# Ανθοκυάνες

## Αριθμός δειγμάτων

Συνολικά συλλέχθηκαν 76 δείγματα, 56 από Λεμεσό και 20 από Πάφο. Όσο αφορά τη ποικιλία, 18 δείγματα Μαραθεύτικο, 28 δείγματα Μαύρο, 12 Cabernet Sauvignon και 18 Ofthalmo

Table 1 Αριθμός δειγμάτων ανά Επαρχία, Έτος, και ποικιλία

|  |  |  |  |
| --- | --- | --- | --- |
|  | Limassol | Paphos | Total |
| 2013 | | | |
| Maratheftiko | 4 | 2 | 6 |
| Mavro | 8 | 2 | 10 |
| Cabernet Sauvignon | 4 | — | 4 |
| OFTHALMO | 4 | 2 | 6 |
| Total | 20 | 6 | 26 |
| 2014 | | | |
| Maratheftiko | 4 | 2 | 6 |
| Mavro | 6 | 4 | 10 |
| Cabernet Sauvignon | 4 | — | 4 |
| OFTHALMO | 4 | 2 | 6 |
| Total | 18 | 8 | 26 |
| 2015 | | | |
| Maratheftiko | 4 | 2 | 6 |
| Mavro | 6 | 2 | 8 |
| Cabernet Sauvignon | 4 | — | 4 |
| OFTHALMO | 4 | 2 | 6 |
| Total | 18 | 6 | 24 |
| Overall | 56 | 20 | 76 |

## Επίπεδο ανθοκυάνων

**Μέση συγκέντρωση ανθοκυάνων ανά ποικιλία**

Το Μαραθεύτικο (Μ.Ο. = 11.9 , Τ.Α. = 0.7) έχει τη *μεγαλύτερη* μέση συγκέντρωση σε **Cyanidol** και το Cabernet Sauvignon (Μ.Ο. = 1.3, Τ.Α. = 0.5) τη μικρότερη

Το Μαραθεύτικο (Μ.Ο. = 6.0 , Τ.Α. = 0.4) έχει τη *μεγαλύτερη* μέση συγκέντρωση σε **Delphinidol** και το Cabernet Sauvignon (Μ.Ο. = 1.6, Τ.Α. = 0.2) τη μικρότερη

Το Μαραθεύτικο (Μ.Ο. = 47.7 , Τ.Α. = 2.9) και το OFTHALMO (Μ.Ο. = 50.3 , Τ.Α. = 1.0) έχουν τη *μεγαλύτερη* μέση συγκέντρωση σε **Peonidol** σε σχέση με τις άλλες ποικιλίες.

Το Mavro (Μ.Ο. = 3.7 , Τ.Α. = 0.7) έχει τη *μικρότερη* μέση συγκέντρωση σε **Petunidol** σε σχέση με τις άλλες ποικιλίες.

Το Mavro (Μ.Ο. = 73.5 , Τ.Α. = 7.3) έχει τη *μικρότερη* μέση συγκέντρωση σε **Malvidol** σε σχέση με τις άλλες ποικιλίες.

Κάτι το οποίο βλέπω εγώ εδώ, είναι οι πολύ μικροί αριθμοί της Τυπικής Απόκλισης . Αυτό σημαίνει ότι οι τιμές συγκέντρωσης των ανθοκυάνων είναι πολύ πολύ κοντά στο μέσο όρο. Κάτι που φαίνεται και από το γράφημα πιο κάτω όπου δεν παρατηρείται μεγάλη διασπορά των συγκεντρώσεων.

Table 2 Μέσο επίπεδο (Τυπική Απόκλιση) Ανθοκυάνες ανά ποικιλία

|  | Maratheftiko,  N = 181 | Mavro,  N = 281 | Cabernet Sauvignon,  N = 121 | OFTHALMO,  N = 181 | p-value2 |
| --- | --- | --- | --- | --- | --- |
| Cyanidol 3G | 11.9 (0.7) | 3.8 (0.5) | 1.3 (0.2) | 3.6 (0.2) | <0.001 |
| Delphinidol 3G | 6.0 (0.4) | 2.6 (0.6) | 1.6 (0.2) | 2.3 (0.2) | <0.001 |
| Peonidol 3G | 47.7 (2.9) | 4.9 (1.1) | 3.7 (0.1) | 50.3 (1.0) | <0.001 |
| Petunidol 3G | 13.7 (1.5) | 3.7 (0.7) | 12.7 (0.6) | 12.2 (0.1) | <0.001 |
| Malvidol 3G | 194.1 (3.4) | 73.5 (7.3) | 166.8 (5.2) | 153.6 (11.2) | <0.001 |
| *1Mean (SD)* | | | | | | |
| *2One-way ANOVA* | | | | | | |

