



Στατιστική επεξεργασία συνθηκών τυχερών παιχνιδιών – Ανάπτυξη εφαρμογής

Πτυχιακή εργασία του φοιτητή
Σκαρλάτου Ηλία 07 / 3186
Επιβλέπων καθηγητής
κ. Σφέτσος Παναγιώτης



ΠΡΟΛΟΓΟΣ (1/2)

Αντικείμενο της εργασίας:

- Η στατιστική επεξεργασία και μελέτη των συνθηκών τυχερών παιχνιδιών
- Εύρεση συνδυασμών από γεννήτρια τυχαίων αριθμών χρησιμοποιώντας στατιστικούς όρους ή βασικές στήλες



ΠΡΟΛΟΓΟΣ (2/2)

Σκοπός της εργασίας:

- Δημιουργία ενός εργαλείου για την στατιστική μελέτη προηγούμενων κληρώσεων.
- Η δυνατότητα στον χρήστη να δημιουργήσει βασικές στήλες και ομάδες βάσει των οποίων θα δημιουργηθούν οι συνδυασμοί.



Συχνότητα εμφάνισης και καθυστέρησης

Η στατιστική ανάλυση χωρίζεται σε δύο
βασικούς κλάδους:

- Συχνότητα εμφάνισης του στατιστικού όρου στους συνδυασμούς των κληρώσεων,
- Συχνότητα καθυστέρησης στην εμφάνιση του στατιστικού όρου στους συνδυασμούς των κληρώσεων.



Μέσος όρος μεταβλητής

Ένα από τα βασικότερα μεγέθη που χρησιμοποιούνται στην στατιστική ανάλυση ενός πληθυσμού είναι ο μέσος όρος. Για τον υπολογισμό του χρησιμοποιείται ο τύπος:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n_1 + n_2 + n_3 + \dots + n_k} = \frac{1}{n} \sum_{i=1}^n x_i$$



Τυπική απόκλιση μεταβλητής

Εξίσου σημαντικό στατιστικό μέγεθος που αναφέρεται στην διασπορά ή μεταβλητότητα ενός πληθυσμού είναι η τυπική απόκλιση. Για τον υπολογισμό της χρησιμοποιείται ο παρακάτω τύπος:

$$s^2 = \frac{1}{n} \sum_{i=1}^k (t_i - \bar{x})^2 = \frac{1}{k} \left\{ \sum_{i=1}^k t_i^2 - \frac{(\sum_{i=1}^k t_i)^2}{k} \right\}$$

$$s = \sqrt{s^2}$$



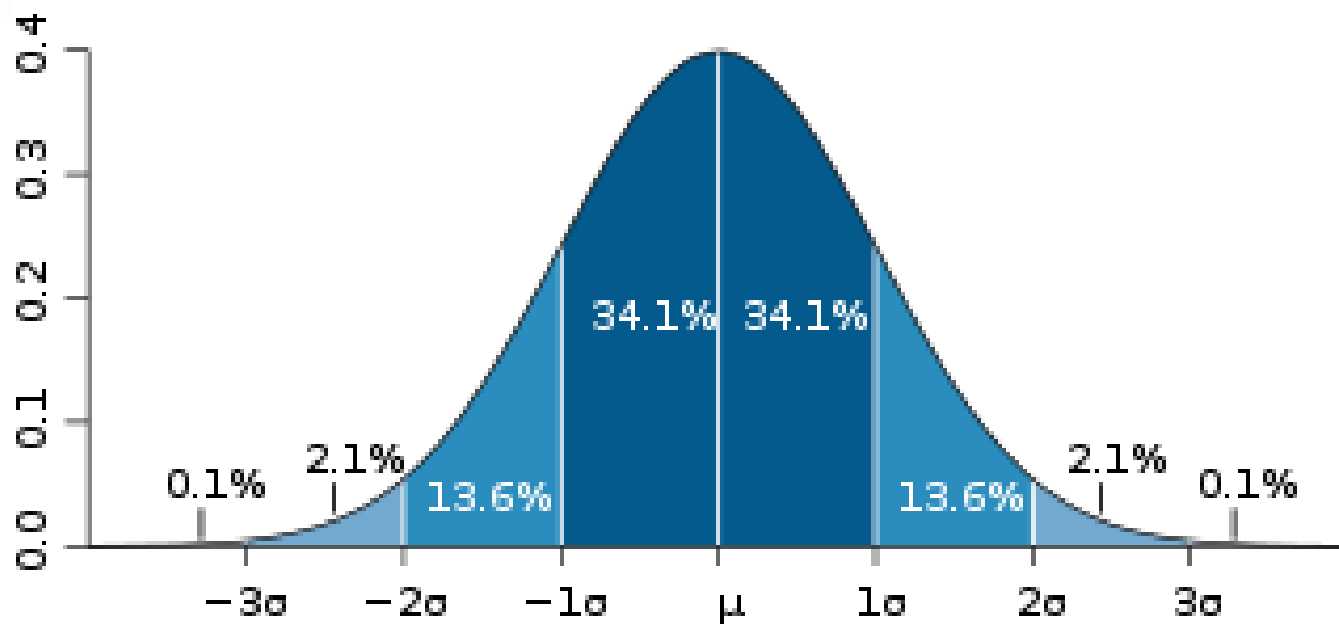
Κώδικας σε JAVA για τον υπολογισμό μέσου όρου και τυπικής απόκλισης

```
public double[] getKanonikotitaOfNumber(int[] tableOfNumbers)
{
    int[][] kanonikotitaResult;
    int max = tableOfNumbers[0];
    int min = tableOfNumbers[0];
    int Sx2v = 0;
    int Sxv = 0;
    int Sv = 0;
    double MOx = 0.0;
    double S2 = 0.0;
    double[] results = new double[2];
    try{
        for (int i = 0; i < tableOfNumbers.length; i++)
        {
            if (tableOfNumbers[i] > max)
                max = tableOfNumbers[i];
            else if (tableOfNumbers[i] < min)
                min = tableOfNumbers[i];
        }
        kanonikotitaResult = new int[max - min + 1][2];
        kanonikotitaResult[0][0] = min;
        kanonikotitaResult[0][1] = 0;
        for (int j = 1; j < kanonikotitaResult.length; j++)
        {
            kanonikotitaResult[j][0] = kanonikotitaResult[j - 1][0] + 1;
            kanonikotitaResult[j][1] = 0;
        }
    }
```

```
        for (int k = 1; k <= kanonikotitaResult.length; k++)
        {
            for (int i = 0; i < kanonikotitaResult.length; i++)
            {
                if (tableOfNumbers[k - 1] == kanonikotitaResult[i][0])
                {
                    kanonikotitaResult[i][1] += 1;
                    break;
                }
            }
        }
        for (int k = 0; k < kanonikotitaResult.length; k++)
        {
            Sv += kanonikotitaResult[k][1];
            Sxv += kanonikotitaResult[k][0] *
kanonikotitaResult[k][1];
            Sx2v +=
(int)java.lang.Math.pow(kanonikotitaResult[k][0], 2) *
kanonikotitaResult[k][1];
        }
        if (Sv != 0)
        {
            MOx = (double)Sxv / (double)Sv;
            S2 = (Sx2v - (java.lang.Math.pow(Sxv, 2) / Sv)) / Sv;
        }
        double mesosOros = java.lang.Math.floor(MOx * 10 +
0.5d)/10;
        double typikiApoklisi
= java.lang.Math.floor(java.lang.Math.sqrt(S2) * 100 + 0.5d)/100;
        results[0] = mesosOros; results[1] = typikiApoklisi;
    }
    catch(Exception e) {
        JOptionPane.showMessageDialog(this,String.valueOf(e.toString()))
, "Προειδοποίηση!!", JOptionPane.ERROR_MESSAGE);
    } return results;}
}
```



