



Σχολή Θετικών Επιστημών και Τεχνολογίας
Συστήματα Κινητού και Διάχυτου Υπολογισμού

Διπλωματική Εργασία

Οπτική Αναλυτική και Εκτίμηση Γεωχωρικών Δεδομένων

Visual Analytics and Evaluation of Geospatial Data

Θεόδωρος Δημητρίου

Επιβλέπουσα καθηγήτρια: Ευγενία Αδαμοπούλου

Πάτρα, Σεπτέμβριος 2022

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Δημητρίου Θεόδωρου («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



Οπτική Αναλυτική και Εκτίμηση Γεωχωρικών Δεδομένων

Visual Analytics and Evaluation of Geospatial Data

Θεόδωρος Δημητρίου

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπουσα Καθηγήτρια:

Ευγενία Αδαμοπούλου

Καθηγήτρια Ε.Α.Π.

Συν-Επιβλέπων Καθηγητής:

Δαμιανός Γαβαλάς

Επιστημονικός συνεργάτης Ε.Α.Π.

Πάτρα, Σεπτέμβριος 2022

Ευχαριστίες

Ευχαριστώ θερμά,

Την καθηγήτρια κ. Αδαμοπούλου Ευγενία για την εμπιστοσύνη που μου έδειξε αναθέτοντας μου το συγκεκριμένο θέμα αλλά και για την συνεργασία μας κατά την εκπόνηση της παρούσης. Ευχαριστώ επίσης τον κ. Γαβαλά Δαμιανό για την συνεργασία μας και τις παρατηρήσεις του.

Περίληψη

Η σωστή διαχείριση των πληροφοριών που σχετίζονται με την γη είναι σημαντική για την βιώσιμη ανάπτυξη και την αποτελεσματική διαμόρφωση πολιτικών, και ειδικότερα για εφαρμογές όπως ο πολεοδομικός σχεδιασμός, η διαχείριση ακινήτων, ο δανεισμός, η φορολογία και οι επενδύσεις. Σημαντικά εργαλεία στον τομέα αυτόν είναι τα Συστήματα Γεωγραφικών Πληροφοριών (GIS), που βοηθούν στην καλύτερη κατανόηση γεωχωρικών δεδομένων και στη συνέχεια σε πιο επιτυχημένες αποφάσεις γύρω από τα ακίνητα. Ωστόσο, διανύουμε την περίοδο των Big Data που χαρακτηρίζεται από την εκτίναξη του όγκου των διαθέσιμων δεδομένων, όπου η δυνατότητα συλλογής δεδομένων δεν έχει ακολουθηθεί από την ικανότητα αξιοποίησής τους. Αυτός είναι ένας από τους λόγους ανάπτυξης του τομέα της Οπτικής Αναλυτικής (Visual Analytics), που χρησιμοποιεί διαδραστική απεικόνιση για να ενσωματώσει την ανθρώπινη κρίση σε αλγοριθμικές διαδικασίες ανάλυσης δεδομένων. Στην εργασία αυτή, μελετάται η Οπτική Αναλυτική σε γεωχωρικά δεδομένα, με σκοπό τον προσδιορισμό της καταλληλότητας γεωγραφικών περιοχών για την ικανοποίηση συγκεκριμένων αναγκών. Αφού πραγματοποιήθηκε βιβλιογραφική επισκόπηση, αναπτύχθηκε διαδραστική εφαρμογή, για την κατανόηση των αναγκών και την υποστήριξη αποφάσεων για την αγορά ή ενοικίαση ακινήτων. Μια σημαντική προδιαγραφή της υλοποίησης, ήταν η χρήση ανοικτών τεχνολογιών και δεδομένων, για τα πολλά πλεονεκτήματα που προσφέρουν στην ανάπτυξη ερευνητικών εφαρμογών.

Λέξεις – Κλειδιά

Οπτική Αναλυτική, Γεωοπτική Αναλυτική, Μεγάλα Δεδομένα, Οπτικοποίηση Πληροφοριών, Διαδραστικές Εφαρμογές, Αντίληψη, Αναλυτικός Συλλογισμός, Συστήματα Γεωγραφικών Πληροφοριών, Εξερευνητική Ανάλυση Γεωχωρικών Δεδομένων, Μοντέλα Αυτοματοποιημένης Αποτίμησης, Συστήματα Υποστήριξης Χωρικών Αποφάσεων, Υποστήριξη Λήψης Αποφάσεων Πολλαπλών Κριτηρίων, Ασαφή Σύνολα, Τεχνολογίες Ανοικτού Κώδικα, Ανοικτά Δεδομένα

Abstract

Proper management of land-related information is important for sustainable development and effective policy making, in particular for applications such as urban planning, property management, lending, taxation and investment. Important tools in this area are Geographic Information Systems (GIS), which help to better understand geospatial data and subsequently to make more successful decisions around real estate. However, we are in the era of Big Data, characterised by the explosion in the volume of data available, where the ability to collect data has not been followed by the ability to use it. This is one of the reasons for the growth of the field of Visual Analytics, which uses interactive visualization to integrate human judgment into algorithmic data analysis processes. In this paper, Visual Analytics is studied on geospatial data, in order to determine the ability of geographic locations to meet specific needs. After conducting a literature review, an interactive application was developed to understand the needs and support decisions for buying or renting real estate. An important specification of the implementation, was the use of open technologies and data, for the many advantages they offer in the development of research applications.

Keywords

Visual Analytics, Geovisual Analytics, Big Data, Information Visualization, Interactive Applications, Perception, Analytic Reasoning, Geographic Information Systems, Exploratory Geospatial Data Analysis, Automated Valuation Models, Spatial Decision Support Systems, Multi-criteria Decision Support Systems, Fuzzy Sets, Open Source Technologies, Open Data

Περιεχόμενα

Ευχαριστίες.....	4
Περίληψη.....	5
Abstract.....	6
Περιεχόμενα.....	7
Κατάλογος Εικόνων / Σχημάτων.....	9
Κατάλογος Πινάκων.....	11
Συντομογραφίες & Ακρωνύμια.....	12
1 Εισαγωγή.....	1
1.1 Δομή εργασίας.....	3
2 Βιβλιογραφική επισκόπηση.....	4
2.1 Συστήματα Γεωγραφικών Πληροφοριών (GIS).....	4
2.2 Συστήματα Υποστήριξης Χωρικών Αποφάσεων (SDSS).....	7
2.3 Οπτική Αναλυτική.....	8
2.4 Γεωοπτική Αναλυτική.....	13
2.4.1 Οπτικοποίηση.....	14
Οπτικές μεταβλητές.....	19
Διαγράμματα.....	21
Χάρτες.....	25
2.4.2 Διάδραση.....	31
2.4.3 Αναλυτικός συλλογισμός (analytical reasoning).....	34
2.4.4 Υπολογιστικές μέθοδοι.....	36
2.4.5 Κλιμακωσιμότητα (Scalability).....	37
2.5 Μοντέλα Αυτοματοποιημένης Αποτίμησης (Automated Valuation Model).....	39
2.6 Γεωχωρικά Δεδομένα.....	41
2.6.1 Διαθεσιμότητα.....	42
2.6.2 Δομή και Μορφή.....	43
2.6.3 Αποδελτίωση (Scraping).....	43
2.7 Χωρικά κριτήρια επιλογής ακινήτου.....	44
2.8 Σχεδιαστικά κριτήρια εφαρμογών.....	46
2.9 Κριτήρια αξιολόγησης εφαρμογών.....	48
2.10 Τεχνολογική στάθμιση.....	50
2.10.1 Εργασίες.....	50
HomeSeeker.....	50
AOI-shapes.....	52
Fuzzy spatio-temporal querying.....	53
Smart Real Estate.....	55
2.10.2 Εργαλεία.....	56
GeoDa.....	56
PostgreSQL/PostGIS.....	57
QGIS.....	58
Apache Superset.....	59
WebGLayer.....	61

HSLayers NG.....	62
3 Εφαρμογή γεωοπτικής αναλυτικής για την αγορά ακινήτων.....	64
3.1 Σκοπός και ερευνητικό ερώτημα.....	64
3.2 Αρχιτεκτονική εφαρμογής.....	65
3.3 Αλγόριθμος υπολογισμού δείκτη χωρικών κριτηρίων.....	67
3.4 Προετοιμασία δεδομένων.....	69
3.4.1 Δημόσια διαθέσιμα δεδομένα για σημεία ενδιαφέροντος.....	70
3.4.2 Δεδομένα από την αγορά ακινήτων.....	71
3.4.3 Παραγωγή πλέγματος διατεταγμένων σημείων.....	71
3.5 Λειτουργία της εφαρμογής.....	74
3.5.1 Εισαγωγή και παραμετροποίηση κριτηρίων.....	75
3.5.2 Εξερευνητική αλληλεπίδραση με τον χάρτη και με τα στοιχεία ελέγχου.....	78
3.6 Αξιολόγηση της εφαρμογής.....	81
3.6.1 Αποτελέσματα.....	82
4 Επίλογος.....	84
4.1 Συνεισφορά.....	84
4.2 Επόμενα βήματα.....	86
Βιβλιογραφικές Αναφορές.....	88

Κατάλογος Εικόνων / Σχημάτων

Εικόνα 2.1: Υπέρθυση γεωχωρικών δεδομένων (Kraak & Ormeling, 2010).....	5
Εικόνα 2.2: Η ροή εργασίας της Οπτικής Αναλυτικής (Keim et al., 2010a).....	11
Εικόνα 2.3: Χάρτης των θανάτων από χολέρα το 1854 στο Λονδίνο από τον John Snow (Wikimedia Commons).....	15
Εικόνα 2.4: Χάρτης της καταστροφικής εκστρατείας του Ναπολέοντα στην Ρωσία το 1812 από τον Charles Joseph Minard (Wikimedia Commons).....	16
Εικόνα 2.5: Η τετράδα του Anscombe.....	17
Εικόνα 2.6: Μια συλλογή παραγόμενων συνόλων δεδομένων, διαφορετικών στην εμφάνιση, αλλά με τα ίδια περιληπτικά στατιστικά, με απόκλιση 2 δεκαδικών ψηφίων. (Matejka & Fitzmaurice, 2017).....	18
Εικόνα 2.7: Το σύστημα των αντιληπτικών μεταβλητών του Bertin (Meirelles, 2013, p127)	20
Εικόνα 2.8: a) Διάγραμμα διασποράς, b) Διάγραμμα διασποράς με ομαδοποίηση σε δοχεία.	22
Εικόνα 2.9: Πίνακας διαγραμμάτων διασποράς με ιστογράμματα στην διαγώνιο.....	22
Εικόνα 2.10: Ραβδογράμματα (Ware, 2020, p430).....	24
Εικόνα 2.11: Διάγραμμα παράλληλων συντεταγμένων (Li, 2019, p47).....	25
Εικόνα 2.12: Σύνθεση θεματικού χάρτη (Kraak & Ormeling, 2010).....	26
Εικόνα 2.13: Χάρτης με σύμβολα τύπου ραντάρ με μορφή πίτας (Li, 2019, p46).....	28
Εικόνα 2.14: Χάρτης με ραβδογράμματα για σύμβολα σημείων (Andrienko et al., 2020, p12).....	28
Εικόνα 2.15: Χάρτες: (a) κουκίδων, (b) χωροπληθής, (c) ισαριθμικός (Kraak & Ormeling, 2010, p142).....	29
Εικόνα 2.16: Θερμικοί χάρτες (Kethireddy et al., 2018).....	30
Εικόνα 2.17: Συνδεδεμένες ή Συγχρονισμένες Προβολές (Andrienko et al., 2020, p84).....	34
Εικόνα 2.18: Προβλήματα κλιμάκωσης της ποσότητας των δεδομένων: (a) υπερβολική χαρτογράφηση (b) υπέρβαση ικανοτήτων χρηστών (c) δυσκολία διάδρασης (Li, 2019, p7).....	39
Εικόνα 2.19: Χάρτες απεικόνισης προφίλ περιοχών (a) χωροπληθείς και (b) κουκίδων (Li, 2019).....	51
Εικόνα 2.20: Οθόνη προαστίων (a) μπάρα πλοήγησης (b) χάρτης κουκίδων (c) προβολή ιστορικού με πολλαπλές γραμμές (d) προβολή ιστορικού με γράφημα ροής (e) διάγραμμα παράλληλων συντεταγμένων με ιστογράμματα (Li, 2019).....	51
Εικόνα 2.21 Οθόνη ακινήτων (a) μπάρα πλοήγησης (b) χάρτης συμβόλων με πολύγωνα (c) πολυδιάστατη προβολή δεδομένων (d) κάρτα με εικόνες (e) σύννεφο λέξεων (f) διάγραμμα αράχνης (Li, 2019).....	52
Εικόνα 2.22: Σχηματισμός συνόρων περιοχής ενδιαφέροντος με τον αλγόριθμο AOI-shapes (Li, 2019).....	53
Εικόνα 2.23: Ασαφής γεωχωροχρονική αναζήτηση (Đuračioná & Faixová Chalachanová, 2017).....	54
Εικόνα 2.24: Δημιουργία υπόβαθρου περιβαλλοντικών παραγόντων (Aydinoglu & Bonkir, 2020).....	55
Εικόνα 2.25: Οθόνη Εξερευνητικής Ανάλυσης Γεωχωρικών Δεδομένων στην εφαρμογή GeoDa (https://geodacenter.github.io).....	57
Εικόνα 2.26: Το QGIS 3.26 Desktop.....	59

Εικόνα 2.27: Superset, συλλογή διαθέσιμων διαγραμμάτων (https://github.com/apache/superset).....	60
Εικόνα 2.28: Superset, σύνθετο και διαδραστικό ταμπλό με χάρτη και διαγράμματα (https://docs.preset.io/docs/dashboard-properties).....	60
Εικόνα 2.29: Οθόνη εφαρμογής που αναπτύχθηκε με το WebGLayer (https://github.com/WebGLayer/webglayer/wiki/User-Guide).....	62
Εικόνα 2.30: Εφαρμογή που αναπτύχθηκε με το HSLayers NG (https://ng.hslayers.org).....	63
Εικόνα 3.1: Διάγραμμα αρχιτεκτονικής της εφαρμογής.....	67
Εικόνα 3.2: Συναρτήσεις ασαφούς συμμετοχής: τμηματική γραμμική (αριστερά) και λογιστική (δεξιά).....	69
Εικόνα 3.3: Απεικόνιση σημείων πλέγματος με θερμικό χάρτη (αριστερά), περιγράμματα (κέντρο) και εξάγωνα σύμβολα (δεξιά).....	73
Εικόνα 3.4: Η οθόνη της εφαρμογής.....	75
Εικόνα 3.5: Μενού εισαγωγής κριτηρίων.....	75
Εικόνα 3.6: Φόρμα παραμετροποίησης κριτηρίου.....	76
Εικόνα 3.7: Σενάριο χρήσης της εφαρμογής.....	77
Εικόνα 3.8: Καρτέλα ακινήτων.....	78
Εικόνα 3.9: Μεταβολή υπόβαθρου δείκτη μετά από εκλέπτυση κριτηρίων.....	79
Εικόνα 3.10: Εμφάνιση σταθμών ΜΕΤΡΟ (πράσινοι ρόμβοι) και σουπερμάρκετ (κόκκινοι κύκλοι).....	79
Εικόνα 3.11: Εμφάνιση αγγελιών ακινήτων σε μορφή κουκίδων.....	80
Εικόνα 3.12: Εμφάνιση αγγελιών ακινήτων σε μορφή ραβδογραμμάτων.....	81

Κατάλογος Πινάκων

Πίνακας 2.1: Σημαντικοί όροι που σχετίζονται με την Οπτική Αναλυτική (Cui, 2019).....	9
---	---

Συντομογραφίες & Ακρωνύμια

AHP	Διαδικασία Αναλυτικής Ιεράρχησης (Analytic Hierarchy Process)
AOI	Περιοχή Ενδιαφέροντος (Area of Interest)
AVM	Μοντέλο Αυτοματοποιημένης Αποτίμησης (Automated Valuation Model)
DBMS	Σύστημα Διαχείρισης Βάσεων Δεδομένων (Database Management System)
ESDA Analysis)	Εξερευνητική Ανάλυση Χωρικών Δεδομένων (Exploratory Spatial Data Analysis)
GIS	Σύστημα Γεωγραφικών Πληροφοριών (Geographic Information System)
GVA	Γεωοπτική Αναλυτική (Geovisual Analytics)
MCDA Analysis)	Ανάλυση Αποφάσεων Πολλαπλών Κριτηρίων (Multiple-criteria Decision Analysis)
OGC	Open Geospatial Consortium
POI	Σημείο Ενδιαφέροντος (Point of Interest)
SDSS System)	Σύστημα Υποστήριξης Χωρικών Αποφάσεων (Spatial Decision Support System)
SQL	Structured Query Language
OA	Οπτική Αναλυτική

Μεταφράσεις όρων

Background map	Χαρτογραφικό υπόβαθρο
Bar chart	Ραβδόγραμμα
Choropleth	Χωροπληθής ή Χωροπληθικός
Data mining	Εξόρυξη δεδομένων
Fuzzy logic	Ασαφής λογική
Geovisulization	Γεωοπτικοποίηση
GIScience	Επιστήμη Γεωγραφικών Πληροφοριών
Glyph	Γλύφος ή Σύμβολο
Heat map	Θερμικός χάρτης
Landmark	Τοπόσημο
Multivariate	Πολυμεταβλητός (λέγεται και Bivariate)
Outlier	Ακραία τιμή

1 Εισαγωγή

Η σωστή διαχείριση των πληροφοριών που σχετίζονται με την γη είναι σημαντική για την βιώσιμη ανάπτυξη και την αποτελεσματική διαμόρφωση πολιτικών, και ειδικότερα για εφαρμογές όπως ο πολεοδομικός σχεδιασμός, η διαχείριση ακινήτων, ο δανεισμός, η φορολογία και οι επενδύσεις. Ειδικότερα, για την αγορά ακινήτων, πρόσφατες μελέτες έχουν δείξει ότι, παρά την αναζήτηση σε διαδικτυακές πλατφόρμες και την αφιέρωση αρκετού χρόνου για έρευνα αγοράς, σχεδόν οι μισοί εξ όσων αγοράζουν ή ενοικιάζουν ακίνητα μετανιώνουν για την επιλογή τους (Li, 2019, p1). Ένας από τους σημαντικότερους παράγοντες για το πρόβλημα αυτό είναι έλλειψη πληροφόρησης και ανάλυσης από τις διαδικτυακές πλατφόρμες, οι οποίες περιορίζονται στην αναζήτηση με βάση την περιοχή ή τα χαρακτηριστικά του ακινήτου, αποτυγχάνοντας να παράσχουν εργαλεία για την ανάλυση, την σύγκριση και την αξιολόγηση των αγορών και των τοπικών χαρακτηριστικών.

Σημαντικά εργαλεία στον τομέα αυτόν είναι τα Συστήματα Γεωγραφικών Πληροφοριών (GIS), που αυξάνουν την ακρίβεια των εκτιμήσεων και αξιολογήσεων της γης και των ακινήτων, και χρησιμοποιούνται λόγω των δυνατοτήτων τους να αναλύσουν διάφορα χαρακτηριστικά του φυσικού, τοπογραφικού και κοινωνικού περιβάλλοντος μέσα από γεωχωρικά δεδομένα, ώστε να βοηθήσουν τον χρήστη στην καλύτερη κατανόηση των δεδομένων και στη συνέχεια σε πιο επιτυχημένες αποφάσεις γύρω από τα ακίνητα.

Ωστόσο, διανύουμε την περίοδο των μεγάλων δεδομένων (Big Data), όπου η ικανότητά μας να συλλέγουμε δεδομένα δεν έχει ακολουθηθεί από την δυνατότητά μας να αποκομίζουμε χρήσιμες πληροφορίες από αυτά. Αυτό το κενό έχει οδηγήσει στην ανάπτυξη του τομέα της Οπτικής Αναλυτικής (Visual Analytics), που χρησιμοποιεί διαδραστική απεικόνιση για να ενσωματώσει την ανθρώπινη κρίση σε αλγοριθμικές διαδικασίες ανάλυσης δεδομένων. Η εξειδίκευση της οπτικής αναλυτικής σε γεωχωρικά δεδομένα είναι ο τομέας της Γεωοπτικής Αναλυτικής (Geovisual Analytics), ο οποίος αναζητά νέα διαδραστικά συστήματα χαρτογράφησης που επιτρέπουν στους χρήστες να εντοπίζουν μοτίβα και να προβλέπουν μελλοντικά αποτελέσματα χρησιμοποιώντας χωρικά δεδομένα.

Στο πλαίσιο αυτό, στην παρούσα εργασία, μελετήθηκε η ανάλυση γεωχωρικών δεδομένων για γεωγραφικούς τόπους, με στόχο την αξιολόγηση της δυνατότητάς τους να ικανοποιήσουν συγκεκριμένες ανθρώπινες ανάγκες, χρησιμοποιώντας Γεωοπτική Αναλυτική. Για τον σκοπό αυτό πραγματοποιήθηκε αρχικά βιβλιογραφική επισκόπηση για την ανακάλυψη της θεωρίας και της επιστημονικής και τεχνολογικής στάθμησης. Στην συνέχεια σχεδιάστηκε και υλοποιήθηκε εφαρμογή γεωοπτικής αναλυτικής, με την οποία ο χρήστης μπορεί να κατανοήσει και να προσδιορίσει τις απαιτήσεις του, και στην συνέχεια να εξερευνήσει την αγορά ακινήτων, χρησιμοποιώντας γεωχωρικά κριτήρια όπως αποστάσεις από σημεία ενδιαφέροντος, εγγύτητα με τοποθεσίες και συμπερίληψη σε περιοχές με συγκεκριμένα χαρακτηριστικά, όπως δήμοι, ζώνες τιμών αντικειμενικών αξιών ή χρήσεις γης.

Για την υλοποίηση χρησιμοποιήθηκαν ανοικτές τεχνολογίες και δεδομένα, για τα πολλά πλεονεκτήματα που προσφέρουν στην ανάπτυξη ερευνητικών εφαρμογών, όπως η απεριόριστη δυνατότητα προσαρμογής, το κόστος, οι ελεύθερες άδειες χρήσεις κτλ. Συγκεκριμένα, η εφαρμογή αναπτύχθηκε με web τεχνολογίες και με το QGIS, ενώ για την εκτέλεση του αλγόριθμου και την επεξεργασία των γεωχωρικών δεδομένων χρησιμοποιήθηκε το σύστημα βάσης δεδομένων PostgreSQL, με εγκατεστημένο το πρόσθετο PostGIS. Τα γεωχωρικά δεδομένα για τα χαρακτηριστικά των περιοχών αντλήθηκαν από το Openstreetmap, ενώ έγιναν προσπάθειες να αξιοποιηθούν και δημόσιες πηγές όπως η ΕΛΣΤΑΤ, το GEODATA.gov.gr, όμως το επίπεδο λεπτομέρειας δεν ήταν κατάλληλο για τους στόχους της εφαρμογής. Ως εξαίρεση, για την επίδειξη της δυνατότητας ενσωμάτωσης με άλλες τεχνολογίες και πλατφόρμες, αλλά και για την καλύτερη αξιολόγηση της εφαρμογής, αντλήθηκαν, με διάφορες τεχνικές, δεδομένα από την τρέχουσα αγορά ακινήτων.

Η εφαρμογή αξιολογήθηκε με πραγματικούς χρήστες με στεγαστικές ανάγκες, με κριτήρια την ευχρηστία της, την βελτίωση της επίγνωσης της επίδρασης της τοποθεσίας και την υποστήριξη στη λήψη αποφάσεων.

Οι κύριες συνεισφορές της εργασίας είναι ένα διαδραστικό εργαλείο και ένας αλγόριθμος. Το διαδραστικό εργαλείο γεωοπτικής αναλυτικής χρησιμοποιεί εξερευνητική ανάλυση γεωχωρικών δεδομένων για την υποστήριξη λήψης αποφάσεων. Ο αλγόριθμος υπολογίζει έναν δείκτη καταλληλότητας τοποθεσιών και καλείται από την εφαρμογή, ώστε τα

αποτελέσματά του να παράγουν άμεση ανατροφοδότηση στον χρήστη, στο πλαίσιο της εξερευνητικής διαδικασίας. Ορισμένα σημαντικά στοιχεία διαφοροποίησης από παρεμφερείς υλοποιήσεις, τα οποία δεν βρίσκονται συνδυαστικά σε άλλες εφαρμογές, είναι: α) η επιλογή των κριτηρίων από τον ίδιο τον χρήστη και η δυνατότητα παραμετροποίησής τους, β) η δυνατότητα επιλογής κριτηρίων αρνητικού συντελεστή, όπως η απόσταση από πηγές ηχορύπανσης, γ) η χρήση βαρών, για ορισμό προτεραιοτήτων στα κριτήρια, δ) ο υπολογισμός του δείκτη για μια ολόκληρη περιοχή και όχι μόνο για συγκεκριμένα σημεία ενδιαφέροντος ε) η απεικόνιση του δείκτη στον χάρτη με χρωματικές διαβαθμίσεις, αντί για γραμμική οριοθέτηση, περιοχών που ικανοποιούν τα κριτήρια και στ) η δυνατότητα προσθήκης της τιμής του δείκτη που παράγεται από τον αλγόριθμο, ως νέο χαρακτηριστικό σε άλλα σύνολα δεδομένων, όπως αγγελίες ακινήτων ή σημεία ενδιαφέροντος.

1.1 Δομή εργασίας

Το παρόν Κεφάλαιο 1 αποτελεί την εισαγωγή, όπου αναφέρονται περιληπτικά το αντικείμενο, η προβληματική, ο σκοπός, η εφαρμογή, η συνεισφορά και η δομή της εργασίας. Στην συνέχεια ακολουθούν άλλα τρία κεφάλαια.

Στο Κεφάλαιο 2 παρουσιάζονται τα Συστήματα Γεωγραφικών Πληροφοριών (GIS), τα Συστήματα Υποστήριξης Χωρικών Αποφάσεων (SDSS), η Οπτική Αναλυτική, η Γεωοπτική Αναλυτική, τα Μοντέλα Αυτοματοποιημένης Αποτίμησης (AVM), τα Γεωχωρικά Δεδομένα, τα χωρικά κριτήρια επιλογής ακινήτων, τα σχεδιαστικά κριτήρια εφαρμογών και τα κριτήρια αξιολόγησης εφαρμογών. Επίσης, αναλύονται οι σημαντικότερες εργασίες και εφαρμογές γεωοπτικής αναλυτικής.

Το Κεφάλαιο 3 είναι αφιερωμένο στην ανάπτυξη της εφαρμογής, όπου περιγράφονται ο σκοπός και το ερευνητικό ερώτημα, η αρχιτεκτονική της εφαρμογής και ο αλγόριθμος που αναπτύχθηκε. Επίσης, περιγράφονται η προετοιμασία των δεδομένων, η λειτουργία της εφαρμογής και η αξιολόγησή της.

Το Κεφάλαιο 4 αποτελεί τον επίλογο, όπου περιλαμβάνονται η συνεισφορά της διπλωματικής εργασίας και προτάσεις για επόμενα βήματα.

Η εργασία ολοκληρώνεται με την παράθεση των βιβλιογραφικών αναφορών.

2 Βιβλιογραφική επισκόπηση

2.1 Συστήματα Γεωγραφικών Πληροφοριών (GIS)

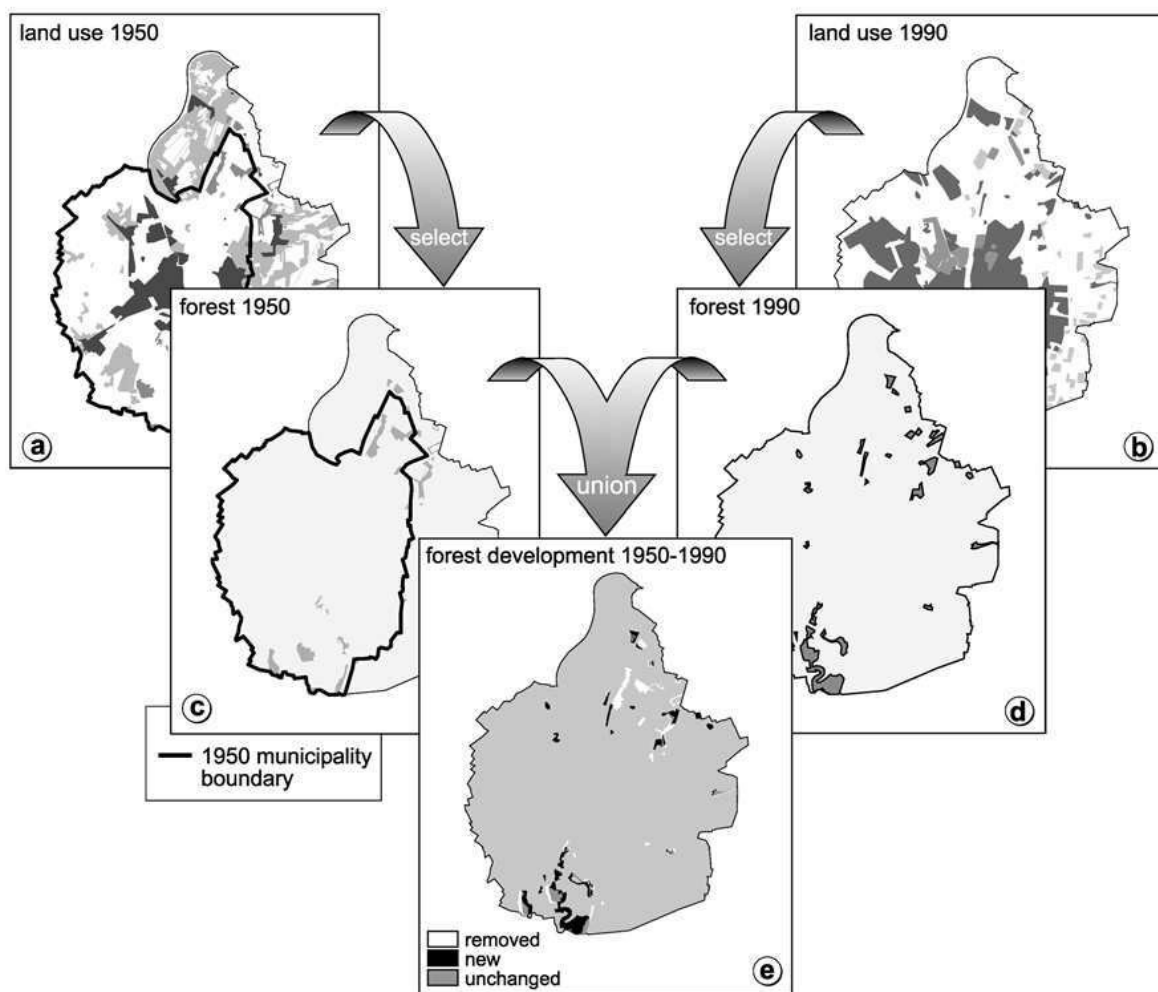
Ένα σύστημα γεωγραφικών πληροφοριών (Geographic Information System) είναι ένα πληροφοριακό σύστημα για τη σύλληψη, αποθήκευση, αναζήτηση, ανάλυση και προβολή γεωχωρικών δεδομένων. Οι συνήθεις διεπαφές χρήστη σε αυτά τα προγράμματα είναι μενού, εικονίδια και γραμμές εντολών, σε λειτουργικά συστήματα όπως Windows ή Linux.

Η απόκτηση δεδομένων είναι συνήθως το πρώτο βήμα στην υλοποίηση ενός έργου GIS. Για τον λόγο αυτό, από τις αρχές της δεκαετίας του 1990, κυβερνητικές υπηρεσίες σε πολλές χώρες, έχουν δημιουργήσει δικτυακούς τόπους για την ανταλλαγή δημόσιων δεδομένων, όπως το Data.gov στις ΗΠΑ και το INSPIRE στην Ευρωπαϊκή Ένωση. Για την αξιοποίηση των δεδομένων χρειάζονται μεταδεδομένα, για την περιγραφή γεωγραφικών πληροφοριών και υπηρεσιών. Τα μεταδεδομένα αυτά σχετίζονται με τον προσδιορισμό, την έκταση, την ποιότητα, τις χωρικές και χρονικές πτυχές, το περιεχόμενο, τη χωρική αναφορά, την απεικόνιση, τη διανομή και άλλες ιδιότητες των ψηφιακών γεωγραφικών δεδομένων και υπηρεσιών.

Τα συστήματα διαχείρισης βάσεων δεδομένων (DBMS) αξιοποιούνται συχνά από τα GIS για τον χειρισμό των διανυσματικών δεδομένων, τα οποία μπορεί να είναι μεγάλα σε μέγεθος. Ο πίνακας χαρακτηριστικών γνωρισμάτων (feature attribute table) είναι ένας σημαντικός τύπος πίνακα, που αναπαριστά τα χωρικά γνωρίσματα (spatial features) ανά γραμμή και τα χαρακτηριστικά (attributes) ανά στήλη.

Υπάρχουν δύο είδη ερωτημάτων στα δεδομένα, που επιτρέπουν την επιλογή υποσυνόλων για πιο λεπτομερή εξέταση. Τα ερωτήματα με βάση τα χαρακτηριστικά, κατά κανόνα χρησιμοποιούν μια γλώσσα βασισμένη στην SQL (structured query language) και επιλέγουν χωρικά γνωρίσματα βάσει τιμών των χαρακτηριστικών τους. Αντίθετα, τα χωρικά ερωτήματα, επιλέγουν γνωρίσματα με βάση τις χωρικές τους σχέσεις και όχι τις τιμές των χαρακτηριστικών τους. Υπάρχουν τριών ειδών σχέσεις μεταξύ γνωρισμάτων, βάσει των οποίων μπορούν να δημιουργηθούν ερωτήματα επιλογής: η συμπερίληψη, η διασταύρωση και η εγγύτητα. (Chang, 2016)

Οι διεργασίες υπέρθεσης (overlay) αποτελούν μέρος των περισσότερων διαδικασιών χωρικής ανάλυσης και συνήθως αποτελούν τον πυρήνα των έργων GIS. Κατά την υπέρθεση, γίνεται συνδυασμός διαφορετικών χαρτών, με αποτέλεσμα να δημιουργούνται νέα χωρικά στοιχεία, που δεν υπήρχαν στους επιμέρους χάρτες (Εικόνα 2.1). Πολλοί χάρτες μπορούν να συνδυαστούν ταυτόχρονα, χρησιμοποιώντας αριθμητικούς, σχεσιακούς ή υπό όρους τελεστές και πολλές διαφορετικές συναρτήσεις.



Εικόνα 2.1: Υπέρθεση γεωχωρικών δεδομένων (Kraak & Ormeling, 2010)

Η ανάλυση περιβαλλουσών ζωνών (buffer analysis) χρησιμοποιείται για τον προσδιορισμό των περιοχών που περιβάλλουν γεωγραφικά χαρακτηριστικά. Η διαδικασία περιλαμβάνει την δημιουργία μιας ζώνης γύρω από τα υπάρχοντα γεωγραφικά στοιχεία και στη συνέχεια τον

εντοπισμό ή την επιλογή στοιχείων με βάση το αν βρίσκονται εντός ή εκτός των ορίων της ζώνης.

Η διαδραστική απεικόνιση μεγάλου όγκου πληροφοριών είναι ένα από τα πιο ενδιαφέροντα και χρήσιμα χαρακτηριστικά των GIS. Οι σύγχρονοι υπολογιστές μπορούν να μεταβάλλουν άμεσα την εμφάνιση αλλά και την ερμηνεία οποιουδήποτε στοιχείου στην οθόνη, δίνοντας στα GIS πολλά πλεονεκτήματα σε σύγκριση με τους παραδοσιακούς χάρτινους χάρτες. Ο χρήστης αποκτά υψηλό επίπεδο ευελιξίας και ισχύος με μια σειρά από διαφορετικές δυνατότητες, όπως η προσθαφαίρεση πεδίων δεδομένων, η αλλαγή του χρωματισμού ή του σχήματος του χάρτη, η προσθήκη κειμένου ή η μετακίνηση συμβόλων. Οι οθόνες του GIS είναι επεκτάσιμες και μετακινούμενες, δίνοντας στον χρήστη νέες προοπτικές, περισσότερες (ή λιγότερες) λεπτομέρειες και μια νέα αντίληψη των πραγμάτων. (Srivastava & Khot, 2018)

Τα διαδικτυακά GIS ή οι διαδικτυακές υπηρεσίες χαρτογράφησης (web GIS ή web mapping) επιτρέπουν στους χρήστες να βρίσκουν δεδομένα, να έχουν πρόσβαση σε υπηρεσίες και να δημοσιεύουν χάρτες για να τους μοιράζονται με άλλους. Κινητές συσκευές, όπως smartphones και tablets με GPS, μπορούν να χρησιμοποιηθούν για την εκτέλεση πολλών εργασιών που σχετίζονται με τα GIS, όπως η λήψη χαρτών από έναν διακομιστή, η συλλογή δεδομένων σε πραγματικό χρόνο στο πεδίο, η επεξεργασία των χαρτών και η αποθήκευση των χαρτών μέσω μιας υπηρεσίας αποθήκευσης στο νέφος (cloud). Πιο πρόσφατη είναι η ανάπτυξη της συνεργατικής διαδικτυακής χαρτογράφησης και των εθελοντικών γεωγραφικών πληροφοριών (VGI), όπως το OpenStreetMap (Chang, 2016).

Η επιστήμη των γεωγραφικών πληροφοριών (GIScience) ασχολείται με τη φύση των γεωγραφικών πληροφοριών και των γεωγραφικών φαινομένων παρέχοντας θεωρητικές βάσεις για τα GIS και τις σχετικές τεχνολογίες γεωγραφικών πληροφοριών. Στόχος της GIScience είναι η τυποποίηση των γεωγραφικών αρχών προκειμένου να διερευνηθούν οι επιστημονικές και πολιτικές εφαρμογές των γεωγραφικών πληροφοριών και να αποκαλυφθούν και να αναλυθούν οι περίπλοκες σχέσεις που έχουν οι άνθρωποι, οι οργανισμοί και η κοινωνία με τις τεχνολογίες γεωγραφικών πληροφοριών. Η GIScience επιδιώκει να απαντήσει σε θεμελιώδη ερωτήματα σχετικά με τη χρήση των GIS ως εργαλεία υποστήριξης αποφάσεων. Παρόλο που τα GIS θεωρούνται συμβατικά ως ένα σύνολο

εργαλείων για την εισαγωγή, αποθήκευση και ανάκτηση, χειρισμό και ανάλυση και έξοδο χωρικών δεδομένων, περιέχουν και ένα σύνολο διαδικασιών για την υποστήριξη δραστηριοτήτων λήψης αποφάσεων (Malczewski & Rinner, 2015).

