ANALYSIS OF NYC SERVICE REQUEST CALLS TO 311

Outline

- Data ETL
- MVA
  - CA, PCA, MCA
  - Clustering, Tree classification
- Conclusions

By: **Cedric Bhihe** <cedric.bhihe@gmail.com> **Santi Calvo** <s.calvo93@gmail.com>

Data from:

- NYC OpenData
- NYPD DB
- IRS
- US Census Office

All material and code available at:
https://www.github.com/Cbhihe/nyc311

1. Manhattan
2. Brooklyn
3. Queens
4. The Bronx
5. Staten Island

5 boroughs:

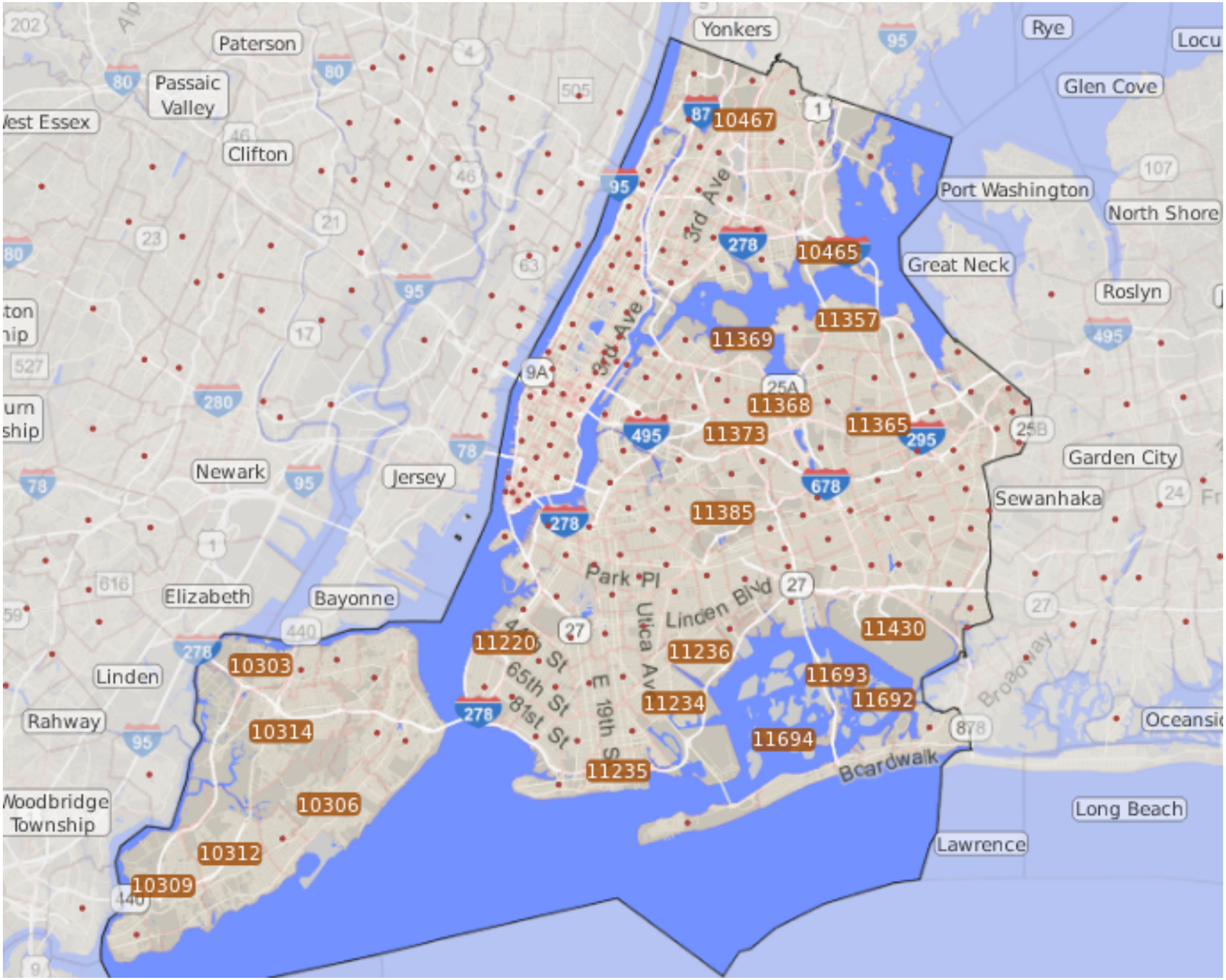
⇒ > 200 ZIPs

⇒ 8.7 M. people

⇒ 100,000 SRC to 311 /month

Objective(s):