Διάγραμμα συχνοτήτων κανονικής κατανομής





Cold – Hot – Neutral

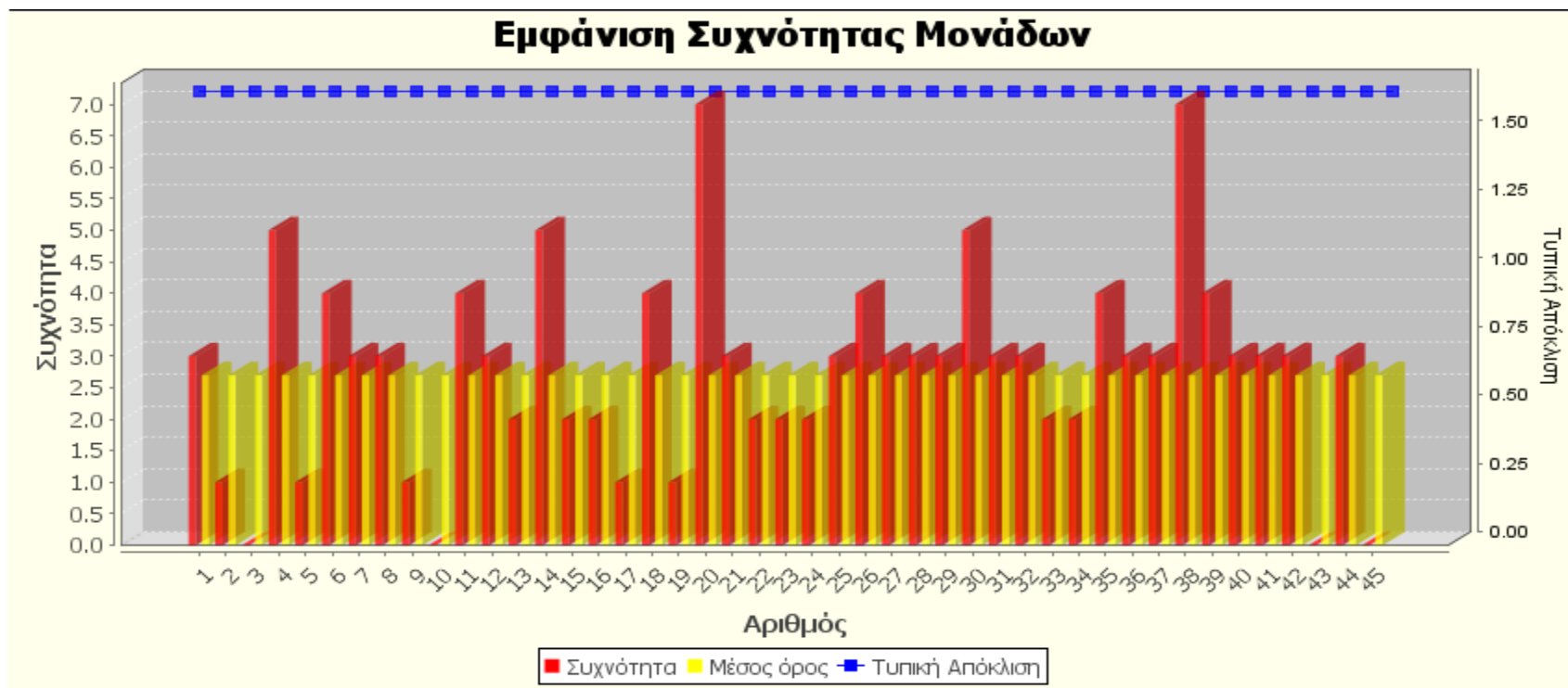
Αν η συχνότητα εμφάνισης κάποιας τιμής του στατιστικού όρου ανήκει στο διάστημα:

- $(\bar{x} - 3s, \bar{x} - 2s)$, τότε η τιμή χαρακτηρίζεται ως “Cold”,
- $(\bar{x} - 2s, \bar{x} - s)$, τότε η τιμή χαρακτηρίζεται ως “Cool”,
- $(\bar{x} - s, \bar{x} + s)$, τότε η τιμή χαρακτηρίζεται ως “Neutral”,
- $(\bar{x} + s, \bar{x} + 2s)$, τότε η τιμή χαρακτηρίζεται ως “Warm”,
- $(\bar{x} + 2s, \bar{x} + 3s)$, τότε η τιμή χαρακτηρίζεται ως “Hot”.



Στατιστικός όρος Μονάδες αριθμών

Αναφέρεται στο σύνολο των αριθμών, μέσα από το οποίο δημιουργούνται οι συνδυασμοί των κληρώσεων. Υπολογίζεται η συχνότητα εμφάνισης και καθυστέρησης αυτών των αριθμών.

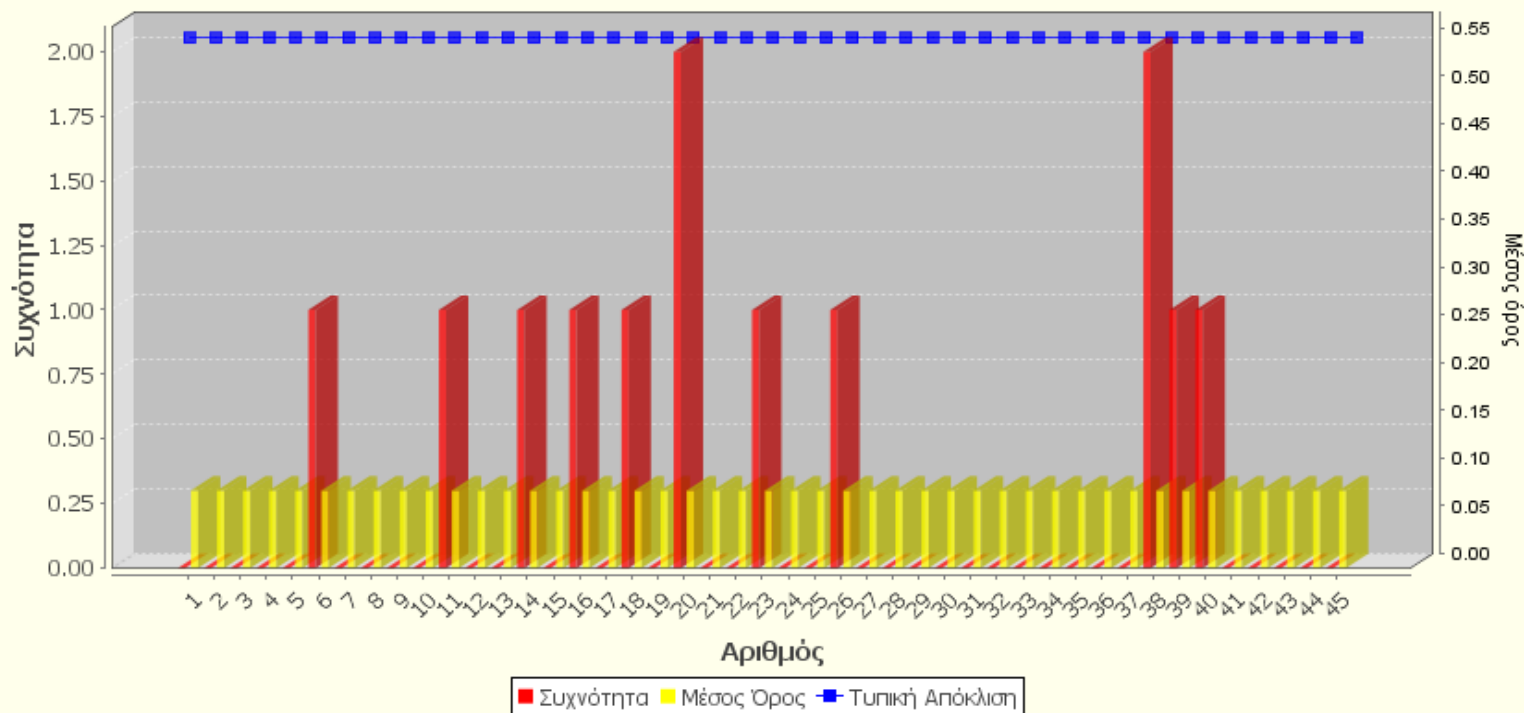




Στατιστικός όρος Επανάληψη Μονάδων αριθμών

Για κάθε έναν αριθμό από το σύνολο των αριθμών των συνδυασμών υπολογίζεται το πλήθος εμφάνισης του αριθμού αυτού δύο συνεχόμενες φορές.