2.2 Συστήματα Υποστήριξης Χωρικών Αποφάσεων (SDSS)

Ένα Σύστημα Υποστήριξης Χωρικών Αποφάσεων (Spatial Decision Support System) είναι ένα διαδραστικό υπολογιστικό σύστημα, που έχει σχεδιαστεί για να υποστηρίζει έναν χρήστη ή μια ομάδα χρηστών στην επίτευξη μεγαλύτερης αποτελεσματικότητας στη λήψη αποφάσεων κατά την επίλυση ενός ημιδομημένου χωρικού προβλήματος λήψης αποφάσεων. Αυτό μπορεί να είναι ένα πρόβλημα χωροθέτησης-κατανομής, αναζήτησης και επιλογής χώρου, αξιολόγησης καταλληλότητας χρήσεων γης, μεταφορών, αξιολόγησης περιβαλλοντικών επιπτώσεων και αξιολόγησης σχεδίων/πολιτικών. Το δομημένο μέρος του ημιδομημένου προβλήματος μπορεί να επιδέχεται αυτοματοποιημένη επίλυση με τη χρήση υπολογιστή, ενώ οι μη δομημένες πτυχές αντιμετωπίζονται από τους υπεύθυνους λήψης αποφάσεων. Κεντρικό στοιχείο της έννοιας του SDSS είναι η αλληλεπίδραση του χρήστη με ένα υπολογιστικό σύστημα που περιέχει ένα σύνολο εργαλείων για την ανάλυση χωρικών και μη χωρικών δεδομένων, και για τη μοντελοποίηση χωρικών προβλημάτων λήψης αποφάσεων (Malczewski & Rinner, 2015). Τα χωρικά χαρακτηριστικά καταγράφουν τις γεωγραφικές συντεταγμένες μιας θέσης και τις χωρικές σχέσεις, όπως εγγύτητα, επικάλυψη, περιορισμός και μοτίβο κατανομής (Keenan & Jankowski, 2019).

Έχουν επινοηθεί διάφορες μαθηματικές θεωρίες και υπολογιστικές τεχνικές για τη βελτιστοποίηση της επιλογής μεταξύ πολλαπλών εναλλακτικών λύσεων προβλημάτων, λαμβάνοντας υπόψη πολλαπλά κριτήρια. Ωστόσο, οι τυπικές μέθοδοι από μόνες τους δεν επαρκούν για την υποστήριξη της λήψης αποφάσεων, λόγω της ύπαρξης σιωπηρών κριτηρίων και γνώσεων, που δεν μπορούν να μεταφραστούν σε αριθμούς, τύπους ή κανόνες. Για το πρόβλημα αυτό έχουν επινοηθεί θεωρίες και τεχνικές από την ερευνητική περιοχή της Ανάλυσης Αποφάσεων Πολλαπλών Κριτηρίων (MCDA). Ένα παράδειγμα είναι η Διαδικασία Αναλυτικής Ιεράρχησης (AHP), κατά την οποία ένας λήπτης αποφάσεων συγκρίνει ζεύγη εναλλακτικών λύσεων και κριτηρίων. Στην MCDA αναγνωρίζεται η σημασία του γόνιμου

διαλόγου μεταξύ ενός υπολογιστικού συστήματος και ενός λήπτη αποφάσεων, όπου το σύστημα παρέχει στον λήπτη αποφάσεων πληροφορίες για εναλλακτικές λύσεις και λαμβάνει από αυτόν πληροφορίες για τις προτιμήσεις του. Για τον διάλογο αυτόν, τα σύγχρονα λογισμικά MCDA συχνά περιλαμβάνουν γραφικές απεικονίσεις όπως δέντρα, διαγράμματα και ραβδογράμματα. Ωστόσο, η ερευνητική κοινότητα της MCDA εστιάζει περισσότερο στα μαθηματικά παρά στις οπτικές ή άλλες διεπαφές, οι οποίες καθιστούν τις μεθόδους προσιτές και αξιοποιήσιμες. Η MCDA αυτή καθαυτή δεν ασχολείται συγκεκριμένα με τα χωρικά προβλήματα λήψης αποφάσεων, αλλά υποστηρίζεται μέσω της ενσωμάτωσης τεχνικών MCDA στα GIS, τα οποία χρησιμοποιούνται για την προετοιμασία των δεδομένων εισόδου των εργαλείων MCDA και για την αναπαράσταση της εξόδου σε χάρτη. Αυτή η προσέγγιση, από μόνο της, δεν επιλύει τα προβλήματα που σχετίζονται με τον όγκο, την πολυπλοκότητα και την ετερογένεια των εναλλακτικών λύσεων, των δεδομένων και της αξιολόγησής τους. Για την επίλυση των προβλημάτων αυτών ταιριάζουν καλύτερα οι προσεγγίσεις της γεωοπτικής αναλυτικής (περιγράφεται σε επόμενη ενότητα), όπου οι μέθοδοι MCDA μπορούν να ενσωματωθούν οργανικά, μέσω διαδραστικών και δυναμικών χωρικών διεπαφών (Andrienko et al., 2007).

2.3 Οπτική Αναλυτική

Η οπτική αναλυτική είναι ένα πολυεπιστημονικό πεδίο, που έχει προκύψει ως συνέχεια των τομέων της επιστημονικής οπτικοποίησης και της οπτικοποίησης πληροφοριών, αλλά σχετίζεται και με άλλους τομείς, όπως η διαχείριση γνώσης, η στατιστική ανάλυση, η γνωστική επιστήμη, η αντίληψη, η αλληλεπίδραση ανθρώπου-υπολογιστή, η αναλυτική λογική, η διαδραστική απεικόνιση, η εξερευνητική ανάλυση δεδομένων και η οπτική εξόρυξη δεδομένων (Cui, 2019). Οι όροι αυτοί συνδέονται σε μεγάλο βαθμό μεταξύ τους και συγγέονται εύκολα. Ο Πίνακας 2.1 εξηγεί τους σημαντικότερους από αυτούς.

Όρος	Επεξήγηση
Οπτικοποίηση	Αναφέρεται στις θεωρίες και τεχνικές δημιουργίας οπτικών αναπαραστάσεων των δεδομένων, που περιλαμβάνει οπτικοποίηση πληροφορικών και επιστημονική οπτικοποίηση.
Οπτικοποίηση πληροφοριών	Αναφέρεται στις θεωρίες και τεχνικές που χρησιμοποιούν διαδραστική οπτική υπολογιστική και αναπαραστάσεις για την ενίσχυση της ανθρώπινη επίγνωσης αφηρημένης πληροφορίας.
Επιστημονική οπτικοποίηση	Αναφέρεται στις θεωρίες και τεχνικές που χρησιμοποιούν οπτικές οθόνες και ρεαλιστικές απεικονίσεις για την ανάκτηση πληροφορίας από χωρικά δεδομένα που σχετίζονται με επιστημονικές διαδικασίες.
Διαδραστική οπτικοποίηση	Αναφέρεται στις θεωρίες και τεχνικές οπτικοποίησης και εξερεύνησης δεδομένων, μέσω χειρισμού του χρώματος, της φωτεινότητας, του μεγέθους και του σχήματος της οπτικής τους αναπαράστασης.
Διάδραση ανθρώπου-υπολογιστή	Αναφέρεται στη μελέτη του τρόπου με τον οποίο η τεχνολογία των υπολογιστών επηρεάζει την εργασία και τις δραστηριότητες των ανθρώπων.
Ανάλυση δεδομένων	Αναφέρεται στην διαδικασία της ανάλυσης των δεδομένων με στόχο την ανακάλυψη χρήσιμων πληροφοριών, εφαρμόζοντας τεχνικές στατιστικής και λογικής. Στην στατιστική, η ανάλυση δεδομένων χωρίζεται σε επιβεβαιωτική και εξερευνητική.
Επιβεβαιωτική ανάλυση δεδομένων	Αναφέρεται στην στατιστική διαδικασία αξιολόγησης καθορισμένων υποθέσεων με χρήση δοκιμών στατιστικής υπόθεσης πάνω σε υπάρχοντα σύνολα δεδομένων.
Εξερευνητική ανάλυση δεδομένων	Αναφέρεται σε προσεγγίσεις ανάλυσης συνόλων δεδομένων, που αποκαλύπτουν κρυμμένες και άγνωστες πληροφορικές από τα δεδομένα, πέρα από αυτές που μπορούν να αποκαλύψουν τα τυπικά μοντέλα ή οι δοκιμές υποθέσεων.
Οπτική εξόρυξη δεδομένων	Αναφέρεται στην διαδικασία διάδρασης και αναλυτικής σκέψης με οπτικές αναπαραστάσεις δεδομένων, που οδηγεί σε οπτική ανακάλυψη εύρωστων μοτίβων, τα οποία αποτελούν πληροφόρηση και γνώση, που στην συνέχεια χρησιμοποιούνται για την λήψη αποφάσεων.
Οπτική αναλυτική	Αναφέρεται στην επιστήμη της αναλυτικής σκέψης που υποστηρίζεται από διαδραστικές οπτικές διεπαφές.

Πίνακας 2.1: Σημαντικοί όροι που σχετίζονται με την Οπτική Αναλυτική (Cui, 2019)

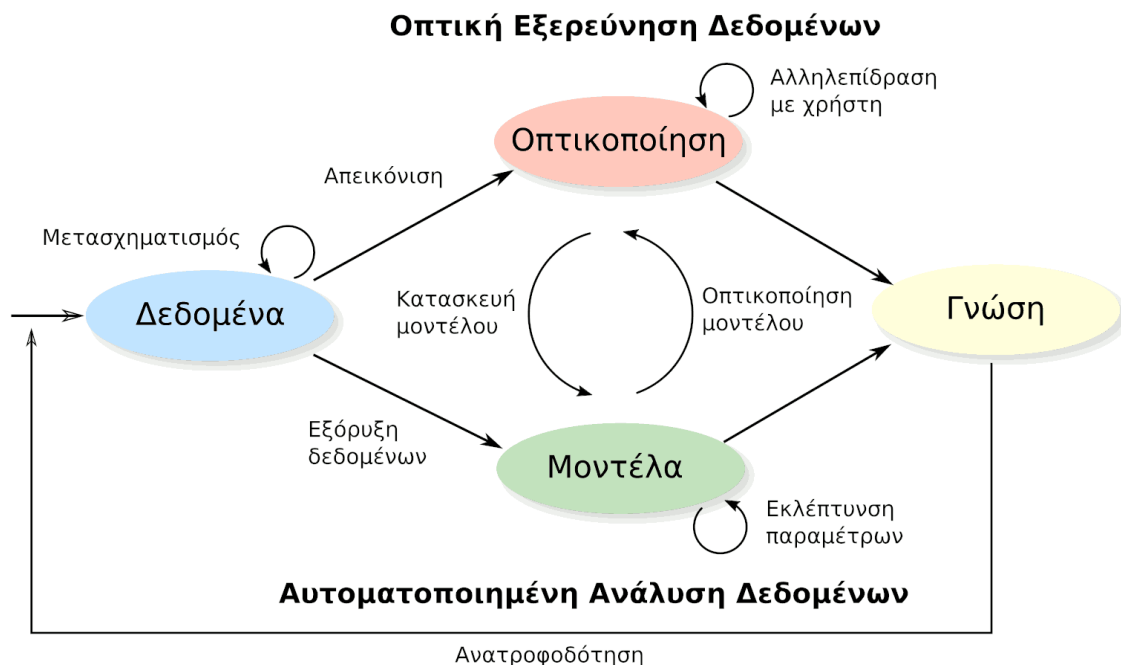
Οι τεχνολογίες της οπτικής αναλυτικής συνδυάζουν τα δυνατά σημεία της ανθρώπινης και της ηλεκτρονικής επεξεργασίας δεδομένων. Η οπτικοποίηση γίνεται το μέσο μιας ημι-αυτοματοποιημένης αναλυτικής διαδικασίας, όπου άνθρωποι και μηχανές συνεργάζονται,

αξιοποιώντας αποτελεσματικά τις αντίστοιχες ιδιαίτερες δυνατότητές τους (Keim et al., 2008). Βασίζεται δε στην υπόθεση ότι ο συνδυασμός των ποσοτικών δυνατοτήτων των υπολογιστών και των γνωστικών ικανοτήτων των ανθρώπων οδηγεί σε ισχυρούς τρόπους παραγωγής νέας γνώσης (Gudivada et al., 2016).

Η οπτική αναλυτική έχει ανακύψει από την ανάγκη αξιοποίησης του τεράστιου όγκου δεδομένων που παράγεται σε καθημερινή βάση από πρακτικά όλους τους κλάδους της βιομηχανίας και των επιχειρήσεων, αλλά και διάφορες πολιτικές ή προσωπικές δραστηριότητες. Ο ρυθμός αύξησης συλλογής των δεδομένων αυτών είναι τέτοιος, που ξεπερνά την ικανότητά μας να τα αξιοποιήσουμε για λήψη αποφάσεων, αφού στην πρωτογενή μορφή τους δεν έχουν αξία, διότι χρειαζόμαστε την πληροφορία που περιέχεται σε αυτά, και την οποία πρέπει να εξάγουμε. Το πρόβλημα αυτό της υπερφόρτωσης πληροφοριών, ενέχει τον κίνδυνο να χαθεί κάποιος μέσα στα δεδομένα, με ενδεχόμενες απώλειες χρόνου, χρήματος και ευκαιριών. Το κυρίαρχο όραμα για την οπτική αναλυτική είναι να μετατρέψει την υπερφόρτωση πληροφοριών σε ανακάλυψη ευκαιριών και παραγωγή γνώσης. (Cui, 2019; Keim et al., 2010a; Wong & Thomas, 2004)

Σύμφωνα με τον Cui (2019), ο James (Jim) Joseph Thomas ήταν αυτός που δημιούργησε, προώθησε και καθιέρωσε τον τομέα της οπτικής αναλυτικής, ενώ η πρώτη εμφάνιση του όρου στην βιβλιογραφία, είναι πολύ πιθανό να ήταν από τους Wong & Thomas (2004), οι οποίοι έδωσαν τον εξής ορισμό: “Η οπτική αναλυτική είναι ο σχηματισμός αφηρημένων οπτικών μεταφορών, σε συνδυασμό με έναν διάλογο (αλληλεπίδραση) του ανθρώπου με την πληροφορία, που επιτρέπει την ανίχνευση του αναμενόμενου και την ανακάλυψη του απροσδόκητου, μέσα σε τεράστιους, δυναμικά μεταβαλλόμενους χώρους πληροφοριών”. Σύμφωνα με τους Thomas & Cook (2005), “η οπτική αναλυτική είναι η επιστήμη του αναλυτικού συλλογισμού, που διευκολύνεται από διαδραστικές οπτικές διεπαφές”. Οι Keim et al. (2008) δίνουν έναν πιο εξειδικευμένο ορισμό, σύμφωνα με τον οποίο “η οπτική αναλυτική συνδυάζει αυτοματοποιημένες τεχνικές ανάλυσης με διαδραστικές απεικονίσεις για αποτελεσματική κατανόηση, συλλογισμό και λήψη αποφάσεων με βάση πολύ μεγάλα και πολύπλοκα σύνολα δεδομένων”.

Στην Εικόνα 2.2 παρουσιάζεται μια αφηρημένη επισκόπηση των διαφόρων σταδίων και των μεταβάσεών τους, στην ροή εργασίας της οπτικής αναλυτικής. Το πρώτο στάδιο είναι ο μετασχηματισμός δεδομένων από ετερογενείς πηγές σε αναπαραστάσεις, κατάλληλες για την περαιτέρω εξερεύνηση. Στην συνέχεια ο αναλυτής επιλέγει μεταξύ οπτικής και αυτοματοποιημένης ανάλυσης. Αν επιλέξει πρώτα την αυτοματοποιημένη ανάλυση, θα κατασκευάσει μοντέλα μέσω εξόρυξης δεδομένων (data mining), και στην συνέχεια θα αλληλεπιδράσει με τα δεδομένα μέσω οπτικοποιήσεων, αξιολογώντας τα ευρήματα, και επαναλαμβάνοντας την διαδικασία μέχρι να εκλεπτύνει τα μοντέλα. Αν επιλέξει πρώτα την οπτική ανάλυση, ο χρήστης θα αλληλεπιδράσει εξερευνητικά με οπτικοποιήσεις των δεδομένων, προκειμένου να ανακαλύψει ενδιαφέρουσες πληροφορίες και να παράξει μια υπόθεση, την οποία στην συνέχεια, θα πρέπει να επαληθεύσει, μέσω αυτοματοποιημένης ανάλυσης. Συνολικά, στην ροή εργασίας της οπτικής αναλυτικής, η γνώση αποκτάται μέσα από την οπτικοποίηση, την αυτόματη ανάλυση και τις αλληλεπιδράσεις μεταξύ του αναλυτή, των οπτικοποιήσεων και των μοντέλων. (Keim et al., 2010a, p10).



Εικόνα 2.2: Η ροή εργασίας της Οπτικής Αναλυτικής (Keim et al., 2010a)

Η οπτική αναλυτική είναι ένας ιδιαίτερα εφαρμοσμένος επιστημονικός τομέας, και οδηγείται από τις πρακτικές απαιτήσεις του εκάστοτε τομέα εφαρμογής. Ενδεικτικά, μπορεί να έχει σημαντική συνεισφορά στους παρακάτω τομείς (Keim et al., 2008):

- Τομείς της *μηχανικής*, όπου η οπτική αναλυτική μπορεί να επιταχύνει την ανάπτυξη προϊόντων, εργαλείων, υλικών και μεθόδων παραγωγής, προσφέροντας αποτελεσματική και ευφυή πρόσβαση σε σύνθετη πληροφορία από πρωτότυπα, πειράματα, ανατροφοδότηση πελατών και διάφορους δείκτες.
- Η *χρηματοοικονομική ανάλυση*, όπου οι αναλυτές έρχονται αντιμέτωποι με ροές ετερογενών πληροφοριών από διαφορετικές πηγές που είναι διαθέσιμες με υψηλούς ρυθμούς ενημέρωσης και ποικίλης αξιοπιστίας, έχει σαν κεντρικό στόχο την ενοποιημένη προβολή των πληροφοριών αυτών σε πραγματικό χρόνο. Για την οπτική αναλυτική είναι μια πολλά υποσχόμενη περιοχή εφαρμογών.
- *Κοινωνικοοικονομικοί* παράγοντες αποτελούν συχνά τη βάση των πολιτικών διαδικασιών λήψης αποφάσεων. Η ανάλυση και η οπτικοποίηση αλληλεπιδράσεων στην σύγχρονη κοινωνία είναι ένα πεδίο εφαρμογής της οπτικής αναλυτικής, με εφαρμογές που θα μπορούσαν να αναπτυχθούν κάνοντας χρήση εργαλείων γεωχωρικής ανάλυσης.
- Η *δημόσια ασφάλεια* είναι ένας ακόμα σημαντικός τομέας, όπου οι αναλυτές πρέπει να παρακολουθούν συνεχώς τεράστιες ποσότητες ετερογενών ροών πληροφοριών, συσχετίζοντας πληροφορίες διαφορετικού βαθμού αφάιρεσης και αξιοπιστίας, όπου θα πρέπει να αξιολογούν το τρέχον, κάθε φορά, επίπεδο ασφάλειας, ενεργοποιώντας συναγερμό σε περίπτωση εντοπισμού απειλητικών καταστάσεων. Η οπτική αναλυτική υπόσχεται την ανάπτυξη πιο αποτελεσματικών διαδραστικών εργαλείων οπτικοποίησης, για την ενοποίηση και τον συσχετισμό δεδομένων στον τομέα αυτό.
- Στον τομέα του *περιβάλλοντος και της κλιματικής αλλαγής*, η εξέταση μακροχρόνιων καταγραφών μετεωρολογικών δεδομένων και δεδομένων από αισθητήρες, είναι απαραίτητη για την αναζήτηση προτύπων συσχετισμού παρατηρήσεων. Η οπτική αναλυτική μπορεί να συμβάλει στην ανάπτυξη συστημάτων τόσο οπτικής και

γραφικής πρόσβασης σε ιστορικά δεδομένα, όσο και προβλέψεων από διάφορα μοντέλα για την αναζήτηση ή την επαλήθευση προτύπων.

2.4 Γεωοπτική Αναλυτική

Η Γεωοπτική Αναλυτική ή Γεωχωρική Οπτική Αναλυτική (Geovisual Analytics ή Geospatial Visual Analytics), αναφέρεται στην επιστήμη του αναλυτικού συλλογισμού με χωρικές πληροφορίες, που διευκολύνεται από διαδραστικές οπτικές διεπαφές (Robinson, 2017). Πρόκειται για εξειδίκευση της οπτικής αναλυτικής σε γεωχωρικά δεδομένα και αποτελεί επέκταση της γεωοπτικοποίησης (geovisualization). Η γεωοπτικοποίηση αρχικά ήταν ένας ελεύθερα οριοθετημένος τομέας που περιλάμβανε την οπτική εξερεύνηση, ανάλυση, σύνθεση και παρουσίαση γεωχωρικών δεδομένων, ενσωματώνοντας προσεγγίσεις της χαρτογραφίας με την οπτικοποίηση, την ανάλυση εικόνας, την εξερευνητική ανάλυση δεδομένων (exploratory data analysis) και την επιστήμη της γεωγραφικής πληροφορίας (GIScience). Ωστόσο, στην βιβλιογραφία φαίνεται ότι ο όρος γεωοπτικοποίηση έχει χάσει πλέον τον παραπάνω ορισμό, και χρησιμοποιείται περισσότερο ως συνώνυμο της χαρτογράφησης (Kraak, 2008). Η γεωοπτική αναλυτική σήμερα είναι το αντικείμενο σημαντικού μέρους έρευνας σχετικά με την χαρτογραφία, τις υπολογιστικές μεθόδους, τον σχεδιασμό διεπαφών και την γνωστική επιστήμη, και αναζητά νέα διαδραστικά συστήματα χαρτογράφησης που επιτρέπουν στους χρήστες να εντοπίζουν μοτίβα και να προβλέπουν μελλοντικά αποτελέσματα χρησιμοποιώντας γεωχωρικά δεδομένα. Ένας σημαντικός παράγοντας για την ανάπτυξη του τομέα, είναι η υποστήριξη αναλυτικών εργασιών που σχετίζονται με ανθρώπινα και περιβαλλοντικά προβλήματα. Αυτό έχει ως αποτέλεσμα, οι προσεγγίσεις της γεωοπτικής αναλυτικής, να έρχονται αντιμέτωπες με το γεγονός ότι τα δεδομένα μπορεί να είναι ακατάστατα και αβέβαια, ότι τα προβλήματα μπορεί να μην έχουν μονοσήμαντες απαντήσεις, και ότι διαφορετικές ομάδες χρηστών μπορεί να έχουν θεμελιωδώς διαφορετικές απαιτήσεις για τους αναλυτικούς τους στόχους. Επιπλέον, ο τομέας επηρεάζεται άμεσα από προκλήσεις των μεγάλων δεδομένων (Big Data), όπως ο όγκος (volume), η ταχύτητα (velocity), η ποικιλότητα (variety) και η αξιοπιστία (veracity) (Robinson, 2017). Παρακάτω αναλύονται οι επιμέρους τομείς που απαρτίζουν την γεωοπτική αναλυτική.

2.4.1 Οπτικοποίηση

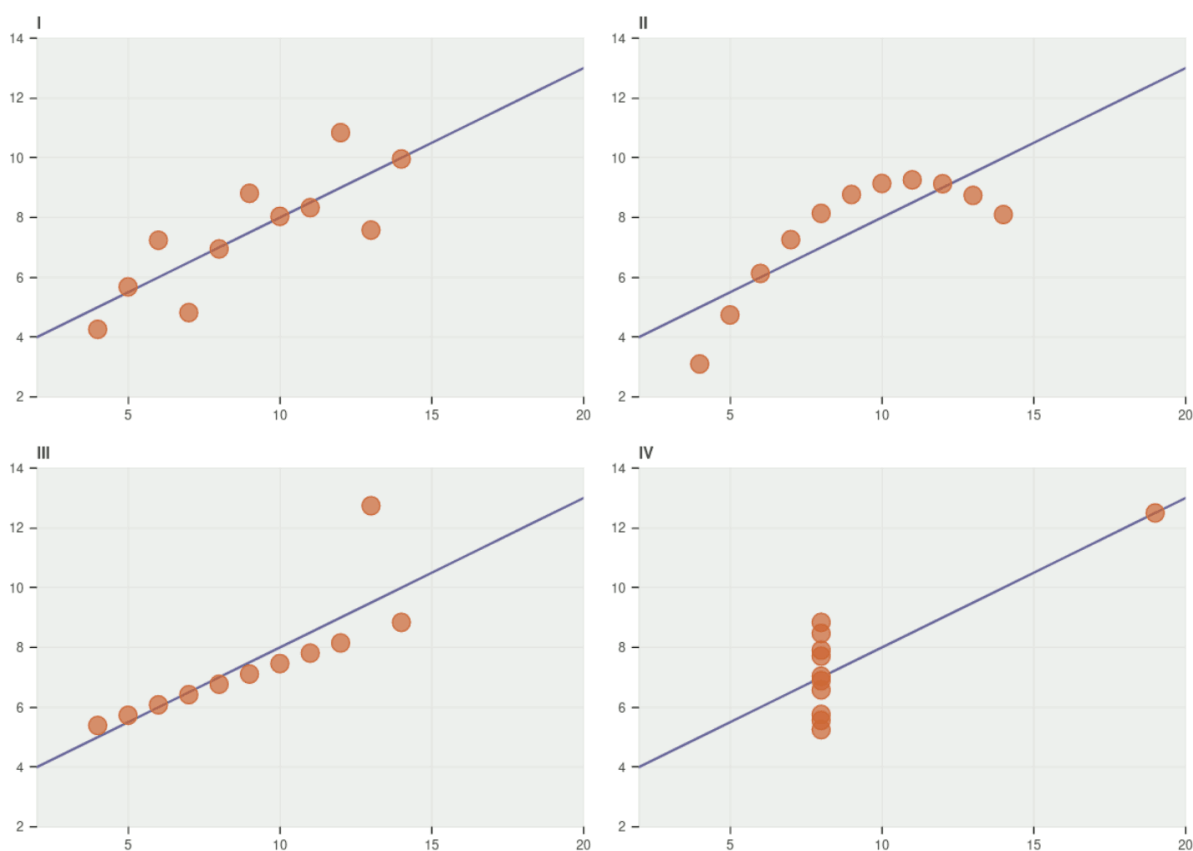
“Η μεγαλύτερη αξία μιας εικόνας είναι όταν αυτή μας αναγκάζει να παρατηρήσουμε αυτό που δεν περιμέναμε ποτέ να δούμε.” (Tukey, 1977)

Οπτικοποίηση ονομάζεται ο τρόπος επικοινωνίας της πληροφορίας, που πραγματοποιείται μέσα από γραφικές αναπαραστάσεις. Ως μορφή επικοινωνίας χρησιμοποιήθηκε πριν από τον γραπτό λόγο. Συχνά αναφέρεται ότι μια εικόνα αξίζει όσο χίλιες λέξεις. Επιπλέον όμως, μια εικόνα μπορεί να αφομοιωθεί πολύ πιο γρήγορα από μια σελίδα αντίστοιχων λέξεων, γιατί η επεξεργασία της στο ανθρώπινο αντιληπτικό σύστημα εκτελείται παράλληλα, ενώ η ανάλυση κειμένου περιορίζεται από την σειριακή διαδικασία της ανάγνωσης. Επίσης οι εικόνες, όπως ένα γράφημα ή ένας χάρτης, μπορούν να γίνουν κατανοητές, χωρίς την ανάγκη ύπαρξης κοινής γλώσσας. Καθημερινά παραδείγματα οπτικοποίησης είναι: ένα διάγραμμα καιρού, ένας χάρτης του μετρό, ένα γράφημα της πορείας μιας μετοχής, οι οπτικές οδηγίες συναρμολόγησης μιας κατασκευής, μια οδική πινακίδα. Χαρακτηριστικά ιστορικά παραδείγματα πρώιμης οπτικοποίησης, είναι ο χάρτης από τον John Snow, που απεικονίζει τους θανάτους από χολέρα, το 1854, στο Λονδίνο (Εικόνα 2.3), καθώς και ο χάρτης από τον Charles Joseph Minard, της καταστροφικής εκστρατείας του Ναπολέοντα στην Ρωσία το 1812 (Εικόνα 2.4). (Ward et al., 2015, p1)



Εικόνα 2.3: Χάρτης των θανάτων από χολέρα το 1854 στο Λονδίνο από τον John Snow (Wikimedia Commons)

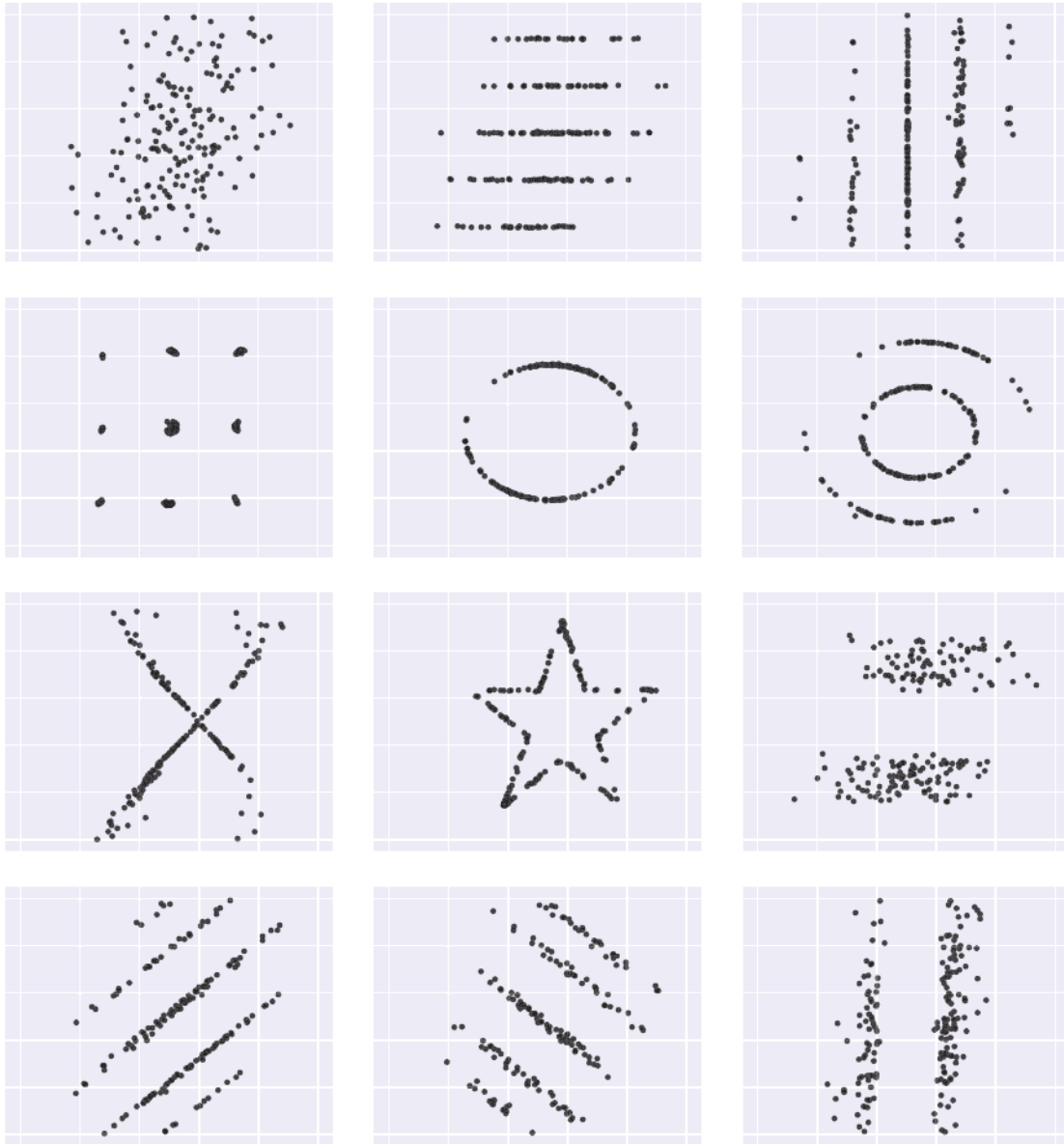
σύνολα δεδομένων, τα οποία έχουν ακριβώς ίσα, συχνά χρησιμοποιούμενα περιληπτικά στατιστικά μέτρα (μέσος όρος, ενδιάμεσος, τυπική απόκλιση, συντελεστής προσδιορισμού r^2), υπαινίσσοντας ότι και οι κατανομές τους είναι παρόμοιες. Ωστόσο τα διαγράμματα διασποράς τους δείχνουν ότι κάτι τέτοιο δεν συμβαίνει. Τα σύνολα δεδομένων αυτά είναι προσεκτικά κατασκευασμένα για να αναδείξουν ότι τα υψηλού επιπέδου περιληπτικά στατιστικά μπορεί να είναι παραπλανητικά, και ότι η οπτικοποίηση μπορεί να προσφέρει μια πολύ πιο αξιοποιήσιμη και ερμηνεύσιμη εικόνα. (Andrienko et al., 2020, p52)



Εικόνα 2.5: Η τετράδα του Anscombe

Σύμφωνα με τους Matejka & Fitzmaurice (2017) η αξία της τετράδας του Anscombe δεν είναι ότι πρόκειται για τέσσερα διαφορετικά σύνολα δεδομένων που τυχαίνει να έχουν τα ίδια περιληπτικά στατιστικά, αλλά ότι τέσσερα ευδιακρίτως διαφορετικά σύνολα δεδομένων μπορούν να παράγουν τα ίδια στατιστικά χαρακτηριστικά. Για να αναδείξουν ακόμα περισσότερο την σημασία της γραφικής απεικόνισης κατά την εξερεύνηση δεδομένων, ανέπτυξαν μια καινοτόμο τεχνική που παράγει σύνολα δεδομένων με σχεδόν πανομοιότυπα

στατιστικά χαρακτηριστικά (με απόκλιση δύο δεκαδικών ψηφίων), αλλά με αρκετά διαφορετικές δομές, που γίνονται προφανείς μετά την οπτικοποίησή τους (Εικόνα 2.6).

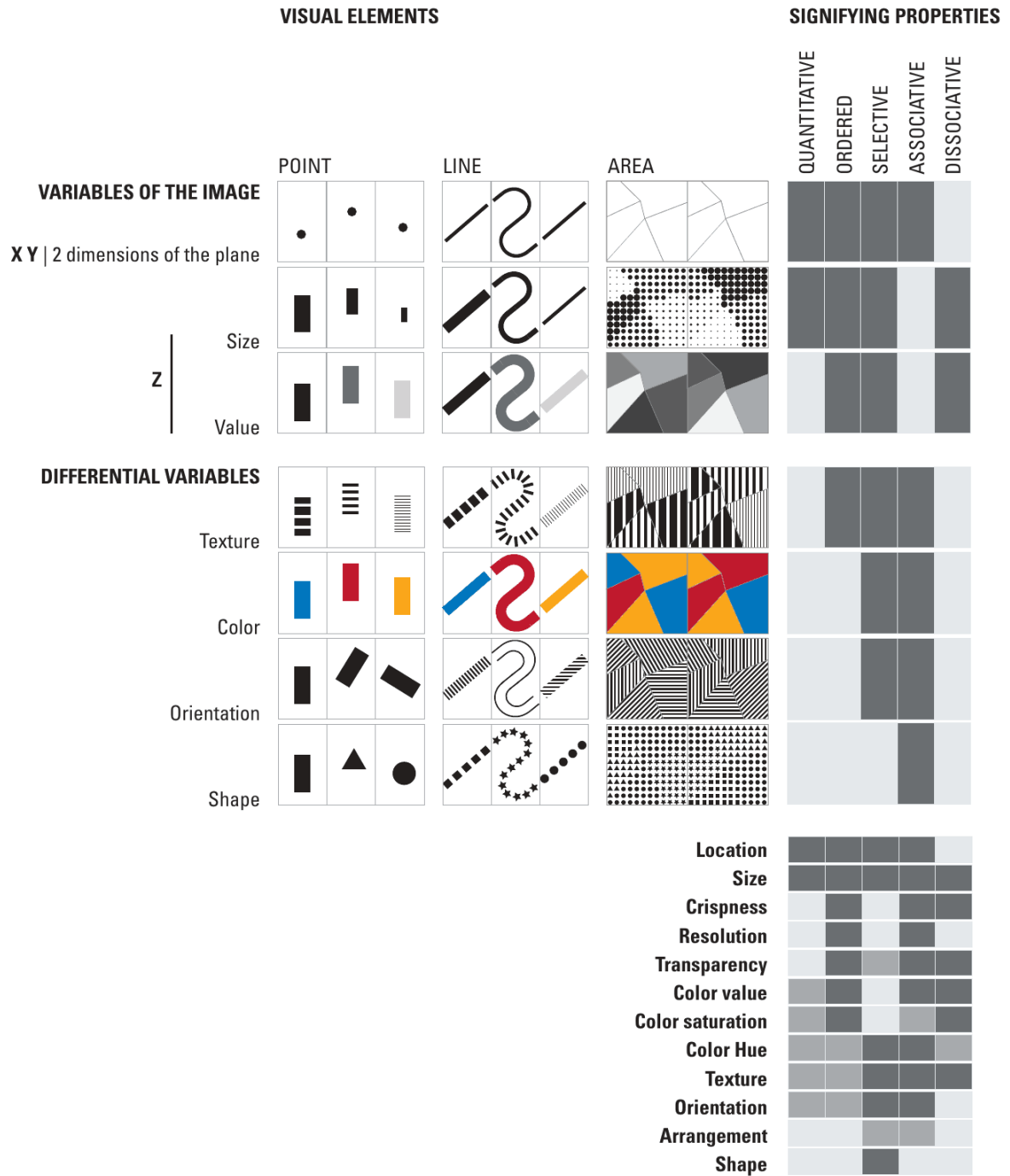


Εικόνα 2.6: Μια συλλογή παραγόμενων συνόλων δεδομένων, διαφορετικών στην εμφάνιση, αλλά με τα ίδια περιληπτικά στατιστικά, με απόκλιση 2 δεκαδικών ψηφίων. (Matejka & Fitzmaurice, 2017)

Οπτικές μεταβλητές

Υπάρχουν οκτώ μεταβλητές με τις οποίες γραφικά αντικείμενα μπορούν να κωδικοποιήσουν πληροφορία: θέση, σχήμα, μέγεθος, φωτεινότητα, χρώμα, προσανατολισμός, υφή και κίνηση. Η θέση των αντικειμένων στον χώρο μίας οθόνης είναι η πιο σημαντική, και δίνει τη δυνατότητα να χαρτογραφηθούν και να αποκαλυφθούν χαρακτηριστικά και ιδιότητες των υποκείμενων δεδομένων, όπως η ομοιότητα και η κατανομή τους. Στην συνέχεια, χρησιμοποιώντας τις υπόλοιπες μεταβλητές, άλλες ιδιότητες μπορούν να αποδοθούν στα γραφικά αντικείμενα, προκειμένου να μεταδώσουν πληροφορία (Ward et al., 2015, p147). Εκτός από τις παραπάνω μεταβλητές, έχουν προταθεί και άλλες όπως η ευκρίνεια, η ανάλυση, η διαφάνεια και διάταξη. Θεμελιώδης θεωρείται η συνεισφορά του Jacques Bertin στις οπτικές μεταβλητές, τις οποίες συνέδεσε με τα βασικά γραφικά στοιχεία και τα είδη των φαινομένων προς απεικόνιση (Εικόνα 2.7). Τα βασικά γραφικά στοιχεία είναι: το σημείο, η γραμμή και η επιφάνεια (Meirelles, 2013, p126).