To Figure 1 παρουσιάζει την κατανομή των συγκεντρώσεων των ανθοκυάνων ανά ποικιλία, όπου φαίνονται οι διαφορές στις συγκεντρώσεις ανάμεσα στις τέσσερις ποικιλίες

Figure 1 Distribution of Anthocyanin level by Variety

Chart, scatter chart

Description automatically generated

## Principal Component Analysis (PCA) (Ανάλυση Κύριων Συνιστωσών)

Πραγματοποιήθηκε ανάλυση κυρίων συνιστωσών (PCA) στα δεδομένα των επιπέδων των 5 ανθοκυάνων. Τα δεδομένα μας είναι επαρκή για διεξαγωγή της ανάλυσης PCA (ΚΜΟ: 0.641) και υπάρχουν επαρκείς συσχετίσεις μεταξύ των 5 μεταβλητών (Bartlett’s test of Sphericity: X2(10) = 555.23, p <0.001).

Εξάχθηκαν 2 κύριες συνιστώσες [Component 1: 49.7%, Component 2: 42.8%) όπου επεξηγούν συνολικά το 92.7% της συνολικής διασποράς του επιπέδου των 5 ανθοκυάνων.

Σύμφωνα με τις φορτίσεις των κυρίως συνιστωσών (Table 5), η 1η κύρια συνιστώσα συσχετίζεται κυρίως με Peonidol 3G, Petunidol 3G, και Malvidol 3G, ενώ η 2η κύρια συνιστώσα συσχετίζεται κυρίως με Cyanidol 3G και Delphinidol 3G

Table 5 Principal Component Factor Loadings

|  |  |  |
| --- | --- | --- |
|  | Component | |
|  | 1 | 2 |
| Cyanidol 3G | 0.258 | 0.961 |
| Delphinidol 3G | 0.243 | 0.959 |
| Peonidol 3G | 0.737 | 0.432 |
| Petunidol 3G | 0.975 | 0.133 |
| Malvidol 3G | 0.931 | 0.278 |

*Extraction Method: Principal Component Analysis.*

*Rotation Method: Varimax with Kaiser Normalization.*

Το Figure 2 παρουσιάζει τις φορτίσεις των 5 μεταβλητών στις δύο κύριες συνιστώσες, όπου φαίνεται και γραφικά ο διαχωρισμός:

- Υψηλές τιμές των τριών μεταβλητών Peonidol 3G, Petunidol 3G, και Malvidol 3G συσχετίζονται με υψηλές τιμές στην συνιστώσα 1 και μέσες τιμές (κοντά στο 0) στην κύρια συνιστώσα 2.

- Υψηλές τιμές των μεταβλητών Cyanidol 3G και Delphinidol 3G συσχετίζονται με υψηλές τιμές στην συνιστώσα 2 και μέσες τιμές (κοντά στο 0 με 0.5) στην κύρια συνιστώσα 1.

Figure 2 Loading Plot

Chart, line chart

Description automatically generated

Το PCA plot (Figure 3) παρουσιάζει το διαχωρισμό της ποικιλίας αναλόγως του επιπέδου των 2 κυρίως συνιστωσών.

Η 1η συνιστώσα διαχωρίζει πλήρως το Μαυρομάτικο από τις άλλες τρεις ποικιλίες, όπου το Μαυρομάτικο έχει χαμηλά επίπεδα Peonidol 3G, Petunidol 3G, και Malvidol 3.

Η 2η συνιστώσα έπειτα διαχωρίζει το Μαραθεύτικο [Υψηλό Cyanidol 3G και Delphinidol 3G] από τις άλλες δύο ποικιλίες.