Εμφάνιση Συχνότητας Επανάληψης Μονάδων

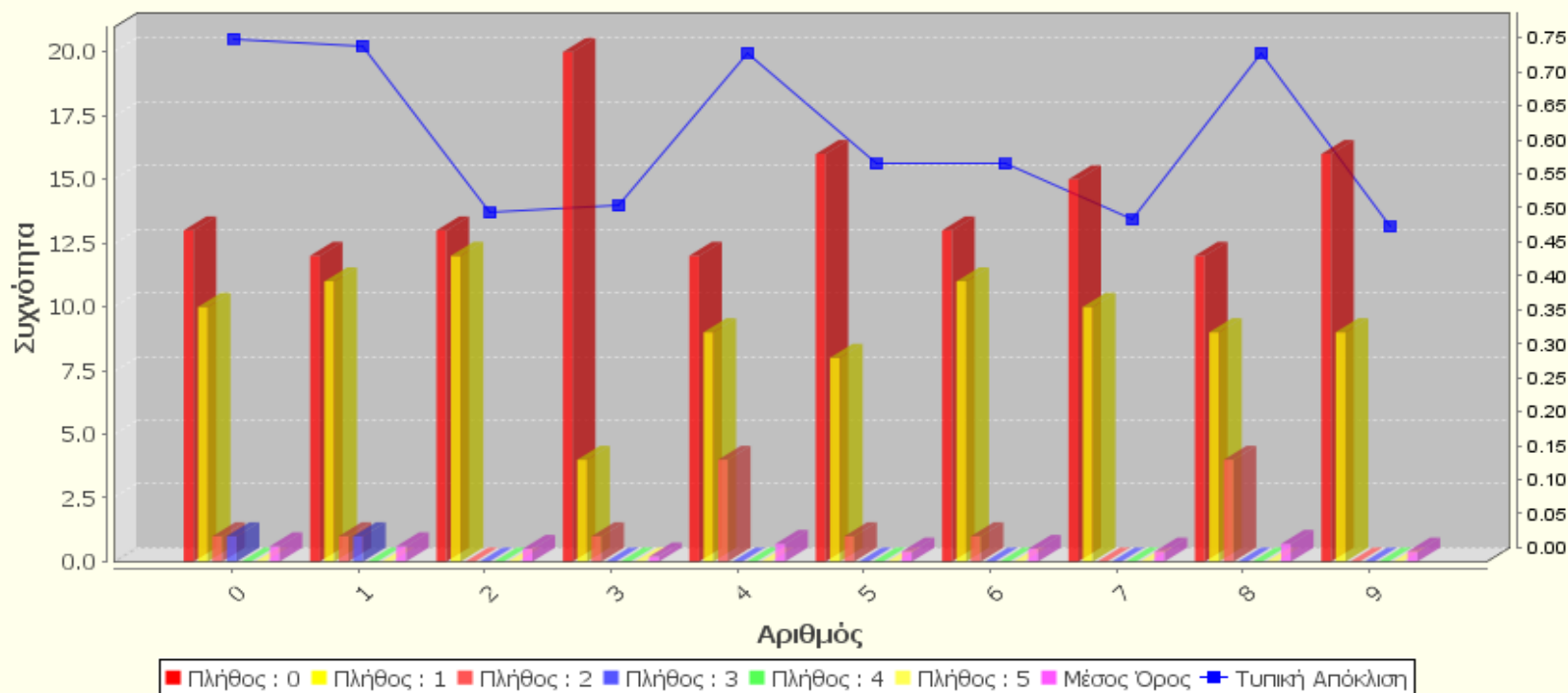




Στατιστικός όρος Λήγοντες αριθμοί

Οι 45 αριθμοί που συμμετέχουν στους συνδυασμούς χωρίζονται σε ομάδες ανάλογα με τον αριθμό στον οποίο λήγουν. Έτσι μελετάται η στατιστική ανάλυση των ομάδων αυτών.

