Applying big data and statistics techniques to create urban archetypes for long-term modelling

Matthieu Denoux[1*], Nadia Maïzi[1]

[1]MINES ParisTech, PSL Research University, CMA - Center for Applied Mathematics, Sophia-Antipolis, France
[*]matthieu.denoux@mines-paristech.fr

SophIA Summit, Sophia-Antipolis and everywhere else - 18/19/20 November 2020

Plan

Introduction and context

Urban planning, long-term model and machine learning

Figure: Classic urban plan

Urban planning, long-term model and machine learning

Figure: Classic urban plan

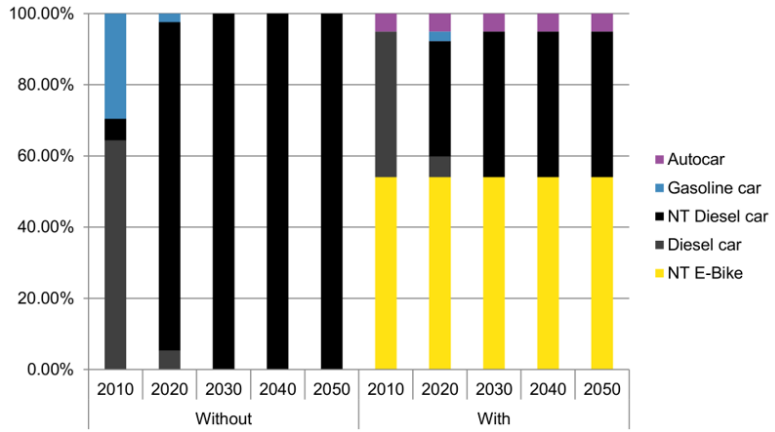


Figure: Typical long-term modeling results.

Urban planning, long-term model and machine learning

Figure: Classic urban plan

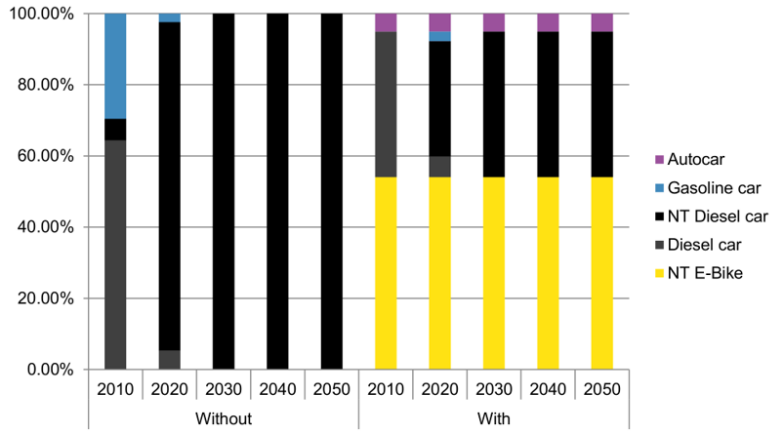


Figure: Typical long-term modeling results.

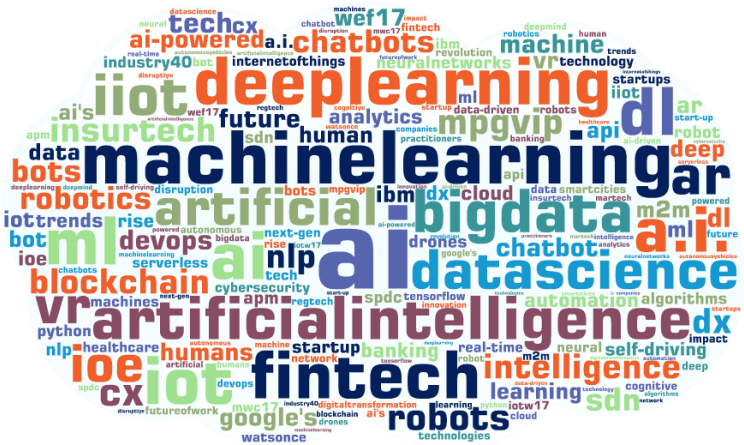
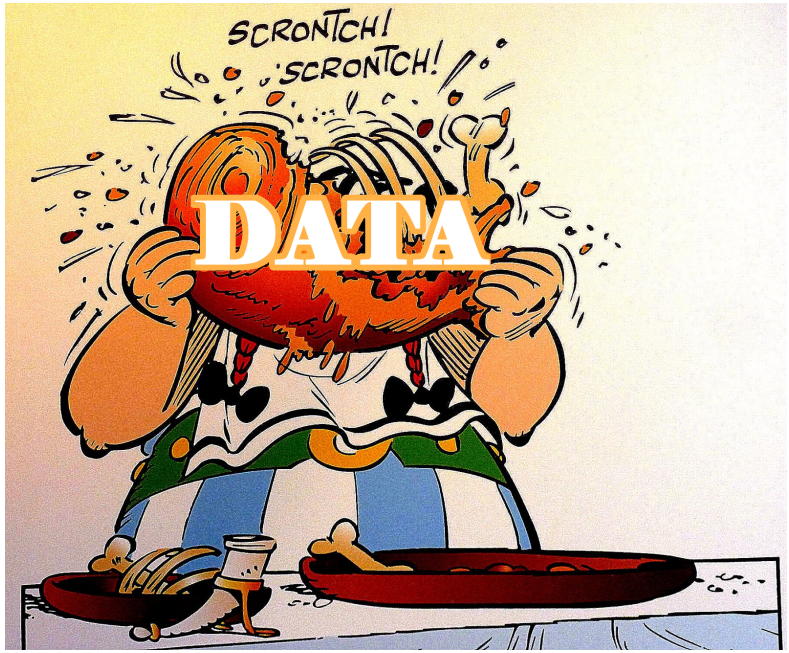


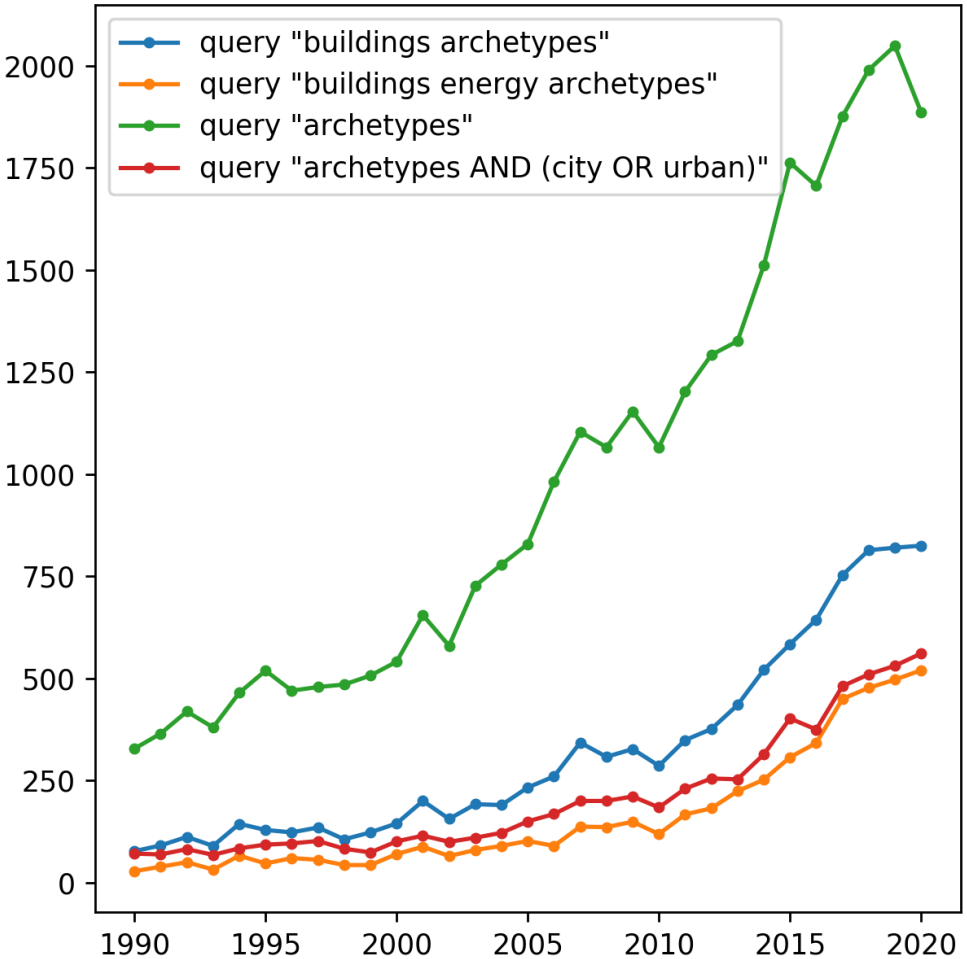
Figure: French platform on OpenData

MINES ParisTech | PSL

Objective



Objective



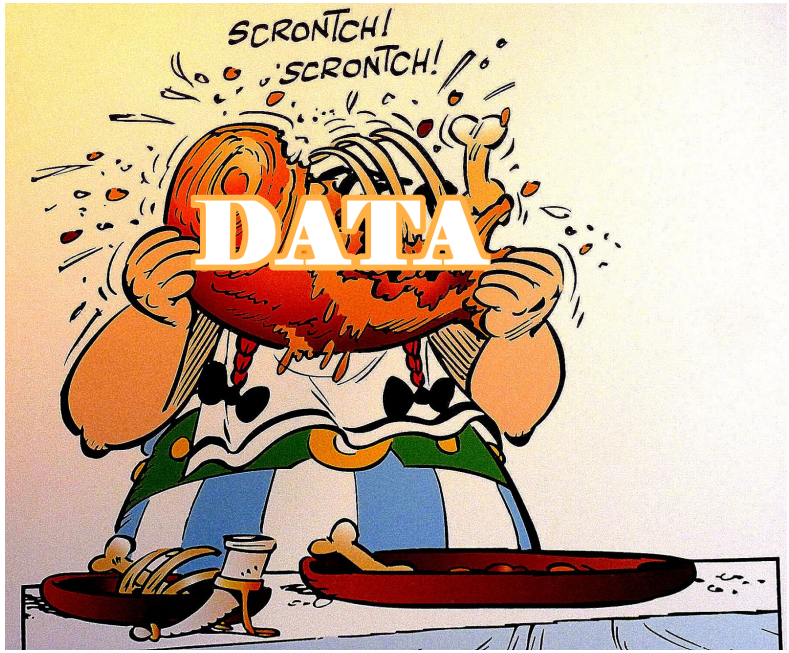


Figure: Results on `ScienceDirect` of various queries around *urban archetypes*

Objective

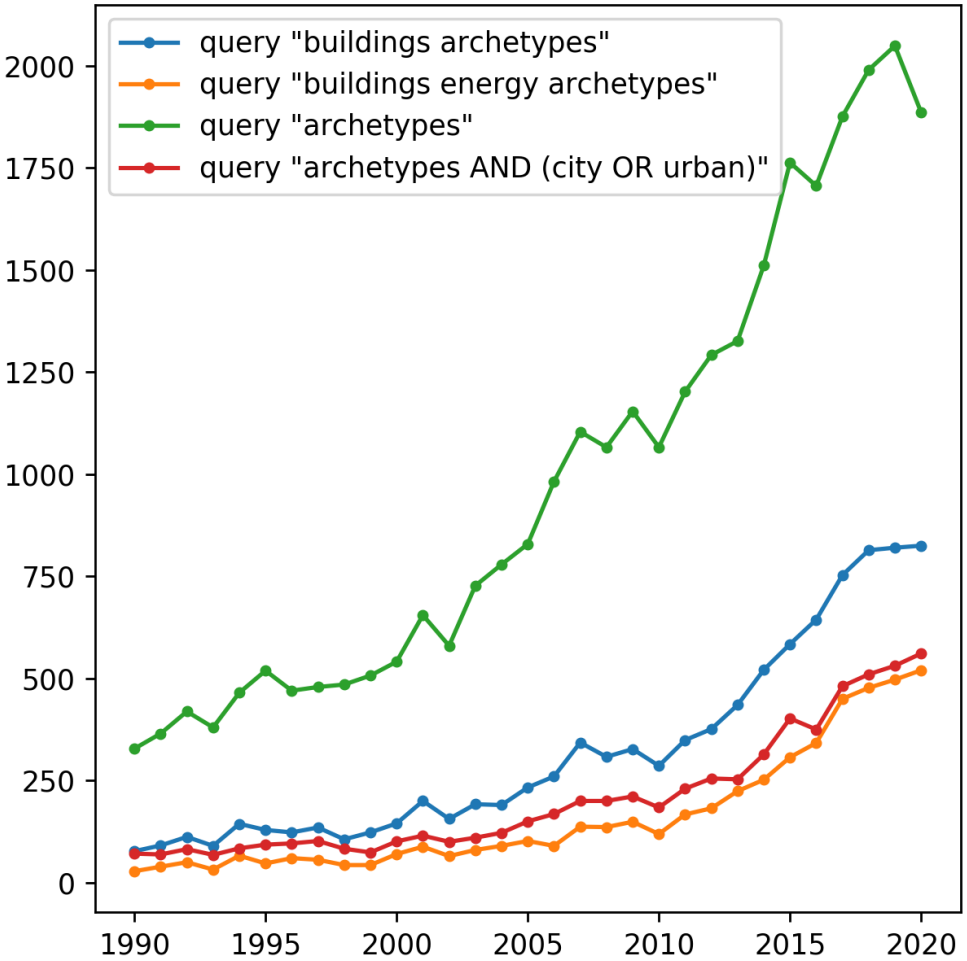
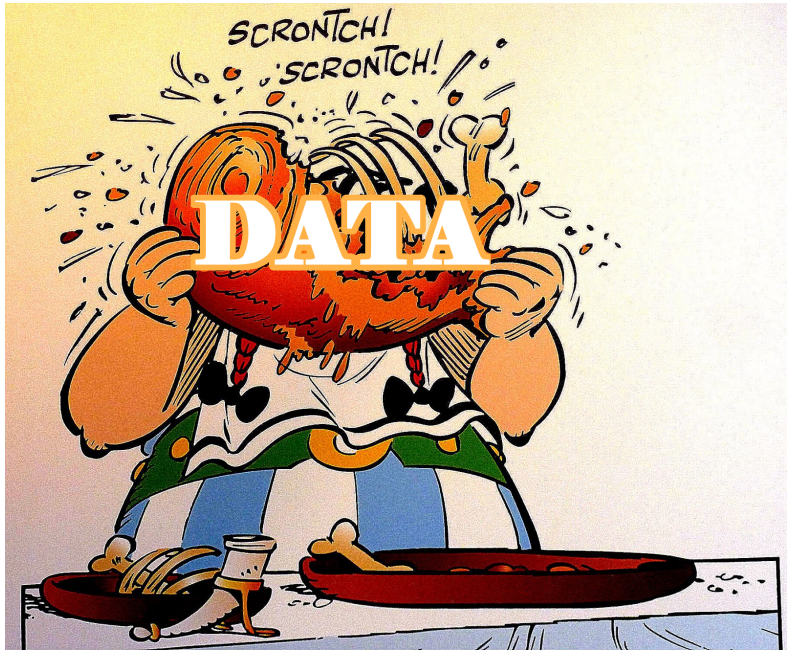


Figure: Results on `ScienceDirect` of various queries around *urban archetypes*



Our goal:
Obtaining a quantitative representation of urban system through *urban shapes* **and** *buildings* **typologies / archetypes.**

A methodology to create archetypes

Methodology proposed and deployed

1. Scale choice

ie.