- to explore data with MVA tools
- to extract features and descriptive information, so we may:
  - detect trends
  - optimize urban resources


Yonkers
Rye
Paterson
Passaic Valley
Glen Cove
Clifton
Port Washington
North Shore
Great Neck
Roslyn
Garden City
Sewanhaka
Newark
Jersey
Elizabeth
Bayonne
Linden
Rahway
Woodbridge Township
Long Beach
Lawrence
10467
10465
11357
11369
11368
11373
11365
11385
11430
11236
11693
11692
11234
11694
11220
11235
10303
10314
10306
10312
10309
3rd Ave
Park Pl
Linden Blvd
Utica Ave
E 19th St
65th St
81st St
Boardwalk
Broadway

Data ETL – extract

ZIP	Date	Complaint	Descriptor	Address	planeX	planeY	GPS
10463	2014-04-01	PAINT/PLASTER	CEILING	1 JACOBUS PLACE	1008544	258129	(40.875145400143914, -73.91215405
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	3631 CORLEAR AVENUE	1011190	261705	(40.884952723690205, -73.90257186
10038	2014-04-01	GENERAL	VENTILATION SYSTEM	41 JOHN STREET	982002	197700	(40.70931729970831, -74.00810843
10463	2014-04-01	FLOORING/STAIRS	FLOOR	110 TERRACE VIEW AVE	1008357	258602	(40.876444151331, -73.9128285342
10306	2014-04-01	Rodent	Condition Attracting Rodents	1934 NORTH RAILROAD	948358	144556	(40.563376021503466, -74.1291780
10007	2014-04-01	UNSANITARY CONDITION	GARBAGE/RECYCLING STORA	23 PARK PLACE	981818	199058	(40.713044638283755, -74.0087726
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	2840 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	HEAT/HOT WATER	APARTMENT ONLY	3810 BAILEY AVENUE	1010532	257794	(40.884220233300874, -73.9049668
10463	2014-04-01	HEAT/HOT WATER	APARTMENT ONLY	225 WEST 232 STREET	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	WATER LEAK	HEAVY FLOW	41 JOHN STREET	1010532	257794	(40.880351339751236, -73.9043180
10038	2014-04-01	PAINT/PLASTER	CEILING	225 WEST 232 STREET	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	1 JACOBUS PLACE	1010709	260028	(40.880351339751236, -73.9043180
10463	2014-04-01	WATER LEAK	SLOW LEAK	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	PLUMBING	BASIN/SINK	1 JACOBUS PLACE	1010709	260028	(40.880351339751236, -73.9043180
10463	2014-04-01	PAINT/PLASTER	WALL	110 TERRACE VIEW AVE	982002	197700	(40.709317299708 31, -74.0081084
10463	2014-04-01	WATER LEAK	HEAVY FLOW	110 TERRACE VIEW AVE	1010709	260028	(40.880351339751236, -73.9043180
10463	2014-04-01	UNSANITARY CONDITION	SLOW LEAK	1 JACOBUS PLACE	1008544	258129	(40.875145400143914, -73.9121540
10463	2014-04-01	HEAT/HOT WATER	MOLD	110 TERRACE VIEW AVE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	110 TERRACE VIEW AVE	1008544	258129	(40.875145400143914, -73.9121540
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	1 JACOBUS PLACE	1008544	258129	(40.875145400143914, -73.9128285342
10463	2014-04-01	HEAT/HOT WATER	APARTMENT ONLY	110 TERRACE VIEW AVE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	Rodent	Rat Sighting	1 JACOBUS PLACE	1008544	258129	(40.876444151331, -73.9128285342
10463	2014-04-01	PAINT/PLASTER	CEILING	110 TERRACE VIEW AVE	1008357	258602	(40.875145400143914, -73.9121540
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	1 JACOBUS PLACE	1008544	258129	(40.876444151331, -73.9128285342
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	140 VAN CORTLANDT A	1013473	261850	(40.885343438331205, -73.8943148
10463	2014-04-01	HEAT/HOT WATER	DOOR	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	Rodent	Condition Attracting Rodents	3555 OXFORD AVENUE	1009450	262094	(40.886025546415 76, -73.9088630
10463	2014-04-01	PAINT/PLASTER	Rat Sighting	3150 BAILEY AVENUE	1011414	259460	(40.878902152578 44, -73.9017708
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	2738 KINGSBRIDGE TER	1010952	256931	(40.871850303847744, -73.9034516
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	110 TERRACE VIEW AVE	1008357	258602	(40.876444151331, -73.9128285342
10463	2014-04-01	HEAT/HOT WATER	ENTIRE BUILDING	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	DOOR/WINDOW	ENTIRE BUILDING	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	Rodent	DOOR	3810 BAILEY AVENUE	1012586	261482	(40.884336286878096, -73.8975242
10306	2014-04-01	Rodent	Condition Attracting Rodents	110 TERRACE VIEW AVE	1008357	258602	(40.876444151331, -73.9128285342
10033	2014-04-01	HEAT/HOT WATER	Rat Sighting	1934 NORTH RAILROAD	1012586	261482	(40.884336286878096, -73.8975242
10463	2014-04-01	Rodent	ENTIRE BUILDING	534 WEST 175TH STREE	1012586	261482	(40.884336286878096, -73.8975242
10306	2014-04-01	Noise - Residential	Rat Sighting	3810 BAILEY AVENUE	1008357	258602	(40.876444151331, -73.9128285342
10306	2014-04-01	Noise - Residential	Loud Music/Party		948358	144556	(40.563376021503466, -74.1291780
11235	2014-04-01	Noise	Banging/Pounding	2341 RICHMOND ROAD	1002201	246971	(40.844535232977 04, -73.9351199
10023	2014-04-01	Noise	Noise: Construction Before/After Hours (NM1)	2775 EAST 12 STREET	1012586	261482	(40.884336286878096, -73.8975242
11230	2014-04-01	Noise	Noise: Alarms (NR3)		954402	142870	(40.558770684605705, -74.1177974
10003	2014-04-01	Noise	Noise: Construction Before/After Hours (NM1)		951527	149691	(40.577482843138505, -74.1074178
					996114	151870	(40.583516220598 98, -73.9572876
					990265	220914	(40.773032188080 42, -73.9782834
					996858	163767	(40.616170043694 14, -73.9545869
					985895	205227	(40.729977235073 79, -73.9940647