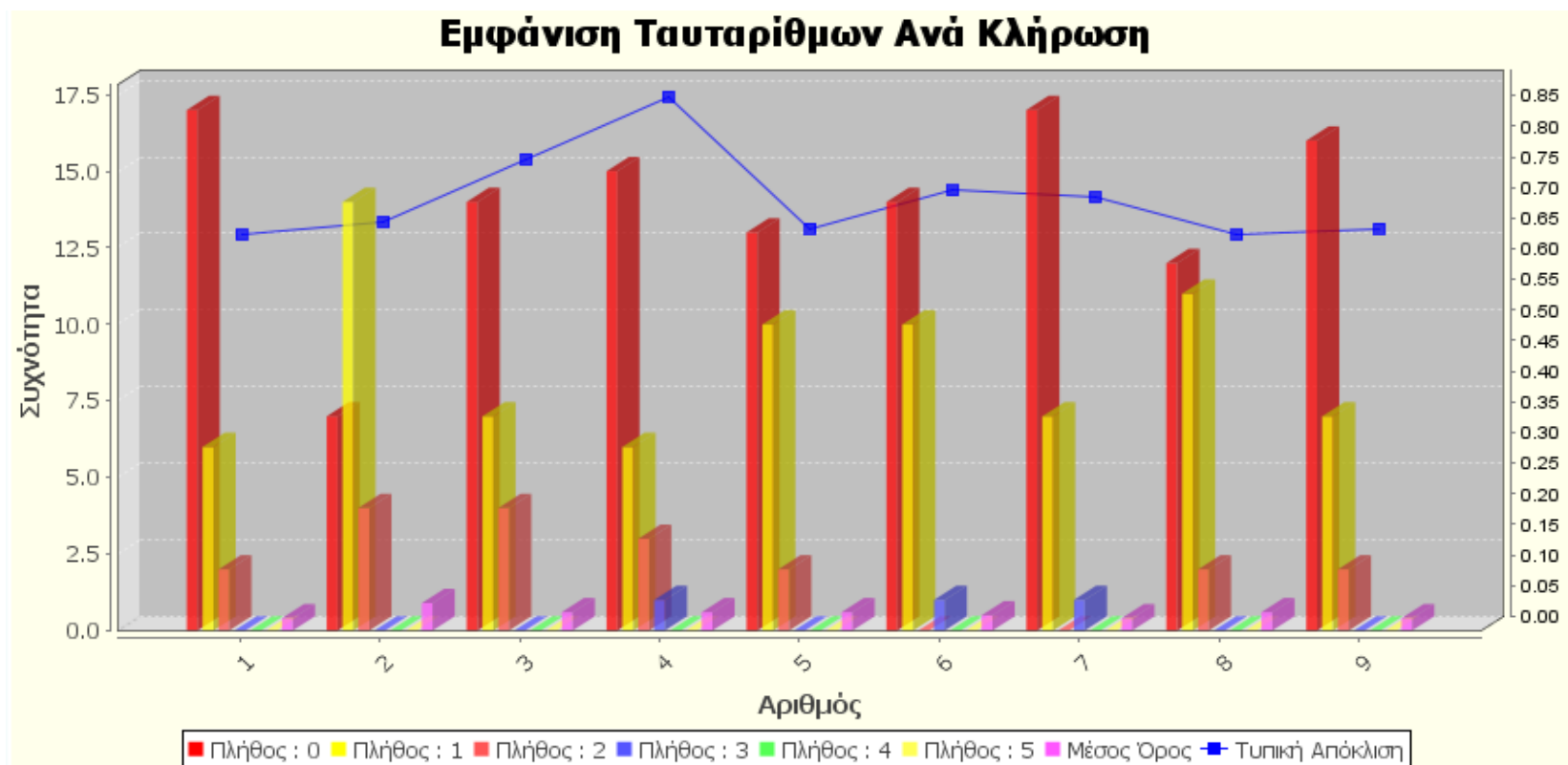
Εμφάνιση Λήγοντων Ανά Κλήρωση





Στατιστικός όρος Ταυτάριθμοι αριθμοί

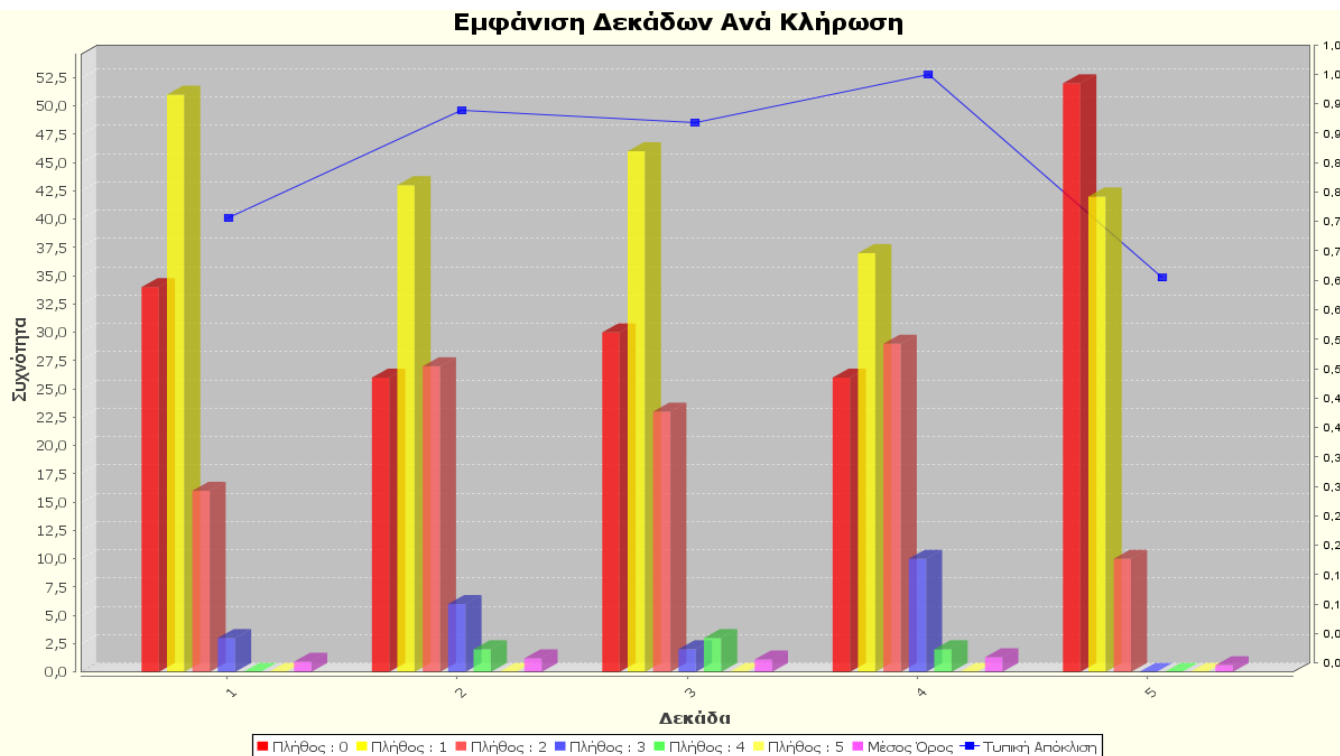
Δημιουργούνται υποσύνολα του συνόλου των αριθμών που συμμετέχουν στους συνδυασμούς βάσει του μονοψήφιου αθροίσματος των ψηφίων κάθε αριθμού ξεχωριστά.





Στατιστικός όρος Δεκάδες αριθμοί

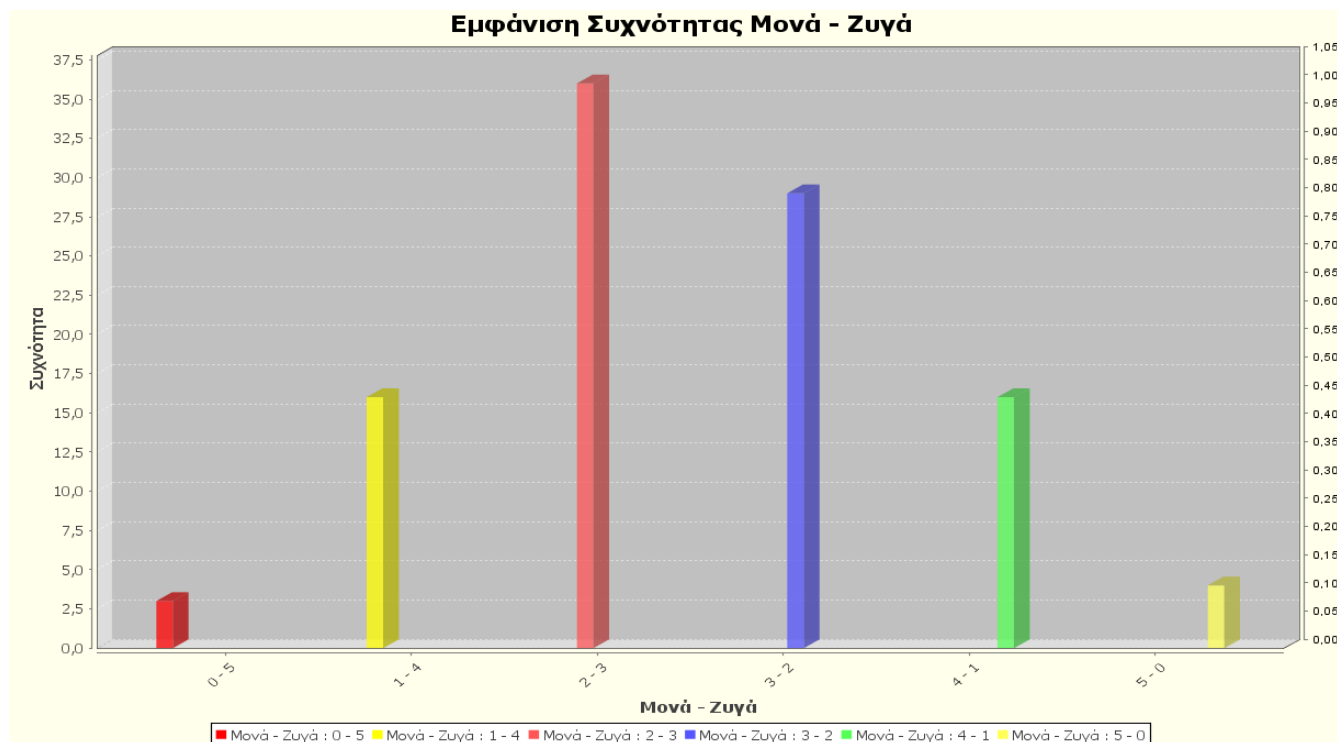
Αφού κατατάξουμε κάθε έναν από τους 45 αριθμούς των συνδυασμών σε κάποιο υποσύνολο, βάσει της δεκάδας στην οποία ανήκουν βρίσκουμε την συχνότητα εμφάνισης και καθυστέρησης, όπως και το πλήθος αυτών.