Ορισμένες μεταβλητές μπορούν να ταξινομηθούν λογικά, όπως η θέση, το μέγεθος και η φωτεινότητα. Άλλες, όπως το χρώμα ή το σχήμα, δεν έχουν λογική ταξινόμηση, αλλά είναι δυνατό να κατασκευαστεί ένα ιεραρχημένο σύνολο από επιλεγμένες τιμές, χρησιμοποιώντας, για παράδειγμα την πυκνότητα των γραμμών ενός σχήματος. Οι τιμές της μεταβλητής του προσανατολισμού μπορούν να αποτυπωθούν ως μοίρες, από 0 ως 360, πάνω σε έναν κύκλο. Εκτός από την ταξινόμηση, την απόσταση και την αναλογία μεταξύ τιμών οπτικών μεταβλητών, σημαντικές ιδιότητες αποτελούν η επιλεξιμότητα και ο συσχετισμός, δηλαδή το πόσο εύκολα, διαφορετικές τιμές μιας μεταβλητής μπορούν να διακριθούν μεταξύ τους, καθώς και πόσο εύκολα ίδιες τιμές μπορούν να εντοπιστούν όλες μαζί και να διακριθούν από τα υπόλοιπα σημάδια. (Andrienko et al., 2020, p55-58)

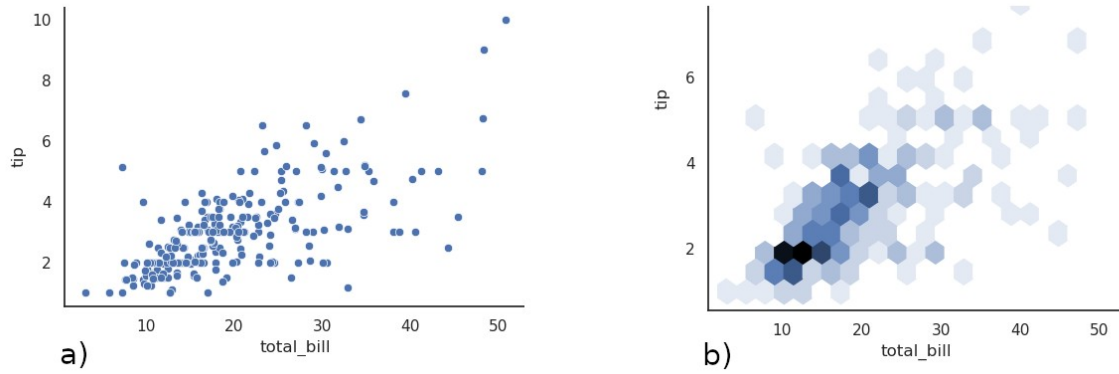


Εικόνα 2.7: Το σύστημα των αντιληπτικών μεταβλητών του Bertin (Meirelles, 2013, p127)

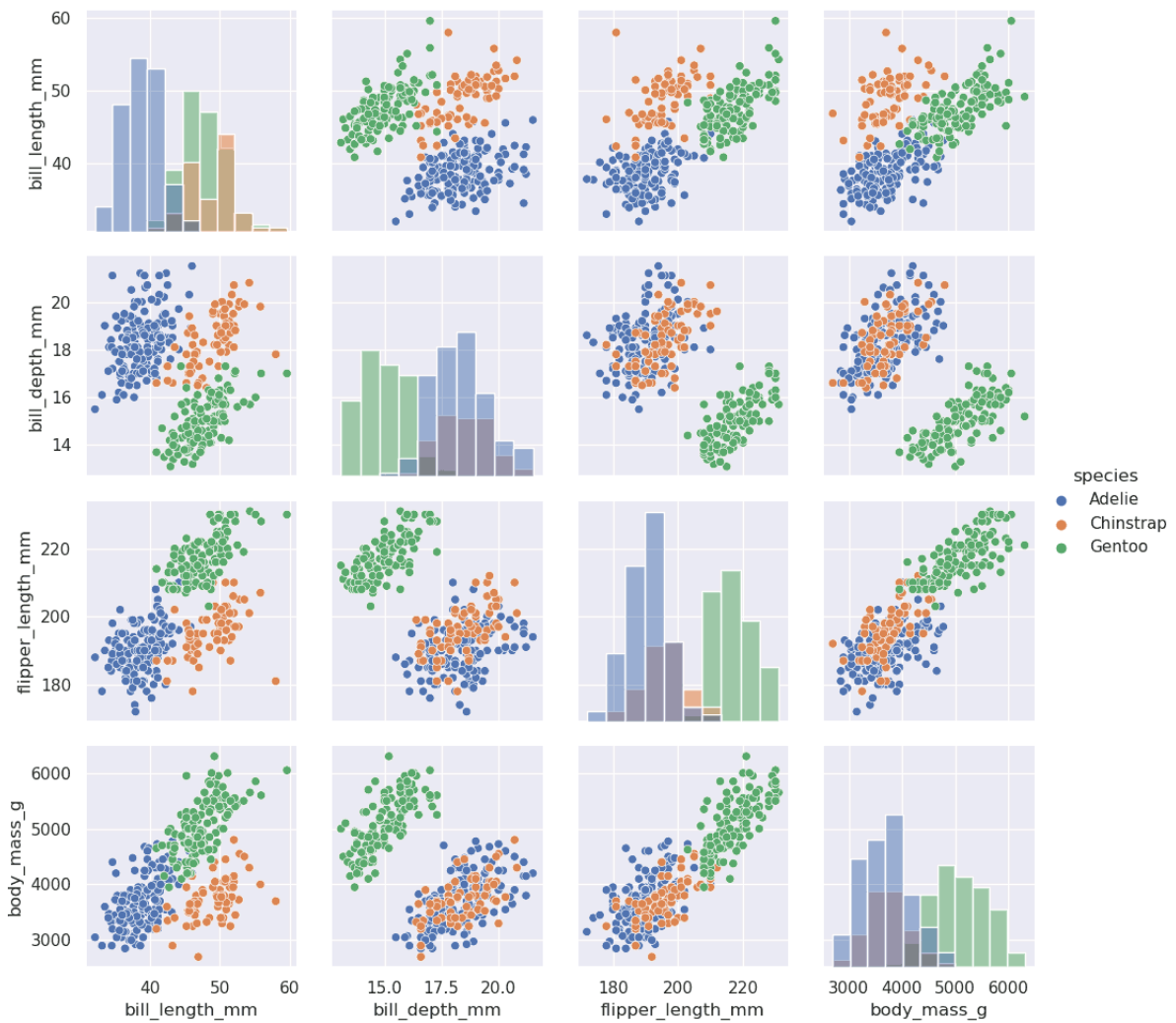
Διαγράμματα

Τα διαγράμματα είναι πολύ συνηθισμένα είδη οπτικοποίησης για την ανάλυση δεδομένων. Γενικά, απαιτούν δεδομένα σε μορφή πίνακα, κάθε γραμμή του οποίου αντιπροσωπεύει μία οντότητα, οι ιδιότητες της οποίας περιέχονται στις στήλες. Συχνά χρησιμοποιούμενα διαγράμματα είναι το διάγραμμα διασποράς (scatter plot), το ιστόγραμμα (histogram), το ραβδόγραμμα (bar chart), το θηκόγραμμα (box plot) και οι παράλληλες συντεταγμένες (parallel coordinates). (Ware, 2020, p429)

Το *διάγραμμα διασποράς (scatter plot)*, που στην πιο συνηθισμένη μορφή του είναι δισδιάστατο, είναι μία αποτελεσματική μέθοδος εξέτασης σχέσεων δύο μεταβλητών, όπως ο βαθμός συσχέτισης (correlation) και η ομαδοποίηση (clustering) (Εικόνα 2.8 a). Ενώ το δισδιάστατο διάγραμμα διασποράς λειτουργεί καλά για δύο μεταβλητές, για μεγαλύτερο αριθμό μεταβλητών πρέπει να επιστρατευτεί μεταβολή άλλων χαρακτηριστικών, όπως το χρώμα και το μέγεθος των σημείων. Για πέντε ή περισσότερες μεταβλητές μπορεί να χρησιμοποιηθεί *πίνακας διαγραμμάτων διασποράς (scatter plot matrix)* ή *διάγραμμα παράλληλων συντεταγμένων* (Ware, 2020, p431). Η Εικόνα 2.9 δείχνει πίνακα διαγραμμάτων διασποράς με ιστογράμματα στην διαγώνιο, ενώ επιστρατεύει και χρώμα για απεικόνιση μιας επιπλέον κατηγορικής μεταβλητής. Όταν μεγαλώνει ο αριθμός των σημείων, αυτά γίνονται δυσδιάκριτα. Σε αυτήν την περίπτωση, μπορεί να χρησιμοποιηθεί *διάγραμμα διασποράς με ομαδοποίηση σε δοχεία (binned scatter plot)* (Εικόνα 2.8 b), όπου τα σημεία αντικαθίστανται από πλέγμα κελιών, για παράδειγμα σε σχήμα εξαγώνου, και χαρακτηριστικά του κελιού, όπως το χρώμα, η φωτεινότητα, το μέγεθος ή συνδυασμός τους, αναπαριστούν την πυκνότητα των σημείων που βρίσκονται στην περιοχή που καταλαμβάνει κάθε κελί (Hao et al., 2017; Carr et al., 1987).



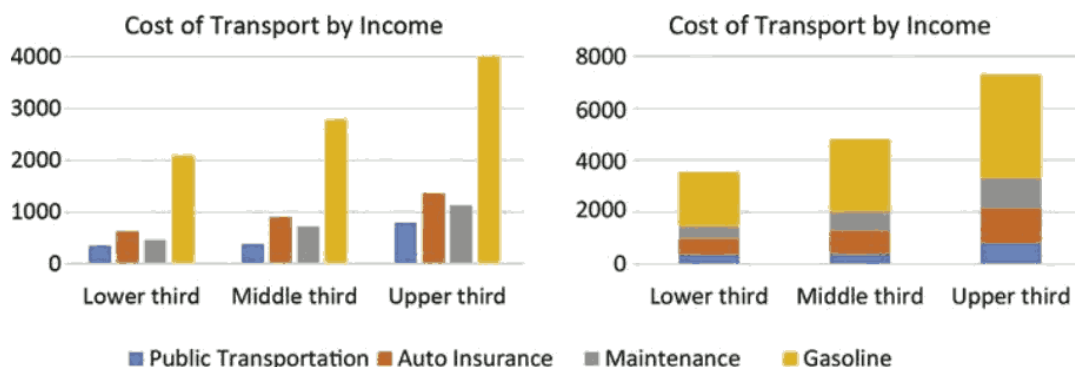
Εικόνα 2.8: α) Διάγραμμα διασποράς, β) Διάγραμμα διασποράς με ομαδοποίηση σε δοχεία



Εικόνα 2.9: Πίνακας διαγραμμάτων διασποράς με ιστογράμματα στην διαγώνιο

Το *ιστόγραμμα συχνότητας (frequency histogram)* απεικονίζει την κατανομή τιμών αριθμητικών χαρακτηριστικών, μοιράζοντας τα δεδομένα σε νοητά δοχεία, κατά μήκος ενός άξονα, και χρησιμοποιώντας ράβδους για να αναπαραστήσει την συχνότητα εμφάνισής τους. Ένας χρήσιμος τρόπος αλληλεπίδρασης με το διάγραμμα είναι η μεταβολή του μεγέθους των δοχείων. Ανταποκρίνεται καλύτερα στην κλιμάκωση από διαγράμματα κουκκίδων, όπως αυτό της διασποράς, γιατί πραγματοποιεί συνάθροιση δεδομένων. Επίσης οι ράβδοι μπορούν να αναπαραστήσουν συχνότητες και για υποσύνολα στοιχείων των δεδομένων, χωρίζοντας την κάθε ράβδο σε τμήματα με διαφορετικές αποχρώσεις, καθένα από τα οποία αναπαριστά μια κατηγορία ή κλάση. (Andrienko et al., 2020, p64)

Το *ραβδόγραμμα (bar chart)* είναι από τους πιο κοινούς τρόπους οπτικοποίησης, όπου ορθογώνιες ράβδοι απεικονίζουν αριθμητικές τιμές. Είναι κατάλληλα για απεικόνιση πολλών ειδών δεδομένων, γιατί αξιοποιούν την οπτική οξύτητα με την οποία οι άνθρωποι συγκρίνουν τα μήκη ευθύγραμμων τμημάτων. Χρησιμοποιείται για μεταβλητές με μικρό αριθμό διακριτών τιμών όπως κατηγορίες ή ακέραιους, που μπορούν να χωρέσουν όλες στο διάγραμμα, διαφορετικά είναι πιο κατάλληλο το ιστόγραμμα που πραγματοποιεί συνάθροιση των δεδομένων και παρουσιάζει μια πιο συνοπτική εικόνα τους. Για πολυμεταβλητά (multivariate) δεδομένα, μια εναλλακτική είναι κάθε ράβδος να αποτελείται από στοίβα επιμέρους ράβδων, κάθε μία από τις οποίες αναπαριστά και μία μεταβλητή (Εικόνα 2.10 δεξιά). Μια άλλη επιλογή, για μικρό αριθμό μεταβλητών, είναι η ομαδοποίηση ράβδων μεταβλητών τοποθετώντας τις τη μία δίπλα στην άλλη τους, σε ομάδες ράβδων με κοινή βάση (Εικόνα 2.10 αριστερά). Και στις δύο παραπάνω περιπτώσεις, για ευκρίνεια, συνηθίζεται η ράβδος για κάθε μεταβλητή να διαφοροποιείται από τις υπόλοιπες με κάποιο χαρακτηριστικό, όπως το χρώμα ή την υφή της. (Ward et al., 2015, p299-300)

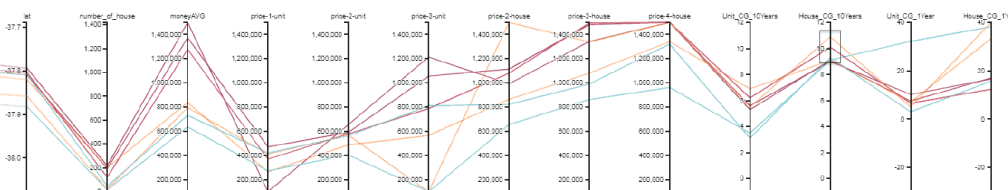


Εικόνα 2.10: Ραβδογράμματα (Ware, 2020, p430)

Το *θηκόγραμμα με κεραίες* (*box and whiskers plot*) συνοψίζει μια κατανομή αριθμητικών τιμών, απεικονίζοντας τα τεταρτημόρια, τα οποία είναι συγκεκριμένες τιμές που διαιρούν τη διατεταγμένη ακολουθία των τιμών που εμφανίζονται στα δεδομένα σε τέσσερα περίπου ίσα μέρη. Το διάγραμμα σχεδιάζεται κατά μήκος ενός άξονα που αντιπροσωπεύει το εύρος τιμών. Το πρώτο και το τρίτο τεταρτημόριο (Q1 και Q3) αναπαρίστανται από δύο πλευρές ενός ορθογώνιου πλαισίου, και το δεύτερο τεταρτημόριο (Q2), που είναι η διάμεσος, από μία γραμμή εντός του πλαισίου. Οι κεραίες μπορούν να συνδέουν τις πλευρές του πλαισίου με διάφορες πιθανές εναλλακτικές τιμές, όπως το ελάχιστο και το μέγιστο των δεδομένων, το 9ο και το 91ο εκατοστημόριο ή το 2ο και το 98ο εκατοστημόριο. Όταν μια κεραία δεν φτάνει στο ελάχιστο ή το μέγιστο, το γράφημα μπορεί να περιλαμβάνει κουκκίδες για να απεικονίσει τις ακραίες τιμές. (Andrienko et al., 2020, p65)

Οι *παράλληλες συντεταγμένες* (*parallel coordinates*) (Εικόνα 2.11) είναι ένα διάγραμμα κατεξοχήν για πολυμεταβλητά δεδομένα, όπου οι μεταβλητές αναπαρίστανται από παράλληλους άξονες και, αντί για κουκκίδες, έχουμε τεθλασμένες γραμμές που συνδέουν τις τιμές των χαρακτηριστικών από την μία διάσταση στην επόμενη. Επειδή τα μοτίβα που εμφανίζονται σε αυτόν τον τύπο διαγράμματος εξαρτώνται από την σειρά τοποθέτησης των αξόνων των μεταβλητών, στις περισσότερες υλοποιήσεις υπάρχει η δυνατότητα αναδιάταξής τους. Είναι σημαντικό να αναφερθεί ότι τα διαγράμματα παράλληλων συντεταγμένων προορίζονται για χρήση με μια τεχνική αλληλεπίδρασης που ονομάζεται βούρτσισμα (*brushing*), σύμφωνα με την οποία επιλέγεται κάποιο διάστημα πάνω σε έναν άξονα,

προκαλώντας τις τεθλασμένες γραμμές που περνάνε από το διάστημα αυτό, να τονιστούν με κάποιον τρόπο. (Ware, 2020, p221)



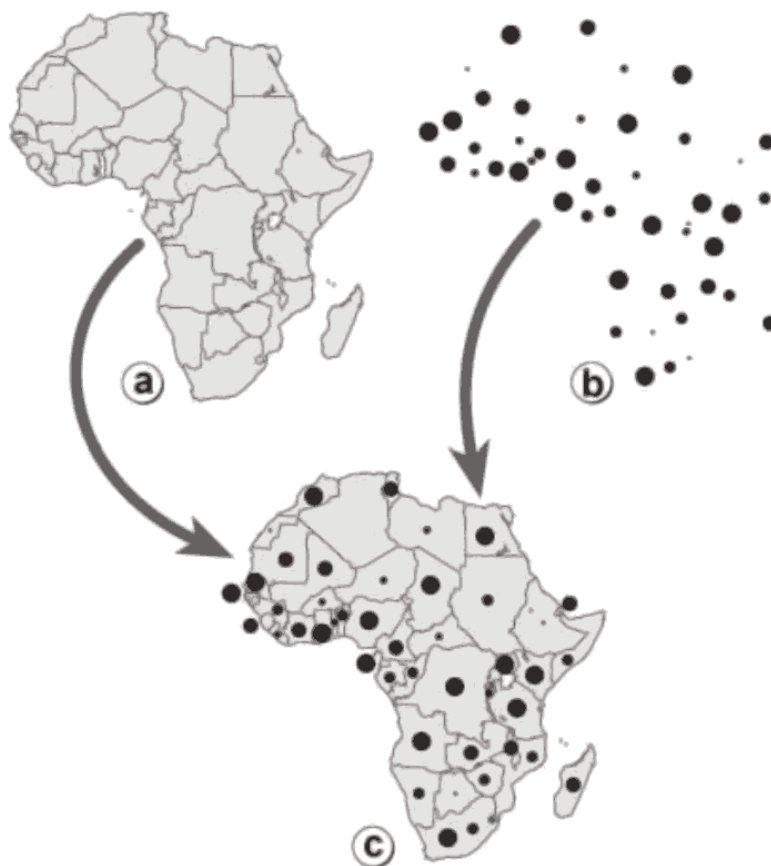
Εικόνα 2.11: Διάγραμμα παράλληλων συντεταγμένων (Li, 2019, p47)

Χάρτες

Ο χάρτης είναι μια αφαιρετική απεικόνιση του κόσμου με σημεία, γραμμές και επίπεδα (Ward et al., 2015). Είναι εργαλείο για την οπτική παρουσίαση επιλεγμένων γεωγραφικών πληροφοριών ή αφηρημένων φαινομένων που μπορούν να εντοπιστούν στο χώρο. Μια σημαντική παράμετρος ενός χάρτη είναι η κλίμακά του, δηλαδή η αναλογία μεταξύ μιας απόστασης στον χάρτη και την αντίστοιχη απόσταση στο έδαφος. (Kraak & Ormeling, 2010, p41). Ο όρος χάρτης χρησιμοποιείται σε πολλούς επιστημονικούς τομείς ως συνώνυμο ενός μοντέλου για την αντίληψη της δομής αυτού που αναπαριστά, οπότε εκτός από την αποτύπωση ενός φαινομένου, βοηθά και στην κατανόησή του. Η αφαιρετική ικανότητα ενός χάρτη είναι αυτή που μας επιτρέπει να παρατηρούμε γεωχωρικές σχέσεις, δομές και μοτίβα. Ο όρος χαρτογραφία, ενώ αρχικά περιέγραφε την διαδικασία δημιουργίας χαρτών, σήμερα τείνει να συνδέεται και με την έννοια της μετάδοσης γεωχωρικών πληροφοριών μέσω χαρτών. (Kraak & Ormeling, 2010, p39)

Παραδοσιακά, οι χάρτες χωρίζονται σε τοπογραφικούς και θεματικούς. Οι τοπογραφικοί χάρτες παρέχουν μια γενική εικόνα της επιφάνειας της Γης, απεικονίζοντας πχ δρόμους, ποταμούς, σπίτια, βλάστηση, ανάγλυφο και ονόματα. Οι θεματικοί χάρτες αναπαριστούν την κατανομή συγκεκριμένων φαινομένων, και απαιτούν τοπογραφική πληροφορία ως βάση, η οποία παρέχεται συχνά με τη μορφή απλουστευμένων τοπογραφικών χαρτών. Στο ψηφιακό περιβάλλον, ο διαχωρισμός μεταξύ θεματικών και τοπογραφικών χαρτών δεν είναι τόσο χρήσιμος, καθώς και οι δύο τύποι αποτελούνται από επίπεδα. Για παράδειγμα ένας τοπογραφικός χάρτης μπορεί να αποτελείται από συνδυασμό πολλών επιμέρους επιπέδων, καθένα από τα οποία είναι ένας θεματικός χάρτης. Αντίθετα, όταν ένα από τα φαινόμενα ενός

τοπογραφικού χάρτη επιλεγεί για ανάδειξη και τα υπόλοιπα υποβιβαστούν σε αναπαράσταση εδάφους, τότε ο χάρτης μετατρέπεται σε θεματικός. Σε γενικές γραμμές, η τοπογραφική βάση που απαιτεί ένας θεματικός χάρτης, μπορεί και να είναι αρκετά απλουστευμένη ή σχηματική, ώστε να μην αποσπά την προσοχή από το φαινόμενο που παρουσιάζει ο θεματικός χάρτης. Στην Εικόνα 2.12 φαίνεται η σύνθεση (c) του θεματικού χάρτη (b) με την τοπογραφική βάση (a). (Kraak & Ormeling, 2010, p42)

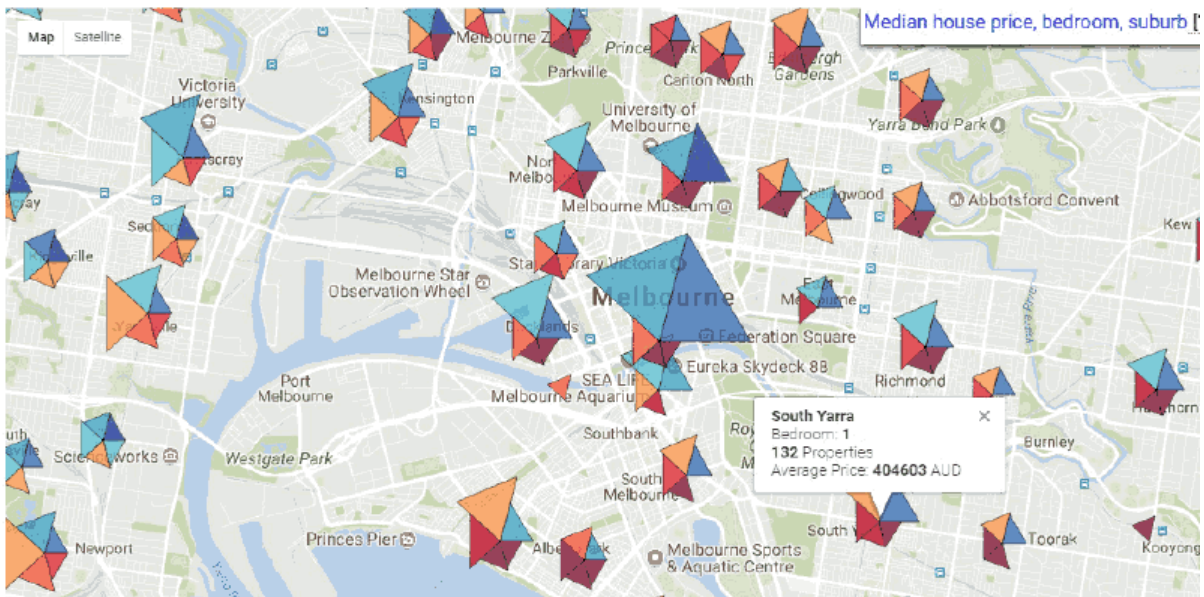


Εικόνα 2.12: Σύνθεση θεματικού χάρτη (Kraak & Ormeling, 2010)

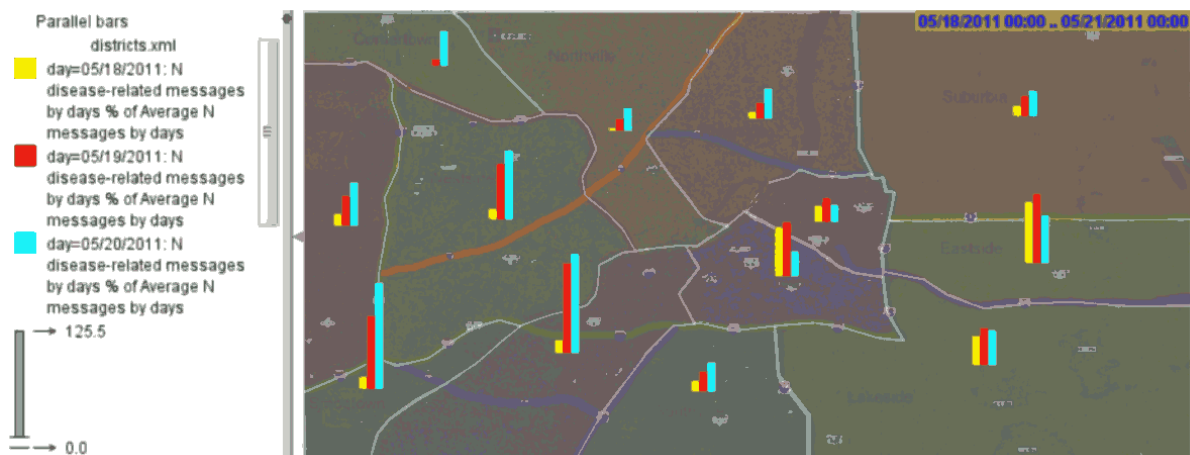
Ο χάρτης κουκίδων (*dot map*) (Εικόνα 2.15 a) χρησιμοποιείται για να αναδείξει την χωρική κατανομή φαινομένων, μέσα από το στοιχειώδες σύμβολο της κουκίδας. Κάθε κουκίδα μπορεί να αναπαριστά την ακριβή τοποθεσία κάθε φαινομένου ή, αν η πυκνότητά τους είναι μεγάλη, αναπαριστά συνάθροιση φαινομένων (πχ 1000 ανά κουκίδα). Ο μονόχρωμος κύκλος είναι το πιο συνηθισμένο σχήμα κουκίδας, αλλά μπορούν να απεικονιστούν κατηγορίες

κάνοντας χρήση άλλων σχημάτων και χρωμάτων, αν και χρειάζεται προσοχή, διότι σε περίπτωση που οι κουκίδες είναι πολύ μικρές, τα χρώματα είναι δυσδιάκριτα. Επίσης χρειάζεται προσοχή στην επιλογή τόσο του μεγέθους της κουκίδας όσο και στον αριθμό των φαινομένων που αντιπροσωπεύει, διότι επηρεάζουν την αντίληψη του πόσο πυκνό ή αραιό είναι το φαινόμενο. (Meirelles, 2013, p130-132)

Σύμφωνα με τους Karimi & Karimi (2017, p85), μία μέθοδος απεικόνισης πολυμεταβλητών δεδομένων πάνω στον χάρτη, είναι με την χρήση *συμβόλων τύπου ραντάρ με μορφή πίτας (pie radar symbols)*. Τα σύμβολα αυτά μπορούν να απεικονίσουν τιμές είτε για μεμονωμένα είτε για συγκεντρωτικά σημεία και η ακτίνα κάθε τομέα είναι ανάλογη της τιμής μία μεταβλητής. Για καλύτερη ευκρίνεια και βελτίωση της εμφάνισης, ταξινομούν τα μικρότερα σύμβολα έτσι ώστε να εμφανίζονται πάνω από τα μεγαλύτερα και να μην κρύβονται, ενώ προσθέτουν ένα ποσοστό διαφάνειας, ώστε να μην κρύβονται τα μεγαλύτερα σύμβολα. Την τεχνική αυτή χρησιμοποιεί ο Li (2019, p46) για να απεικονίσει συγκεντρωτικά τρεις μεταβλητές ολοκληρωμένων πωλήσεων σπιτιών πάνω στον χάρτη (Εικόνα 2.13). Συγκεκριμένα ο κάθε τομέας αφορά διαφορετικό αριθμό δωματίων, το μέγεθος του τομέα σχετίζεται με τον αριθμό των σπιτιών και το χρώμα του με την μέση τιμή πώλησης. Οι Andrienko et al. (2020, p12) περιγράφουν μια παραλλαγή της τεχνικής, όπου αντί για σύμβολα πίτας χρησιμοποιούν ραβδογράμματα (Εικόνα 2.14).



Εικόνα 2.13: Χάρτης με σύμβολα τύπου ραντάρ με μορφή πίτας (Li, 2019, p46)



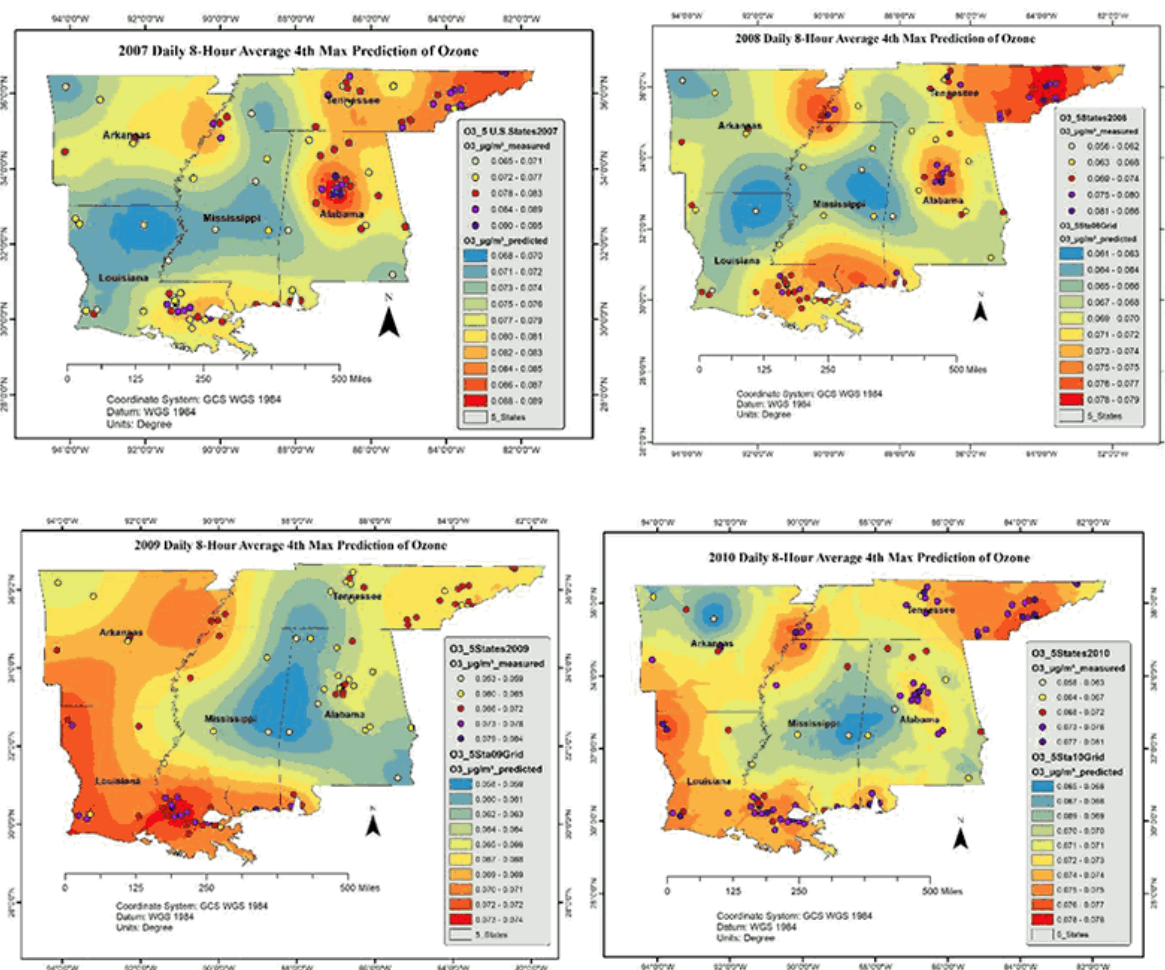
Εικόνα 2.14: Χάρτης με ραβδογράμματα για σύμβολα σημείων (Andrienko et al., 2020, p12)

Ο χωροπληθής (*choropleth*) χάρτης (Εικόνα 2.15 b) χρησιμοποιείται ευρέως για την απεικόνιση γεωγραφικών πληροφοριών ανά περιοχή, μαζί με συναφή δεδομένα, όπως ο πληθυσμός της. Τα δεδομένα που απεικονίζονται είναι συγκεντρωτικά και συνήθως αφορούν διοικητικές ενότητες, όπως νομούς ή δήμους. Ωστόσο, οι διοικητικές ενότητες συχνά δεν είναι ιδανικές, καθώς μπορεί να διαφέρουν έντονα σε μέγεθος και δεν αντιπροσωπεύουν απαραίτητα φυσικά όρια, και τα αποτελέσματα μπορούν να παρερμηνευτούν σε περίπτωση ελλιπών ή αραιών δεδομένων.



Εικόνα 2.15: Χάρτες: (a) κουκίδων, (b) χωροπληθής, (c) ισarithμικός (Kraak & Ormeling, 2010, p142)

Ο θερμικός χάρτης (*heatmap*) (Εικόνα 2.16) χρησιμοποιείται επίσης για την εμφάνιση γεωγραφικών πληροφοριών με βάση την περιοχή. Σε αντίθεση με τους χωροπληθείς χάρτες, προϋποθέτουν ότι τα δεδομένα στην επιφάνεια που απεικονίζουν είναι συνεχής, και έχουν πιο ακριβή αποτελέσματα. Χρησιμοποιούν διακριτές ή συνεχείς χρωματικές κλίμακες για τη χαρτογράφηση είτε ονομαστικών είτε ποσοτικών δεδομένων. Επειδή τα δεδομένα δεν είναι συνήθως διαθέσιμα για κάθε σημείο του χάρτη, οι τιμές που λείπουν συχνά παρεμβάλλονται. (Preim & Lawonn, 2020, p556)



Εικόνα 2.16: Θερμικοί χάρτες (Kethireddy et al., 2018)

Ο *ισαριθμικός (isarithmetic)* χάρτης (Εικόνα 2.15 c) απεικονίζει ένα συνεχές φαινόμενο μέσα από γραμμές περιγράμματος (contour lines), κάθε μία από τις οποίες αντιστοιχεί σε μία τιμή του φαινομένου. Αν το φαινόμενο μετριέται από ένα σημείο, όπως η θερμοκρασία από έναν αισθητήρα θερμοτήτας, τότε ονομάζεται *ισομετρικός (isometric)* χάρτης. Αν το φαινόμενο δεν μετριέται από ένα σημείο, όπως ο πληθυσμός μιας περιοχής, και η μέτρηση αφορά το κέντρο βάρους (centroid) του φαινομένου, τότε ονομάζεται *ισοπληθής (isopleth)* χάρτης. Σε κάθε περίπτωση, για την παραγωγή των περιγραμμάτων χρησιμοποιούνται τεχνικές παρεμβολής (interpolation). (Ward et al., 2015, p239)

Σε αντίθεση με τους χωροπληθείς και τους θερμικούς χάρτες, οι ισარიθμικοί και οι χάρτες κουκίδων, είναι αραιές αναπαραστάσεις και αφήνουν χώρο για την απεικόνιση των υποκείμενων γεωγραφικών δεδομένων (Preim, 2020).

2.4.2 Διάδραση

Παρά την μεγάλη ποσότητα πληροφορίας που μπορεί να μεταδώσει μια εικόνα, μια απλή στατική εικόνα είναι συχνά ανεπαρκής για την κατανόηση και την έγκυρη ανάλυση ενός πολύπλοκου θέματος. Συνήθως ένας αναλυτής χρειάζεται να εξετάσει τα δεδομένα από διαφορετικές οπτικές γωνίες, καθώς και να δει διαφορετικές πτυχές ή μέρη των δεδομένων. Για να το πετύχει αυτό, πρέπει να αλληλεπιδρά με τα δεδομένα και με το σύστημα που παράγει οπτικές απεικονίσεις των δεδομένων, δηλαδή να επιλέγει στοιχεία και υποσύνολα δεδομένων για προβολή, να επιλέγει και να συντονίζει τεχνικές οπτικοποίησης, να μετασχηματίζει τις προβολές, να μετασχηματίζει τα δεδομένα κτλ. Στην συνέχεια παρατίθενται οι τύποι αλληλεπίδρασης με γραφικές αναπαραστάσεις.