Data ETL – reduce

Period	Raw data's obs number	Obs # with missing ZIP	Obs # missing all location info	Service requests' modalities #	Unique ZIP
April 2014	81645	3206	2740	170	278
April 2015	101890	4231	3069	178	260

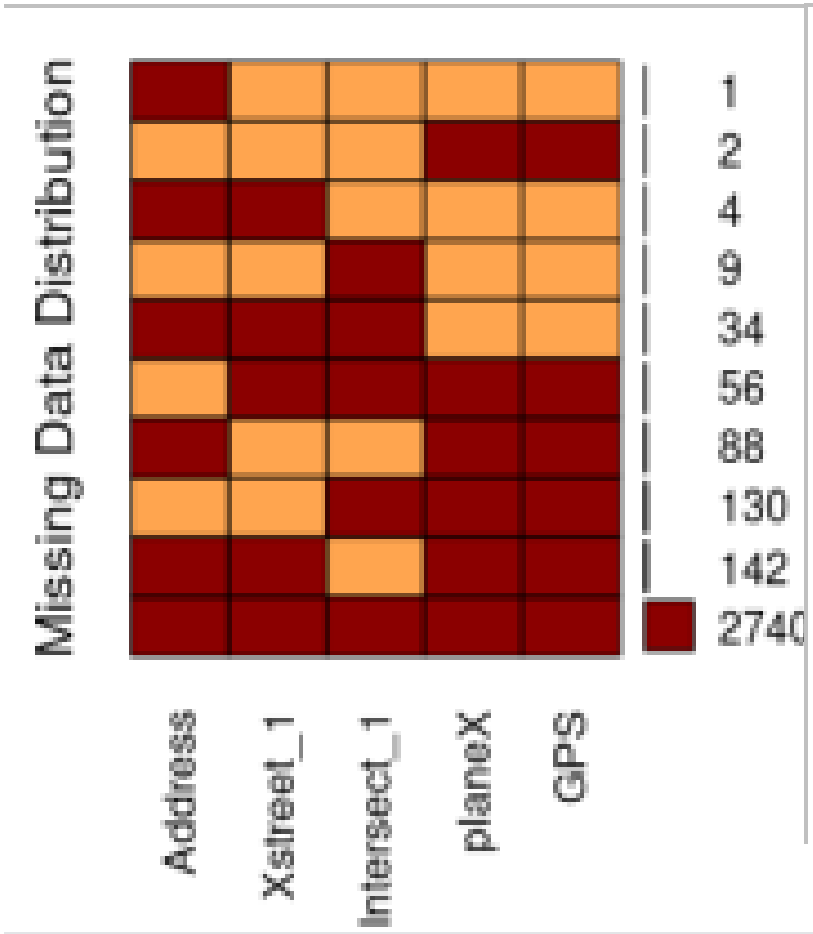
- Between 150 and 200 different SRCs' raw features
- Reduce to 13 features by consolidating calls' objects