Στατιστικός όρος Μονά – Ζυγά αριθμών

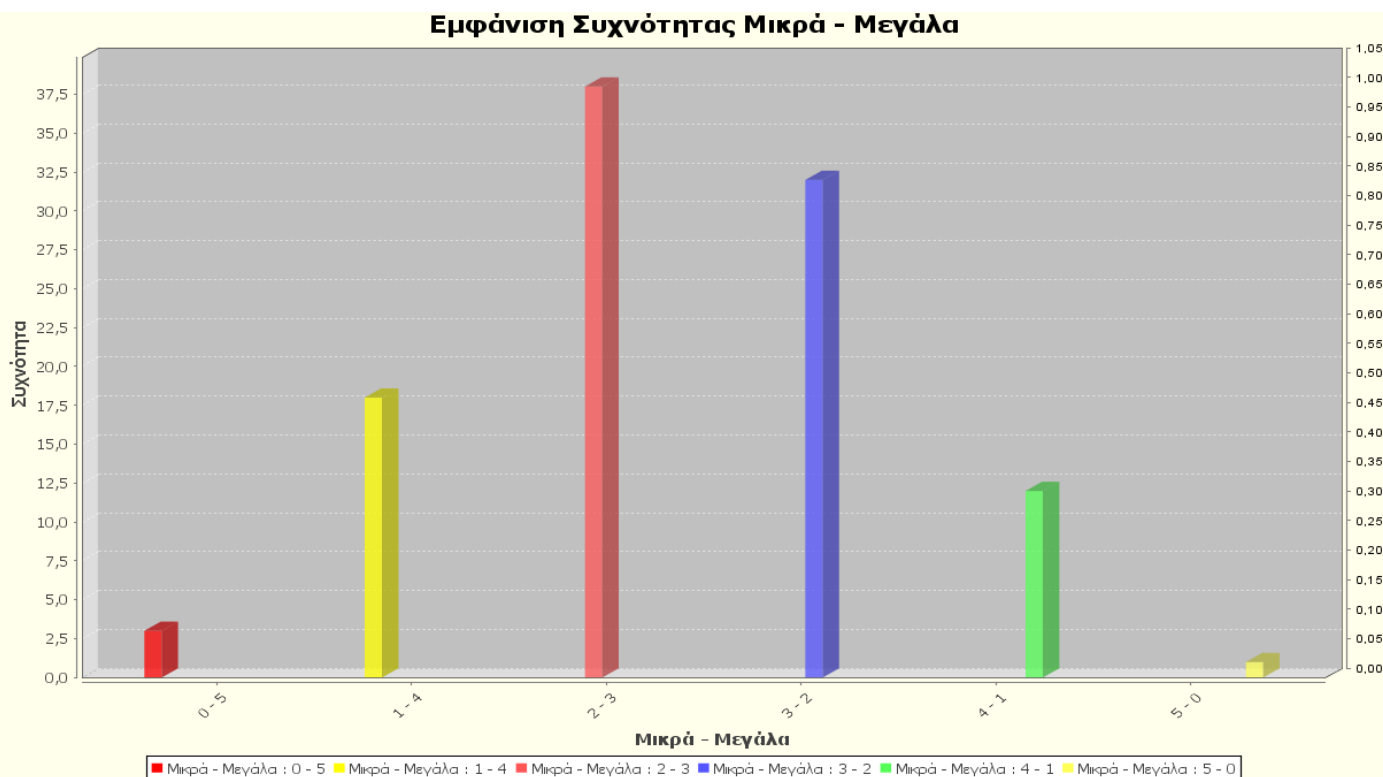
Ο συγκεκριμένος στατιστικός όρος αναφέρεται στη συχνότητα εμφάνισης ή καθυστέρησης του πλήθους των αριθμών ενός συνδυασμού που μπορεί να είναι μονοί ή ζυγοί.





Στατιστικός όρος Μικροί – Μεγάλοι αριθμών

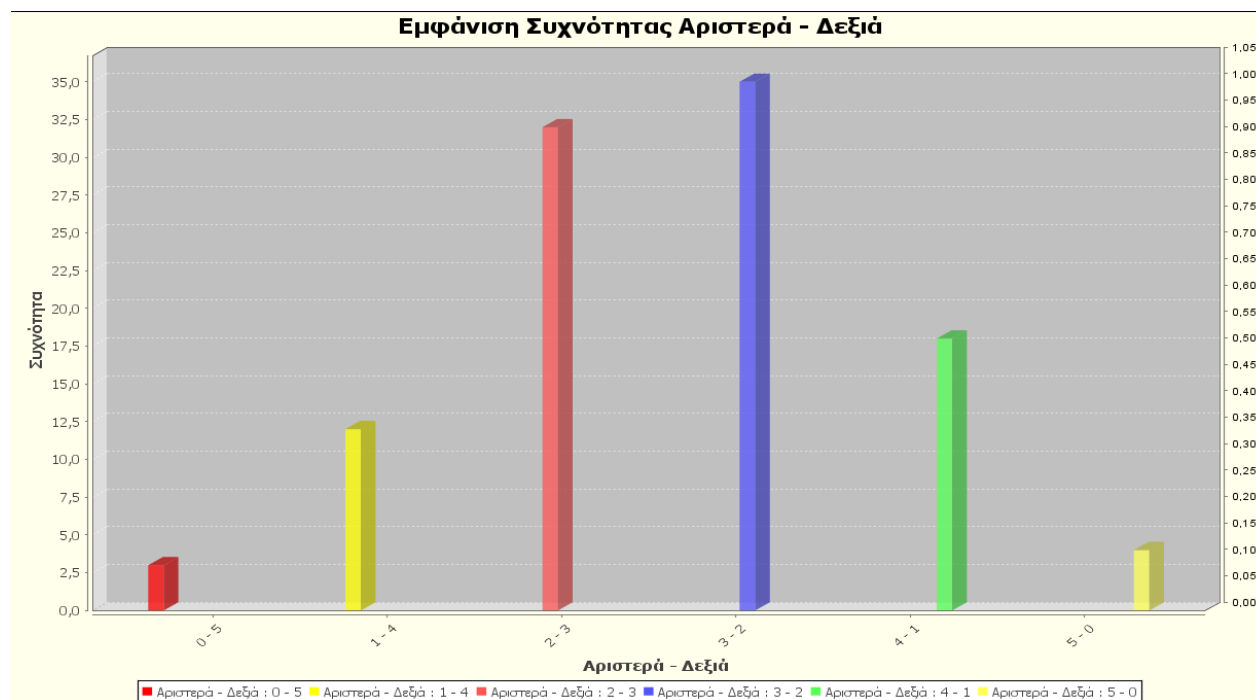
Ένας αριθμός μπορεί να χαρακτηριστεί «Μικρός» αν ανήκει στο πρώτο μισό των συνολικών αριθμών, δηλαδή στο σύνολο $[1, 22]$, ενώ αντίθετα «Μεγάλος» ονομάζεται ένας αριθμός που ανήκει στο δεύτερο μισό των συνολικών αριθμών, στο σύνολο $[23, 45]$.





