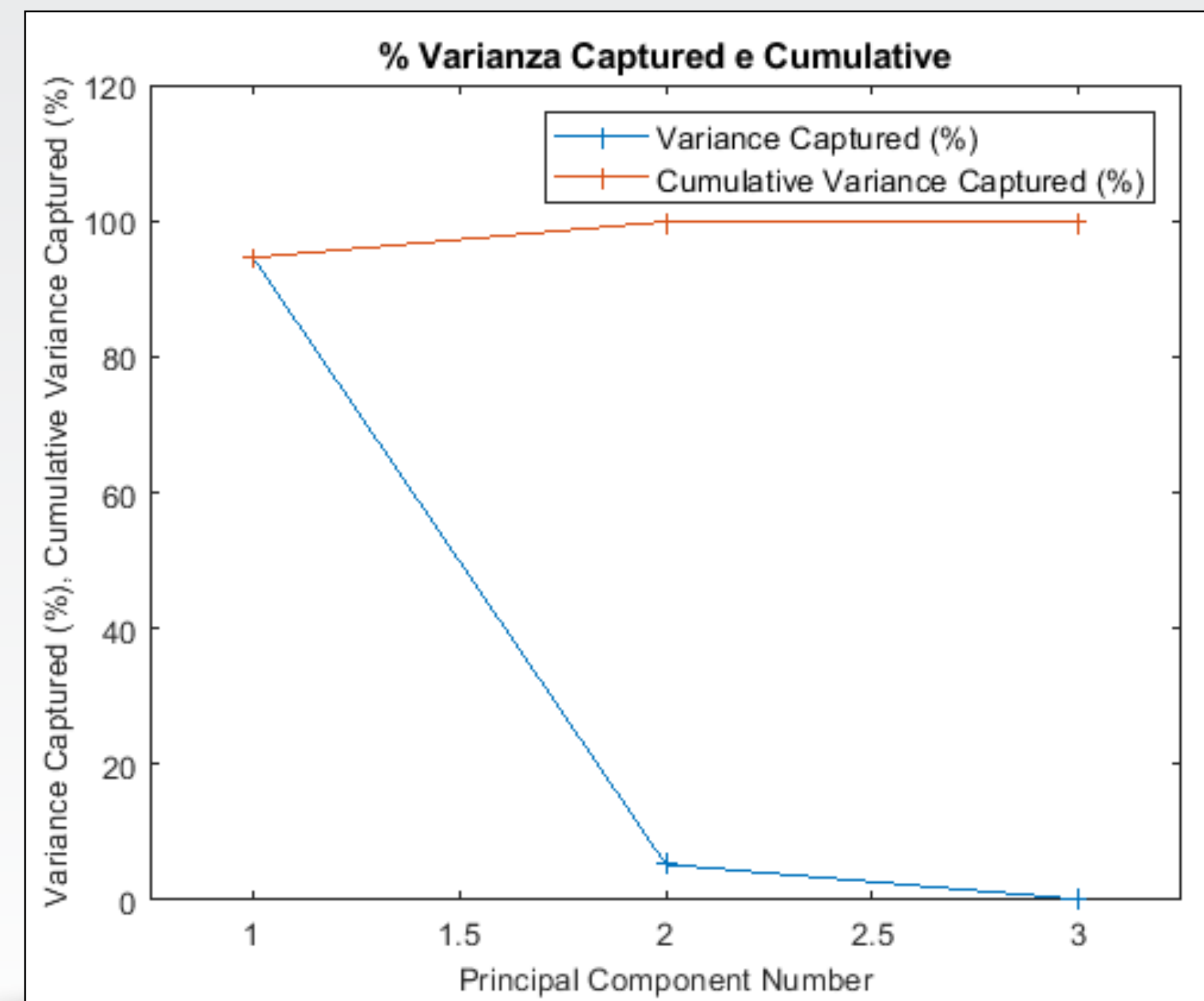
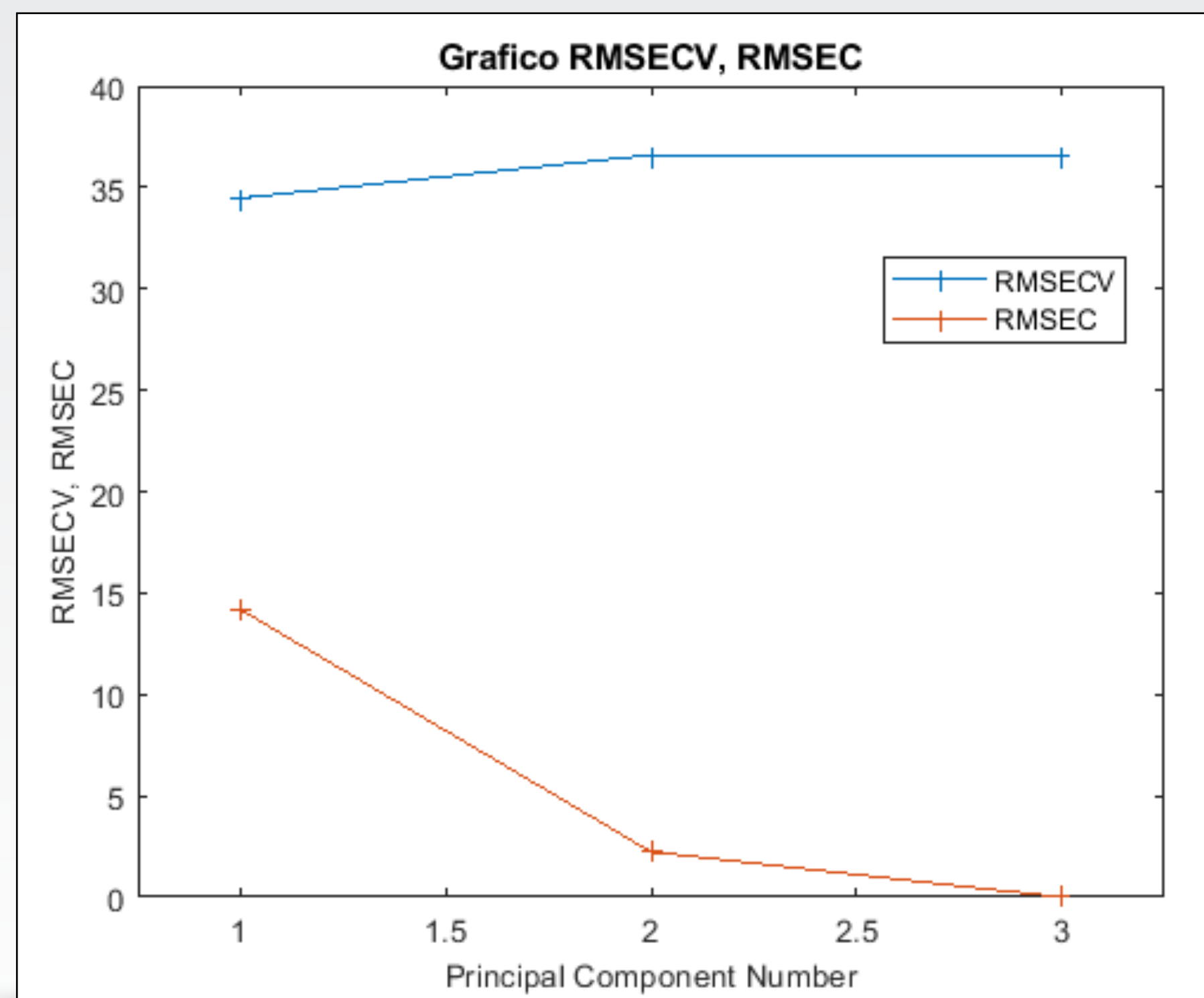
**Numero PC ottimale trovato  
è 1**

**Scelta numero delle componenti  
modello PCA**

## 4. PCA esplorativa

Numero PC ottimale trovato  
è 1

Scelta numero delle componenti  
modello PCA

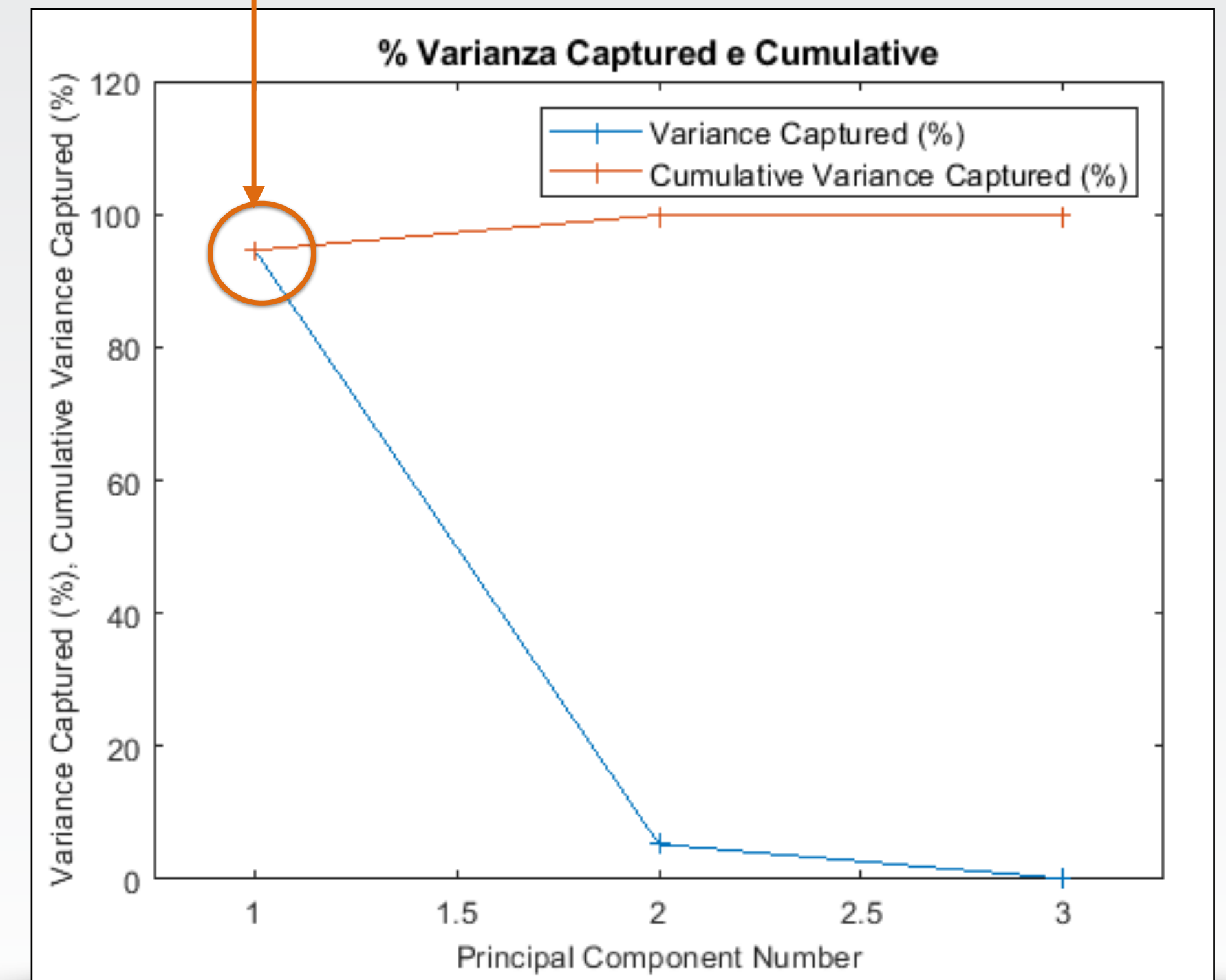
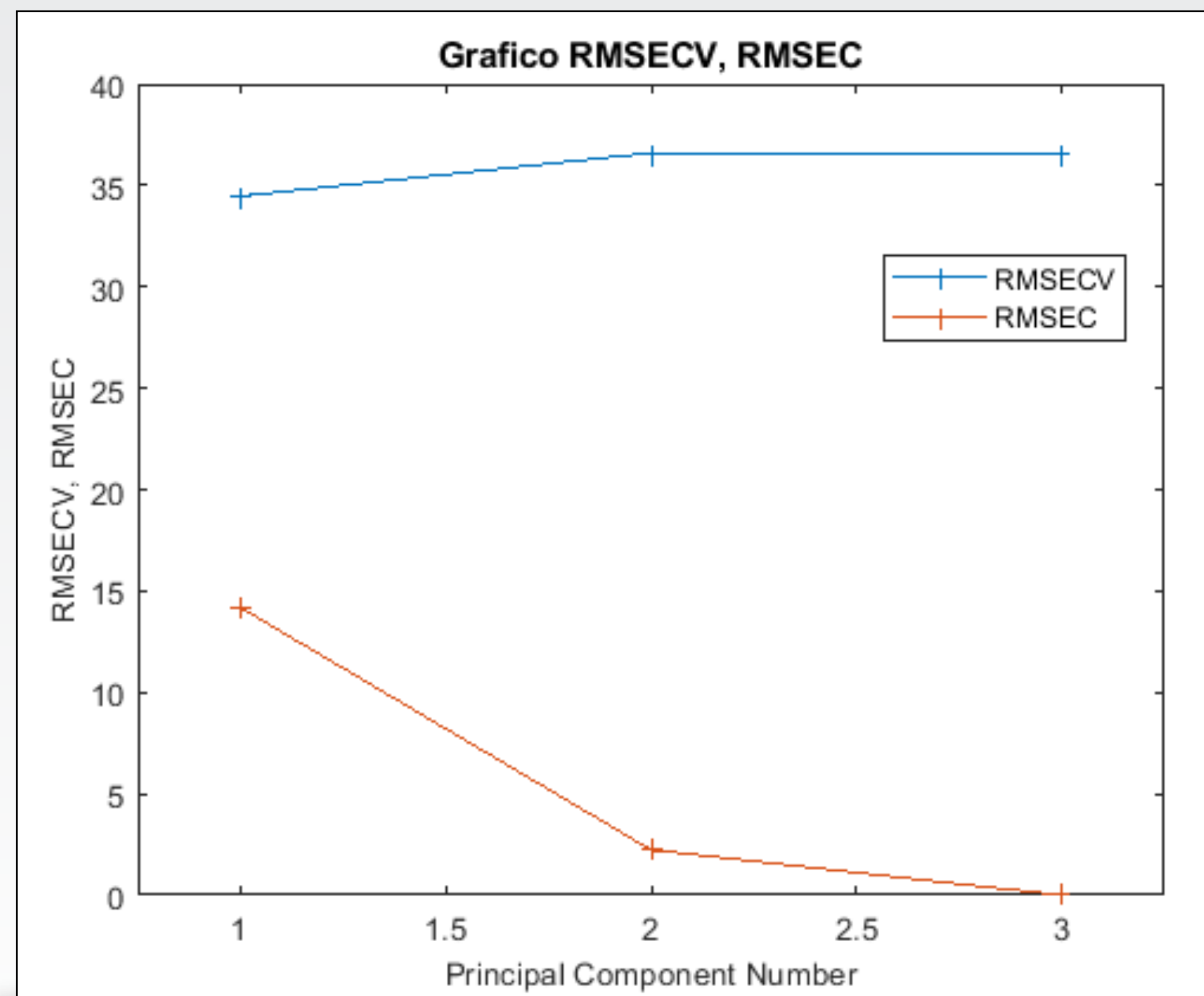


## 4. PCA esplorativa

Numero PC ottimale trovato  
è 1

Scelta numero delle componenti  
modello PCA

PC1 cattura una **varianza** del  
**94.58%**



## 4. PCA esplorativa

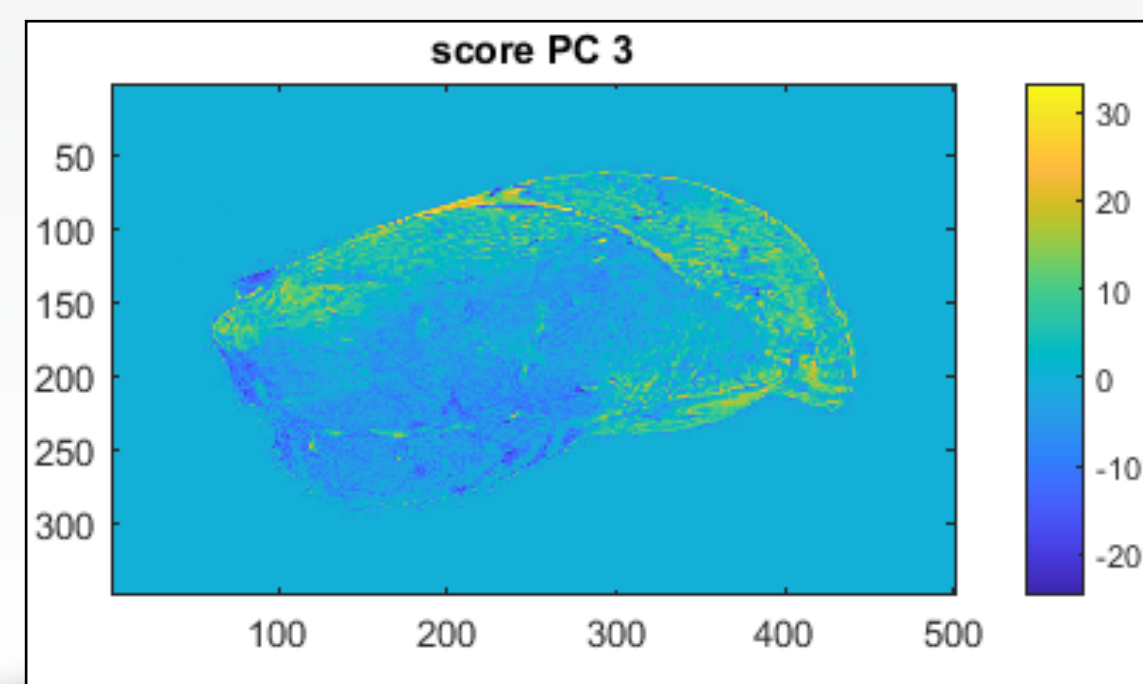
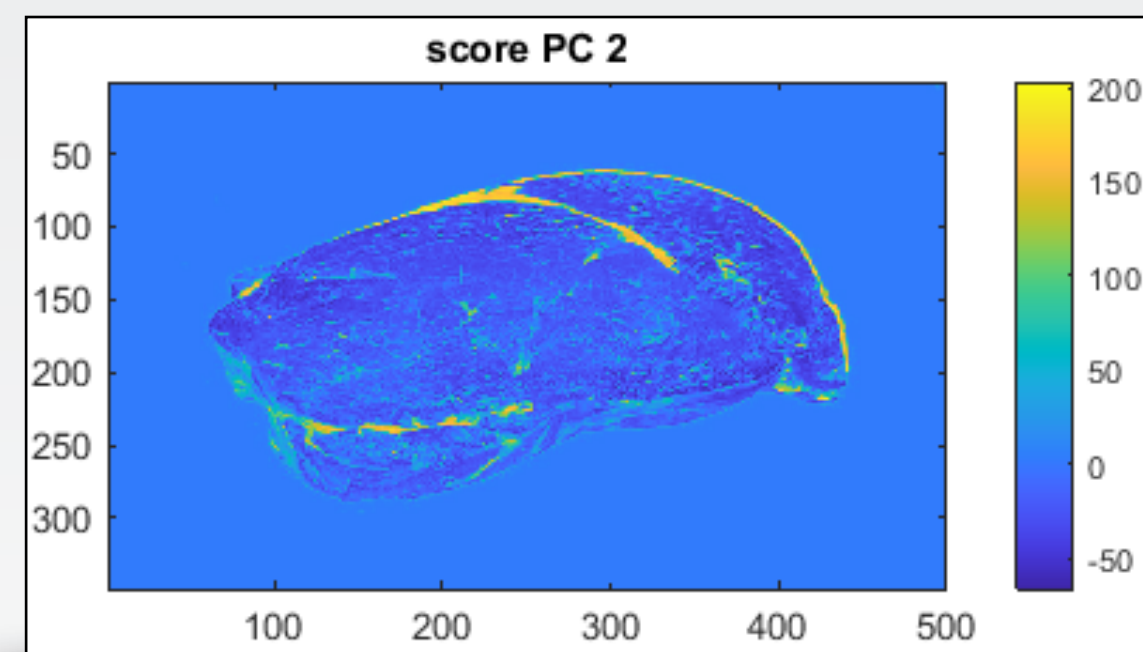
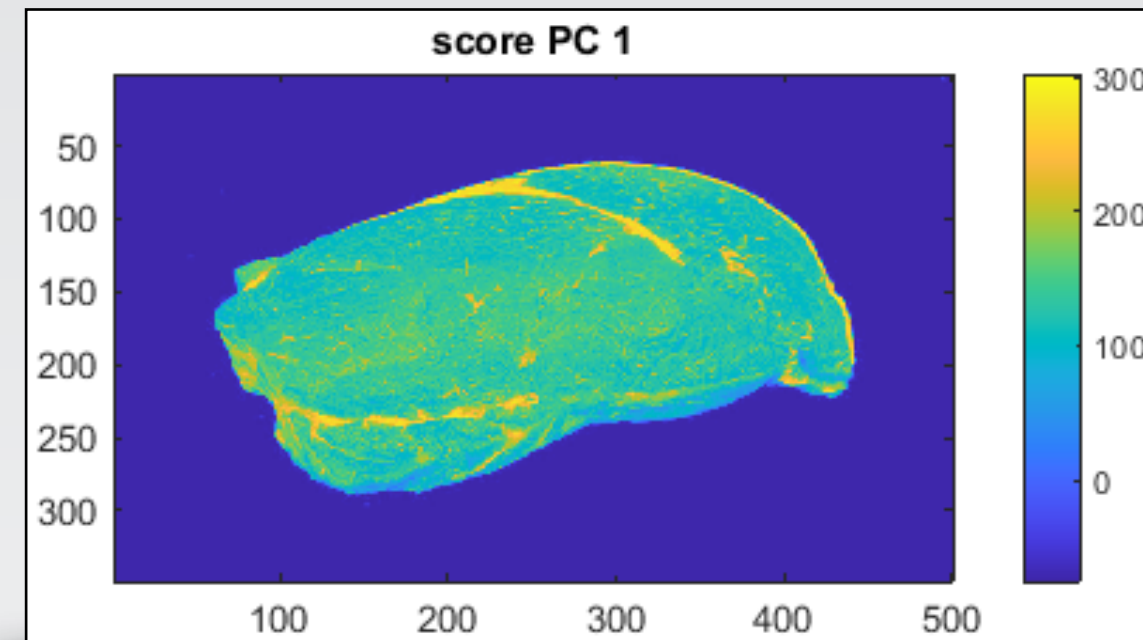
**Scelta numero delle componenti  
modello PCA**

Immagini degli **score** per ogni PC.

## 4. PCA esplorativa

Scelta numero delle componenti  
modello PCA

Immagini degli **score** per ogni PC.

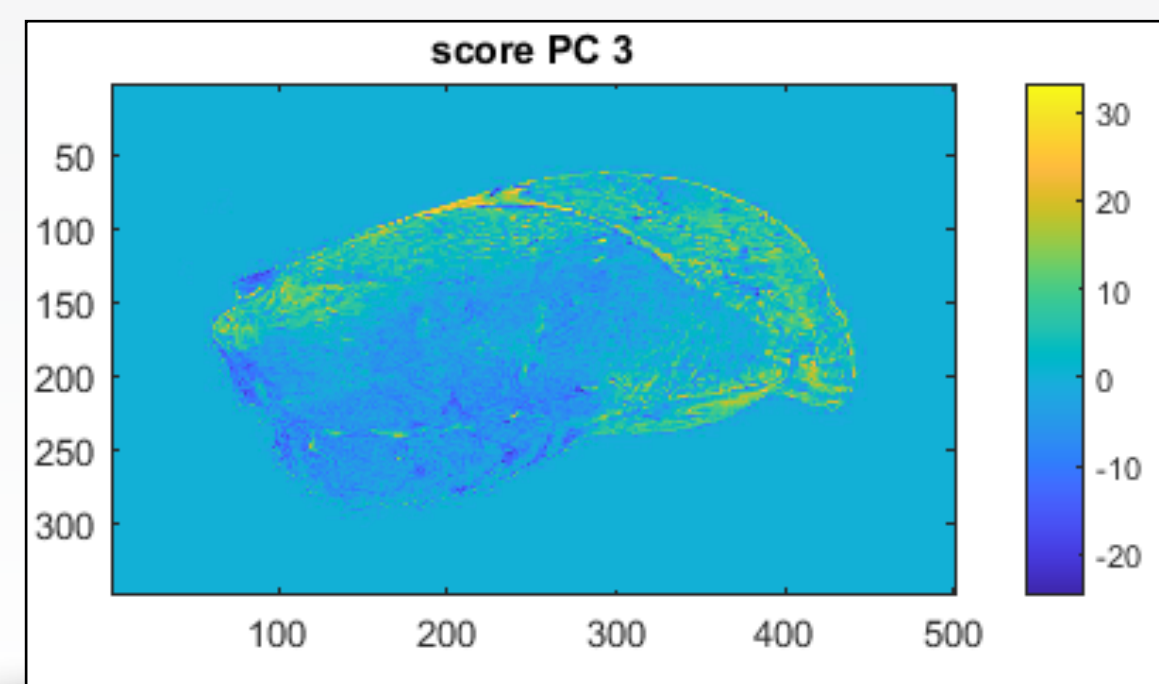
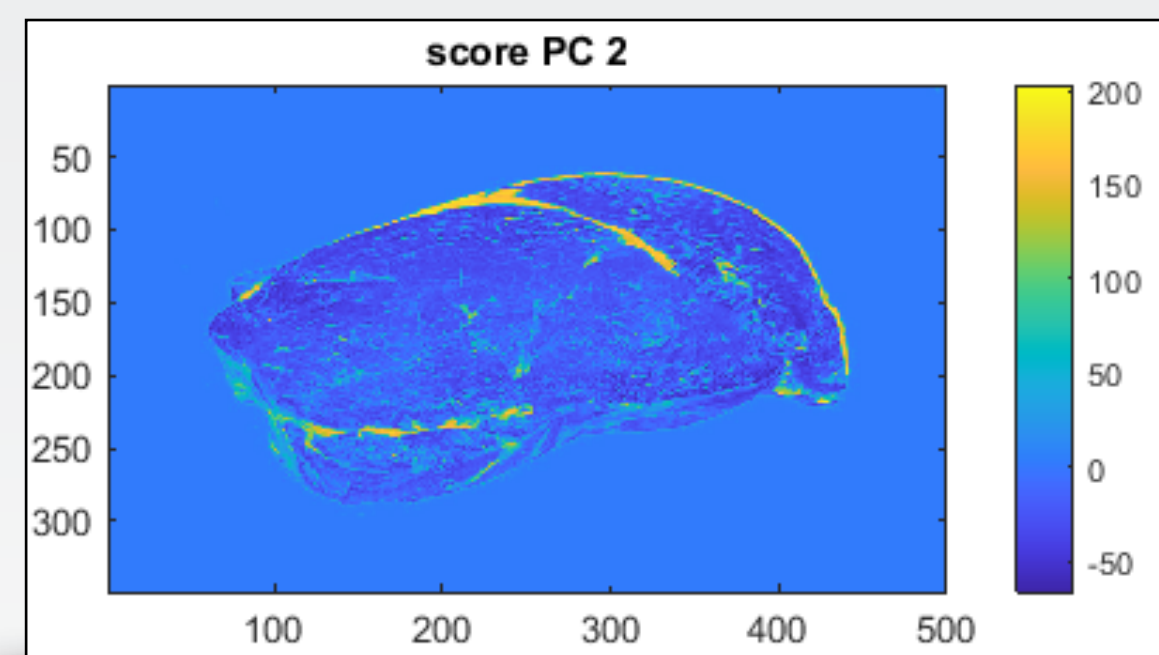
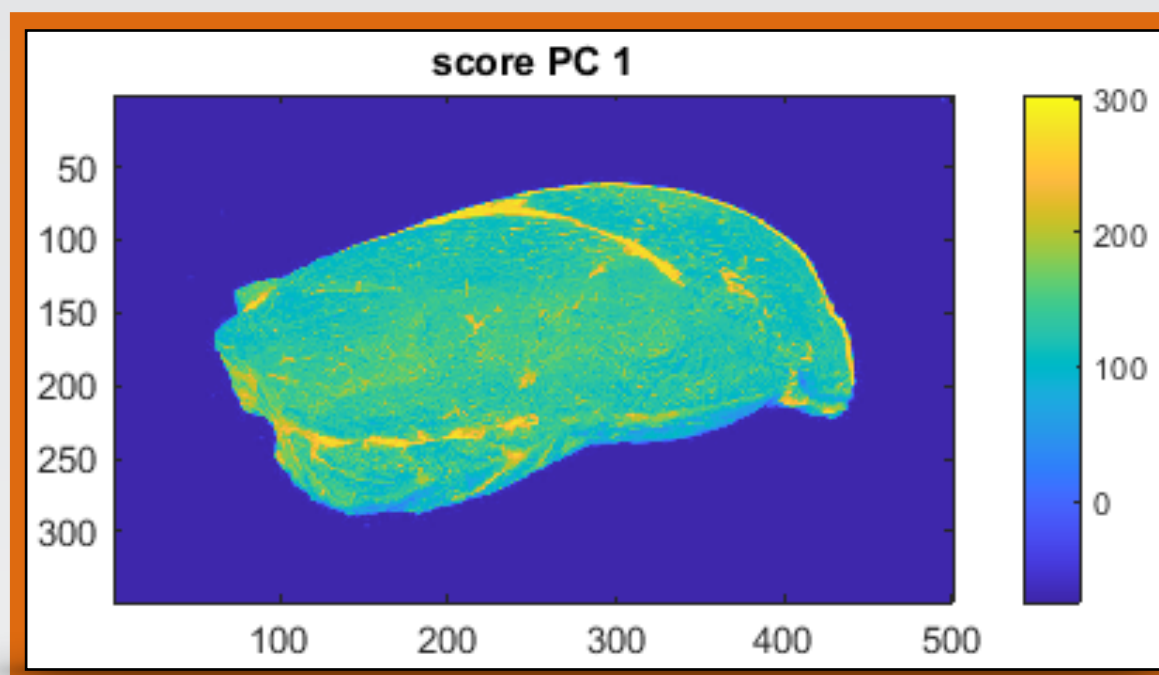


## 4. PCA esplorativa

### Scelta numero delle componenti modello PCA

Immagini degli **score** per ogni PC.