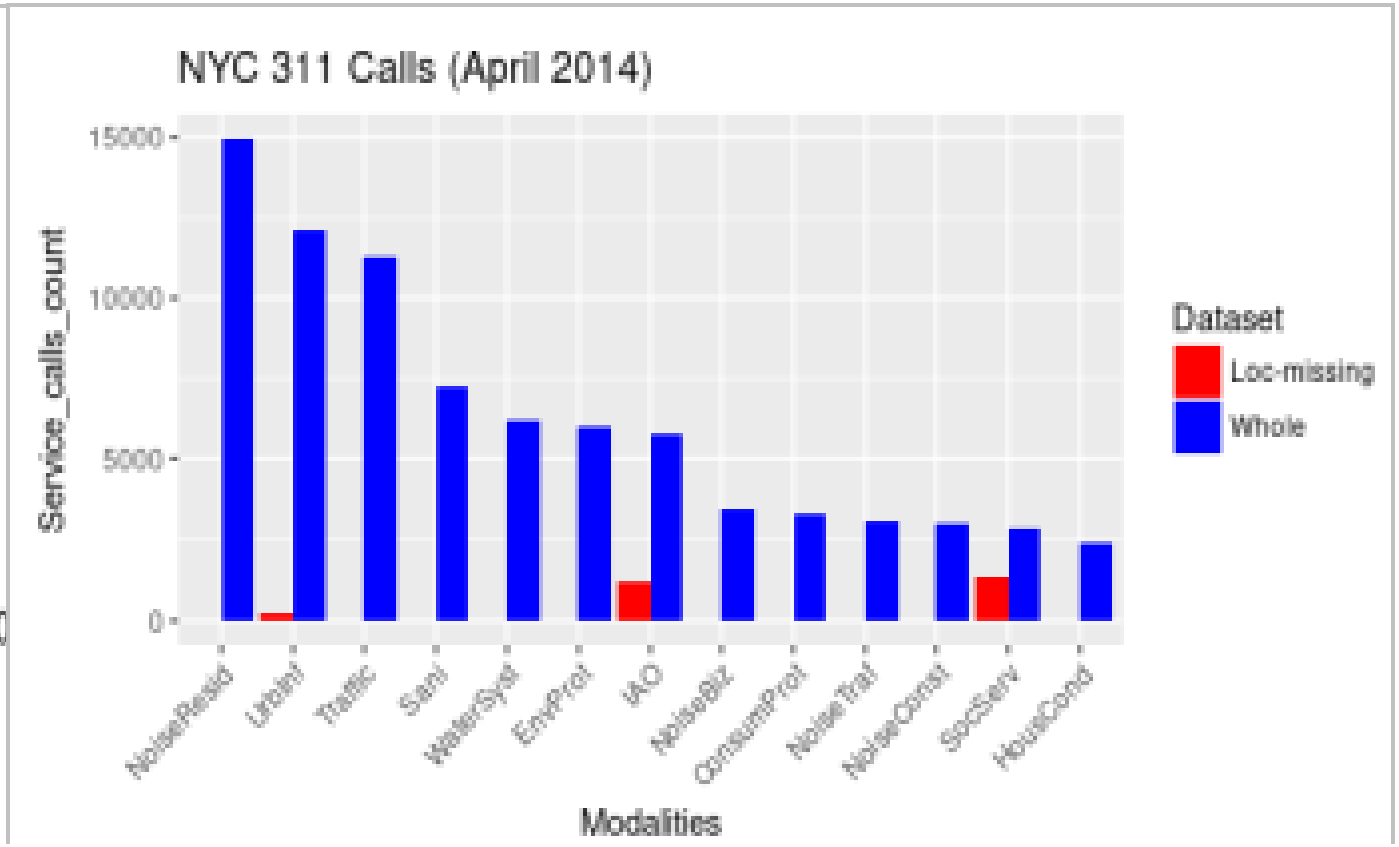
Data ETL – reduce

Service request calls' modalities	Modality description	Service request call frequencies		Change in rank from 2014 to 2015
		April 2014	April 2015	
NoiseResid	Residential Noise	19.00%	17.50%	–
UrbInf	Urban Infrastructure	15.00%	13.40%	↘
Traffic	Traffic related Issues	14.30%	17.20%	↗
Sani	Unsanitary Conditions	9.20%	10.50%	–
WaterSyst	Water Systems	7.80%	7.60%	–
EnvProt	Environmental Protection	7.60%	5.90%	–
IAO	Inspect., Audit, Order	5.80%	5.20%	↘
