

Διαδικτυακό Μάθημα:

**Ανάλυση των
Δεδομένων της
Πλατφόρμας
ENIRISST**

Με τη Χρήση
Μεθόδων
Χωρικής
Ανάλυσης και ΣΓΠ



ENIRISST
<http://www.enirisst.gr>



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΙΓΑΙΟΥ



ΕΠΑνεΚ 2014-2020
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ • ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΤΗΤΑ • ΚΑΙΝΟΤΟΜΙΑ



ΕΣΠΑ
2014-2020
Partnership Agreement
2014 - 2020

03: R και Βασικά GIS

Μάριος Μπάτσαρης, Δημήτρης Καβρουδάκης

03.1 Προετοιμασία

Δημιουργήστε ένα φάκελο στον προσωπικό σας υπολογιστή με όνομα `rspatial_2`.

Κατεβάστε και αποσυμπιέστε τα δεδομένα στο φάκελο `rspatial_2`.

Στη συνέχεια στο Rstudio ακολουθήστε τα βήματα της προετοιμασίας (διαφάνειες 01.1 και 01.2) από την παρουσίαση `01_R_Introduction.pptx` για τη δημιουργία και αποθήκευση ενός νέου script.

Το νέο script να ονομαστεί `rspatial_2` και αποθηκευτεί στον αντίστοιχο φάκελο `rspatial_2`

03.2 Προετοιμασία

Η δομή του συστήματος αρχείων που απαιτούνται για το μάθημα αυτό θα πρέπει να έχει την εξής μορφή.

```
earning > tests > rspatial_2
```

Όνομα

data

rspatial_2

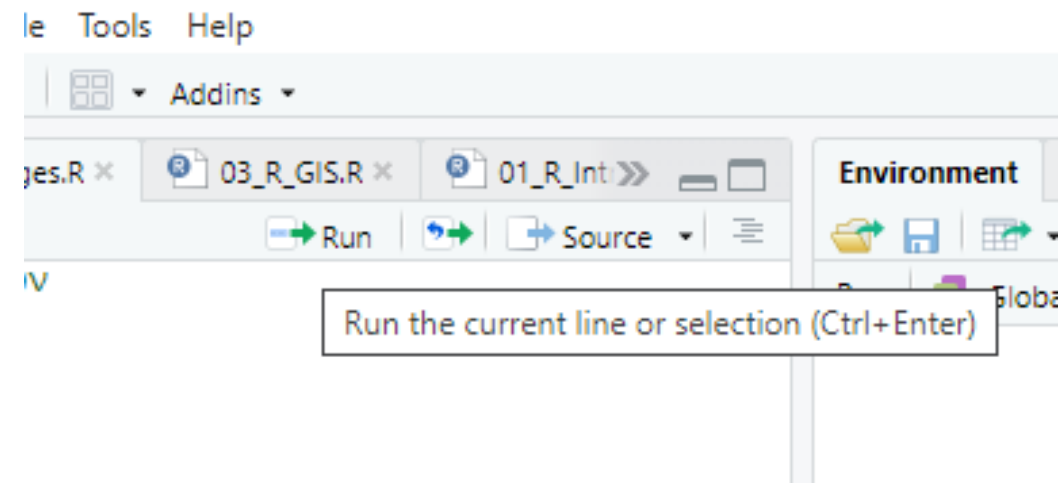
03.3 Προετοιμασία

Το script 00_install_r_packages.R περιλαμβάνει τις εντολές για την εγκατάσταση των απαραίτητων βιβλιοθηκών. Εάν έχετε ήδη εγκαταστήσει τις βιβλιοθήκες αγνοήστε αυτό το βήμα.

Υπάρχουν δύο τρόποι για να τρέξετε τον κώδικα γραμμή – γραμμή.

Ο πρώτος είναι:

Με τοποθέτηση του κερσορα του ποντικιού στην αντίστοιχη γραμμή και κλικ στην επιλογή «Run»



03.4 Προετοιμασία

Το script 00_install_r_packages.R περιλαμβάνει τις εντολές για την εγκατάσταση των απαραίτητων βιβλιοθηκών.