Con **PC1** è possibile distinguere le  
categorie “sfondo”, “carne” e  
“grasso”





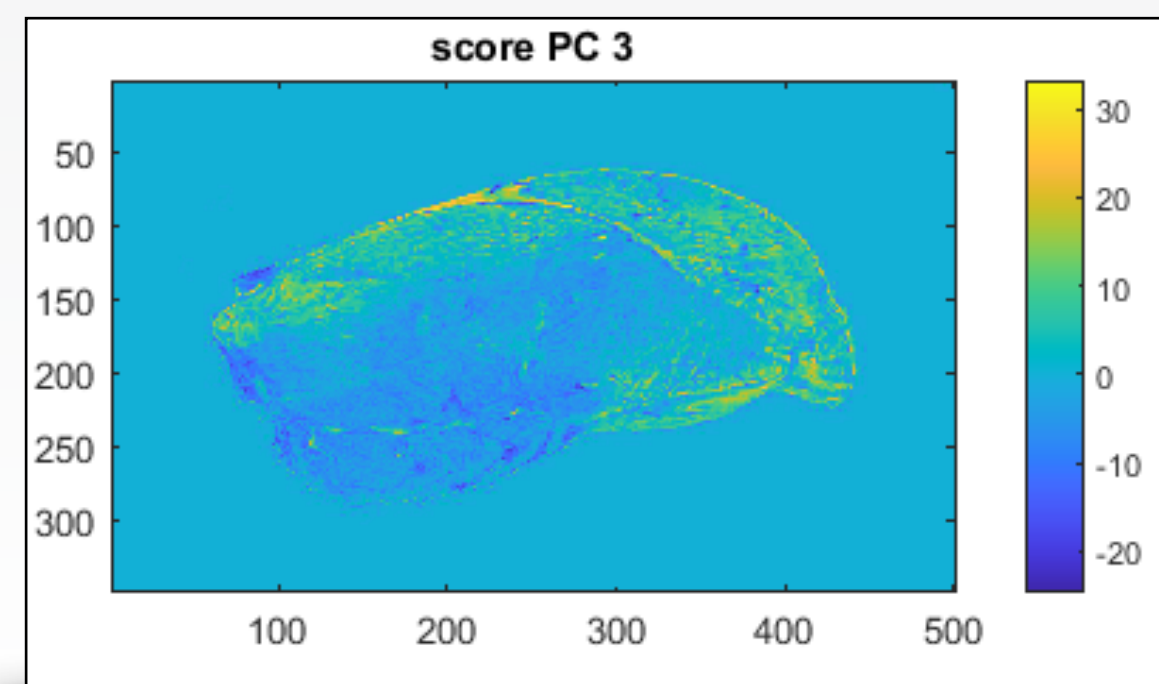
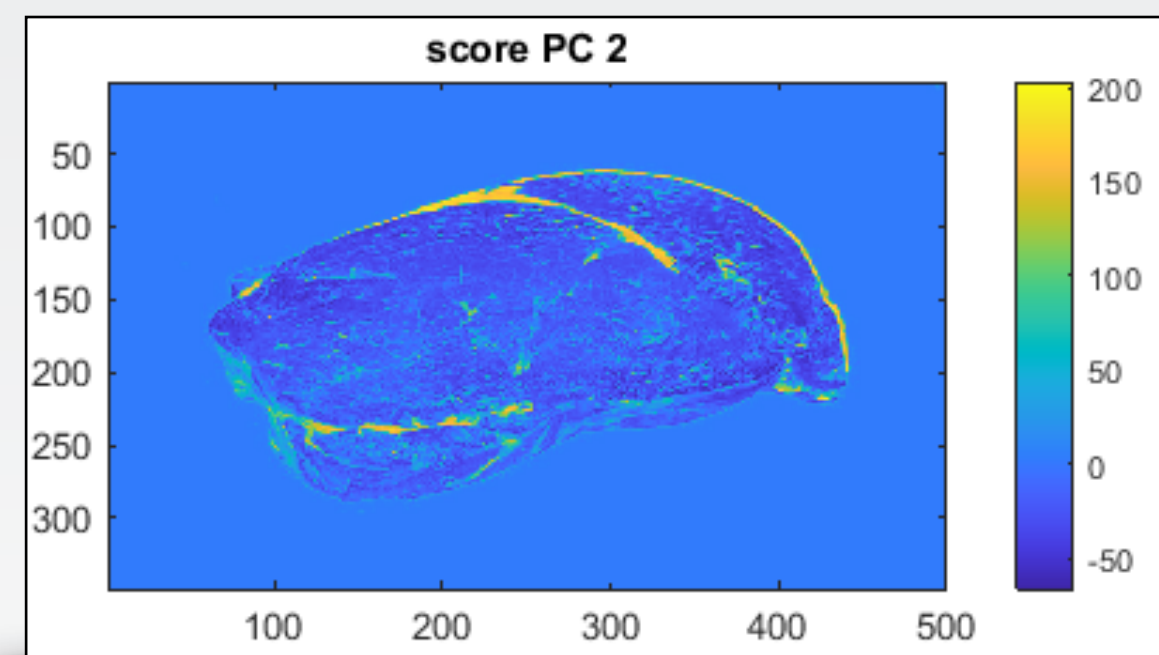
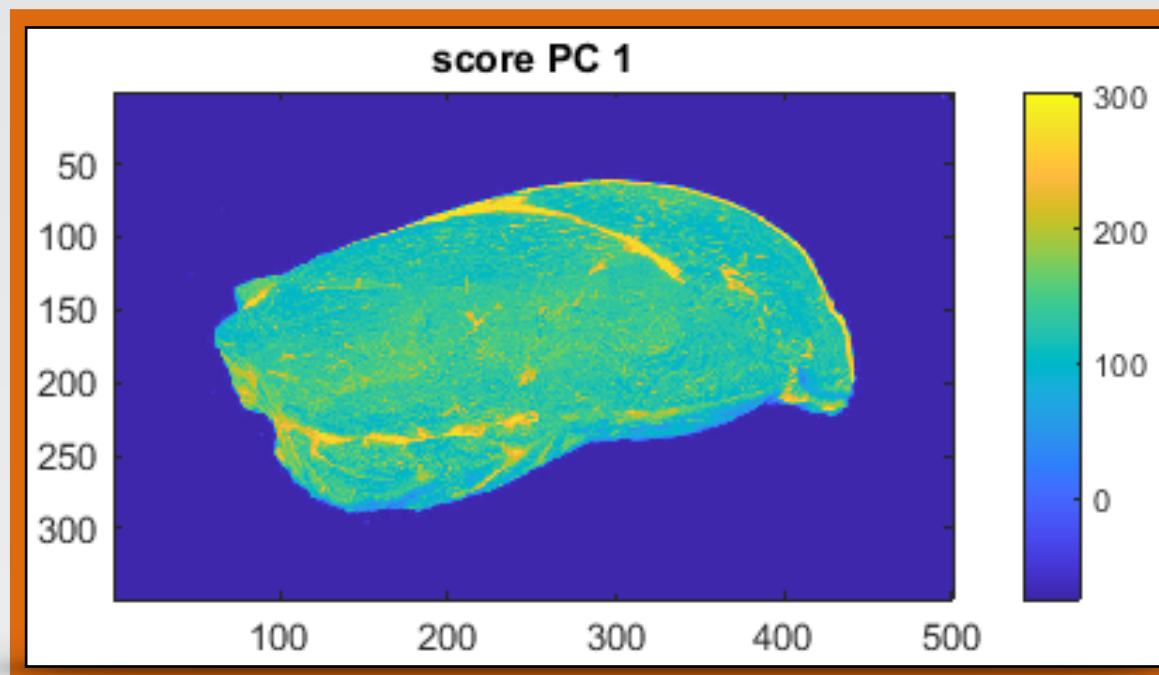
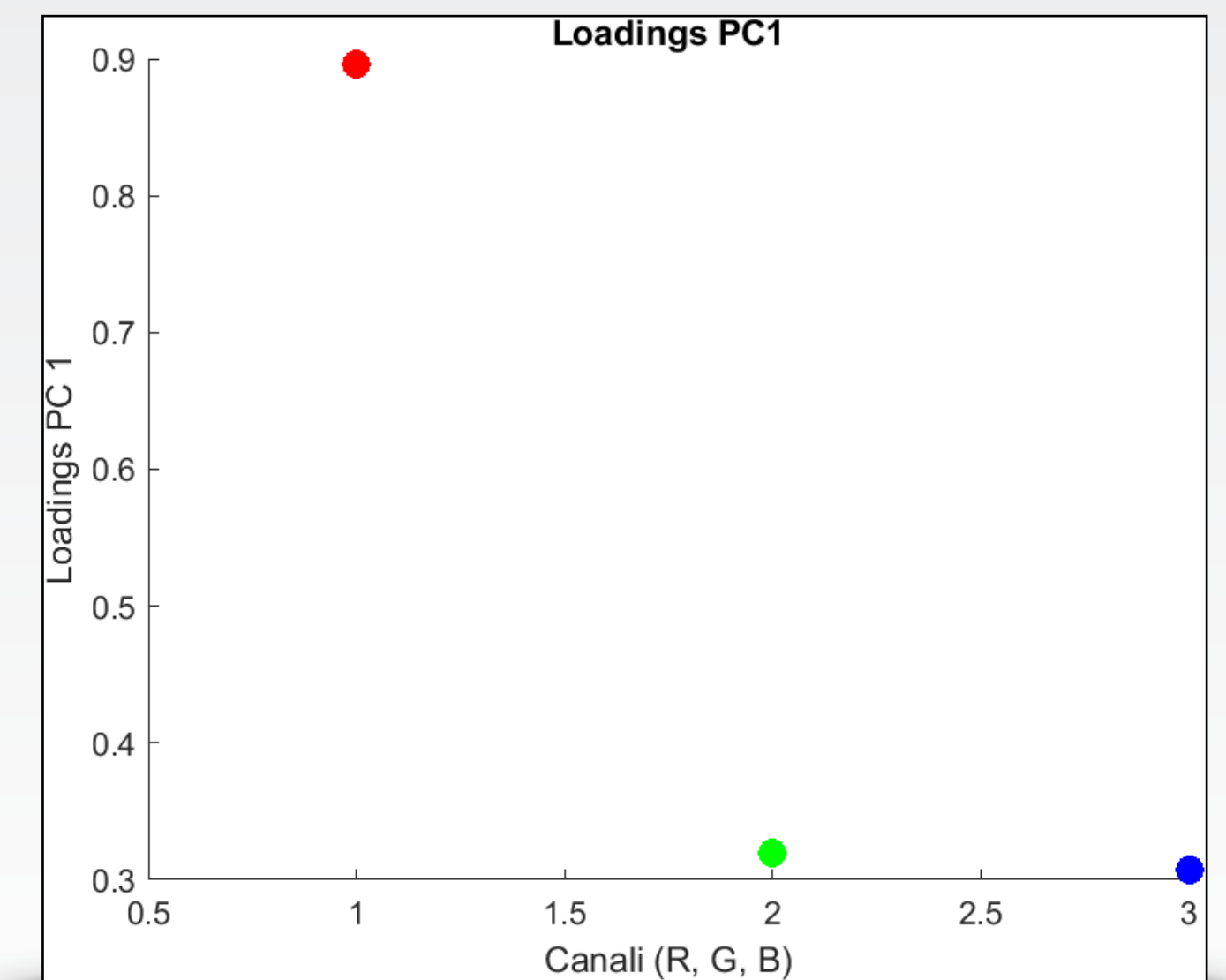
## 4. PCA esplorativa

### Scelta numero delle componenti modello PCA

Immagini degli **score** per ogni PC.

Con **PC1** è possibile distinguere le  
categorie “sfondo”, “carne” e  
“grasso”

### Loadings PC1



## 4. PCA esplorativa

## 4. PCA esplorativa

**Residui modello PCA**

## 4. PCA esplorativa

### Residui modello PCA

Immagine score PC1

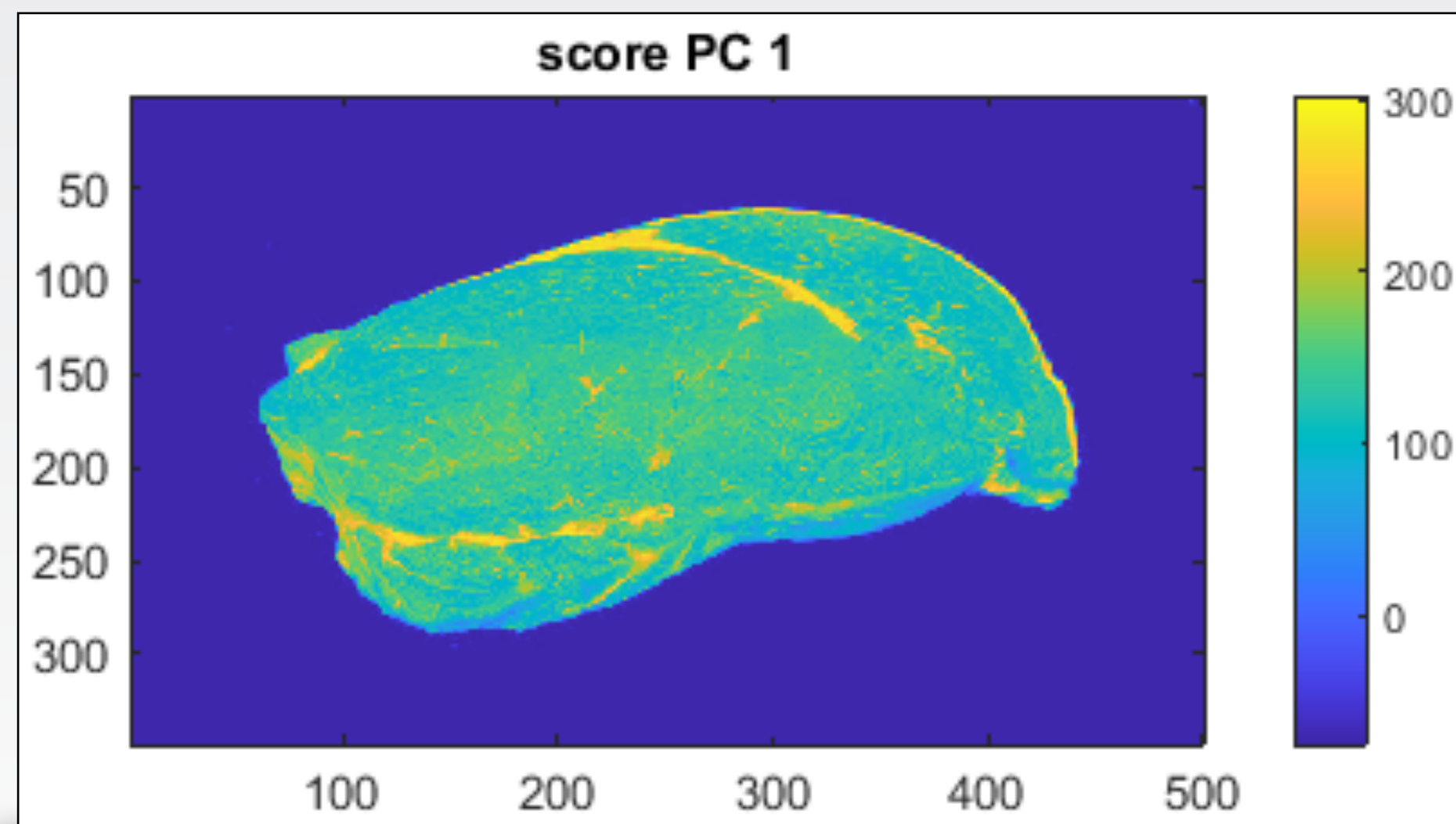
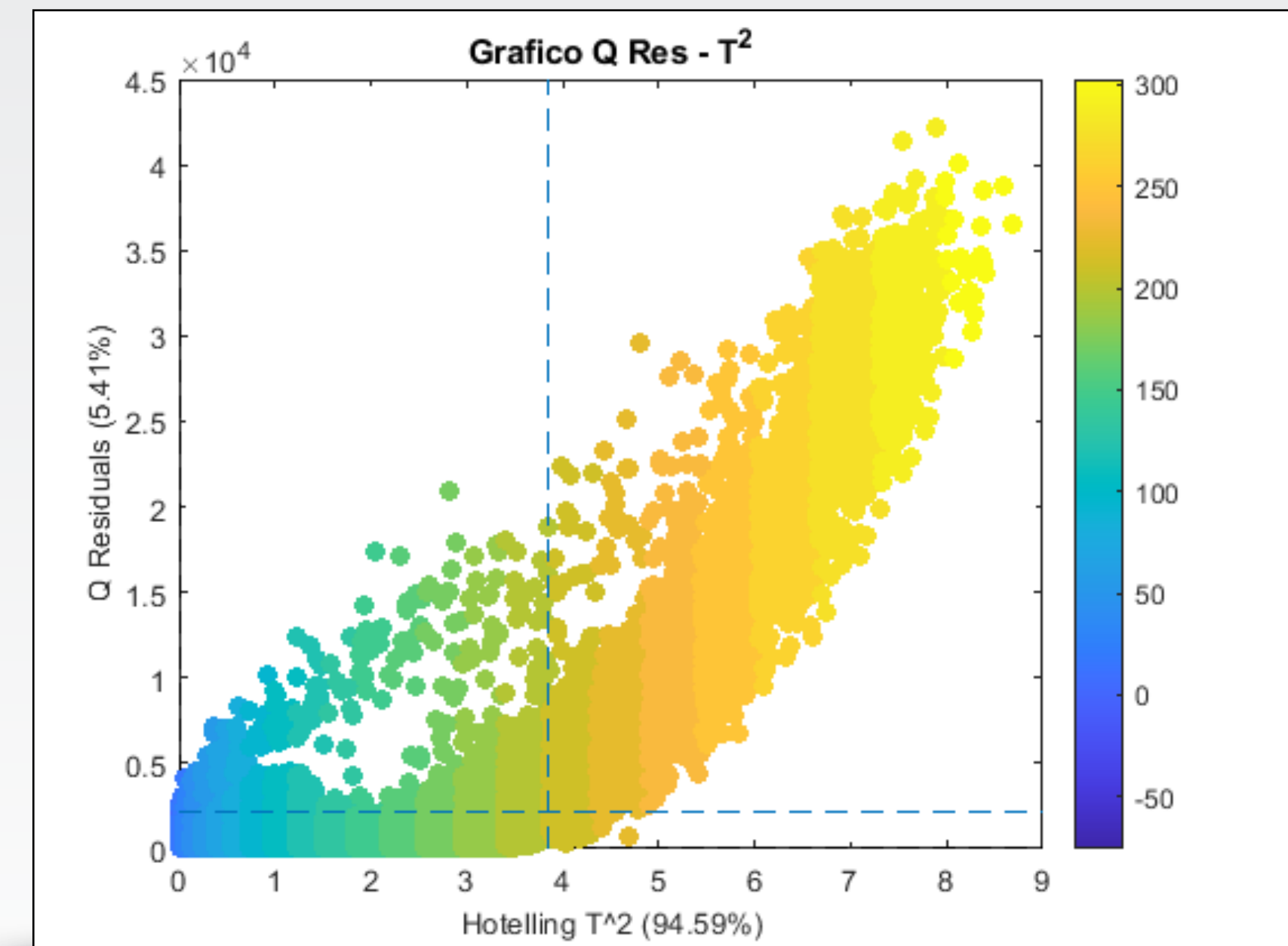