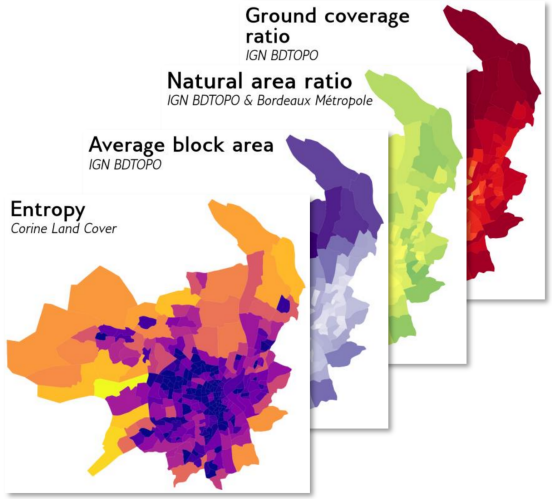
Selecting territory and disaggregation boundaries

2. Data selection

ie.

Choosing and processing parameters for both categories

(buildings and urban shapes)

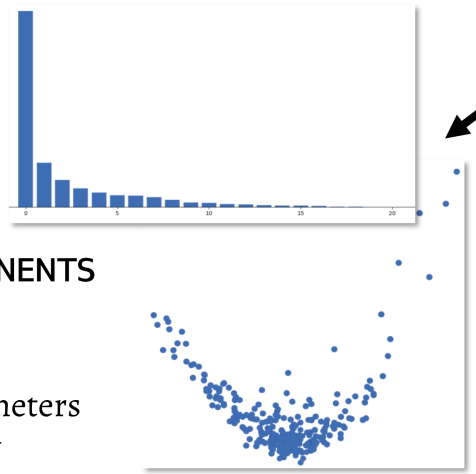


3. Principal components analysis

ie.

Extracting main parameters for each category

(buildings and urban shapes)

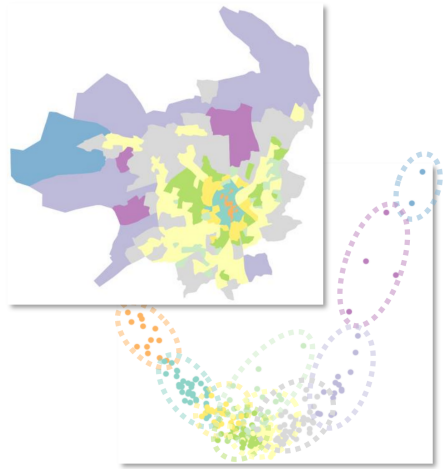


4. K-means

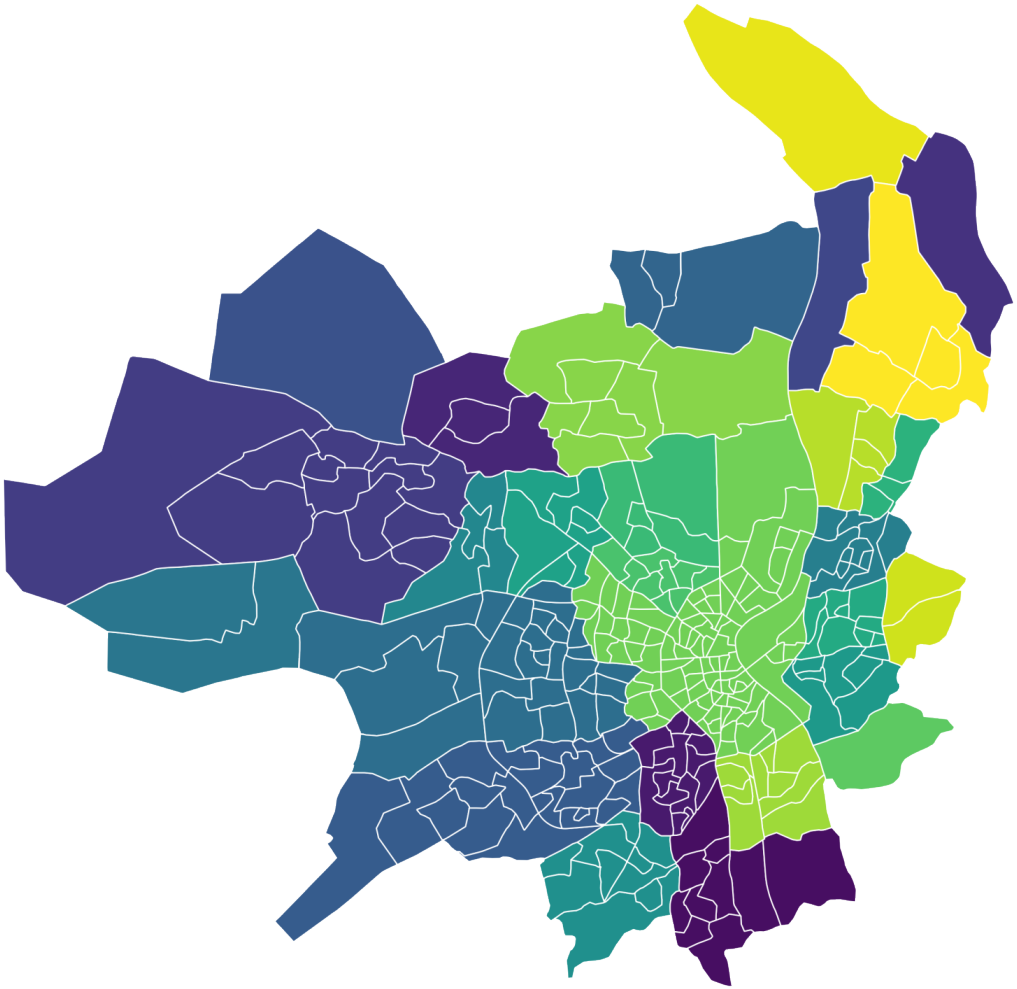
ie.

Creating archetypes for each category

(buildings and urban shapes)



From Bordeaux Métropole to the IRIS layer



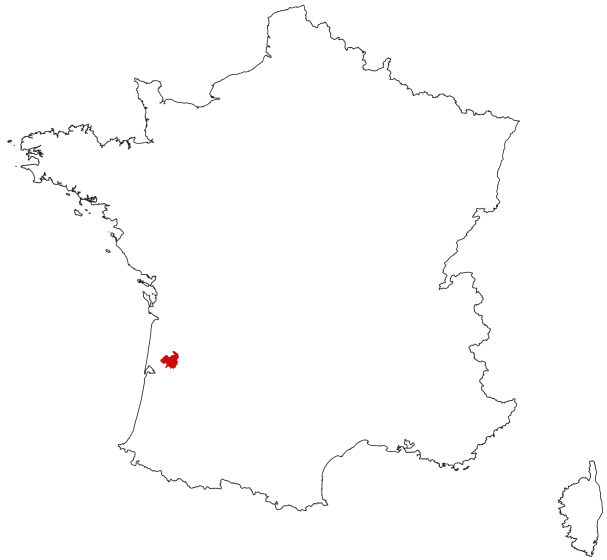


Figure: Bordeaux Métropole - gathering 27 cities, one per color.

From Bordeaux Métropole to the IRIS layer

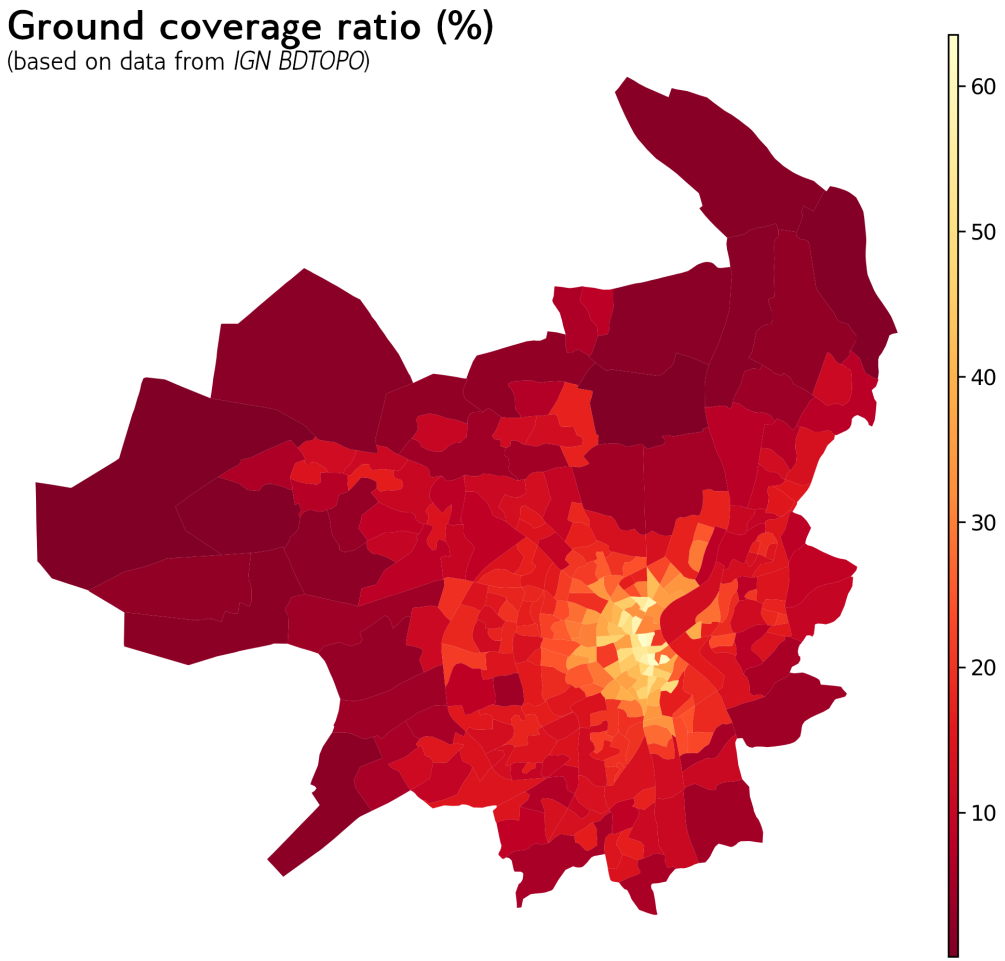
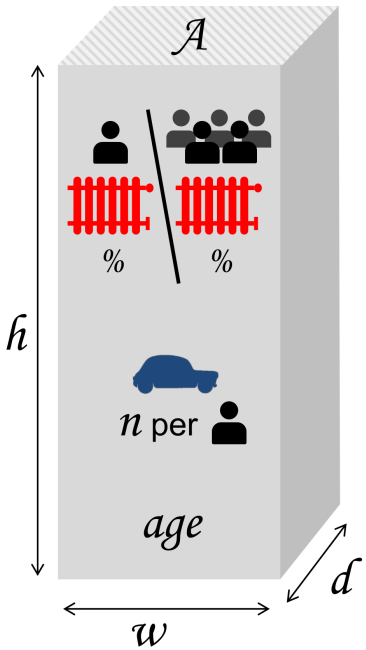


Figure: Ground coverage ratio map.

Which parameters?

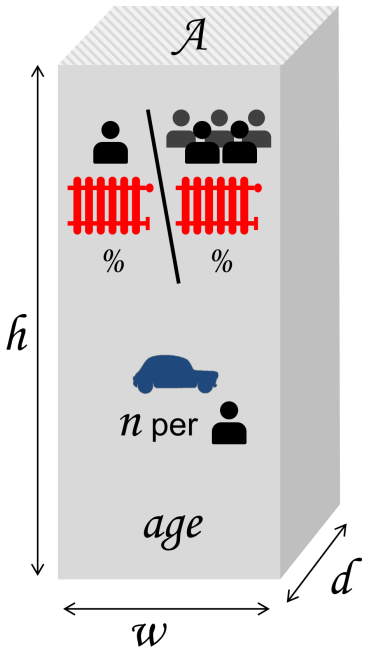
Data sources: IGN, INSEE, Bordeaux Métropole, Copernicus, DGFIP (age)...

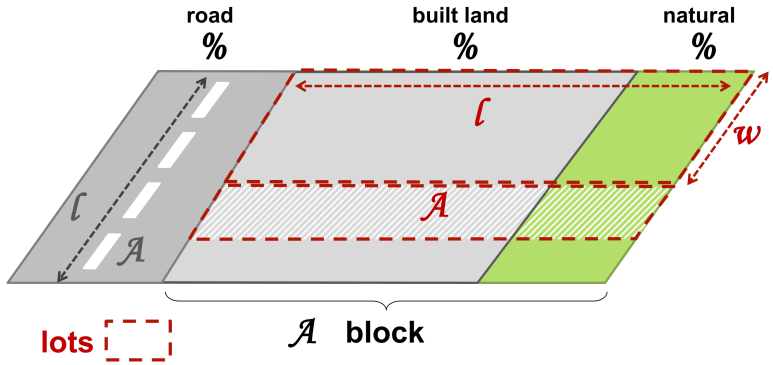


According to Bonhomme, 2013, Adam et al., 2005, Cerezo et al., 2017, Gonçalves et al., 2017 et Delmastro et al., 2016 among others.