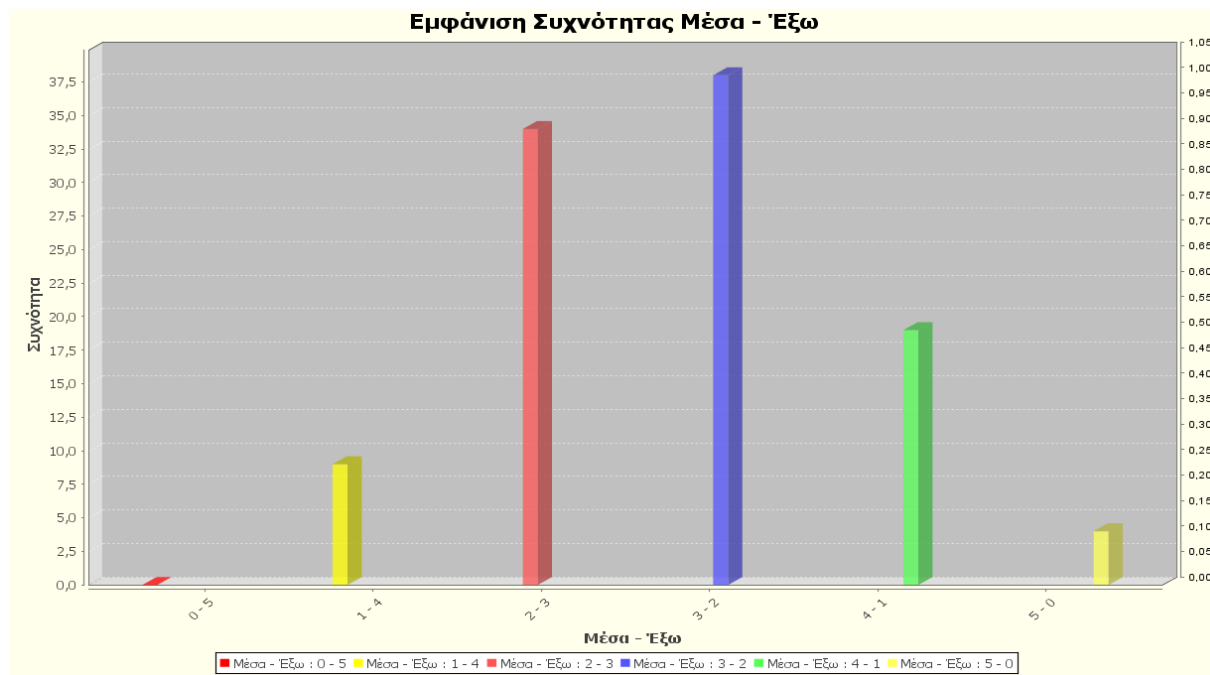
Στατιστικός όρος Αριστεροί – Δεξιοί αριθμών

Με τον όρο «Αριστερά» αναφερόμαστε στους αριθμούς των οποίων ο λήγοντας αριθμός ανήκει στο σύνολο $\{1, 2, 3, 4, 5\}$, ενώ δεξιοί είναι οι αριθμοί που ο λήγοντάς τους αριθμός ανήκει στο σύνολο $\{6, 7, 8, 9, 10\}$.



Στατιστικός όρος Μέσα – Έξω αριθμών

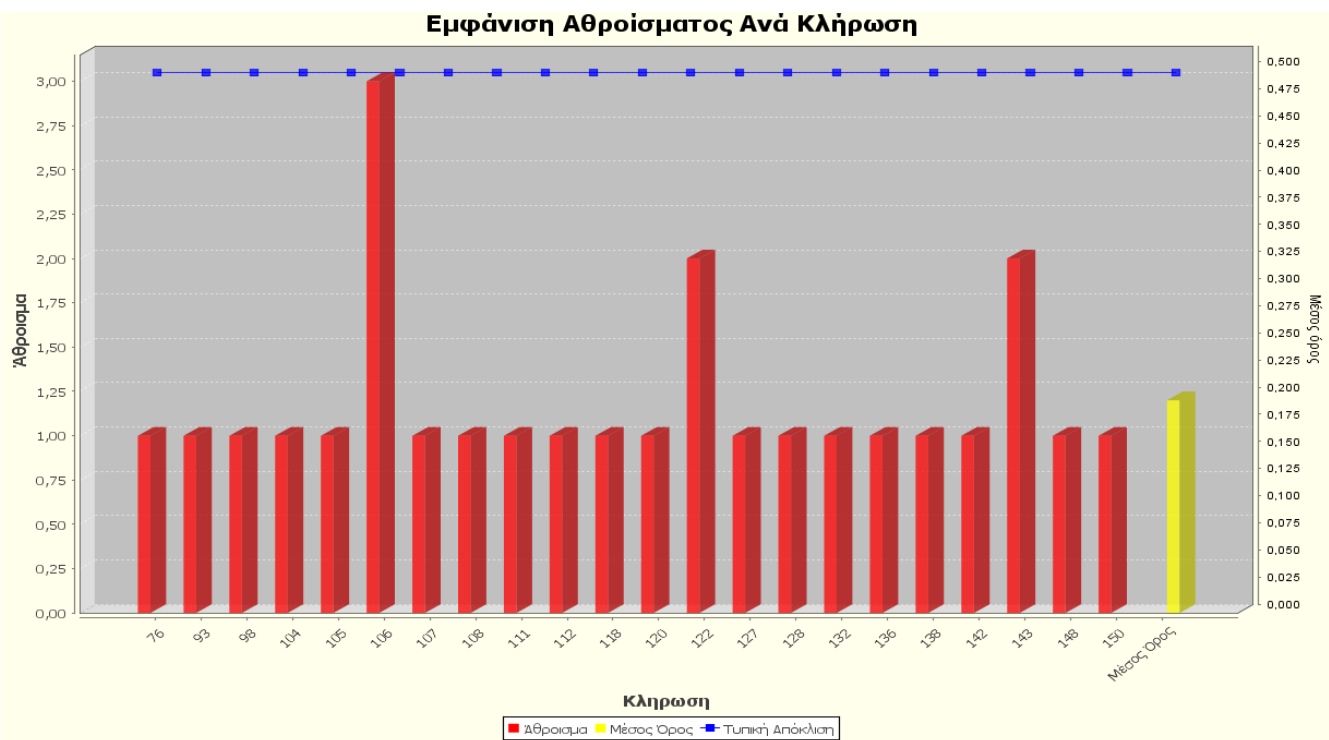
- Ένας αριθμός χαρακτηρίζεται ως «Μέσα» αν ανήκει στην ομάδα των εσωτερικών αριθμών, δηλαδή στο σύνολο $[12, 19] \cup [22, 29] \cup [32, 35]$. Αντίθετα «Έξω» αριθμός ονομάζεται ένας αριθμός που ανήκει στην ομάδα των εξωτερικών αριθμών, δηλαδή στο σύνολο $[1, 11] \cup [20, 21] \cup [30, 31] \cup [36, 45]$.