Grafico  $Q-T^2$   
Ogni elemento è colorato  
in base al valore di **score**



## 4. PCA esplorativa

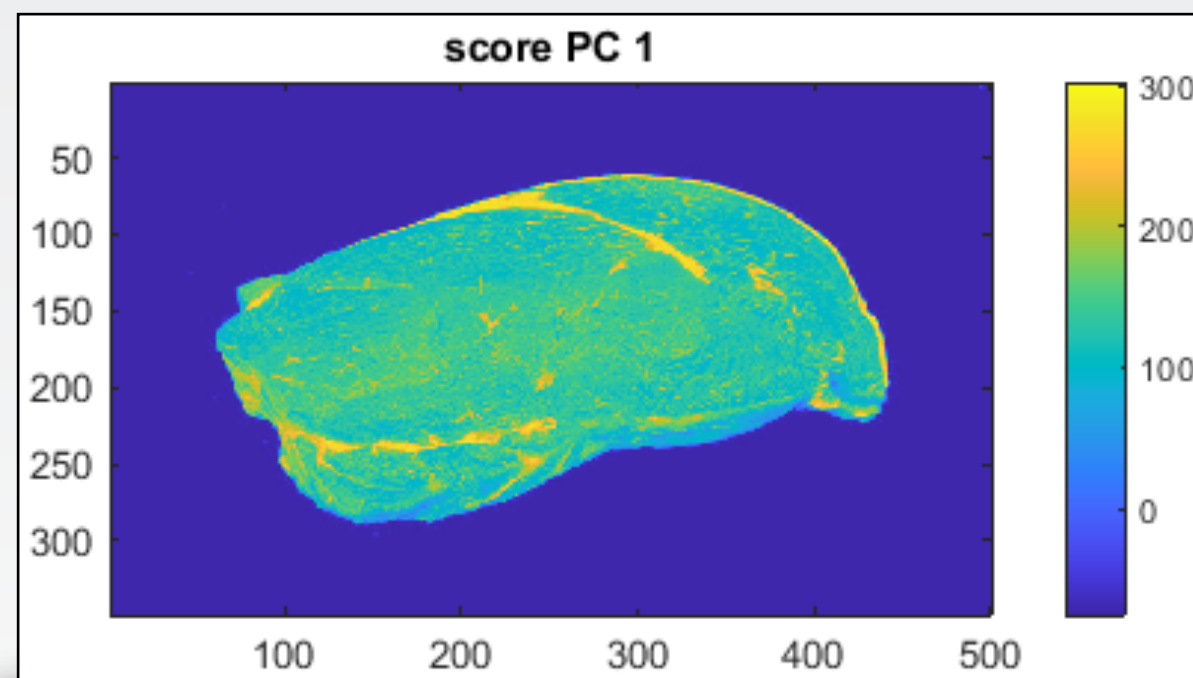
## 4. PCA esplorativa

**Selezione soglie score di PC1 per distinguere le categorie**

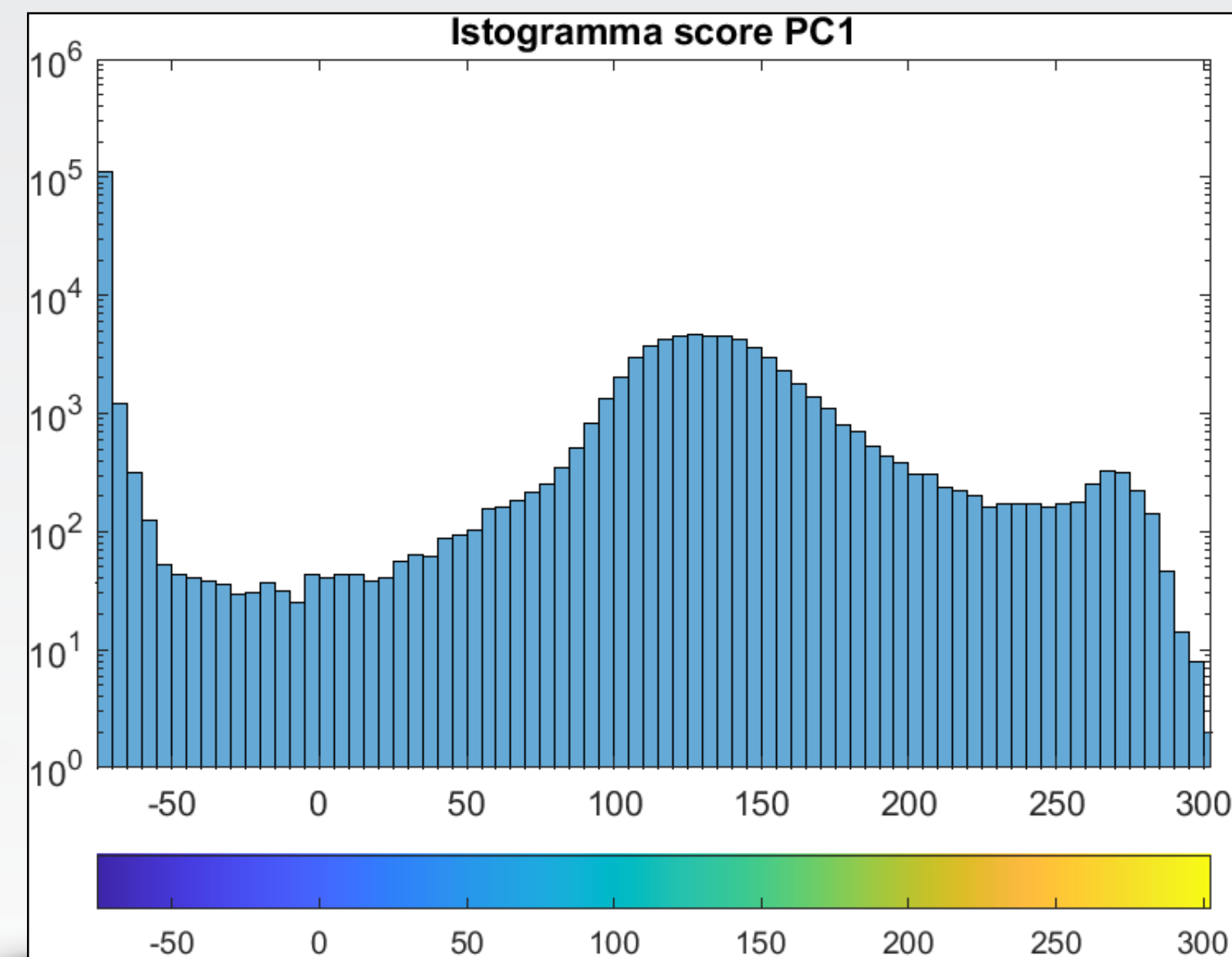
## 4. PCA esplorativa

Selezione soglie score di PC1 per distinguere le categorie

Immagine score PC1



Istogramma score PC1

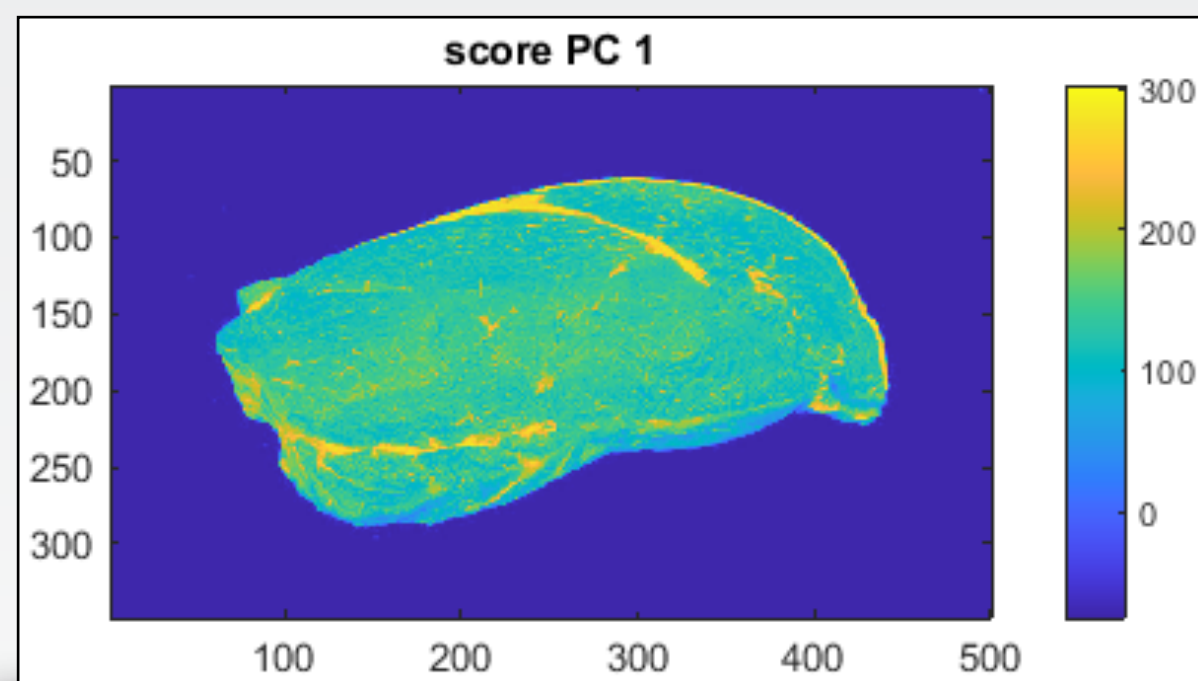




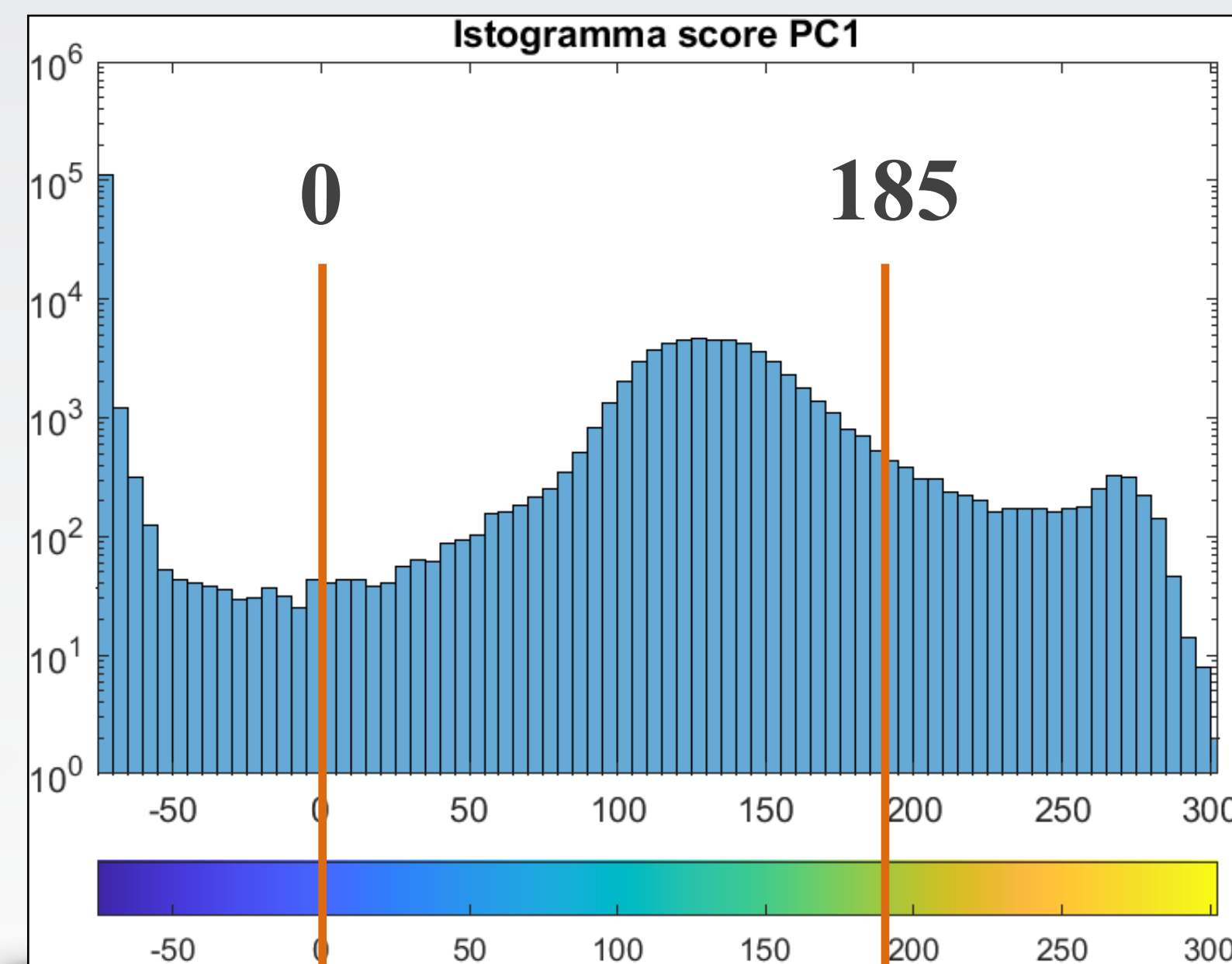
## 4. PCA esplorativa

Selezione soglie score di PC1 per distinguere le categorie

Immagine score PC1



Istogramma score PC1



← ← → →

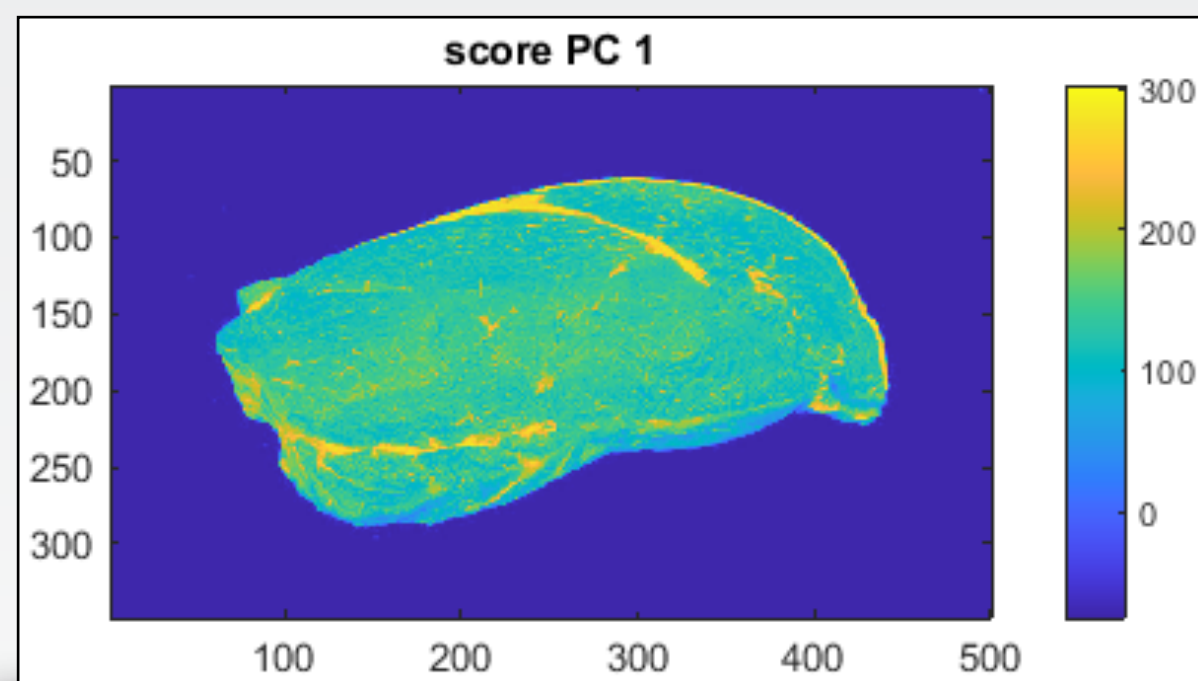
**Sfondo** **Carne** **Grasso**



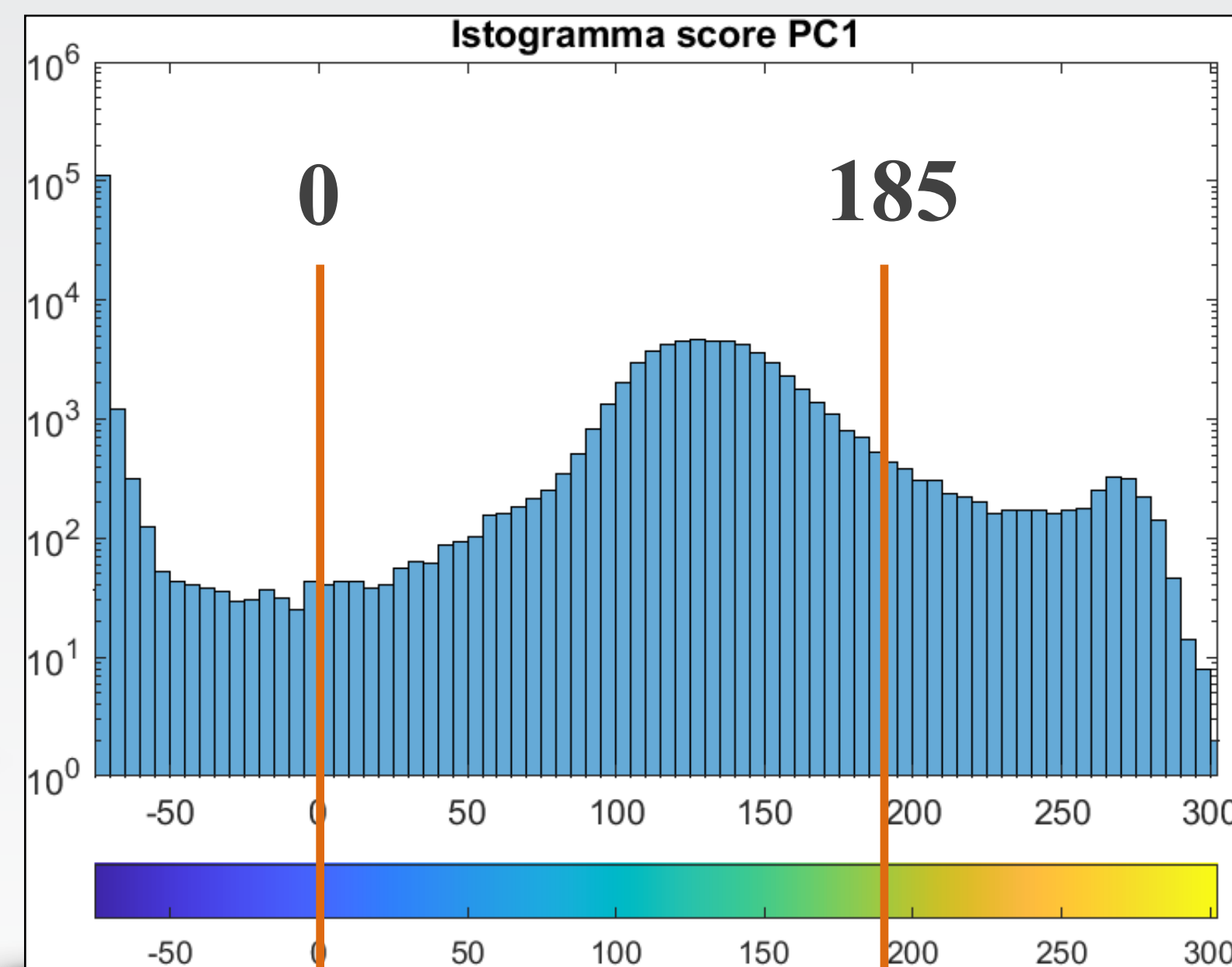
## 4. PCA esplorativa

Selezione soglie score di PC1 per distinguere le categorie

Immagine score PC1



Istogramma score PC1



**Sfondo** **Carne** **Grasso**

Risultato

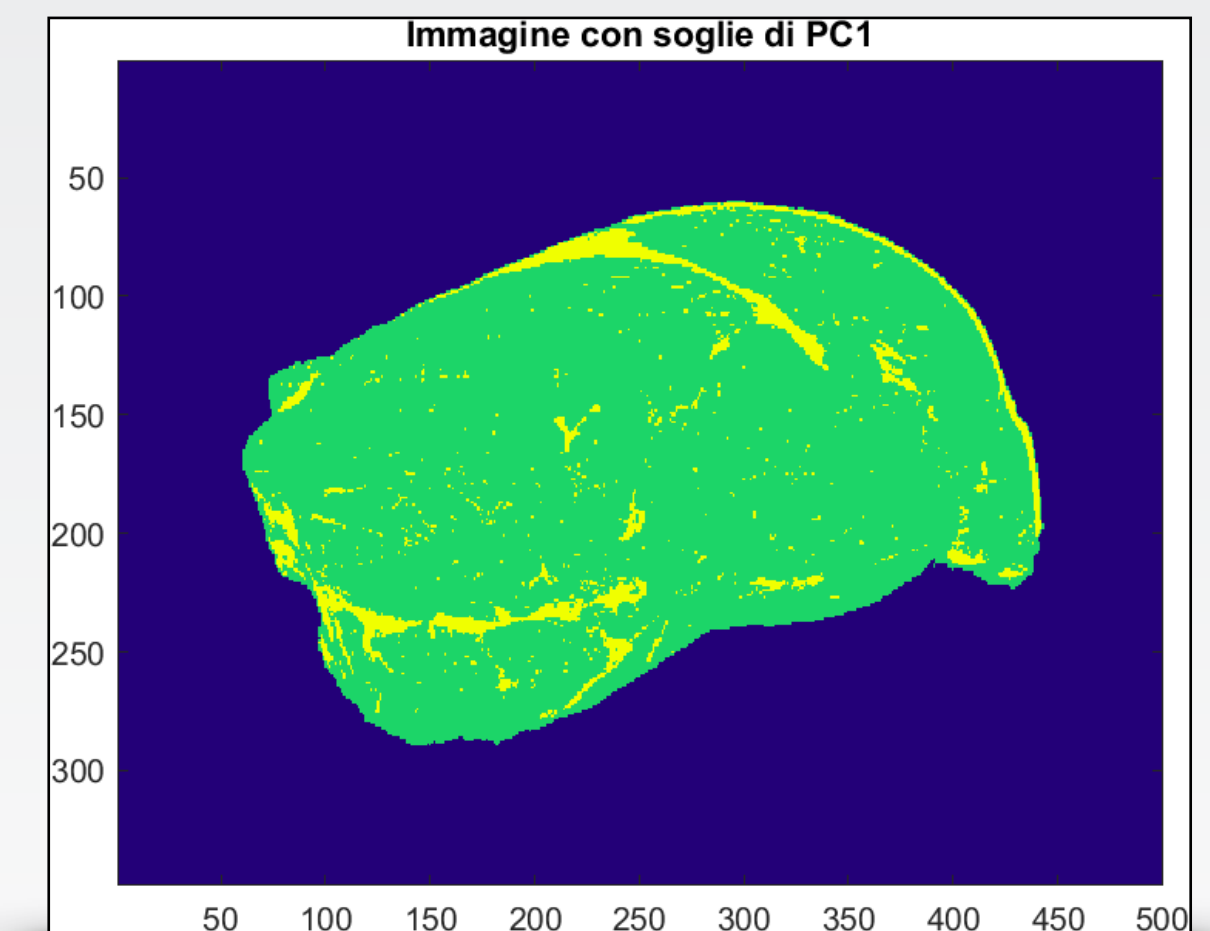


Immagine creata tenendo in considerazione i valori di soglia

**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

**5. Clustering**

**6. Previsione grasso**

# 5. Clustering

## 5. Clustering

Per questa analisi sono stati presi in considerazione 2 metodi di clustering:

- **KMeans**
- **DBScan**

## 5. Clustering

Per questa analisi sono stati presi in considerazione 2 metodi di clustering:

- **KMeans**
- **DBScan**

**OBIETTIVO:** individuare i cluster di “sfondo”, “carne”, “grasso”

## 5. Clustering - KMeans

## 5. Clustering - KMeans

### Parametri KMeans:

- $K = 3$
- Distance = EuclideanSq / Manhattan



## 5. Clustering - KMeans

### Parametri KMeans:

- $K = 3$
- Distance = EuclideanSq / Manhattan

**Obiettivo:** confronto cluster KMeans con diversi metodi per il calcolo della distanza (EuclideanSq e Manhattan)

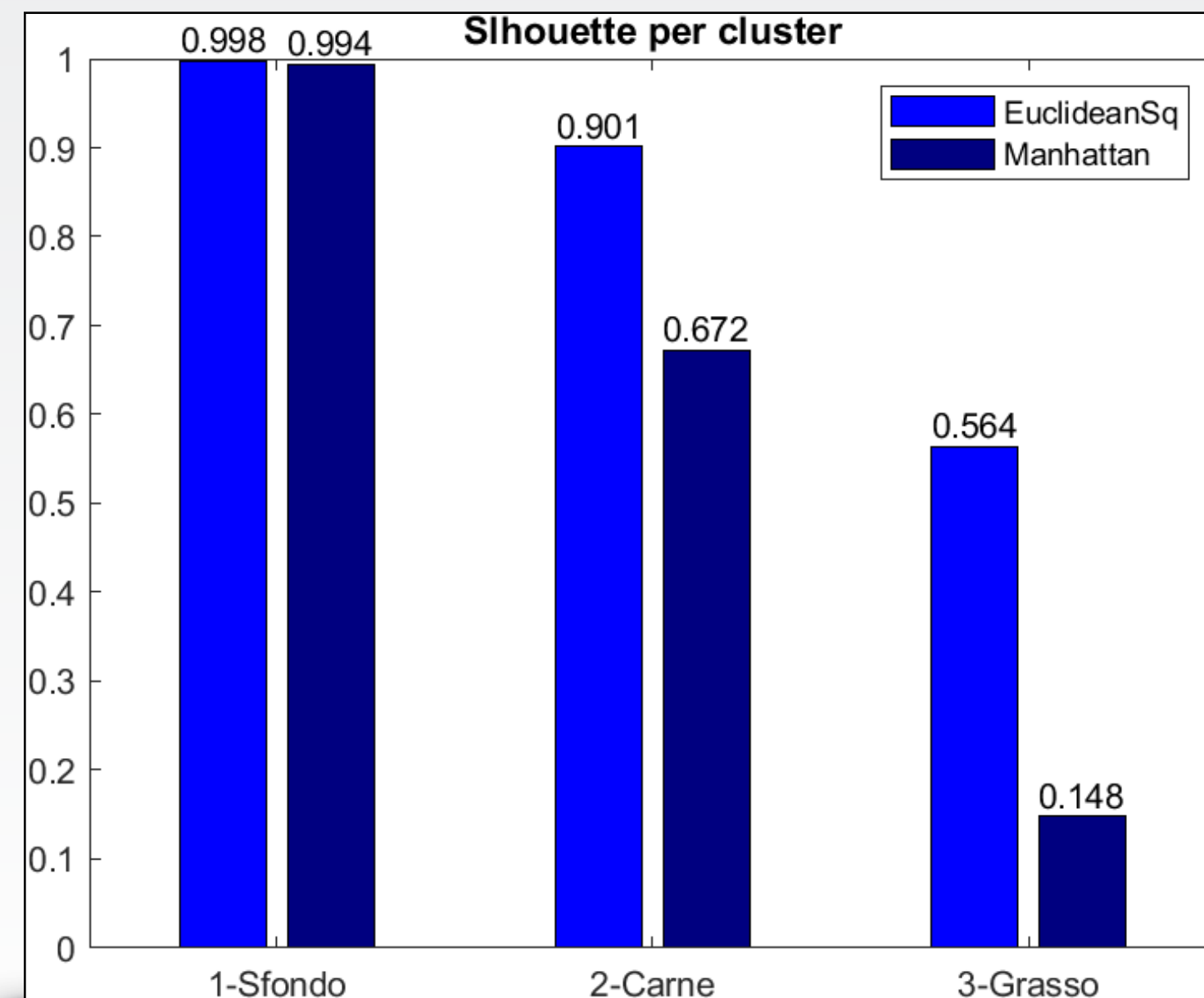
---

## 5. Clustering - KMeans

### Parametri KMeans:

- $K = 3$
- Distance = EuclideanSq / Manhattan

**Obiettivo:** confronto cluster KMeans con diversi metodi per il calcolo della distanza (EuclideanSq e Manhattan)

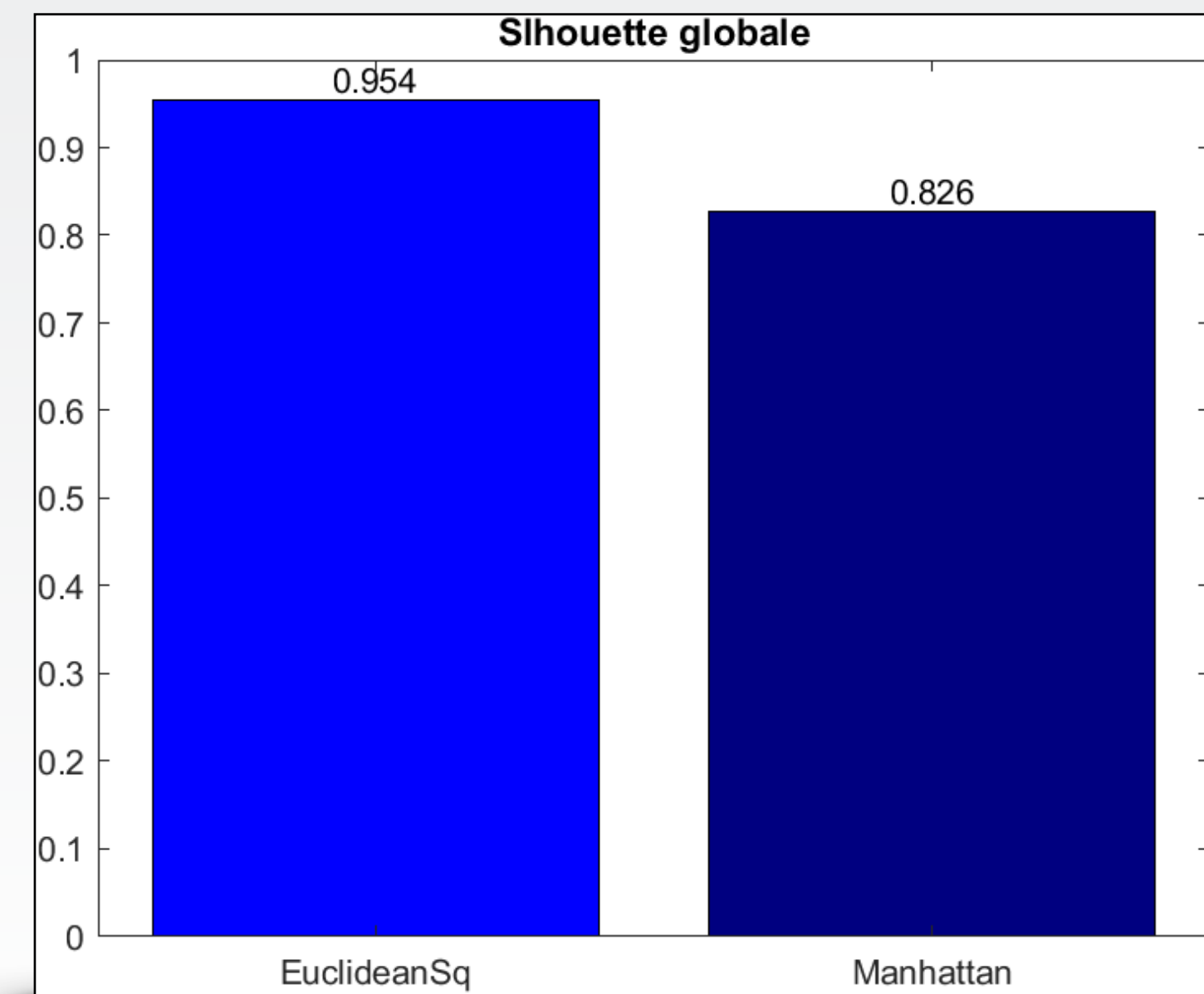
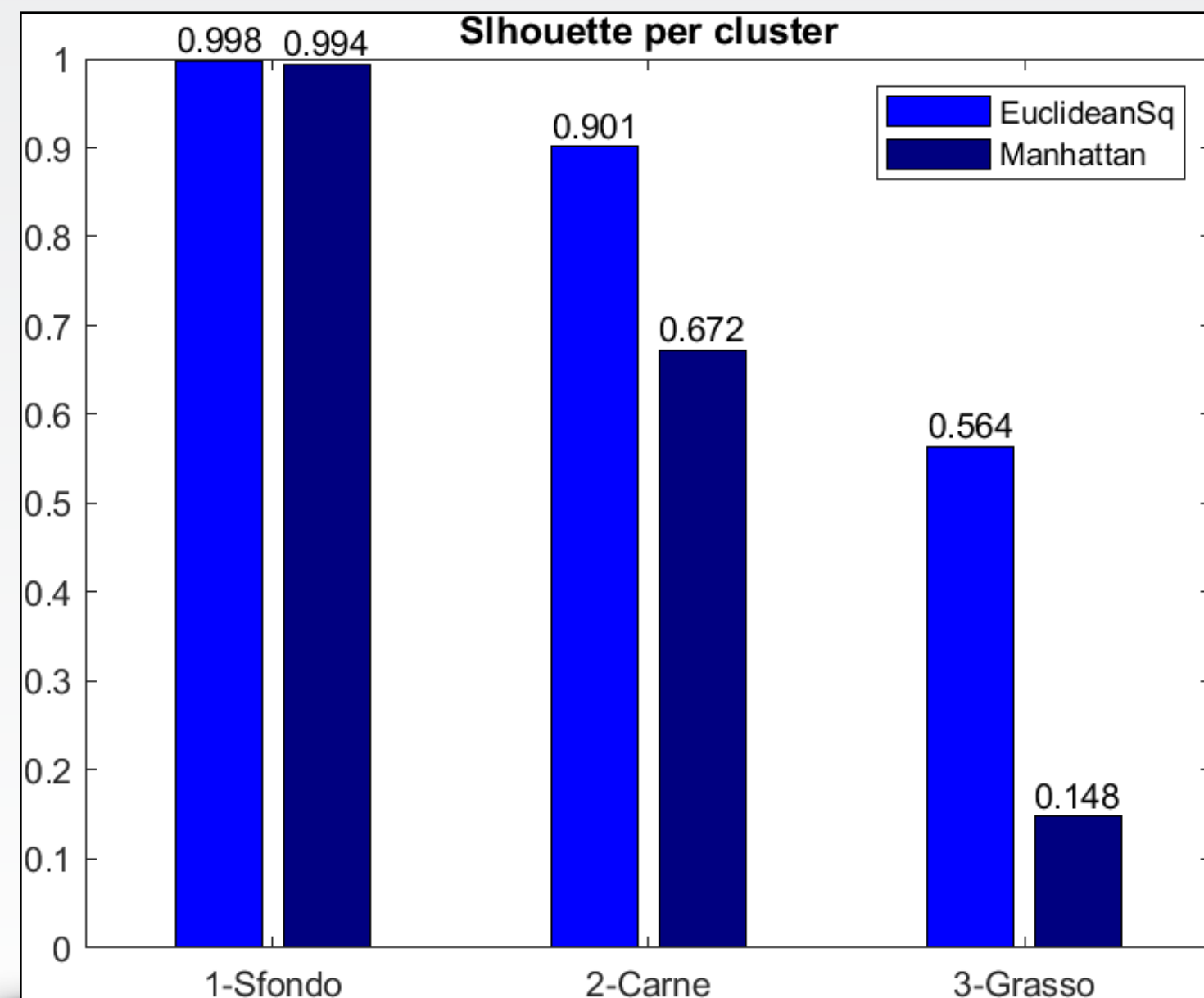


## 5. Clustering - KMeans

### Parametri KMeans:

- $K = 3$
- Distance = EuclideanSq / Manhattan

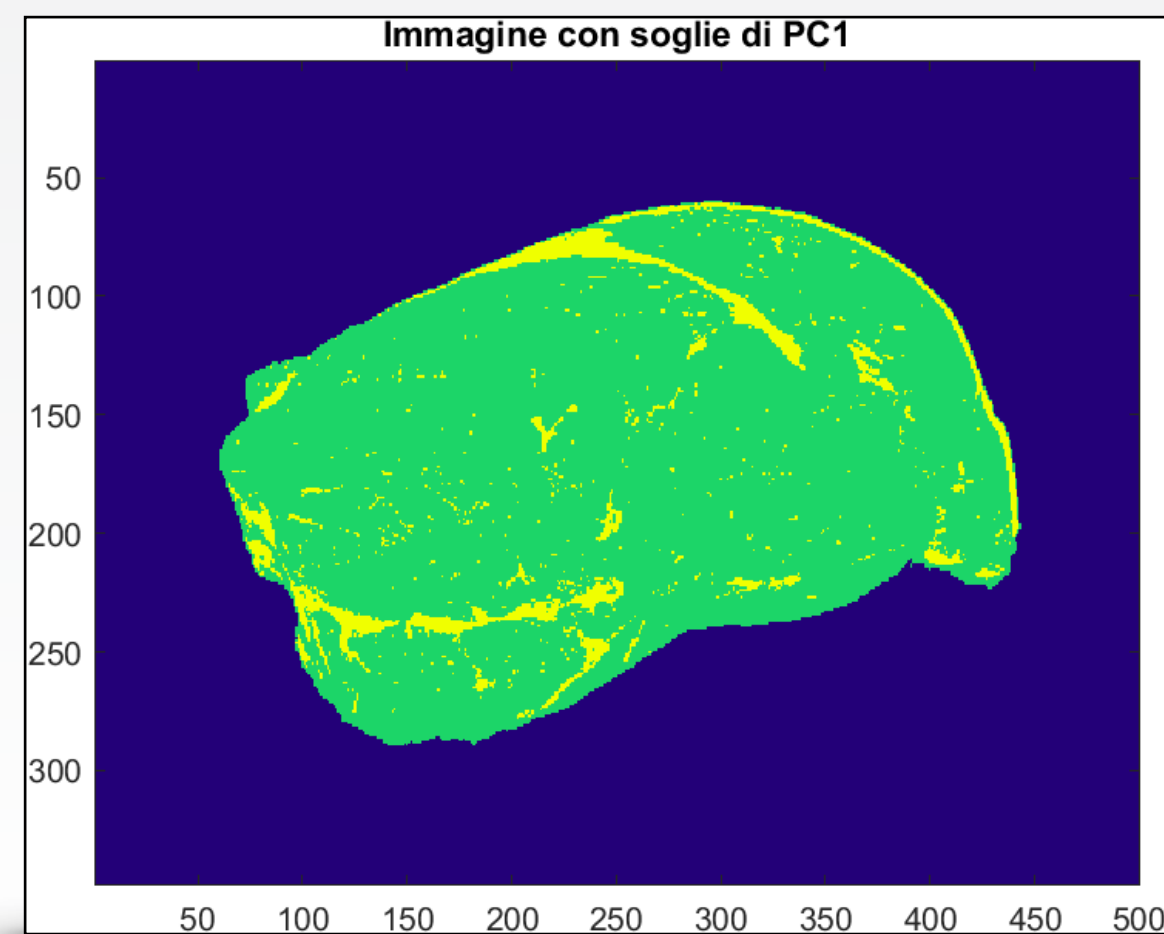
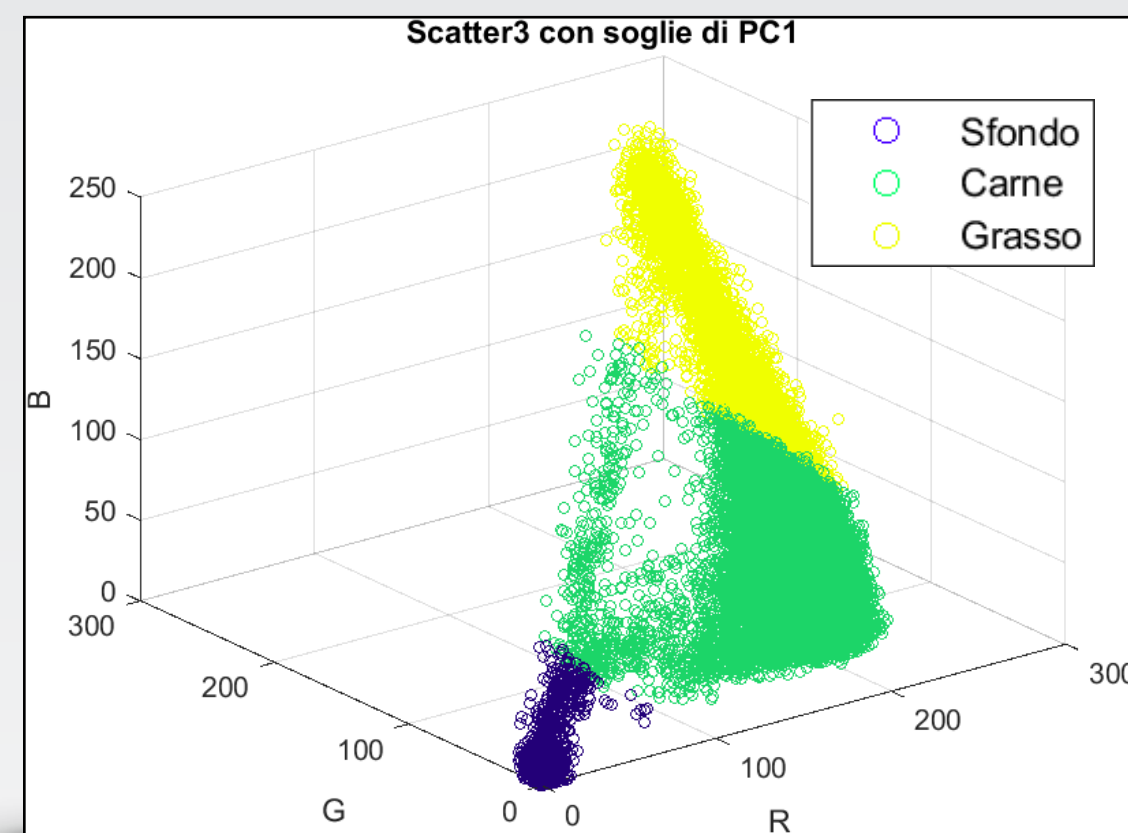
**Obiettivo:** confronto cluster KMeans con diversi metodi per il calcolo della distanza (EuclideanSq e Manhattan)



## 5. Clustering - KMeans

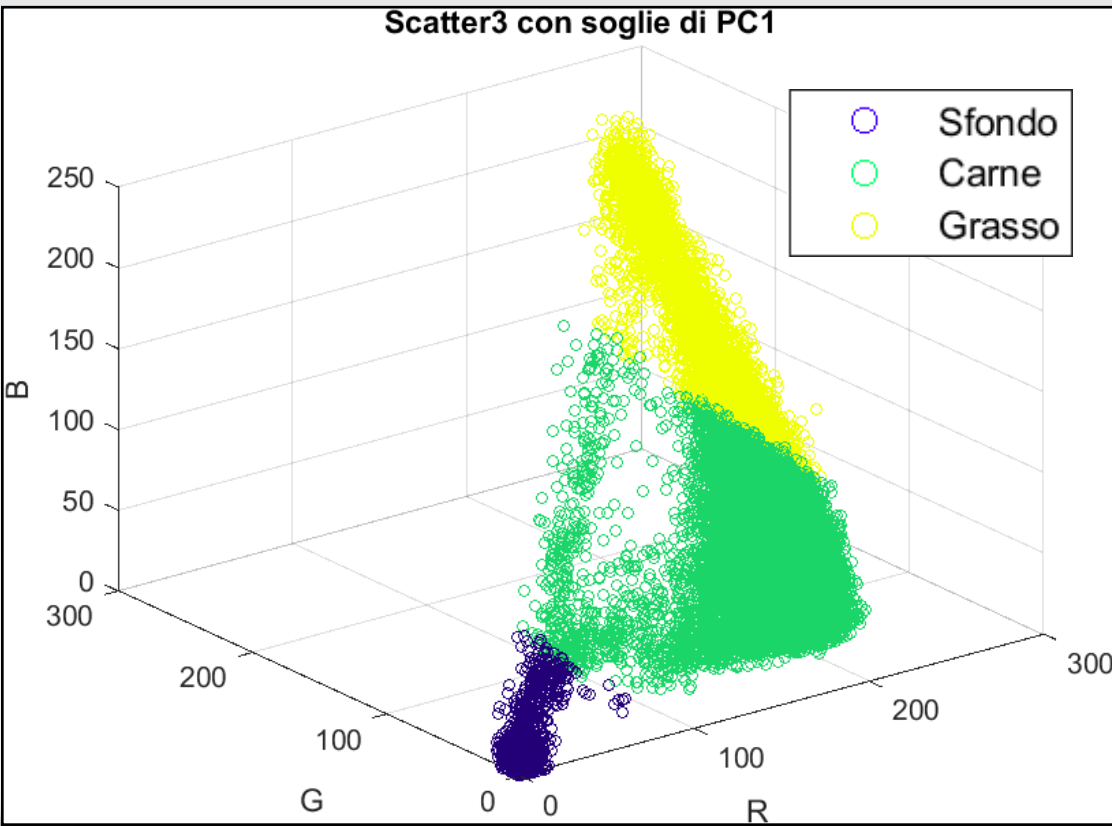
# 5. Clustering - KMeans

PC1 con soglie

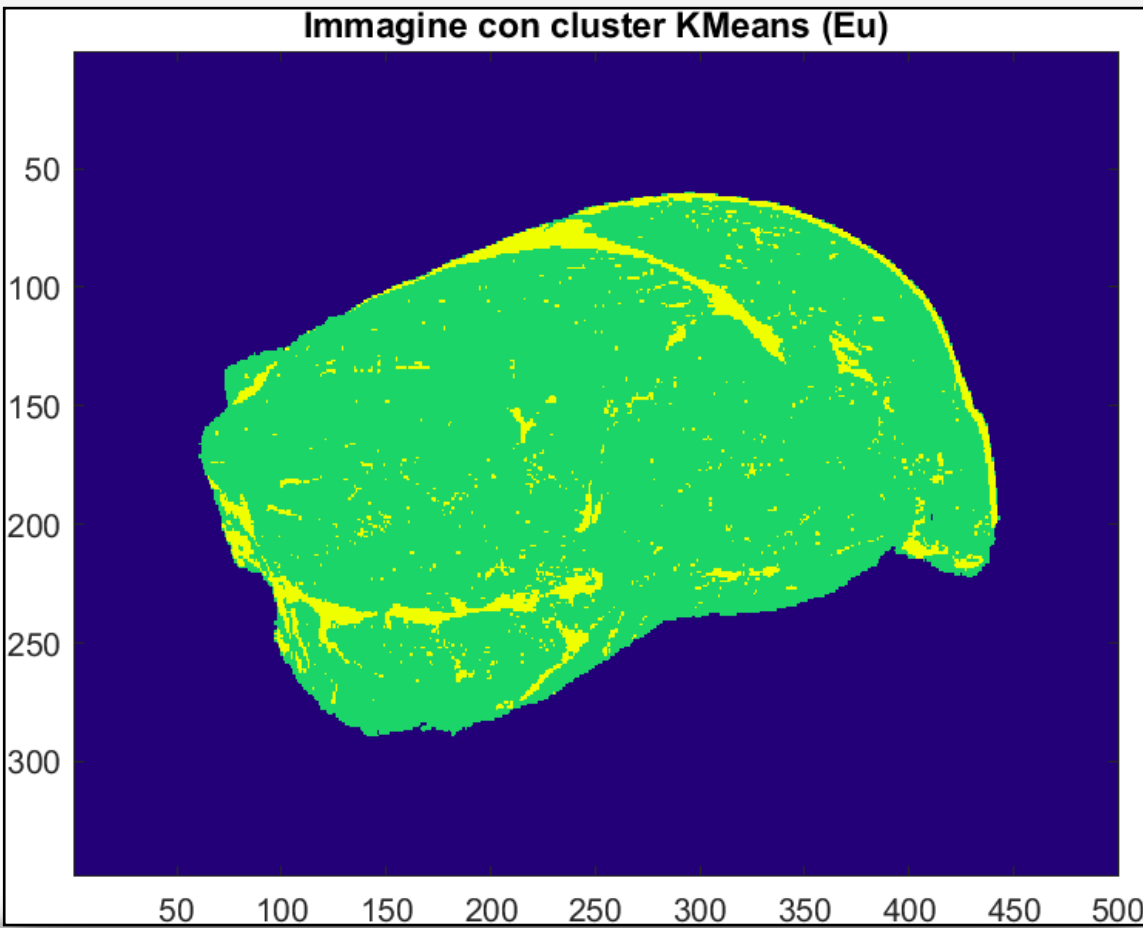
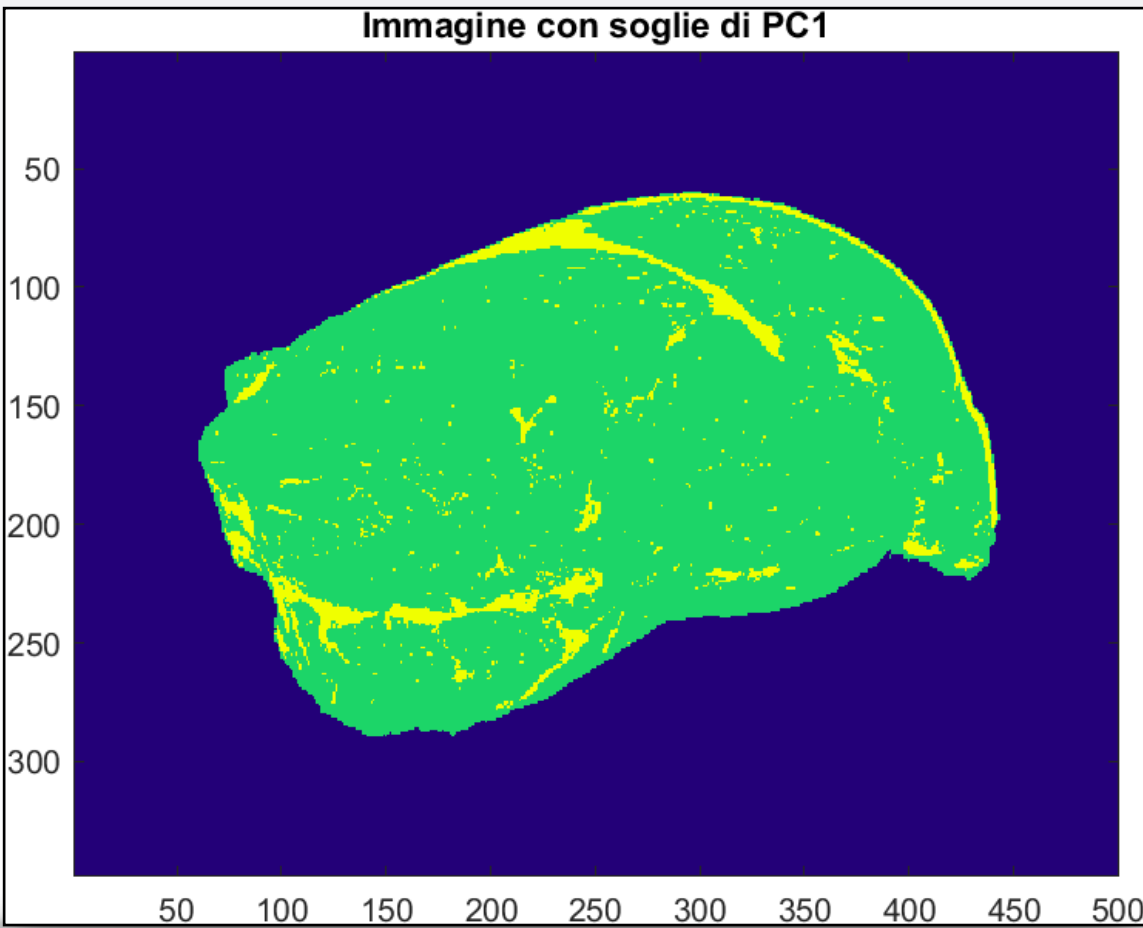
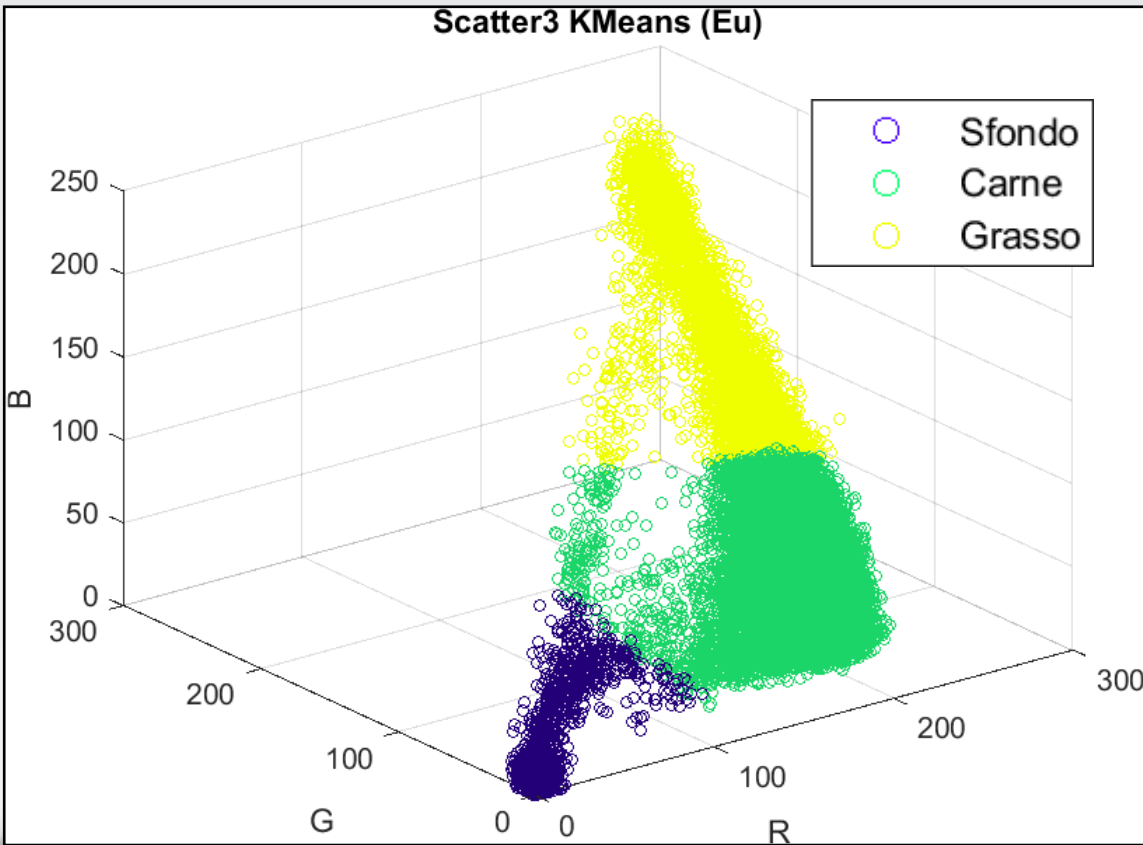