Which parameters?

Data sources: IGN, INSEE, Bordeaux Métropole, Copernicus, DGFIP (age)...





According to Bonhomme, 2013, Adam et al., 2005, Cerezo et al., 2017, Gonçalves et al., 2017 et Delmastro et al., 2016 among others.

Data presentation

Correlations examples

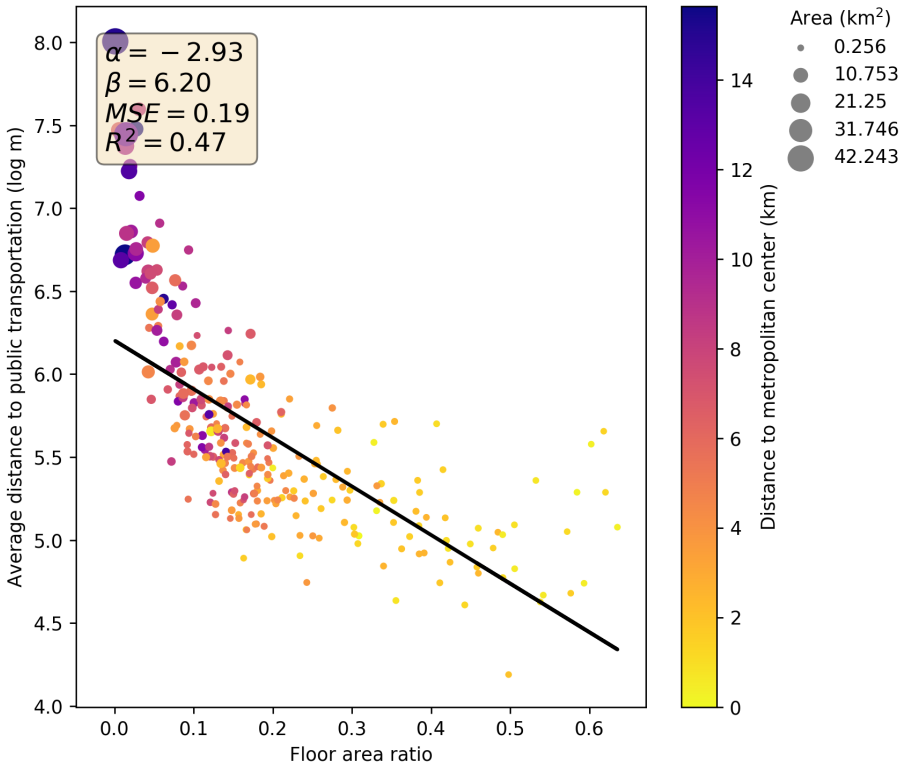


Figure:
Correlation graph between *Floor area ratio* compared to *Average distance to public transportation (log m)*

Correlations examples

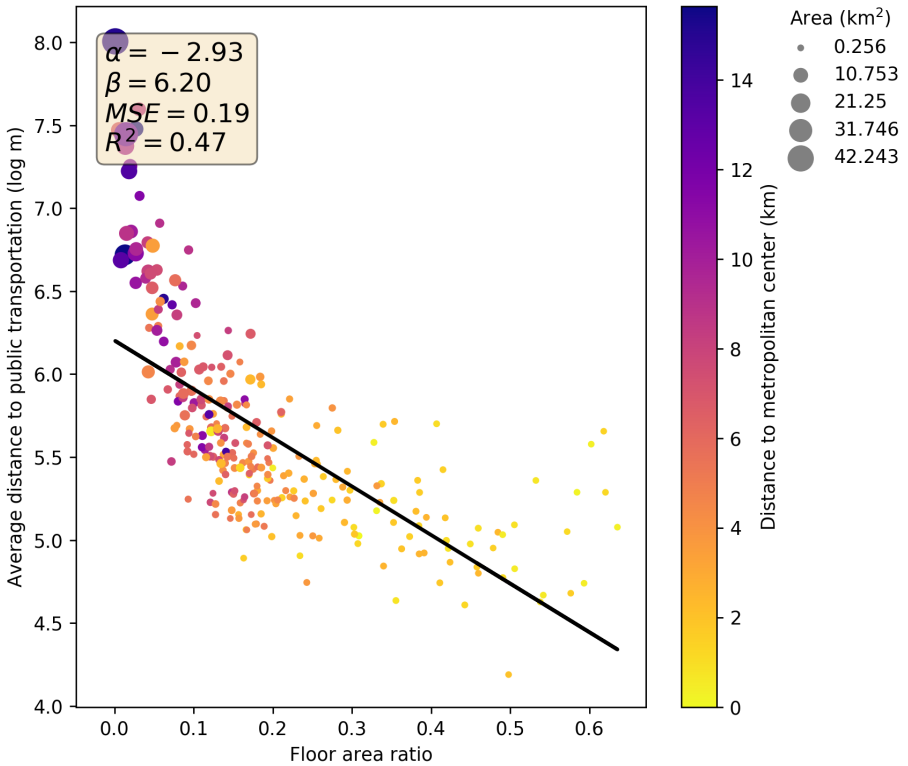


Figure:
Correlation graph between *Floor area ratio* compared to *Average distance to public transportation (log m)*

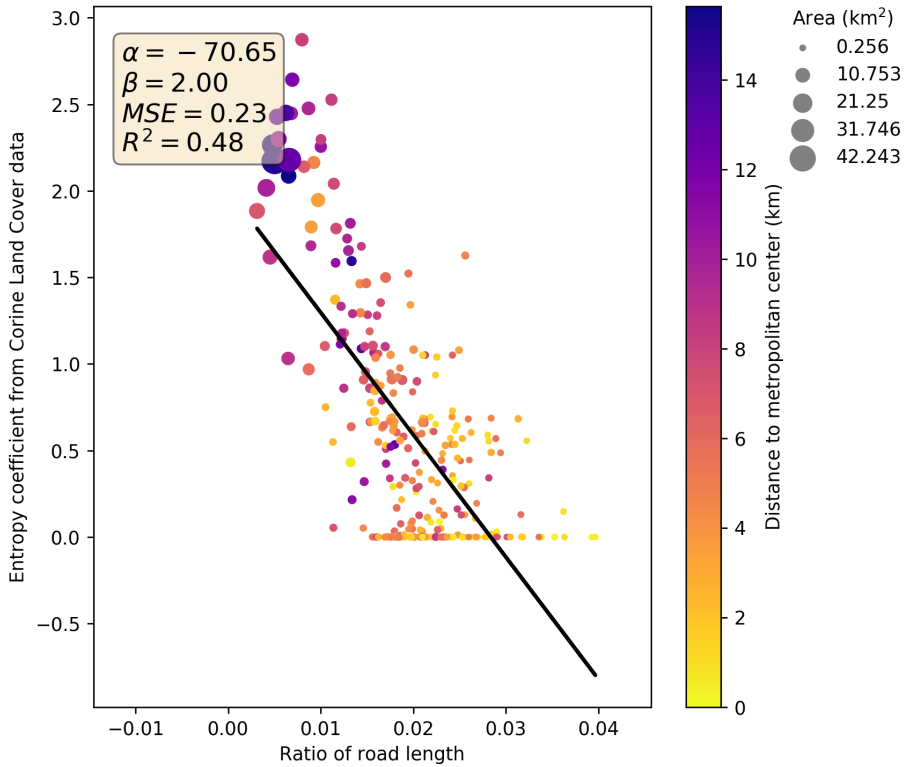


Figure:
Correlation graph between *Ratio of road length* compared to *Entropy coefficient from Corine Land Cover data*

Application of the Principal Component Analysis

From **35 parameters** (*buildings*) or **23** (*urban shapes*) to most significant axes

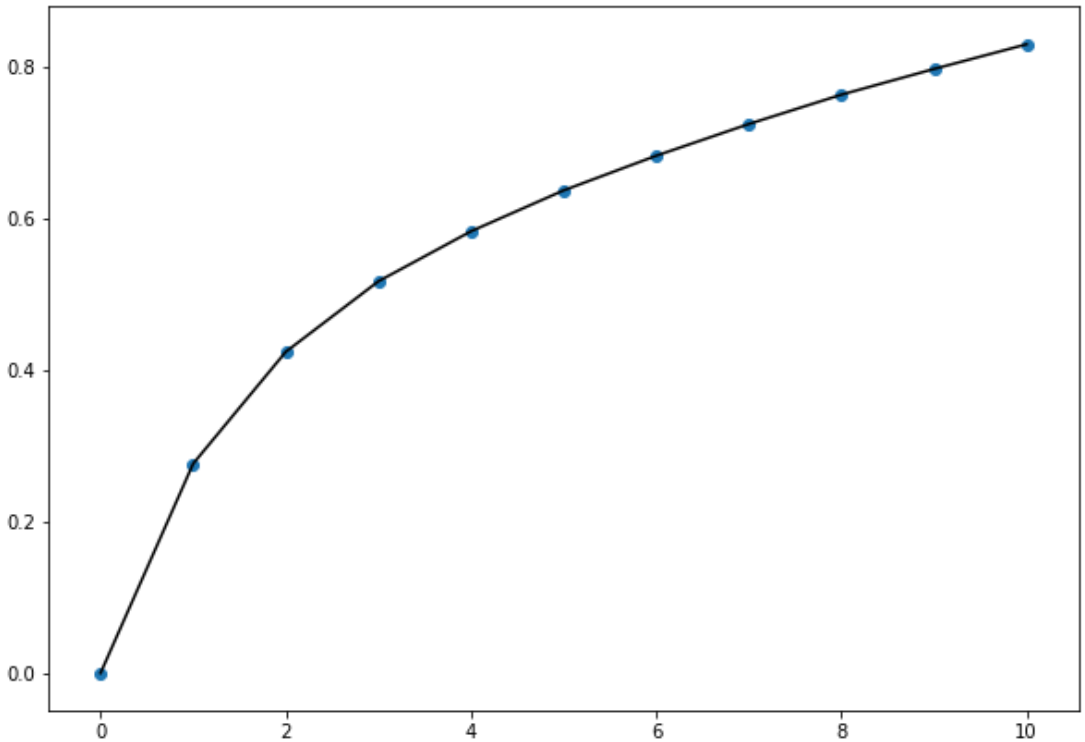


Figure: Cumulated variance ratio for all PCA axes.

First axis for buildings

Parameter	Eigen value
Density of dwellings per m^2 of floor	0.294
Density of population per m^2 of floor	0.286
% of households with a parking space	-0.282
Number of buildings	-0.280
Average height of buildings	0.276
% of households with two cars or more	-0.275
% of collective central heating among the main residences	0.274
% of households with one car or more	-0.271
% of individual central heating among the main residences	-0.252
Average length of buildings	0.231
Buildings' width average	0.222
Average form factor of buildings	0.221
Volume occupied by the buildings present in the parcel	0.217

First axis for buildings

Parameter	Eigen value
Density of dwellings per m^2 of floor	0.294
Density of population per m^2 of floor	0.286
% of households with a parking space	-0.282
Number of buildings	-0.280
Average height of buildings	0.276
% of households with two cars or more	-0.275
% of collective central heating among the main residences	0.274
% of households with one car or more	-0.271
% of individual central heating among the main residences	-0.252
Average length of buildings	0.231
Buildings' width average	0.222
Average form factor of buildings	0.221
Volume occupied by the buildings present in the parcel	0.217

\+

- housing density
- population density
- height, width, length, shape form, volume
- collective heating

\−

- garage
- number of buildings
- cars
- individual heating

Archetypes creation results

Results on urban shapes I

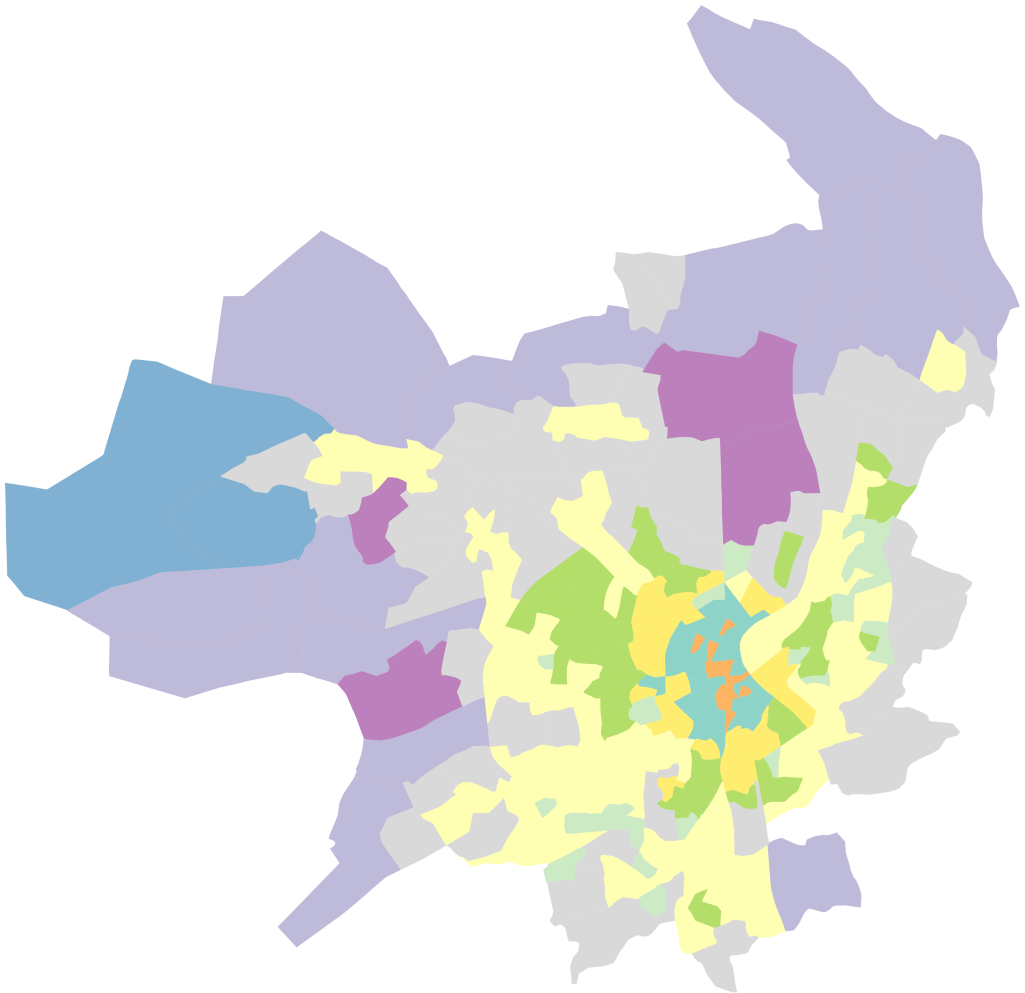


Figure: Clusters for urban shapes archetypes.