Στατιστικός όρος Άθροισμα στήλης

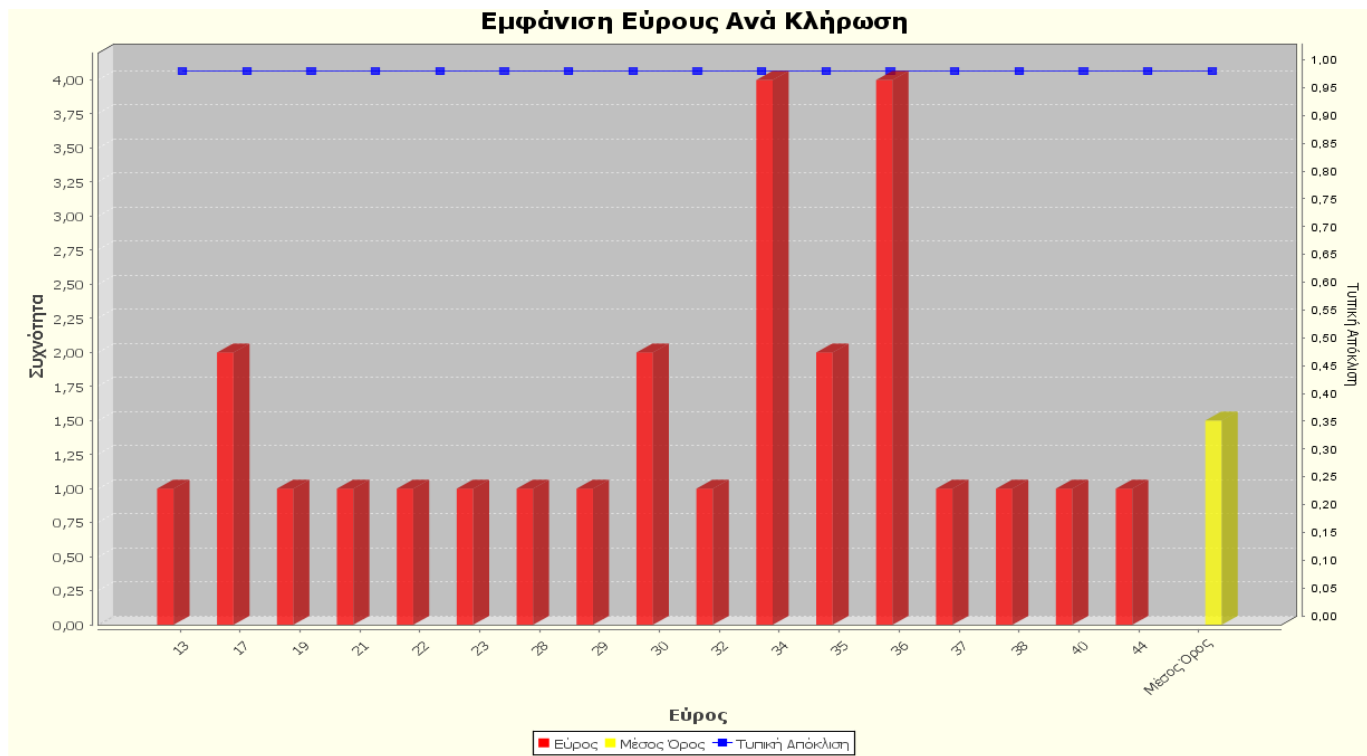
Άθροισμα στήλης είναι ο όρος εκείνος, ο οποίος αναλύει στατιστικά την συχνότητα εμφάνισης και καθυστέρησης των αθροισμάτων των αριθμών ανά κλήρωση.





Στατιστικός όρος Εύρος στήλης

Εύρος στήλης είναι διαφορά του μεγαλύτερου αριθμού της στήλης από τον μικρότερο. Προφανώς, για την υλοποίηση αυτού, θα πρέπει για κάθε στήλη να ταξινομηθούν κατά αύξουσα τάξη οι αριθμοί κάθε μίας κλήρωσης.





Χρήση βάσης δεδομένων για την αποθήκευση και διαχείριση βάσης δεδομένων

Οι λόγοι για τους οποίους επιλέχθηκε η χρήση βάσης δεδομένων για την αποθήκευση και διαχείριση των δεδομένων από τους συνδυασμούς των κληρώσεων είναι οι παρακάτω:

- Δυνατότητα αποθήκευσης μεγάλου όγκου πληροφορίας.
- Έλεγχος και ασφάλεια των δεδομένων.
- Εγκυρότητα των δεδομένων.
- Εύκολα επαναχρησιμοποιούμενη πληροφορία.
- Λειτουργικός τρόπος διαχείρισης της πληροφορίας με χρήση sql commands.



Χρήση της βιβλιοθήκης SQLite

- Η βιβλιοθήκη SQLite είναι ελεύθερο λογισμικό, που χρησιμοποιήθηκε η αντικειμενοστραφής γλώσσα προγραμματισμού C++ για την ανάπτυξή της.
- Έναρξη διαδικασίας ανάπτυξής της τον Απρίλιο του 2004 με την πρώτη έκδοσή της τον Ιούνιο του ίδιου έτους.
- Πρόσφατη δοκιμαστική έκδοση 3.7.7.1 τον Απρίλιο του 2011.
- Μεγάλες εταιρίες χρηματοδοτούν το project, όπως η ORACLE, ADOBE, NOKIA, κ.ά. με αποτέλεσμα την συνεχή ανάπτυξη της βιβλιοθήκης.



Bloomberg



Γιατί να προτιμηθεί η SQLite;

Οι λόγοι που μας οδήγησαν στην επιλογή αυτής της βιβλιοθήκης είναι οι εξής:

- Ανεξάρτητη από εξυπηρετητές βάσεων δεδομένων μηχανής, όπως MS SQL server 2008, Oracle server, DBII server, κ.ά.
- Πλήρης λειτουργικότητα πινάκων, triggers, indexes, views, stored procedure, κ.ά.
- Χρήση ενός μικρού αρχείου .db για την δημιουργία τοπικής βάσης δεδομένων, με ελάχιστες απαιτήσεις μνήμης.
- Δεν απαιτείται εγκατάσταση και ρυθμίσεις διαχείρισης και παραμετροποίησης της.



Διαδικασία σύνδεσης με την τοπική βάση μέσω JDBC driver

- Φόρτωση του κατάλληλου αρχείου .jar, το sqlitejdbc-v056.jar μέσα στο project έτσι, ώστε να έχουμε τη δυνατότητα χρήσης του API της βιβλιοθήκης.
- Φόρτωση του αντίστοιχου οδηγού επικοινωνίας με την βάση με την εντολή: `Class.forName("org.sqlite.JDBC");`
- Εντολή σύνδεσης με την τοπική βάση, με το όνομα αρχείου sample.db: `Connection connection = DriverManager.getConnection ("jdbc:sqlite:sample.db");`



Δημιουργία sql commands

- Ορισμός και αρχικοποίηση του αντικειμένου τύπου Statement, το οποίο αλληλεπιδράει μέσω του driver με την βάση δεδομένων:

```
Statement statement = connection.createStatement();
```

- Ορισμός και αρχικοποίηση παραμετρικών ερωτημάτων τύπου PreparedStatement, για την αλληλεπίδραση με την βάση δεδομένων.

```
PreparedStatement insertSixnotita =  
connection.prepareStatement("insert into (SixnotitaTable, isJoker,  
FromDate, ToDate, Sixnotita, MOEmfanisis, TAEmfanisis, IsHot)  
values ( ?, ?, ?, ?, ?, ?, ?, ?);");
```

- Χρήση των μεθόδων set για την καταχώρηση τιμών σε παράμετρο και get για το διάβασμα των τιμών ενός ερωτήματος.
- Αποσύνδεση από την βάση με την εντολή:

```
connection.close();
```



Εκτέλεση sql commands

Εκτέλεση του command μέσω δύο εντολών:

1. `executeUpdate()`, για εντολές που δημιουργούν ή ενημερώνουν τα περιεχόμενα των πινάκων,
2. `executeQuery()`, για ερωτήματα τύπου `SELECT`. Τα αποτελέσματα του ερωτήματος επιστρέφονται σε ένα αντικείμενο τύπου `ResultSet`.



Ενημέρωση και διαχείριση της βάσης δεδομένων

Για την ενημέρωση της τοπικής βάσης με τους συνδυασμούς των κληρώσεων χρησιμοποιούνται τρεις τρόποι.

- Χειροκίνητα
- Μέσω αρχείου κειμένου ή excel
- Απευθείας από το επίσημο διαδικτυακό τόπο του ΟΠΑΠ Α.Ε.



Ενημέρωση βάσης δεδομένων με χρήση αρχείου κειμένου

- Για την ενημέρωση της τοπική βάσης, ο χρήστης μπορεί να φορτώσει κάποιο αρχείο κειμένου, όπως .txt, .doc ή .csv, στα οποία οι αριθμοί είναι χωρισμένοι μεταξύ τους με κάποιον από τους χαρακτήρες ‘,’ , ‘;’ , ‘_’ ή ‘.’.