Υπάρχουν δύο τρόποι για να τρέξετε τον κώδικα γραμμή – γραμμή.

Ο δεύτερος είναι:

Με τοποθέτηση του κέρσορα του ποντικιού στην αντίστοιχη γραμμή και «Ctrl + Enter» από το πληκτρολόγιο

03.5 Προετοιμασία

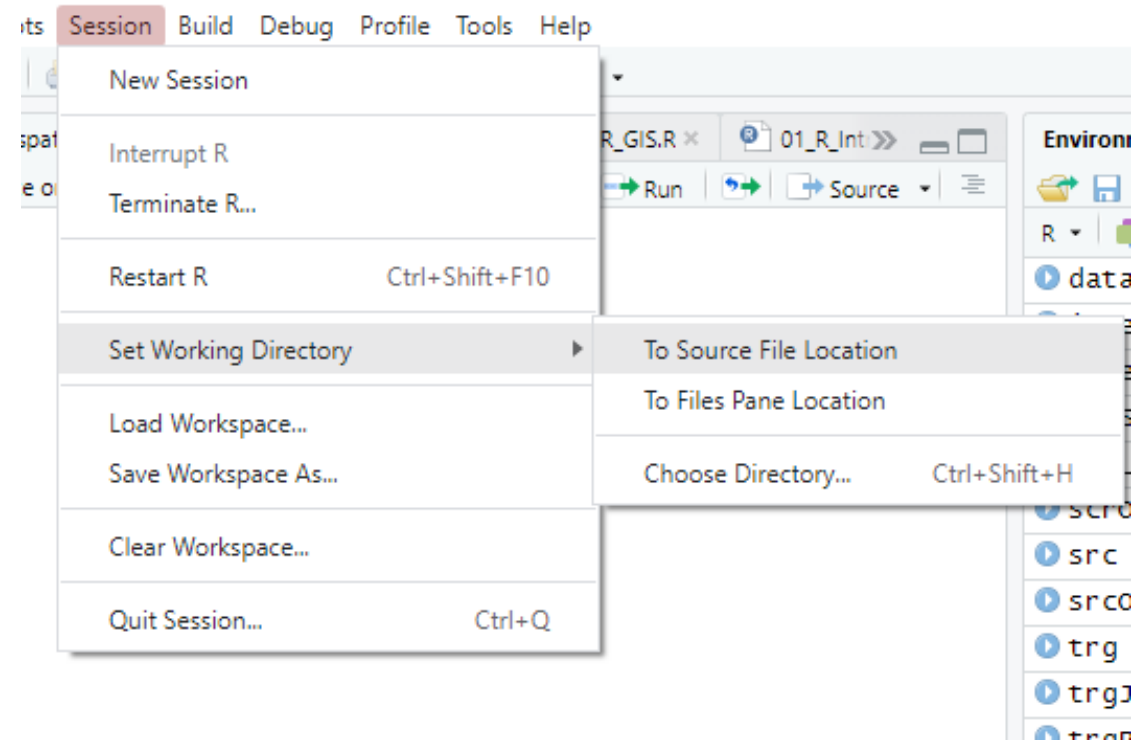
Επιστροφή το νέο script που δημιουργήσατε:

Από το μενού επιλογών του RStudio ακολουθήστε τα παρακάτω βήματα:

Session ->

Set Working Directory ->

To Source File Location



03.6 Προετοιμασία

Με αυτόν τον τρόπο καθορίζεται ότι ο φάκελος `rspatial_1` στον οποίο βρίσκεται και το script `rspatial_1.R` θα είναι το περιβάλλον εργασίας.

Έτσι διευκολύνεται σε μεγάλο βαθμό η διαδικασία φόρτωσης των απαραίτητων δεδομένων.