Results on urban shapes II

Name	Description	Average blocks area (log m^2)	Ratio of vegetalized area	Average distance to public transportation (log m)	Floor area ratio	Ratio of road surface	...
Dense close city 1	Bordeaux's highly dense center (small lots, few green spaces, few variety of land cover, high CES, much road surface)	8.83395 (0.29333)	0.07602 (0.05215)	5.04190 (0.26440)	0.39233 (0.08076)	0.11276 (0.01698)	...
Suburbs / smaller towns 1	Similar lots of suburbs/smaller towns (average CES, small lots, few road surface)	9.44613 (0.29550)	0.21096 (0.09864)	5.63650 (0.25539)	0.13672 (0.03198)	0.07890 (0.01465)	...
Rural area 1	Rural areas (few schools, larger portion of green areas, diverse land-use, very little built or road surface, very large lots)	11.15463 (0.49406)	0.51442 (0.23598)	6.92366 (0.33461)	0.02419 (0.01181)	0.02604 (0.00756)	...
Saint-Médard-en-Jalles	Saint-Médard-en-Jalles (specific part with small amount of buildings, schools, road, huge lots very diversely used)	11.72924 (0.23985)	0.90987 (0.04691)	7.87007 (0.19080)	0.00357 (0.00377)	0.01925 (0.00342)	...
Historical centre	**Highly dense historical center (small lots, few green spaces, high CES, much road surface, many schools/health infrastructure)**	**8.28203 (0.30990)**	**0.03221 (0.02826)**	**5.06009 (0.39676)**	**0.56311 (0.05299)**	**0.14258 (0.01892)**	**...**
Dense close city 2	Bordeaux's dense diverse further surroundings (variety of land-use, fewer public equipments, smaller CES, rather dense and small lots)	9.25435 (0.27048)	0.10974 (0.04940)	5.40891 (0.24011)	0.19502 (0.03861)	0.08626 (0.01635)	...
Suburbs / smaller towns 2	Similar lots of peri-urban small towns/suburbs (average/small CES, few equipments, normal/large lots, few road surface)	9.97472 (0.36878)	0.32657 (0.14023)	6.13602 (0.32896)	0.08354 (0.03100)	0.05366 (0.01116)	...
Activity zone 1	**Activity zones around the metropolitan territory**	**11.11342 (1.05265)**	**0.31171 (0.27149)**	**7.22860 (0.37593)**	**0.02914 (0.01757)**	**0.03147 (0.01531)**	**...**
Activity zone 2	Very similar neighborhood of dense fabric built or industrial/activity zone (few housing, high ratio of roads, average/large lots)	8.79446 (0.24010)	0.16334 (0.07863)	5.40508 (0.34289)	0.15497 (0.04279)	0.11542 (0.01747)	...
Dense close city 3	Bordeaux's denser diverse surroundings (variety of land-use, fewer public equipments, high CES, rather dense and small lots)	9.10579 (0.22896)	0.10384 (0.07679)	5.31922 (0.25042)	0.28212 (0.06293)	0.09553 (0.01618)	...

...

Results on buildings I

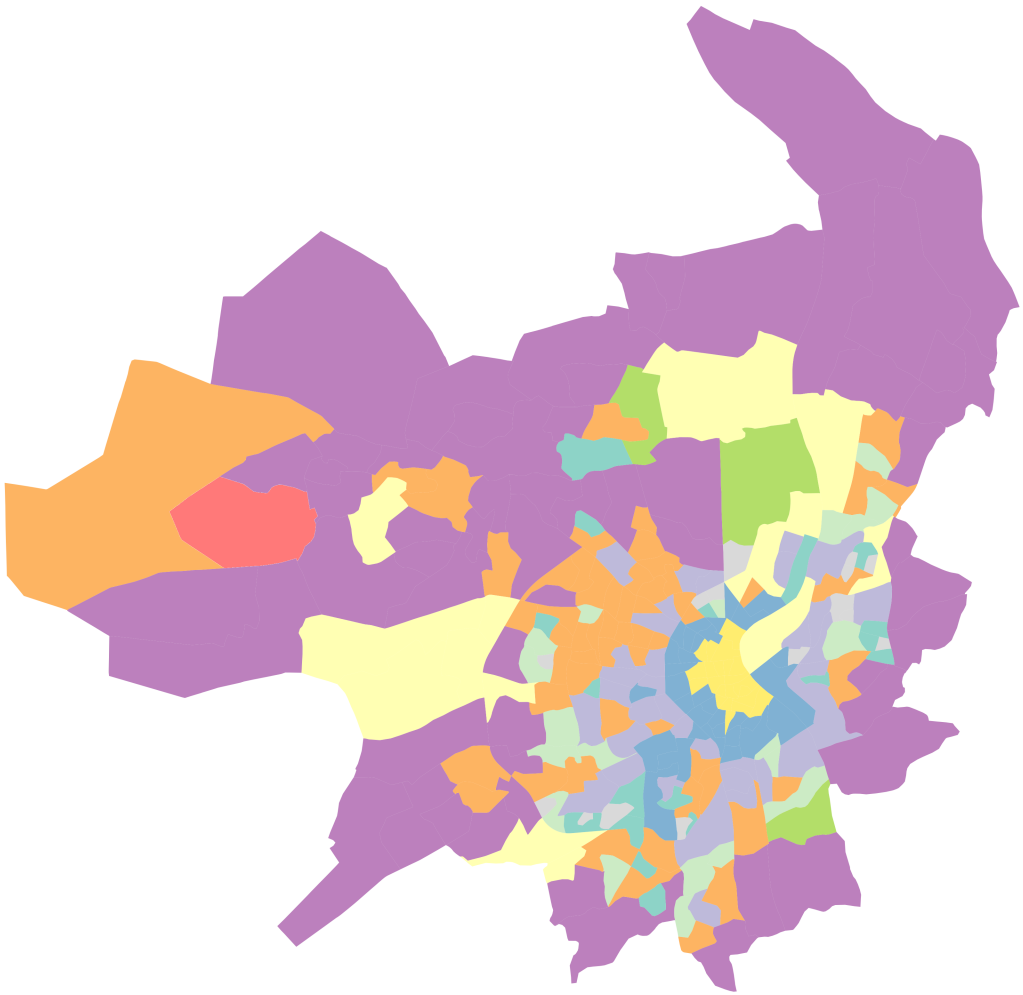


Figure: Clusters for buildings archetypes.

Results on buildings II

Name	Description	Number of jobs per m^2 for company buildings only	Number of average floors	Density of population per m^2 of floor (population / CES / area)	Proportion of buildings built before 1915	% of households with one car or more	...
Large activity buildings	Large activity buildings (since 80-90, small population, large number of employees)	0.03416 (0.01867)	2.47714 (0.51356)	0.00083 (0.00070)	0.07677 (0.13238)	0.94891 (0.06293)	...
Medium density zones	Medium density zones ("grands ensembles" and small houses, no jobs, post WWII-80', collective heating, one car)	0.01486 (0.01396)	2.46905 (0.69367)	0.03441 (0.01189)	0.01968 (0.02516)	0.75182 (0.10446)	...
New periurban suburbs housing	**New periurban suburbs housing (many houses, very small buildings, cars, since 2006, no jobs, not dense)**	**0.01531 (0.00950)**	**1.94982 (0.17071)**	**0.01555 (0.00630)**	**0.06552 (0.05051)**	**0.90019 (0.05051)**	**...**
Activity zones with offices	Activity zones with offices (high buildings, high number of employees, mostly recent buildings)	0.09177 (0.03405)	3.08766 (1.19544)	0.01979 (0.00951)	0.11652 (0.10490)	0.68981 (0.16494)	...
High density zones	**High density zones (towers, not much jobs, post WWII-80', collective heating, one car)**	**0.01845 (0.02601)**	**4.40631 (1.50829)**	**0.06725 (0.01885)**	**0.00698 (0.01057)**	**0.64375 (0.11914)**	**...**
Close suburbs	Close suburbs (small housing, old buildings, individual or two stores, no car, average jobs)	0.02166 (0.01923)	3.15456 (0.53618)	0.02404 (0.00598)	0.46656 (0.15570)	0.69539 (0.06815)	...
ZA de Saint-Médard-en-Jalles	ZA de Saint-Médard-en-Jalles	0.22219 (0.00000)	2.48303 (0.00000)	0.00094 (0.00000)	0.50000 (0.00000)	0.88052 (0.00000)	...
Periurban suburbs housing	Periurban suburbs housing (many houses, very small buildings, cars, between 75-00, no jobs, not dense)	0.01125 (0.00882)	1.77626 (0.12941)	0.01502 (0.00357)	0.06326 (0.06498)	0.93829 (0.03193)	...
Old urban suburbs housing	Old urban suburbs housing (many houses, very small buildings, between 15-48, no jobs, not dense)	0.01060 (0.00804)	2.11898 (0.33796)	0.02305 (0.00565)	0.10197 (0.09758)	0.82176 (0.06818)	...
Historical center	Historical center (old, dense, high, no car, elec heating, average nb jobs)	0.03718 (0.03169)	5.22854 (1.18080)	0.03216 (0.01067)	0.72736 (0.15306)	0.46322 (0.09589)	...
...							

Conclusion

5 Conclusion

Conclusion and perspectives

Conclusion:

- We obtained archetypes...
  - *10 urban shapes*,
  - *10 buildings*,
- ...*based on* statistical data on Bordeaux Métropole...
- ...from open data sources;
- Interesting correlations;
- Realistic PCA axes;
- Coherent results (spatial distribution, can be associated with real urban shapes...).

Conclusion and perspectives

Conclusion:

- We obtained archetypes...
  - *10 urban shapes*,
  - *10 buildings*,
- ...*based on* statistical data on Bordeaux Métropole...
- ...from open data sources;
- Interesting correlations;
- Realistic PCA axes;
- Coherent results (spatial distribution, can be associated with real urban shapes...).

Perspectives:

- Apply to other French cities;
- Compare with archetype databases (TABULA/EPISCOPE);
- Take more parameters into account?

MINES ParisTech | PSL

Thank you for your attention!

Matthieu Denoux, *PhD Student*

`matthieu.denoux@mines-paristech.fr`

supervised by *Nadia Maïzi*

References I

Adam, Claire and Muriel Moulinié (June 5, 2005). "Appréhender la densité 2 : Les indicateurs de densité". In: *Note rapide sur l'occupation du sol* 383, p. 4. URL: https://www.iau-idf.fr/savoir-faire/nos-travaux/edition/apprehender-la-densite-2.html (visited on 10/12/2018) (cit. on pp. 14, 15).

Bonhomme, Marion (Dec. 11, 2013). "Contribution à la génération de bases de données multi-scalaires et évolutives pour une approche pluridisciplinaire de l'énergétique urbaine". Thèse de doctorat. Université de Toulouse. 557 pp. URL: http://lra.toulouse.archi.fr/lra/productions/theses-et-hdr-soutenues/theses-soutenues/Marion_Bonhomme/memoire_v10 (visited on 10/11/2017) (cit. on pp. 14, 15).

Cerezo, Carlos, Julia Sokol, Saud AlKhaled, Christoph Reinhart, Adil Al-Mumin, and Ali Hajiah (Nov. 1, 2017). "Comparison of Four Building Archetype Characterization Methods in Urban Building Energy Modeling (UBEM): A Residential Case Study in Kuwait City". In: *Energy and Buildings* 154 (Supplement C), pp. 321–334. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2017.08.029. URL: http://www.sciencedirect.com/science/article/pii/S0378778817314743 (visited on 10/12/2017) (cit. on pp. 14, 15).

Delmastro, Chiara, Guglielmina Mutani, and Laura Schranz (Sept. 2016). "The Evaluation of Buildings Energy Consumption and the Optimization of District Heating Networks: A GIS-Based Model". In: *International Journal of Energy and Environmental Engineering* 7.3, pp. 343–351. ISSN: 2008-9163, 2251-6832. DOI: 10.1007/s40095-015-0161-5. URL: http://link.springer.com/10.1007/s40095-015-0161-5 (visited on 06/18/2018) (cit. on pp. 14, 15).

Gonçalves, Jorge, Marta Castilho Gomes, Sofia Ezequiel, Francisco Moreira, and Isabel Loupa-Ramos (Apr. 1, 2017). "Differentiating Peri-Urban Areas: A Transdisciplinary Approach towards a Typology". In: *Land Use Policy* 63, pp. 331–341. ISSN: 0264-8377. DOI: 10.1016/j.landusepol.2017.01.041. URL: http://www.sciencedirect.com/science/article/pii/S0264837717300960 (visited on 01/31/2018) (cit. on pp. 14, 15).