Figure 3 Scatterplot between the 2 principal components of the PCA analysis

Chart, scatter chart

Description automatically generated

## Cluster Analysis

Πραγματοποιήθηκε ανάλυση συστάδων (Cluster Analysis) για τη διερεύνηση πιθανού διαχωρισμού των μετρήσεων, σε συστάδες. Πραγματοποιηθήκαν 2 μέθοδοι:

### Α) Kmeans

Η μέθοδος αυτή στοχεύει στο να διαχωρίσει τις παρατηρήσεις σε k ομάδες, όπου η κάθε παρατήρηση ανήκει σε μια ομάδα στην οποία βρίσκεται πιο κοντά σε απόσταση από το κέντρο της. Οι αποστάσεις των παρατηρήσεων μετριούνται με την ευκλείδεια απόσταση και η μέθοδος αυτή ελαχιστοποιεί τις αποστάσεις των παρατηρήσεων που ανήκουν στην ίδια ομάδα. Ο αριθμός των ομάδων στην k-means πρέπει να καθοριστεί από την αρχή.

Στη δημιουργία των ομάδων (clusters) συμμετέχουν οι 5 μεταβλητές των ανθοκυάνων Cyanidol 3G, Delphinidol 3G, Peonidol 3G, Petunidol 3G, και Malvidol 3G. Ο αριθμός των ομάδων καθορίστηκε στις 4, στη προσπάθεια να διερευνηθεί κατά πόσο η ομαδοποίηση με το cluster analysis θα συνάδει με τις ποικιλίες που προέρχονται τα δείγματά μας.

Αποτελέσματα

Η k-means έδειξε ομαδοποίηση η οποία συμφωνεί σε 100% με την ποικιλία από την οποία προέρχονται τα δείγματα (Table 6). Η ομάδα 1 περιέχει τα 28 δείγματα του Μαύρο, η ομάδα 2 περιέχει και τα 18 δείγματα του Μαραθεύτικο, η ομάδα 3 περιέχει και τα 12 δείγματα του Cabernet Sauvignon, και η ομάδα 4 περιέχει και 18 δείγματα του Ofthalmo.

Table 6 Cluster solution frequencies and Variety

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Custer | | | |  |
| Variety | 1 | 2 | 3 | 4 | Total |
| Maratheftiko | 0 | 18 | 0 | 0 | 18 |
| Mavro | 28 | 0 | 0 | 0 | 28 |
| Cabernet Sauvignon | 0 | 0 | 12 | 0 | 12 |
| OFTHALMO | 0 | 0 | 0 | 18 | 18 |
| Total | 28 | 18 | 12 | 18 | 76 |
| Average Silhouette Score | 0.91 | 0.85 | 0.88 | 0.66 |  |