# 5. Clustering - KMeans

PC1 con soglie

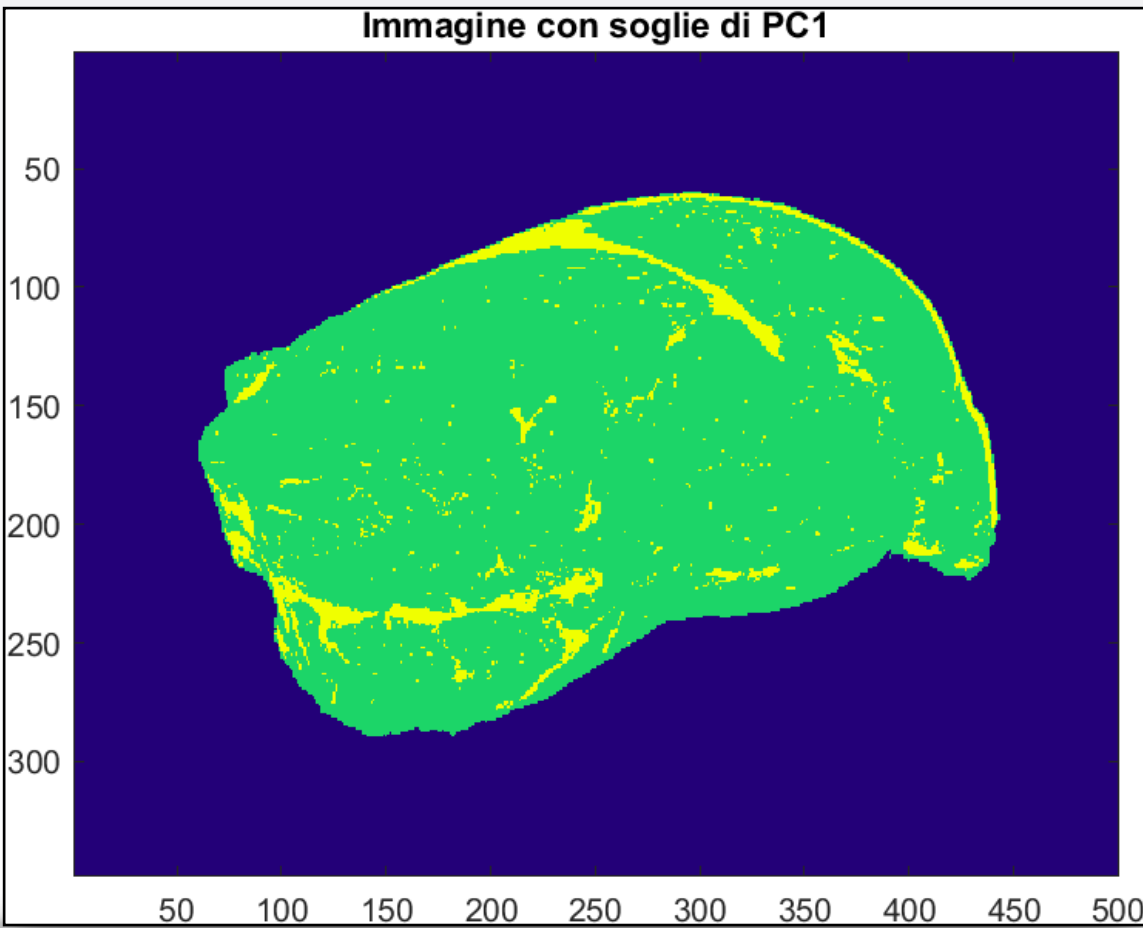
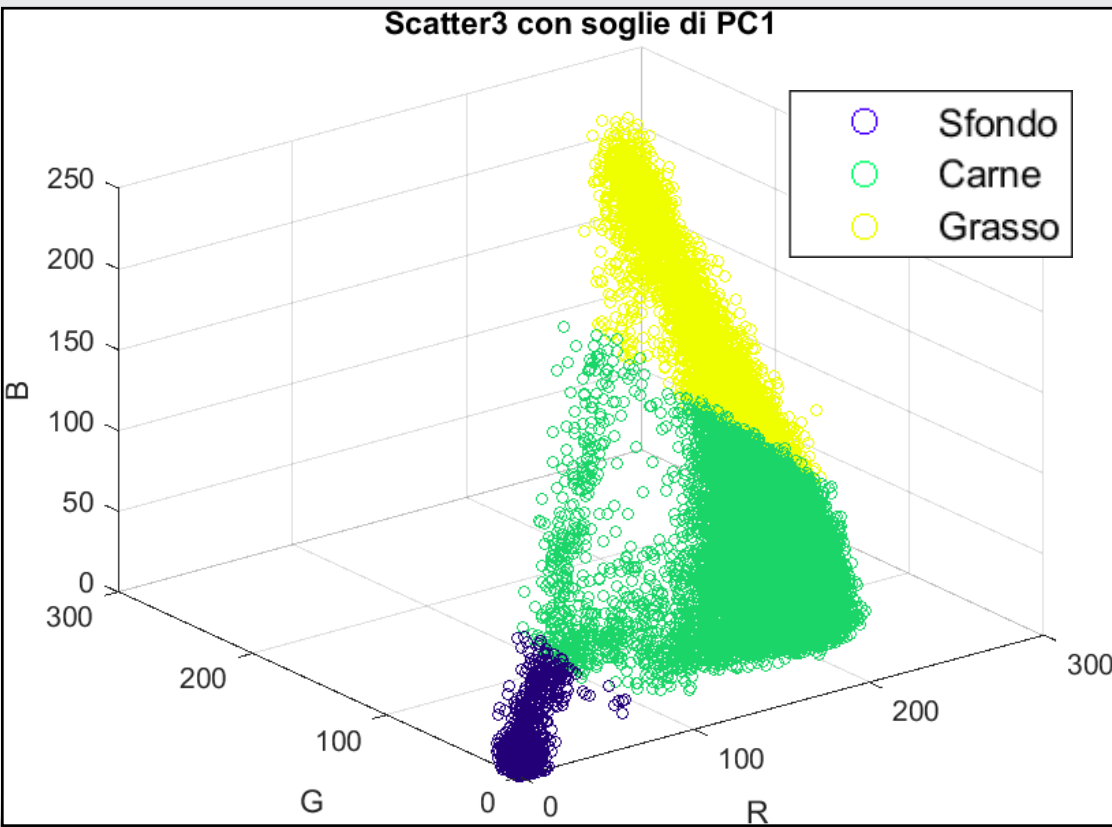


KMeans (EuclideanSq)

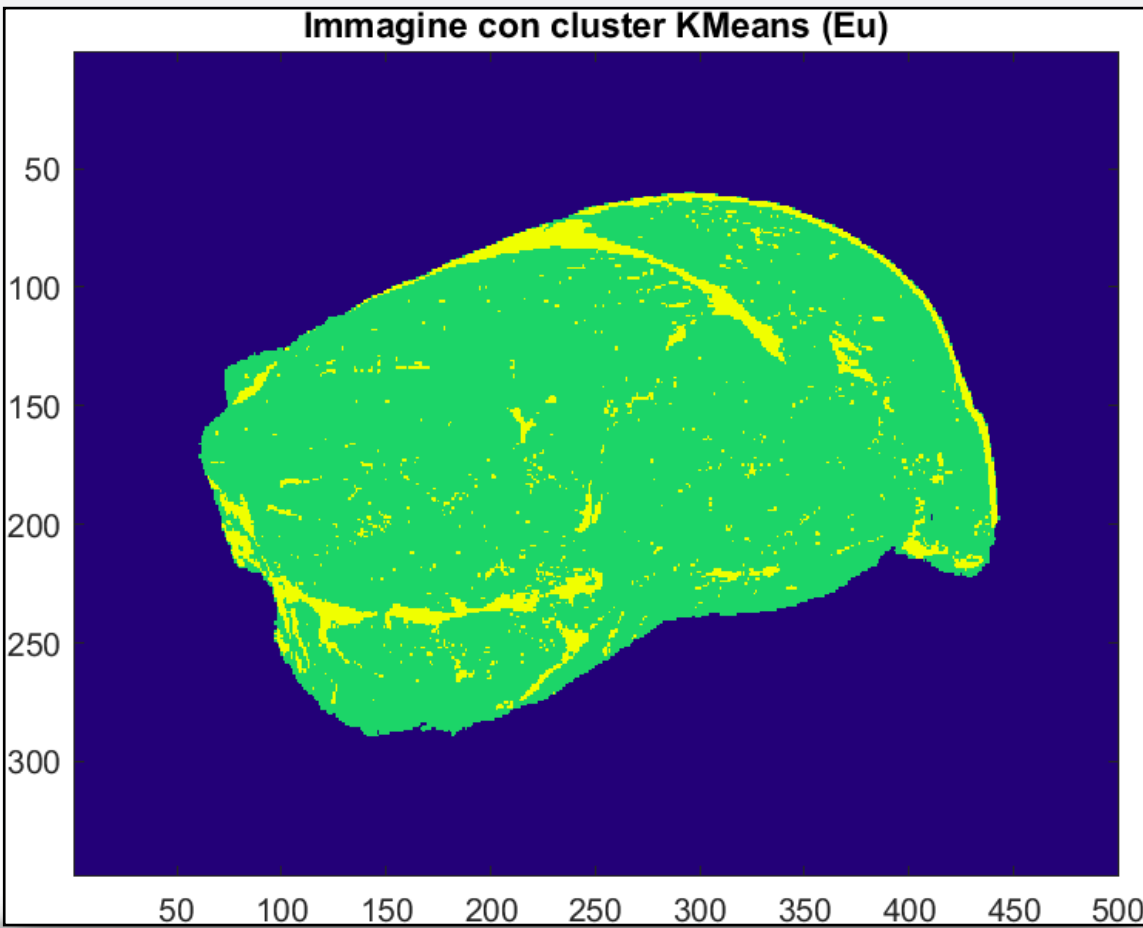
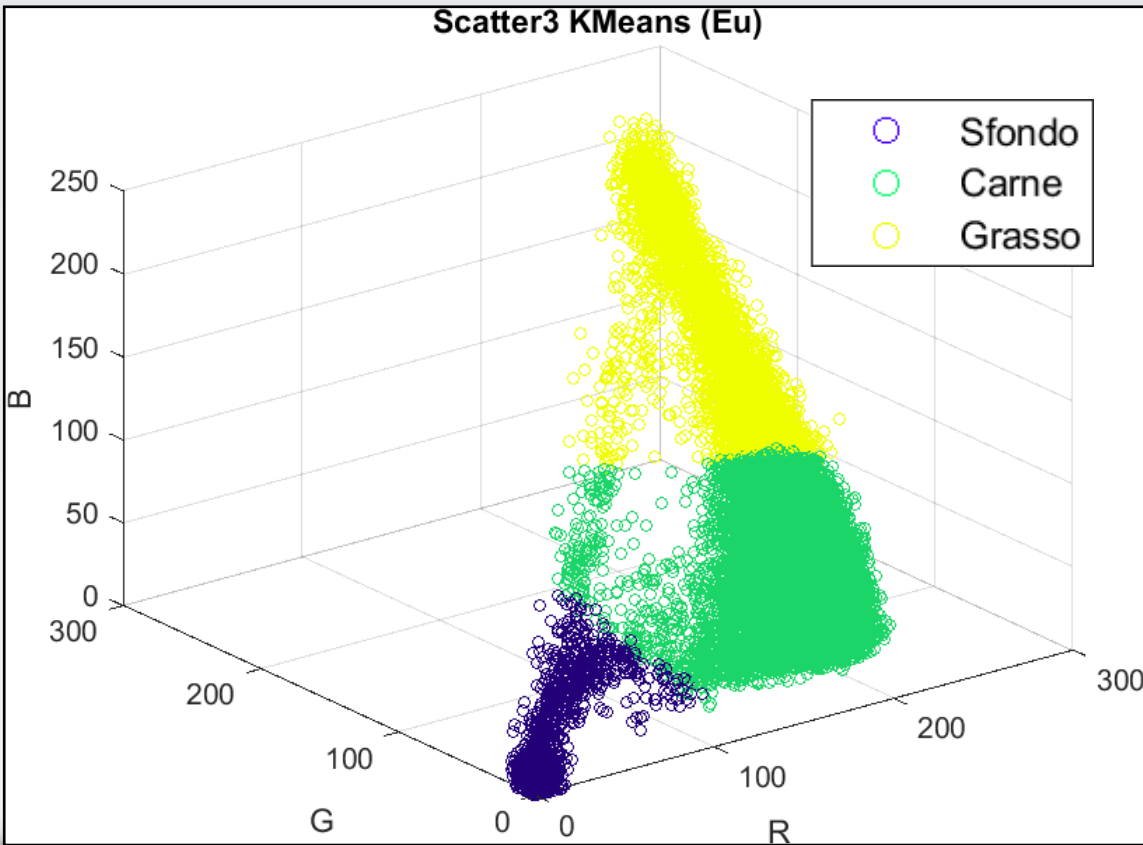


# 5. Clustering - KMeans

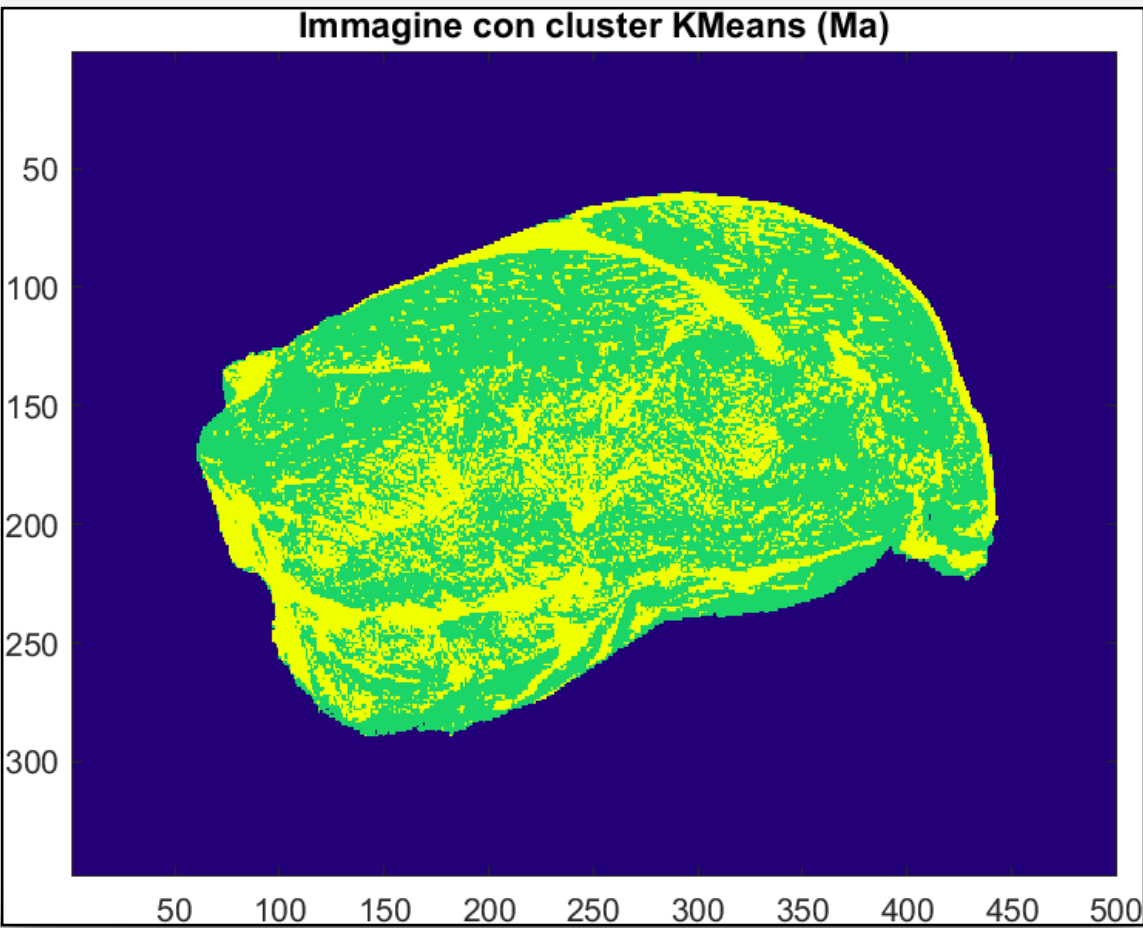
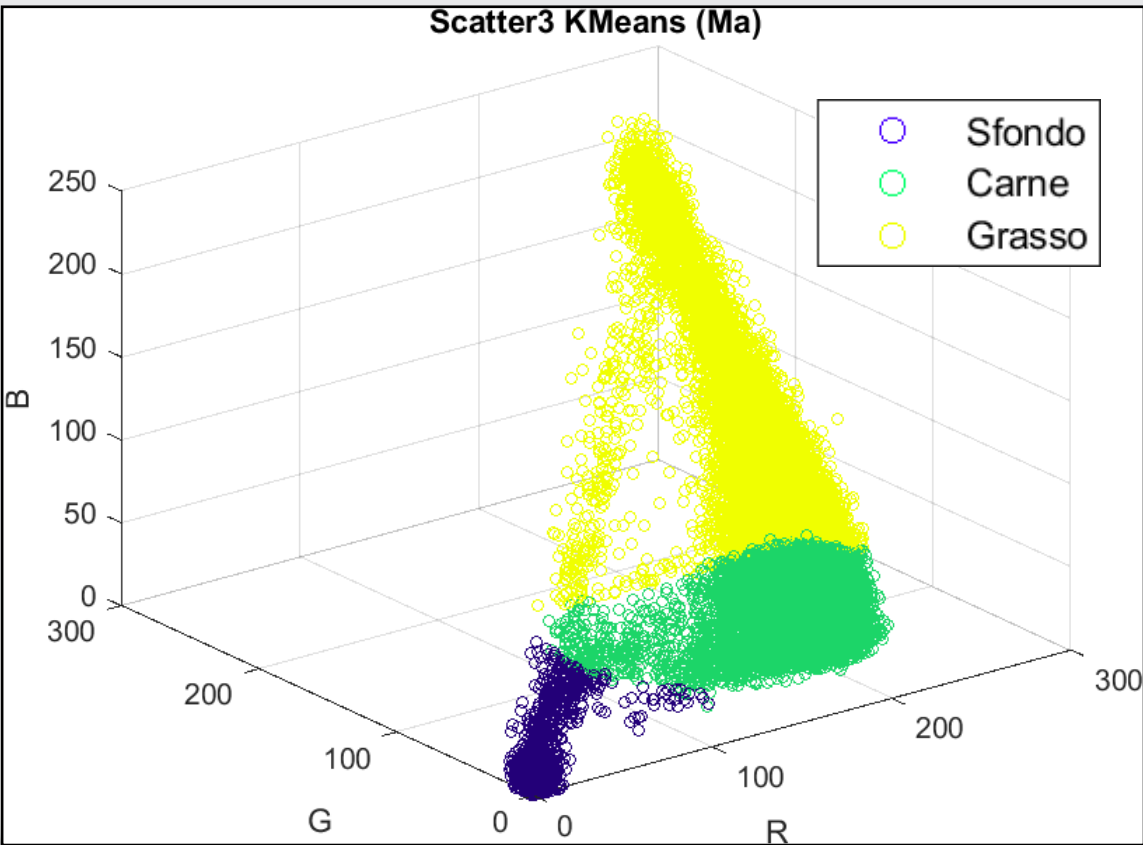
PC1 con soglie



KMeans (EuclideanSq)



KMeans (Manhattan)





## 5. Clustering - DBScan

## 5. Clustering - DBScan

### Parametri

- **distance** = Manhattan
- **minPts** = 6
- **eps** = [3, 4, 5]

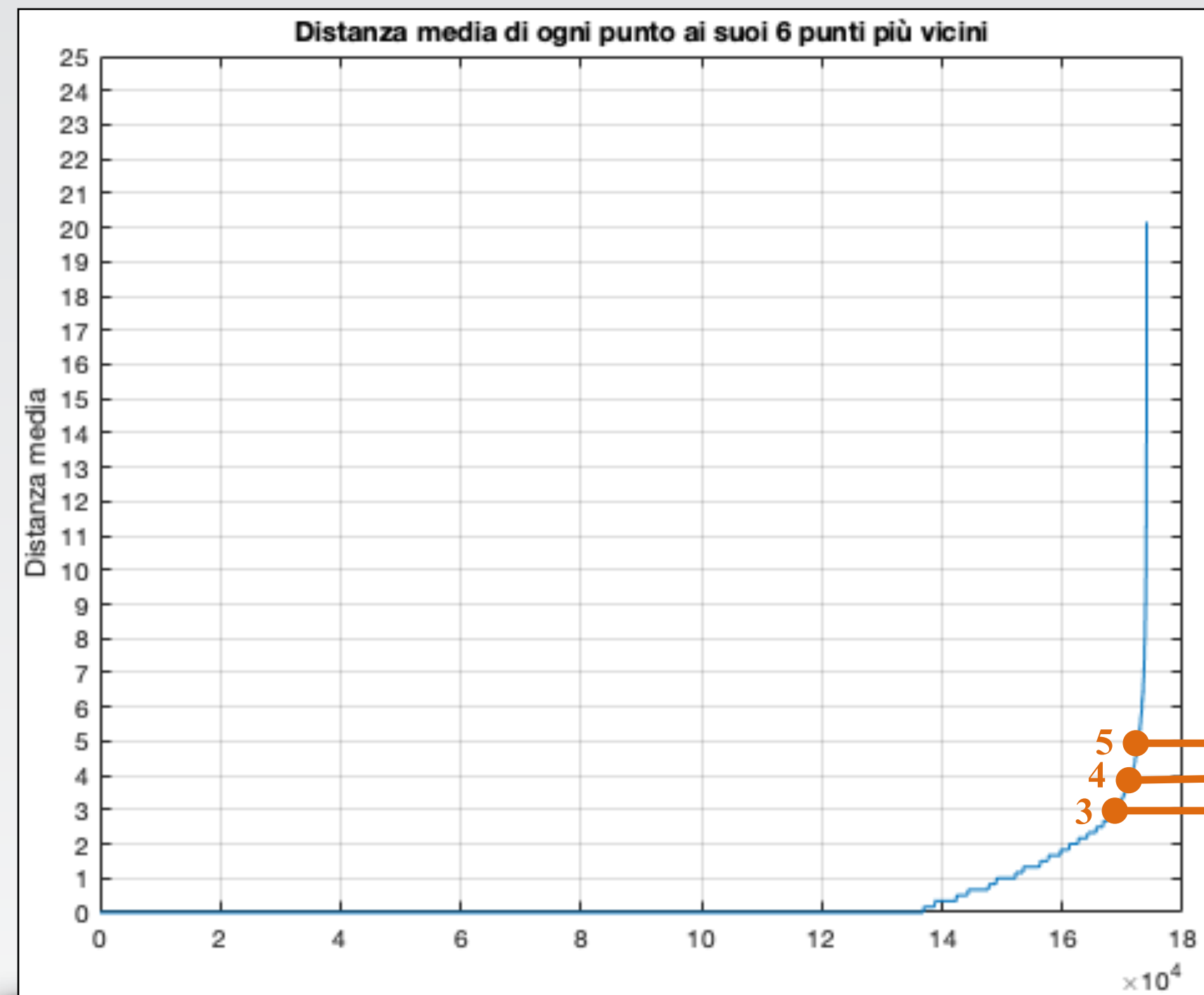
## 5. Clustering - DBScan

### Parametri

- **distance** = Manhattan
- **minPts** = 6
- **eps** = [3, 4, 5]

2 \* dimensionalità dataset

## 5. Clustering - DBScan



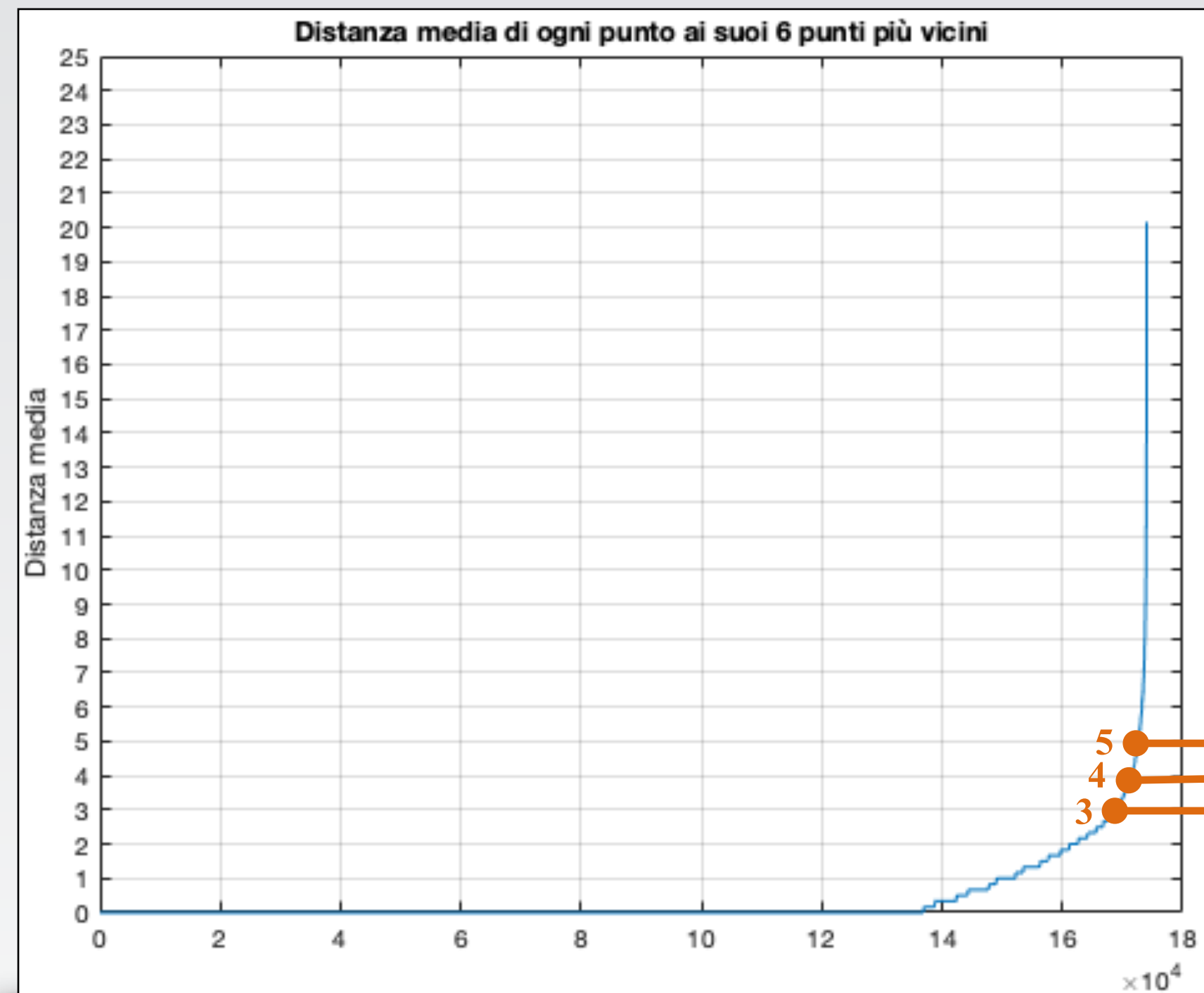
K Distance Graph

### Parametri

- distance = Manhattan
- minPts = 6
- eps = [3, 4, 5]

2 \* dimensionalità dataset

## 5. Clustering - DBScan



K Distance Graph

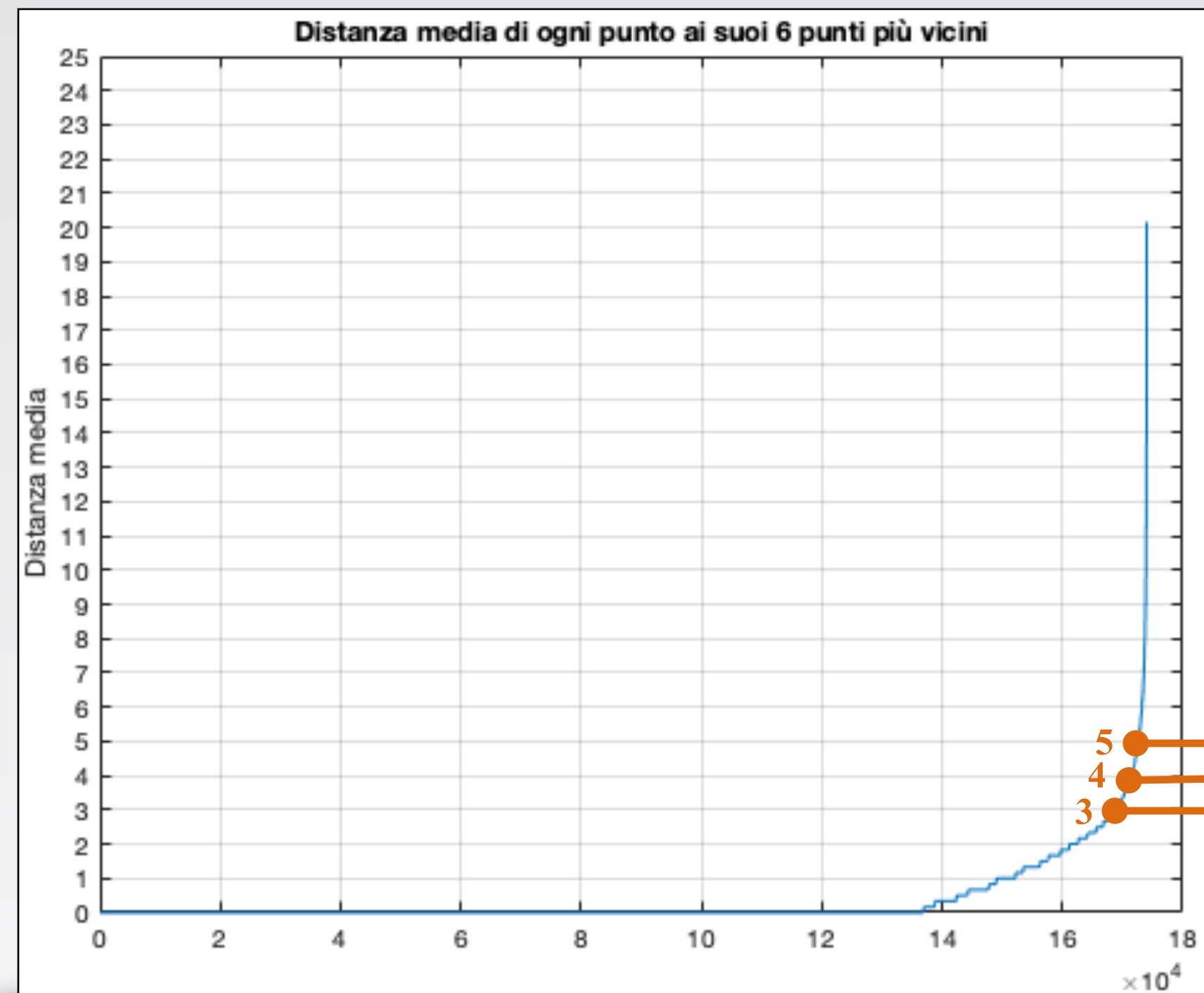
### Parametri

- distance = Manhattan
- minPts = 6
- eps = [3, 4, 5]