Correlations examples

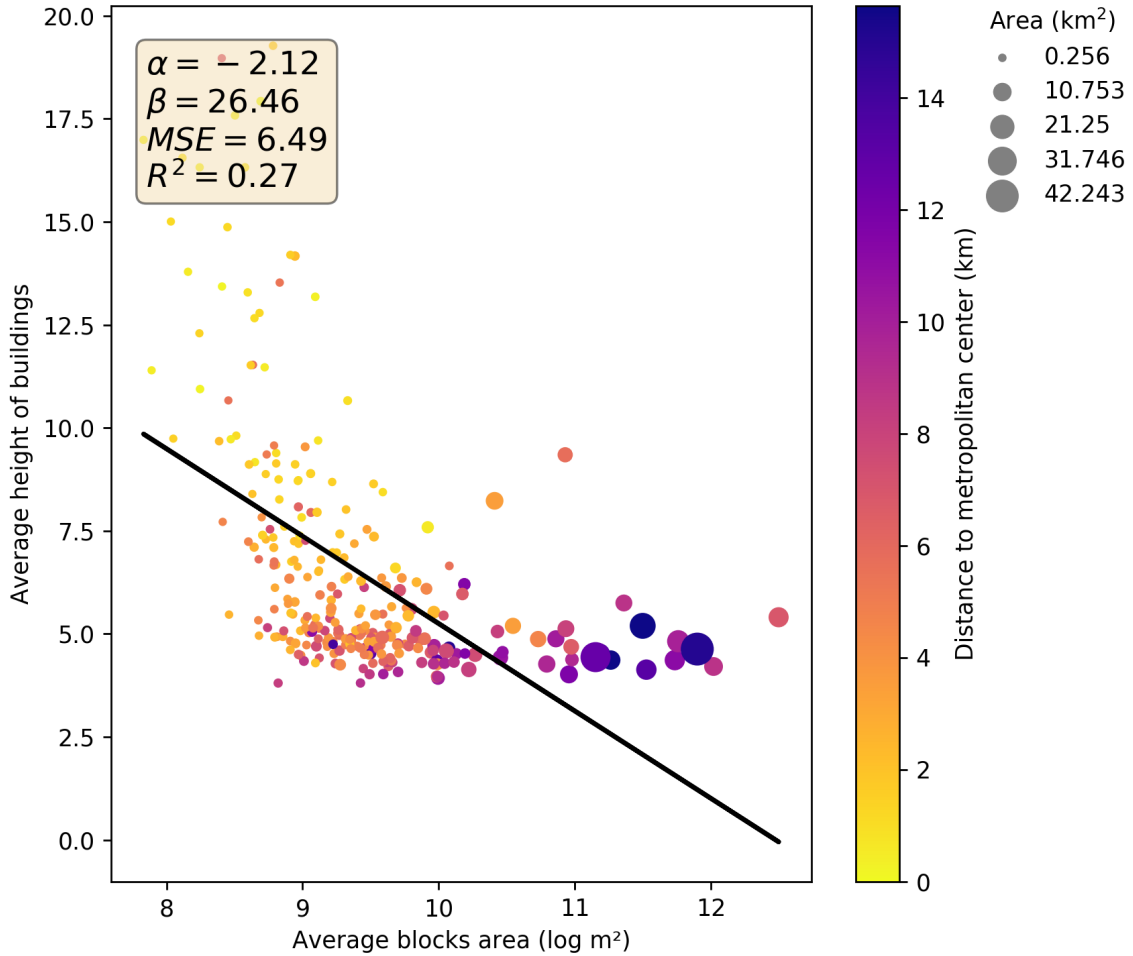


Figure: Correlation graph between *Average blocks area (log m^2)* compared to *Average height of buildings*

Correlations examples

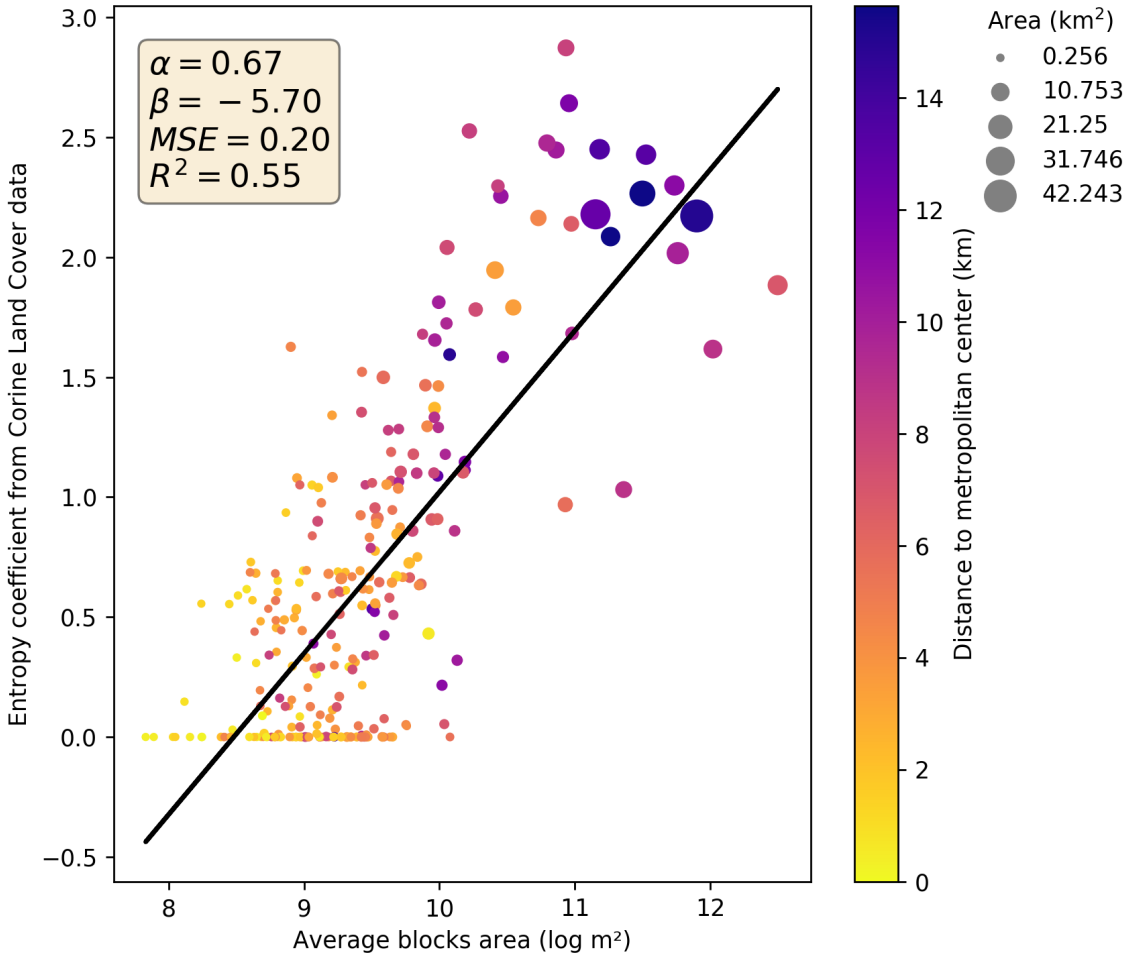


Figure: Correlation graph between *Average blocks area (log m^2)* compared to *Entropy coefficient from Corine Land Cover data*

Correlations examples

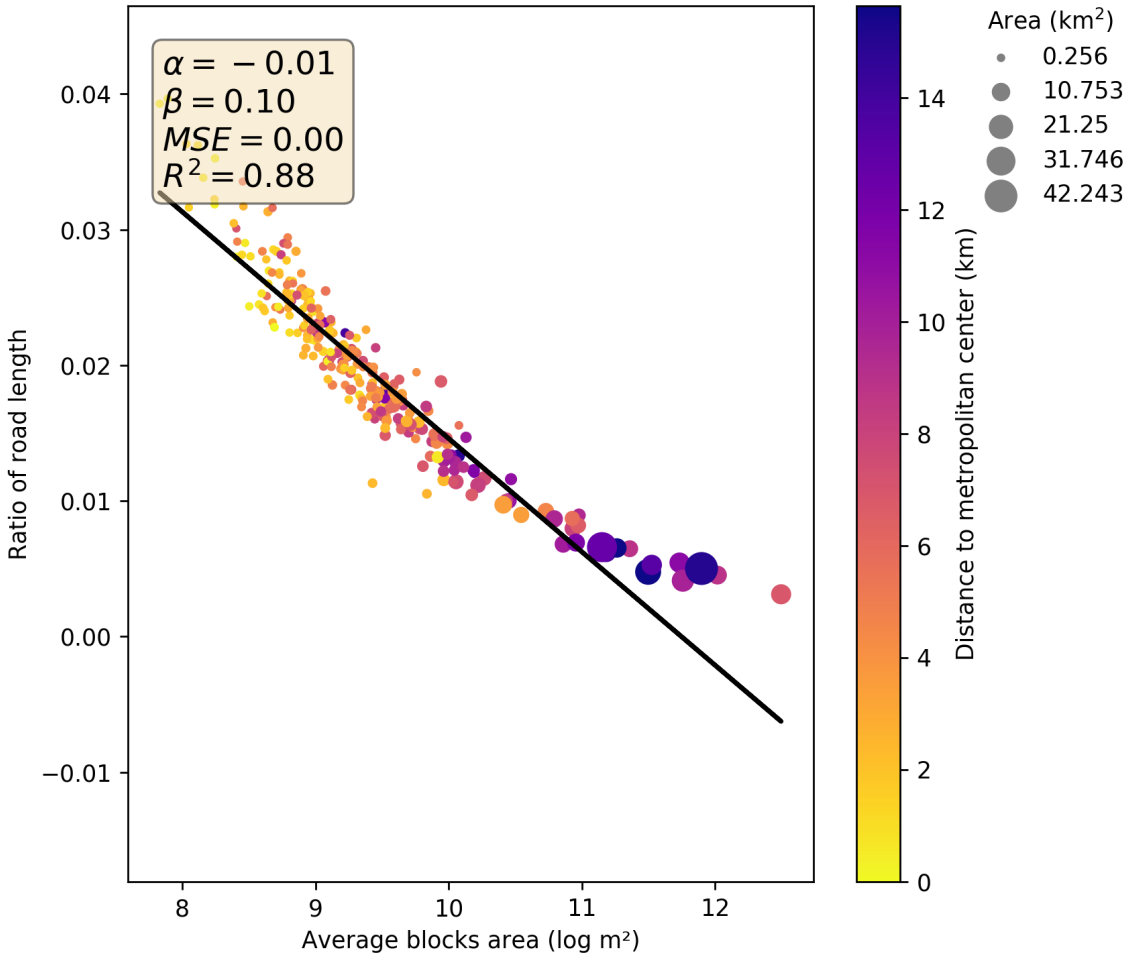


Figure: Correlation graph between *Average blocks area (log m^2)* compared to *Ratio of road length*

Correlations examples

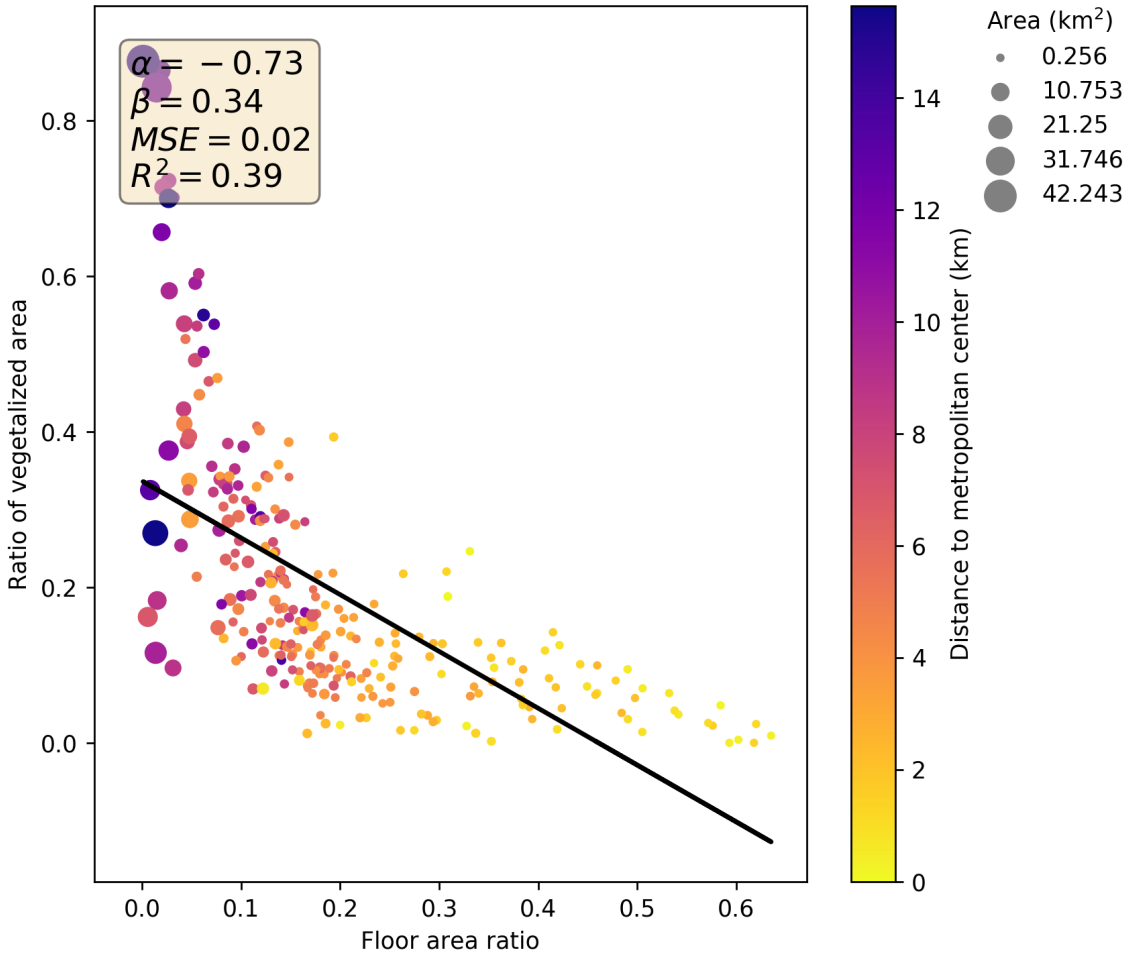


Figure: Correlation graph between *Floor area ratio* compared to *Ratio of vegetalized area*