03.7 Φόρτωση Βιβλιοθηκών

Εφόσον έχει γίνει η εγκατάσταση των απαραίτητων βιβλιοθηκών επόμενο βήμα είναι η φόρτωση τους στο script. Για τη φόρτωση μιας βιβλιοθήκης χρησιμοποιείται η εντολή `library()` όπως φαίνεται στην παρακάτω εικόνα. Η βιβλιοθήκες που θα χρησιμοποιηθούν είναι η `popular` και η `sf`

<https://cran.r-project.org/package=populR>

<https://journal.r-project.org/archive/2018/RJ-2018-009/RJ-2018-009.pdf>

<https://cran.r-project.org/web/packages/sf/vignettes/sf1.html>

```
1 # Φόρτωση των απαραίτητων βιβλιοθηκών
2 # φόρτωση των απαραίτητων βιβλιοθηκών
3 library(populR)
4 library(sf)
5
```

03.8 Φόρτωση Δεδομένων

Εφόσον έχει γίνει η φόρτωση των απαραίτητων βιβλιοθηκών επόμενο βήμα είναι η φόρτωση των δεδομένων.

Τα δεδομένα που θα χρησιμοποιηθούν ανήκουν στη βιβλιοθήκη popuIR και αφορούν υποσύνολο των οικοδομικών τετραγώνων (src) και κτιρίων (trg) της πόλης της Μυτιλήνης.

Τα οικοδομικά τετράγωνα (ΟΤ) περιλαμβάνουν:

- sid: αναγνωριστικό των ΟΤ
- pop: defacto πληθυσμός της απογραφής του 2011
- geometry: στοιχεία του αρχέτυπου αλλά και των συντεταγμένων των σημείων

03.9 Φόρτωση Δεδομένων

Τα κτιρια περιλαμβάνουν:

- tid: αναγνωριστικό των κτιρίων
- floors: αριθμός ορόφων για κάθε κτήριο
- rf: εκτίμηση πληθυσμού για κάθε κτήριο
- geometry: στοιχεία του αρχέτυπου αλλά και των συντεταγμένων των σημείων

03.10 Φόρτωση Δεδομένων

Τα δεδομένα προσφέρονται από τη βιβλιοθήκη popular για εξερεύνηση των δυνατοτήτων της και άρα φορτώνονται με διαφορετικό τρόπο.

```
6 # χρήση δεδομένων βιβλιοθήκης popular
7 # τα δεδομένα αφορούν ένα μικρό τμήμα της πόλης
8 # της Μυτιλήνης, Λέσβου. Το ένα σετ δεδομένων
9 # (src) αποτελείται από οικοδομικά τετράγωνα
10 # ενώ το άλλο σετ (trg) αποτελείται από
11 # κτίρια εντός των οικοδομικών τετραγώνων
12 data("src")
13 data("trg")
14
```

03.11 Επισκόπηση Δεδομένων

Είναι σημαντικό να γίνεται επισκόπηση των δεδομένων για την καλύτερη δυνατή επεξεργασία και διαχείρισή τους.

Πληροφορίες για τα ονόματα `names()` των στηλών των δεδομένων, τον αριθμό των στηλών `ncol()` και γραμμών `nrow()` του πίνακα χαρακτηριστικών κάθε αρχείου αντλούνται ως εξής:

```
15 # αριθμός οικοδομικών τετραγώνων
16 nrow(src)
17
18 # αριθμός κτιρίων
19 nrow(trg)
20
21 # ονόματα στηλών οικοδομικών τετραγώνων
22 names(src)
23 # sid: αναγνωριστικός κωδικός οικοδομικού τετραγώνου
24 # pop: πληθυσμός οικοδομικού τετραγώνου (απογραφής 2011)
25
26 # ονόματα στηλών κτιρίων
27 names(trg)
28 # tid: αναγνωριστικός κωδικός κτιρίου
29 # floors: αριθμός ορόφων κτιρίου
30 # rf: πληθυσμός σε επίπεδο κτιρίου (εκτίμηση)
31
```

03.12 Επισκόπηση Δεδομένων

Σε αυτό το στάδιο γίνεται επιβεβαίωση ότι τα δεδομένα είναι της ίδιας κλάσης class() όπως επίσης ότι μοιράζονται το ίδιο προβολικό σύστημα.