2 \* dimensionalità dataset

minPts = 6	
eps	Numero Cluster
3	159
4	58
5	24

## 5. Clustering - DBScan



K Distance Graph

### Parametri

- distance = Manhattan
- minPts = 6
- eps = [3, 4, 5]

2 \* dimensionalità dataset

Troppi cluster trovati

↓  
I parametri (minPts e eps)  
sembrano non essere  
adatti al dataset.

minPts = 6	
eps	Numero Cluster
3	159
4	58
5	24

## 5. Clustering - DBScan



## 5. Clustering - DBScan

### Parametri

- **distance** = Manhattan
- **minPts** = 50
- **eps** = [7, 8, 9]

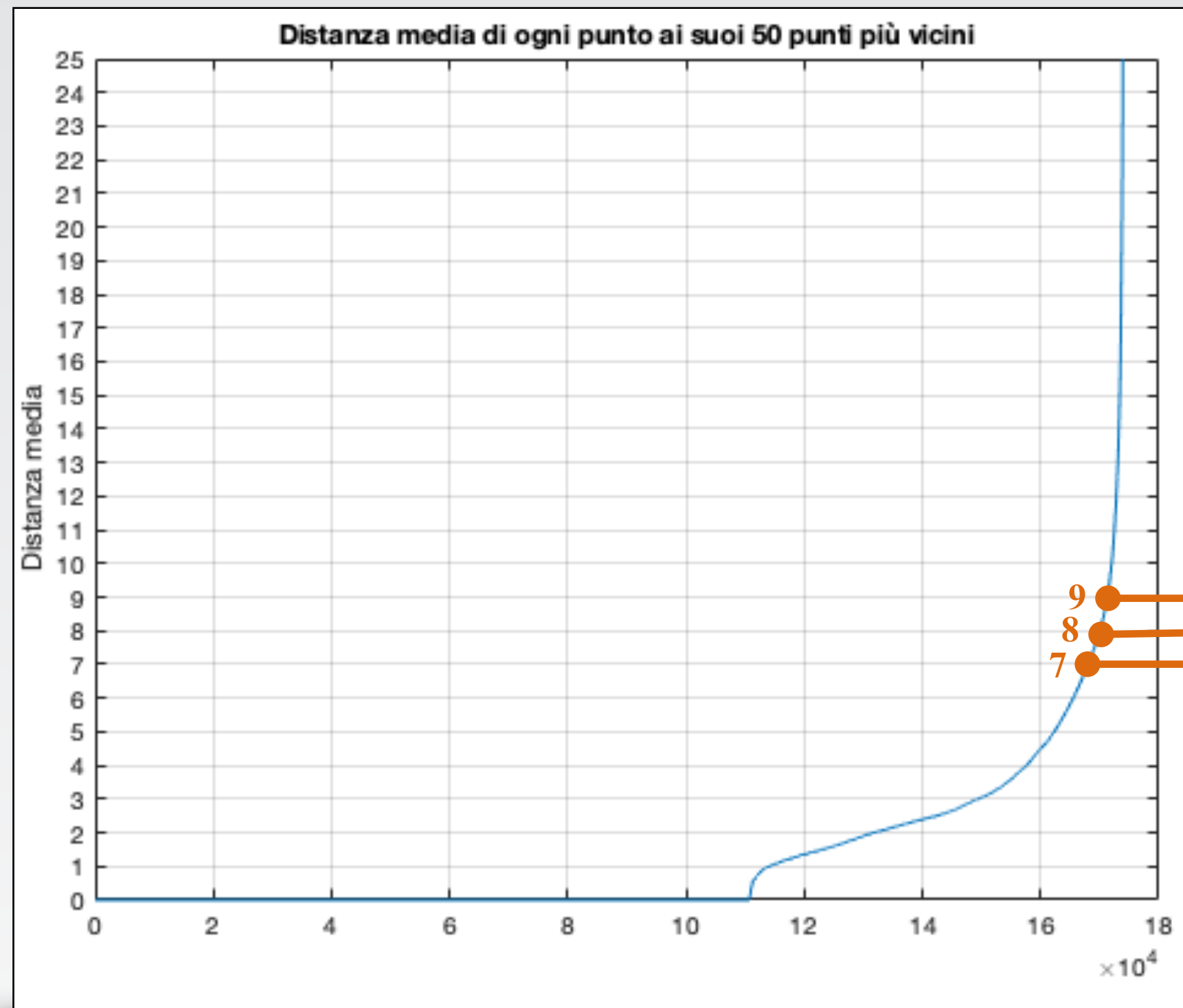
## 5. Clustering - DBScan

### Parametri

- **distance** = Manhattan
- **minPts** = 50
- **eps** = [7, 8, 9]

**Dataset soggetto ad avere duplicati e rumore**

## 5. Clustering - DBScan



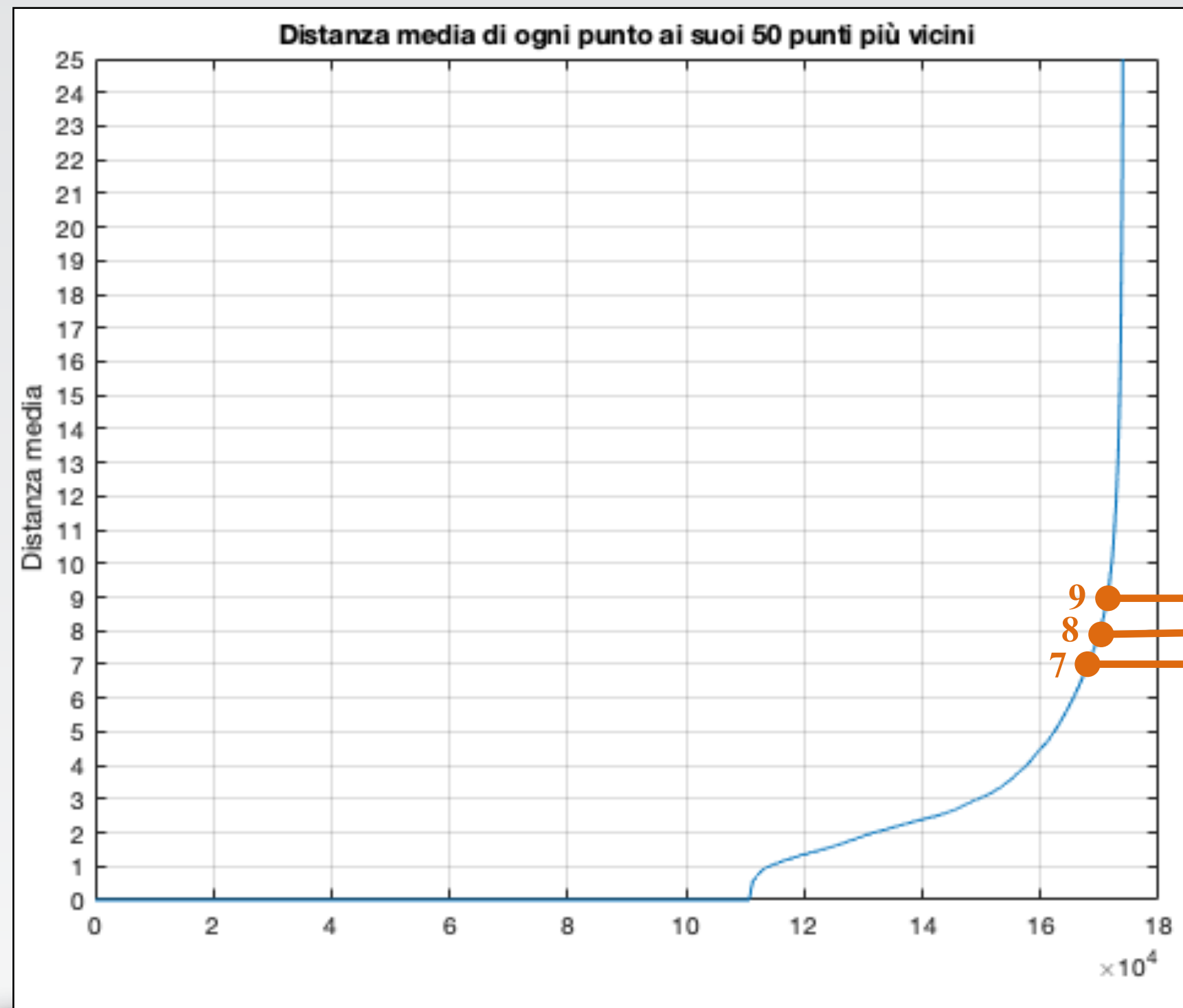
K Distance Graph

### Parametri

- distance = Manhattan
- minPts = 50
- eps = [7, 8, 9]

Dataset soggetto ad avere duplicati e rumore

## 5. Clustering - DBScan



K Distance Graph

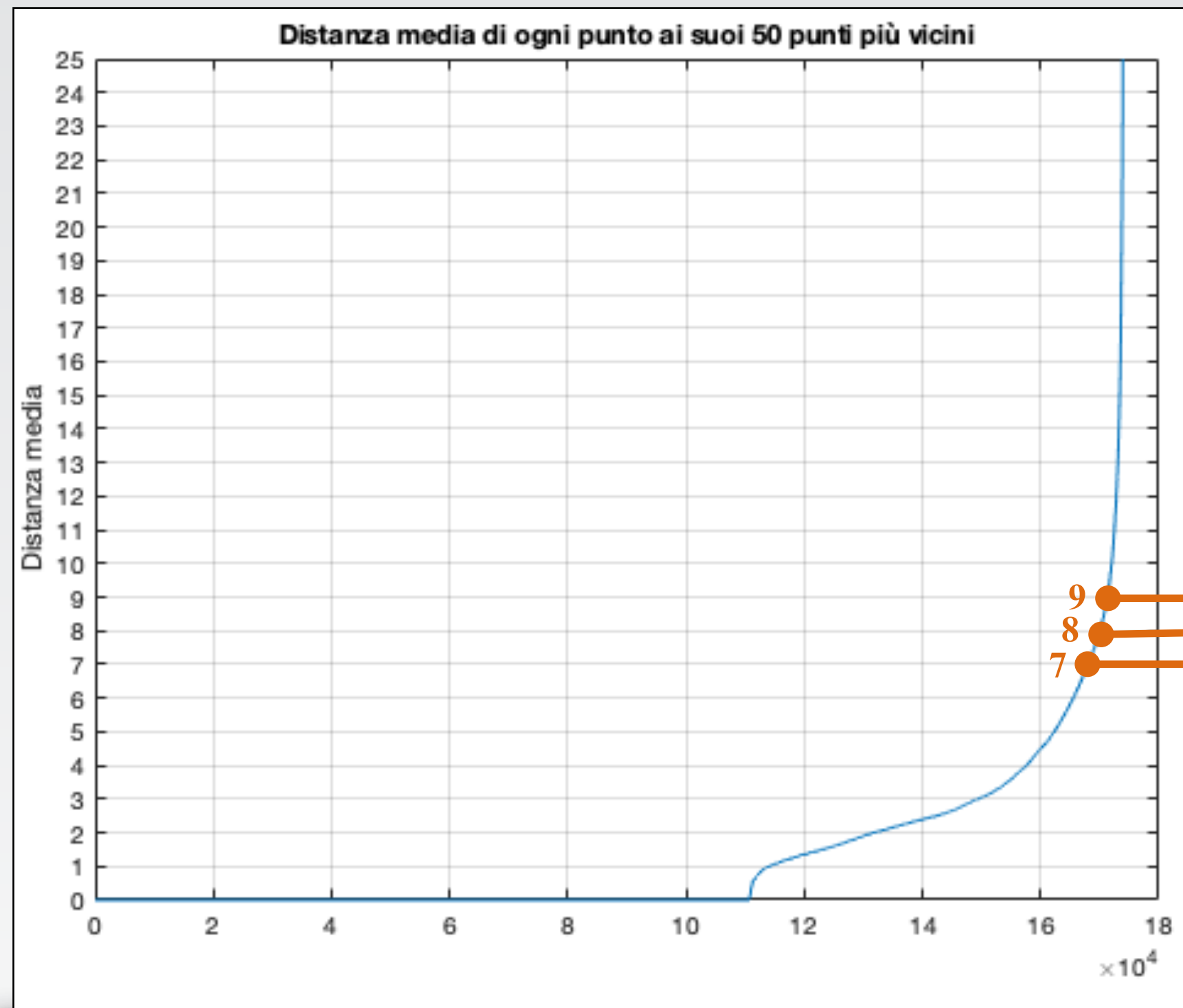
### Parametri

- distance = Manhattan
- minPts = 50
- eps = [7, 8, 9]

Dataset soggetto ad avere duplicati e rumore

minPts = 50	
eps	Numero Cluster
7	4
8	4
9	5

## 5. Clustering - DBScan



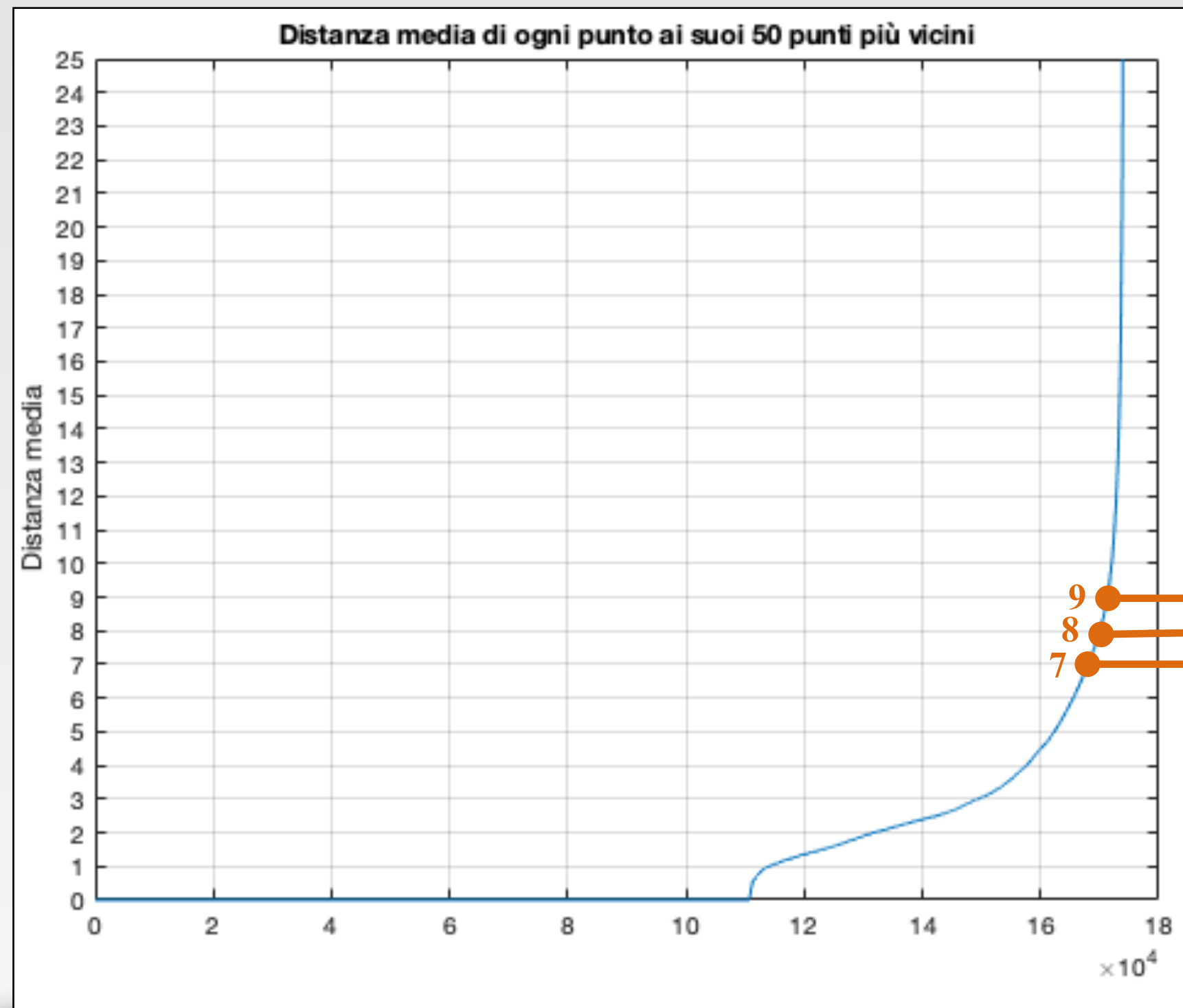
K Distance Graph

### Parametri

- distance = Manhattan
- minPts = 50
- eps = [7, 8, 9]

Dataset soggetto ad avere duplicati e rumore

## 5. Clustering - DBScan

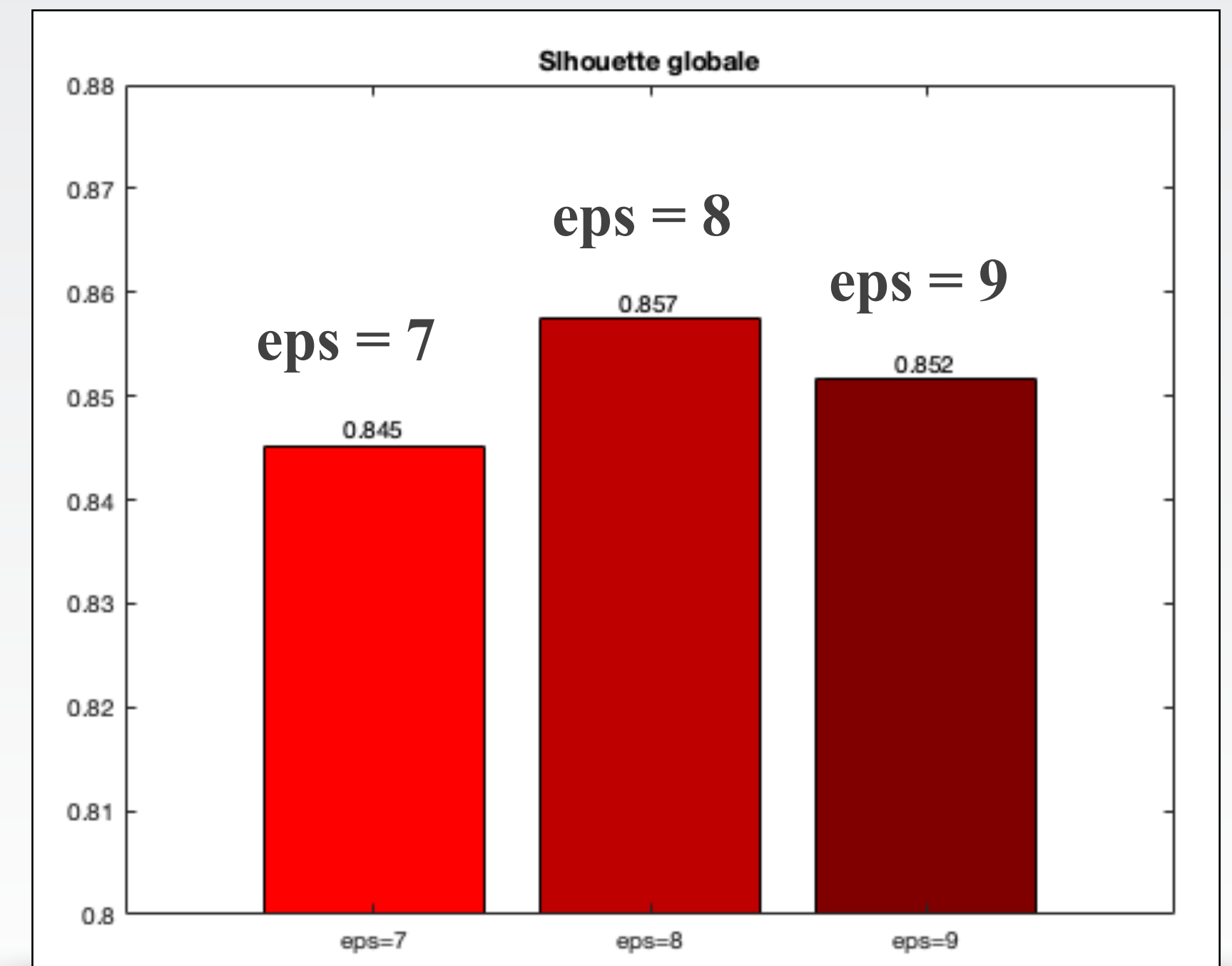


K Distance Graph

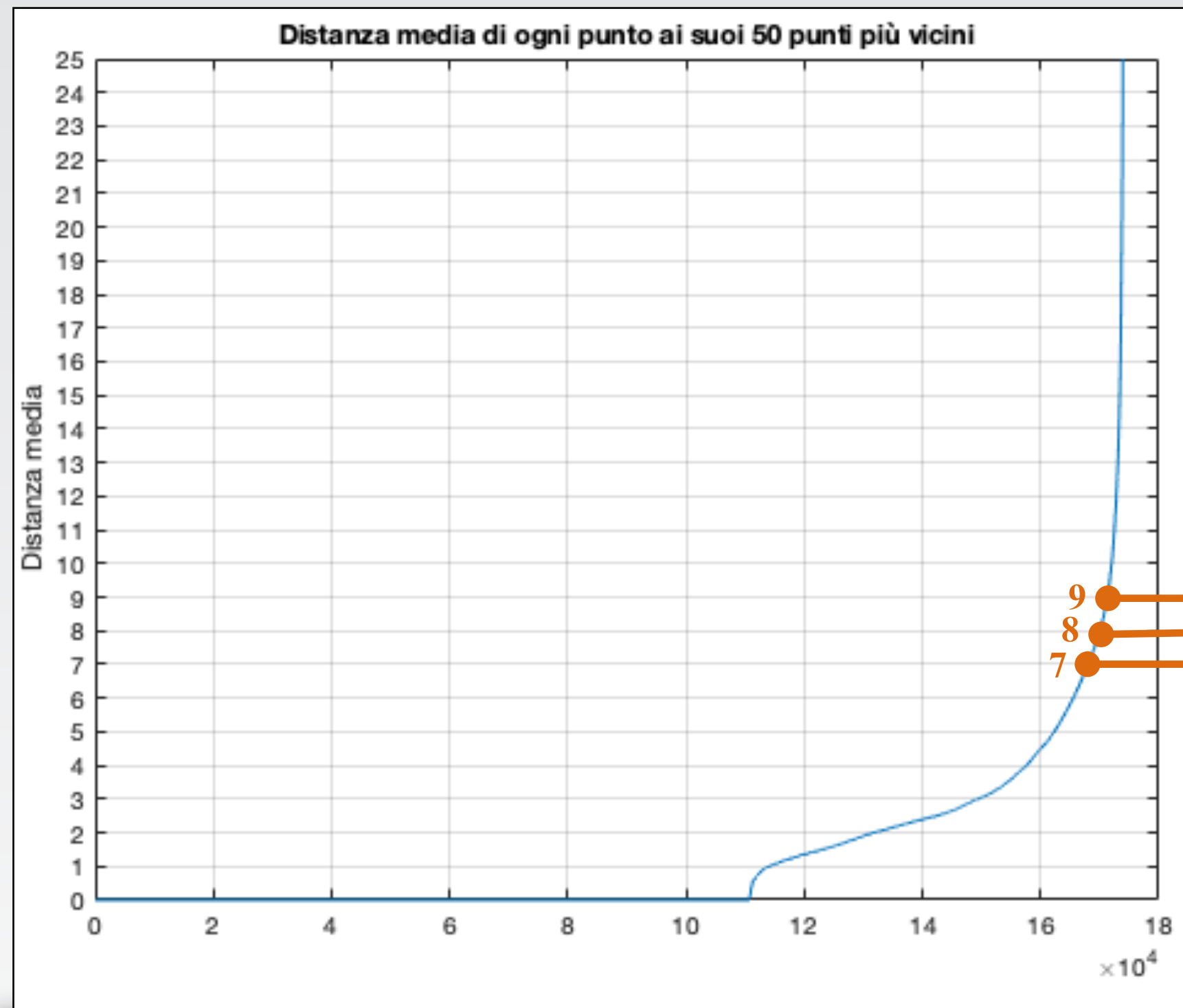
### Parametri

- distance = Manhattan
- minPts = 50
- eps = [7, 8, 9]

Dataset soggetto ad avere duplicati e rumore



## 5. Clustering - DBScan

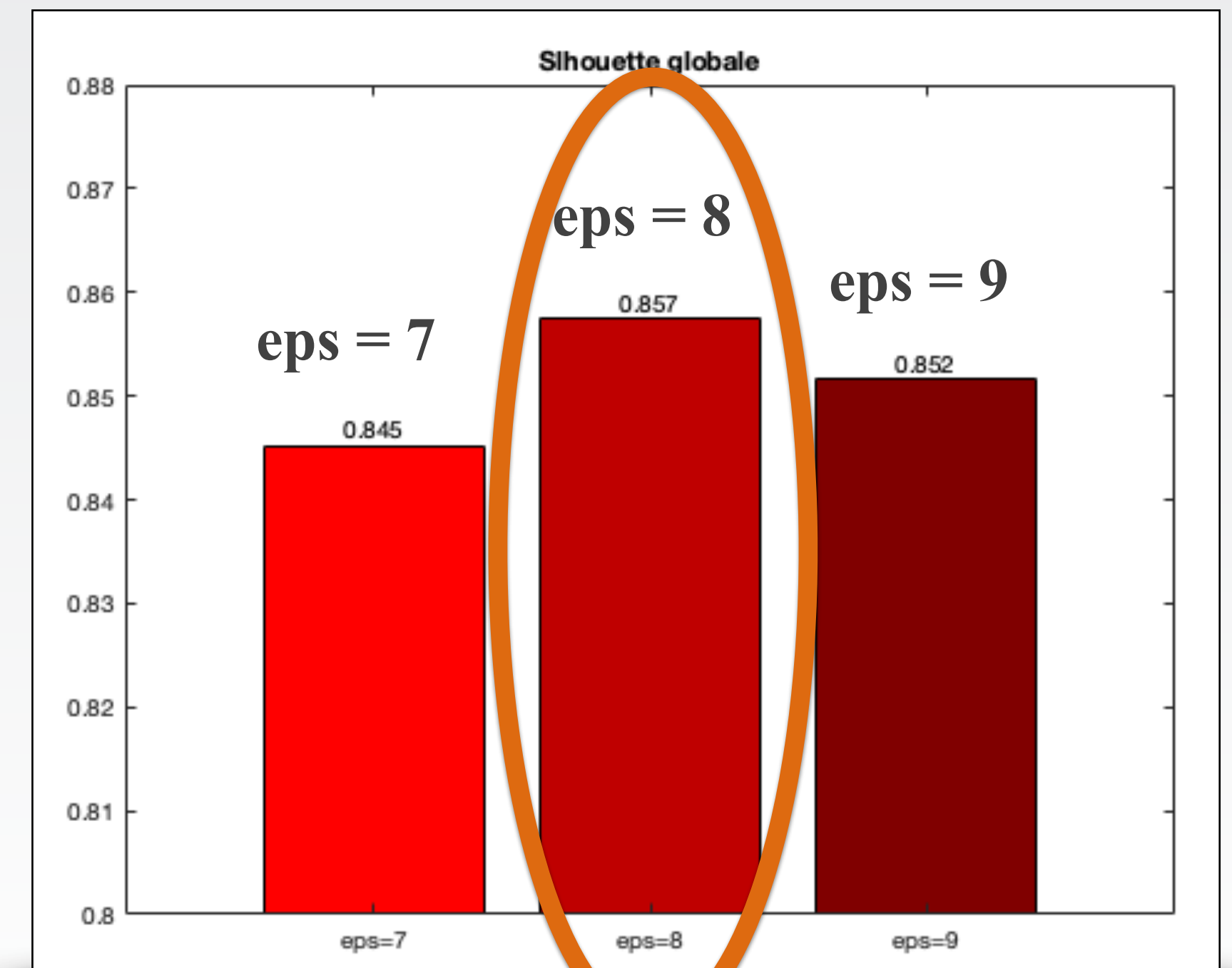


K Distance Graph

### Parametri

- distance = Manhattan
- minPts = 50
- eps = [7, 8, 9]

Dataset soggetto ad avere duplicati e rumore





## 5. Clustering - DBScan

# 5. Clustering - DBScan

## Parametri migliori trovati

- **distance** = Manhattan
- **minPts** = 50
- **eps** = 8

## 5. Clustering - DBScan

### Parametri migliori trovati

- **distance** = Manhattan
- **minPts** = 50
- **eps** = 8

Scatter 3 del cluster

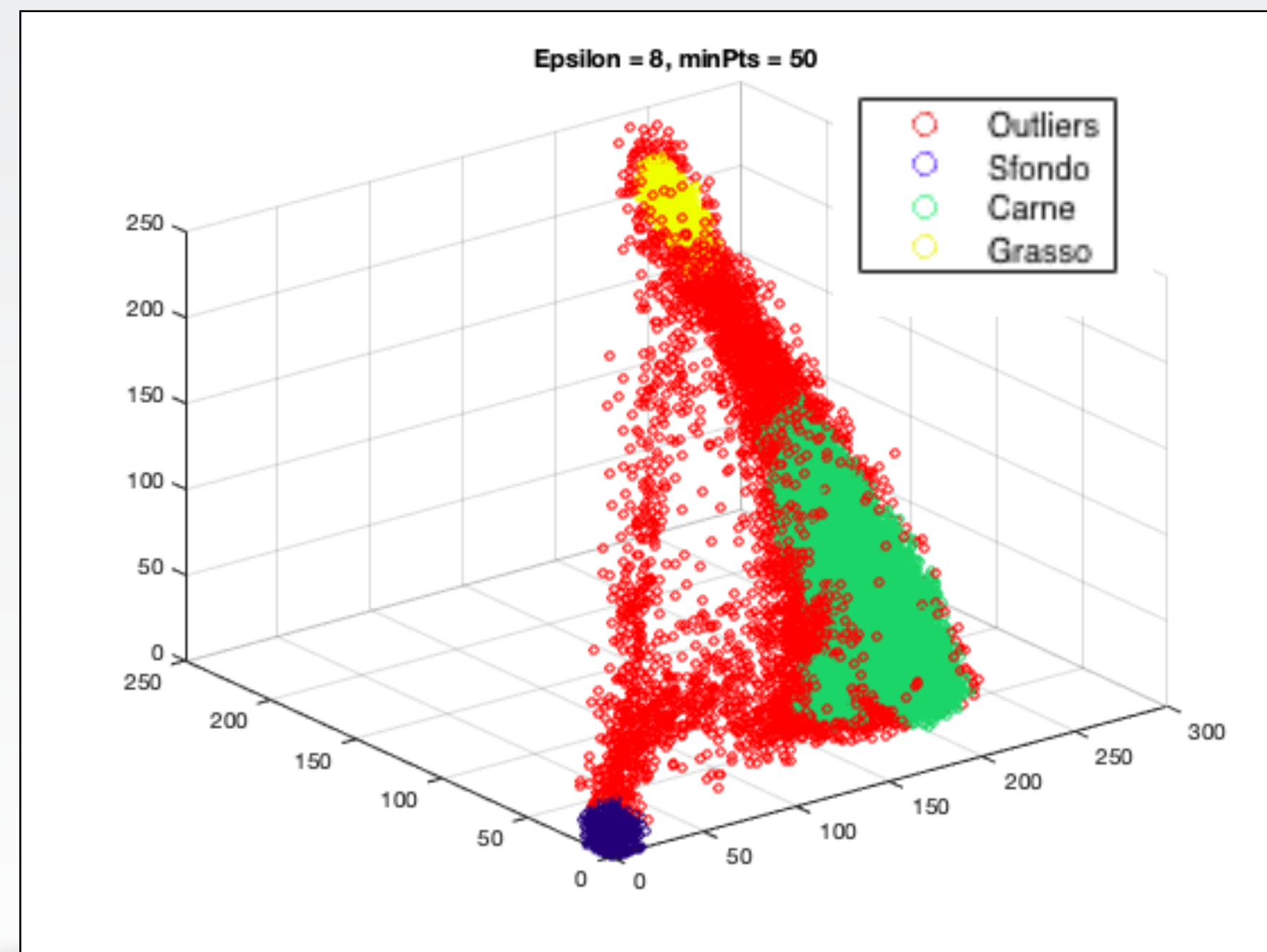
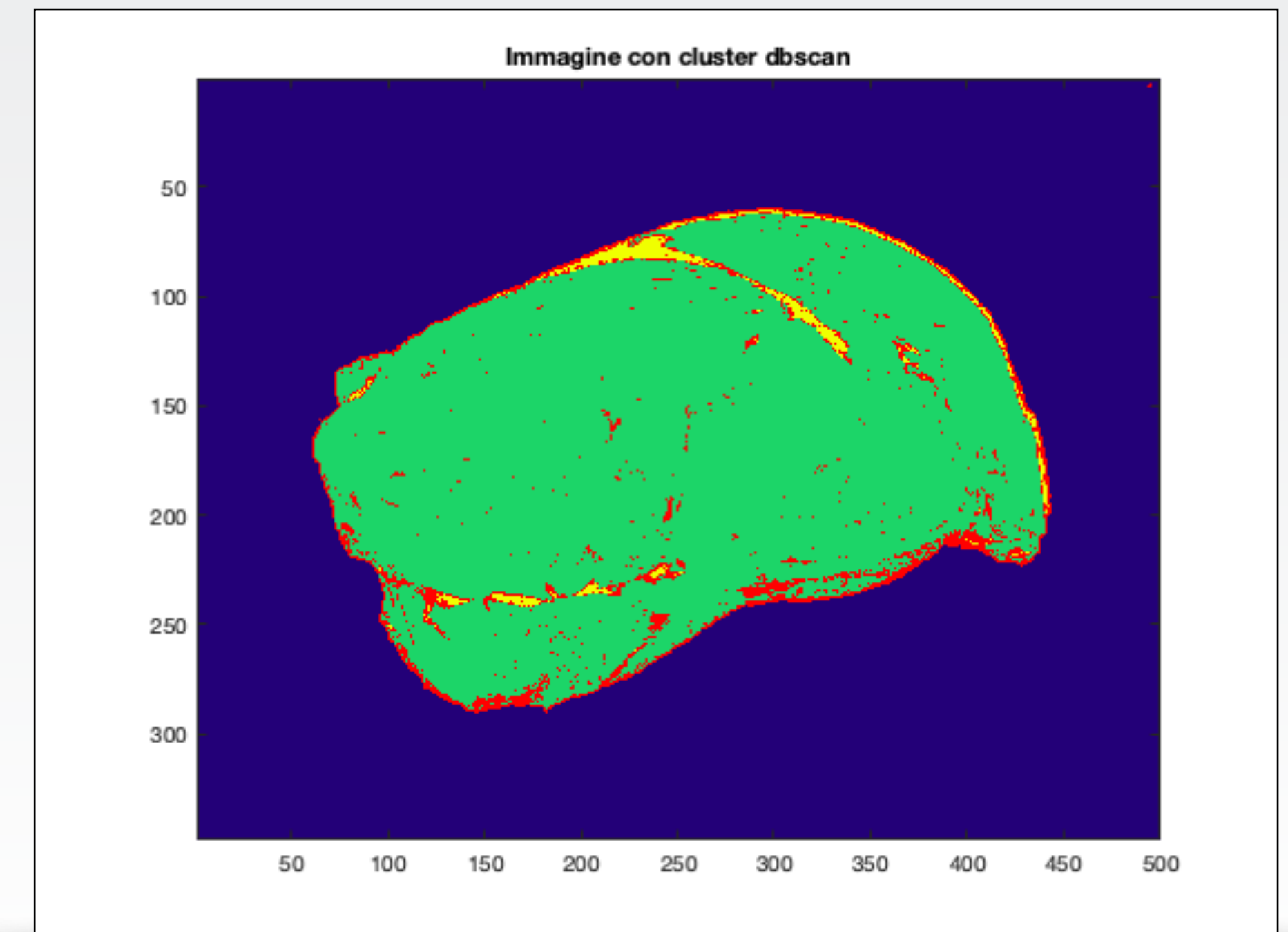