Correlations examples

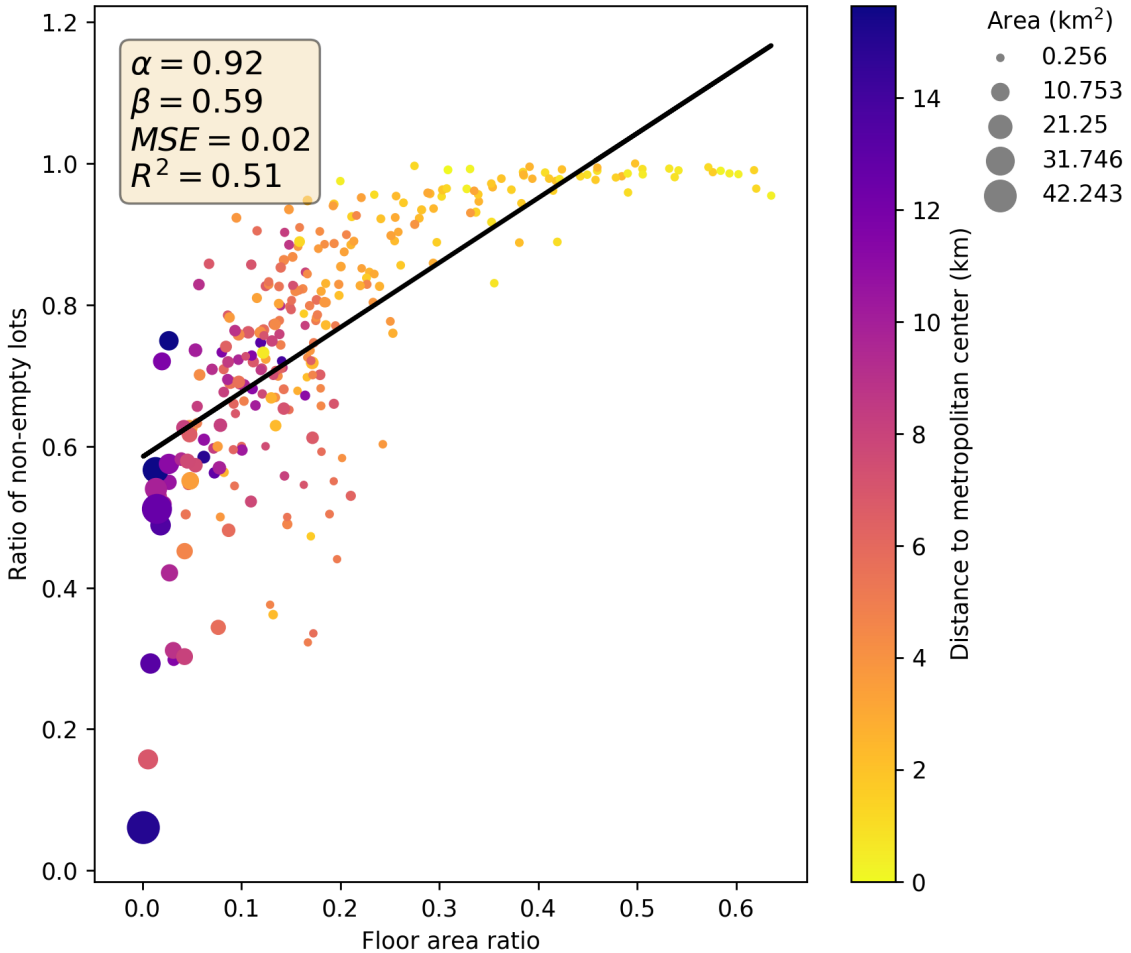


Figure: Correlation graph between *Floor area ratio* compared to *Ratio of non-empty lots*

Correlations examples

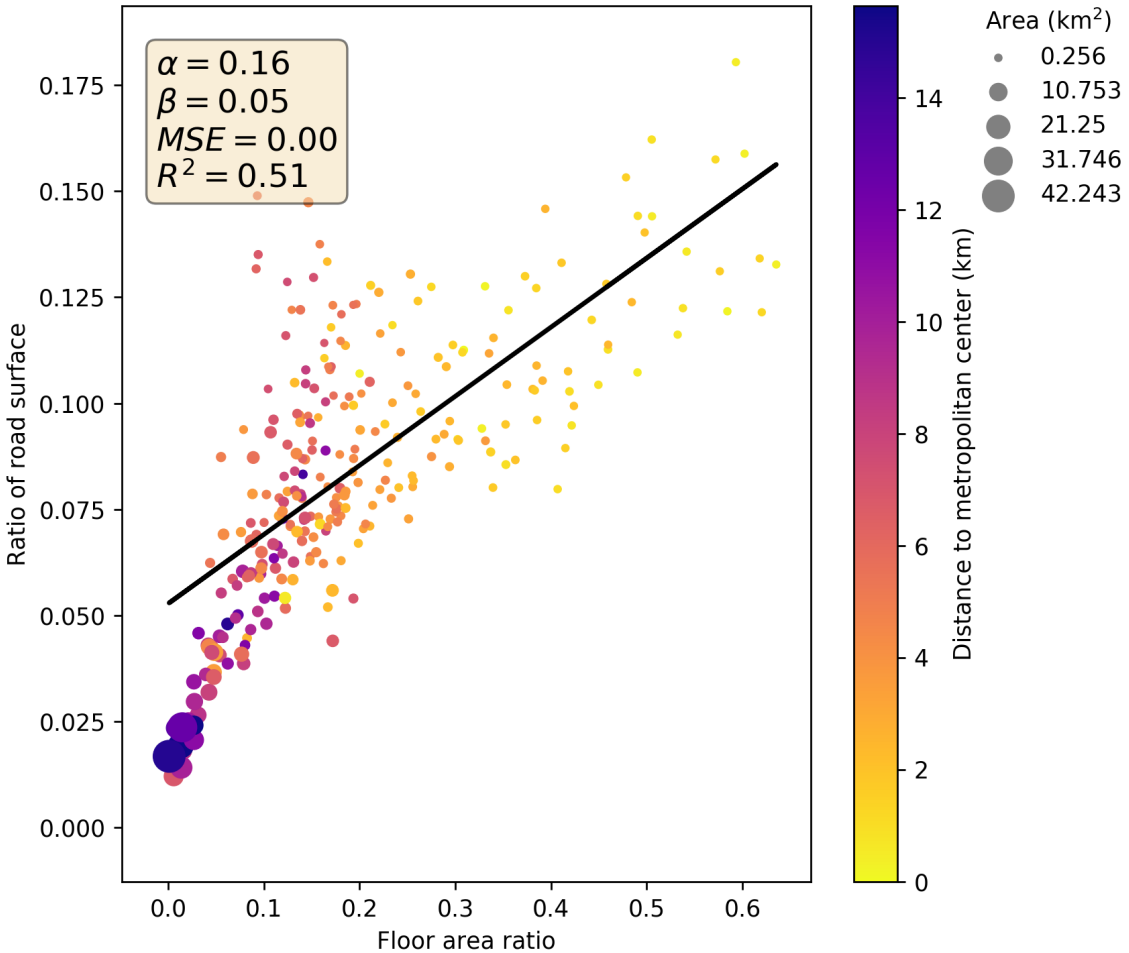


Figure: Correlation graph between *Floor area ratio* compared to *Ratio of road surface*

Correlations examples

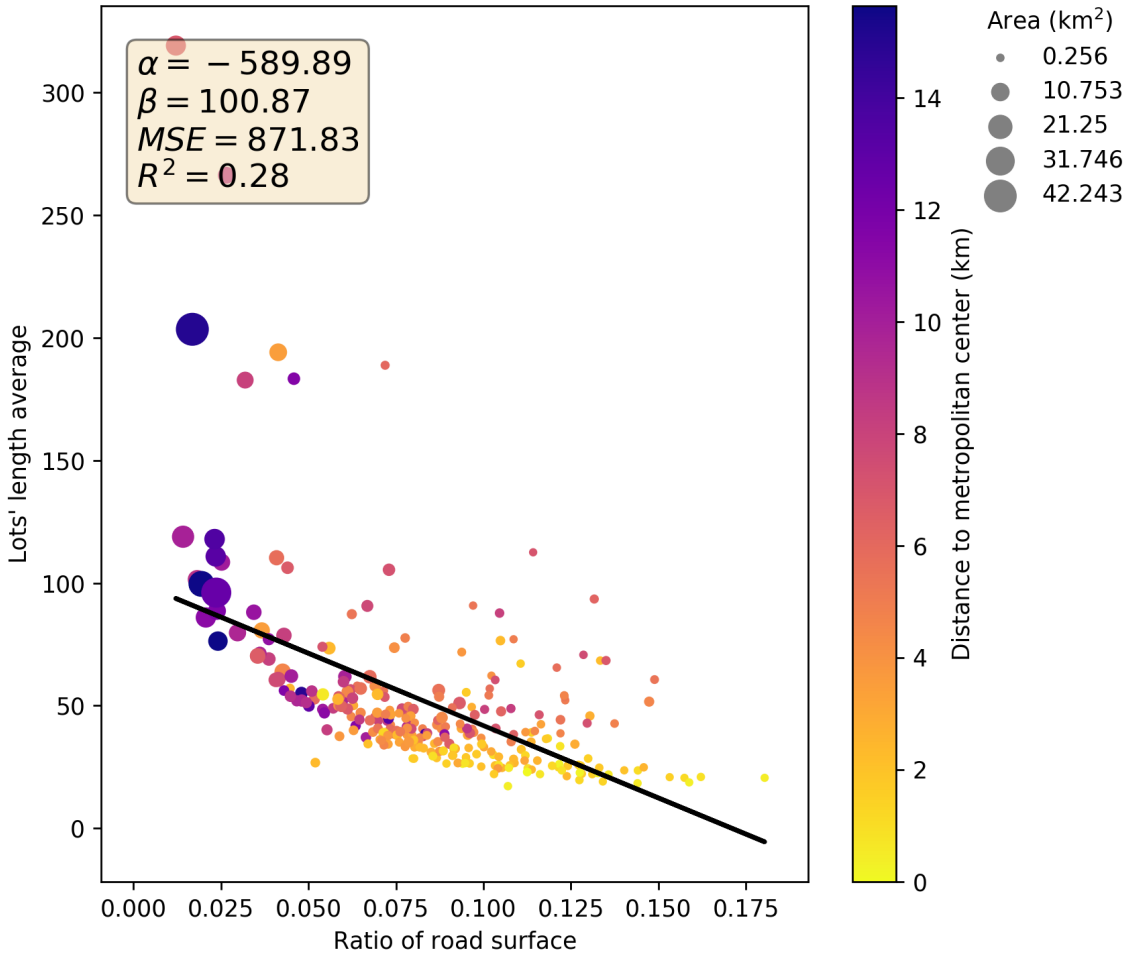


Figure: Correlation graph between *Ratio of road surface* compared to *Lots' length average*

Correlations examples

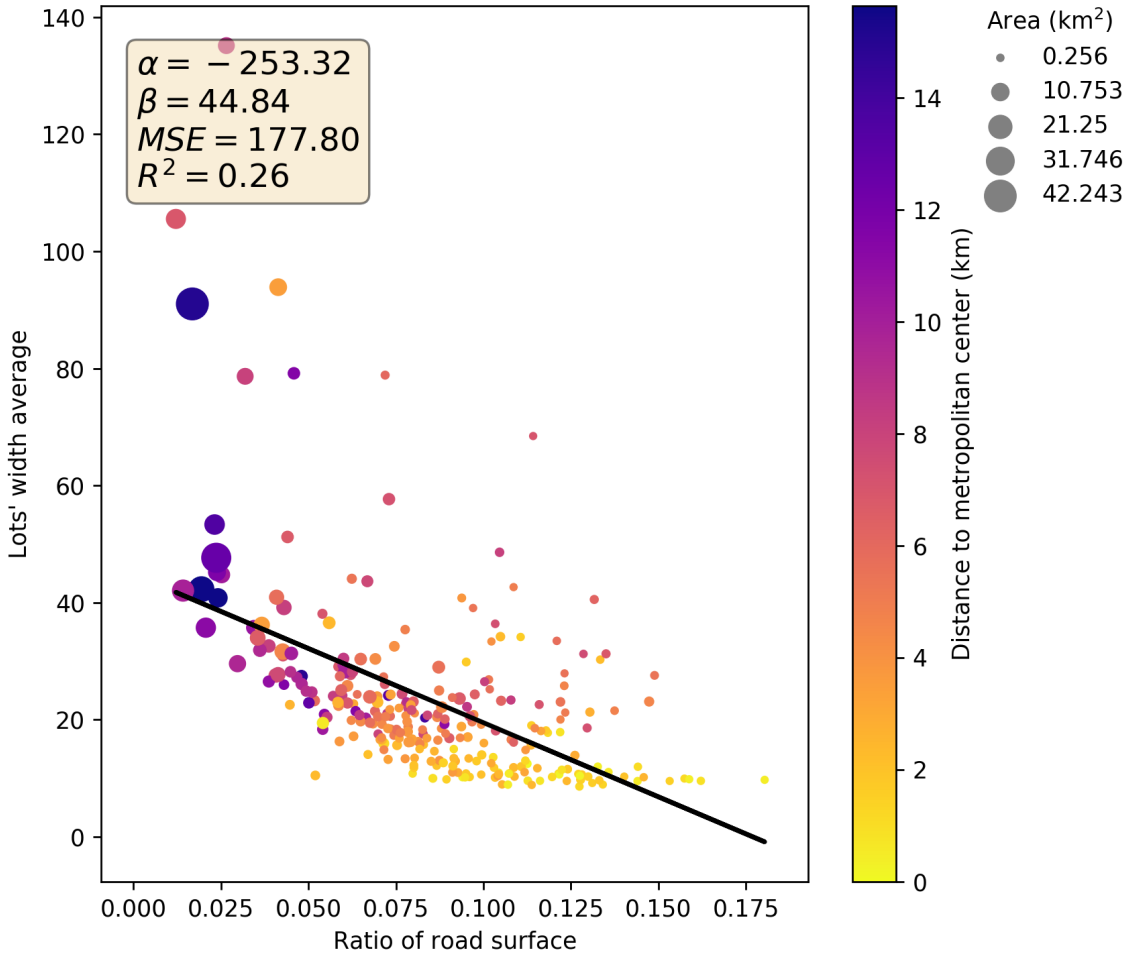


Figure: Correlation graph between *Ratio of road surface* compared to *Lots' width average*