Η *αλλαγή τρόπου αναπαράστασης* είναι χρήσιμη για την παρατήρηση των ίδιων δεδομένων αλλά από διαφορετικές προοπτικές, ανάλογα με την επιτελούμενη εργασία. Για να επιτευχθεί η αλλαγή τρόπου αναπαράστασης στα δεδομένα, καθώς και στα χαρακτηριστικά του μέσου απεικόνισης και της αναλυτικής εργασίας, εφαρμόζονται τροποποιήσεις όπως αλλαγές στο μέγεθος και την αναλογία του πλαισίου απεικόνισης, καθώς και στην κλίμακα, στα χρώματα και στα διαστήματα διαχωρισμού μεταξύ των κλάσεων, ή και σε άλλες παραμέτρους της απεικόνισης.

Η *εστίαση και λήψη λεπτομερειών* υλοποιείται με διάφορες τεχνικές. Η λήψη συγκεκριμένων τιμών και η προβολή λεπτομερειών ή σχετιζόμενων πληροφοριών, πατώντας πάνω σε σύμβολα, είναι πλέον αναμενόμενη δυνατότητα των γραφικών οθονών. Η εστίαση (zooming) και η ολίσθηση (panning) ενός συνεχόμενου τμήματος της οθόνης, επιτρέπει την παρουσίαση μέρους των δεδομένων σε αυξημένη οπτική ανάλυση. Η ανάθεση μιας κλίμακας χρωμάτων σε ένα μέρος του εύρους των δεδομένων και η απόκρυψη των τιμών εκτός του εύρους, παρέχει καλύτερη ευκρίνεια στο επιλεγμένο εύρος δεδομένων. Η ταξινόμηση χρησιμοποιείται συχνά σε πίνακες για την ανάδειξη συσχετισμών μεταξύ χαρακτηριστικών,

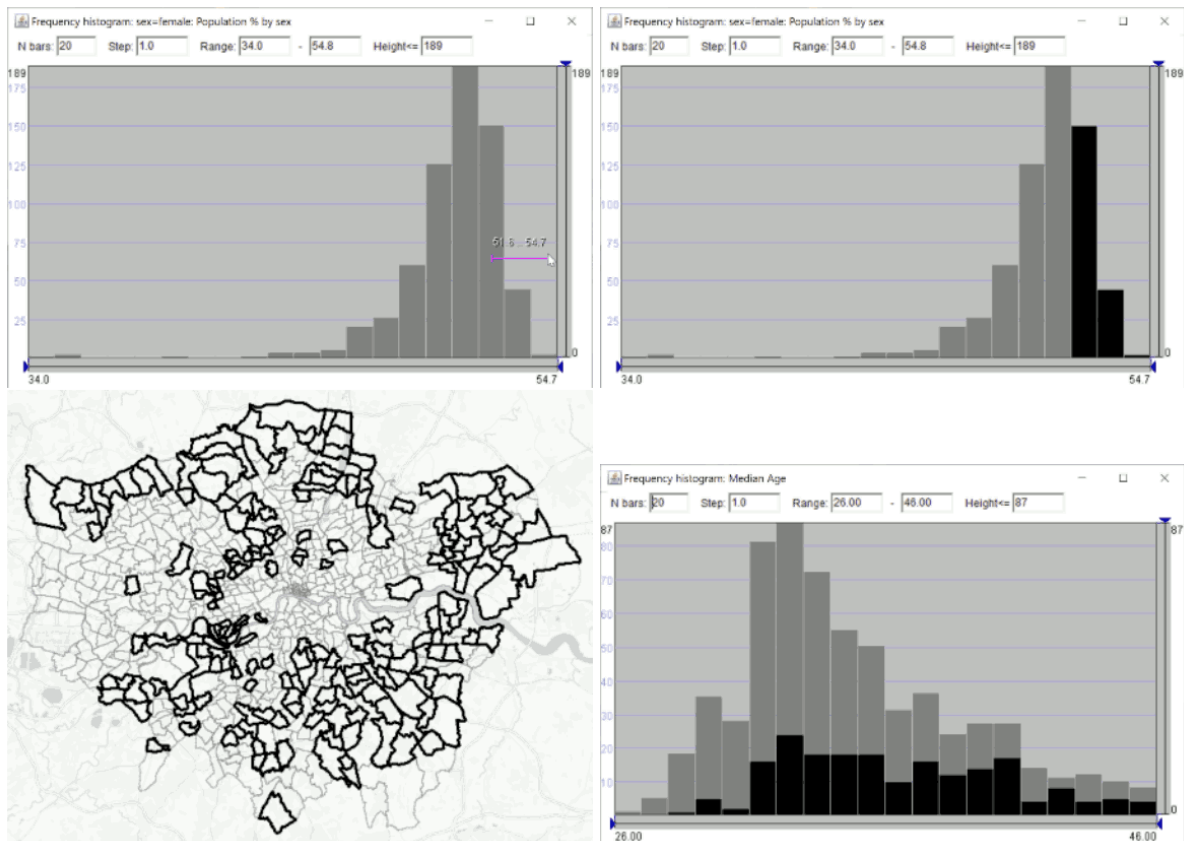
αλλά μπορεί να εφαρμοστεί και σε άξονες διαγραμμάτων όπως των παράλληλων συντεταγμένων και των πινάκων διαγραμμάτων διασποράς.

Μέσω της εξερεύνησης των δεδομένων συχνά στοχεύουμε να αποκτήσουμε μια πιο καθαρή εικόνα, να απλοποιήσουμε την απεικόνιση και να μειώσουμε την ποσότητα των δεδομένων ή των λεπτομερειών τους. Αυτό μπορούμε να το πετύχουμε *μετασχηματίζοντας τα δεδομένα*. Με την διακριτοποίηση (discretisation) μπορούμε να χωρίσουμε μια συνεχή μεταβλητή, πχ με ένα εύρος πραγματικών αριθμών, σε κλάσεις διαστημάτων. Μετατρέποντας τις τιμές μιας μεταβλητής σε λογαριθμική κλίμακα, σε περιπτώσεις που η πλειονότητα των τιμών είναι μικρές και υπάρχουν λίγες τιμές που είναι πολύ μεγάλες, μπορούμε να διακρίνουμε καλύτερα τις μικρές τιμές, αν και η ερμηνεία της απεικόνισης γίνεται δυσκολότερη. Η συνάθροιση δεδομένων με βάση κάποιο κριτήριο (πχ κατηγορίες, κλάσεις, χρονικά ή χωρικά διαστήματα) σε ένα περιληπτικό στατιστικό όπως ο μέσος όρος, βοηθά στην μείωση της ποσότητας της πληροφορίας και στην απόκτηση συνολικής άποψης ενός θέματος. Όταν κάποια αναλυτική διεργασία δεν απαιτεί δεδομένα μεγάλης λεπτομέρειας, μπορεί να γίνει συγχώνευση σε αθροίσματα, ομαδοποιώντας τα δεδομένα με βάση κάποιο χαρακτηριστικό, πχ πωλήσεις ομάδων προϊόντων αντί για πωλήσεις κάθε μεμονωμένου προϊόντος. Σε πολυμεταβλητά δεδομένα, προκειμένου να μειώσουμε το φορτίο της απεικόνισης, μπορούμε να επιλέξουμε να αναπαραστήσουμε μόνο το κυρίαρχο ή το πιο αντιπροσωπευτικό, από μια ομάδα χαρακτηριστικών. Η εξομάλυνση (smoothing) δεδομένων που περιέχουν χρονοσειρές ή γενικότερα μεταβλητές με συνεχόμενη κατανομή τιμών, μειώνει τις σύντομες αυξομειώσεις, και έτσι βοηθά στον καλύτερο εντοπισμό μοτίβων. Ένας ακόμη μετασχηματισμός χρονικών δεδομένων είναι η απεικόνιση της μεταβολής τους σε σχέση με μια αντίστοιχη προγενέστερη περίοδο, για παράδειγμα, η μεταβολή μηνιαίων δεδομένων σε σχέση με τα αντίστοιχα του προηγούμενου έτους.

Η *επιλογή και το φιλτράρισμα* είναι δύο πολύ χρήσιμες λειτουργίες της εξερευνητικής ανάλυσης δεδομένων. Κατά την επιλογή, λαμβάνονται τα επιθυμητά δεδομένα και απεικονίζονται, ενώ κατά το φιλτράρισμα, απορρίπτονται από την απεικόνιση τα αδιάφορα δεδομένα. Και τα δύο, πραγματοποιούνται βάσει κριτηρίων όπως οι τιμές χαρακτηριστικών, η θέση στον χώρο ή στον χρόνο, και η σχέση με άλλες οντότητες και δεδομένα.

Πραγματοποιείται είτε με γραφικά στοιχεία ελέγχου, όπως μοχλούς κύλισης, είτε απευθείας στην απεικόνιση των δεδομένων, όπως χάρτες, είτε πάνω σε διαγράμματα κατανομής, όπως ιστογράμματα συχνότητας. Μια κοινή τεχνική επιλογής, είναι το βούρτσισμα (brushing), που γίνεται σύροντας το ποντίκι πάνω από την οθόνη απεικόνισης, αλλά μπορεί να γίνει και επιλέγοντας μεμονωμένα στοιχεία. Συχνά, τα επιλεγμένα δεδομένα απεικονίζονται μαζί με τα υπόλοιπα, αλλά αναδεικνύονται με χρήση διαφορετικών οπτικών χαρακτηριστικών. Χρησιμοποιώντας χρώματα ή στυλ, είναι δυνατό να απεικονιστούν ταυτόχρονα πολλά διαφορετικά επιλεγμένα υποσύνολα δεδομένων. Εκτός από την ρητή επιλογή, είναι δυνατό να επιλεγεί ένα υποσύνολο δεδομένων με κριτήριο την εγγύτητά τους σε ένα σημείο, ενός χώρου που μπορεί να αποτελείται από πολλές διαστάσεις.

Μια κοινή προσέγγιση αλληλεπίδρασης για την ανάλυση δεδομένων, που πρέπει να παρατηρηθούν από διαφορετικές προοπτικές ή που δεν μπορούν να αναπαρασταθούν όλα τα στοιχεία τους σε μία απεικόνιση, είναι οι *συνδεδεμένες ή συγχρονισμένες προβολές (linked or coordinated views)* (Εικόνα 2.17). Τεχνικές όπως η επιλογή, το φιλτράρισμα και το βούρτσισμα, που αναφέρθηκαν παραπάνω, χρησιμοποιούνται για να εντοπίσουν αντιστοιχίσεις τμημάτων της πληροφορίας στις οθόνες απεικόνισής τους. Ένας μηχανισμός συντονισμού προβολών, είναι ο *κοινός συμβολισμός (common symbolisation)*, όπου οπτικά χαρακτηριστικά συμβόλων, όπως το χρώμα, διαδίδονται από μία γραφική απεικόνιση στα αντίστοιχα σύμβολα άλλης απεικόνισης. Μια άλλη τεχνική συγχρονισμού, είναι η δημιουργία πολλαπλών προβολών, μία για κάθε κατηγορία ή υποσύνολο των δεδομένων που απεικονίζονται σε άλλη προβολή. Παρόλο που η προσέγγιση των συνδεδεμένων προβολών είναι πολύ χρήσιμη, και σε κάποιες περιπτώσεις η μόνη λύση για την ανάλυση σύνθετων δεδομένων, κάποιες φορές μπορεί να αποσπάσει την προσοχή του αναλυτή ή να είναι υπολογιστικά απαιτητική. Σε αυτές τις περιπτώσεις, ο τελικός χρήστης πρέπει να έχει τον έλεγχο ενεργοποίησης της δυνατότητας αυτής στις προβολές της επιλογής του. (Andrienko et al., 2020, p75-87)



Εικόνα 2.17: Συνδεδεμένες ή Συγχρονισμένες Προβολές (Andrienko et al., 2020, p84)

2.4.3 Αναλυτικός συλλογισμός (analytical reasoning)

Η επιστήμη του αναλυτικού συλλογισμού παρέχει το συλλογιστικό πλαίσιο στο οποίο μπορεί κανείς να οικοδομήσει στρατηγικές και τακτικές τεχνολογίες οπτικής αναλυτικής, και είναι κεντρικής σημασίας στο έργο του αναλυτή, το οποίο είναι να εφαρμόζει την ανθρώπινη κρίση για την εξαγωγή συμπερασμάτων από συνδυασμό στοιχείων και υποθέσεων (Thomas & Cook, 2005, p33). Πολλά προβλήματα που προσπαθεί να επιλύσει η οπτική αναλυτική αφορούν τη διαχείριση και την κατανόηση της τεράστιας ροής δεδομένων. Μια σημαντική πτυχή της επιστήμης του αναλυτικού συλλογισμού είναι η δημιουργία τρόπων αναπαράστασης δεδομένων σε μορφές που επιτρέπουν την αλληλεπίδραση και τη μετάβαση των διαδικασιών σκέψης από τα δεδομένα σε πληροφορίες, από τις πληροφορίες σε νόημα και από το νόημα στην κατανόηση. Άλλα προβλήματα έχουν να κάνουν με τη βελτίωση της

οπτικής αναπαράστασης, που σχετίζονται με τις βασικές αρχές της αντίληψης και της νόησης (Thomas & Cook, 2005, p50).

Ο όρος νόηση αναφέρεται στον τρόπο με τον οποίο οι άνθρωποι αποκτούν και χρησιμοποιούν τη γνώση μέσω των αισθήσεών τους, μαθαίνουν από τις αλληλεπιδράσεις και τις εμπειρίες στο περιβάλλον τους, και αποκτούν και βελτιώνουν την ικανότητά τους να εκτελούν λειτουργίες όπως το περπάτημα, η ομιλία, η οδήγηση αυτοκινήτου και η επίλυση προβλημάτων. Εικάζεται ότι η νόηση ενεργοποιείται από τις λειτουργίες ανώτερου επιπέδου του εγκεφάλου. Μια νοητική λειτουργία αναφέρεται στα συγκεκριμένα βήματα που χρησιμοποιεί ο εγκέφαλος για την επιτέλεση εργασιών, όπως η αντίληψη, ο σχεδιασμός, η εκμάθηση γλώσσας και η σκέψη. Οι νοητικές λειτουργίες διαφέρουν από τους ντετερμινιστικούς αλγορίθμους, καθώς αντιμετωπίζουν με ευχέρεια δεδομένα που είναι διαφορετικά, αβέβια, ελλιπή και ασυνεπή, χρησιμοποιώντας πιθανοτικούς αλγορίθμους. (Gudivada et al., 2016, p190)

Ο διάλογος ανθρώπου-πληροφορίας στην οπτική αναλυτική πραγματοποιείται επιτυχημένα όταν οι μηχανισμοί πρόσβασης και χειρισμού των εργαλείων της χάνονται στο παρασκήνιο και επιτρέπουν την απρόσκοπτη ροή επίλυσης προβλημάτων. Η έννοια της ροής έχει τις ρίζες της στην ψυχολογία και η εφαρμογή των αρχών της σε διαδραστικά συστήματα δεν έχει ακόμη επιτευχθεί και αποτελεί μείζονα ερευνητική πρόκληση της οπτικής αναλυτικής.

Ένα βασικό πρόβλημα για την οπτική αναλυτική προκύπτει από τις περιορισμένες ικανότητες της ανθρώπινης αντίληψης και νόησης, π.χ. τα όρια της βραχυπρόθεσμης μνήμης. Για να παρακάμψουμε αυτά τα όρια, χρησιμοποιούμε εξωτερικά βοηθήματα, για να μπορέσουμε να αντιμετωπίσουμε τον όγκο και την πολυπλοκότητα των πληροφοριών που θέλουμε να χρησιμοποιήσουμε. Η οπτική αναλυτική είναι ακριβώς ένα τέτοιο εξωτερικό βοήθημα. Για να επιτύχουμε τη ροή του αναλυτικού διαλόγου, πρέπει να κατανοήσουμε καλύτερα την αλληλεπίδραση μεταξύ αντίληψης και νόησης, και πώς αυτές επηρεάζονται όταν εργαζόμαστε με ένα δυναμικό εξωτερικό βοήθημα. (Thomas & Cook, 2005, p50)

Ως εξειδίκευση της οπτικής αναλυτικής, η γεωοπτική αναλυτική εστιάζει στον αναλυτικό συλλογισμό, δηλαδή την διαδικασία της εξέτασης πληροφοριών για την εύρεση μοτίβων. Έτσι, αντί να επικεντρώνεται στην οπτικοποίηση, η γεωοπτική αναλυτική έχει ως στόχο να

υποστηρίζει τον άνθρωπο στο να αναλύει και να κατανοεί τις γεωχωρικές πληροφορίες, χρησιμοποιώντας οπτικές διεπαφές. Επομένως, απαιτούνται νέες προσεγγίσεις όχι μόνο για τις οπτικές διεπαφές για τις γεωχωρικές πληροφορίες, αλλά και για τους υποκείμενους μηχανισμούς της ανθρώπινης συλλογιστικής και κατανόησης σχετικά με τις γεωχωρικές πληροφορίες. Η υποστήριξη του αναλυτικού συλλογισμού στη γεωοπτική αναλυτική συχνά έρχεται με τη μορφή υπολογιστικών μεθόδων που χρησιμοποιούνται για την ανίχνευση μοτίβων ή/και την πρόβλεψη αποτελεσμάτων. Μπορεί επίσης να είναι εμφανής στις διεπαφές με τις γεωχωρικές πληροφορίες, που είναι προσαρμοσμένες ώστε να υποστηρίζουν τις αναλυτικές διαδικασίες και τις ροές εργασίας που χρησιμοποιούν οι αναλυτές για τη διεξαγωγή της εργασίας τους. (Robinson, 2017).

2.4.4 Υπολογιστικές μέθοδοι

Οι προσεγγίσεις της οπτικής αναλυτικής συνδυάζουν διαδραστικές οπτικοποιήσεις με τη χρήση υπολογιστικών τεχνικών για την επεξεργασία και την ανάλυση δεδομένων. Ο συνδυασμός οπτικοποίησης και υπολογισμού έχει δύο πλευρές. Η μία πλευρά είναι η υπολογιστική υποστήριξη της οπτικής ανάλυσης, όπου τα αποτελέσματα των υπολογισμών εισέρχονται στην ανθρώπινη νόηση μέσω οπτικής αναπαράστασης. Η άλλη πλευρά είναι η οπτική υποστήριξη των υπολογιστικών μεθόδων, η οποία περιλαμβάνει την οπτική εξερεύνηση των ιδιοτήτων των δεδομένων για την προετοιμασία τους για υπολογισμούς, για την αξιολόγηση των αποτελεσμάτων των υπολογισμών και για τη σύγκριση των αποτελεσμάτων διαφορετικών εκτελέσεων υπολογιστικών τεχνικών. Ένας σημαντικός κοινός σκοπός της χρήσης υπολογιστικών μεθόδων στην οπτικοποίηση είναι η δυνατότητα επισκόπησης ογκωδών και πολύπλοκων δεδομένων.

Η σωστή χρήση των υπολογιστικών μεθόδων απαιτεί καταρχήν την επιλογή της κατάλληλης μεθόδου για την εκάστοτε αναλυτική εργασία, και στην συνέχεια την προετοιμασία των δεδομένων για την μέθοδο που έχει επιλεγεί. Μέρος της προετοιμασίας των δεδομένων είναι η διεξοδική έρευνα και καλή κατανόηση των ιδιοτήτων τους, η οποία δύσκολα μπορεί να επιτευχθεί χωρίς την βοήθεια της οπτικοποίησης. Δύο τέτοιες ιδιότητες των δεδομένων που μπορούν να αποκαλυφθούν με την βοήθεια της οπτικοποίησης είναι η ύπαρξη ακραίων τιμών

και ανόμοιων υποσυνόλων. Οι *ακραίες τιμές (outliers)* μπορούν να επηρεάσουν σημαντικά τα αποτελέσματα των υπολογιστικών μεθόδων, καθιστώντας τα άχρηστα ή παραπλανητικά. Συνήθως οι ακραίες τιμές εξαλείφονται, αλλά σε περίπτωση που χρειαστεί να διερευνηθεί και να κατανοηθεί η επίδρασή τους στα αποτελέσματα των υπολογισμών, η οπτικοποίηση είναι ένα κατάλληλο εργαλείο για το έργο αυτό. Όταν ένα σύνολο δεδομένων αποτελείται από πολύ *ανόμοια υποσύνολα*, συνήθως αντιμετωπίζουμε το κάθε υποσύνολο ξεχωριστά, εμπλέκοντας διαφορετικά τμήματά τους στην ανάλυση ή/και εφαρμόζοντας διαφορετικές ρυθμίσεις παραμέτρων ή ακόμη και διαφορετικές μεθόδους ανάλυσης για το καθένα.

Για την υπολογιστική υποστήριξη της οπτικής ανάλυσης οι δύο γενικές προσεγγίσεις είναι η χωροθέτηση, η οποία επιτυγχάνεται μέσω τεχνικών ενσωμάτωσης δεδομένων ή αλλιώς μείωσης των διαστάσεών τους, και η ομαδοποίηση, η οποία επιτυγχάνεται με τη χρήση αλγορίθμων συσταδοποίησης. Συναρτήσεις απόστασης είναι βοηθητικές και για τις δύο περιπτώσεις, και παρέχουν αριθμητικές εκτιμήσεις της ανομοιότητας μεταξύ στοιχείων των δεδομένων. (Andrienko et al., 2020, p89)

2.4.5 Κλιμακωσιμότητα (Scalability)

Σήμερα, ο ρυθμός παραγωγής δεδομένων αυξάνεται με ιλιγγιώδη ρυθμό, και παρόλο που οι αναλυτές έχουν στην διάθεσή τους τεράστιες ποσότητες πληροφορίας από πολλές πηγές, η σημαντική πληροφορία μπορεί να κρύβεται σε μερικά ψήγματα. Ορισμένες βασικές κατηγορίες δεδομένων από τις οποίες δημιουργείται αυτή η υπερφόρτωση πληροφορίας είναι τα κείμενα (πχ έγγραφα, ομιλίες, νέα, email, ιστοσελίδες), οι βάσεις δεδομένων (δημόσιες και ιδιωτικές), οι εικόνες (πχ δορυφορικές φωτογραφίες της γης), δεδομένα αισθητήρων (πχ θέσης, θερμοκρασίας, ακτινοβολίας) και τα βίντεο. Αποτελεί πρόκληση η ανάπτυξη μεθόδων που θα επιτρέπουν την εξέταση μεγάλων ποσοτήτων, πολυδιάστατων και μεταβαλλόμενων δεδομένων από πολλές πηγές, ώστε να μπορούν οι αναλυτές να λαμβάνουν αποτελεσματικές αποφάσεις σε κρίσιμες χρονικά καταστάσεις. Η πρόκληση δεν αφορά μόνο την ποσότητα, την ποικιλότητα και τον δυναμικό χαρακτήρα των δεδομένων, αλλά και το γεγονός ότι μπορεί να είναι ασαφή ή αβέβαια, και πιθανώς εσκεμμένα παραπλανητικά. Επίσης, στο πλαίσιο της υπερφόρτωσης πληροφοριών, οι ανθρώπινες δεξιότητες και ικανότητες δεν

αλλάζουν σημαντικά με την πάροδο του χρόνου, με αποτέλεσμα η πρόσβαση σε ποσότητα πληροφοριών πολύ μεγαλύτερη από αυτήν που μπορεί να διαχειριστεί ο άνθρωπος να οξύνει ακόμα περισσότερο το πρόβλημα. Όμως η κλίμακα της πληροφορίας που είναι διαθέσιμη έχει και κάποια θετικά στοιχεία, όπως την μεγαλύτερη πληρότητα δεδομένων και την μείωση της αβεβαιότητας για καταστάσεις που μελετούνται (Thomas & Cook, 2005, p29).

Η κλιμακωσιμότητα στην οπτική αναλυτική ορίζεται ως η δυνατότητα αποτελεσματικής αναπαράστασης ογκωδών δεδομένων, με τον όγκο να αφορά τόσο την ποσότητα όσο και τον αριθμό των διαστάσεων τους (Cui, 2019).

Για την αντιμετώπιση της αύξησης του αριθμού των διαστάσεων, οι Ward et al. (2015, p286) αναφέρουν τις εξής λύσεις για απεικόνιση με σημειακά διαγράμματα: 1) την επιλογή υποσυνόλου μόνο των πιο χρήσιμων διαστάσεων, 2) την μείωση των διαστάσεων χρησιμοποιώντας τεχνικές όπως την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis), 3) την απεικόνιση μικρού αριθμού επιπλέον διαστάσεων σε άλλες οπτικές μεταβλητές εκτός από την θέση, όπως το χρώμα, το μέγεθος και το σχήμα, 4) την προσθήκη οθονών διαγραμμάτων είτε με υπέρθεση, είτε με αντιπαράθεση, με το κάθε διάγραμμα να αναπαριστά ορισμένες από τις διαστάσεις. Επίσης αναφέρουν την χρήση γραμμικών διαγραμμάτων όπως των παράλληλων συντεταγμένων και των γραφημάτων κυκλικής γραμμής (circular line graph), ενώ αναφέρουν και χωρικά διαγράμματα όπως ραβδογράμματα, ιστογράμματα και διαγράμματα πινάκων.

Με την αύξηση της ποσότητας των δεδομένων, οι υπάρχουσες μέθοδοι οπτικοποίησης πάσχουν από 1) χαρτογράφηση υπερβολικά μεγάλου αριθμού στοιχείων (Εικόνα 2.18 a), 2) υπέρβαση των αντιληπτικών και γνωστικών ικανοτήτων των χρηστών (Εικόνα 2.18 b) και 3) από δυσκολία υποστήριξης διαδραστικής εξερεύνησης (Εικόνα 2.18 c), αφού η αλληλεπίδραση με δεδομένα μεγάλης κλίμακας θα οδηγούσε σε υψηλές καθυστερήσεις και μειωμένο βαθμό ανταπόκρισης. Για την αντιμετώπιση της μεγάλης ποσότητας δεδομένων έχουν προταθεί δύο είδη μεθόδων: 1) τεχνικής οπτικοποίησης όπως η προοδευτική οπτικοποίηση, η οπτικοποίηση εικονοστοιχείων, η χωρική εκτόπιση, η προβλεπτική οπτική αναλυτική και η παράλληλη επεξεργασία και 2) μέθοδοι ελάττωσης δεδομένων, όπως το φιλτράρισμα, ο δειγματισμός και η συσσώρευση σε δοχεία (Li, 2019, p6,7).



Εικόνα 2.18: Προβλήματα κλιμάκωσης της ποσότητας των δεδομένων: (a) υπερβολική χαρτογράφηση (b) υπέρβαση ικανοτήτων χρηστών (c) δυσκολία διάδρασης (Li, 2019, p7)

2.5 Μοντέλα Αυτοματοποιημένης Αποτίμησης (Automated Valuation Model)

Το Μοντέλο Αυτοματοποιημένης Αποτίμησης (AVM) είναι “ένα μαθηματικό πρόγραμμα λογισμικού υπολογιστή που χρησιμοποιούν οι αναλυτές της αγοράς για να παράγουν μια εκτίμηση της αγοραίας αξίας με βάση την ανάλυση της αγοράς όσον αφορά την τοποθεσία, τις συνθήκες της αγοράς και τα χαρακτηριστικά των ακινήτων, από πληροφορίες που είχαν συλλεχθεί ανεξάρτητα και σε προηγούμενο χρόνο”. Η αξιοπιστία ενός AVM εξαρτάται από τα δεδομένα που χρησιμοποιούνται και τις δεξιότητες του μελετητή που παράγει το μοντέλο, ο οποίος θα πρέπει να είναι καταρτισμένος αναλυτής της αγοράς, π.χ. εκτιμητής/αξιολογητής, που χρησιμοποιεί εφαρμογές που βασίζονται σε στατιστικά στοιχεία για να αναλύουν τα δεδομένα και να επιλέγουν την καλύτερη προσομοίωση της δραστηριότητας της αγοράς για την ανάλυση της τοποθεσίας, των συνθηκών της αγοράς και των χαρακτηριστικών του ακινήτου από δεδομένα που έχουν συλλεχθεί προηγουμένως. Τα AVM έχουν σχεδιαστεί για να παράγουν εκτιμήσεις αξίας για ακίνητα σε συγκεκριμένες χρονικές στιγμές του παρελθόντος ή του μέλλοντος, ανάλογα με τις απαιτήσεις του πελάτη (ΙΑΑΟ, 2018).

Ο κύριος σκοπός ενός AVM είναι να παρέχει αποτελεσματικά, μια ακριβή, ομοιόμορφη, και δίκαιη εκτίμηση της αγοραίας αξίας ενός ακινήτου. Πέρα από τον σκοπό αυτό της αποτίμησης, υπάρχουν και δευτερεύοντες σκοποί του AVM, οι οποίοι μπορεί να είναι: φορολογική εκτίμηση, αξιολόγηση κινδύνου χαρτοφυλακίου ακινήτων ή ασφάλισης ή

δανεισμού, αναπροσαρμογή της γης πχ αναδασμός, επένδυση σε ακίνητα, προσδιορισμός περιθωρίου διαπραγμάτευσης. Τα AVM μπορούν να χρησιμοποιηθούν για εκτίμηση γης, οικιστικών και μη οικιστικών ακινήτων (Glumac & Des Rosiers, 2020).

Η μαζική αποτίμηση είναι η διαδικασία εκτίμησης της αγοραίας αξίας μιας ομάδας ακινήτων ίδιου τύπου, με τη χρήση μοντέλων αυτοματοποιημένης αποτίμησης (AVM), δηλαδή υπολογιστικά μοντέλα που βασίζονται κυρίως στην ανάλυση πολλαπλής παλινδρόμησης (multiple regression analysis, MRA) και μερικές φορές μπορεί να βασίζονται σε ένα σύστημα γεωγραφικών πληροφοριών (GIS) ή σε άλλες τεχνικές τεχνητής νοημοσύνης. Η ανάπτυξη ενός AVM περιλαμβάνει τρία βασικά στάδια: τον καθορισμό των προδιαγραφών του μοντέλου, την βαθμονόμισή του και την δοκιμή και διασφάλιση της ποιότητάς του. Ο καθορισμός των προδιαγραφών αφορά α) τον ορισμό της μεθόδου εκτίμησης, πχ με βάση το κόστος, με βάση την αγορά ή με βάση το εισόδημα, που στην συνέχεια μετατρέπεται σε μια μαθηματική φόρμουλα που μπορεί να είναι προσθετική, πολλαπλασιαστική ή υβριδική και β) την επιλογή των ανεξάρτητων μεταβλητών που θα πρέπει να συμπεριληφθούν στο μοντέλο ως παράγοντες πρόβλεψης της αγοραίας αξίας, δηλαδή της εξαρτημένης μεταβλητής. Η βαθμονόμιση είναι η διαδικασία δοκιμασίας της δομής του μοντέλου για την εκτίμηση των συντελεστών και των παραμέτρων των μεταβλητών, με την χρήση ενός διαφορετικού συνόλου δεδομένων από εκείνο του προηγούμενου σταδίου. Τα δύο αυτά στάδια, δηλαδή του ορισμού των προδιαγραφών και της βαθμονόμισης, επαναλαμβάνονται μέχρι οι δείκτες απόδοσης του μοντέλου να φτάσουν σε αποδεκτά επίπεδα. Το τρίτο στάδιο, της δοκιμής και της διασφάλισης της ποιότητας, αφορά την δοκιμή των επιδόσεων του μοντέλου που πραγματοποιείται με ένα υποσύνολο των ακινήτων, το οποίο δεν χρησιμοποιήθηκε στο προηγούμενο στάδιο της βαθμονόμισης. Για την εκτίμηση των επιδόσεων χρησιμοποιούνται διάφορα βασικά μέτρα όπως το επίπεδο εκτίμησης (μέσος όρος, διάμεσος, σταθμισμένος μέσος όρος) για την ακρίβεια, την μεταβλητότητα-ομοιομορφία που αντανάκλα τη συνέπεια, και την αξιοπιστία (Demetriou, 2018).

Οι Nitschke et al. (2022) αναφέρουν δύο είδη μοντέλων AVM, τα συγκριτικά (comparables) και τα ωφελμιστικά (hedonic).

Τα ωφελμιστικά μοντέλα βασίζονται στην παραδοχή ότι η αξία ενός ακινήτου είναι συνάρτηση των επιμέρους χαρακτηριστικών του και αποτυπώνουν πώς τα διάφορα αυτά χαρακτηριστικά, όπως ο τύπος ακινήτου, το εμβαδόν, το έτος κατασκευής, ο αριθμός δωματίων και άλλα, επηρεάζουν την αξία του, σε μια δεδομένη χρονική στιγμή και για μια δεδομένη γεωγραφική περιοχή. Για την δημιουργία του μοντέλου επιλέγεται μια μαθηματική εξίσωση, στην οποία εισάγονται παράμετροι για κάποια από τα χαρακτηριστικά των ακινήτων. Οι παράμετροι αυτές εκτιμούνται από ένα σύνολο δεδομένων ακινήτων βαθμονόμησης, για τα οποία τα χαρακτηριστικά τους, η τιμή τους και τα χαρακτηριστικά της περιοχής στην οποία βρίσκονται, είναι γνωστά. Τελικά, η αξία ενός ακινήτου υπολογίζεται μέσα από τη λύση μιας εξίσωσης που περιέχει τα χαρακτηριστικά του, και ενδεχομένως βασίζεται στην γενική τοποθεσία του.

Τα συγκριτικά μοντέλα βασίζονται στην παραδοχή ότι η αξία ενός ακινήτου μπορεί να εκτιμηθεί βέλτιστα με βάση μια εξατομικευμένη επιλογή συγκριτικών στοιχείων ακινήτων, και συνήθως περιλαμβάνουν δύο στάδια. Στο πρώτο στάδιο γίνεται αυτόματη επιλογή κατάλληλων συγκριτικών ακινήτων από βάσεις δεδομένων, ανάλογα με τα χαρακτηριστικά του ακινήτου προς εκτίμηση (π.χ. τύπος ακινήτου, εμβαδόν, έτος κατασκευής) και την τοποθεσία του (π.χ. εγγύτητα, συγκρισιμότητα της τοποθεσίας). Στο δεύτερο στάδιο χρησιμοποιούνται οι τιμές αυτών των συγκριτικών ακινήτων για την αποτίμηση του, προς εκτίμηση, ακινήτου. Ανάλογα με την πολυπλοκότητα του μοντέλου, τα βήματα αυτά μπορούν να επαναληφθούν και να ληφθούν υπόψη πρόσθετα δεδομένα. (Nitschke et al., 2022)

2.6 Γεωχωρικά Δεδομένα

Σύμφωνα με την Ανοικτή Γεωχωρική Κοινοπραξία (Open Geospatial Consortium, OGC), που είναι κορυφαίος φορέας στην ανάπτυξη προτύπων για GIS, πρόκειται για ιδιότητες τοποθεσιών, που σχετίζονται με οποιοδήποτε επίγειο χαρακτηριστικό ή φαινόμενο. Οι ιδιότητες τοποθεσίας μπορεί να περιλαμβάνουν οποιοσδήποτε πληροφορίες σχετικά με τοποθεσίες ή περιοχές και τις σχέσεις μεταξύ τους, καθώς και περιγραφικές πληροφορίες σχετικά με γεωγραφικά χαρακτηριστικά και φαινόμενα. Αυτό περιλαμβάνει δεδομένα τηλεπισκόπησης, δεδομένα διανυσματικών χαρτών, διευθύνσεις, συντεταγμένες κτλ. Σε

πολλές περιπτώσεις, σαν όρος, το “γεωχωρικά δεδομένα” είναι πιο ακριβής από το “γεωγραφικά δεδομένα”, επειδή τα γεωχωρικά δεδομένα χρησιμοποιούνται συχνά με τρόπους που δεν περιλαμβάνουν γραφική αναπαράσταση, ή χάρτη, που δημιουργήθηκε από τα δεδομένα. (<https://www.ogc.org/ogc/glossary/g>)

Η OGC έχει συντάξει δύο πολύ σημαντικά πρότυπα υπηρεσιών για την ανταλλαγή γεωχωρικών δεδομένων, το WMS (Web Map Specification) και το WFS (Web Feature Specification). Πρόκειται για πρότυπα που έχουν σημαντικές δυνατότητες για την ομαλοποίηση και συνεπώς τη βελτίωση του τρόπου με τον οποίο τα δεδομένα διαμοιράζονται μέσω του Διαδικτύου. Η υπηρεσία WMS περιγράφει μηχανισμούς επικοινωνίας που επιτρέπουν σε ασύνδετα προϊόντα λογισμικού να ζητούν και να παρέχουν προκατασκευασμένες εικόνες χάρτη, οι οποίες μπορεί να περιέχουν διανυσματικά (vector) και εικονιστικά (raster) δεδομένα. Η WFS, επιτρέπει αιτήματα για διανυσματικά δεδομένα και επιστρέφει αποκλειστικά διανυσματικά δεδομένα, χωρίς απώλεια πληροφορίας κατά την μετατροπή σε εικόνα, όπως το WMS. Το χαρακτηριστικό (feature) στην WFS, αναφέρεται σε διανυσματικό αντικείμενο, όπως σημείο, γραμμή ή πολύγωνο. Η WMS χρησιμοποιεί XML για την απαρίθμηση των διαθέσιμων δεδομένων στον διακομιστή ή για τα αιτήματα για εικόνες χάρτη. Η WFS κάνει πιο εκτεταμένη χρήση XML, τόσο για τα αιτήματα για απαρίθμηση των διαθέσιμων δεδομένων, αλλά και για τα ίδια τα διανυσματικά δεδομένα που επιστρέφονται, σε μια διάλεκτο XML την GML. Το πιο γνωστό πακέτο λογισμικού που υλοποιεί WFS και WMS είναι το MapServer, το οποίο λειτουργεί τόσο ως πελάτης όσο και ως διακομιστής για δεδομένα WFS και WMS. (Michaelis & Ames, 2008)

Επιπλέον, υπάρχουν ορισμένοι κοινοί τύποι αρχείων διαμοιρασμού γεωχωρικών δεδομένων, όπως τα GeoPackage, Shapefile και GeoJSON, αλλά και απλοί τύποι αρχείων κειμένου, όπως το CSV, που μπορούν να περιέχουν πεδία γεωχωρικών δεδομένων σε μορφή Well-known text (WKT) (Mooney & Minghini, 2022).