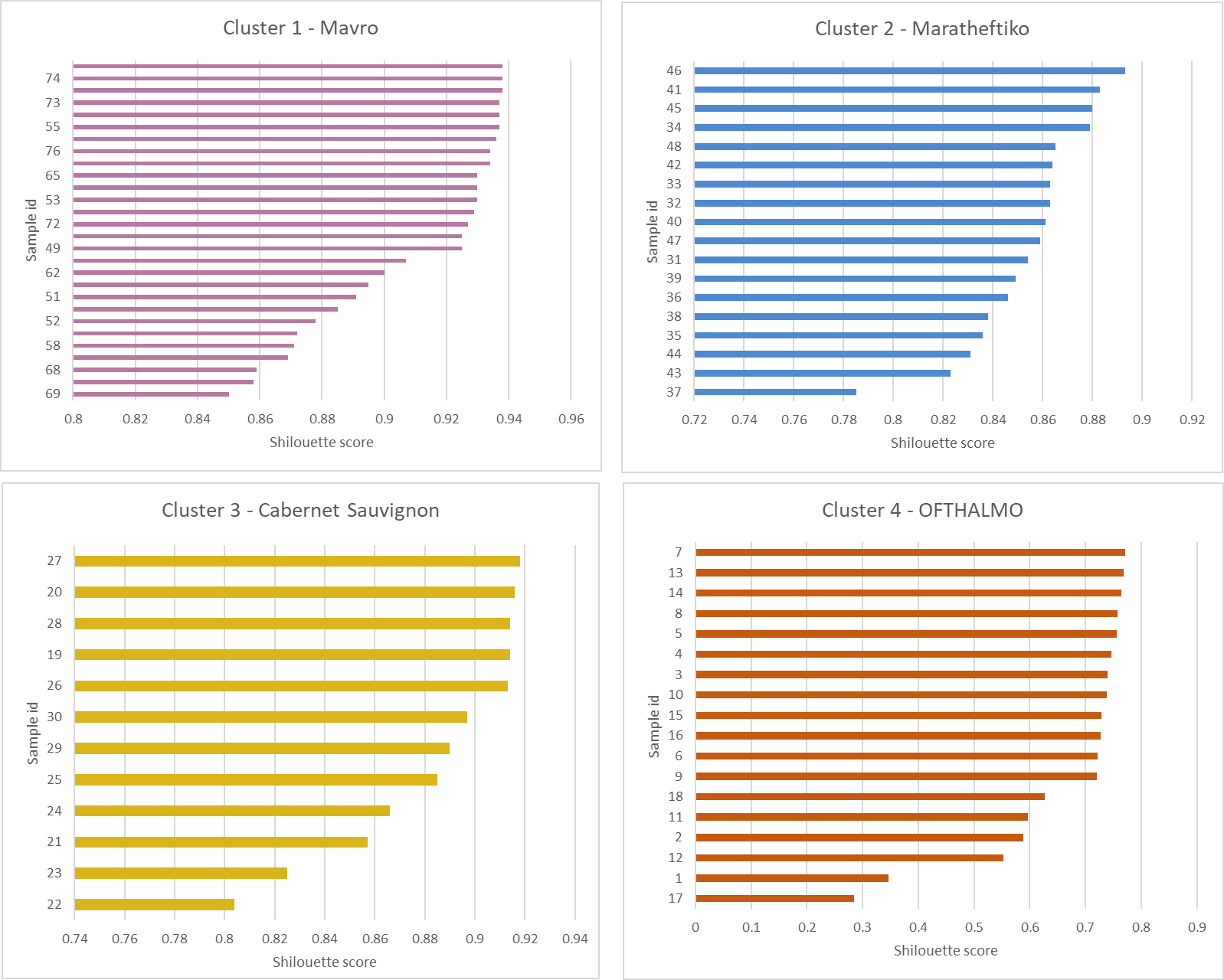
Συνοχή και διαχωρισμός ομάδων

Η συνοχή (cohesion) και διαχωρισμός (separation) των ομάδων, μελετήθηκε με το δείκτη Silhouette. Ο δείκτης silhouette υπολογίζεται για κάθε παρατήρηση του δείγματος και δείχνει κατά πόσο η παρατήρηση είναι παρόμοια με τις υπόλοιπες παρατηρήσεις της ομάδας (Συνοχή) σε σχέση με τις άλλες παρατηρήσεις (separation).

Ο δείκτης silhouette κυμαίνεται από −1 έως +1, όπου ψηλές τιμές υποδεικνύουν ότι η παρατήρηση είναι πολύ παρόμοια με τις παρατηρήσεις της ομάδας του και πολύ ανόμοια με τις παρατηρήσεις των άλλων ομάδων. Ενώ, αν ο δείκτης είναι χαμηλός ή ακόμα αρνητικός, τότε σημαίνει ότι η παρατήρηση δεν ταιριάζει και πολύ με την ομάδα, αλλά ούτε και διαχωρίζεται με τις υπόλοιπες παρατηρήσεις των άλλων ομάδων.

Το Figure 4 παρουσιάζει τα silhouette scores όλων των παρατηρήσεων σε κάθε ομάδα. Παρατηρούνται υψηλά score σε όλες τις ομάδες. Τα μέσο silhouette score σε κάθε ομάδα είναι: Cluster 1- 0.91, Cluster 2: 0.85, Cluster 3: 0.88, Cluster : 0.66

Figure 4 Silhouette scores for cluster cohesion and separation



### Α) Hierarchical Clustering

Αυτή είναι μέθοδος «από κάτω προς τα επάνω» ("bottom-up") όπου ο αλγόριθμος ξεκινά θεωρώντας τη κάθε παρατήρηση να αποτελεί από μόνη της ένα cluster. Στη συνέχεια στη συνέχεια ζευγάρια παρατηρήσεων ομαδοποιούνται σε περισσότερες ομάδες. Η μέθοδος που ενώνονται οι παρατηρήσεις είναι με βάση ένα πίνακα διαφοροποιήσεων (dissimilarity matrix) από τις υπόλοιπες παρατηρήσεις. Επαναλαμβάνεται η διαδικασία αυτή, μέχρι όλες οι παρατηρήσεις να αποτελούν ένα και μόνο cluster.