Correlations examples

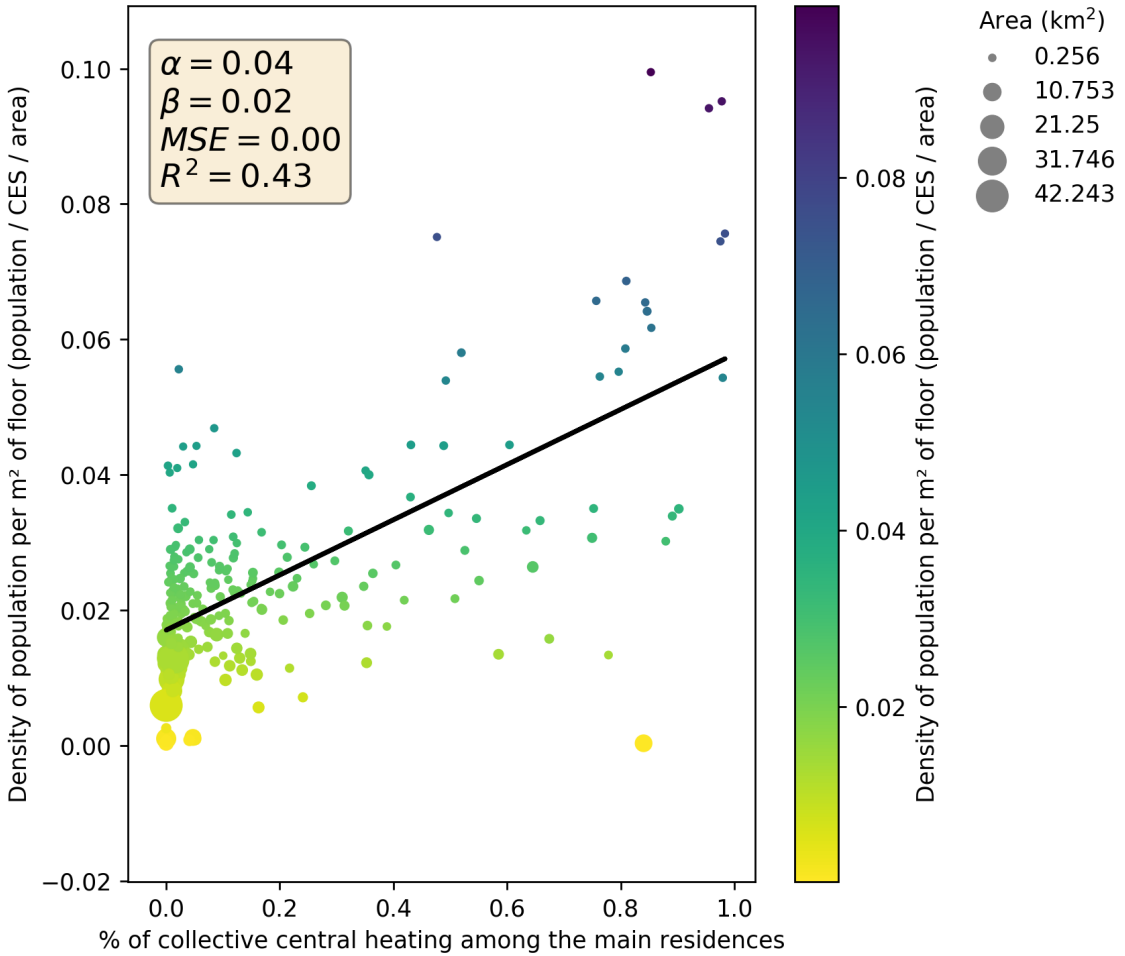


Figure: Correlation graph between *% of collective central heating among the main residences* compared to *Density of population per m^2 of floor (population / CES / area)*

Correlations examples

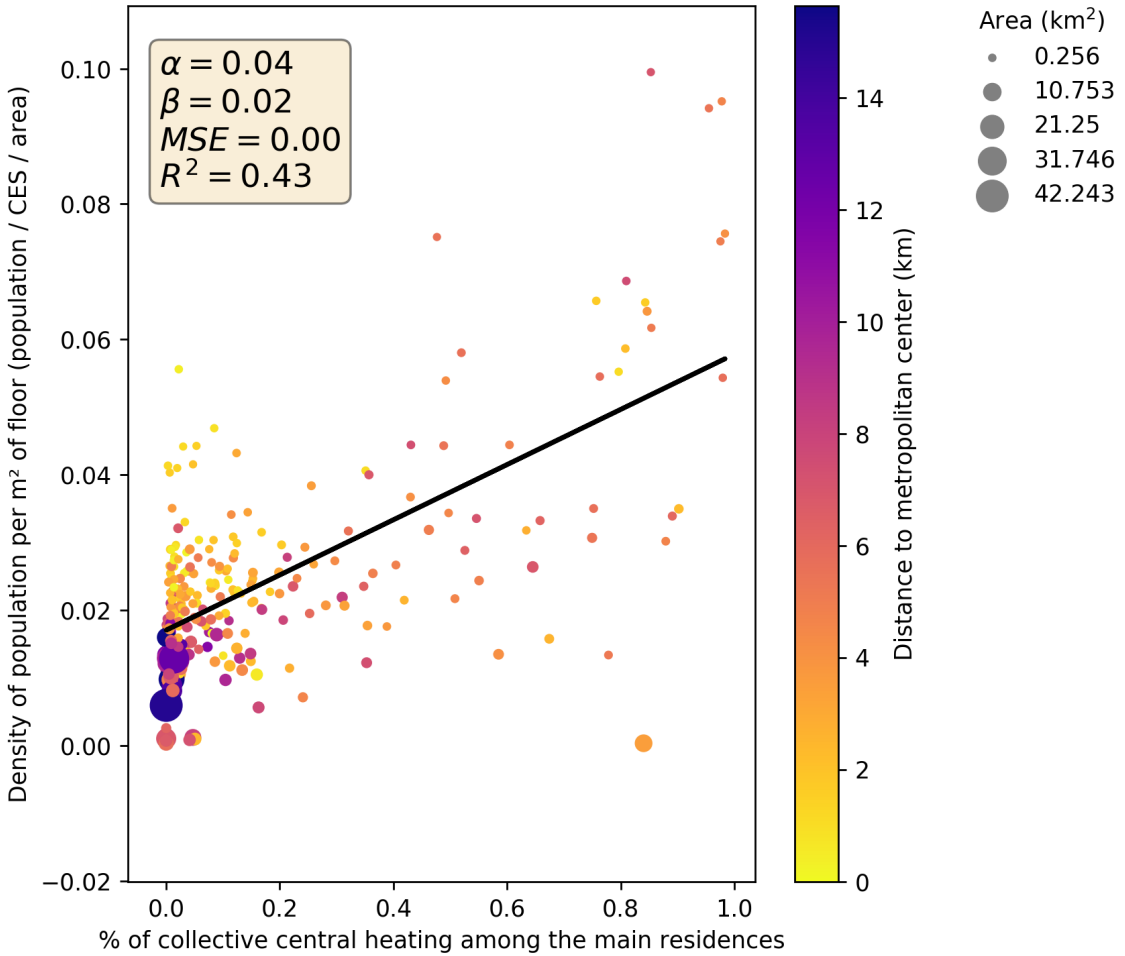


Figure: Correlation graph between *% of collective central heating among the main residences* compared to *Density of population per m^2 of floor (population / CES / area)*

Correlations examples

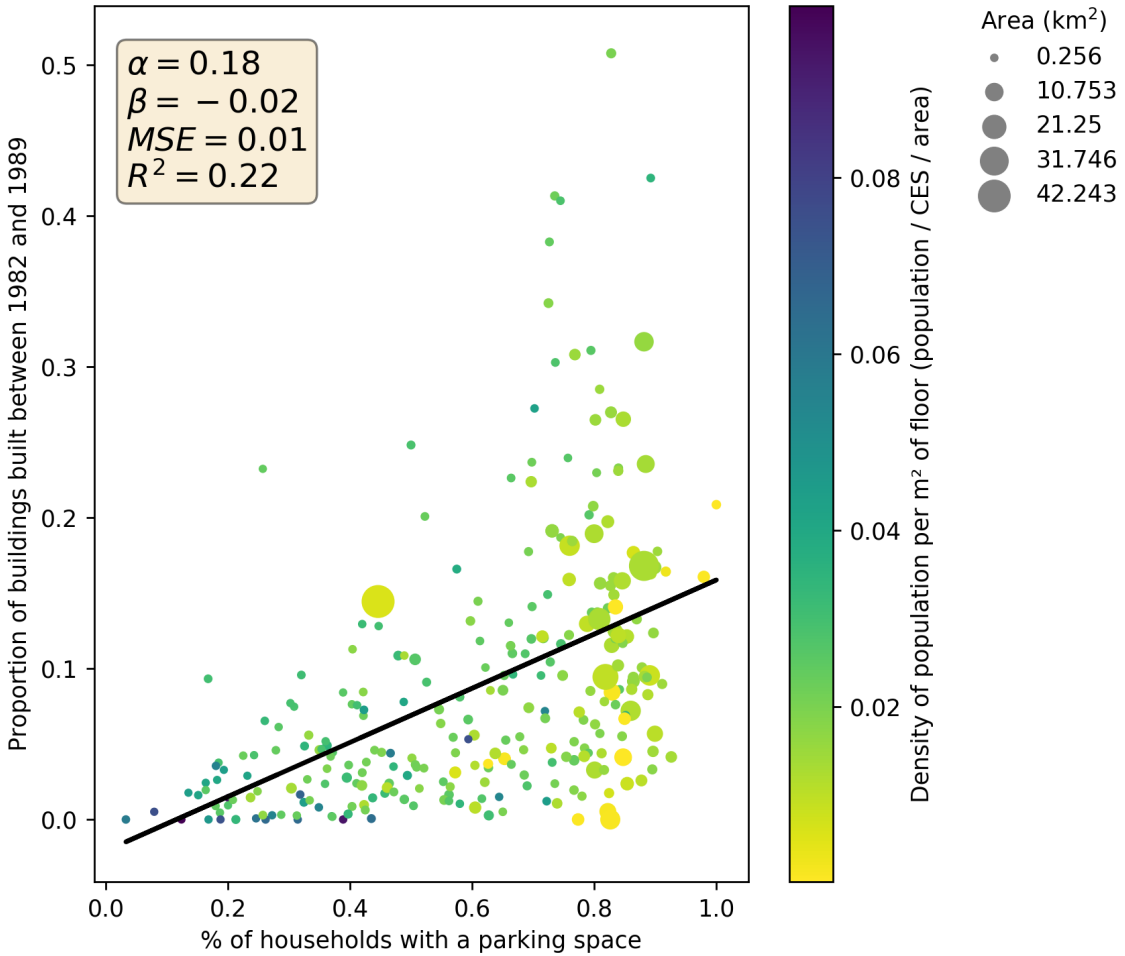


Figure: Correlation graph between *% of households with a parking space* compared to *Proportion of buildings built between 1982 and 1989*

Correlations examples

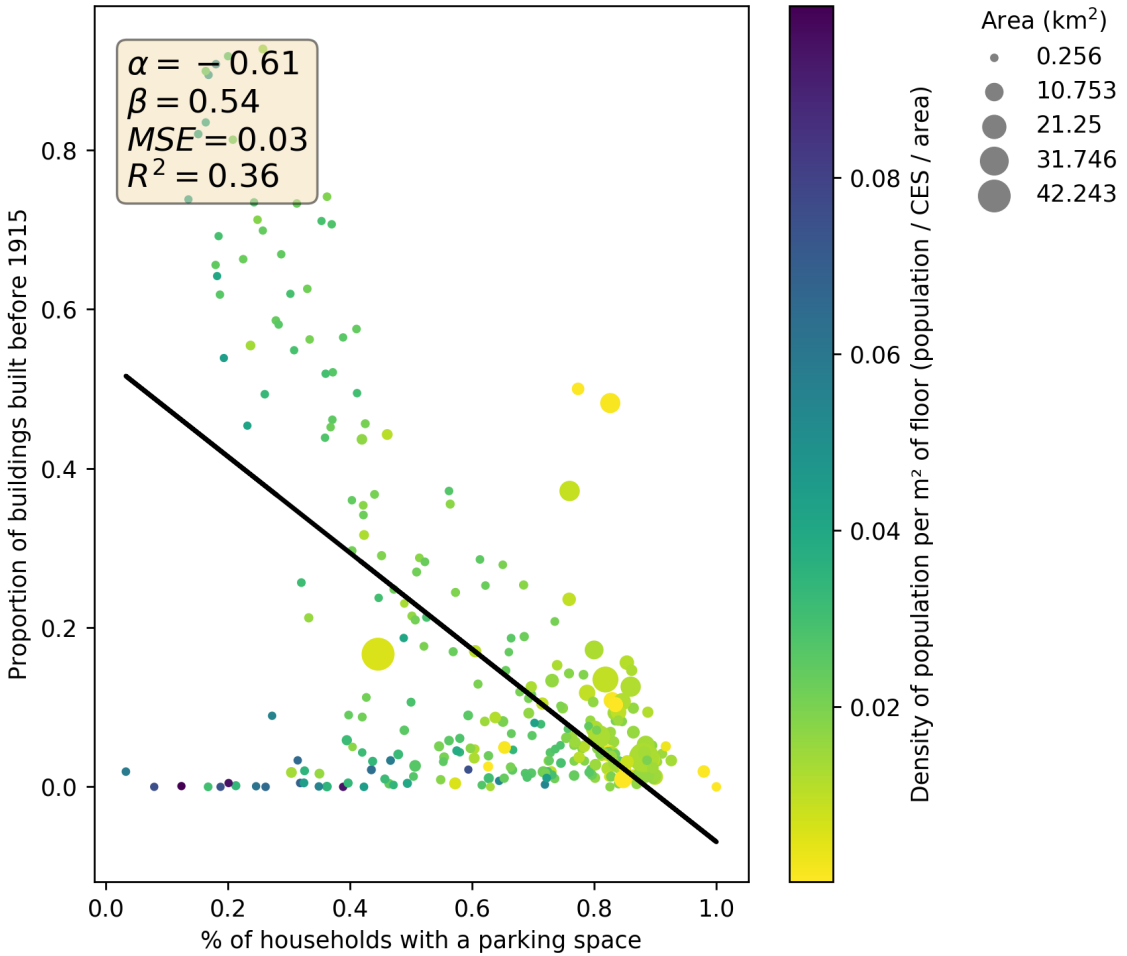


Figure: Correlation graph between *% of households with a parking space* compared to *Proportion of buildings built before 1915*