2.6.1 Διαθεσιμότητα

Τα περισσότερα δεδομένα δεν είναι άμεσα διαθέσιμα στην μορφή που χρειάζεται για να αξιοποιηθούν. Αρχικά μπορεί να απαιτηθεί κάποια διαδικασία για να εξαχθούν από την πηγή

τους όπως κατέβασμα και αποσυμπίεση ή αποδελτίωση, στην συνέχεια μπορεί να χρειαστεί να εφαρμοστεί κάποιος μετασχηματισμός, και τέλος, καταχώρηση σε κάποιο αποθετήριο όπως μια βάση δεδομένων. Η διαδικασία αυτή είναι γνωστή ως Εξαγωγή, Μετασχηματισμός, Εισαγωγή (Extract, Transform, Load), δηλαδή εξαγωγή δεδομένων από την πηγή, μετατροπή στην κατάλληλη μορφή και εισαγωγή στο αποθετήριο. Συνήθως συνδυάζει πολλές πηγές ταυτόχρονα και εφαρμόζει τεχνικές διαδικτυακής εξόρυξης (web mining). (Muggenhuber, 2019)

2.6.2 Δομή και Μορφή

Τα πηγαία δεδομένα μπορεί να είναι δομημένα, ημιδομημένα ή αδόμητα. Στα δομημένα δεδομένα συγκαταλέγονται πηγές με αυστηρά καθορισμένο σχήμα (schema), όπως γεωχωρικές υπηρεσίες και βάσεις δεδομένων, καθώς και αρχεία χαρακτήρων κειμένου όπως XML, JSON, CSV κτλ, οι οποίες μπορούν να χρησιμοποιηθούν ως έχουν ή με μικρές επεξεργασίες ή μετατροπές, όπως κανονικοποίηση ή αλλαγή τύπων, ονομάτων πεδίων και συστήματος αναφοράς. Τα ημιδομημένα δεδομένα συνήθως βρίσκονται σε πίνακες εντός κειμένων φυσικής γλώσσας, όπως ιστοσελίδες, email, pdf και αρχεία κειμενογράφου, όπου η δομή και ο τύπος δεν είναι σαφώς ορισμένα, αλλά υπονοούνται, και θα πρέπει να καταβληθεί προσπάθεια για την ανίχνευσή τους. Τα αδόμητα είναι ουσιαστικά γλωσσολογικά δεδομένα, σε μορφή κειμένου, ήχου ή εικόνας, όπου θα πρέπει να χρησιμοποιηθούν τεχνικές σημασιολογικής ανάλυσης, όπως εργαλεία Επεξεργασίας Φυσικής Γλώσσας (Natural Language Processing), για να εξαχθούν πληροφορίες από αυτά. (Thomas & Cook, 2005)

2.6.3 Αποδελτίωση (Scraping)

Με τον όρο αποδελτίωση εννοούμε την αυτόματη ανάκτηση δεδομένων από το διαδίκτυο, με χρήση εργαλείων, τα οποία είτε υπάρχουν ήδη είτε αναπτύσσονται κατά περίπτωση. Χρησιμοποιείται συχνά για την αξιοποίηση του ολοένα και αυξανόμενου και ήδη τεράστιου όγκου δεδομένων, σε διάφορα έργα, ερευνητικά και μη, παρέχοντας δυνατότητες καλύτερης κατανόησης φαινομένων. Πρόκειται για σχετικά νέα διαδικασία και, παρά τα μεγάλα οφέλη, κατατάσσεται σε μια γκριζα ζώνη, επειδή έχει ορισμένες προβληματικές ηθικές και νομικές

πτυχές, ανάλογα με το πώς πραγματοποιείται, πώς χρησιμοποιούνται τα αποκτηθέντα δεδομένα, καθώς και τους όρους χρήσης της πηγής. Στο πλαίσιο αυτό, οι Krotov & Silva (2018) μελετούν τις πτυχές αυτές και προτείνουν επτά ερωτήσεις που θα πρέπει να απαντηθούν αρνητικά, ώστε να καταστήσουν το έργο νόμιμο και ηθικό. Οι ερωτήσεις είναι οι εξής:

- Απαγορεύεται ρητά το web crawling ή η αποδελτίωση (scraping) από την πολιτική «όρων χρήσης» του ιστότοπου;
- Τα δεδομένα του ιστότοπου προστατεύονται ρητά από πνευματικά δικαιώματα;
- Το έργο περιλαμβάνει παράνομη ή δόλια χρήση των δεδομένων;
- Μπορεί το web crawling και η αποδελτίωση να προκαλέσουν υλική ζημιά στον ιστότοπο ή στον διακομιστή που φιλοξενεί τον δικτυακό τόπο;
- Μπορούν τα δεδομένα που λαμβάνονται από τον ιστότοπο να θέσουν σε κίνδυνο το προσωπικό απόρρητο;
- Μπορούν τα δεδομένα που λαμβάνονται από τον ιστότοπο να αποκαλύψουν εμπιστευτικές πληροφορίες σχετικά με τις λειτουργίες του οργανισμού που παρέχει τα δεδομένα ή την εταιρεία που κατέχει τον ιστότοπο;
- Μπορεί το έργο που απαιτεί τα δεδομένα να μειώσει δυνητικά την αξία της υπηρεσίας που παρέχεται από τον δικτυακό τόπο;

Στο αντικείμενο έρευνας της παρούσας εργασίας, η αποδελτίωση, για την πρόσβαση σε δεδομένα από την αγορά ακινήτων, φαίνεται να είναι συνηθισμένη πρακτική. Ενδεικτικά, την τεχνική αυτή αναφέρουν ότι την χρησιμοποιούν στις εργασίες τους οι Li, (2019), Fu et al. (2014) και Muggenhuber (2019).

2.7 Χωρικά κριτήρια επιλογής ακινήτου

Η επιλογή ακινήτου για αγορά ή ενοικίαση είναι μια σημαντική απόφαση με πολλά κριτήρια. Από οικονομικής άποψης, η αγορά ακινήτου μπορεί να είναι η μεγαλύτερη επένδυση ενός ατόμου και συνεπώς ενέχει υψηλό οικονομικό κίνδυνο (Rabiei-Dastjerdi et al., 2021).

Ωστόσο, η διαδικασία λήψης απόφασης είναι διαφορετική από εκείνη της αγοράς ενός τυπικού προϊόντος. Για παράδειγμα, οι αγοραστές κατοικιών δεν αποσκοπούν μόνο στην απόκτηση χρηστικότητας από ένα σπίτι, αλλά επιδιώκουν και την αύξηση της αξίας μεταπώλησης και του κεφαλαίου τους (Fu et al., 2014). Επίσης, για τους οικονομικούς πόρους που διαθέτουν, προτιμούν ακίνητα με τα μέγιστα διαθέσιμα αγαθά και υπηρεσίες. Για να ελαχιστοποιήσουν τους χρηματοοικονομικούς κινδύνους, οι αγοραστές κατοικιών χρησιμοποιούν διάφορες πηγές και μηχανισμούς, όπως το κοινωνικό τους δίκτυο και τη γνώση τους για την τοπική αγορά, ώστε να αποκτήσουν πληροφορίες για να αξιολογήσουν τις πιθανές εναλλακτικές λύσεις από άποψη τιμής, ποιότητας και τοποθεσίας. Υπάρχει μεγάλος αριθμός παραγόντων που επηρεάζουν την επιλογή κατοικίας, ωστόσο, αυτοί μπορούν γενικά να κατηγοριοποιηθούν σε τρεις αλληλένδετες κατηγορίες: τις ανάγκες των νοικοκυριών, τις κτιριακές εγκαταστάσεις και τα χαρακτηριστικά της τοποθεσίας ή της γειτονιάς (Rabiei-Dastjerdi et al., 2021).

Όσον αφορά τις γεωγραφικές εξαρτήσεις της αξίας ενός ακινήτου, αυτές μπορεί να αφορούν (1) τα χαρακτηριστικά της γειτονιάς του, (2) τις αξίες των γειτονικών ακινήτων του και (3) την ευημερία της παρακείμενης επιχειρηματικής περιοχής. Η εξάρτηση της αξίας ενός ακινήτου από τα γεωγραφικά χαρακτηριστικά της γειτονιάς του (ατομική εξάρτηση) αφορά την συσχέτιση του ακινήτου με την αστική γεωγραφία (π.χ. στάσεις λεωφορείων, σταθμοί μετρό, εισοδοί οδικού δικτύου και σημεία ενδιαφέροντος), καθώς και με τα μοτίβα της ανθρώπινης κινητικότητας. Η εξάρτηση της αξίας ακινήτων από κοντινά τους ακίνητα (ομότιμη εξάρτηση) μπορεί να αποτυπωθεί με τη συγκριτική ανάλυση ακινήτων. Ένα διαισθητικό παράδειγμα της εξάρτησης αυτής, είναι ότι ένα ακίνητο συχνά έχει υψηλή αξία, εάν τα γύρω ακίνητά του έχουν και αυτά υψηλή αξία. Η εξάρτηση της αξίας ενός ακινήτου από την ευημερία της παρακείμενης επιχειρηματικής περιοχής (εξάρτηση ζώνης) βασίζεται στο ότι κάθε ακίνητο σχετίζεται με παρακείμενες επιχειρηματικές περιοχές. Μια ευημερούσα επιχειρηματική περιοχή αντιπροσωπεύει ένα σύμπλεγμα υψηλής πυκνότητας ανθρώπινων δραστηριοτήτων, εμπορικών δραστηριοτήτων και ακινήτων. Οπότε, όσο περισσότερο μια επιχειρηματική περιοχή χαρακτηρίζεται από ευημερία, τόσο πιο πιθανό είναι να υπάρχουν παρακείμενα ακίνητα υψηλής αξίας (Fu et al., 2014).

Έχουν γίνει διάφορες προσπάθειες να προσδιοριστούν και να κατηγοριοποιηθούν τα κυριότερα χωρικά κριτήρια για την επιλογή και την αξία ενός ακινήτου. Συγκεκριμένα οι Aydınoglu & Bonkir (2020) αναφέρουν 20 κριτήρια, τα οποία κατηγοριοποιούν σε 6 θεματικές ομάδες παραγόντων (Περιβαλλοντικοί παράγοντες, Δημόσιες υπηρεσίες, Πολιτιστικοί παράγοντες, Μεταφορές, Αξιοποίηση, Κοινωνικοοικονομικοί παράγοντες). Ο Li (2019, p36) αναφέρει 16 κριτήρια, τα οποία θεωρεί κρίσιμα για τις επιλογές των αγοραστών ακινήτων και τα οποία κατηγοριοποιεί σε βασικές πληροφορίες, σε περιβαλλοντικές πληροφορίες και σε 4 προφίλ δεδομένων (Μεταφορικό, Περιφερειακό, Εκπαιδευτικό, Εγκαταστάσεων). Οι Zheng et al. (2016) εντάσσουν τους πιο συχνά χρησιμοποιούμενους τύπους δεδομένων για οπτικοποίηση, στο πλαίσιο επίλυσης αστικών προκλήσεων με χρήση οπτικής αναλυτικής, σε 5 κατηγορίες (Ανθρώπινης κινητικότητας, Κοινωνικών δικτύων, Γεωγραφικά, Περιβαλλοντικά, Υγειονομικής περίθαλψης, Άλλα). Οι Rabiei-Dastjerdi et al., (2021) έχουν αναπτύξει μια λίστα από τα σημαντικότερα τοπικά χαρακτηριστικά για την επιλογή περιοχών και ακινήτων, τα οποία ομαδοποιούν σε 7 κατηγορίες (αστικές εγκαταστάσεις και υπηρεσίες, χρόνος/κόστος ταξιδιού, τιμή, κύρος και αισθητική, Συστάσεις περιοχής, περιβαλλοντικές συνθήκες, ασφάλεια και προστασία, κοινωνικοοικονομικές συνθήκες).

2.8 Σχεδιαστικά κριτήρια εφαρμογών

Οι Rabiei-Dastjerdi et al. (2021), προτείνουν να λαμβάνονται υπόψη τα ακόλουθα κρίσιμα ζητήματα για διαδικτυακούς τόπους ακινήτων:

1. Οι απαιτήσεις όλων των εμπλεκόμενων μερών πρέπει να γίνουν κατανοητές πριν από την ανάπτυξη.
2. Οι στόχοι πρέπει να καθοριστούν σαφώς πριν από την ανάπτυξη.
3. Η βιωσιμότητα του εργαλείου ακινήτων, από οικονομική και τεχνική άποψη, πρέπει να γίνει κατανοητή πριν από την ανάπτυξη.

4. Η εγκυρότητα και η διαθεσιμότητα των δεδομένων πρέπει να είναι γνωστές προκειμένου να γίνει κατανοητή η ακρίβεια και η πληρότητά τους τόσο για τους προγραμματιστές όσο και για τους τελικούς χρήστες.
5. Θα πρέπει να παρουσιάζονται τα κίνητρα των προγραμματιστών του διαδικτυακού τόπου, ώστε οι τελικοί χρήστες να μπορούν να εμπιστευτούν την εφαρμογή.
6. Πρέπει να εξεταστούν και να αντιμετωπιστούν ηθικά ζητήματα που σχετίζονται με τη διαθεσιμότητα και τη χρήση των δεδομένων, εντός του διαδικτυακού τόπου.

Επίσης, οι Rabiei-Dastjerdi et al. (2021) θεωρούν ότι, για την υποστήριξη της λήψης αποφάσεων, είναι σημαντικός ο σχεδιασμός διεπαφών και οι τεχνικές παρουσίασης και διερεύνησης γεωχωρικών δεδομένων. Με το σκεπτικό αυτό, ταξινομούν τις ιδιότητες αποτελεσματικών διαδικτυακών τόπων ακινήτων σε τέσσερις κατηγορίες:

1. Οπτικοποίηση των δεδομένων: Οι δικτυακοί τόποι πρέπει να έχουν τη δυνατότητα να επιτυγχάνουν επτά κύριες εργασίες: επισκόπηση, μεγέθυνση, φιλτράρισμα, λεπτομέρειες κατά απαίτηση, συσχέτιση και παροχή ιστορικού. Εκτός από την ταυτόχρονη αναπαράσταση ενός ή περισσότερων μεγεθών, θα πρέπει να εξετάζεται η οπτικοποίηση δεδομένων συνάφειας και η δυνατότητα εισαγωγής χωρικών δεδομένων και χαρτών.
2. Εξερεύνηση των δεδομένων: Απαιτείται η διαθεσιμότητα λειτουργιών εξερεύνησης δεδομένων και λειτουργιών χωροχρονικής ανάλυσης, χειρισμού και υπολογισμού διαφορετικών μεγεθών, φιλτραρίσματος και συνάθροισης.
3. Δομή των δεδομένων: Απαιτείται υποστήριξη για επτά τύπους δεδομένων: μονοδιάστατα, δισδιάστατα και τρισδιάστατα δεδομένα, χρονικά και πολυδιάστατα δεδομένα και δεδομένα δέντρων και δικτύων, καθώς και υποστήριξη για αυτόματη γενίκευση και πολλαπλές αναπαραστάσεις.
4. Γραφική διεπαφή χρήστη: Τα εργαλεία θα πρέπει να είναι διαδραστικά, εύχρηστα με δυνατότητα συνδυασμού μοντέλων και δεδομένων και να υποστηρίζουν την ανάκαμψη από ενέργειες.

Για την ευχρηστία διαδραστικών αναλυτικών εφαρμογών που βασίζονται σε χάρτες, οι Poppelka et al. (2019, p18), αναφέρουν τα εξής σχεδιαστικά κριτήρια:

1. Τα σημαντικά χαρακτηριστικά (τίτλος, υπόμνημα) πρέπει να τοποθετούνται εντός της γραφικής διεπαφής στην αριστερή πλευρά.
2. Εάν ένας χρήστης επιλέξει ένα χρονικό εύρος ή ένα εύρος ενός άλλου χαρακτηριστικού, το εύρος τιμών αυτής της επιλογής θα πρέπει να επισημαίνεται. Η τιμή της επιλογής θα πρέπει να εμφανίζεται στις περισσότερες δυνατές διαθέσιμες προβολές δεδομένων και η τιμή θα πρέπει να απεικονίζεται με αρκετά μεγάλη, διακριτική γραμματοσειρά και με χρώμα με αρκετή αντίθεση.
3. Για τα σύμβολα σημείων στο χάρτη θα πρέπει να υπάρχει δυνατότητα αλληλεπίδρασης.
4. Για την επιλογή τιμών σε ραβδογράμματα ή ιστογράμματα, θα πρέπει να υποστηρίζονται, τόσο το κλικ όσο και το σύρσιμο με το ποντίκι.
5. Το διάγραμμα παράλληλων συντεταγμένων θα πρέπει να σχεδιάζεται και να υλοποιείται προσεκτικά. Συνιστάται να υπάρχει λειτουργία βοήθειας ή επεξηγήσεις με τη μορφή tooltips. Το διάγραμμα παράλληλων συντεταγμένων χρησιμεύει για τον εντοπισμό ομάδων με βάση τις τιμές των χαρακτηριστικών, οπότε είναι σκόπιμο αυτές οι ομάδες να χρωματίζονται.

2.9 Κριτήρια αξιολόγησης εφαρμογών

Οι Forsell & Johansson (2010), προτείνουν 10 κριτήρια ευρετικής αξιολόγησης για τις πολύπλοκες διαδραστικές οθόνες οπτικοποίησης πληροφοριών. Τα 10 αυτά κριτήρια προέκυψαν ως εκλέπτυνση 63 ευρετικών κριτηρίων από 6 προηγουμένως δημοσιευμένα σύνολα ευρετικών, σχετικών με την διαδραστική οπτικοποίηση πληροφοριών (InfoVis). Βασίστηκαν στην προσέγγιση του Nielsen, όταν παρουσίαζε τα γνωστά 10 ευρετικά κριτήρια ευχρηστίας για την αλληλεπίδραση ανθρώπου-υπολογιστή (HCI). Το σύνολο έπρεπε να είναι αρκετά μικρό ώστε να είναι αποτελεσματικό στην εφαρμογή του και αρκετά γενικό ώστε να

καλύπτει και να εξηγεί κοινά και σημαντικά ζητήματα ευχρηστίας, που αναμένεται να προκαλέσουν προβλήματα σε μια τεχνική InfoVis. Η μελέτη τους δημιούργησε τον ακόλουθο κατάλογο των 10 επικρατέστερων ευρετικών, που παρείχαν την καλύτερη επεξηγηματική κάλυψη από όλους τους δυνατούς συνδυασμούς των 63 υπονήφιων ευρετικών:

1. Κωδικοποίηση πληροφοριών. Η αντίληψη των πληροφοριών εξαρτάται άμεσα από την αντιστοίχιση δεδομένων σε οπτικά αντικείμενα. Αυτό θα πρέπει να ενισχύεται με τη χρήση ρεαλιστικών χαρακτηριστικών ή τεχνικών, ή με τη χρήση πρόσθετων συμβόλων.
2. Ελάχιστες ενέργειες. Αφορά τον φόρτο εργασίας σε σχέση με τον αριθμό των ενεργειών που απαιτούνται για την επίτευξη ενός στόχου ή μιας δραστηριότητας.
3. Ευελιξία. Η ευελιξία αντικατοπτρίζεται στον αριθμό των πιθανών τρόπων επίτευξης ενός συγκεκριμένου στόχου. Αναφέρεται στα διαθέσιμα μέσα προσαρμογής στις στρατηγικές εργασίας, στις συνήθειες και τις απαιτήσεις του έργου.
4. Προσανατολισμός και βοήθεια. Λειτουργίες όπως η υποστήριξη για τον έλεγχο των επιπέδων λεπτομέρειας, η επανάληψη/αναίρεση ενεργειών και η αναπαράσταση πρόσθετων πληροφοριών.
5. Χωρική οργάνωση. Αφορά τον προσανατολισμό των χρηστών στο χώρο των πληροφοριών, την διαρρύθμιση των στοιχείων, την ακρίβεια και την αναγνωσιμότητα, την αποτελεσματικότητα στη χρήση του χώρου και την αλλοίωση των οπτικών στοιχείων.
6. Συνέπεια. Αναφέρεται στον τρόπο με τον οποίο οι σχεδιαστικές επιλογές διατηρούνται σε παρόμοιες περιπτώσεις και είναι διαφορετικές σε διαφορετικές περιπτώσεις.
7. Αναγνώριση και όχι ανάκληση. Ο χρήστης δεν θα πρέπει να χρειάζεται να απομνημονεύει πολλές πληροφορίες για την εκτέλεση εργασιών.
8. Προτροπή. Αναφέρεται σε όλα τα μέσα που βοηθούν στον εντοπισμό όλων των εναλλακτικών λύσεων, όταν είναι δυνατές πολλές ενέργειες, ανάλογα με το πλαίσιο εργασίας.

9. Αφαίρεση του περιττού. Αφορά το κατά πόσον οποιαδήποτε επιπλέον πληροφορία μπορεί να αποσπάσει την προσοχή και να απομακρύνει το βλέμμα από τη θέαση των δεδομένων ή την πραγματοποίηση συγκρίσεων.
10. Μείωση του συνόλου των δεδομένων. Αφορά τις παρεχόμενες δυνατότητες σμίκρυνσης συνόλων δεδομένων, καθώς και την αποτελεσματικότητα και την ευκολία χρήσης τους.

2.10 Τεχνολογική στάθμιση

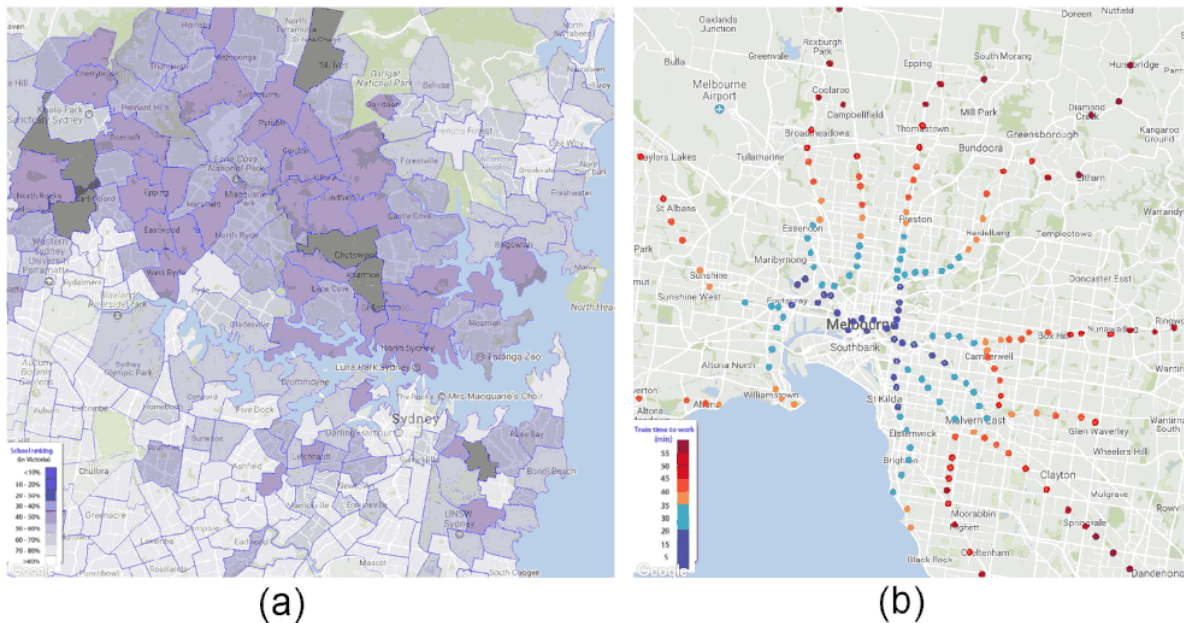
Στην ενότητα αυτή περιγράφονται και συγκρίνονται διάφορα έργα γεωοπτικής αναλυτικής από την βιβλιογραφία, καθώς και εργαλεία που έχουν φτιαχτεί για οπτική και γεωοπτική αναλυτική ή που μπορούν να τις υποστηρίξουν.

2.10.1 Εργασίες

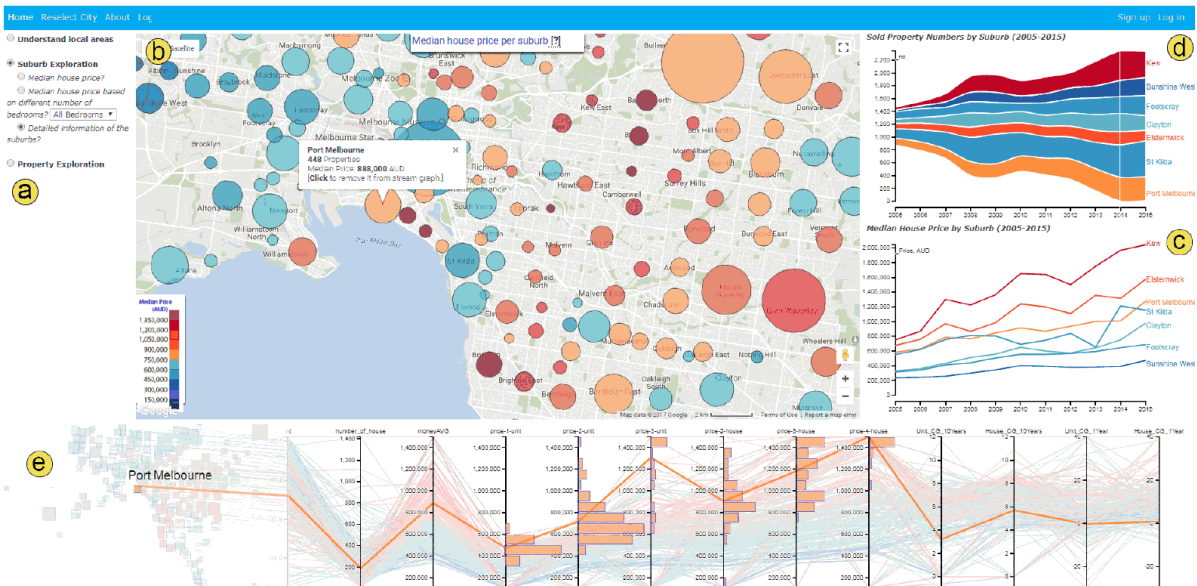
HomeSeeker

Ο Li (2019) έχει αναπτύξει την εφαρμογή οπτικής αναλυτικής HomeSeeker, που είναι ένα σύστημα οπτικής αναλυτικής πολυδιάστατων γεωχωρικών δεδομένων στον τομέα των ακινήτων, το οποίο συμπληρώνει τα υπάρχοντα εμπορικά συστήματα βοηθώντας τους χρήστες να κατανοήσουν την τοπική αγορά ακινήτων, να καταλάβουν καλύτερα τι είδους ακίνητα προτιμούν και να βρουν ακίνητα με βάση τις προσωπικές τους προτιμήσεις, τα οποία μπορούν να αντέξουν και οικονομικά. Χρησιμοποιεί τρεις διαφορετικές οπτικοποιήσεις για να βοηθήσει τους χρήστες να κατανοήσουν τα δεδομένα από διαφορετικές οπτικές. Η οπτικοποίηση με βάση το προφίλ των περιοχών (Εικόνα 2.19), βοηθά τους χρήστες να κατανοήσουν τους διαφορετικούς τρόπους ζωής της κάθε περιοχής, χρησιμοποιώντας χωροπληθείς χάρτες και χάρτες κουκίδων για να απεικονίσουν πληροφορίες για την εκπαίδευση και τα μέσα μεταφοράς αντίστοιχα. Η οπτικοποίηση με βάση τα προάστια (Εικόνα 2.20), βοηθά τους χρήστες να κατανοήσουν μέσα στον χρόνο, τις τάσεις των τιμών και των αριθμό πωλήσεων ακινήτων ανά προάστιο, χρησιμοποιώντας χάρτη κουκίδων, διαμόρφωση των συμβόλων πάνω στον χάρτη, διάγραμμα πολλαπλών γραμμών, γράφημα

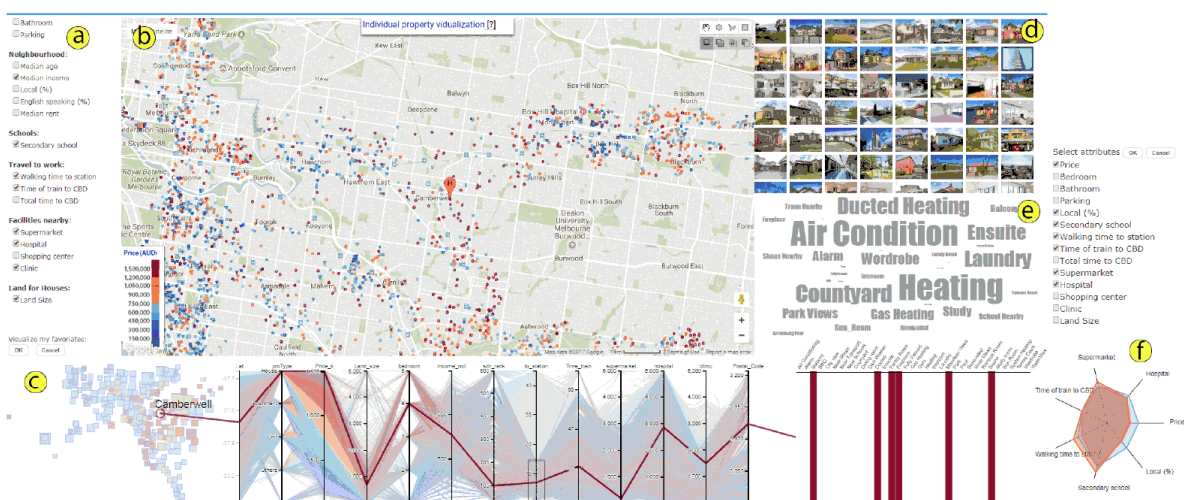
ροής και διάγραμμα παράλληλων συντεταγμένων με ιστογράμματα. Η οθόνη με βάση τα ακίνητα (Εικόνα 2.21) επιτρέπει στους χρήστες να εξερευνήσουν την αγορά ακινήτων, μέσα από μια σειρά συνδεδεμένων οπτικοποιήσεων με τις οποίες μπορούν να αλληλεπιδράσουν.



Εικόνα 2.19: Χάρτες απεικόνισης προφίλ περιοχών (a) χωροπληθείς και (b) κουκίδων (Li, 2019)



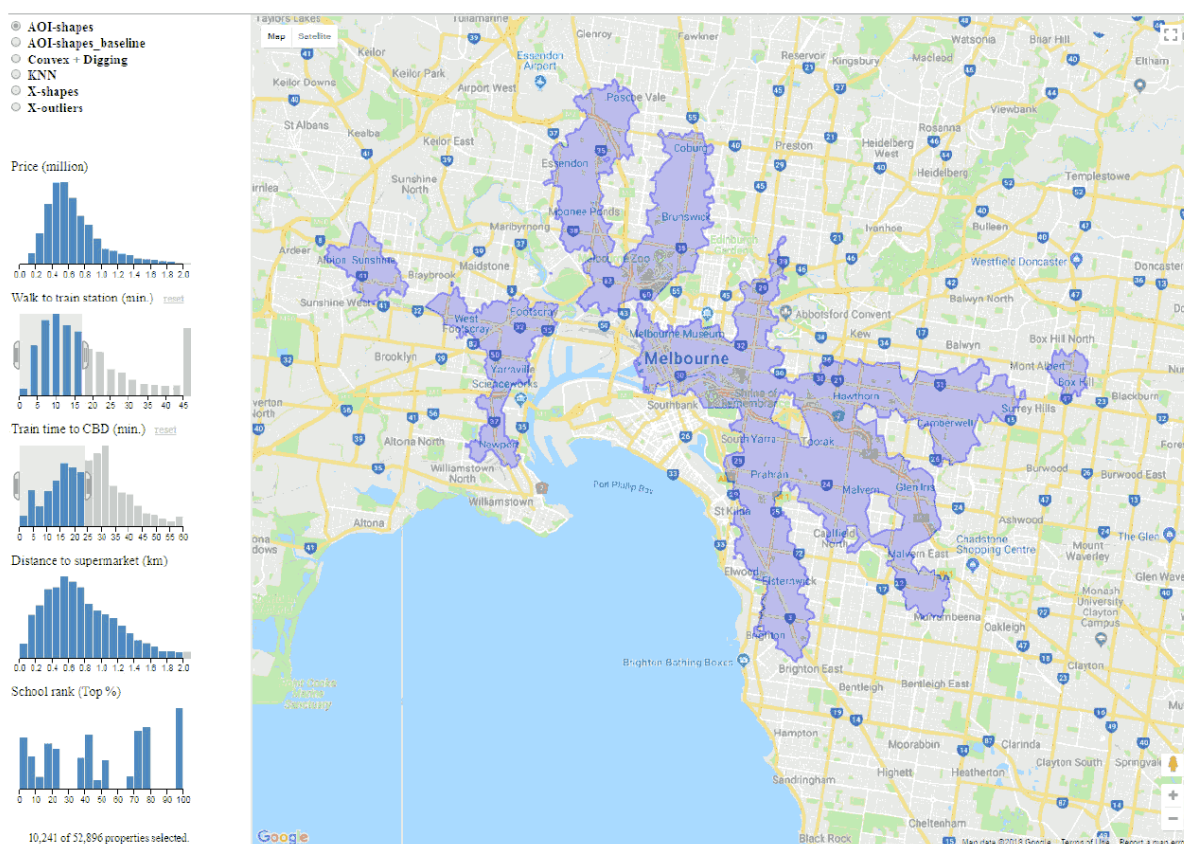
Εικόνα 2.20: Οθόνη προαστίων (a) μάρα πλοήγησης (b) χάρτης κουκίδων (c) προβολή ιστορικού με πολλαπλές γραμμές (d) προβολή ιστορικού με γράφημα ροής (e) διάγραμμα παράλληλων συντεταγμένων με ιστογράμματα (Li, 2019)



Εικόνα 2.21 Οθόνη ακινήτων (a) μπάρα πλοήγησης (b) χάρτης συμβόλων με πολύγωνα (c) πολυδιάστατη προβολή δεδομένων (d) κάρτα με εικόνες (e) σύννεφο λέξεων (f) διάγραμμα αράχνης (Li, 2019)

AOI-shapes

Ο Li (2019) έχει επίσης αναπτύξει τον AOI-shapes, που είναι ένας αποδοτικός και αποτελεσματικός αλγόριθμος οριοθέτησης περιοχών ενδιαφέροντος (Areas Of Interest, AOI) από υποσύνολα σημείων που έχει επιλέξει ο χρήστης, για την οπτικοποίησή τους σε εφαρμογές γεωοπτικής αναλυτικής. Αφού μελέτησαν την τεχνολογική στάθμιση σε μεθόδους αποτυπώματος (footprint), ανέπτυξαν μια καινοτόμο μέθοδο για να προσδιορίσουν το περίγραμμα αστικών περιοχών ενδιαφέροντος που έχει επιλέξει ο χρήστης. Σε αντίθεση με τα σαφή όρια διοικητικών περιοχών, για τις περιοχές ενδιαφέροντος δεν υπάρχουν ξεκάθαρα όρια, ενώ το σχήμα τους μπορεί να είναι αόριστο. Ο πιο συχνός τρόπος χαρακτηρισμού του σχήματος μιας περιοχής ενδιαφέροντος είναι η κατασκευή συνόρων με βάση τα σημεία ενδιαφέροντος (Points Of Interest, POI) που ικανοποιούν τα κριτήρια του χρήστη. Στην Εικόνα 2.22 φαίνεται ο σχηματισμός της περιοχής ενδιαφέροντος από σημεία ακινήτων με τον αλγόριθμο AOI-shapes, καθώς ο χρήστης αλλάζει τις τιμές των κριτηρίων αναζήτησης ακινήτων.