NoiseBiz	Commercial Noise	4.40%	4.90%	↘
ConsumProt	Comsumer Protection	4.20%	3.40%	↘
NoiseTraf	Traffic Noise	3.90%	5.40%	↗↗
NoiseConst	Construction Noise	3.80%	3.70%	↗
HousCond	Housing Conditions	3.10%	3.40%	–
SocServ	Social Services	1.90%	1.90%	–
	Total number of SRCs	78825	98649	↗↗

Data ETL – impute missings

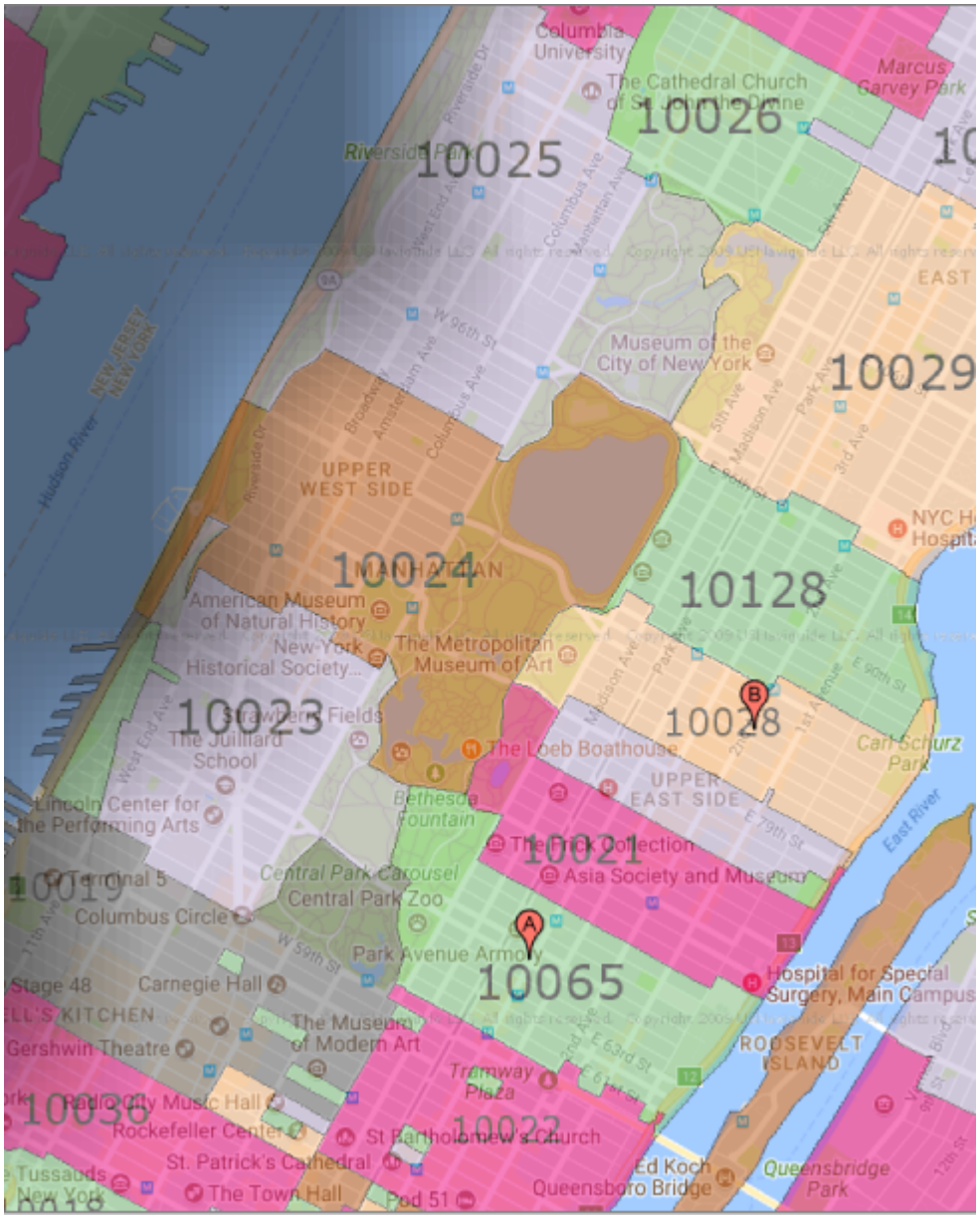
Period	Raw data's obs number	Obs # with missing ZIP	Obs # missing all location info	Service requests' modalities #	Unique ZIP
April 2014	81645	3206	2740	170	278