Με αυτόν τον τρόπο περιορίζονται τα σφάλματα που μπορεί να προκύψουν κατά την επεξεργασία ή/και διαχείριση των δεδομένων.

```
37  
38 # προβολή της κλάσης των δεδομένων  
39 class(trg)  
40 class(src)  
41  
42 # εκτύπωση συστήματος συντεταγμένων  
43 st_crs(trg)$srId  
44 st_crs(src)$srId  
45
```

03.13 Οπτικοποίηση Δεδομένων

Η οπτικοποίηση των δεδομένων μπορεί να είναι από απλή (όπως στο προηγούμενο μάθημα) ή πιο σύνθετη. Σε αυτό το σημείο θα προστεθεί τίτλος, χρώματα και διαχείριση των ορίων των πολυγώνων κατά τη δημιουργία ενός χάρτη. Επιπλέον, μέσω της υπέρθεσης σε ένα χάρτη θα παρουσιάζονται τα ΟΤ και τα κτίρια με ευδιάκριτα χρώματα για την καλύτερη δυνατή κατανόηση του χάρτη.

```
45
46 # οπτικοποίηση των δεδομένων
47 # main είναι ο τίτλος του χάρτη
48 # χρωματισμός με τη μέθοδο hexagonal
49 # χωρίς όριο (border = NA)
50 # τα οικοδομικά τετράγωνα εμφανίζονται πίσω
51 # ενώ τα κτίρια εμφανίζονται μπροστά με τη
52 # χρήση της παραμέτρου add = T (όπου T = TRUE)
53 plot(src['geometry'], col = "#634B56", border = NA, main = "Χάρτης Μυτιλήνης")
54 plot(trg['geometry'], col = "#FD8D3C", border = NA, add = T)
55
```

03.14 Οπτικοποίηση Δεδομένων

Τα αποτελέσματα της οπτικοποίησης:

Χάρτης Μυτιλήνης



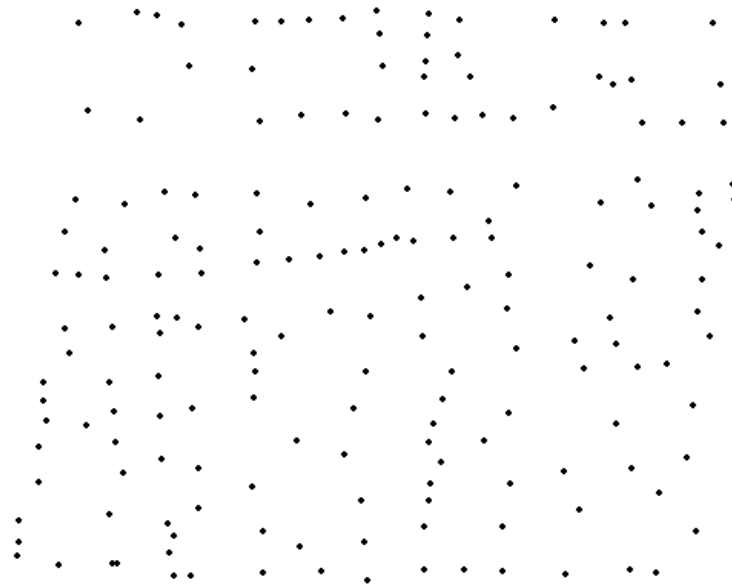
03.15 Γεωμετρικός Μετασχηματισμός

Σε ένα GIS παρέχεται η δυνατότητα μετατροπής μιας γεωμετρίας από τον έναν αρχέτυπο στον άλλον. Για παράδειγμα ένα τετράγωνο σχήμα μπορεί να αποτυπωθεί ως 4 σημεία (ένα για κάθε γωνία) ή ως 1 σημείο (κέντρο βάρους – κεντροειδές). Ωστόσο, η μετατροπή ενός σημείου σε γραμμή ή πολύγωνο είναι αδύνατη. Με την εντολή `st_centroid()` της `sf` είναι δυνατή η μετατροπή ενός πολυγώνου σε σημείο (κεντροειδές) όπως φαίνεται παρακάτω:

```
55  
56 # μετατροπή κτιρίων από πολύγωνα σε σημεία  
57 # με τη χρήση του κεντροειδούς  
58 trgPoint <- st_centroid(trg)  
59  
60 # οπτικοποίηση αποτελέσματος  
61 plot(trgPoint["geometry"], pch = 19, cex = 0.5)  
62
```