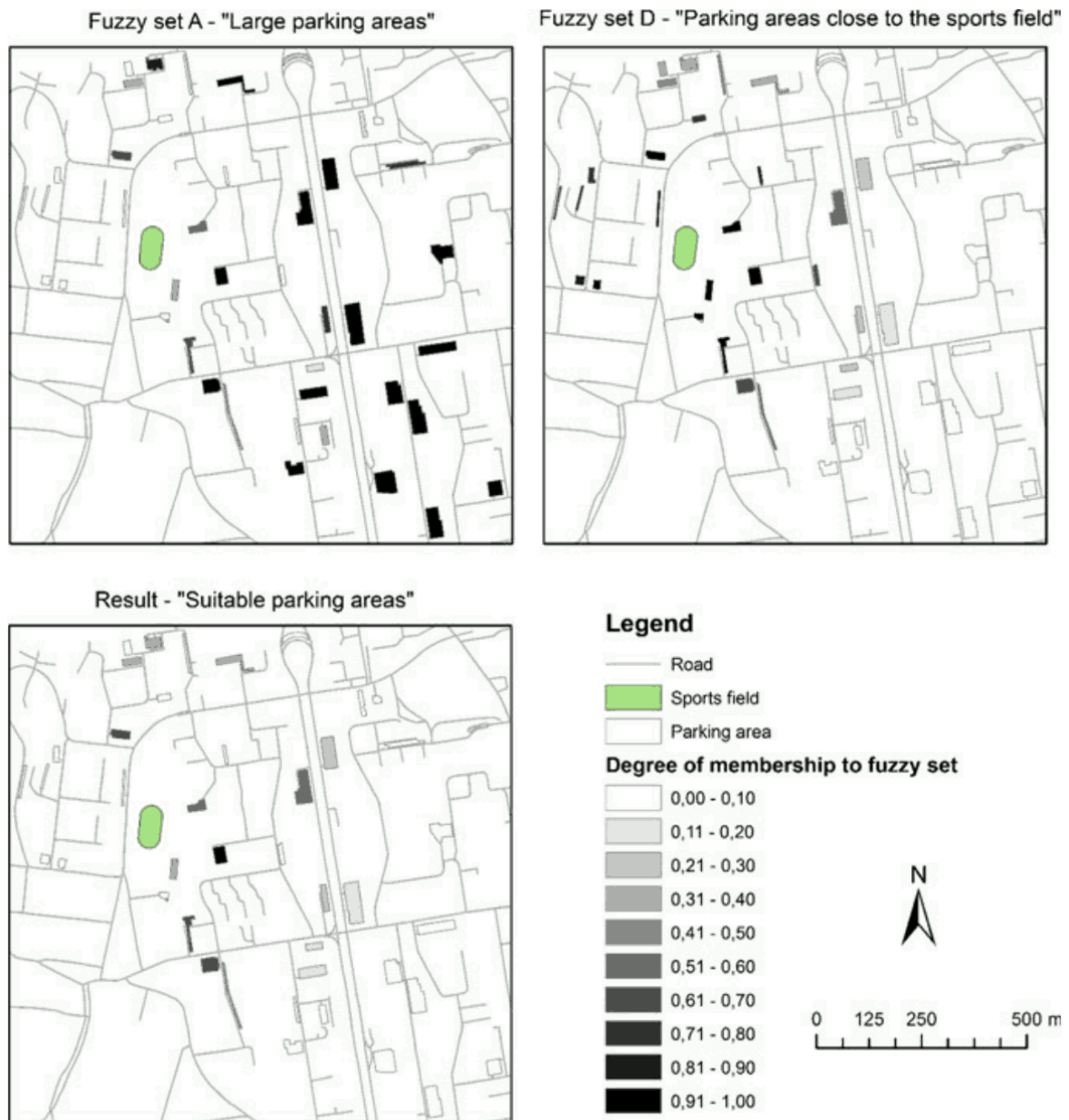


Εικόνα 2.22: Σχηματισμός συνόρων περιοχής ενδιαφέροντος με τον αλγόριθμο AOI-shapes (Li, 2019)

Fuzzy spatio-temporal querying

Στην εργασία τους οι Đuračić & Faixová Chalachanová (2017) έχουν εφαρμόσει την θεωρία ασαφών συνόλων (fuzzy set theory) για την πολυκριτηριακή αναζήτηση σε γεωχωροχρονικά δεδομένα. Συγκεκριμένα, για την αναζήτηση σε SQL βάσεις δεδομένων έχουν χρησιμοποιήσει την PostgreSQL, στην οποία υλοποιούν τις ασαφείς συναρτήσεις συμμετοχής με τμηματικές γραμμικές συναρτήσεις, και τον ασαφή τελεστή AND με την συνάρτηση LEAST, η οποία επιστρέφει την τιμή του ελάχιστου από τα ορίσματά του. Στην συνέχεια, ως επίδειξη της λειτουργίας του τρόπου αναζήτησης που ανέπτυξαν, δίνουν το εξής παράδειγμα: “Αναζήτησε όλους τους μεγάλους χώρους στάθμευσης που είναι κοντά σε έναν συγκεκριμένο αθλητικό χώρο και έχουν ολοκληρωθούν περίπου δύο χρόνια πριν”. Για την πραγματοποίηση της αναζήτησης κατασκεύασαν ένα SQL ερώτημα, όπου χρησιμοποίησαν τις ασαφείς συναρτήσεις που δημιούργησαν στην PostgreSQL. Για τις ασαφείς συναρτήσεις, έπρεπε να ερμηνεύσουν τα ασαφή λεκτικά κριτήρια (μεγάλοι χώροι στάθμευσης, περίπου

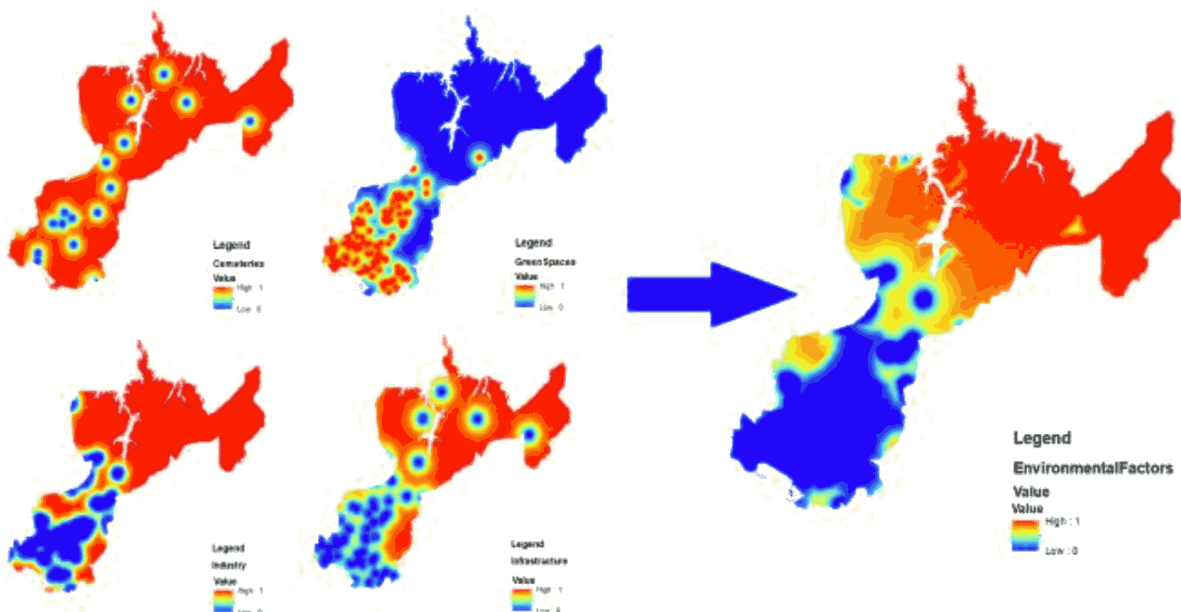
δύο χρόνια πριν, κοντά στον αθλητικό χώρο) και να τα αποδώσουν σε αριθμητικές τιμές, τις οποίες χρησιμοποίησαν ως ορίσματα στις συναρτήσεις. Στην Εικόνα 2.23 φαίνεται το αποτέλεσμα της αναζήτησης, όπως απεικονίζεται σε ένα σύστημα GIS.



Εικόνα 2.23: Ασαφής γεωχωροχρονική αναζήτηση (Đuračićová & Faixová Chalachanová, 2017)

Smart Real Estate

Οι Aydinoglu & Bovkir (2020) ανέπτυξαν το Smart Real Estate ως πρότυπο εφαρμογής έξυπνου τηλεφώνου, το οποίο μπορεί να χρησιμοποιηθεί ελεύθερα ως βάση για την ανάπτυξη παρόμοιων εφαρμογών. Επίσης ανέπτυξαν ένα γεωγραφικό υπόβαθρο για την εφαρμογή, το οποίο παρέχει πληροφορίες για την αγορά ακινήτων, βάσει ενός θεματικού γεωγραφικού υπόβαθρου, το οποίο κατασκευάσαν από ανοικτά δημόσια γεωχωρικά δεδομένα. Για την παραγωγή του υπόβαθρου χρησιμοποίησαν 20 κριτήρια από 6 θεματικές ομάδες παραγόντων (περιβαλλοντικοί, δημόσιες υπηρεσίες, πολιτιστικοί, μεταφορές, αξιοποίηση, κοινωνικοοικονομικοί), που επηρεάζουν την αξία ενός ακινήτου. Συγκεκριμένα, χρησιμοποιούν ασαφή λογική (fuzzy logic) για να παράξουν ένα εικονιστικό (raster) υπόβαθρο, όπου η τιμή κάθε εικονοστοιχείο είναι το αποτέλεσμα εφαρμογής ασαφών τελεστών στο σύνολο των αντίστοιχων εικονοστοιχείων των επιπέδων των κριτηρίων. Στην Εικόνα 2.24 φαίνεται η δημιουργία του υπόβαθρου των περιβαλλοντικών παραγόντων ως το αποτέλεσμα εφαρμογής των τελεστών στα επίπεδα των επιμέρους κριτηρίων της περιβαλλοντικής θεματικής ομάδας.



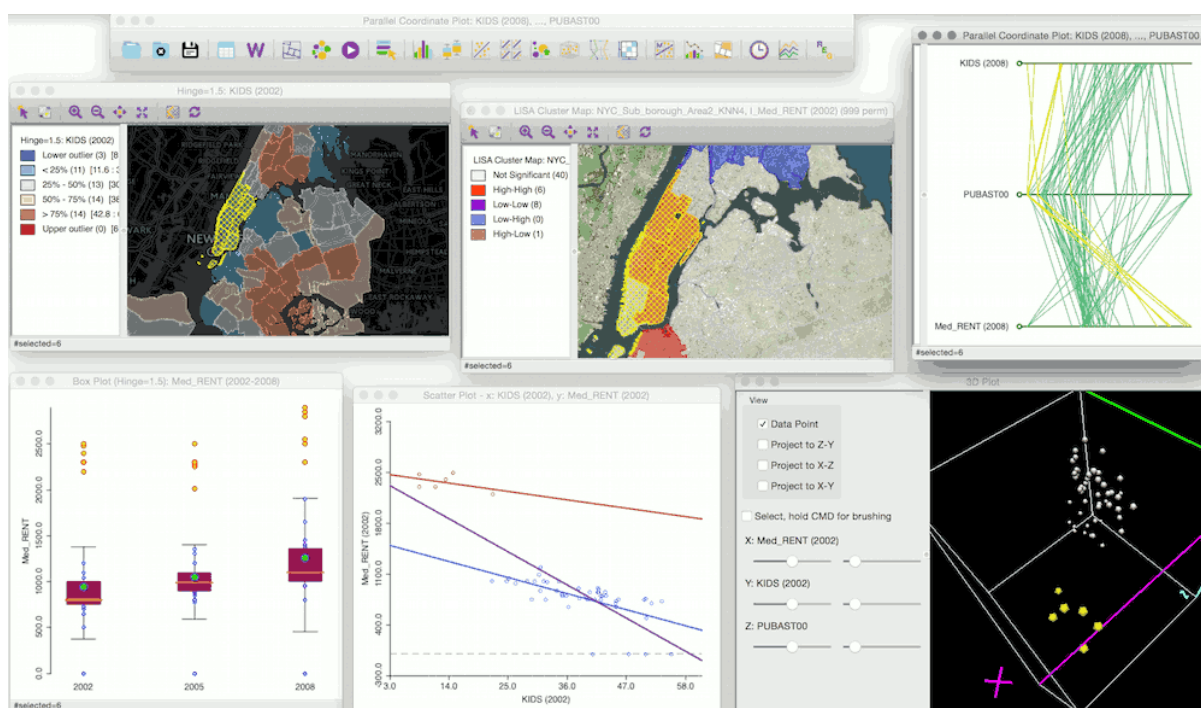
Εικόνα 2.24: Δημιουργία υπόβαθρου περιβαλλοντικών παραγόντων (Aydinoglu & Bovkir, 2020)

2.10.2 Εργαλεία

GeoDa

Το GeoDa σχεδιάστηκε για να διευκολύνει την εξερεύνηση και την ανάλυση γεωχωρικών δεδομένων, από την απλή περιγραφή και οπτικοποίηση στη δομημένη εξερεύνηση και τη μοντελοποίηση. Είναι πρόγραμμα ανοικτού κώδικα, γραμμένο στην γλώσσα C++ και εκτελείται σε πολλές πλατφόρμες, όπως Linux, Windows και MacOS. Οι κύριες κατηγορίες λειτουργιών του είναι: εισαγωγή και εξαγωγή δεδομένων, δημιουργία χωρικών δεδομένων, αναζήτηση δεδομένων, μετασχηματισμοί μεταβλητών, χωροπληθική χαρτογράφηση, στατιστικοί χάρτες, εξομαλυμένοι χάρτες ποσοστών, εξερευνητική ανάλυση δεδομένων, χωρική αυτοσυσχέτιση, χωροχρονική εξερεύνηση, εξερεύνηση ετερογένειας και χωρική παλινδρόμηση. Επίσης χρησιμοποιεί δυναμική σύνδεση και βούρτσισμα (brushing) για όλα τα ανοικτά παράθυρα (χάρτες, πίνακες και στατιστικά γραφήματα) (Anselin, 2017).

Η τελευταία έκδοση (1.20) περιέχει υποστήριξη πολλαπλών επιπέδων, πολλά νέα χαρακτηριστικά τοπικής ομαδοποίησης. Υλοποιεί επίσης διάφορες κλασικές μη χωρικές τεχνικές ομαδοποίησης (ανάλυση κύριων συνιστωσών, κ-μέσων και ιεραρχική), καθώς και το HDBScan (<https://geodacenter.github.io>). Η Εικόνα 2.25 δείχνει μια τυπική οθόνη εξερεύνησης των αποτελεσμάτων στατιστικών δοκιμών και μοντέλων μέσω συνδεδεμένων χαρτών και διαγραμμάτων.



Εικόνα 2.25: Οθόνη Εξερευνητικής Ανάλυσης Γεωχωρικών Δεδομένων στην εφαρμογή GeoDa (<https://geodacenter.github.io>)

PostgreSQL/PostGIS

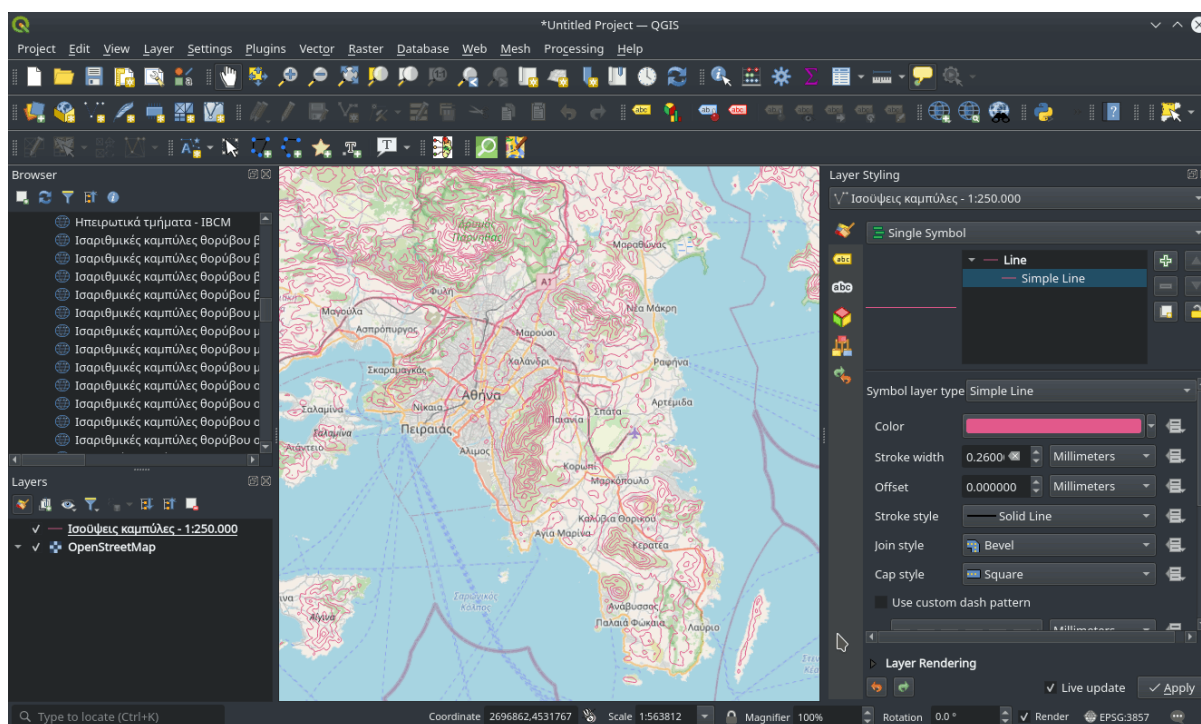
Η PostgreSQL είναι ένα σύστημα διαχείρισης αντικειμενο-σχεσιακών βάσεων δεδομένων (object-relational database management system, ORDBMS), υποστηρίζει ένα μεγάλο μέρος του προτύπου SQL και προσφέρει πολλά σύγχρονα χαρακτηριστικά όπως: πολύπλοκα ερωτήματα, ξένα κλειδιά (foreign keys), εναύσματα (triggers), προβολές με δυνατότητα ενημέρωσης (updatable views), ακεραιότητα συναλλαγών (transactional integrity), έλεγχο ταυτόχρονης λειτουργίας πολλαπλών εκδοχών (multiversion concurrency control). Επίσης, η PostgreSQL μπορεί να επεκταθεί από τον χρήστη με πολλούς τρόπους, για παράδειγμα με την προσθήκη νέων τύπων δεδομένων, συναρτήσεων, τελεστών, συναρτήσεων συνάθροισης, μεθόδων ευρετηρίου και διαδικαστικών γλωσσών (Perl, Java, C++, JavaScript, .NET, Tcl, Go, ODBC, Python). Λόγω της ελεύθερης άδειας χρήσης, η PostgreSQL μπορεί να χρησιμοποιηθεί, να τροποποιηθεί και να διανεμηθεί από οποιονδήποτε δωρεάν για οποιονδήποτε σκοπό, είτε πρόκειται για ιδιωτικό είτε για εμπορικό είτε για ακαδημαϊκό σκοπό (PostgreSQL, 2022).

Το PostGIS είναι μια χωρική επέκταση της PostgreSQL, που επιτρέπει την αποθήκευση GIS αντικειμένων στην βάση δεδομένων. Το PostGIS περιλαμβάνει υποστήριξη για χωρικά ευρετήρια R-Tree με βάση το GiST, καθώς και συναρτήσεις για την ανάλυση και την επεξεργασία αντικειμένων GIS (PostGIS, 2022).

Το PostGIS βασίζεται σε άλλα υποστηρικτικά έργα όπως: το Proj (για την υποστήριξη προβολής), το Geometry Engine Open Source (GEOS) για προηγμένη υποστήριξη επεξεργασίας γεωμετρίας, το Geospatial Data Abstraction Library (GDAL) για πολλά προηγμένα χαρακτηριστικά εικονιστικής (raster) επεξεργασίας, την βιβλιοθήκη αλγορίθμων υπολογιστικής γεωμετρίας (CGAL/SFCGAL) για προηγμένη τρισδιάστατη ανάλυση. Τα περισσότερα από αυτά τα έργα, όπως και το PostGIS, πλέον υπάγονται κάτω από αιγίδα του Open Source Geospatial Foundation (OSGeo) (Hsu & Obe, 2021).

QGIS

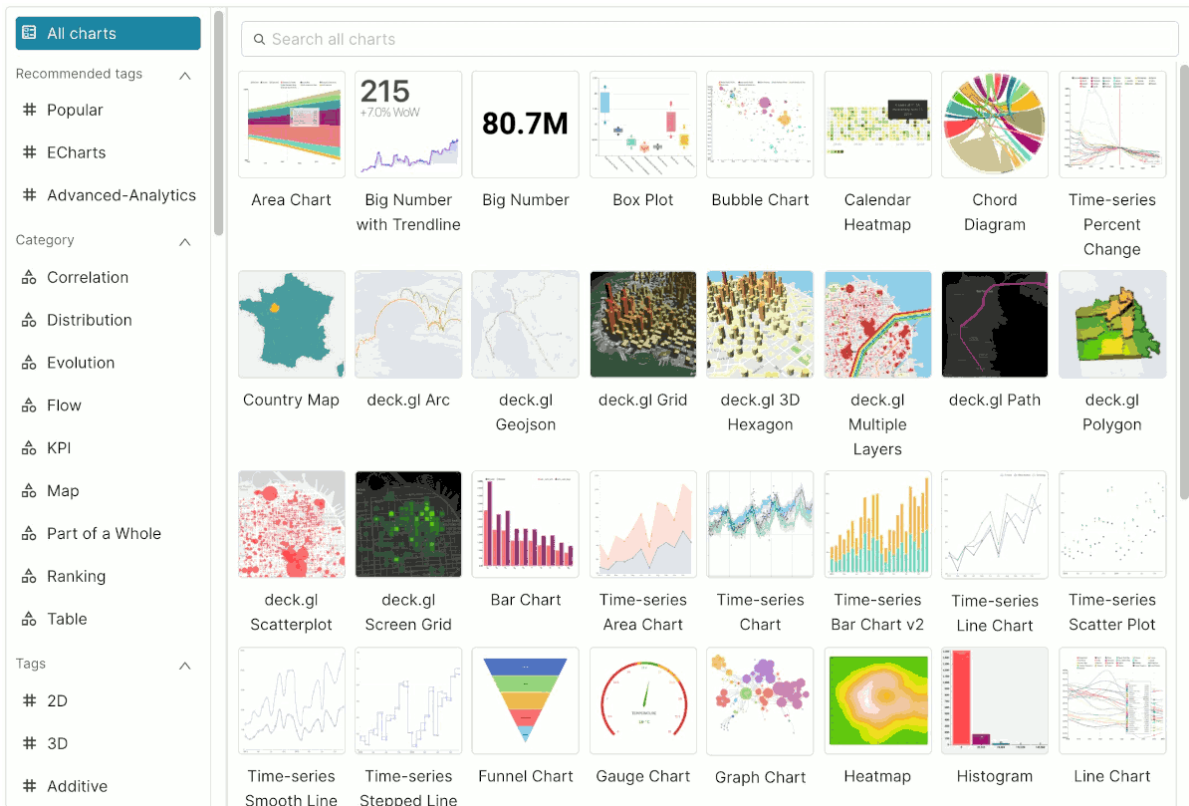
Το QGIS (Εικόνα 2.26) είναι ένα Σύστημα Γεωγραφικών Πληροφοριών Ανοικτού Κώδικα, το οποίο ξεκίνησε τον Μάιο του 2002, Το QGIS τρέχει επί του παρόντος στις περισσότερες πλατφόρμες Unix, στα Windows και στο MacOS. Έχει αναπτυχθεί χρησιμοποιώντας την εργαλειοθήκη Qt (<https://www.qt.io>) και τη C++, που του δίνουν ταχύτητα και ένα ευχάριστο, εύχρηστο γραφικό περιβάλλον χρήστη (GUI). Αν και είχε ξεκινήσει ως ένα πρόγραμμα προβολής δεδομένων GIS, έχει δυνατότητες καταχώρησης, προηγμένης ανάλυσης και δημιουργίας εξελιγμένων χαρτών, ατλάντων και εκθέσεων. Υποστηρίζει πληθώρα εικονιστικών (raster) και διανυσματικών δεδομένων, και έχει τη δυνατότητα να υποστηρίξει νέα, μέσω της αρχιτεκτονικής πρόσθετων (plugins). Επειδή κυκλοφορεί υπό την ανοικτή άδεια GNU GPL, ο πηγαίος κώδικας είναι διαθέσιμος για επιθεώρηση, τροποποίηση, και το πρόγραμμα είναι διαθέσιμο δωρεάν. (QGIS, 2022)



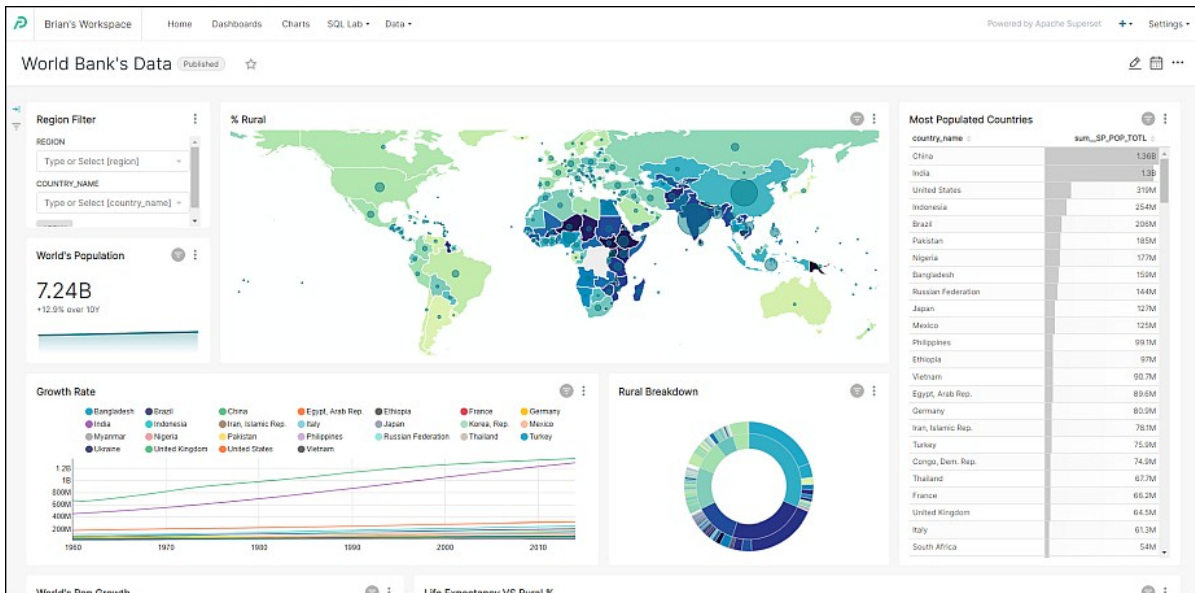
Εικόνα 2.26: Το QGIS 3.26 Desktop

Apache Superset

Το Apache Superset είναι μια πλατφόρμα απεικόνισης και εξερεύνησης δεδομένων ανοικτού κώδικα, που χρησιμοποιείται συνήθως για την επιχειρηματική ευφυΐα. Παρέχει μια διαισθητική διεπαφή για τη δημιουργία διαδραστικών οπτικών ταμπλό σε συνδυασμό με ένα διαδραστικό περιβάλλον ανάπτυξης SQL. Το Apache Superset προσφέρει μια σειρά από διαγράμματα οπτικοποίησης, βασισμένα στο D3.js ή στο deck.gl, τα οποία μπορούν να συνδεθούν δυναμικά για τη δημιουργία διαδραστικών οπτικών ταμπλό (Navarra et al., 2020). Διαθέτει ένα μεγάλο αριθμό από διαδραστικά διαγράμματα και χάρτες (Εικόνα 2.27) (<https://github.com/apache/superset>), και επιτρέπει την δημιουργία σύνθετων ταμπλό με τα οποία ο χρήστης μπορεί να αλληλεπιδράσει εξερευνητικά (Εικόνα 2.28) (<https://docs.preset.io/docs/dashboard-properties>).



Εικόνα 2.27: Superset, συλλογή διαθέσιμων διαγραμμάτων (<https://github.com/apache/superset>)



Εικόνα 2.28: Superset, σύνθετο και διαδραστικό ταμπλό με χάρτη και διαγράμματα (<https://docs.preset.io/docs/dashboard-properties>)

WebGLayer

Το WebGLayer είναι μια βιβλιοθήκη ανοικτού κώδικα βασισμένη στην JavaScript και στο WebGL, για πολλαπλές συνδεδεμένες οπτικοποιήσεις. Η βιβλιοθήκη βασίζεται στην WebGL και αξιοποιεί την κάρτα γραφικών της συσκευής για τη γρήγορη απεικόνιση και το φιλτράρισμα των δεδομένων. Μπορεί να απεικονίσει δεδομένα σε χάρτη που παρέχεται από άλλες βιβλιοθήκες (π.χ. OpenLayers, Leaflet, GoogleMap API). Η βιβλιοθήκη επικεντρώνεται σε χωρικά δεδομένα και μεγάλα σύνολα δεδομένων, μέχρι δεκάδες εκατομμύρια εγγραφές. Αναπτύχθηκε και υποστηρίχθηκε από το έργο OpenTransportNet του προγράμματος-πλαίσιου ανταγωνιστικότητας και καινοτομίας της ΕΕ, CIP (Competitiveness and Innovation Framework Programme) για τις κυκλοφοριακές ροές και τα τροχαία ατυχήματα, ενώ αργότερα επεκτάθηκε και βελτιώθηκε στο πλαίσιο του ευρωπαϊκού έργου PoliVisu του προγράμματος Horizon 2020 για τη διευκόλυνση ενός ευρύτερου φάσματος οπτικοποιήσεων που ανταποκρίνονται στις ανάγκες των πόλεων όταν εργάζονται με την έξυπνη πολιτική κινητικότητας στην εποχή των μεγάλων δεδομένων.

Το WebGLayer επιτρέπει την ανάπτυξη διαδραστικών απεικονίσεων θερμικών χαρτών μεγάλων συνόλων δεδομένων, με την υλοποίηση πολλαπλών συνδεδεμένων προβολών για την παρουσίαση των δεδομένων (Εικόνα 2.29). Κάθε μία από τις προβολές επιτρέπει διαφορετικές αλληλεπιδράσεις (όπως βούρτσισμα, φιλτράρισμα ή ανάλυση σχέσεων) που προκαλούν άμεση ενημέρωση των άλλων προβολών. Έτσι, οι χρήστες επωφελούνται από την άμεση και δυναμική απεικόνιση των δεδομένων, αποκτούν καλύτερη κατανόηση των δεδομένων εφαρμόζοντας φίλτρα, και αναπτύσσουν την ευκαιρία να ανακαλύψουν σχέσεις και μοτίβα στα δεδομένα. (Popelka et al., 2019)



Εικόνα 2.29: Οθόνη εφαρμογής που αναπτύχθηκε με το WebGLayer (<https://github.com/WebGLayer/webglayer/wiki/User-Guide>)

HSLayers NG

Το HSLayers NG είναι μια βιβλιοθήκη γραμμένη σε JavaScript, που επεκτείνει τη λειτουργικότητα του OpenLayers 6 παρέχοντας ένα θεμέλιο για τη δημιουργία διαδραστικού χάρτη (Εικόνα 2.30) με πρόσθετα στοιχεία, όπως διαχειριστή επιπέδων, δημιουργία μόνιμων συνδέσμων, μορφοποίηση διανυσματικών χαρακτηριστικών, συμπεριλαμβανομένων των επιπέδων υπηρεσιών WMS, με φιλικό προς το χρήστη τρόπο (<https://ng.hslayers.org>). Χρησιμοποιεί σύγχρονες βιβλιοθήκες JavaScript, αντί του ExtJS 3. Το HSLayers είναι κατασκευασμένο με αρθρωτό τρόπο, ο οποίος επιτρέπει την ελεύθερη προσάρτηση και αφαίρεση των λειτουργικών μονάδων, εφόσον ικανοποιούνται οι εξαρτήσεις για καθεμία από αυτές. Ο πυρήνας του πλαισίου αναπτύσσεται με τη χρήση των AngularJS, RequireJS και Bootstrap. Η βιβλιοθήκη διανέμεται με ανοικτή άδεια χρήσης MIT. (Popelka et al., 2019)

enabling The Best Practices Atlas

UNSELECT ALL

AUT BEL BGR CAN CZE DEU FRA GBR
 IRL ITA NLD NOR

BEST PRACTICE
Bio product from grape processing – the Polyphenols practice
Emilia-Romagna, Italy
 Example of biorefinery able to use its agricultural residues to produce a varied range of bio-based products.

BEST PRACTICE
Fish-tomato farm partnership
Deinze, Belgium
 The tomato farm supplies electricity, rain water and residual heat to fish nursery while receiving back filtered water as a reciprocal ser-vice.

BEST PRACTICE
Soil Nutrients from Wool
Piedmont, Italy
 An "Interreg-europe" project has developed activities for the re-use of remaining coarse wool from sheep farming, which resulted unserviceable for further textile application.

BEST PRACTICE
Bio-fertilisers from agri production
Plovdiv, Bulgaria
 A biorefinery to convert residues into bio-based product.

BEST PRACTICE
Mitigation of phosphate surplus from pig manure
Son, the Netherlands
 The digestion process produces renewable gas for households.

BEST PRACTICE
Thermoplastic from grass fibre
Odenwald, Germany
 A German company has developed an innovative fibre-reinforced thermoplastic for injection moulding and extrusion.

Εικόνα 2.30: Εφαρμογή που αναπτύχθηκε με το HSLayers NG (<https://ng.hslayers.org>)

3 Εφαρμογή γεωοπτικής αναλυτικής για την αγορά ακινήτων

Υπάρχουν αρκετές εφαρμογές που σχετίζονται με ακίνητα, όπως ιστοσελίδες αγγελιών για αγορά και ενοικίαση, ή για ανεύρεση τουριστικών καταλυμάτων. Ωστόσο, η αναζήτηση τις περισσότερες φορές δεν περιλαμβάνει κριτήρια για τα χαρακτηριστικά της περιοχής που βρίσκεται το ακίνητο, κάτι που πολλές φορές μπορεί να είναι και πιο σημαντικό από τα χαρακτηριστικά του ίδιου του ακινήτου. Για παράδειγμα, στην περίπτωση της αναζήτησης τουριστικών καταλυμάτων, πολλές φορές οι ανάγκες του χρήστη έχουν να κάνουν περισσότερο με τον τουριστικό χαρακτήρα της περιοχής ή με την απόσταση του καταλύματος από αξιοθέατα ή δραστηριότητες. Έτσι, ένας τουρίστας μπορεί πχ να αναζητά περιοχές που να είναι κοντά σε θάλασσα, αρχαιολογικά αξιοθέατα και ποδηλατικές διαδρομές. Επίσης οι Rabiei-Dastjerdi et al. (2021) έχουν διαπιστώσει ότι, ακόμα και σε περιπτώσεις που υπάρχουν διαθέσιμα δεδομένα για τα χαρακτηριστικά περιοχών, που μπορούν να υποστηρίξουν τη λήψη αποφάσεων, αυτά δεν παρουσιάζονται από ιστοσελίδες αγγελιών ακινήτων. Υποθέτουν ότι αυτό γίνεται είτε λόγω τεχνικών προκλήσεων της μετατροπής χωρικών δεδομένων σε χρήσιμες πληροφορίες, είτε λόγω απροθυμίας και αντίστασης των ιστοσελίδων να δημοσιεύουν δεδομένα που σχετίζονται με την ανάλυση της τοποθεσίας ή της γειτονιάς ακινήτων, και τα οποία μπορεί να εκληφθούν ως αρνητικά.

3.1 Σκοπός και ερευνητικό ερώτημα

Στο πλαίσιο της εργασίας υλοποιήθηκε εφαρμογή γεωοπτικής αναλυτικής για να βοηθήσει τον χρήστη στην κατανόηση των αναγκών του και για να υποστηρίξει την λήψη βέλτιστων και ικανοποιητικών αποφάσεων σχετικά με την αγορά ακινήτων. Η εφαρμογή πραγματοποιήθηκε στο πλαίσιο ενός σημαντικού προβλήματος που εντοπίστηκε στην βιβλιογραφία. Το πρόβλημα αφορά τον μεγάλο χρόνο έρευνας και αναζήτησης ακινήτων, και των άστοχων, τελικά, αποφάσεων αγοράς ή ενοικίασης, για τις οποίες οι αγοραστές και ενοικιαστές μετανιώνουν, σε ποσοστό περίπου 45%. Ο κυριότερος λόγος για αυτό είναι η έλλειψη παροχής πληροφοριών και αναλύσεων από τα υφιστάμενα εμπορικά συστήματα της αγοράς ακινήτων.

Σε αυτά τα συστήματα, συνήθως οι χρήστες μπορούν να πραγματοποιήσουν αναζητήσεις, καθορίζοντας την περιοχή και τα χαρακτηριστικά του ακινήτου που τους ενδιαφέρουν, με τα αποτελέσματα να εμφανίζονται σε μορφή λίστας ή πάνω σε χάρτη. Ωστόσο, αυτό δεν είναι αρκετό, γιατί συχνά δεν είναι εξοικειωμένοι με τις ιδιαιτερότητες της κάθε περιοχής και των τοπικών αγορών, και οι πλατφόρμες αυτές δεν παρέχουν εργαλεία που να τους βοηθούν να αναλύουν και να συγκρίνουν τα δεδομένα των περιοχών και των ακινήτων (Li, 2019). Για την επίλυση του προβλήματος αυτού έχουν προταθεί, μεταξύ άλλων, λύσεις οπτικής αναλυτικής, που αποσκοπούν στο να παρέχουν εξερευνητικά εργαλεία κατανόησης των ιδιαίτερων χαρακτηριστικών των περιοχών και της αγοράς ακινήτων από τους χρήστες.