Immagine colorata per cluster



# 5. Clustering - DBScan

## Parametri migliori trovati

- **distance** = Manhattan
- **minPts** = 50
- **eps** = 8

# Outliers

4219

# Sfondo

113841

# Carne

54732

# Grasso

1208

Scatter 3 del cluster

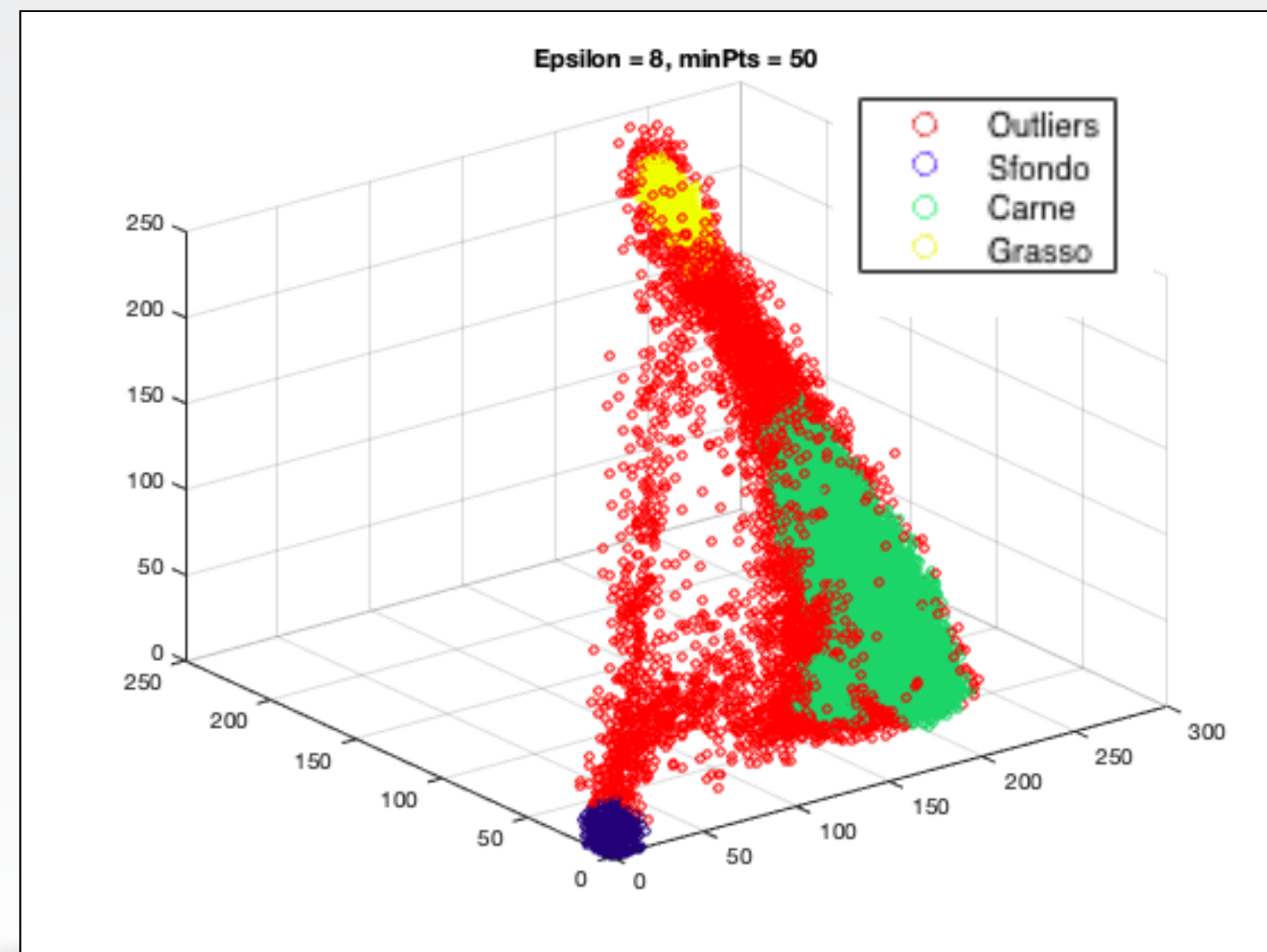
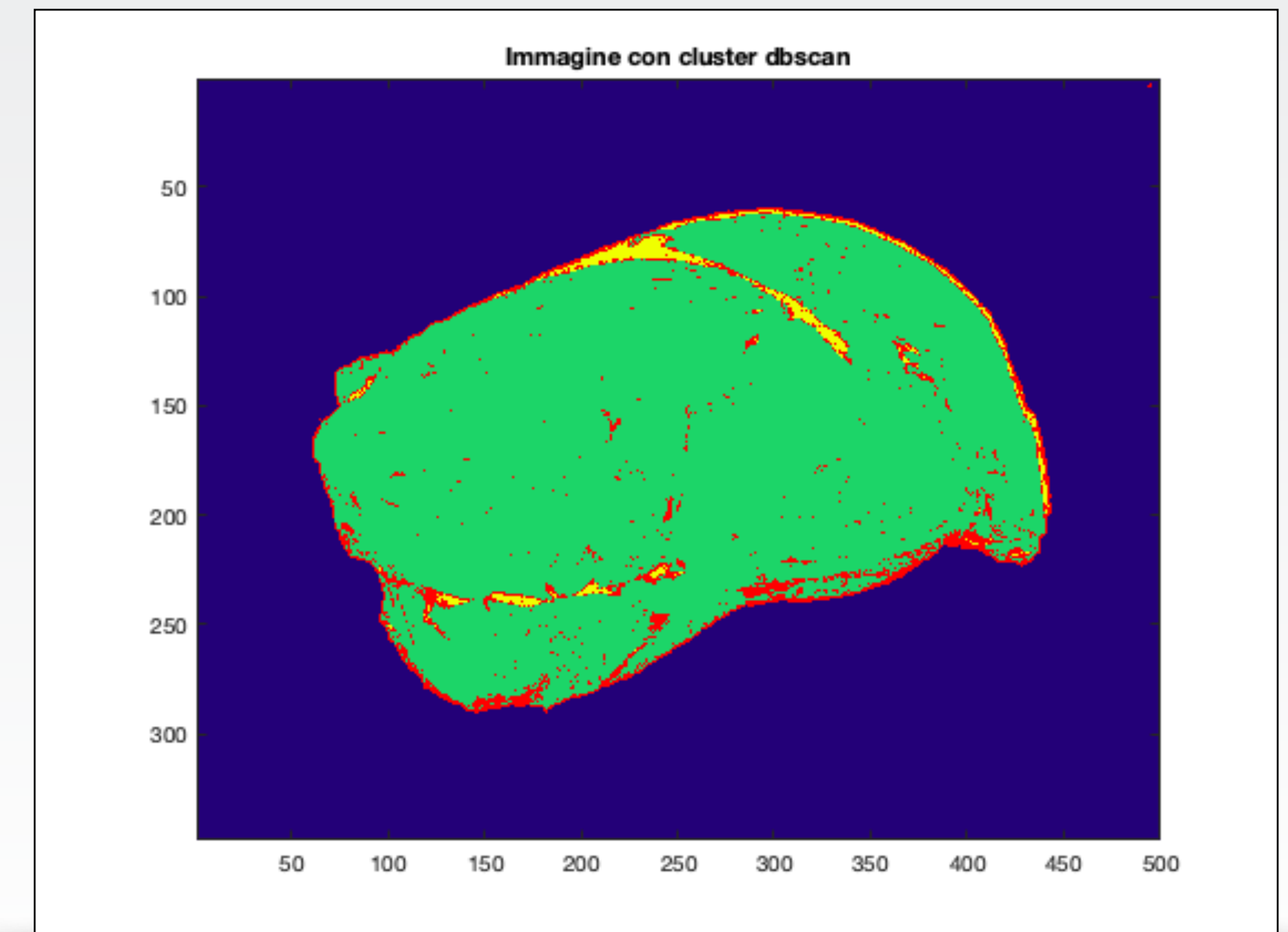


Immagine colorata per cluster



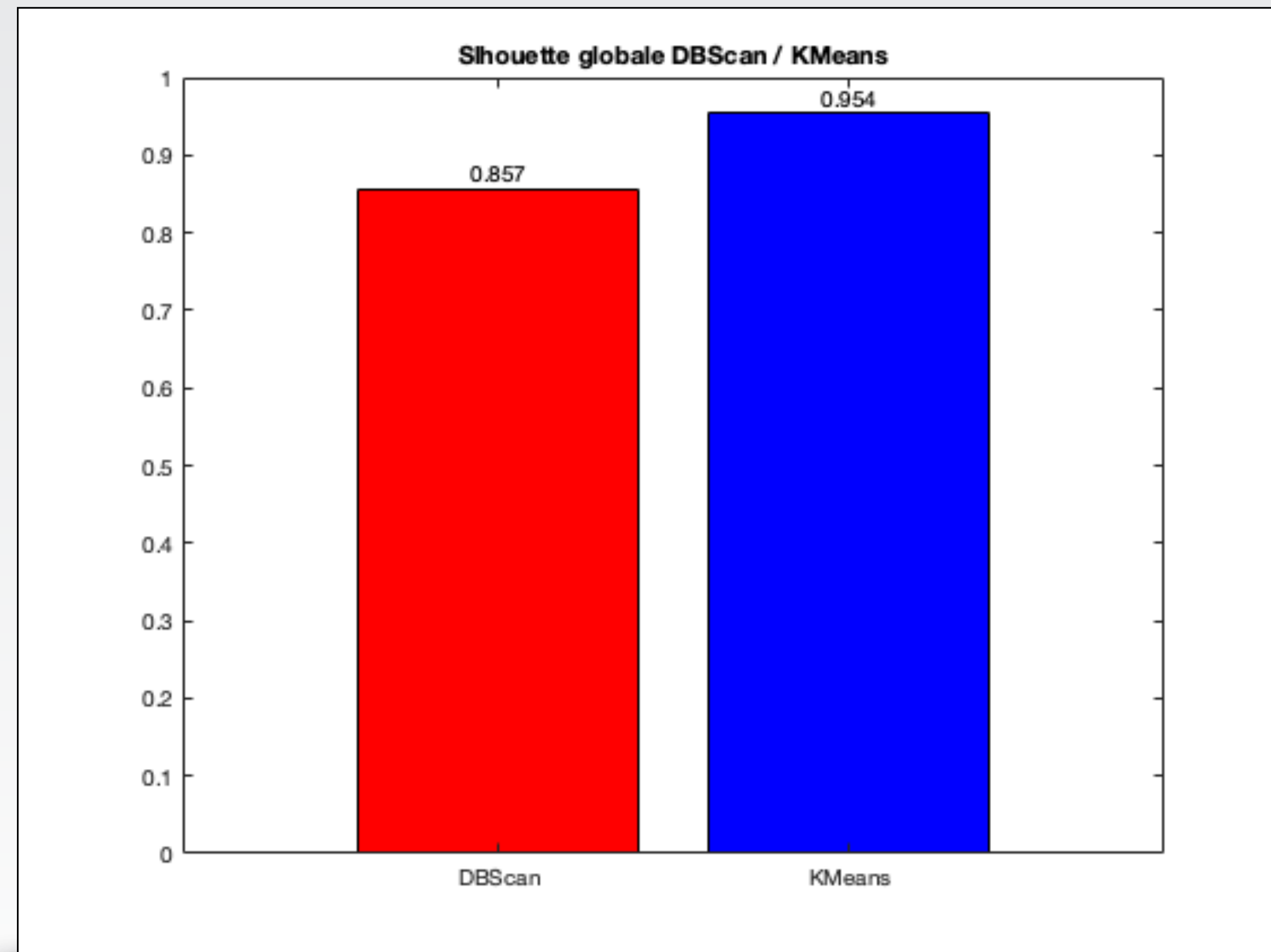
# 5. Clustering

# 5. Clustering

## Confronto tra KMeans e DBScan

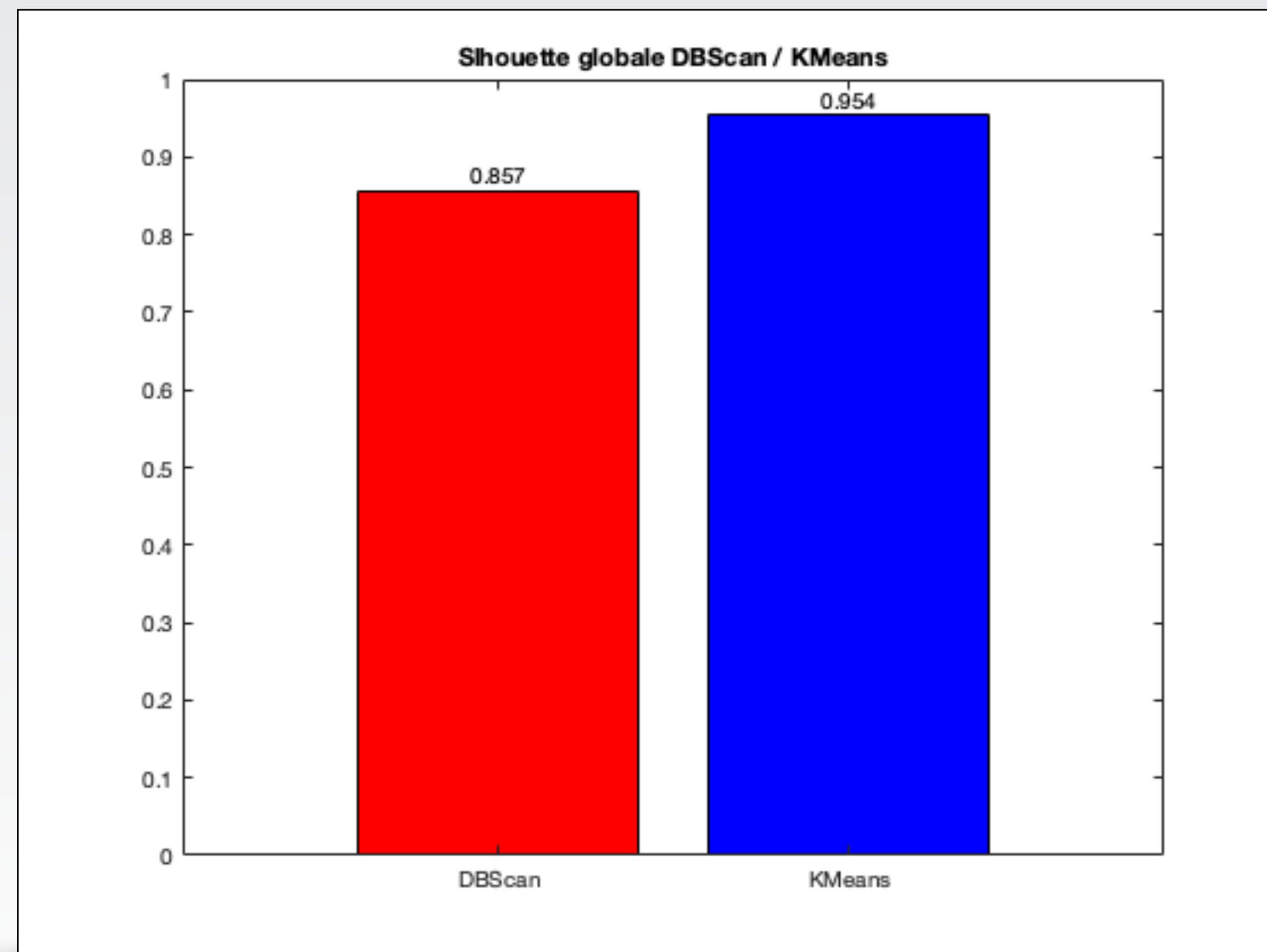
# 5. Clustering

## Confronto tra KMeans e DBScan



# 5. Clustering

## Confronto tra KMeans e DBScan



Il valore di **Silhouette del KMeans** è **maggiore** rispetto a quello del DBScan

**KMeans** riesce a creare **cluster migliori** rispetto al DBScan



**1. Descrizione dataset**

**2. Obiettivi del progetto**

**3. Analisi esplorativa**

**4. PCA esplorativa**

**5. Clustering**

**6. Previsione grasso**

## 6. Previsione grasso

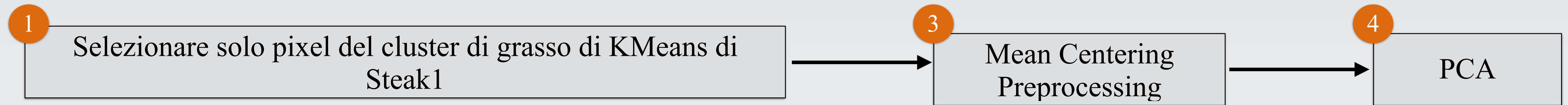
L'obiettivo è adesso la previsione del grasso su una nuova immagine (steak2). Per farlo sono stati seguiti i seguenti passaggi:

- A. Creazione di un modello PCA sulla base dei soli pixel del grasso trovati dal metodo di clustering migliore (KMeans in questo caso) su steak1
- B. Proiezione dei pixel di steak2 nello spazio PCA
- C. Tutti quei punti con valore  $Q < Q_{lim}$  e  $T^2 < T^2_{lim}$  sono considerati come grasso

## 6. Previsione grasso

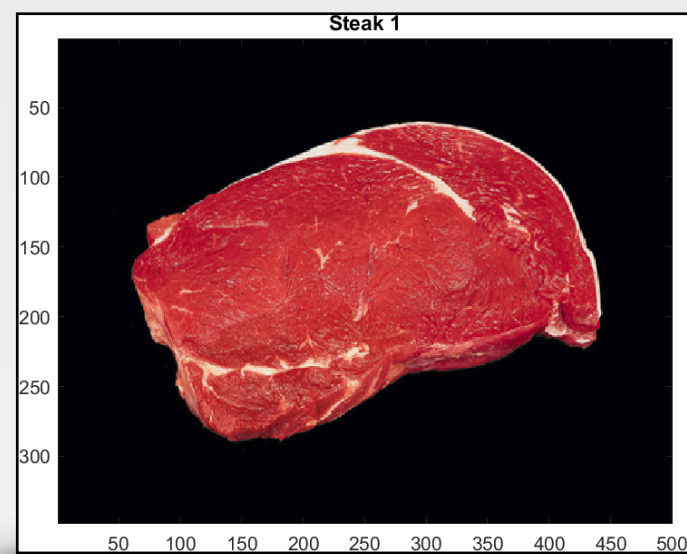
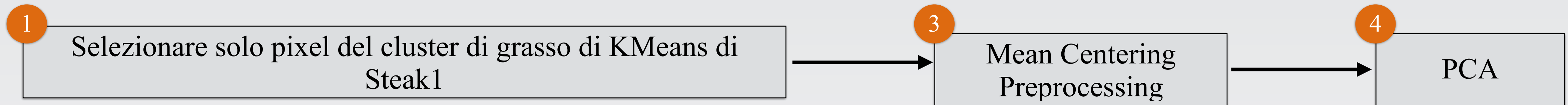
## 6. Previsione grasso

### A. Creazione modello PCA

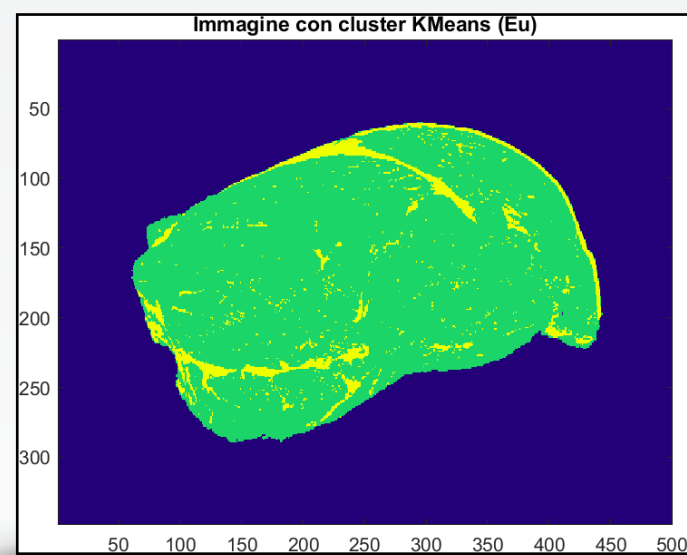


# 6. Previsione grasso

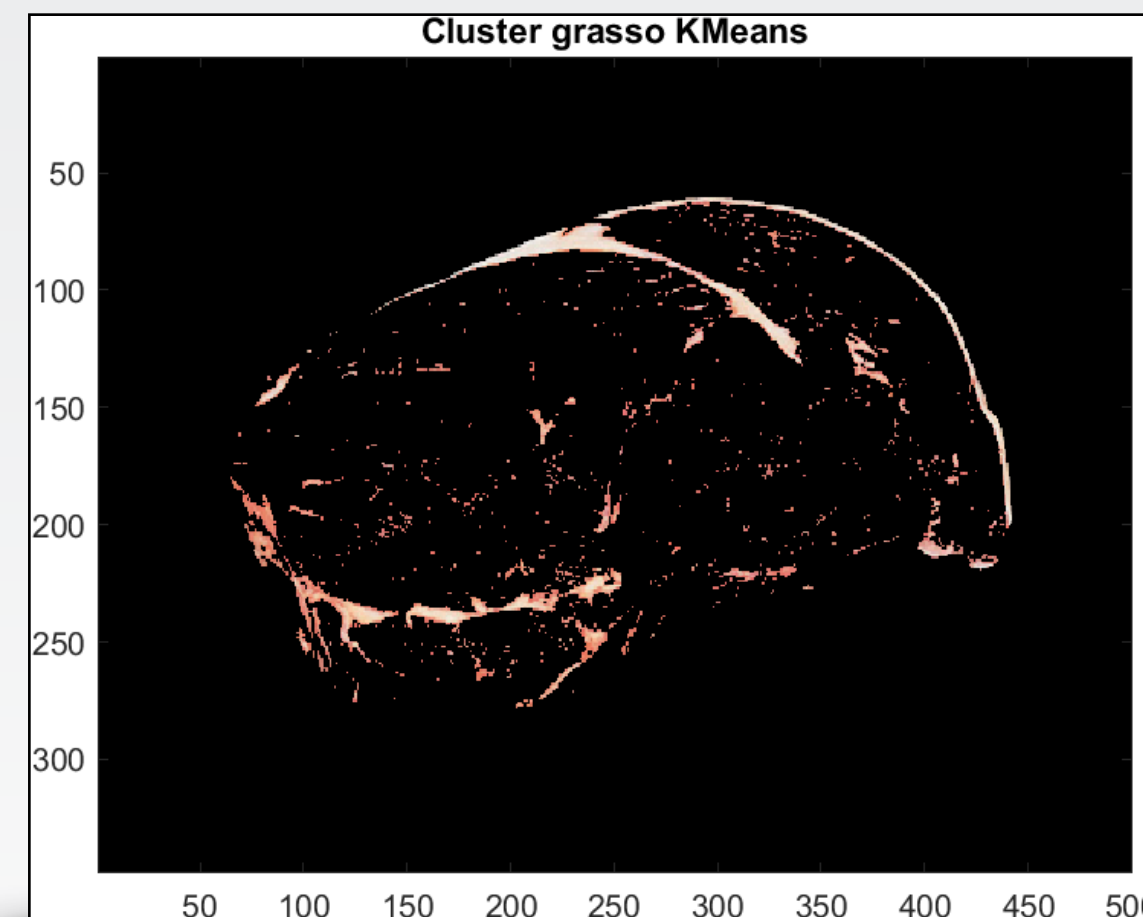
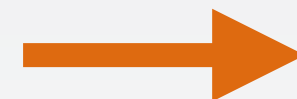
## A. Creazione modello PCA



MASK



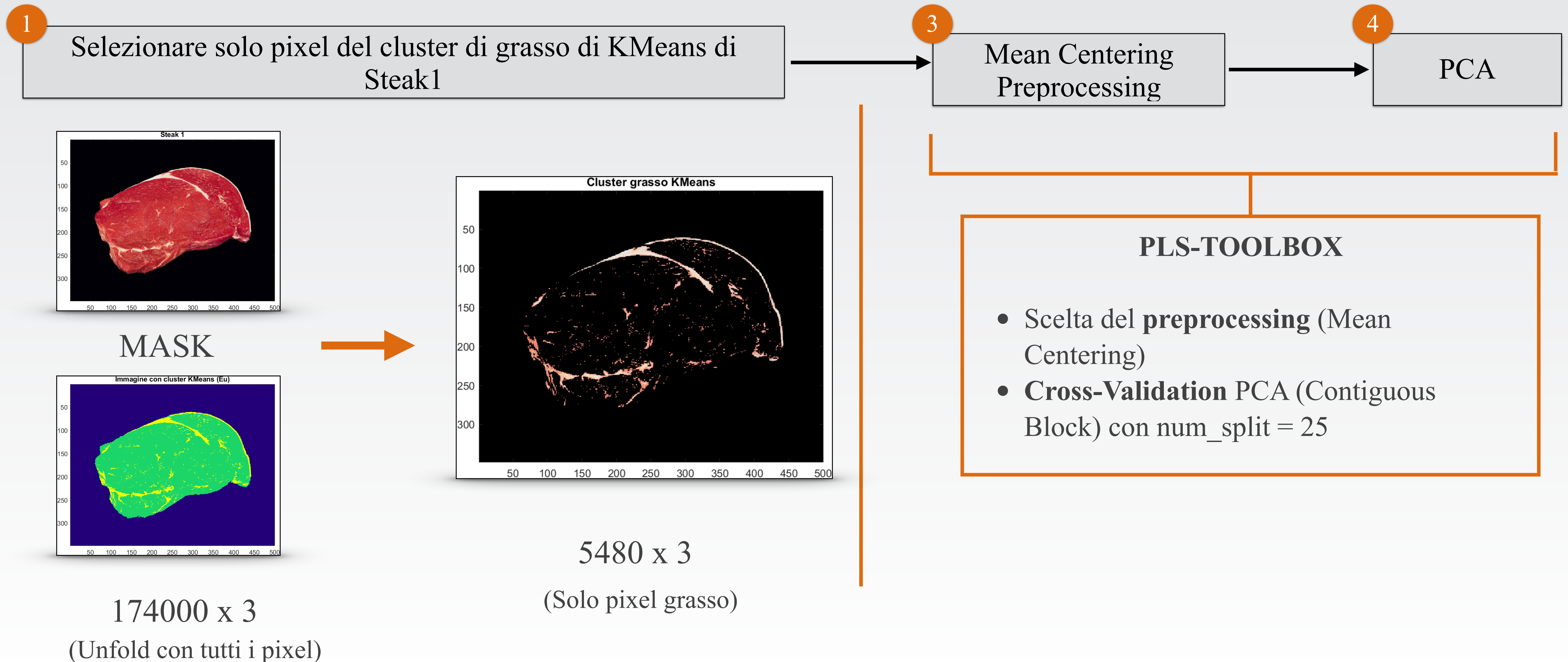
174000 x 3  
(Unfold con tutti i pixel)



5480 x 3  
(Solo pixel grasso)

# 6. Previsione grasso

## A. Creazione modello PCA



## 6. Previsione grasso

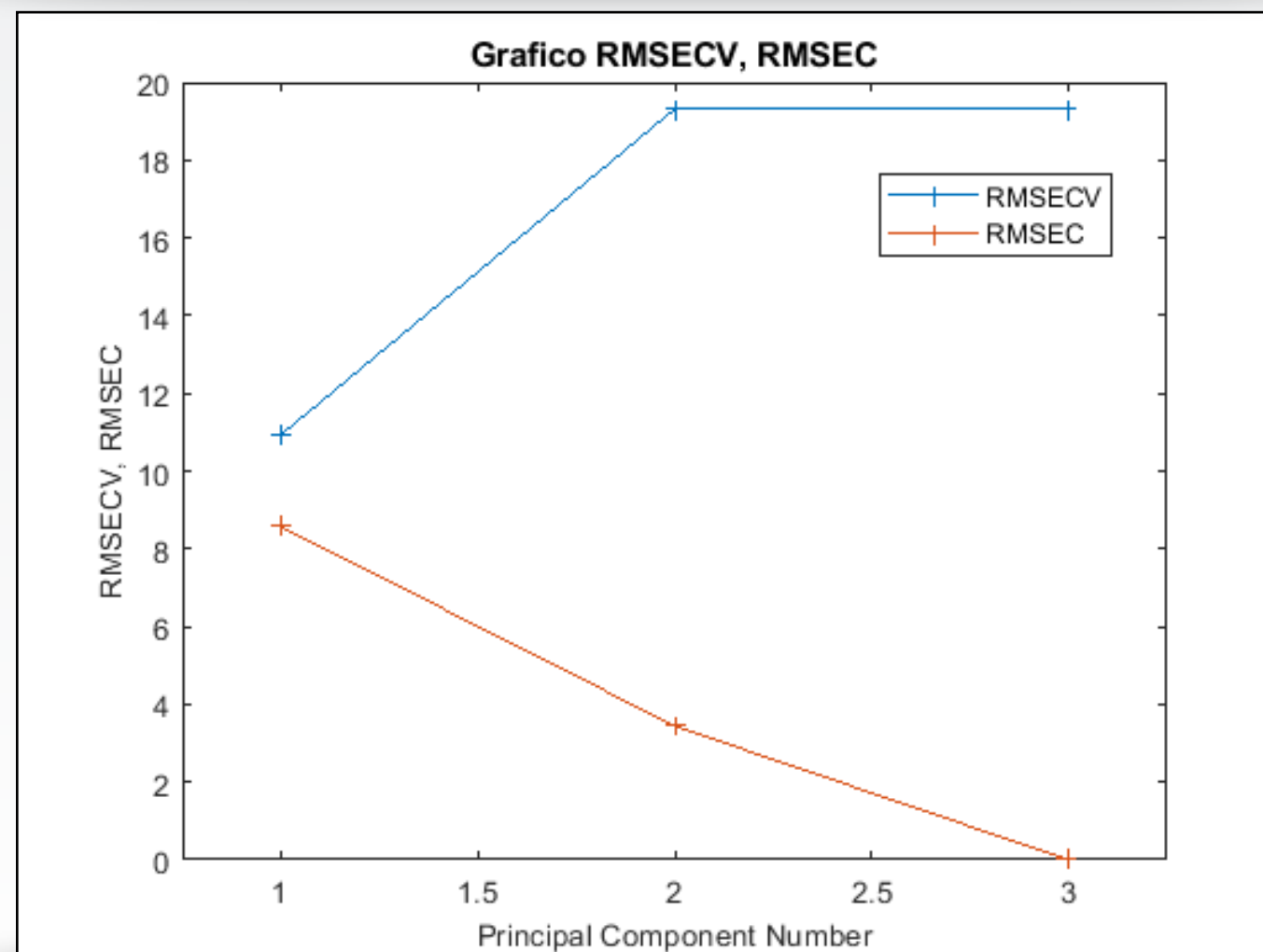
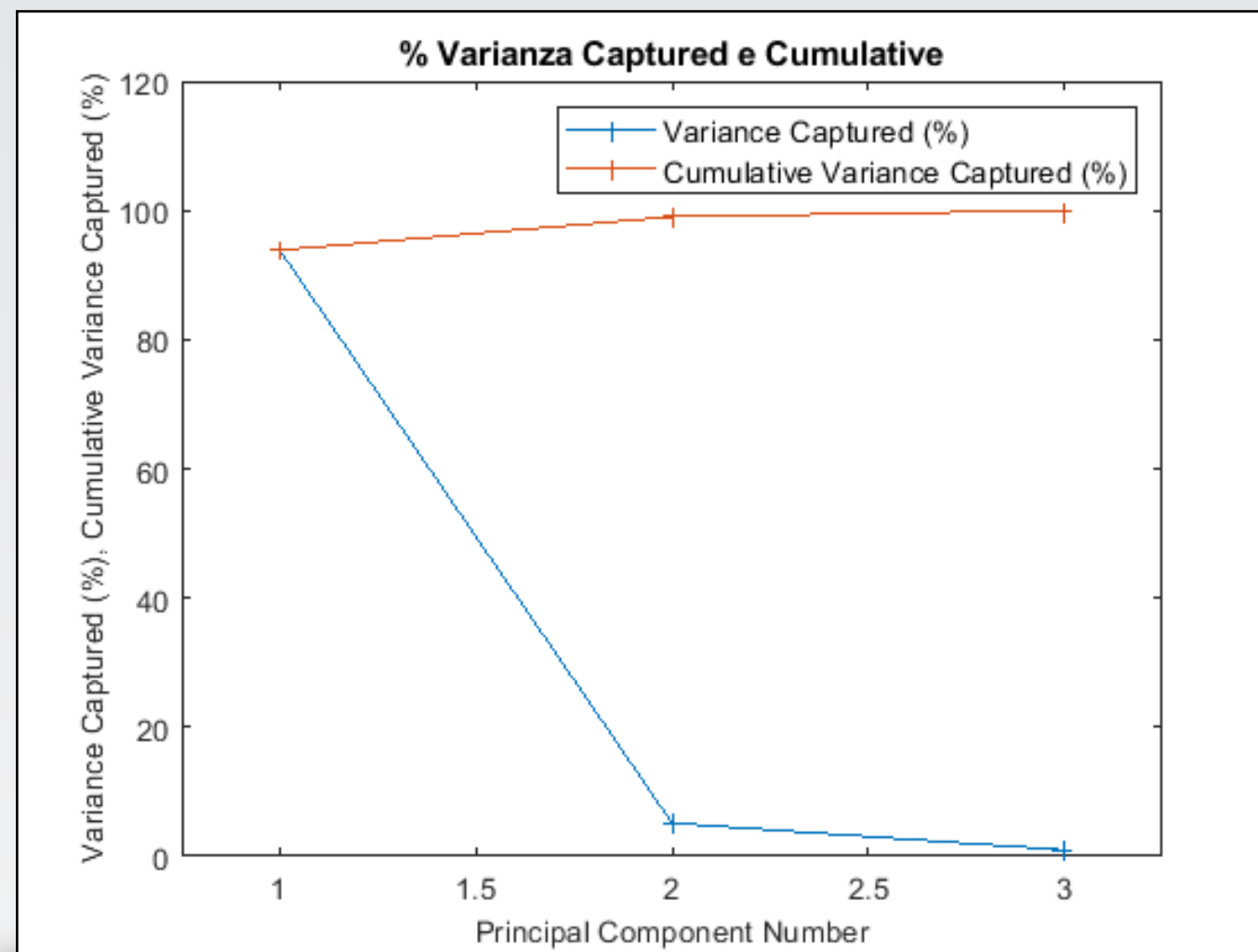
## **6. Previsione grasso**