Correlations examples

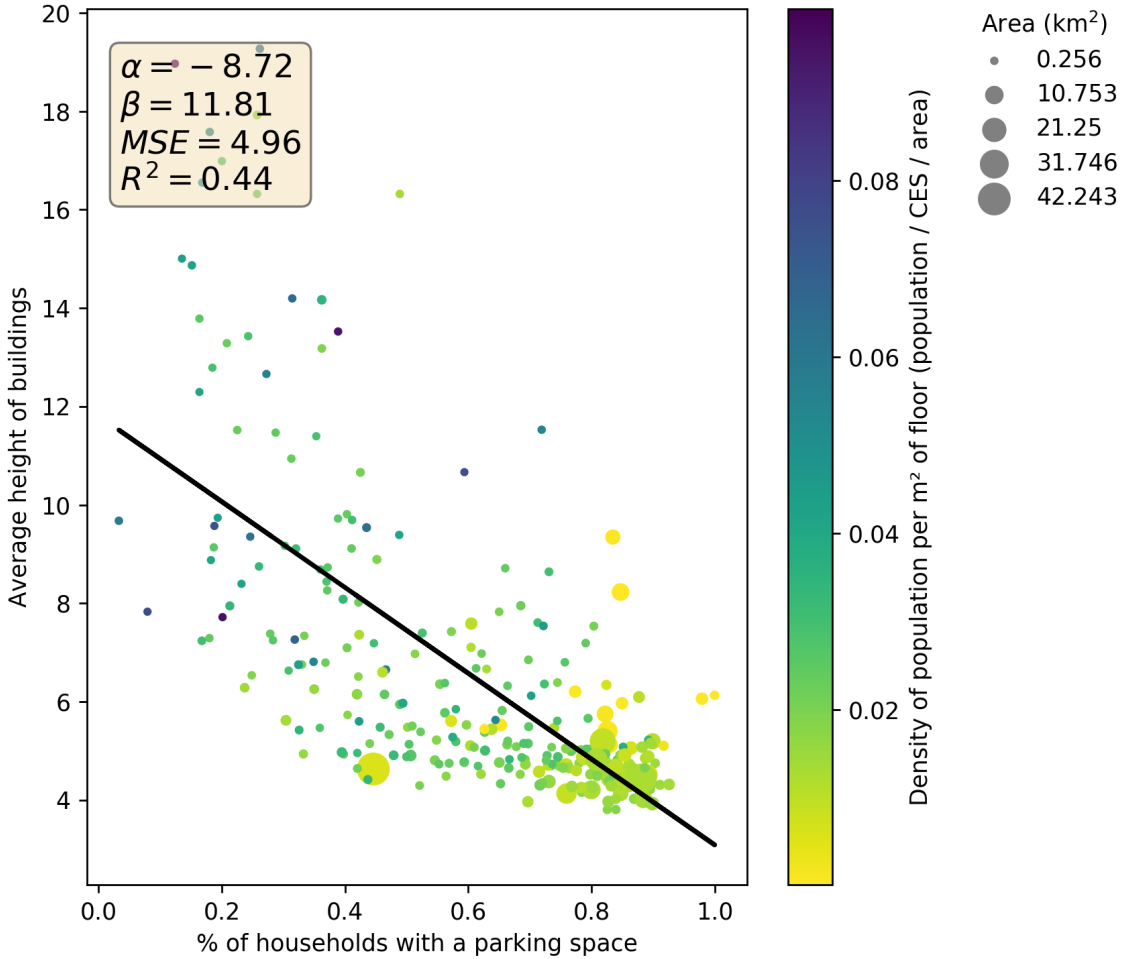


Figure: Correlation graph between *% of households with a parking space* compared to *Average height of buildings*

Correlations examples

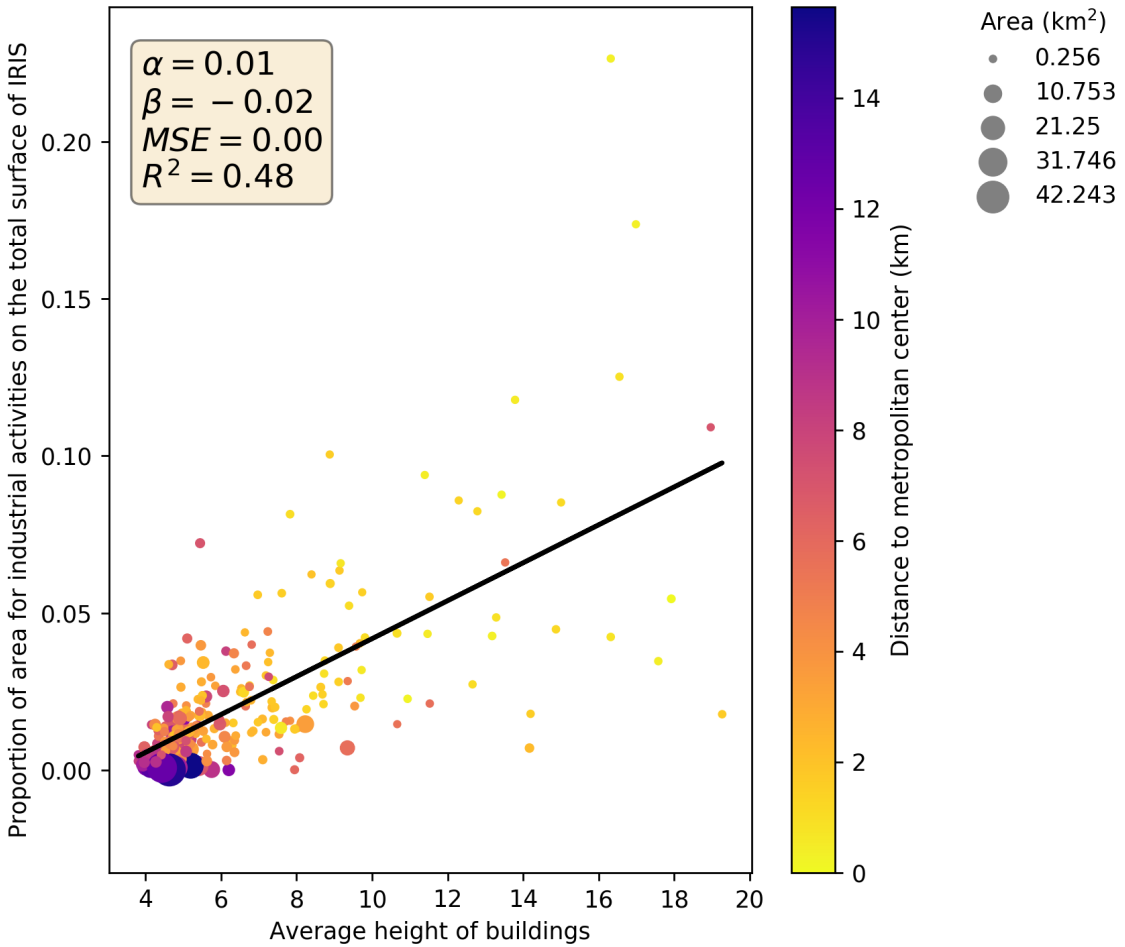


Figure: Correlation graph between *Average height of buildings* compared to *Proportion of area for industrial activities on the total surface of IRIS*

Correlations examples

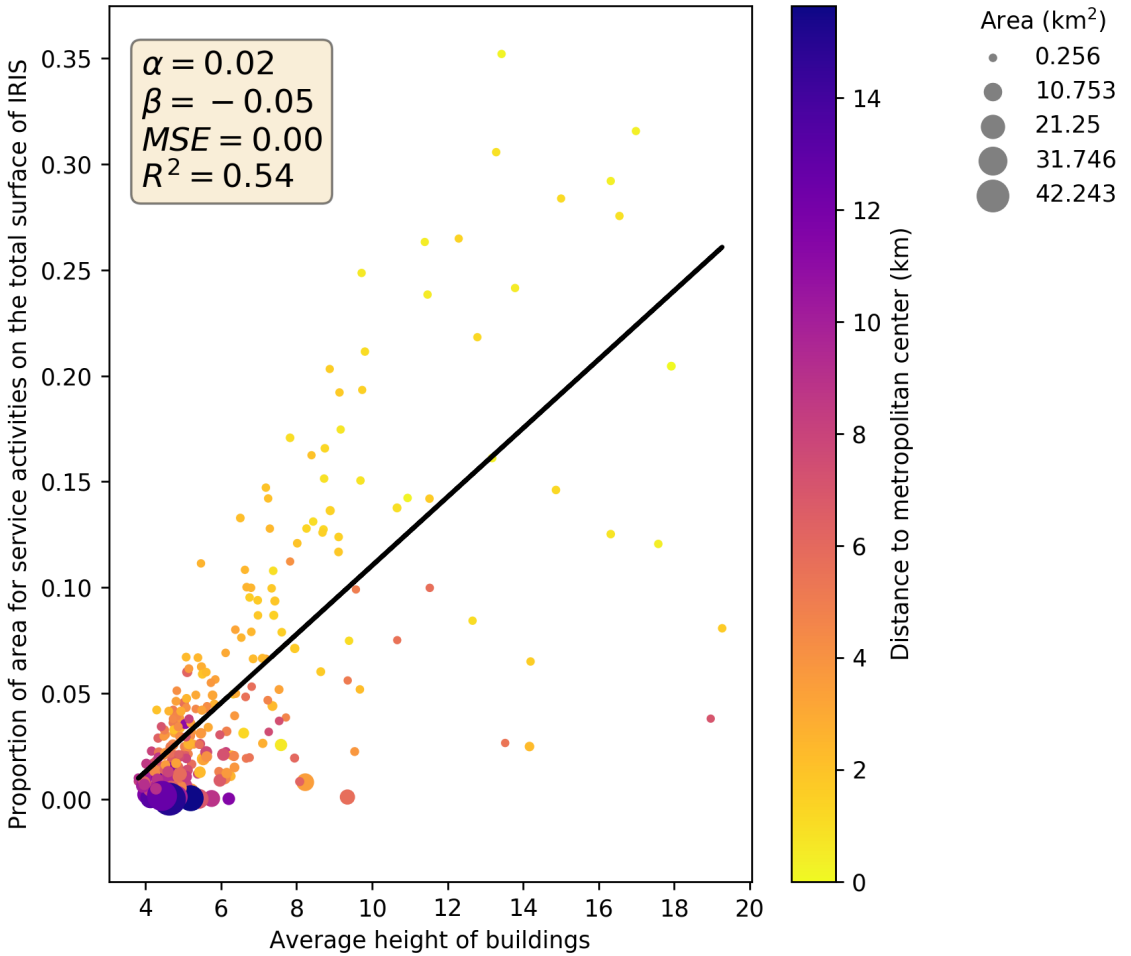


Figure: Correlation graph between *Average height of buildings* compared to *Proportion of area for service activities on the total surface of IRIS*

Correlations examples

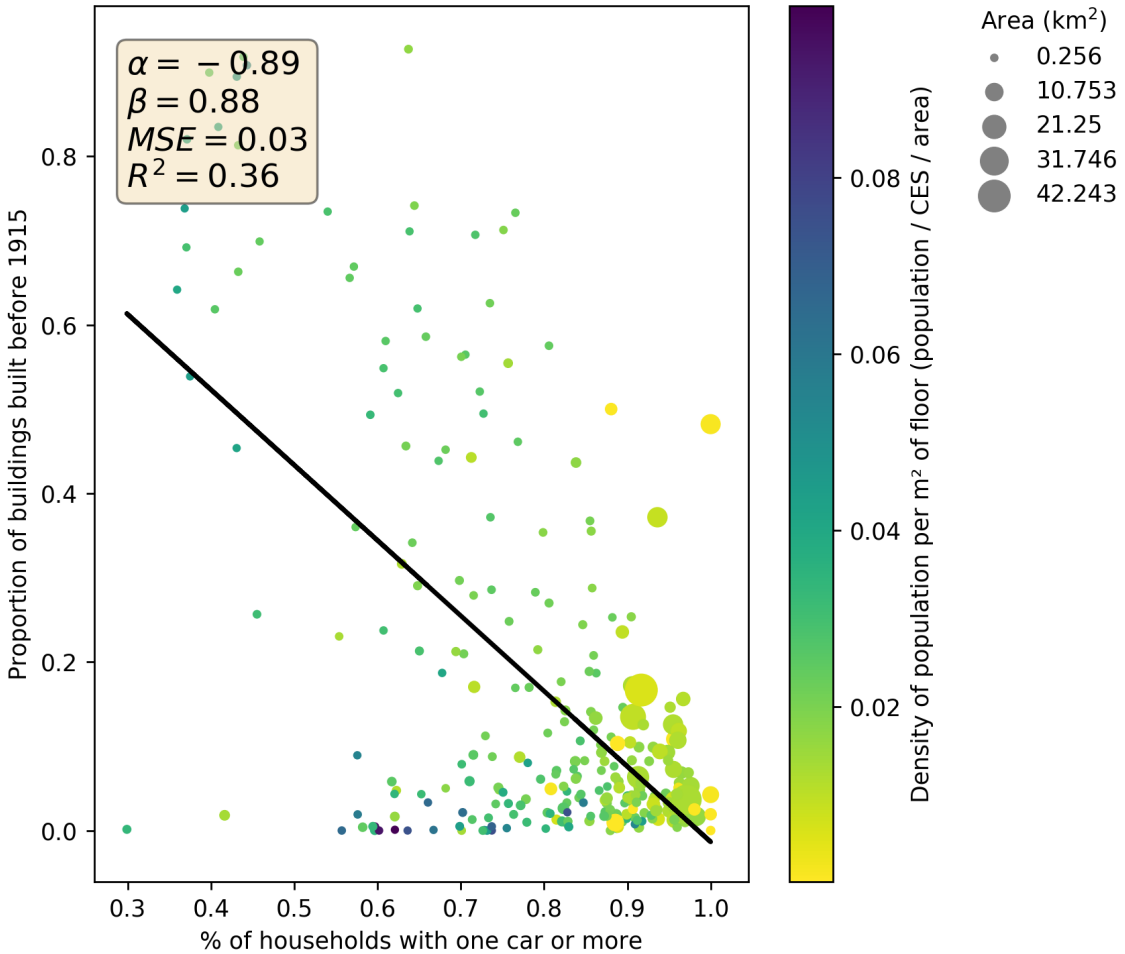


Figure: Correlation graph between *% of households with one car or more* compared to *Proportion of buildings built before 1915*