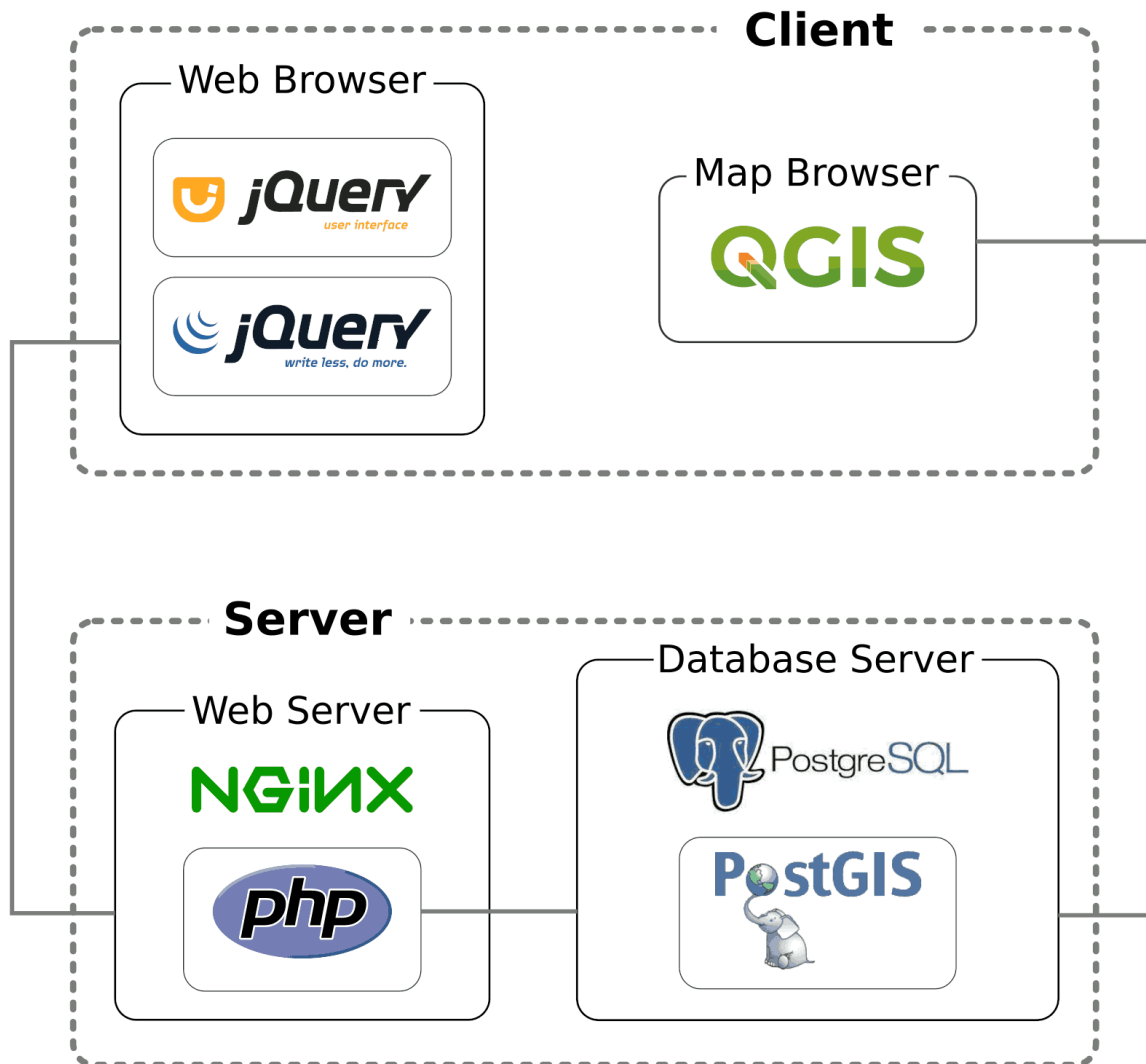
Η παρούσα εφαρμογή βασίζεται σε καλά μελετημένες τεχνικές και δοκιμασμένα εργαλεία, και συνεισφέρει έναν νέο τρόπο εξερεύνησης γεωχωρικών δεδομένων, που βασίζεται σε έναν αλγόριθμο, ο οποίος υπολογίζει έναν δείκτη καταλληλότητας περιοχών, με βάση τα κριτήρια των απαιτήσεων του χρήστη. Με την εφαρμογή αυτή, η εργασία προσπαθεί να απαντήσει το ερευνητικό ερώτημα: Πώς μπορεί η Γεωοπτική Αναλυτική, με τα διαδραστικά οπτικά εργαλεία εξερευνητικής ανάλυσης και παρουσίασης γεωχωρικών δεδομένων, να βοηθήσει στην κατανόηση τοπικών χαρακτηριστικών και στη αποσαφήνιση ανθρώπινων στεγαστικών αναγκών, και στην συνέχεια να υποστηρίξει την λήψη επιτυχημένων αποφάσεων γύρω από τα ακίνητα;

3.2 Αρχιτεκτονική εφαρμογής

Η εφαρμογή ακολουθεί την αρχιτεκτονική πελάτη-διακομιστή (Εικόνα 3.1). Ο χρήστης αλληλεπιδρά με την εφαρμογή μέσω διεπαφής, η οποία εκτελείται εν μέρει στον περιηγητή και εν μέρει στο QGIS. Συγκεκριμένα, μόνο το στοιχείο του χάρτη εκτελείται στο QGIS, ενώ τα υπόλοιπα στοιχεία, όπως διαγράμματα και κουμπιά, εκτελούνται στον περιηγητή. Οι χειρισμοί του χρήστη στα στοιχεία του περιηγητή μετατρέπονται σε εντολές επικαιροποίησης των γεωγραφικών υπόβαθρων, και μεταδίδονται στον Nginx web server, ο οποίος, μέσω του πρωτοκόλλου FastCGI, μεταφέρει τα δεδομένα αυτά σε ένα πρόγραμμα της γλώσσας PHP. Το PHP πρόγραμμα επικοινωνεί με την PostgreSQL, στην οποία πυροδοτεί την εκτέλεση αλγόριθμων ενημέρωσης των γεωχωρικών δεδομένων, χρησιμοποιώντας τις εντολές

επικαιροποίησης. Το QGIS έχει μόνιμη σύνδεση με την PostgreSQL και ενημερώνεται αυτόματα για τα νέα δεδομένα, επικαιροποιώντας τα υπόβαθρα τα οποία, στην συνέχεια, απεικονίζονται στην οθόνη του χάρτη.

Η ιστοσελίδα με τα στοιχεία αλληλεπίδρασης έχει υλοποιηθεί ως εφαρμογή μία σελίδας (Single Page Application, SPA), όπου οι χειρισμοί του χρήστη δεν καλούν νέες σελίδες, αλλά επικοινωνούν με τον Nginx με την τεχνική Ajax (Asynchronous JavaScript and XML). Η βιβλιοθήκη jQuery χρησιμοποιείται για την ασύγχρονη επικοινωνία και η jQueryUI για προηγμένα διαδραστικά στοιχεία. Επειδή δεν είναι δυνατή η απευθείας επικοινωνία από την ιστοσελίδα με την βάση δεδομένων, μεσολαβεί η PHP, η οποία διαθέτει οδηγό αλληλεπίδρασης με την PostgreSQL, για να εκτελέσει τις εντολές που έχουν προετοιμαστεί από την ιστοσελίδα. Η PostgreSQL έχει εγκατεστημένο το άρθρωμα PostGIS, ώστε να μπορέσει να υποστηρίξει την αποθήκευση και επεξεργασία γεωχωρικών δεδομένων, καθώς και την εκτέλεση των αλγόριθμων που έχουν αναπτυχθεί για την εργασία. Το QGIS συνδέεται με εξωτερικές υπηρεσίες γεωχωρικών υπόβαθρων για χάρτες φόντου, και έχει μόνιμη σύνδεση με την PostgreSQL, από όπου αντλεί τα δεδομένα για τα υπόβαθρα, που παράγονται από τους αλγόριθμους της εργασίας.



Εικόνα 3.1: Διάγραμμα αρχιτεκτονικής της εφαρμογής

3.3 Αλγόριθμος υπολογισμού δείκτη χωρικών κριτηρίων

Για της ανάγκες της εργασίας αναπτύχθηκε αλγόριθμος που παράγει έναν δείκτη καταλληλότητας μιας περιοχής βάσει κάποιων κριτηρίων εισόδου. Συγκεκριμένα, λαμβάνει σαν είσοδο ένα σημείο και μία ή περισσότερες κατηγορίες στοιχείων ενδιαφέροντος (πχ σουπερμάρκετ, σταθμοί μέσων μεταφοράς, σχολεία, αυτοκινητόδρομοι, πάρκα), μαζί με παραμέτρους απόστασης, την βαρύτητα του κριτηρίου και την φορά της προτίμησης, δηλαδή αν προτιμά να είναι κοντά ή μακριά από το στοιχείο. Στην έξοδό του, ο αλγόριθμος παράγει

τον δείκτη καταλληλότητας, που είναι ένας αριθμός από το κλειστό διάστημα $[0,1]$, ο οποίος αναπαριστά, πόσο καλά το σημείο ικανοποιεί τα κριτήρια εισόδου, με το 1 να σημαίνει απόλυτα και το 0, καθόλου. Η τιμή του δείκτη Δ , είναι ο σταθμισμένος μέσος όρος των επιμέρους δεικτών δ , των κριτηρίων, όπου β είναι το βάρος κάθε κριτηρίου:

$$\Delta = \frac{\sum_{i=1}^n \beta_i \delta_i}{\sum_{i=1}^n \beta_i} \quad (1)$$

Οι επιμέρους δείκτες δ των κριτηρίων, υπολογίζονται από μια συνάρτηση ασαφούς συμμετοχής (fuzzy membership function). Συχνά χρησιμοποιούμενες συναρτήσεις ασαφούς συμμετοχής είναι οι σιγμοειδείς και οι γραμμικές, με τις γραμμικές να είναι αποτελεσματικές, απλές στην υλοποίηση και με μικρές υπολογιστικές απαιτήσεις (Đurađioná & Faixoná Chalachanová, 2017). Η συνάρτηση που χρησιμοποιήθηκε τελικά είναι μία τμηματική γραμμική (Εικόνα 3.2 αριστερά), η οποία έχει δύο παραμέτρους, α και β , που αντιστοιχούν στα όρια μεταβολής της συμμετοχής, και η οποία ορίζεται ως εξής:

$$f(x) = \begin{cases} 0, & x < \alpha \\ \frac{x-\alpha}{\beta-\alpha}, & \alpha \leq x \leq \beta \\ 1, & x > \beta \end{cases} \quad (2)$$

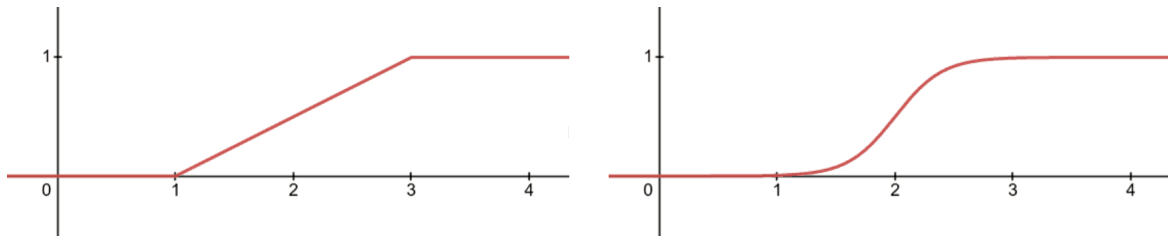
Μια άλλη μορφή της παραπάνω συνάρτησης (2), η οποία χρησιμοποιείται για την υλοποίηση σε SQL είναι η εξής:

$$f(x) = \min\left(1, \max\left(0, \frac{x-\alpha}{\beta-\alpha}\right)\right) \quad (3)$$

Για σύγκριση, χρησιμοποιήθηκε και μία λογιστική συνάρτηση (Εικόνα 3.2 δεξιά), από την ομάδα των σιγμοειδών, με τον εξής ορισμό:

$$f(x) = \frac{1}{1 + e^{\left(10 \left(\frac{x-a}{b-a} - 0.5\right)\right)}} \quad (4)$$

η οποία όμως μείωνε αισθητά την απόκριση της εφαρμογής, οπότε δεν επιλέχθηκε στην τελική έκδοση.



Εικόνα 3.2: Συναρτήσεις ασαφούς συμμετοχής: τμηματική γραμμική (αριστερά) και λογιστική (δεξιά)

Στις περισσότερες περιπτώσεις, η χρήση των συναρτήσεων ασαφούς συμμετοχής συνοδεύεται από ενοποίηση των συνόλων με ασαφείς τελεστές (ασαφείς εκδόσεις των AND, OR και NOT της άλγεβρας Μπουλ). Η χρήση των τελεστών αυτών όμως, αποκλείει την χρήση βαρών. Επίσης στην περίπτωση του τελεστή AND, χρησιμοποιείται για τον συγκερασμό ετερογενών κριτηρίων, δηλαδή όταν θέλουμε το σημείο να πληροί όλα τα κριτήρια ταυτόχρονα, και επιστρέφεται απλά το όρισμα με την μικρότερη τιμή. Σε αυτήν την περίπτωση, το αποτέλεσμα όχι μόνο δεν προκύπτει από την ίση συμμετοχή των κριτηρίων, αλλά καθορίζεται αποκλειστικά από ένα μόνο κριτήριο (αυτό με την μικρότερη τιμή), το οποίο ενδέχεται να είναι αυτό με την μικρότερη βαρύτητα.

Η μέθοδος του σταθμισμένου μέσου όρου ή σταθμισμένου γραμμικού συνδυασμού (weighted linear combination, WLC), που τελικά χρησιμοποιήθηκε, είναι η πιο συχνά χρησιμοποιούμενη μέθοδος συνάθροισης κριτηρίων, που λαμβάνει υπόψη τα βάρη τους, και συνιστάται για τα αρκετά πλεονεκτήματά της, σε περιπτώσεις αξιολόγησης και επιλογής τοποθεσιών (Aliani et al., 2017; Malczewski & Rinner, 2015).

3.4 Προετοιμασία δεδομένων

Για την λειτουργία της εφαρμογής ήταν απαραίτητο να συλλεχθούν και να προετοιμαστούν τα δεδομένα πάνω στα οποία δρα η εφαρμογή. Συγκεκριμένα, για τον υπολογισμό του δείκτη καταλληλότητας, χρειάστηκε να αντληθούν δημόσια διαθέσιμα δεδομένα για σημεία ενδιαφέροντος, και για την παρουσίαση των ακινήτων χρειάστηκε να αντληθούν δεδομένα από την αγορά ακινήτων. Επιλέχθηκε η περιοχή του κέντρου της Αθήνας, λόγω οικειότητας.

Επίσης, για την εκτέλεση του αλγόριθμου της εργασίας, χρειάστηκε να παραχθεί ένα πλέγμα από γεωχωρικά σημεία, σε κυψελοειδή διάταξη, για καθένα από τα οποία υπολογίζεται ο δείκτης καταλληλότητας βάσει των κριτηρίων του χρήστη.

3.4.1 Δημόσια διαθέσιμα δεδομένα για σημεία ενδιαφέροντος

Για τον υπολογισμό του δείκτη καταλληλότητας απαιτήθηκαν δύο είδη γεωχωρικών δεδομένων. Το πρώτο είδος αφορά σημεία ενδιαφέροντος όπως κεντρικοί δρόμοι, σουπερμάρκετ, σχολεία και στάσεις μέσων μεταφοράς, από τα οποία ο χρήστης μπορεί να θέλει να βρίσκεται κοντά ή μακριά. Το δεύτερο αφορά φαινόμενα όπως η εγκληματικότητα, η ρύπανση του περιβάλλοντος, οι υψομετρικές ζώνες και η κλίση και ο προσανατολισμός του εδάφους, τα οποία είτε είναι ενιαίας μορφής, ή καθορίζονται από όρια όπως οι δήμοι, οι ζώνες τιμών αντικειμενικών αξιών και οι χρήσεις γης, για τα οποία ο χρήστης θέλει να γνωρίζει την τιμή τους σε κάθε σημείο που μελετά.

Για την πρώτη περίπτωση δεδομένων, τις ανάγκες κάλυψε το Openstreetmap, το οποίο διαθέτει πολύ λεπτομερή στοιχεία και χαρακτηρισμούς για όλα τα παραπάνω σημεία ενδιαφέροντος, σε ενιαία και προσβάσιμα μορφή. Τα δεδομένα αφορούσαν σημεία, γραμμές και πολύγωνα. Αντλήθηκαν για την περιοχή του κέντρου της Αθήνας, μέσα από το πρόσθετο QuickOSM του QGIS, από το οποίο, στην συνέχεια, εισήχθησαν σε πίνακες της PostgreSQL, με χρήση του εργαλείου DB Manager του QGIS.

Για την δεύτερη περίπτωση, χρήσιμα δεδομένα για την εφαρμογή, θα ήταν αυτά της εγκληματικότητας και των ζωνών τιμών αντικειμενικών αξιών, όμως δεν κατέστη δυνατό να αντληθούν σε αξιοποιήσιμη μορφή, καθώς τα δεδομένων εγκληματικότητας ήταν συγκεντρωτικά, σε επίπεδο αστυνομικής διεύθυνσης νομού, και οι ζώνες τιμών αντικειμενικών αξιών ήταν σε μορφή υπηρεσίας εικονιστικών στοιχείων (tile server) ή σε περιγραφικούς πίνακες δρόμων με τα όρια των ζωνών. Επίσης για την κλίση και τον προσανατολισμό δεν βρέθηκαν απευθείας γεωχωρικά υπόβαθρα, είναι όμως δυνατόν να παραχθούν μετασχηματίζοντας υψομετρικά δεδομένα. Τέτοια υψομετρικά δεδομένα εντοπίστηκαν στην Διαδικτυακή Πύλη Γεωχωρικών Πληροφοριών του ΥΠΕΝ (<http://mapsportal.ypen.gr/>), αλλά είναι χαμηλής ανάλυσης και ο μετασχηματισμός τους

απαιτεί εφαρμογή μη τετριμμένων αλγορίθμων, που είναι εκτός του αντικειμένου της παρούσας εργασίας.

3.4.2 Δεδομένα από την αγορά ακινήτων

Για την ολοκλήρωση της εφαρμογής με την αγορά ακινήτων, ήταν απαραίτητη η άντληση δεδομένων ακινήτων από την τρέχουσα αγορά. Αυτό πραγματοποιήθηκε με την λήψη αγγελιών ακινήτων στις περιοχές που μελετήθηκαν. Συγκεκριμένα τα δεδομένα συλλέχθηκαν μέσα από τις υπολογιστικές διεπαφές API των χαρτών ιστοσελίδων, από τις υπηρεσίες ειδοποίησης μέσω email, καθώς και από αυτοματοποιημένη αποθήκευση κατά την περιήγηση. Στην συνέχεια τα δεδομένα υπέστησαν επεξεργασία και συσχετίστηκαν, ώστε να είναι κατάλληλα για χρήση στους αλγόριθμους της εφαρμογής.

Μια ιδιομορφία των δεδομένων από την αγορά ακινήτων είναι ότι περιέχουν ορισμένες ακραίες τιμές (outliers), οι οποίες, αν δεν αντιμετωπιστούν, δημιουργούν στρεβλώσεις τόσο στα στατιστικά αποτελέσματα, όσο και στην διάδραση και στις απεικονίσεις. Πρόκειται για μη φυσιολογικές εγγραφές, οι οποίες, ως επί το πλείστον, έχουν προκύψει από λάθος καταχώρηση, η οποία μερικές φορές είναι ακούσια, αλλά πολλές φορές είναι σκόπιμη. Στην περίπτωση του σκόπιμου λάθους, ο καταχωρητής συμπληρώνει κάποιο χαρακτηριστικό με πλασματική τιμή, με στόχο, είτε την ανάδειξη του ακινήτου σε ταξινομημένες λίστες, είτε τη μη αποκάλυψη της πραγματικής τιμής του χαρακτηριστικού, σε περιπτώσεις που η συμπλήρωσή του είναι υποχρεωτική. Ο εντοπισμός και ο περιορισμός της επίδρασης των ακραίων αυτών περιπτώσεων αντιμετωπίζονται με τεχνικές της οπτικής αναλυτικής, όπως αναφέρθηκε σε προηγούμενη ενότητα.

3.4.3 Παραγωγή πλέγματος διατεταγμένων σημείων

Για την περιοχή που εξετάζεται, αρχικά, παράγεται ένα πλέγμα από γεωχωρικά σημεία σε κυψελοειδή διάταξη. Τα σημεία αυτά, εκτός από τις γεωγραφικές τους συντεταγμένες, περιέχουν πεδία για την αποθήκευση της απόστασής τους από το κοντινότερο σημείο ενδιαφέροντος, καθώς και ένα πεδίο για την αποθήκευση της τιμής του δείκτη καταλληλότητας, που υπολογίζεται από τα κριτήρια που ορίζει κάθε φορά ο χρήστης.

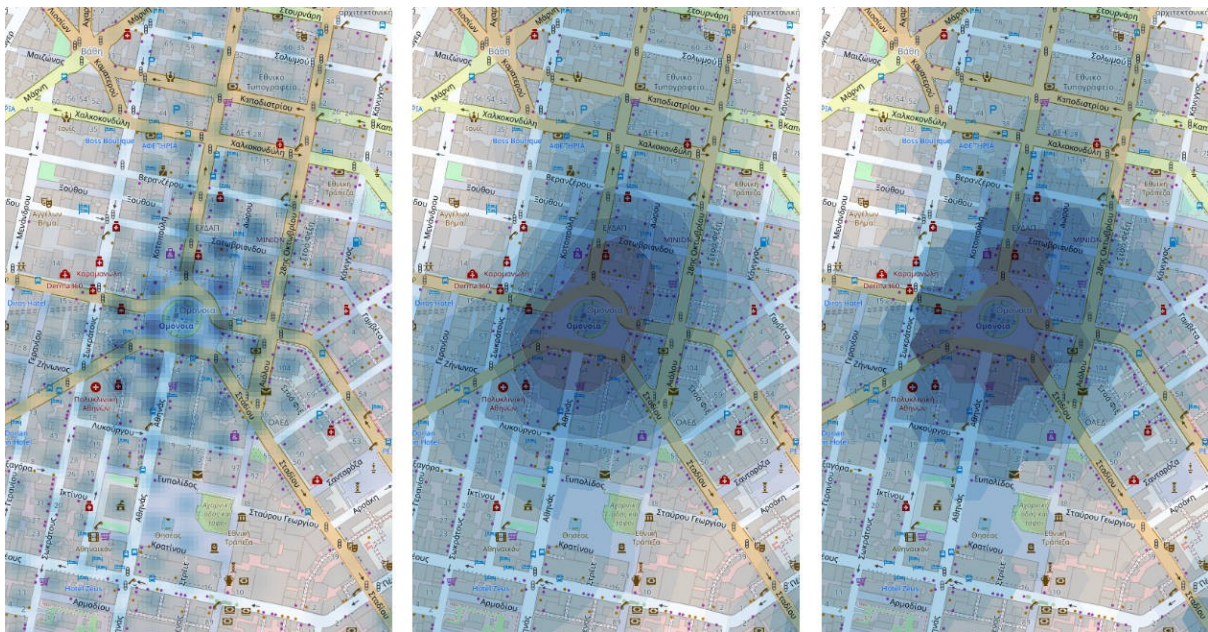
Για το πλέγμα, δοκιμάστηκαν τόσο η ορθογώνια όσο και η κυψελοειδής διάταξη, αλλά επικράτησε η κυψελοειδής λόγω της καλύτερης οπτικής ευκρίνειας (Birch et al., 2007). Για την δημιουργία του πλέγματος χρησιμοποιήθηκε το εργαλείο “Vector > Research Tools > Create Grid” του QGIS, με παράμετρο “Grid type”=Hexagon. Η απόσταση μεταξύ των εξαγώνων ορίστηκε με εμπειρική εκτίμηση στα 50 μέτρα, ώστε κάθε σημείο να χαρακτηρίζει μια αρκετά μικρή περιοχή, αλλά όχι τόσο μικρή που να επηρεάσει την απόκριση της εφαρμογής, λόγω μεγάλου όγκου δεδομένων και υπολογιστικών απαιτήσεων. Επειδή το εργαλείο αυτό δημιουργεί ένα γεωγραφικό επίπεδο με πολύγωνα, στο παραγόμενο πλέγμα εφαρμόζεται το εργαλείο “Vector > Geometry Tools > Centroids“, που κατασκευάζει ένα νέο γεωγραφικό επίπεδο με πλέγμα σημείων από τα κέντρα των εξαγώνων πολυγώνων.

Τα σημεία του πλέγματος, από το γεωγραφικό επίπεδο, εισάγονται στην PostgreSQL, με το εργαλείο DB Manager του QGIS, το οποίο δημιουργεί ένα νέο πίνακα στην βάση, με δύο πεδία: ένα αναγνωριστικό (id) και των γεωγραφικών συντεταγμένων (geom). Για κάθε κατηγορία σημείου ενδιαφέροντος που μπορεί να χρειαστεί να ληφθεί ως κριτήριο (πχ σουπερμάρκετ, σχολεία, πάρκα, κεντρικές οδοί, σταθμοί μετρό και λεωφορείων), προστίθεται από ένα πεδίο ακεραίας τιμής. Τα πεδία αυτά χρησιμοποιούνται για να αποθηκεύεται η απόσταση μεταξύ του σημείου του πλέγματος και του κοντινότερου σημείου της κατηγορίας σημείων ενδιαφέροντος.

Για τον υπολογισμό της κοντινότερης απόστασης από τα σημεία ενδιαφέροντος, εκτελείται ένα SQL ερώτημα για κάθε κατηγορία σημείων ενδιαφέροντος. Το ερώτημα συνδυάζει τον πίνακα του πλέγματος με τους πίνακες των σημείων ενδιαφέροντος και, για κάθε σημείο του πλέγματος, χρησιμοποιεί τον τελεστή απόστασης “<->” για βρει το κοντινότερο σημείο ενδιαφέροντος, αποθηκεύοντας, στην συνέχεια, την μεταξύ τους απόσταση στο αντίστοιχο πεδίο της κατηγορίας σημείων ενδιαφέροντος. Η διαδικασία αυτή είναι χρονοβόρα, αλλά χρειάζεται να πραγματοποιηθεί μόνο μία φορά για κάθε περιοχή και για όλα τα πιθανά κριτήρια. Οπότε οι υπολογισμοί αυτοί γίνονται πριν χρησιμοποιηθεί η εφαρμογή. Ενδεικτικά, για το κέντρο της Αθήνας και για πλέγμα με απόσταση 50 μέτρων μεταξύ των σημείων, ο υπολογισμός των αποστάσεων διήρκεσε περίπου 30 δευτερόλεπτα έως 2 λεπτά για κάθε

κατηγορία σημείων ενδιαφέροντος, ανάλογα με τον αριθμό των σημείων ενδιαφέροντος και με το αν είναι σε μορφή σημείων, γραμμών ή επιπέδων.

Για να αξιοποιηθούν τα σημεία του πλέγματος έπρεπε να απεικονιστούν στον χάρτη. Αυτό δοκιμάστηκε με τρεις μεθόδους απεικόνισης φαινομένων: με την χρήση συμβόλων για κάθε σημείο, με θερμικό χάρτη και με περιγράμματα. Η απεικόνιση με σύμβολα (Εικόνα 3.3 δεξιά) έγινε με παραμετροποίηση των στυλ των σημείων. Συγκεκριμένα επιλέχθηκε το σύμβολο του εξαγώνου, με διάμετρο τέτοια που να εφάπτονται τα περιγράμματα των γειτονικών εξαγώνων. Η απεικόνιση με θερμικό χάρτη (Εικόνα 3.3 αριστερά) έγινε με την επιλογή Heatmap για είδος συμβολισμού των σημείων, με επιλογή του πεδίου του δείκτη για την παράμετρο της βαρύτητας (weight points by), και με διάμετρο 50 μέτρων. Η απεικόνιση με ισομετρικά περιγράμματα (Εικόνα 3.3 κέντρο) έγινε με χρήση του QGIS πρόσθετο “Countour plugin”, το οποίο εφαρμόζεται στα σημεία του πλέγματος και παράγει ένα νέο γεωγραφικό επίπεδο με τα αποτελέσματα.



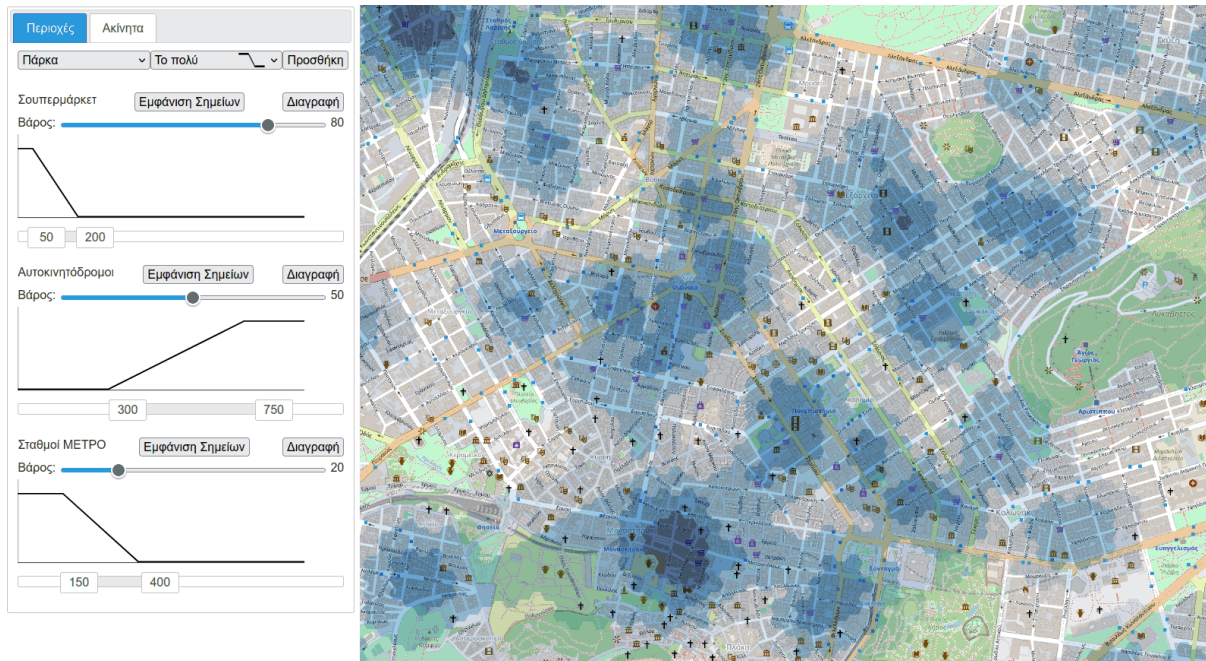
Εικόνα 3.3: Απεικόνιση σημείων πλέγματος με θερμικό χάρτη (αριστερά), περιγράμματα (κέντρο) και εξαγώνα σύμβολα (δεξιά)

Ο θερμικός χάρτης παράγεται από το QGIS σε πραγματικό χρόνο για τα σημεία που αναπαριστά. Χρησιμοποιείται κυρίως για προβολή κατανομών σημείων στον χώρο, απεικονίζοντας μεγαλύτερη πυκνότητα σημείων με πιο έντονο χρώμα. Επειδή τα σημεία

έχουν ομοιόμορφη κατανομή στον χώρο, η μέθοδος αυτή δεν λειτουργεί αποτελεσματικά. Αυτό συμβαίνει γιατί η απεικόνιση του φαινομένου θα πρέπει να βασίζεται στην τιμή των σημείων και όχι στον βαθμό συγκέντρωσής τους. Αντίθετα, οι απεικονίσεις με περιγράμματα και με εξάγωνα σύμβολα αποτυπώνουν καλά το φαινόμενο. Όμως η χρήση της μεθόδου των περιγραμμάτων απαιτεί ένα παραπάνω υπολογιστικό βήμα, το οποίο πρέπει να εφαρμοστεί μετά τον υπολογισμό των δεικτών του συνόλου των σημείων, επιβαρύνοντας την ταχύτητα απόκρισης της εφαρμογής. Συνεπώς, επιλέχθηκε η μέθοδος απεικόνισης με τα εξάγωνα σύμβολα.

3.5 Λειτουργία της εφαρμογής

Η οθόνη της εφαρμογής (Εικόνα 3.4) είναι χωρισμένη σε δύο ενότητες. Στην αριστερή ενότητα βρίσκονται τα εργαλεία προσδιορισμού και παραμετροποίησης των κριτηρίων, ενώ στην δεξιά ενότητα, που καταλαμβάνει και το μεγαλύτερο μέρος της οθόνης, εμφανίζεται ένας διαδραστικός χάρτης. Ο χρήστης αλληλεπιδρά με τα στοιχεία των εργαλείων (κουμπιά, μοχλούς κύλισης, πτυσσόμενα μενού, κουτάκια επιλογής), με τα οποία επιλέγει και παραμετροποιεί τα κριτήρια των προτιμήσεών του. Τα αποτελέσματα των ενεργειών του, στην συνέχεια, εμφανίζονται στον χάρτη, τον οποίο μπορεί να κυλίσει και να εστιάσει. Επίσης, μπορεί να αλληλεπιδράσει με τα στοιχεία του χάρτη, πατώντας ή περνώντας από πάνω τους με το ποντίκι, για να εμφανιστούν περισσότερες επιλογές και πληροφορίες.

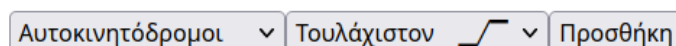


Εικόνα 3.4: Η οθόνη της εφαρμογής

Μέσα από την αλληλεπίδραση με τα εργαλεία και τον χάρτη, ο χρήστης εξερευνά τοποθεσίες και την αγορά ακινήτων. Με κάθε επανάληψη της εναλλαγής μεταξύ εκλέπτυνσης των κριτηρίων και εξερεύνησης στον χάρτη, ο χρήστης ανακαλύπτει περιοχές και γειτονιές και εξοικειώνεται με τα ιδιαίτερα χαρακτηριστικά τους. Την ίδια στιγμή, εντοπίζει μοτίβα στην αγορά ακινήτων, ενώ προσδιορίζει ή αναθεωρεί τις ανάγκες του και τις ιεραρχεί με βάση τις δυνατότητες και τους περιορισμούς του.

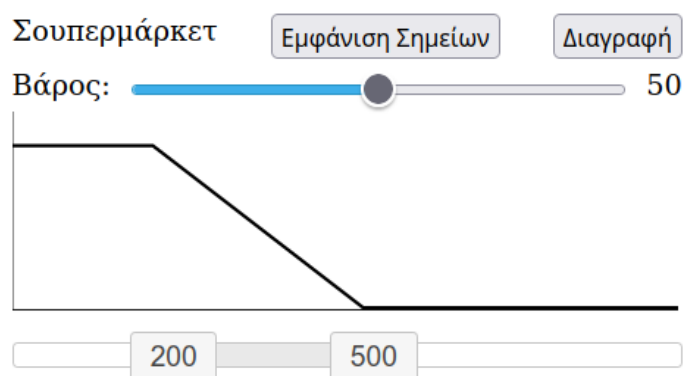
3.5.1 Εισαγωγή και παραμετροποίηση κριτηρίων

Από το μενού εισαγωγής κριτηρίων, στην αριστερή ενότητα, στο πάνω μέρος της καρτέλας “Περιοχές” (Εικόνα 3.5), ο χρήστης εισάγει κριτήρια για τις απαιτήσεις του. Για κάθε απαίτηση επιλέγει από τα αντίστοιχα μενού την κατηγορία και την φορά του κριτηρίου, και πατάει το κουμπί προσθήκη για να εμφανιστεί από κάτω η φόρμα παραμετροποίησής του.



Εικόνα 3.5: Μενού εισαγωγής κριτηρίων

Στην φόρμα παραμετροποίησης (Εικόνα 3.6) εμφανίζονται το όνομα του κριτηρίου, το κουμπί για την εμφάνιση των σημείων της κατηγορίας στον χάρτη, το κουμπί διαγραφής του κριτηρίου, η γραφική παράσταση της συνάρτησης και τα στοιχεία ελέγχου για την παραμετροποίησή του. Τα στοιχεία ελέγχου είναι ένας μοχλός κύλισης για τον καθορισμό του βάρους και ένας διπλός μοχλός κύλισης για τον καθορισμό των ορίων της συνάρτησης συμμετοχής. Η τιμή που έχει επιλεγεί με τον μοχλό κύλισης του βάρους εμφανίζεται στα δεξιά του και ενημερώνεται καθώς ο χρήστης μετακινεί τον μοχλό. Οι τιμές των μοχλών καθορισμού των ορίων της συνάρτησης εμφανίζονται πάνω στους ίδιους τους μοχλούς, η μετακίνηση των οποίων ενημερώνει αυτόματα και την γραφική παράσταση της συνάρτησης.

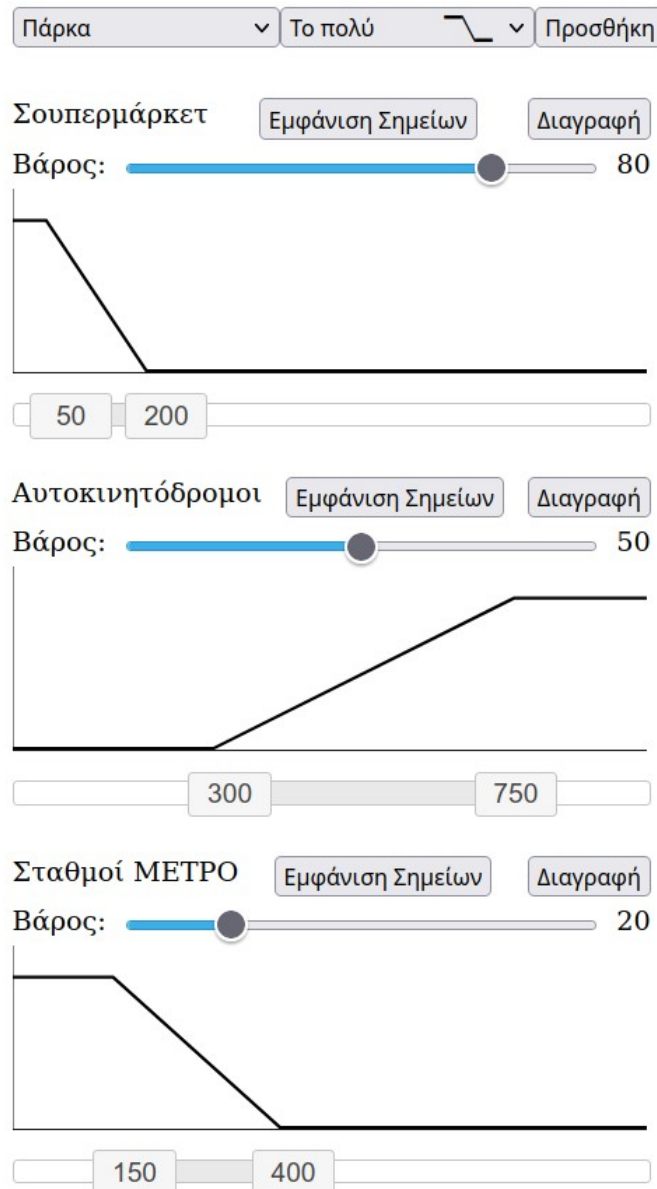


Εικόνα 3.6: Φόρμα παραμετροποίησης κριτηρίου

Κάθε νέο κριτήριο που εισάγεται, αμέσως συμμετέχει στον υπολογισμό του δείκτη. Κάθε μεταβολή στα όρια ή στο βάρος της συνάρτησης με τα στοιχεία ελέγχου, πυροδοτεί εκ νέου τον υπολογισμό του δείκτη, λαμβάνοντας σαν είσοδο όλα τα κριτήρια που έχουν εισαχθεί μέχρι εκείνη την στιγμή, και τα οποία εμφανίζονται στην αριστερή ενότητα της εφαρμογής.