Ο αριθμός των ομάδων δεν καθορίζεται από την αρχή, αλλά ο ερευνητής καλείται να επιλέξει τον αριθμό των ομάδων μετά από διερεύνηση της ιεράρχησης.

Το πιο κάτω δενδρόγραμμα παρουσιάζει τις πιθανές ομαδοποιήσεις των ανθοκυάνων.

Παρατηρείται όταν επιλέξουμε 4 clusters τότε μέσα στο κάθε cluster έχουμε ακριβώς την κάθε ποικιλία. ΑΝ επιλέξουμε 3 cluster τότε το Μαραθεύτικο και το Cabernet Sauvignon πέφτουν στο ίδιο cluster και οι άλλες 2 ποικιλίες σε άλλα 2 cluster.

Figure 5 Dendrogram – Hierarchical clustering – Διαχωρισμός 4, 3 και 2 ομάδων. Ονοματισμός των παρατηρήσεων με βάση τη ποικιλία

Chart, schematic

Description automatically generated

Figure 6 Dendrogram – Hierarchical clustering – Διαχωρισμός 4, 3 και 2 ομάδων. Ονοματισμός των παρατηρήσεων με βάση την τη ποικιλία

A picture containing text, device

Description automatically generated

### Κατανομή ομάδων-cluster και χαρακτηριστικών

Οι πιο κάτω πίνακες παρουσιάζουν την κατανομή της ποικιλίας και της Περιοχής, στη λύση των 4άρων και 3ων cluster έτσι όπως εξάχθηκαν από το Hierarchical Clustering

Πίνακας 1 Κατανομή ποικιλίας στη λύση των 4αρων cluster (Hierarchical Clustering)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Cluster Groups (k=4) | | | |  |
| Variety | 1 | 2 | 3 | 4 | Total |
| Maratheftiko | 18 | 0 | 0 | 0 | 18 |
| Mavro | 0 | 28 | 0 | 0 | 28 |
| Cabernet Sauvignon | 0 | 0 | 12 | 0 | 12 |
| OFTHALMO | 0 | 0 | 0 | 18 | 18 |
| Total | 18 | 28 | 12 | 18 | 76 |

Πίνακας 2 Κατανομή ποικιλίας στη λύση των 3ων cluster (Hierarchical Clustering)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Cluster Groups (k=3) | | |  |
| Variety | 1 | 2 | 3 | Total |
| Maratheftiko | 18 | 0 | 0 | 18 |
| Mavro | 0 | 28 | 0 | 28 |
| Cabernet Sauvignon | 0 | 0 | 12 | 12 |
| OFTHALMO | 18 | 0 | 0 | 18 |
| Total | 36 | 28 | 12 | 76 |

Πίνακας 3 Κατανομή Περιοχής στη λύση των 4αρων cluster (Hierarchical Clustering)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Cluster Groups (k=4) | | | |  |
| Region | 1 | 2 | 3 | 4 | Total |
| VP | 6 | 6 | 0 | 6 | 18 |
| LAONA | 6 | 6 | 0 | 6 | 18 |
| VASA | 0 | 6 | 6 | 0 | 12 |
| OMODOS | 6 | 6 | 6 | 0 | 18 |
| MALLIA | 0 | 4 | 0 | 6 | 10 |
| Total | 18 | 28 | 12 | 18 | 76 |

Πίνακας 4 Κατανομή Περιοχής στη λύση των 3ων cluster (Hierarchical Clustering)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Cluster Groups (k=3) | | |  |
| Region | 1 | 2 | 3 | Total |
| VP | 12 | 6 | 0 | 18 |
| LAONA | 12 | 6 | 0 | 18 |
| VASA | 0 | 6 | 6 | 12 |
| OMODOS | 6 | 6 | 6 | 18 |
| MALLIA | 6 | 4 | 0 | 10 |
| Total | 36 | 28 | 12 | 76 |