03.16 Γεωμετρικός Μετασχηματισμός

Η οπτικοποίηση του γεωμετρικού μετασχηματισμού:



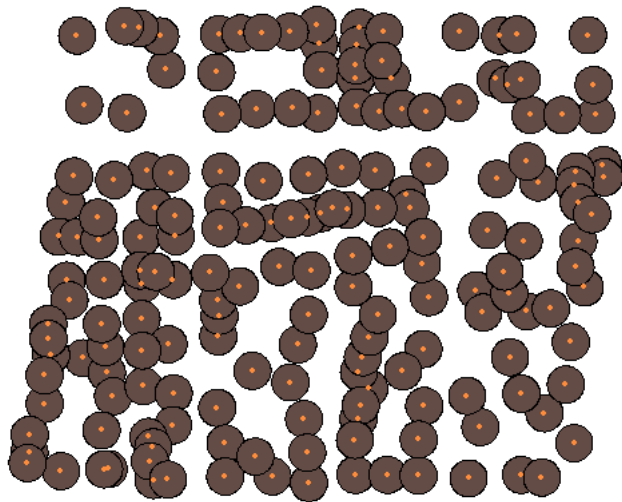
03.17 Δημιουργία Ζώνης (Buffer Zone)

Σε πολλές περιπτώσεις κρίνεται αναγκαία η δημιουργία μιας ζώνης με συγκεκριμένη απόσταση γύρω από ένα γεωμετρικό αρχέτυπο (σημείο, γραμμή ή πολύγωνο). Η δημιουργία ζώνης υλοποιείται με την εντολή `st_buffer()` της `sf`.

```
62
63 # δημιουργία ζώνης (buffer) 8 μέτρων σε σημειακά δεδομένα
64 bufferZonePoint <- st_buffer(trgPoint, 8)
65
66 # οπτικοποίηση αποτελέσματος
67 plot(bufferZonePoint["geometry"], col = "#634B46")
68 plot(trgPoint["geometry"], col = "#FD8D3C", pch = 19, cex = 0.5, add = T)
69
70 # δημιουργία ζώνης (buffer) 5 μέτρων σε δεδομένα επιφάνειας
71 bufferZonePoygon <- st_buffer(trg, 5)
72
73 # οπτικοποίηση αποτελεσμάτων
74 plot(bufferZonePoygon["geometry"], col = "#634B46")
75 plot(trg["geometry"], col = "#FD8D3C", add = T)
76
77
```

03.18 Δημιουργία Ζώνης (Buffer Zone)

Τα αποτελέσματα δημιουργίας ζώνης σε σημεία και πολύγωνα μοιάζει ως εξής:



03.19 Χωρική Διασταύρωση (Intersection)

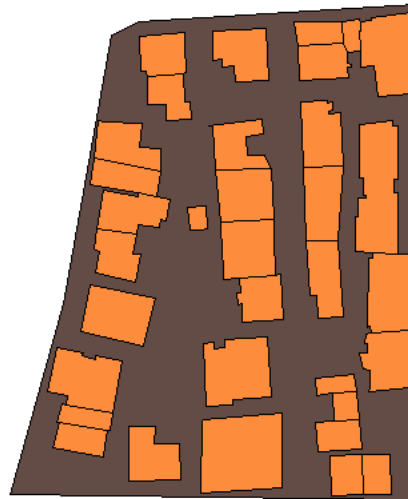
Είναι σύνηθες σε μια ανάλυση να μελετάται ένα υπο-σύνολο των δεδομένων.