### **A. Creazione modello PCA**



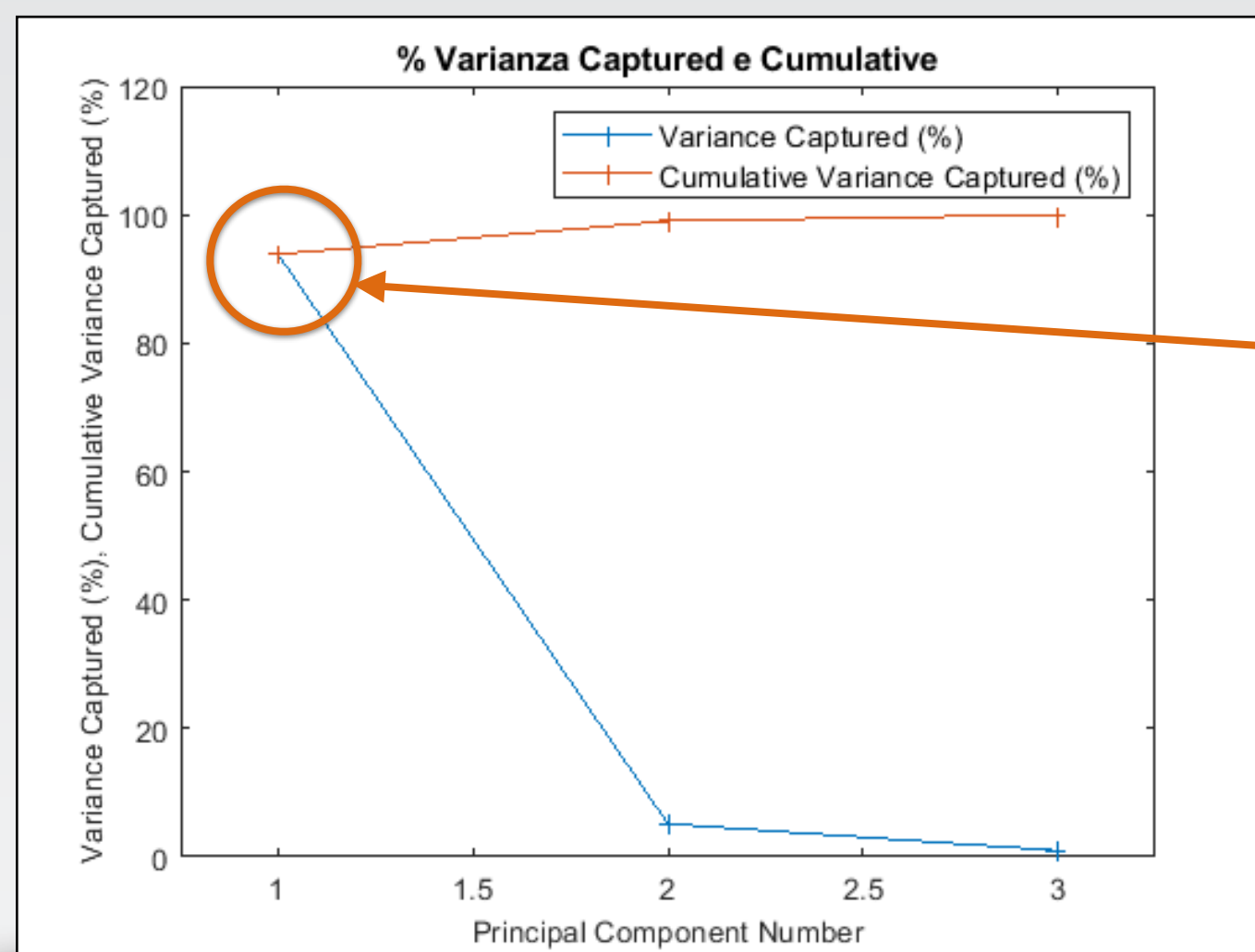
## 6. Previsione grasso

### A. Creazione modello PCA



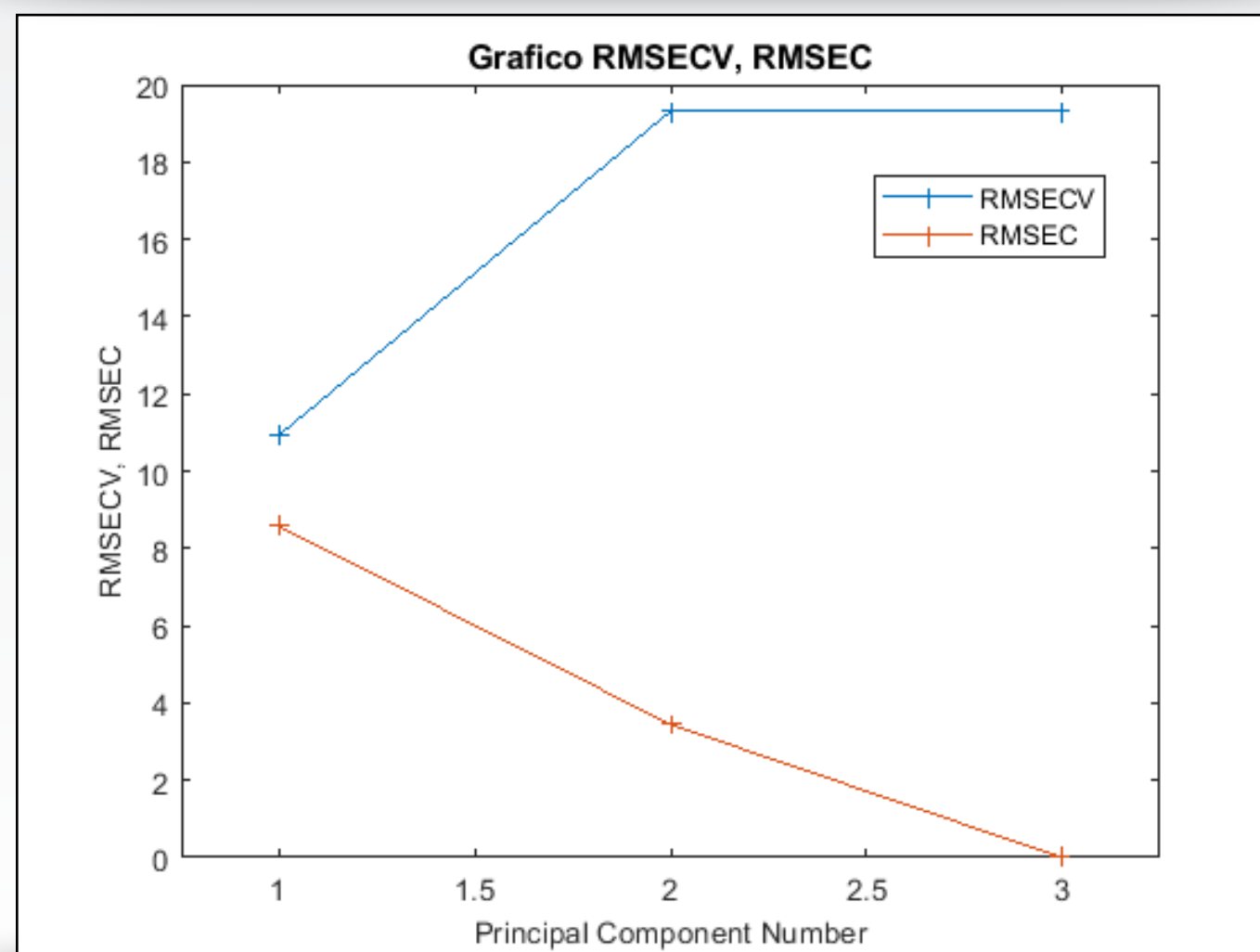
## 6. Previsione grasso

### A. Creazione modello PCA



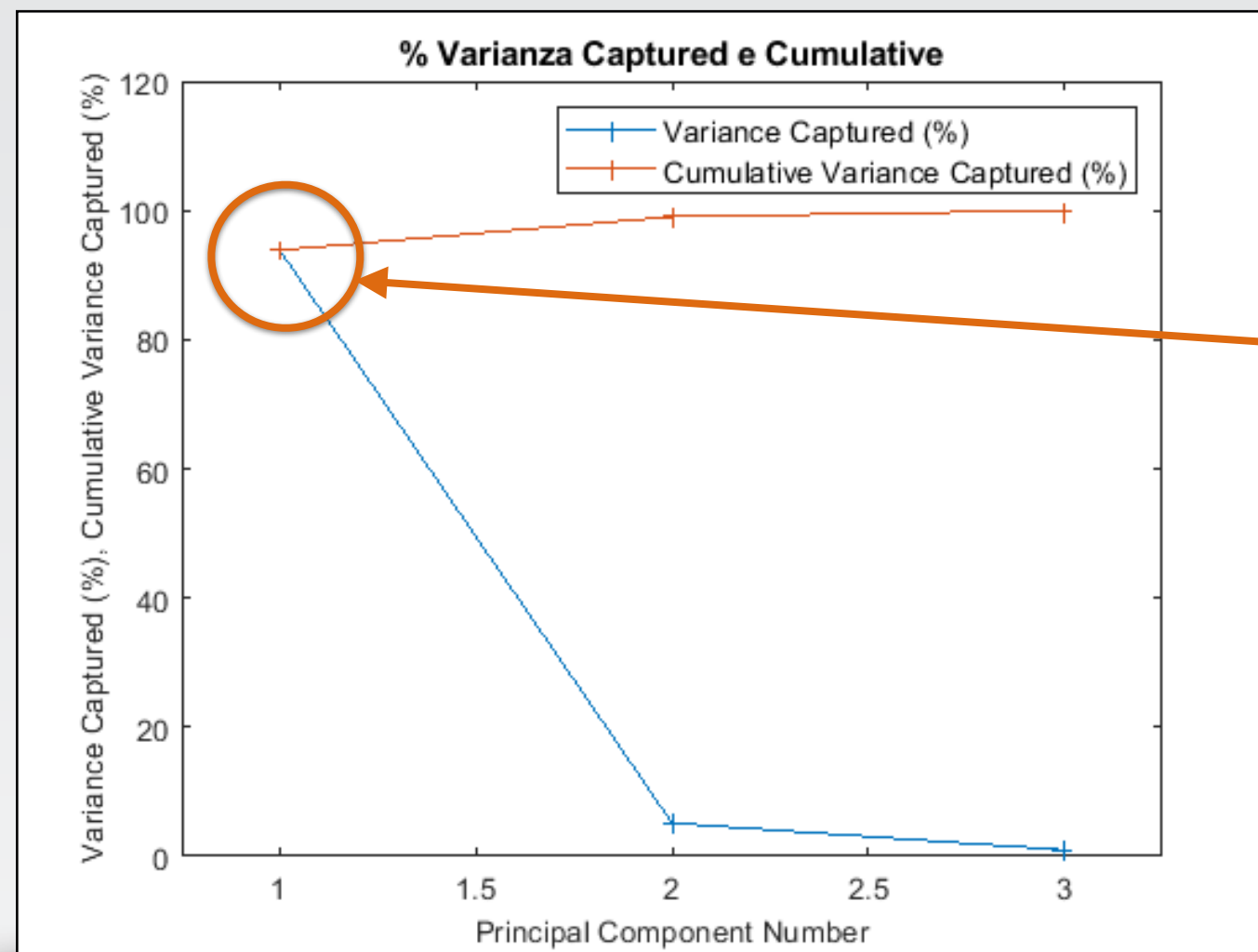
**Numero PC = 1**

PC1 cattura più del 95% di varianza



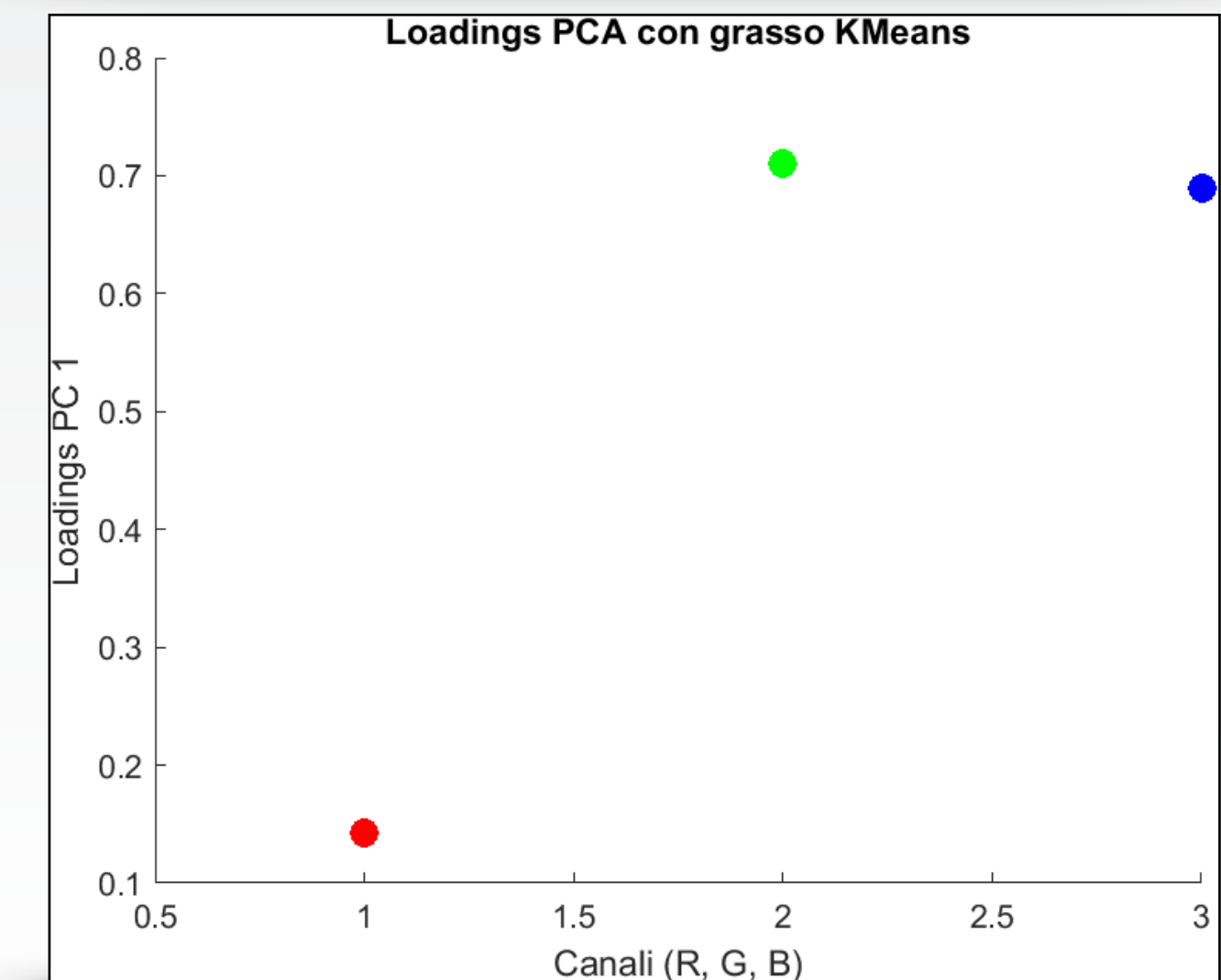
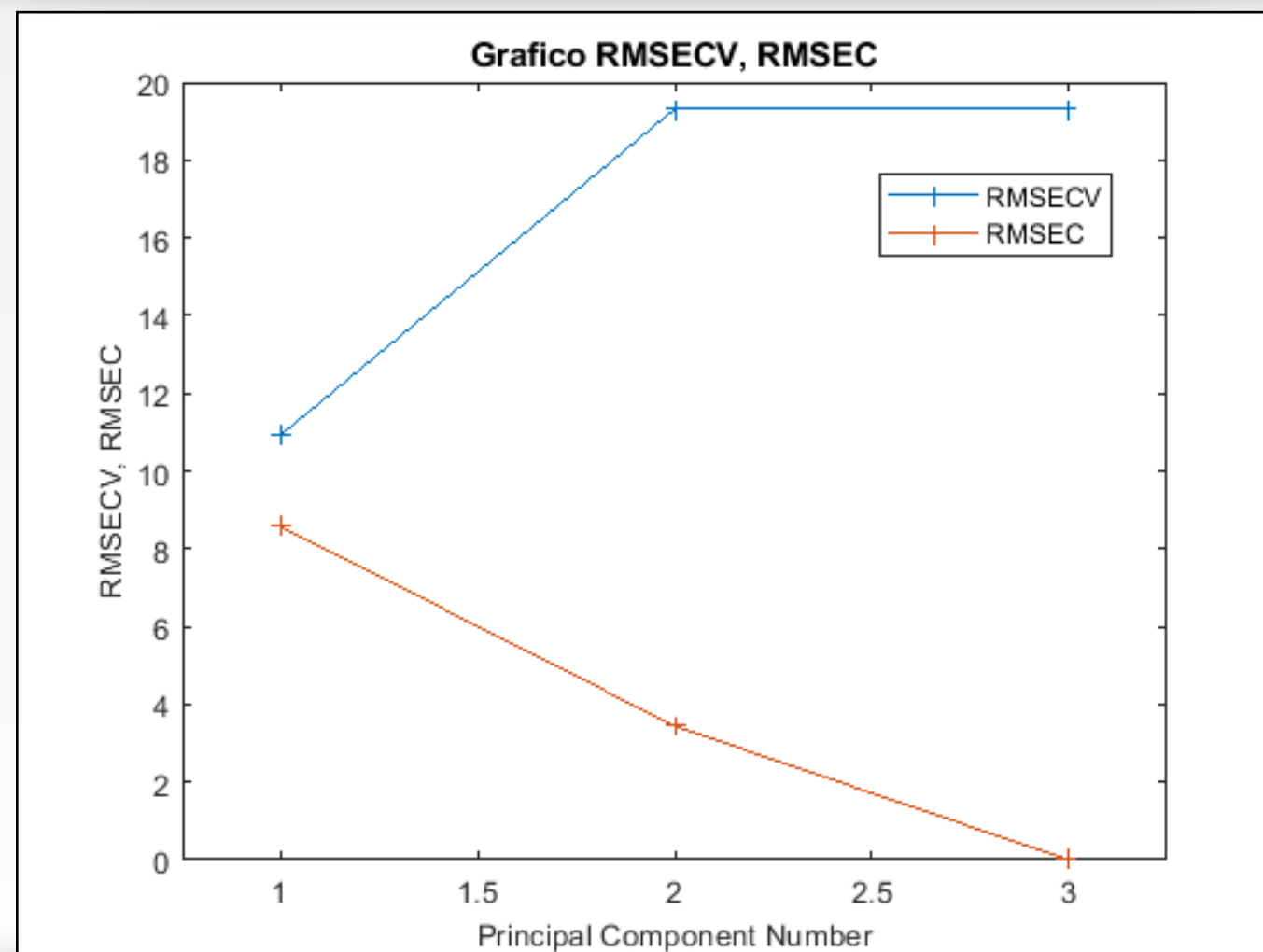
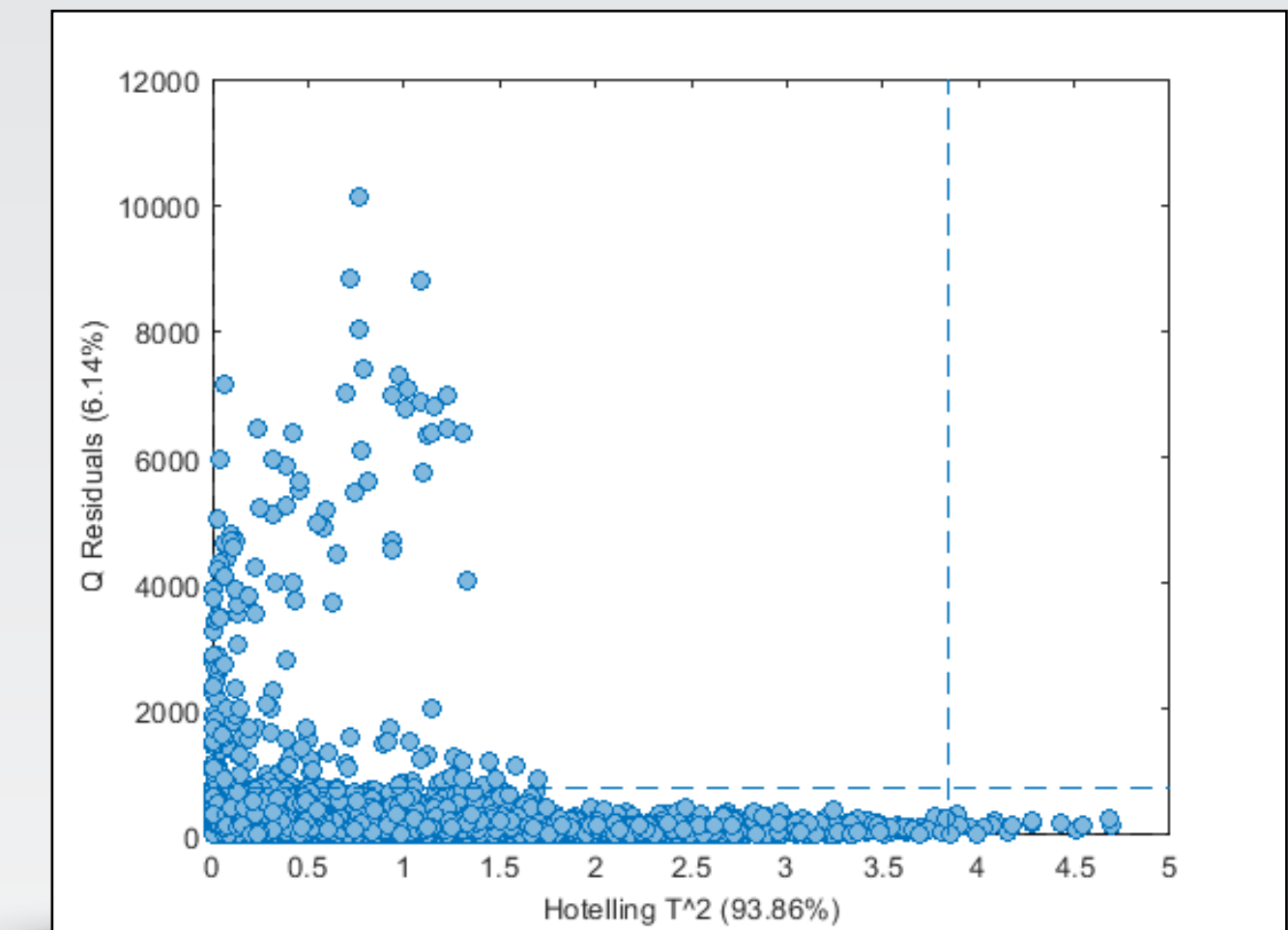
# 6. Previsione grasso

## A. Creazione modello PCA



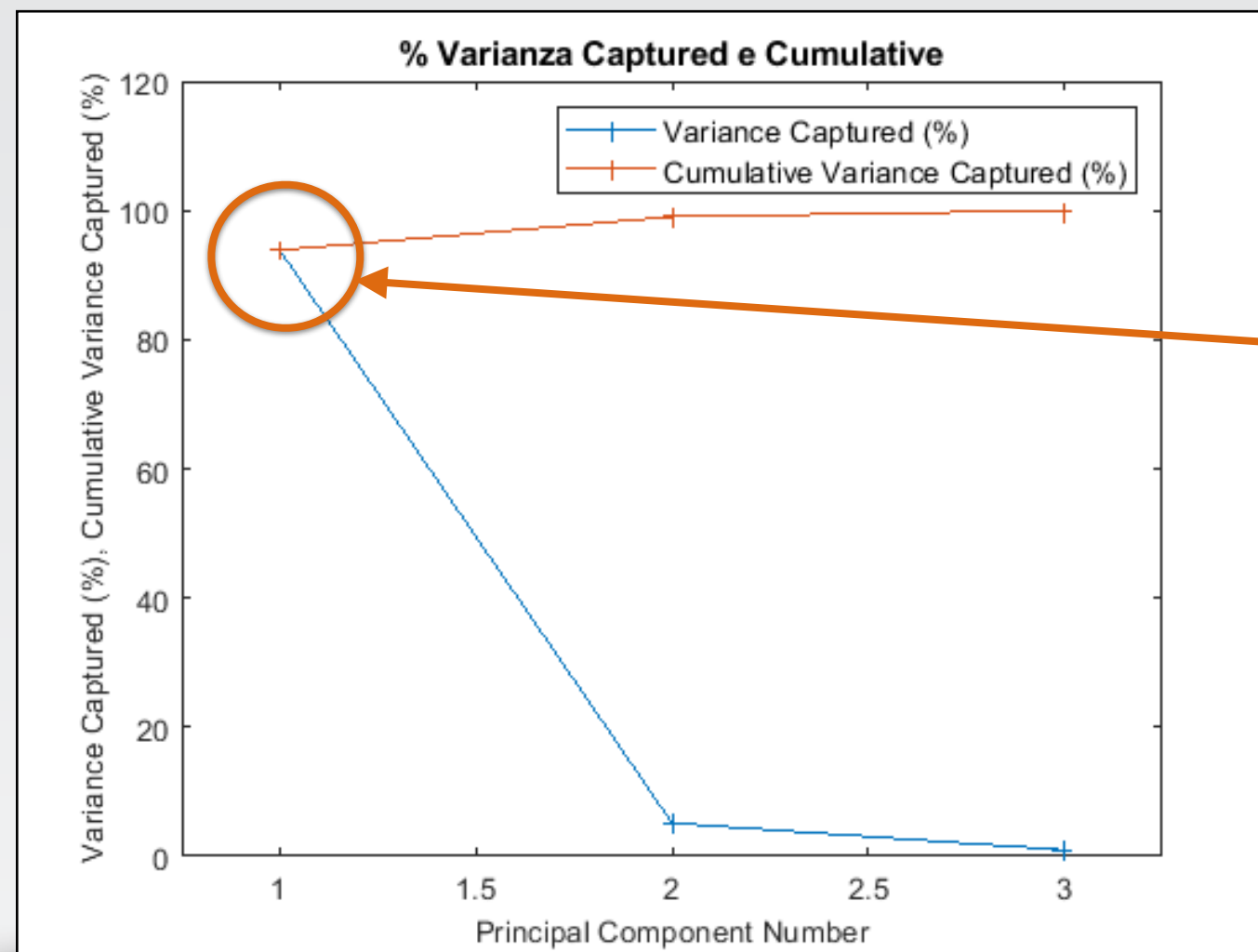
**Numero PC = 1**

PC1 cattura più del 95% di varianza



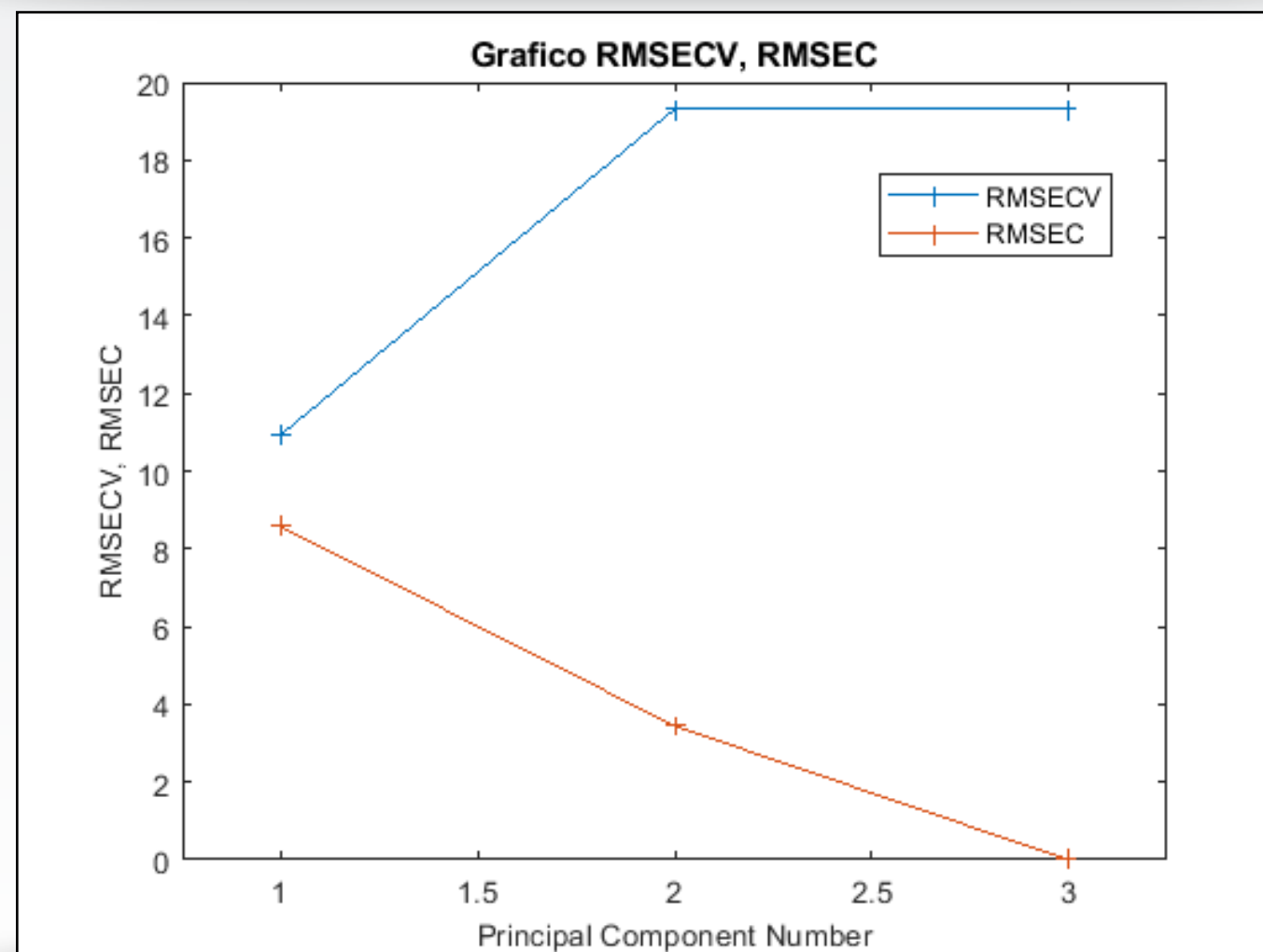
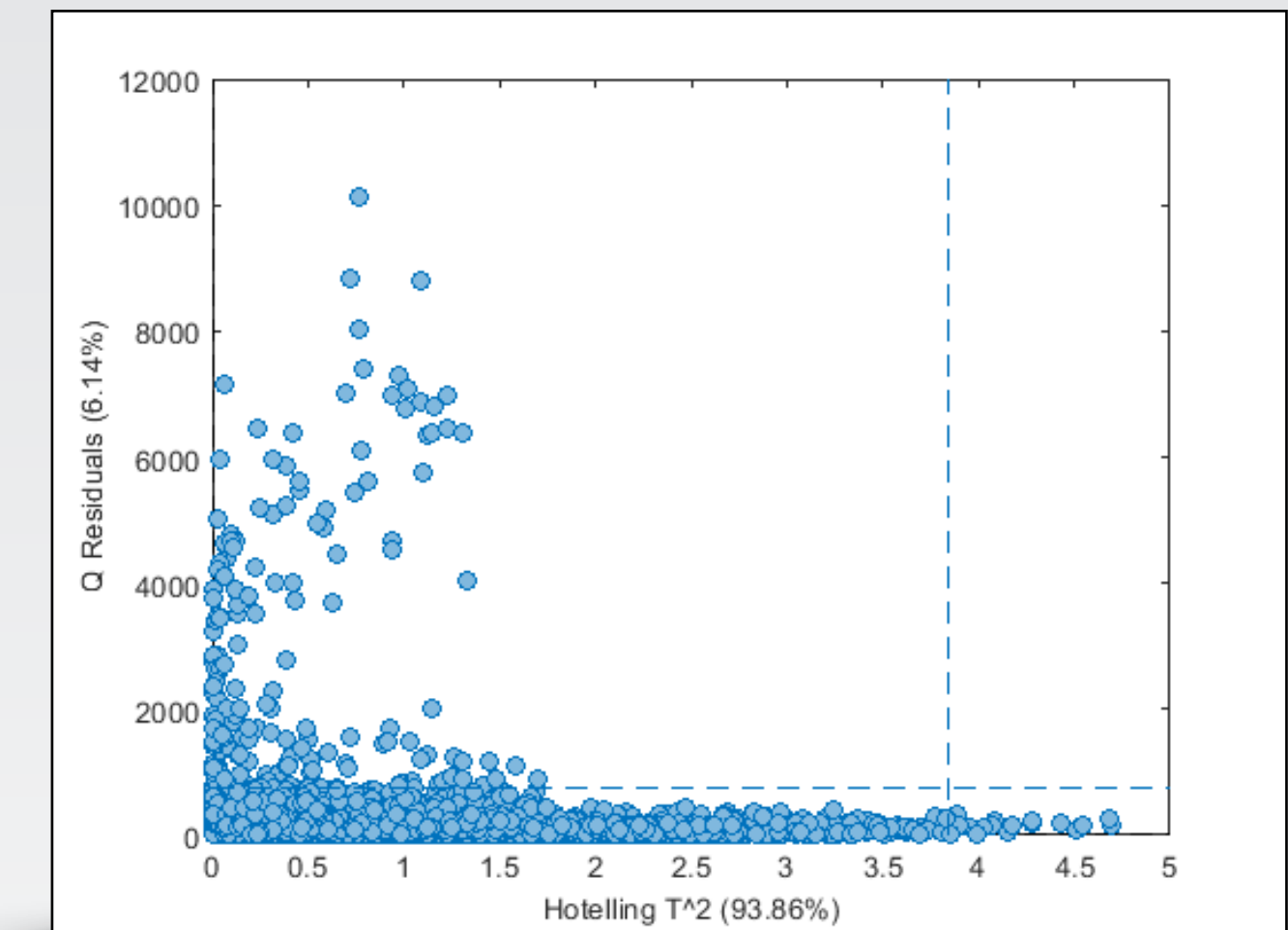
## 6. Previsione grasso