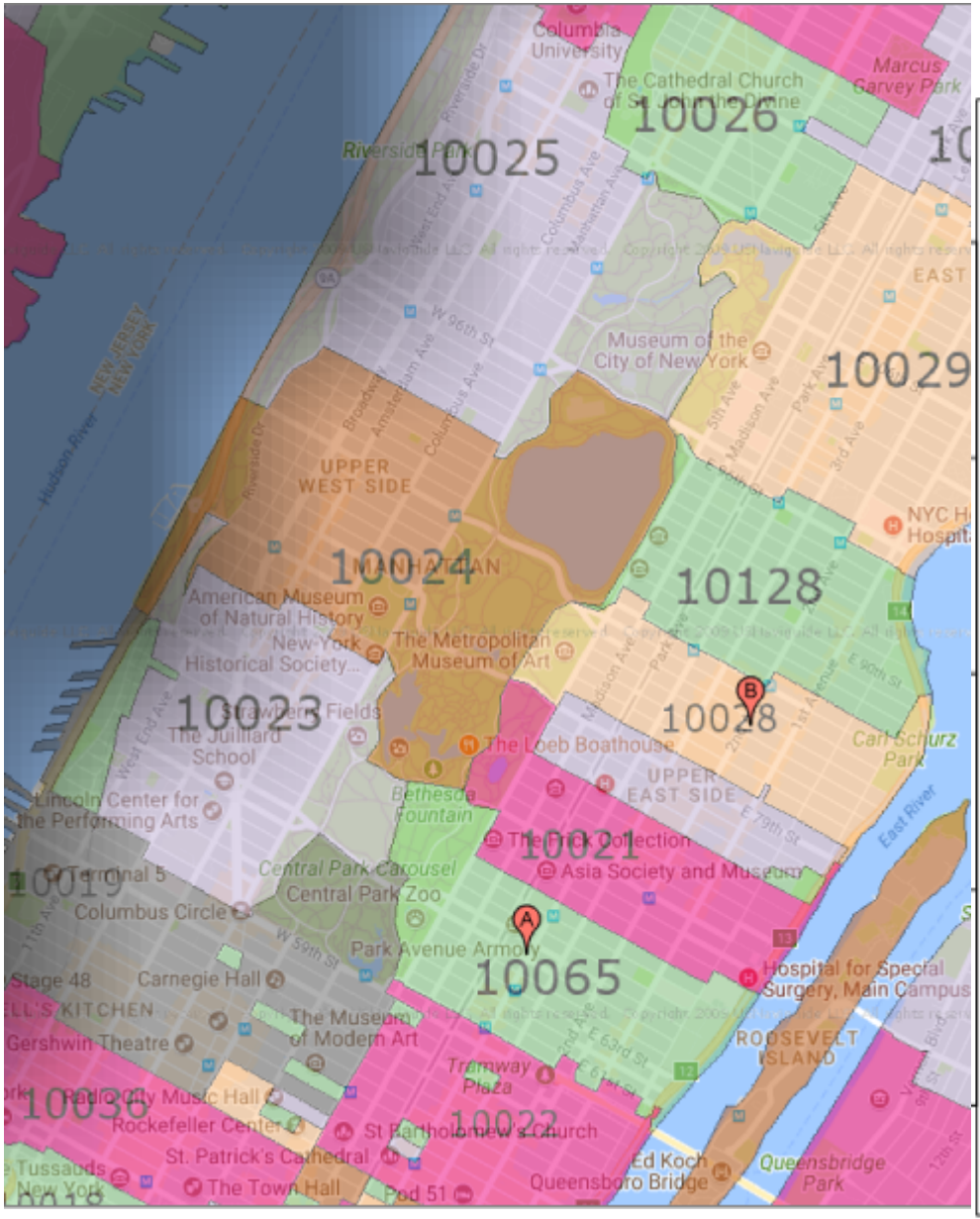
Data ETL – clean

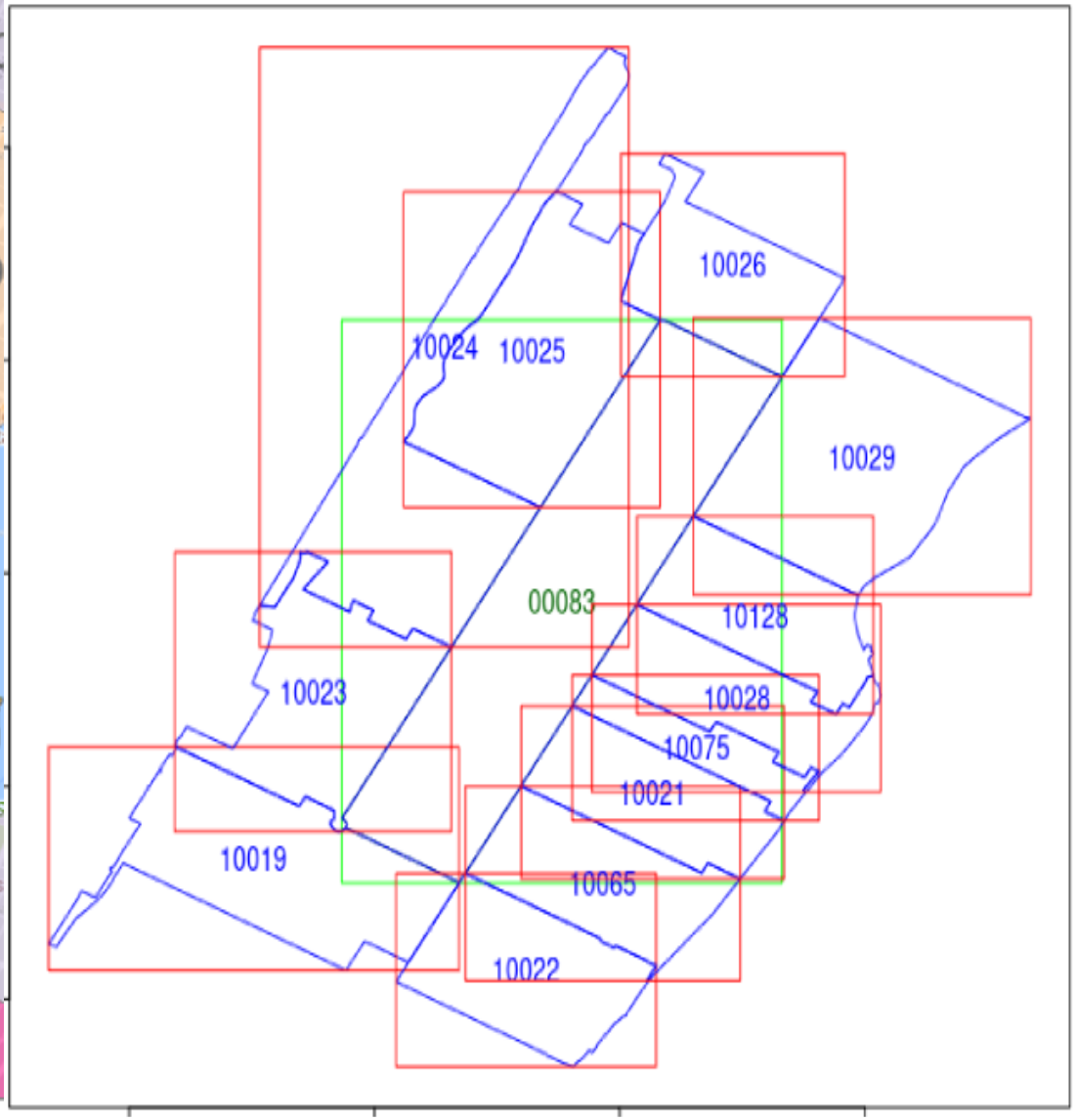
ZIP code: 00083