Παρακάτω παρουσιάζεται ένα τέτοιο παράδειγμα. Επιλέγεται το ΟΤ με αναγνωριστικό 1 και στη συνέχεια διασταυρώνοντας τα δεδομένα (intersection) με την εντολή `st_intersect()` της `sf` παίρνουμε ως αποτέλεσμα τα κτίρια που ανήκουν απόκλειστικά στο πρώτο ΟΤ.

```
77
78 # επιλέγω το οικοδομικό τετράγωνο με sid 1 (υπο-σύνολο).
79 # τα κτίρια βρίσκονται εντός των οικοδομικών τετραγώνων
80 # εφαρμόζοντας τη διασταύρωση (intersection) σε ένα οικοδομικό τετράγωνο
81 # τότε το αποτέλεσμα θα είναι τα πολύγωνα των κτηρίων που διασταυρώνονται
82 # μόνο με το αντίστοιχο πολύγωνο του επιλεγμένου οικοδομικού τετραγώνου
83 srcone <- src[src$sid == 1, ]
84 inter <- st_intersection(srcone, trg)
85
86 # οπτικοποίηση των αποτελεσμάτων
87 plot(srcone["geometry"], col = "#634B46")
88 plot(inter["geometry"], col = "#FD8D3C", add = T)
89
```

03.20 Χωρική Διασταύρωση (Intersection)

Τώρα η ανάλυση μπορεί να υλοποιηθεί μόνο στο 1^ο ΟΤ και στα κτήρια που ανήκουν στο πρώτο ΟΤ. Παρακάτω παρουσιάζεται η οπτικοποίηση του αποτελέσματος.



03.21 Έλεγχος Διασταύρωσης

Πολλές είναι οι περιπτώσεις όπου ένας χρήστης GIS επιθυμεί να μετρήσει π.χ. πόσα κτίρια ανήκουν σε ένα οικοδομικό τετράγωνο. Ο έλεγχος διασταύρωσης μπορεί να βοηθήσει σε αυτή την κατεύθυνση. Ο έλεγχος διασταύρωσης εξετάζει ποια κτίρια του παραδείγματος τέμνουν κάθε οικοδομικό τετράγωνο κι επιστρέφει τα αποτελέσματα με τη μορφή μιας λίστας. Κάθε αντικείμενο της λίστας είναι ένα διάνυσμα με τιμές κτιρίων που τέμνουν το αντίστοιχο ΟΤ.

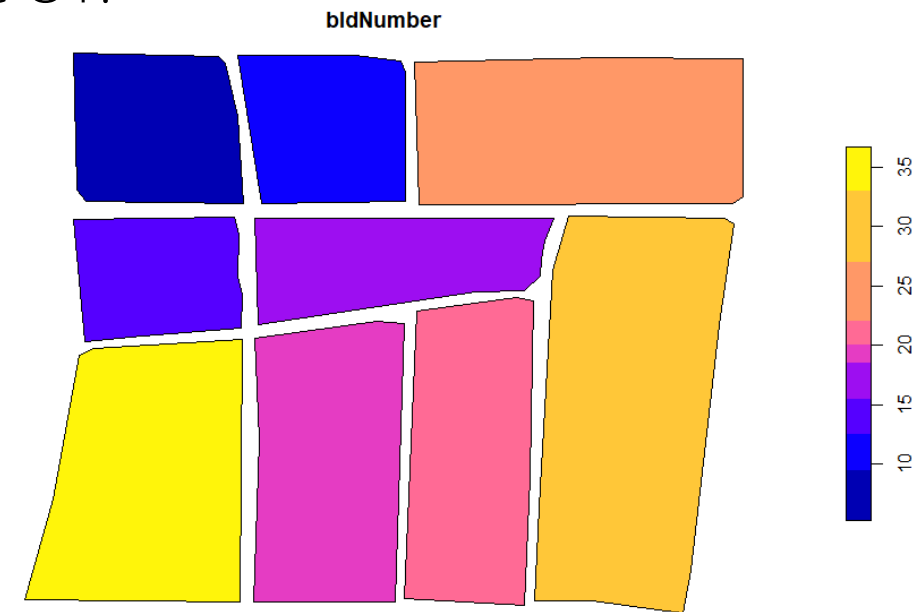
```
89
90 # καταμέτρηση κτιρίων σε κάθε οικοδομικό τετράγωνο
91 # για να γίνει αυτό γίνεται έλεγχος διασταύρωσης (st_intersection)
92 # το αποτέλεσμα αυτής της διαδικασίας είναι μια λίστα.
93 # τα αντικείμενα της λίστας είναι όσα τα οικοδομικά τετράγωνα
94 # για κάθε τετράγωνο επιστρέφει τους κωδικούς των κτιρίων
95 # που τέμνονται με το αντίστοιχο οικοδομικό τετράγωνο
96 # παρακάτω φαίνεται ποια σημεία περιλαμβάνονται στο πρώτο
97 # οικοδομικό τετράγωνο
98 intersection <- st_intersects(src, trg)
99 intersection[[1]]
100
101 # αντίστοιχα για το δεύτερο κ.ο.κ.
102 intersection[[2]]
103
```

```
> # παρακάτω φαίνεται ποια σημεία περιλαμβάνονται στο πρώτο
> # οικοδομικό τετράγωνο
> intersection <- st_intersects(src, trg)
> intersection[[1]]
[1] 7 15 25 34 36 37 38 39 40 41 42 43 44 45 46 106 107 108 109 110 129 130 131 132 133 134 135 136 137
[30] 145 146 154 155 156 157
> # αντίστοιχα για το δεύτερο κ.ο.κ.
> intersection[[2]]
[1] 17 18 19 90 91 92 93 94 95 96 105 111 128 149 150 153 158 159 160
> |
```