### A. Creazione modello PCA

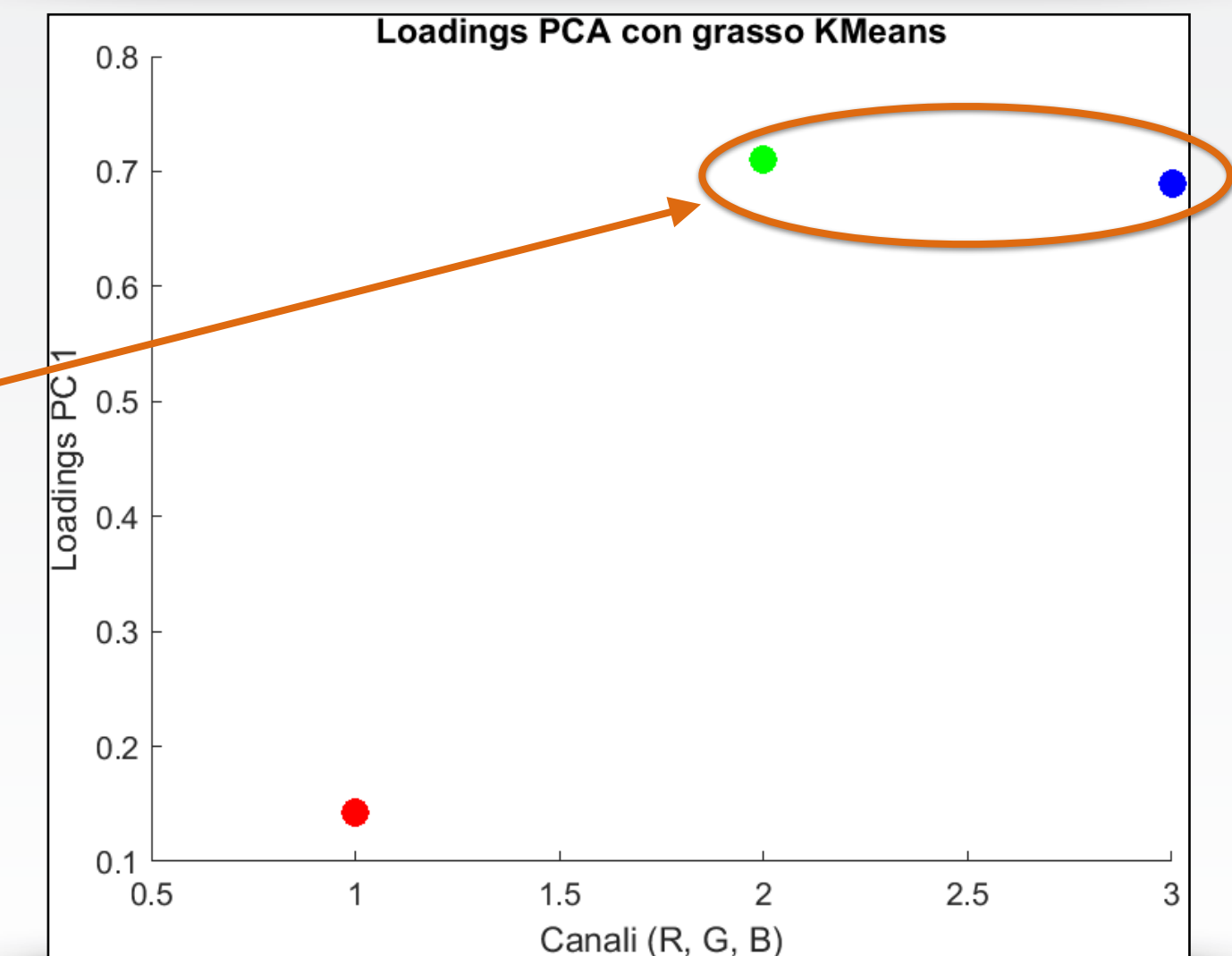


**Numero PC = 1**

PC1 cattura più del 95% di varianza



Loadings delle variabili G e B  
maggiori rispetto a R



## 6. Previsione grasso

## **6. Previsione grasso**

### **B. Proiezione steak2**

## 6. Previsione grasso

### B. Proiezione steak2

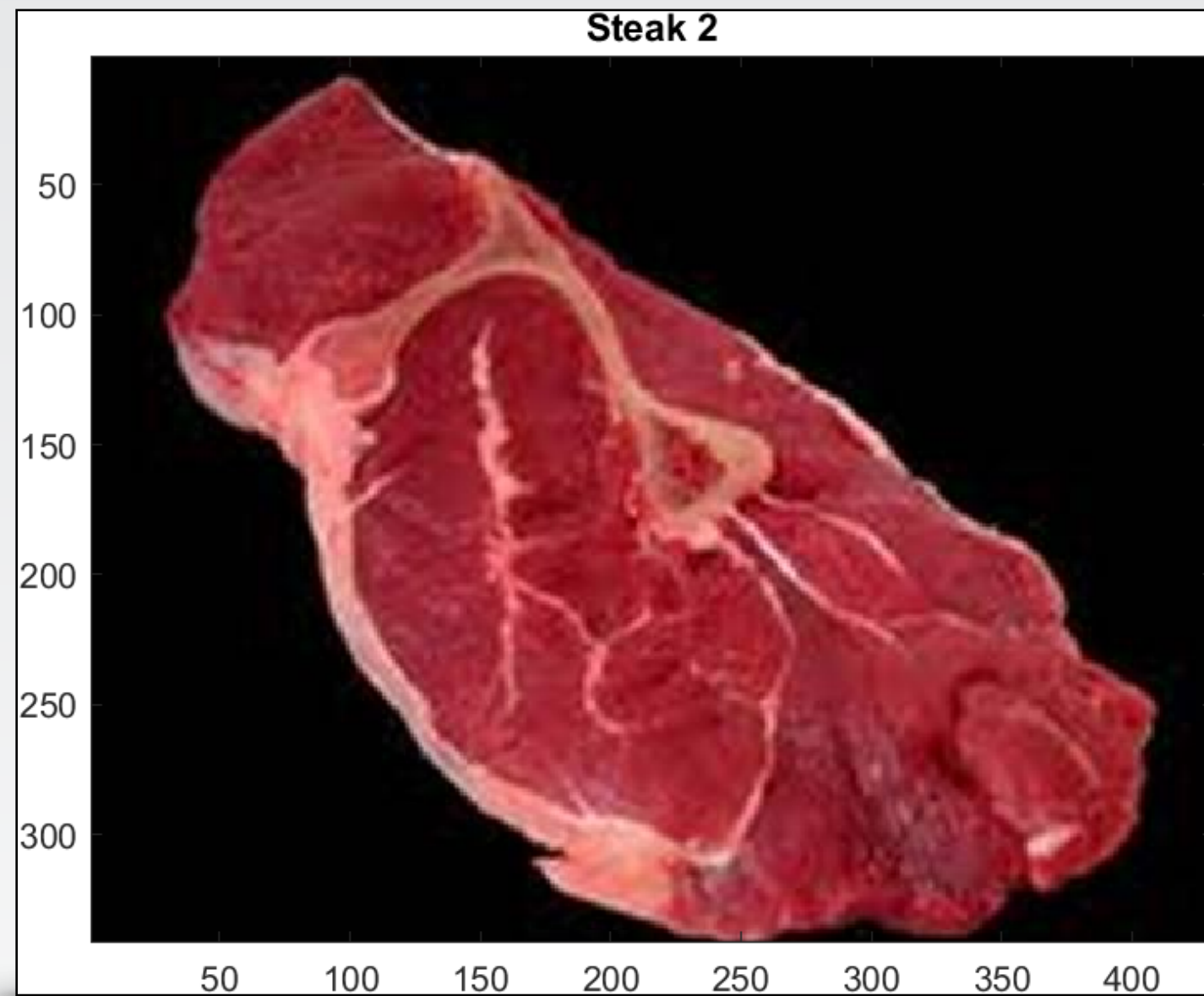


Immagine originale



## 6. Previsione grasso

### B. Proiezione steak2

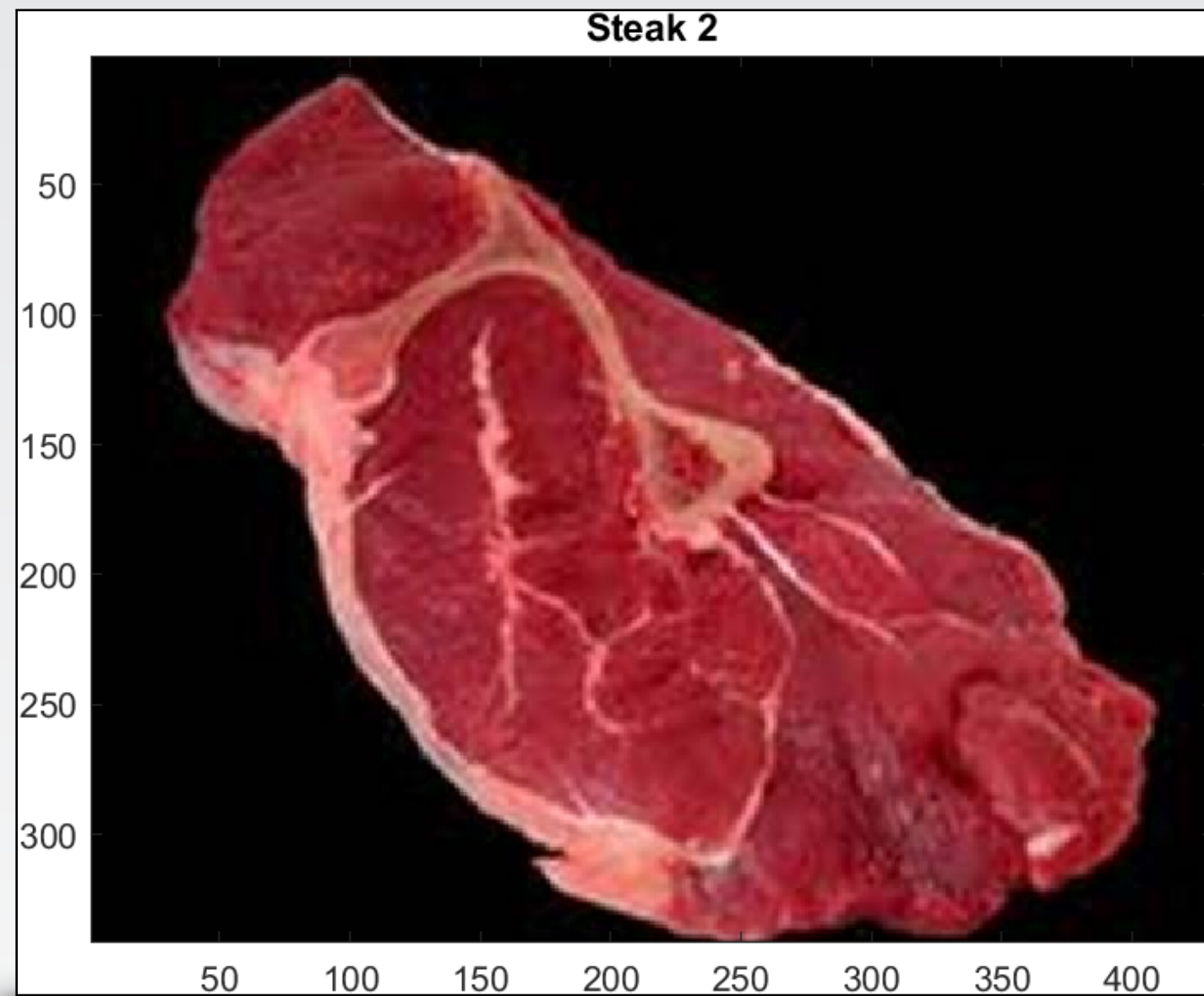


Immagine originale

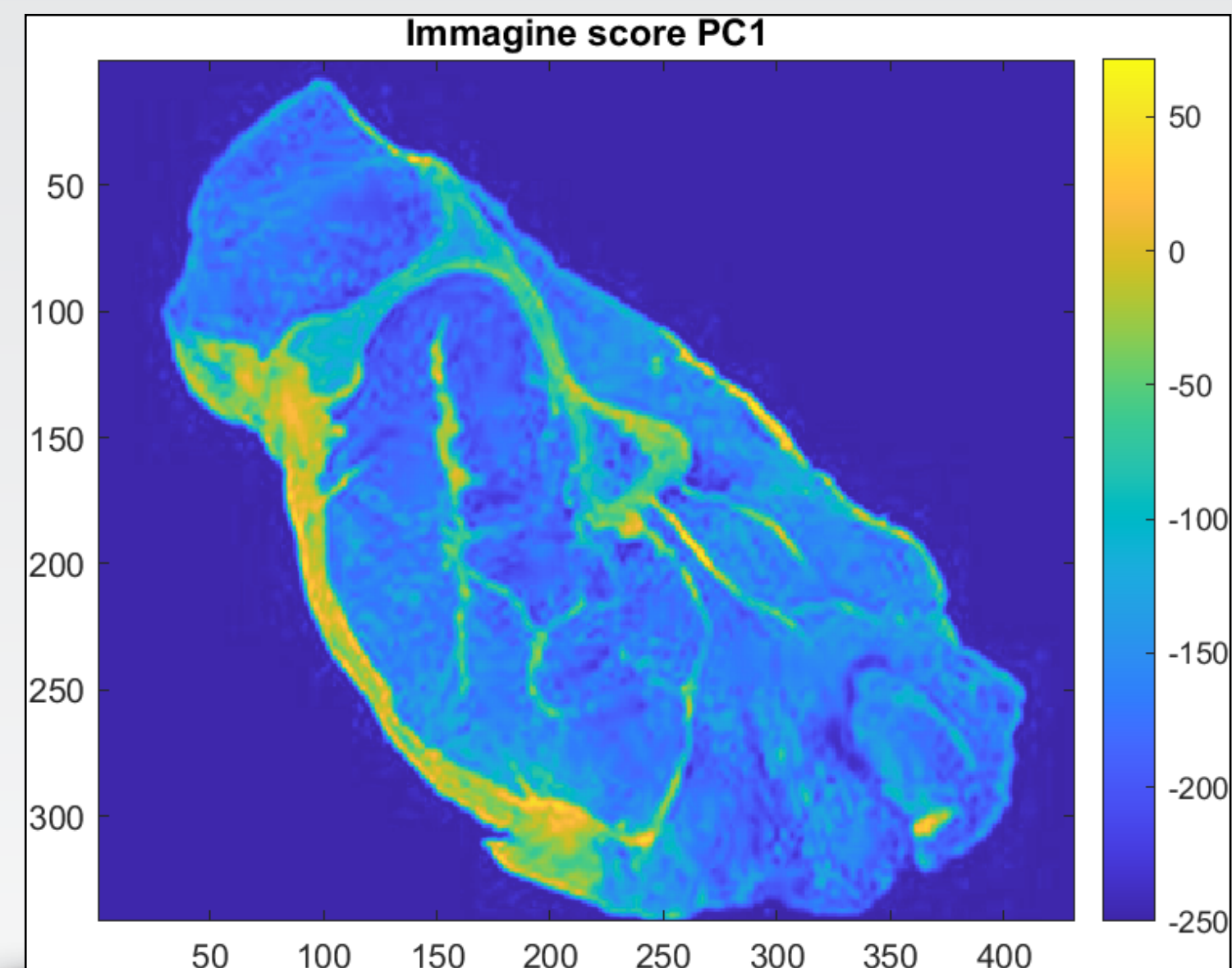


Immagine proiettata  
(Score PC1)



## 6. Previsione grasso

### B. Proiezione steak2

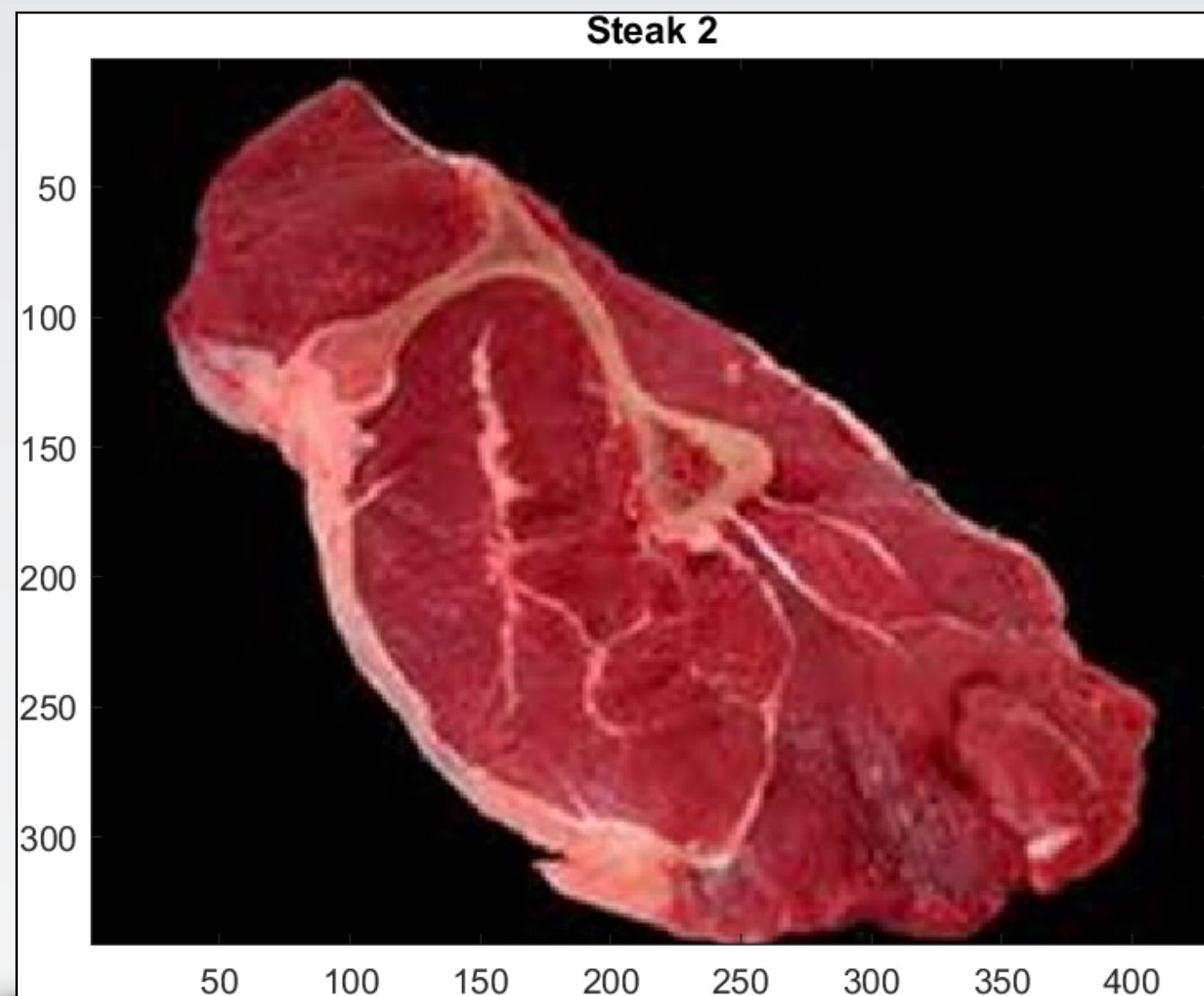


Immagine originale

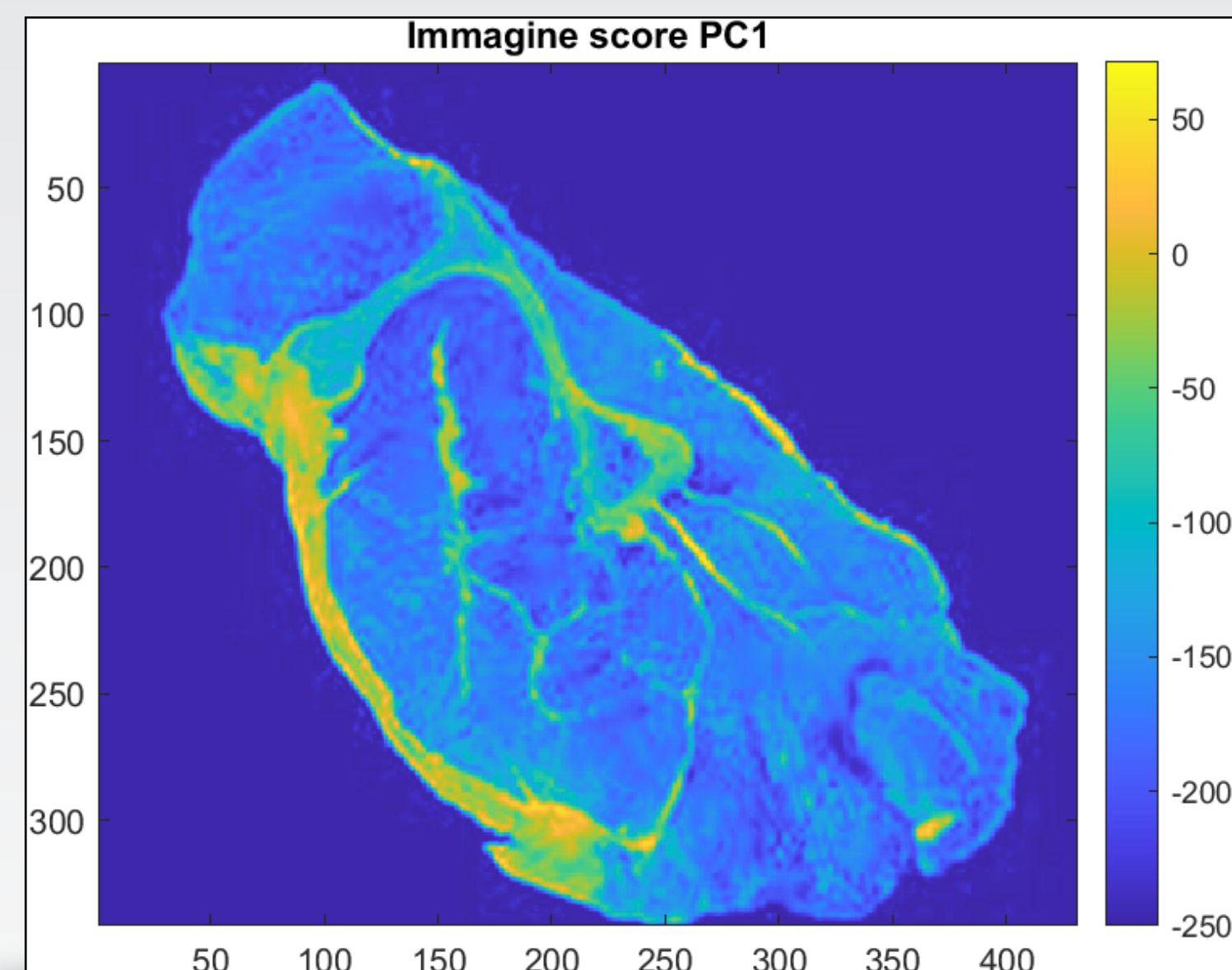
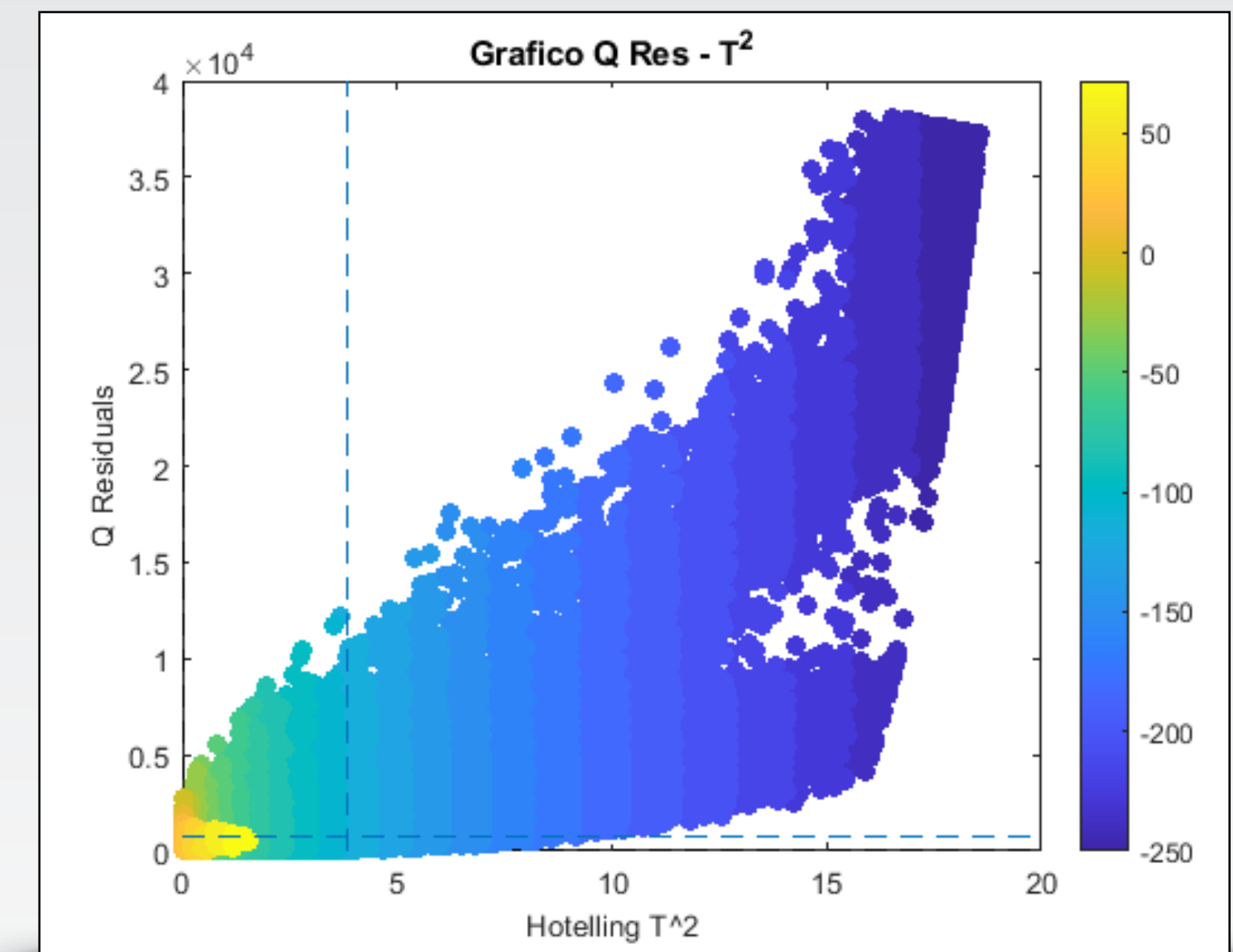


Immagine proiettata  
(Score PC1)



Q- $T^2$  plot colorato in  
base al valore degli scores

## 6. Previsione grasso

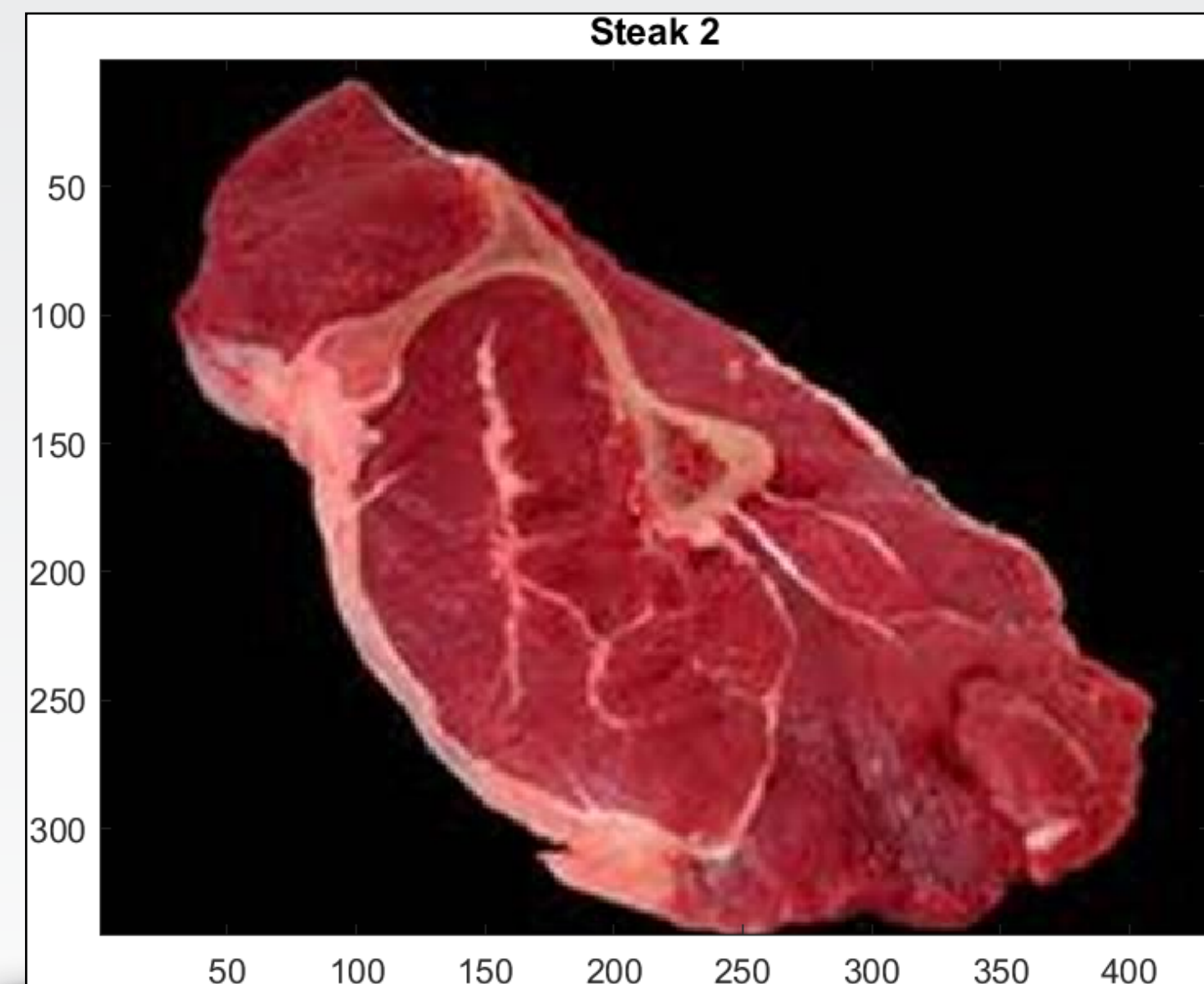
## 6. Previsione grasso

### C. Previsione

## 6. Previsione grasso

### C. Previsione

I punti di steak2 che hanno valori  $Q < Q_{\text{limit}}$  e  $T^2 < T^2_{\text{limit}}$  sono considerati come **grasso** e sono riportati di seguito in giallo



**Grazie per l'attenzione**

---