Για παράδειγμα, στην Εικόνα 3.7 απεικονίζεται ένα απλό σενάριο με τρία κριτήρια, για την αναζήτηση περιοχών που μπορούν να τα ικανοποιήσουν. Στο σενάριο αυτό, ο χρήστης ψάχνει για ένα γεωγραφικό σημείο που, ιδανικά, θα έχει απόσταση από το κοντινότερο σουπερμάρκετ, το πολύ 50 μέτρα και σίγουρα όχι περισσότερα από 200 μέτρα, ενώ το βάρος του κριτηρίου είναι 80. Επίσης, το σημείο αυτό θα πρέπει να βρίσκεται τουλάχιστον 300 μέτρα μακριά από τον κοντινότερο αυτοκινητόδρομο και ιδανικά, από 750 μέτρα και πάνω, ενώ το βάρος του κριτηρίου είναι 50. Τέλος, το σημείο αυτό, ιδανικά, θα έχει απόσταση από

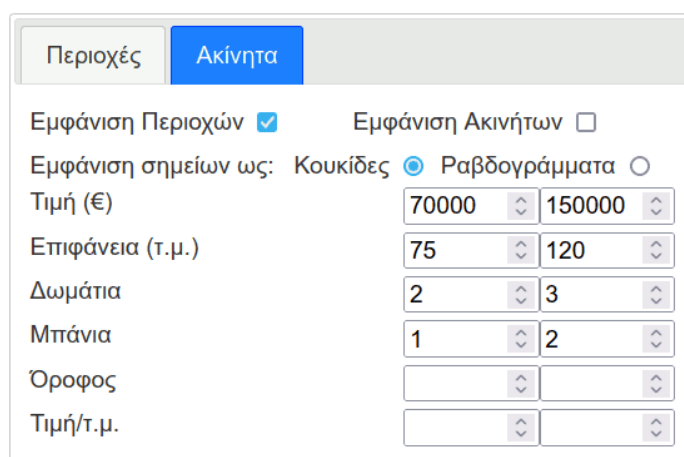
τον κοντινότερο σταθμό ΜΕΤΡΟ, το πολύ 150 μέτρα και σίγουρα όχι περισσότερα από 400 μέτρα, ενώ το βάρος του κριτηρίου είναι 20. Επίσης, στο μενού εισαγωγής κριτηρίων (Εικόνα 3.7 κορυφή), φαίνεται ότι ο χρήστης ετοιμάζεται να προσθέσει ένα νέο κριτήριο μέγιστης απόστασης από πάρκα.



Εικόνα 3.7: Σενάριο χρήσης της εφαρμογής

3.5.2 Εξερευνητική αλληλεπίδραση με τον χάρτη και με τα στοιχεία ελέγχου

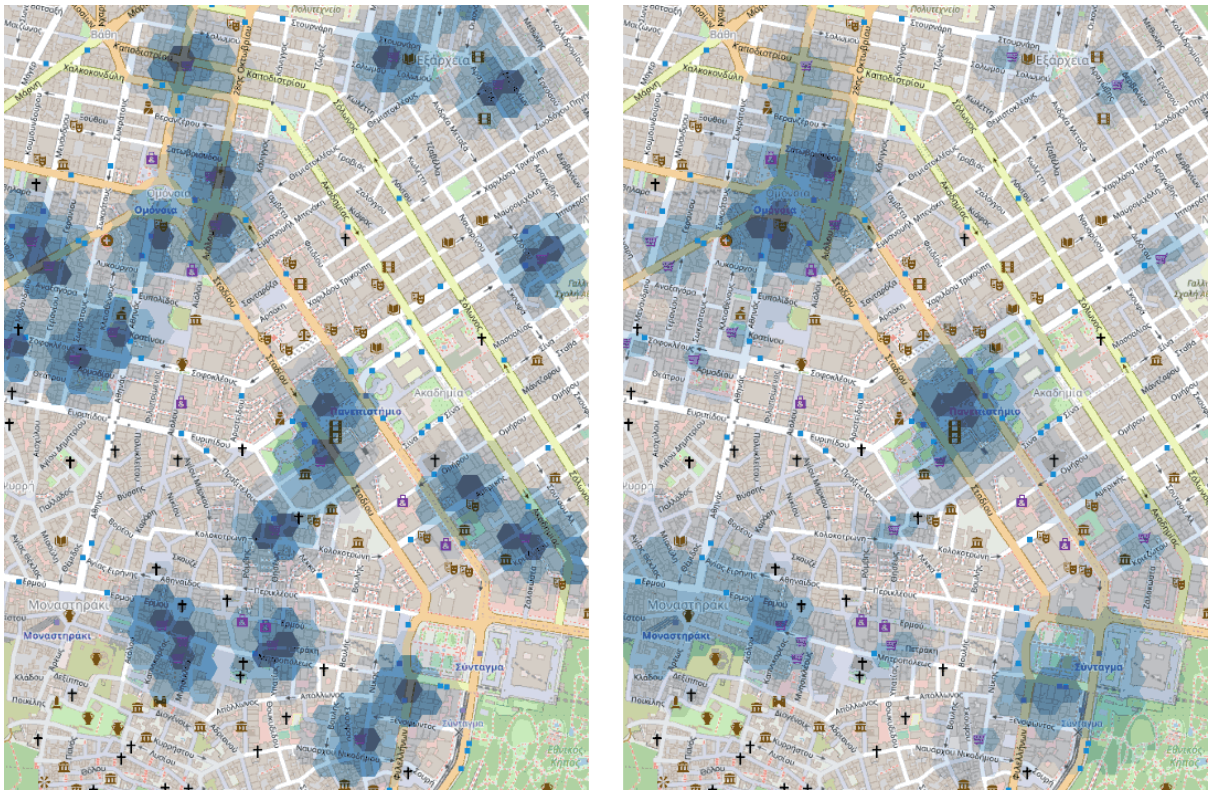
Αφού έχει εισάγει και παραμετροποιήσει τα πρώτα κριτήρια, ο χρήστης περιεργάζεται τον χάρτη για να κατανοήσει πώς αυτά αποτυπώνονται στις διάφορες περιοχές και γειτονιές. Στην συνέχεια επιστρέφει στα εργαλεία, όπου μπορεί να εκλεπτύνει τα κριτήρια και να εμφανίσει τα σημεία τους στον χάρτη, πατώντας το κουμπί “Εμφάνιση Σημείων” για την κατηγορία σημείων ενδιαφέροντος που επιθυμεί. Επίσης, μπορεί να μεταβεί στην καρτέλα των ακινήτων (Εικόνα 3.8), από όπου μπορεί να επιλέξει να τα εμφανίσει στον χάρτη, πατώντας στο κουτάκι επιλογής “Εμφάνιση Ακινήτων”, να επιλέξει την μορφή των συμβόλων τους, αλλά και να εισάγει αριθμητικά κριτήρια για τα χαρακτηριστικά τους, ώστε να περιορίσει τον αριθμό των αποτελεσμάτων.



Περιοχές	Ακίνητα
Εμφάνιση Περιοχών <input checked="" type="checkbox"/>	Εμφάνιση Ακινήτων <input type="checkbox"/>
Εμφάνιση σημείων ως: Κουκίδες <input checked="" type="radio"/>	Ραβδογράμματα <input type="radio"/>
Τιμή (€)	70000 <input type="text"/> 150000 <input type="text"/>
Επιφάνεια (τ.μ.)	75 <input type="text"/> 120 <input type="text"/>
Δωμάτια	2 <input type="text"/> 3 <input type="text"/>
Μπάνια	1 <input type="text"/> 2 <input type="text"/>
Όροφος	<input type="text"/> <input type="text"/>
Τιμή/τ.μ.	<input type="text"/> <input type="text"/>

Εικόνα 3.8: Καρτέλα ακινήτων

Εκτός από την εισαγωγή και διαγραφή κριτηρίων, ο χρήστης μπορεί να τα εκλεπτύνει, χρησιμοποιώντας τα γραφικά στοιχεία της φόρμας του κάθε κριτηρίου, με τα αποτελέσματα των ενεργειών του να εμφανίζονται αμέσως στον χάρτη. Για παράδειγμα, στην Εικόνα 3.9, αριστερά, φαίνεται το υπόβαθρο καταλληλότητας, με μοναδικό κριτήριο την απόσταση από σουπερμάρκετ και στην Εικόνα 3.9, δεξιά, το ίδιο υπόβαθρο μετά την προσθήκη του κριτηρίου της απόστασης από ΜΕΤΡΟ, με το ίδιο βάρος. Επίσης, για όποιο κριτήριο επιθυμεί, μπορεί να εμφανίσει ή να αποκρύψει τα σημεία του στον χάρτη, πατώντας το κουμπί “Εμφάνιση Σημείων”. Στην Εικόνα 3.10 εμφανίζονται σημεία για τους σταθμούς ΜΕΤΡΟ (πράσινοι ρόμβοι) και για τα σουπερμάρκετ (κόκκινοι κύκλοι).

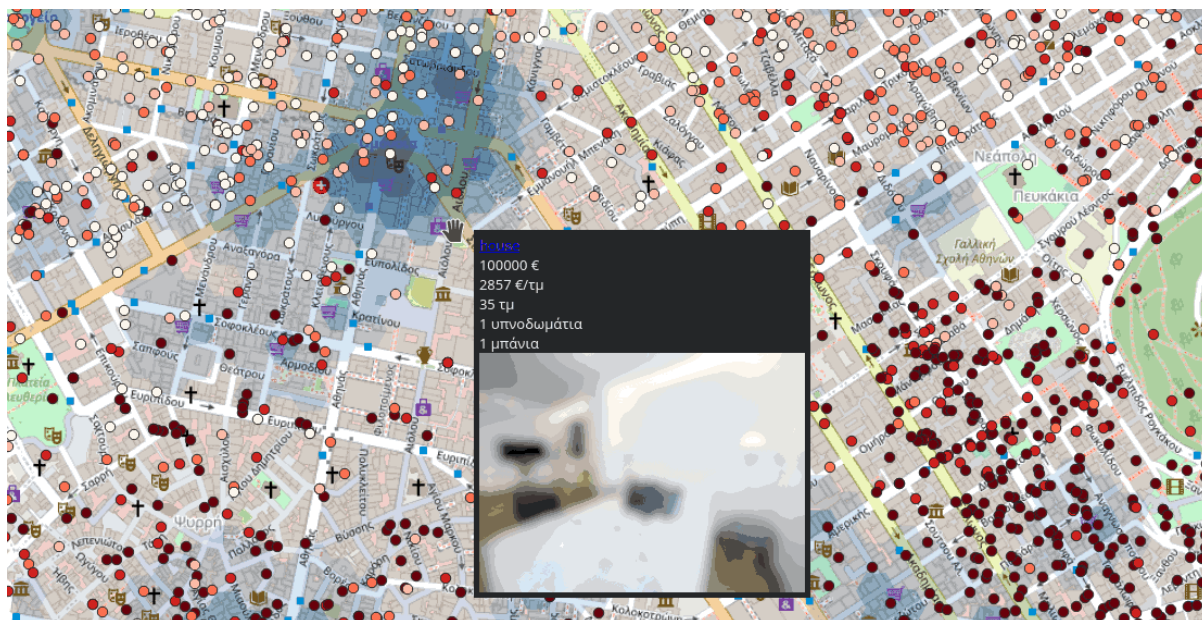


Εικόνα 3.9: Μεταβολή υπόβαθρου δείκτη μετά από εκλέπτυνση κριτηρίων

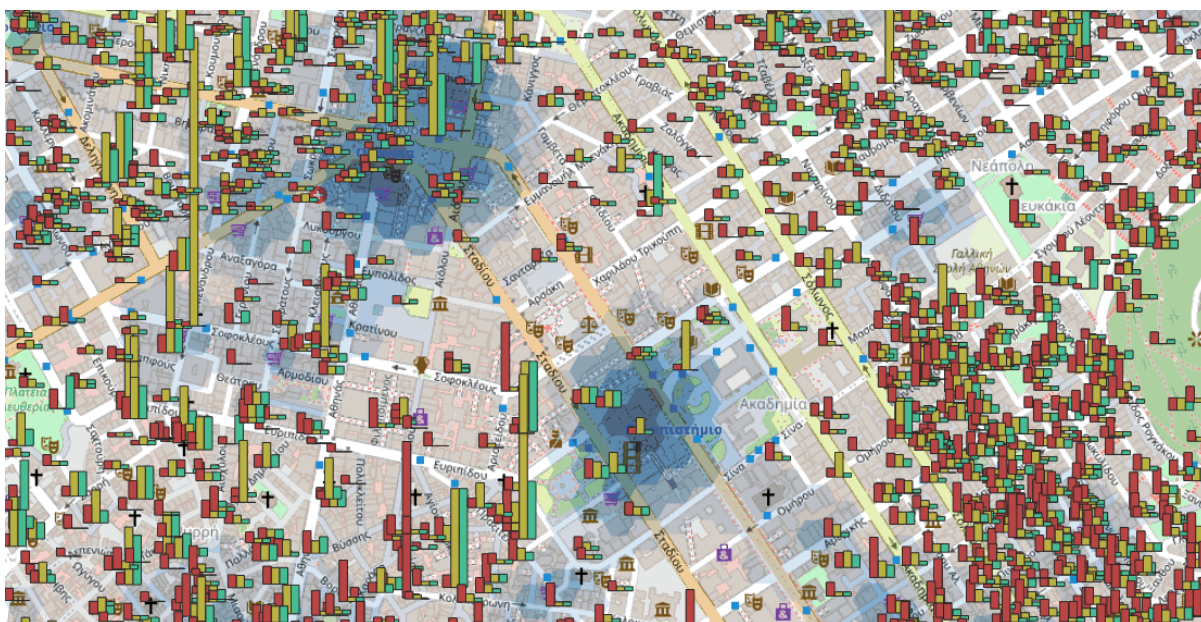


Εικόνα 3.10: Εμφάνιση σταθμών ΜΕΤΡΟ (πράσινοι ρόμβοι) και σουπερμάρκετ (κόκκινοι κύκλοι)

Μεταβαίνοντας στην καρτέλα των ακινήτων, ο χρήστης έχει την δυνατότητα να τα εμφανίσει στον χάρτη, επιλέγοντας αν θα εμφανιστούν ως κουκίδες ή ως ραβδογράμματα και, μετακινώντας το ποντίκι πάνω από κάποιο ακίνητο, εμφανίζεται καρτέλα με τα στοιχεία του και μια φωτογραφία. Αν επιλέξει να εμφανίζονται ως κουκίδες εμφανίζονται στον χάρτη μικροί κύκλοι, που τα χαρακτηριστικά τους (χρώμα, διάμετρος, διαφάνεια κτλ) αναπαριστούν τιμές χαρακτηριστικών των ακινήτων (Εικόνα 3.11). Στην συγκεκριμένη περίπτωση, χρησιμοποιείται μόνο το χρώμα τους, για να αναπαραστήσει την τιμή ανά τετραγωνικό μέτρο, σε μία κλίμακα 5 αποχρώσεων μεταξύ κόκκινου και άσπρου. Αν επιλέξει να εμφανίζονται ως ραβδογράμματα, για κάθε ακίνητο εμφανίζεται ένα ραβδόγραμμα με τρεις κάθετες ράβδους διαφορετικών χρωμάτων που, με το ύψος τους, αναπαριστούν την τιμή ανά τετραγωνικό μέτρο, τον αριθμό δωματίων και τον αριθμό των μπάνιων (Εικόνα 3.12). Σε κάθε περίπτωση, μπορεί να εισάγει τιμές στα αριθμητικά κριτήρια για τα χαρακτηριστικά τους, ώστε να μειώσει τον αριθμό των ακινήτων που εμφανίζονται στον χάρτη.



Εικόνα 3.11: Εμφάνιση αγγελιών ακινήτων σε μορφή κουκίδων



Εικόνα 3.12: Εμφάνιση αγγελιών ακινήτων σε μορφή ραβδογραμμάτων

Μέσα από επαναλαμβανόμενες εναλλαγές μεταξύ εκλέπτυνσης των κριτηρίων και αλληλεπίδρασης με τον χάρτη, ο χρήστης εξοικειώνεται με τα ιδιαίτερα χαρακτηριστικά των τοποθεσιών και κατανοεί τις σχέσεις επίδρασης με τους διάφορους παράγοντες. Η τεχνική αυτή της γεωοπτικής αναλυτικής επιστρατεύει τις αντιληπτικές και συλλογιστικές ικανότητες του χρήστη, υποστηρίζοντάς τες με διαδραστικά εργαλεία και υπολογιστικές και απεικονιστικές τεχνικές.

3.6 Αξιολόγηση της εφαρμογής

Η αξιολόγηση της εφαρμογής πραγματοποιήθηκε με τρεις χρήστες, ηλικίας 35-70 ετών, οι οποίοι είχαν προηγουμένως εκφράσει την ανάγκη τους για αγορά κατοικίας. Η αξιολόγηση πραγματοποιήθηκε με την δημοφιλή τεχνική της παρατήρησης των χρηστών, την ώρα που αλληλεπιδρούν με αυτήν (Dix et al., 2003, p343). Αφού τους περιγράφηκε ο σκοπός και η λειτουργία της εφαρμογής, τους ζητήθηκε να εξερευνήσουν μια περιοχή με στόχο να κατανοήσουν τα χαρακτηριστικά της και στην συνέχεια να αναζητήσουν ακίνητα που θα μπορούσαν να ικανοποιήσουν τις ανάγκες τους. Το πλαίσιο της αξιολόγησης βασίστηκε στα 10 κριτήρια που έθεσαν οι Forsell & Johansson (2010) για πολύπλοκες διαδραστικές οθόνες

οπτικοποίησης πληροφοριών, τα οποία αναφέρονται στην ενότητα της βιβλιογραφικής επισκόπησης. Επειδή η απλή παρατήρηση είναι σπάνια επαρκής για να προσδιοριστεί πόσο καλά το σύστημα ανταποκρίνεται στις απαιτήσεις των χρηστών, δεδομένου ότι δεν παρέχει πάντα εικόνα των διαδικασιών λήψης αποφάσεων ή της συμπεριφοράς τους, τους ζητήθηκε να αναλύσουν τις ενέργειές τους “σκεπτόμενοι δυνατά”. Μετά την ολοκλήρωση της εργασίας, πραγματοποιήθηκε συνέντευξη με τους χρήστες, για την κατανόηση των κινήτρων πίσω από τις ενέργειες που πραγματοποίησαν κατά την διάρκεια της αλληλεπίδρασης με την εφαρμογή (Dix et al., 2003, p343).

3.6.1 Αποτελέσματα

Οι περισσότεροι χρήστες βρήκαν την εφαρμογή εύκολη και διαισθητική στην χρήση και κατατοπιστική. Τους βοήθησε να αποκτήσουν πληροφόρηση και γνώση για τις περιοχές που εξερεύνησαν, καθώς και για την κατανόηση των φαινομένων και των μοτίβων, αλλά και τις σχέσεις με τα αίτια που τα παράγουν. Τους κίνησε την περιέργεια και τους γέννησε νέα ερωτήματα, τα οποία ήθελαν να ερευνήσουν περαιτέρω. Τα εργαλεία της γεωοπτικής αναλυτικής τους βοήθησαν να εξερευνήσουν περιοχές, με τρόπο που δεν θα μπορούσαν να πραγματοποιήσουν με εργαλεία που παρέχονται από ιστοσελίδες αγγελιών ακινήτων. Τα σημαντικότερα ευρήματα για την βελτίωση της εφαρμογής αναφέρονται παρακάτω.

Οι κυριότερες παρατηρήσεις για τις φόρμες παραμετροποίησης των χωρικών κριτηρίων και των ακινήτων ήταν οι εξής:

- Θα ήθελαν το κουμπί για ενεργοποίηση περιοχών να βρίσκεται στην καρτέλα των περιοχών.
- Κάποιοι δυσκολεύτηκαν να κατανοήσουν την σχέση της γραφικής παράστασης με τον ορισμό των ορίων απόστασης, ιδιαίτερα κατά την εισαγωγή του κριτηρίου, όταν έπρεπε να επιλέξουν την φορά από το μενού εισαγωγής.
- Σε κάποιους δεν ήταν ξεκάθαρα τα όρια της κάθε φόρμας κριτηρίου, δηλαδή συγγέανε ποιο κουμπί άνηκε σε ποια φόρμα.

- Ορισμένοι, θα προτιμούσαν το κουμπί “Εμφάνιση Σημείων” κάθε κριτηρίου να είναι με κουτάκι επιλογής αντί για κουμπί, όπως στην καρτέλα των ακινήτων.
- Κάποιοι θα προτιμούσαν οι παράμετροι αναζήτησης ακινήτων να εισάγονται με διπλούς μοχλούς κύλισης, όπως γίνεται στις φόρμες για τις περιοχές.

Οι κυριότερες παρατηρήσεις για τον χάρτη ήταν οι εξής:

- Στο πλαίσιο που αναδύεται όταν μετακινούν το ποντίκι πάνω από ένα σημείο του υπόβαθρου καταλληλότητας, εκτός από την τιμή του δείκτη για το σημείο, θα ήθελαν να έβλεπαν τις επιμέρους αποστάσεις που παράγουν τον δείκτη, δηλαδή πόσο μακριά βρίσκεται από τις κατηγορίες σημείων ενδιαφέροντος, είτε ως πίνακας αριθμών είτε ως ραβδόγραμμα.
- Όταν εμφάνισαν όλες τις πληροφορίες στον χάρτη, δηλαδή το υπόβαθρο του δείκτη, τα σημεία για όλες τις κατηγορίες σημείων ενδιαφέροντος και τα ακίνητα, μερικοί χρήστες θεώρησαν ότι υπήρχε υπερβολική πληροφορία και, ειδικά σε κάποια επίπεδα ζουμ, τα σύμβολα αλληλοεπικαλύπτονταν και γίνονταν δυσδιάκριτα.
- Θα ήθελαν να υπάρχει η δυνατότητα επιλογής δύο ή περισσότερων ακινήτων από τον χάρτη, ώστε να μπορούν να συγκριθούν τα χαρακτηριστικά τους σε πίνακα.

4 Επίλογος

4.1 Συνεισφορά

Η κύρια συνεισφορά της παρούσας εργασίας είναι ένα εργαλείο γεωοπτικής αναλυτικής που χρησιμοποιεί εξερευνητική ανάλυση γεωχωρικών δεδομένων για την υποστήριξη λήψης αποφάσεων, με στόχο την κατανόηση και τον προσδιορισμό των απαιτήσεων από την πλευρά του χρήστη, και στην συνέχεια την εύρεση ακινήτων που ικανοποιούν τις απαιτήσεις αυτές. Η εφαρμογή ξεχωρίζει για την δυνατότητα που παρέχει στον χρήστη να καθορίζει ο ίδιος τα κριτήρια και την βαρύτητά τους, για αναζήτηση γεωγραφικών τόπων με βάση τις απαιτήσεις του. Επίσης, ο αλγόριθμος υπολογισμού του δείκτη καταλληλότητας, συνυπολογίζει και αρνητικούς γεωγραφικούς παράγοντες όπως την ηχορύπανση, που αυξάνουν – αντί να μειώνουν – την τιμή του δείκτη, με την αύξηση της απόστασης από αυτούς.

Ενώ τα περισσότερα εργαλεία αναζήτησης ακινήτων επιτρέπουν στον χρήστη να αναζητήσει ακίνητα με βάση τα χαρακτηριστικά τους, η εφαρμογή του δίνει επιπλέον την δυνατότητα να εξερευνήσει και περιοχές με βάση τις προτιμήσεις του. Έτσι, μπορεί να πραγματοποιεί επιλογές βάσει δεδομένων και, τελικά, να λαμβάνει τεκμηριωμένες αποφάσεις.

Η παρούσα εργασία εστιάζει στην ανάπτυξη εργαλείων που θα βοηθήσουν τον χρήστη να ανακαλύψει εξερευνητικά, περιοχές που μπορούν να ικανοποιήσουν τις απαιτήσεις του. Θεωρεί δηλαδή ως ζητούμενο, τον προσδιορισμό ή την εξειδίκευση τόσο των απαιτήσεων, όσο και των περιοχών ενδιαφέροντος του χρήστη. Παρέχει στον χρήστη εργαλεία εξερεύνησης, με τα οποία πλοηγείται στα δεδομένα, δοκιμάζοντας σενάρια, αναθεωρώντας προαιρετικά και επαναλαμβάνοντας την διαδικασία, μέχρι να καταλήξει σε ικανοποιητικό προσδιορισμό των αναγκών του. Το τελικό αποτέλεσμα της παραπάνω διαδικασίας είναι ένα χαρτογραφικό υπόβαθρο, το οποίο αναπαριστά έναν δείκτη που ενσωματώνει τα κριτήρια που επιλέγει ο χρήστης. Το υπόβαθρο αυτό μπορεί να χρησιμοποιηθεί με διάφορους τρόπους, ενδεικτικά:

- απευθείας ως χαρτογραφική απεικόνιση, σε συνδυασμό με άλλα υπόβαθρα
- ως φίλτρο, για να περιορίσει την εμφάνιση σημείων ενδιαφέροντος

- με ενσωμάτωση, ως νέο χαρακτηριστικό, σε σημεία ενδιαφέροντος

Επίσης, το πεδίο τιμών του αλγόριθμου δεν είναι δυαδικό, αλλά παίρνει τιμές από το κλειστό διάστημα $[0,1]$, σε αντίθεση με πολλές περιπτώσεις φιλτραρίσματος, που έχουν για πεδίο τιμών το σύνολο $\{0,1\}$. Έτσι, οι περιοχές που σχηματίζονται ως αποτέλεσμα του φιλτραρίσματος δεν έχουν αυστηρά όρια, αλλά παράγουν σχηματισμούς βαθμίδων καταλληλότητας, που βοηθούν στην καλύτερη αντίληψη των φαινομένων, αφού απεικονίζουν την τάση τους, σε αντίθεση με ένα αυστηρό όριο, πριν και μετά από το οποίο, ο χρήστης δεν μπορεί να δει τι ισχύει για την περιοχή που εξετάζει. Επίσης, σε αντίθεση με άλλες εφαρμογές, ο δείκτης υπολογίζεται για ένα πλέγμα σημείων μιας ολόκληρης περιοχής και όχι μόνο για επιλέξιμα σημεία, όπως τοποθεσίες ακινήτων, επιτρέποντας στον χρήστη να μελετήσει και να εξερευνήσει περιοχές, ανεξάρτητα από την ύπαρξη επιλέξιμων σημείων.

Μία ακόμη συνεισφορά της εργασίας, είναι το ότι δίνονται βάρη στις απαιτήσεις, ώστε ο δείκτης τελικά να αναπαριστά καλύτερα τις προτεραιότητες του χρήστη. Αυτό αποδίδει καλύτερα από ένα απλό φίλτρο απόρριψης, αφού, τελικά, συμπεριλαμβάνονται στα αποτελέσματα περισσότερα σημεία που ικανοποιούν τις απαιτήσεις του χρήστη και λιγότερα που δεν τις ικανοποιούν.

Μια ακόμη σημαντική συνεισφορά της εργασίας είναι ότι, όχι μόνο δίνει στον χρήστη την δυνατότητα να επιλέξει ο ίδιος κριτήρια, αλλά και να επιλέξει το ίδιο κριτήριο περισσότερες της μίας φορές, ενδεχομένως και με αντίθετες παραμέτρους. Αρχικά αυτό μπορεί να φαίνεται αντιφατικό, όμως υπάρχουν έγκυρες περιπτώσεις που χρειάζεται κάτι τέτοιο. Ένα παράδειγμα αποτελεί η απόσταση από κεντρικούς δρόμους. Είναι απολύτως κατανοητό ένας χρήστης που ψάχνει για σπίτι, να θέλει να είναι αρκετά μακριά από κεντρικούς δρόμους, ώστε να αποφύγει την ηχορύπανση και το καυσαέριο, ταυτόχρονα όμως, να μην είναι και πολύ μακριά, ώστε να ικανοποιεί το κριτήριο της προσβασιμότητας.

Επιπλέον, ο αλγόριθμος μπορεί να εφαρμοστεί τόσο σε μία περιοχή, με υπολογισμό του δείκτη σε κάθε σημείο του πλέγματος, όσο και σε σημεία ενδιαφέροντος όπως διαθέσιμα ακίνητα προς πώληση ή ενοικίαση. Δεν αποκλείεται δε και ο συνδυασμός των παραπάνω, δηλαδή η δημιουργία ενός χρωματικού υπόβαθρου για τα χαρακτηριστικά μιας περιοχής, πάνω από το οποίο θα εμφανίζεται άλλο υπόβαθρο με σημεία για διαθέσιμα ακίνητα, με

κάποιο χαρακτηριστικό του συμβόλου να απεικονίζει τον δείκτη, πχ χρώμα, σχήμα ή μέγεθος.

4.2 Επόμενα βήματα

Το πρωτότυπο της εφαρμογής υλοποιήθηκε στο QGIS, λόγω των γραφικών εργαλείων ταχείας ανάπτυξης, που επέτρεψε την δοκιμή πολλών σεναρίων σε σύντομο χρόνο. Το περιβάλλον αυτό όμως, δεν είναι κατάλληλο για διαχείριση από το κοινό στο οποίο απευθύνεται, καθώς απαιτεί πολύπλοκη εγκατάσταση επιμέρους εφαρμογών, και έχει απότομη καμπύλη εκμάθησης. Έτσι, σε δεύτερο χρόνο θα μπορούσε να υλοποιηθεί εξολοκλήρου ως web εφαρμογή, με στόχο την εύκολη πρόσβαση σε ένα ενιαίο περιβάλλον. Στην έκδοση αυτή, το στοιχείο του χάρτη θα μπορούσε να επικοινωνεί αμφίδρομα με την υπόλοιπη εφαρμογή, ώστε η διάδραση του χρήστη με αυτόν, να ενημερώνει και τα άλλα στοιχεία αλληλεπίδρασης, όπως τα διαγράμματα, τους μοχλούς κύλισης κτλ.

Οι δείκτες που παράγονται από τον αλγόριθμο που περιγράφεται στην παρούσα εργασία, θα μπορούσαν να χρησιμοποιηθούν ως επιπλέον είσοδοι σε μοντέλα αυτοματοποιημένης αποτίμησης ακινήτων, με στόχο την βελτίωση της ακρίβειας των αποτελεσμάτων τους.

Μελλοντικά, οι αποστάσεις από ορισμένα σημεία ενδιαφέροντος, θα μπορούσαν να υπολογιστούν με αλγόριθμους δρομολόγησης ή ακόμα και ως χρόνοι μετάβασης, ανάλογα με το μέσο που θα επιλέξει ο χρήστης.

Η τεχνική που αναπτύχθηκε στην παρούσα εργασία δεν περιορίζεται μόνο στην αγορά και ενοικίαση ακινήτων, αλλά θα μπορούσε να εφαρμοστεί και σε άλλους τομείς. Μια χαρακτηριστική περίπτωση είναι η αναζήτηση τουριστικών προορισμών. Η εξερευνητική αναζήτηση μπορεί να γίνει είτε με βάση θεματικούς τομείς, όπως πολιτισμός, αθλητισμός φύση και διασκέδαση, είτε με βάση ιδιαίτερα ενδιαφέροντα και απαιτήσεις του χρήστη, όπως το να κάνει ταξίδι αναψυχής σε περιοχή που να έχει κοντά παραλίες, δάσος και ποδηλατικές διαδρομές. Σε αυτήν την περίπτωση, θα πρέπει να εισαχθούν στην εφαρμογή νέα γεωχωρικά δεδομένα που να καλύπτουν τον εκάστοτε τομέα, καθώς και τα ανάλογα πεδία αποστάσεων στην βάση δεδομένων και τα κριτήρια στην διεπαφή χρήστη. Ταυτόχρονα θα πρέπει να

ληφθεί υπόψη ότι η κάθε περίπτωση μπορεί να απαιτεί διαφορετική κλίμακα αναζήτησης, οπότε θα πρέπει να υλοποιηθεί η ανάλογη δυνατότητα προσαρμογής της εφαρμογής.

Βιβλιογραφικές Αναφορές

- Aliani, H., BabaieKafaky, S., Saffari, A., & Monavari, S. M. (2017). Land evaluation for ecotourism development—an integrated approach based on FUZZY, WLC, and ANP methods. *International Journal of Environmental Science and Technology*, 14(9), 1999-2008.
- Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, A., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International journal of geographical information science*, 21(8), 839-857.
- Andrienko, N., Andrienko, G., Fuchs, G., Slingsby, A., Turkay, C., & Wrobel, S. (2020). *Visual analytics for data scientists*. Basingstoke, UK: Springer International Publishing.
- Anscombe, F. J. (1973). Graphs in statistical analysis. *The american statistician*, 27(1), 17-21.
- Anselin, L. (2017). *The GeoDa Book Exploring Spatial Data*.
- Aydinoglu, A. C., & Bovkir, R. (2020). Developing a Mobile Application for Smart Real Estate Information. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 44, 89-94.
- Birch, C. P., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological modelling*, 206(3-4), 347-359.
- Carr, D. B., Littlefield, R. J., Nicholson, W. L., & Littlefield, J. S. (1987). Scatterplot matrix techniques for large N. *Journal of the American Statistical Association*, 82(398), 424-436.
- Chang, K. T. (2016). Geographic information system. *International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology*, 1-9.
- Cui, W. (2019). Visual analytics: A comprehensive overview. *IEEE Access*, 7, 81555-81573.
- Demetriou, D. (2018). Automating the land valuation process carried out in land consolidation schemes. *Land Use Policy*, 75, 21-32.
- Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2003). *Human-computer interaction*. Pearson Education.

Forsell, C., & Johansson, J. (2010, May). An heuristic set for evaluation in information visualization. In Proceedings of the International Conference on Advanced Visual Interfaces (pp. 199-206).

Fu, Y., Xiong, H., Ge, Y., Yao, Z., Zheng, Y., & Zhou, Z. H. (2014, August). Exploiting geographic dependencies for real estate appraisal: A mutual perspective of ranking and clustering. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1047-1056).

Glumac, B., & Des Rosiers, F. (2020). Practice briefing—Automated valuation models (AVMs): Their role, their advantages and their limitations. *Journal of Property Investment & Finance*.

Gudivada, V. N., Irfan, M. T., Fathi, E., & Rao, D. L. (2016). Cognitive analytics: Going beyond big data analytics and machine learning. In *Handbook of statistics* (Vol. 35, pp. 169-205). Elsevier.

Hsu, L. S., & Obe, R. (2021). *PostGIS in action*. Simon and Schuster.

IAAO (2018). Standard on Automated Valuation Models (AVMs). International Association Of Assessing Officers. Retrieved from https://www.iaao.org/media/standards/Standard_on_Automated_Valuation_Models.pdf

Hao, M. C., Dayal, U., Keim, D., Janetzko, H., & Hill, W. (2017). *Generalized Scatter Plots*.

Karimi, H. A., & Karimi, B. (Eds.). (2017). *Geospatial data science techniques and applications*. CRC Press.

Keenan, P. B., & Jankowski, P. (2019). Spatial decision support systems: Three decades on. *Decision Support Systems*, 116, 64-76.

Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual analytics: Definition, process, and challenges. In *Information visualization* (pp. 154-175). Springer, Berlin, Heidelberg.

Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010a). *Mastering the information age: solving problems with visual analytics*.

Kethireddy, S. R., Adegoye, G. A., Tchounwou, P. B., Tuluri, F., Ahmad, H. A., Young, J. H., & Zhang, L. (2018). The status of geo-environmental health in Mississippi: Application of spatiotemporal statistics to improve health and air quality. *AIMS environmental science*, 5(4), 273.

Kraak, M. J. (2008). Editorial from geovisualisation toward geovisual analytics. *The cartographic journal*, 45(3), 163-164.

Kraak, M. J., & Ormeling, F. (2010). *Cartography: visualization of geospatial data*. CRC Press.

Krotov, V., & Silva, L. (2018). *Legality and ethics of web scraping*.

Li, M. (2019). *Visual analytics of geo-related multidimensional data* (Doctoral dissertation, RMIT University).

Malczewski, J., & Rinner, C. (2015). *Multicriteria decision analysis in geographic information science* (Vol. 1, pp. 55-77). New York: Springer.

Matejka, J., & Fitzmaurice, G. (2017, May). Same stats, different graphs: generating datasets with varied appearance and identical statistics through simulated annealing. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 1290-1294).

Meirelles, I. (2013). *Design for information: an introduction to the histories, theories, and best practices behind effective information visualizations*. Rockport publishers.

Michaelis, C. D., & Ames, D. P. (2008). *Web Feature Service (WFS) and Web Map Service (WMS)*.

Mooney, P., & Minghini, M. (2022). *Geospatial Data Exchange Using Binary Data Serialization Approaches*. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 4.

Muggenhuber, G. (2019). *Geospatial Data Mining and Analytics for Real-Estate Applications*. In *Service-Oriented Mapping* (pp. 225-240). Springer, Cham.

Navarra, C., Opach, T., Vrotsou, K., Joling, A., Wilk, J., & Neset, T. S. (2020). *Visual exploration of climate-related volunteered geographic information*. In *Workshop on*

Visualisation in Environmental Sciences (EnvirVis). Eurographics-European Association for Computer Graphics.

Nitschke, S., Biguzzi, A., Bücken, A., Magnolfi, S., Pellemans, T., Selleri, F., & Vetrano, P. (2022). European standards for statistical valuation methods for residential properties. Tech. rep. European AVM Alliance.

Popelka, S., Herman, L., Řezník, T., Pařilová, M., Jedlička, K., Bouchal, J., ... & Charvát, K. (2019). User evaluation of map-based visual analytic tools. ISPRS International Journal of Geo-Information, 8(8), 363.

PostGIS (2022). PostGIS 3.3.2dev Manual. <https://postgis.net/stuff/postgis-3.3.pdf>

PostgreSQL (2022). PostgreSQL 14.5 Documentation. The PostgreSQL Global Development Group. <https://www.postgresql.org/files/documentation/pdf/14/postgresql-14-A4.pdf>

QGIS (2022). Desktop 3.22 User Guide. QGIS Project. <https://docs.qgis.org/3.22/pdf/en/QGIS-3.22-DesktopUserGuide-en.pdf>

Rabiei-Dastjerdi, H., McArdle, G., Matthews, S. A., & Keenan, P. (2021). Gap analysis in decision support systems for real-estate in the era of the digital earth. International Journal of Digital Earth, 14(1), 121-138.

Robinson, A. (2017). Geovisual analytics. The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2017 Edition), edited by John P. Wilson. UCGIS. <https://doi.org/10.22224/gistbok/2017.3>, 6.

Srivastava, S., & Khot, P. G. (2018). GIS-based Computational Tools & Techniques for Multidimensional Data Analysis & Visualization. International Journal of Applied Engineering Research, 13(15), 11770-11775.

Thomas, J. J. & Cook, K. A. (2005). Illuminating the path: The research and development agenda for visual analytics (No. PNNL-SA-45230). Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Tukey, J. W. (1977). Exploratory data analysis

Ward, M. O., Grinstein, G., & Keim, D. (2015). Interactive data visualization: foundations, techniques, and applications. AK Peters/CRC Press.

Ware, C. (2020). Information visualization: perception for design. Morgan Kaufmann.

Wong, P. C., & Thomas, J. (2004). Guest Editors' Introduction--Visual Analytics. IEEE Computer Graphics and Applications, 24 (5): 20-21, 24(PNNL-SA-41935).

Zheng, Y., Wu, W., Chen, Y., Qu, H., & Ni, L. M. (2016). Visual analytics in urban computing: An overview. IEEE Transactions on Big Data, 2(3), 276-296.

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.