03.22 Έλεγχος Διασταύρωσης

Εάν μετρήσουμε το μέγεθος κάθε αντικειμένου της λίστας τότε σαν αποτέλεσμα λαμβάνουμε τον αριθμό κτιρίων που ανήκουν σε κάθε ΟΤ.

```
103
104 # επόμενο βήμα είναι η δημιουργία μιας νέας στήλης με τιμή 0
105 # για όλα τα οικοδομικά τετράγωνα. η στήλη αυτή θα φιλοξενήσει τον
106 # αριθμό των κτιρίων.
107 src$bidNumber <- 0
108 src$bidNumber <- lengths(intersection)
109
110 # εκτύπωση των αποτελεσμάτων
111 src$bidNumber
112
113 # οπτικοποίηση
114 plot(src["bidNumber"])
115
```



03.23 Έλεγχος Διασταύρωσης

Η ίδια διαδικασία μπορεί να γίνει με σημεία, γραμμές ή πολύγωνα. Παρακάτω παρουσιάζεται ακόμα ένα παράδειγμα καταμέτρηση των κτιρίων ως σημεία σε κάθε ΟΤ. Αυτή τη φορά ο έλεγχος διασταύρωσης και η καταμέτρηση γίνεται σε ένα βήμα. Συγκρίνοντας τα αποτελέσματα με το προηγούμενο παράδειγμα (καταμέτρηση πολυγώνων σε πολύγωνα) παρατηρούμε ότι το αποτέλεσμα είναι το ίδιο.

```
116 # καταμέτρηση σημείων σε πολύγωνα
117 # για την καταμέτρηση ουσιαστικά χρησιμοποιείται
118 # η ίδια διαδικασία
119 # αυτή τη φορά θα υλοποιηθεί σε ένα βήμα
120 # με συνδυσμό των μεθόδων
121 src$blDnumberTwo <- lengths(st_intersects(src, trgPoint))
122
123 # εκτύπωση των αποτελεσμάτων στην κονσόλα
124 src$blDnumberTwo
125
126 # σύγκριση με τα αποτελέσματα της διαδικασίας με πολύγωνα
127 src$blDnumber
128
```

```
> # αυτή τη φορά θα υλοποιηθεί σε ένα βήμα
> # με συνδυσμό των μεθόδων
> src$blDnumberTwo <- lengths(st_intersects(src, trgPoint))
> # εκτύπωση των αποτελεσμάτων στην κονσόλα
> src$blDnumberTwo
[1] 35 19 21 18 13 31 23 12 7
> # σύγκριση με τα αποτελέσματα της διαδικασίας με πολύγωνα
> src$blDnumber
[1] 35 19 21 18 13 31 23 12 7
> |
```

03.24 Χωρική Σύζευξη

Είναι κοινό σε Βάσεις Δεδομένων ή σε Υπολογιστικά φύλλα να γίνεται σύζευξη βάση ενός κοινού πεδίου μεταξύ δύο σετ δεδομένων. Η χωρική σύζευξη είναι μια διαδικασία κατά την οποία η σύζευξη των περιγραφικών χαρακτηριστικών των χωρικών οντοτήτων γίνεται με κριτήριο τη θέση τους στο χώρο.