- Ορισμός και αρχικοποίηση αντικειμένου τύπου FileInputStream για την σύνδεση του thread με το αντίστοιχο αρχείο κειμένου:

```
FileInputStream fis = new FileInputStream(chooser.getSelectedFile());
```

- Ορισμός και αρχικοποίηση αντικειμένου τύπου DataInputStream για την ανάγνωση των δεδομένων από το αρχείο κειμένου:

```
DataInputStream dis = new DataInputStream(new  
BufferedInputStream(fis));
```



Ενημέρωση της βάσης δεδομένων με χρήση αρχείου excel

- Το αρχείο excel που θα διαλέξει ο χρήστης για να διαβάσει τους συνδυασμούς των κληρώσεων και να τους αποθηκεύσει στη βάση θα πρέπει να είναι format 1997– '03.
- Για την λειτουργία ανάγνωσης από αρχείο excel χρησιμοποιήθηκε μία άλλη βιβλιοθήκη ανοικτού κώδικα, όπως η HSSF – POI της εταιρίας ανάπτυξης λογισμικού Apache POI project.
- Για την χρήση της βιβλιοθήκης αυτής απαιτείται η φόρτωση του αντίστοιχου αρχείου .jar έτσι, ώστε να επιτρέπεται η χρήση του API του συγκεκριμένου project.



The Apache POI Project

<http://poi.apache.org>



Διαδικασία ανάγνωσης και εγγραφής σε αρχείο excel μέσω HSSF POI

- Ορισμός και αρχικοποίηση αντικειμένου τύπου HSSFWorkbook:
`HSSFWorkbook saveWorkbook = new HSSFWorkbook();`
- Ορισμός και αρχικοποίηση αντικειμένου τύπου HSSFSheet:
`HSSFSheet saveDataSheet = saveWorkbook.createSheet("Sheet1");`
- Ανάγνωση όλων των γραμμών του φύλλου excel μέσω του αντικειμένου τύπου HSSFRow:
`HSSFRow dataRow = saveDataSheet.createRow(i);`
- Ανάγνωση των κελιών ανά γραμμή μέσω του αντικειμένου τύπου HSSFCell:
`HSSFCell cell = dataRow.getCell(c);`



Μέθοδοι ανάγνωση δεδομένων από κελί αρχείου excel με HSSF POI

- Αν είναι αριθμητικού τύπου:
`cell.getNumericCellValue();`
- Αν είναι αλφαριθμητικού τύπου:
`cell.getStringCellValue();`
- Αν είναι τύπου formula:
`cell.getCellFormula();`



Ενημέρωση της βάσης δεδομένων απευθείας από τα αρχεία του ΟΠΑΠ Α.Ε

- Ο πιο εύχρηστος και ασφαλής τρόπος για την ενημέρωση των δεδομένων στην τοπική βάση είναι η απευθείας από το site του ΟΠΑΠ Α.Ε.
- Αρχικά ορίζουμε και αρχικοποιούμε ένα αντικείμενο τύπου URL:
`url = new URL(numberOfExcel);`
Όπου `numberOfExcel` είναι το πλήρες url στο οποίο βρίσκεται το αντίστοιχο αρχείο excel στον εξυπηρετητή που φιλοξενεί το site. Π.χ.
http://media.opap.gr/Excel/5104/Joker_2012.xls
- Στη συνέχεια ανοίγουμε την σύνδεση με το συγκεκριμένο url με τη εντολή `url.openConnection()`
- Τέλος ορίζουμε και αρχικοποιούμε ένα αντικείμενο τύπου `InputStream` με την εντολή:
`InputStream is = url.openStream();`
και έπειτα μπορούμε να ακολουθήσουμε τις διαδικασίες ανάγνωσης αρχείου excel όπως ορίστηκαν προηγουμένως.



Εργαλεία ανάπτυξης του λογισμικού

Γλώσσα προγραμματισμού

- Για γλώσσα προγραμματισμού χρησιμοποιήθηκε η Java, η οποία υποστηρίζει τον αντικειμενοστραφή προγραμματισμό.
- Είναι προϊόν ελεύθερου λογισμικού με πολλές δυνατότητες και λειτουργίες.
- Υποστηρίζεται από πλήρες και τεκμηριωμένο API.
- Για τα γραφικά της εφαρμογής χρησιμοποιήθηκε το GUI της Swing, το οποίο είναι ένα ολοκληρωμένο πακέτο γραφικών συστατικών components (Java GUI widget toolkit) για java εφαρμογές.
- Ο κώδικας αναπτύχθηκε στο περιβάλλον του Netbeans



Εργαλεία ανάπτυξης του λογισμικού

Δημιουργία διαγραμμάτων 1/2

- Για την δημιουργία των γραφημάτων αναπαράστασης της συχνότητας εμφάνισης και καθυστέρησης των στατιστικών όρων χρησιμοποιήθηκε η κλάση JFreeChart.java, ελεύθερου λογισμικού και προϊόν του David Gilbert από τον Φεβρουάριο του 2000.
- Αρχικά ορίζεται ένα αντικείμενο τύπου CategoryDataset:
`CategoryDataset dataset1;`
- Για την αρχικοποίηση του χρησιμοποιείται η μέθοδος `createDataset(data)`, με παράμετρο τα δεδομένα που θα εμφανίζονται στο γράφημα. Η μέθοδος αυτή καταχωρεί στο αντικείμενο `dataset1` τις τιμές των δεδομένων μέσω της εντολής `addValue(value1, value2, value3)`, όπου `value1` η τιμή στον άξονα y , `value2` το label του άξονα y και `value3` η τιμή στον άξονα x .



Εργαλεία ανάπτυξης του λογισμικού

Δημιουργία διαγραμμάτων 2/2

Ορισμός και αρχικοποίηση του αντίστοιχου γραφήματος με την μέθοδο `createBarChart3D` του αντικειμένου `ChartFactory`.

```
JFreeChart chart = ChartFactory.createBarChart3D(  
    title, // chart title  
    xLabel, // domain axis label  
    yLabel, // range axis label  
    dataset1, // data  
    PlotOrientation.VERTICAL,  
    true, // include legend  
    true,  
    false  
);
```



Συμπεράσματα

Η διαδικασία επιλογής των αριθμών που θα ορίσουν τον νικηφόρο συνδυασμό είναι μία περίπλοκη κατάσταση. Η επιτυχία είναι προϊόν μιας σύνθετης διαδικασίας ελέγχου, μελέτης και τύχης. Τον βασικό ρόλο παίζει φυσικά η τύχη του χρήστη, ωστόσο η προσεκτική μελέτη μαθηματικών στατιστικών μοντέλων και αναλύσεων μπορεί να οδηγήσει σε συγκεκριμένους αριθμούς, που έχουν περισσότερες πιθανότητες επιτυχίας. Η εφαρμογή αυτή προσπαθεί να καλύψει το κομμάτι του μαθηματικού στατιστικού μοντέλου με όσο το δυνατόν μεγαλύτερη ευστοχία. Ελπίζω να φέρει αποτελέσματα και ευτυχία στους χρήστες του!!



Βιβλιογραφία

- Herb Schilddt: JAVA, Έτοιμες συνταγές, Εκδόσεις: Μ. Γκιούρδας.
- Herb Schilddt: Οδηγός της Java 2, Τρίτη έκδοση, Εκδόσεις: Μ. Γκιούρδας.
- Ηλεκτρονική βιβλιογραφία
 1. <http://www.anendotos.gr/Tzoker/Tzoker.asp?stID=91>
 2. <http://docs.oracle.com/javase/tutorial/uiswing/components/index.html>
 3. <http://www.jfree.org/jfreechart/>
 4. <http://poi.apache.org/spreadsheet/index.html>
 5. <http://www.sqlite.org/>
 6. <http://www.lottomania2000.net/>
 7. http://www.lotterypowerpicks.com/tx6_hotcold.htm
 8. <http://netbeans.org/>