Η χωρική σύζευξη μπορεί να επιτευχθεί με διάφορους τρόπους. Για παράδειγμα εάν ένα πολύγωνο περιέχεται εντός άλλου ή αν τμήμα του περιέχεται, ή έχει απόσταση μικρότερη των 10 μέτρων από ένα σημείο κ.α. Εφόσον πληρούνται αυτά τα κριτήρια τότε η πληροφορία μεταφέρεται από το ένα σετ δεδομένων στο άλλο.

Η χωρική σύζευξη στην R επιτυγχάνεται με την εντολή `st_join()` της `sf`

03.25 Χωρική Σύζευξη

Παρακάτω παρουσιάζεται ένα παράδειγμα χωρικής σύζευξης στην R. Στην προκειμένη περίπτωση κάθε κτίριο πρόκειται να κληρονομήσει τα χαρακτηριστικά του ΟΤ το οποίο τέμνει.

```
R 4.2.2 · D:/sagisrs/Projects/ENIRISST_onGoing_finished/e-learning/practice/
> # χωρική σύζευξη
> # εκτύπωση των στηλών του πίνακα χαρακτηριστικών των κτιρίων
> names(trg)
[1] "tid"      "floors"   "rf"       "geometry"
> # η χωρική σύζευξη (spatial join) είναι μια διαδικασία
> # η οποία βάση της θέσης των αντικειμένων στο χώρο
> # μεταφέρει τα χαρακτηριστικά από τα οικοδομικά τετράγωνα
> # στα κτήρια. Έτσι θα είναι γνωστό ποιο κτήριο ανήκει σε
> # ποιο οικοδομικό τετράγωνο.
> # η χωρική σύζευξη μπορεί να υλοποιηθεί με διάφορους τρόπους.
> # με διασταύρωση, με συμπερίληψη, εντός συγκεκριμένης απόστασης κ.α.
> trgJoin <- st_join(trg, src, join = st_intersects)
> # εκτύπωση των στηλών του πίνακα χαρακτηριστικών των κτιρίων που
> # έχουν πλέον κληρονομήσει τα στοιχεία των ΟΤ
> names(trgJoin)
[1] "tid"      "floors"   "rf"       "sid"      "pop"      "bldNumber" "bldNumberTwo"
[8] "geometry"
> |
```

```
128
129 # χωρική σύζευξη
130 # στήλες του πίνακα χαρακτηριστικών των κτιρίων
131 head(trg)
132
133 # η χωρική σύζευξη (spatial join) είναι μια διαδικασία
134 # η οποία βάση της θέσης των αντικειμένων στο χώρο
135 # μεταφέρει τα χαρακτηριστικά από τα οικοδομικά τετράγωνα
136 # στα κτήρια. Έτσι θα είναι γνωστό ποιο κτήριο ανήκει σε
137 # ποιο οικοδομικό τετράγωνο.
138 # η χωρική σύζευξη μπορεί να υλοποιηθεί με διάφορους τρόπους.
139 # με διασταύρωση, με συμπερίληψη, εντός συγκεκριμένης απόστασης κ.α.
140 trgJoin <- st_join(trg, src, join = st_intersects)
141
142 # επισκόπηση του πίνακα χαρακτηριστικών των κτιρίων που
143 # έχουν πλέον κληρονομήσει τα στοιχεία των ΟΤ
144 head(trgJoin)
145
146 # οπτικοποίηση του οικοδομικού τετραγώνου των κτιρίων
147 plot(trgJoin["sid"])
148 |
```

03.25 Χωρική Σύζευξη

Οπτικοποίηση των αποτελεσμάτων της χωρικής σύζευξης.