Data ETL – clean

ZIP code: 00083





Data ETL – impute / clean

ZIP code: 10020

MVA – ZIP-SRC contingency table *(April 2014)*

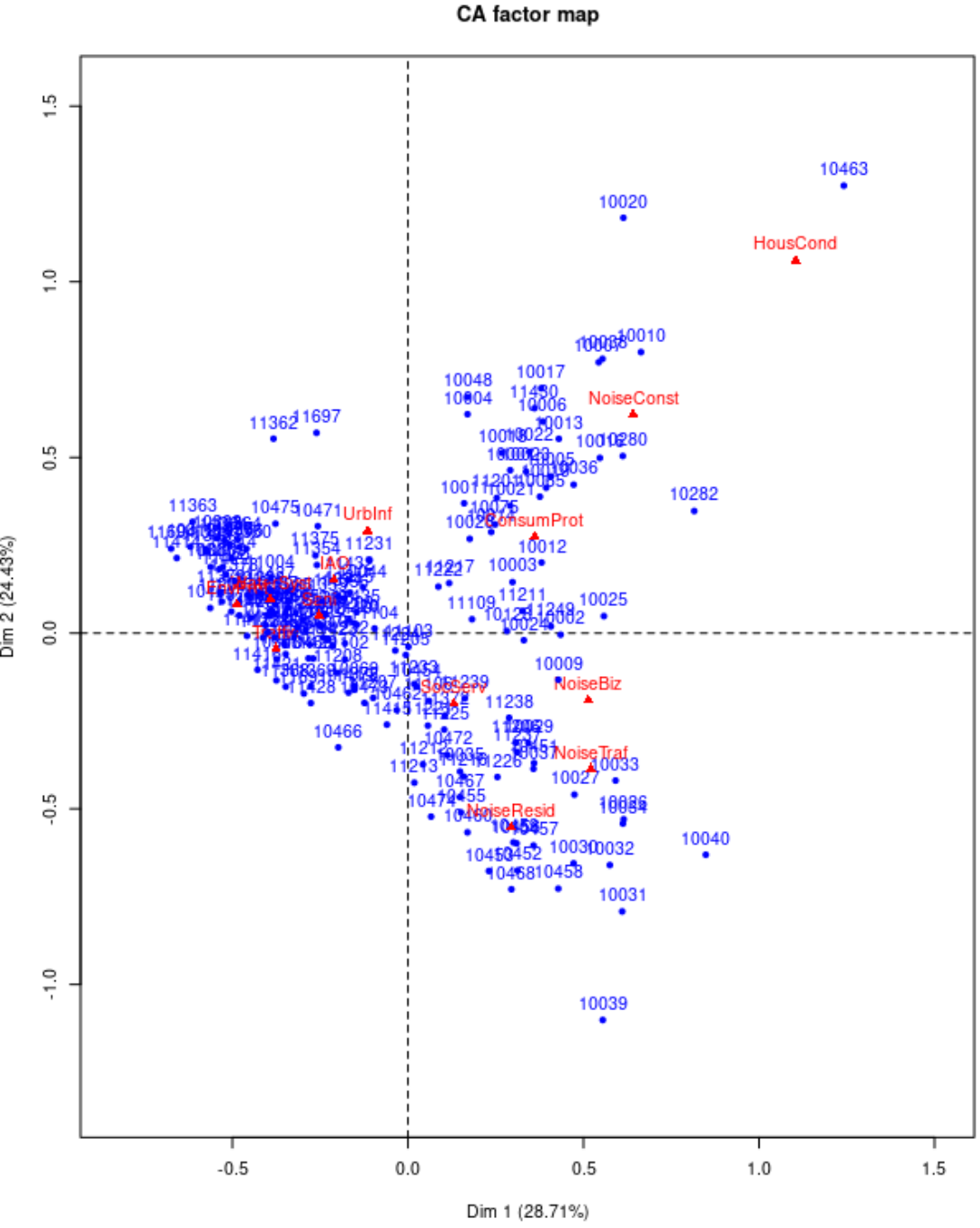
- Build frequency table
  (row profiles for ZIPs, column profiles for SRCs)

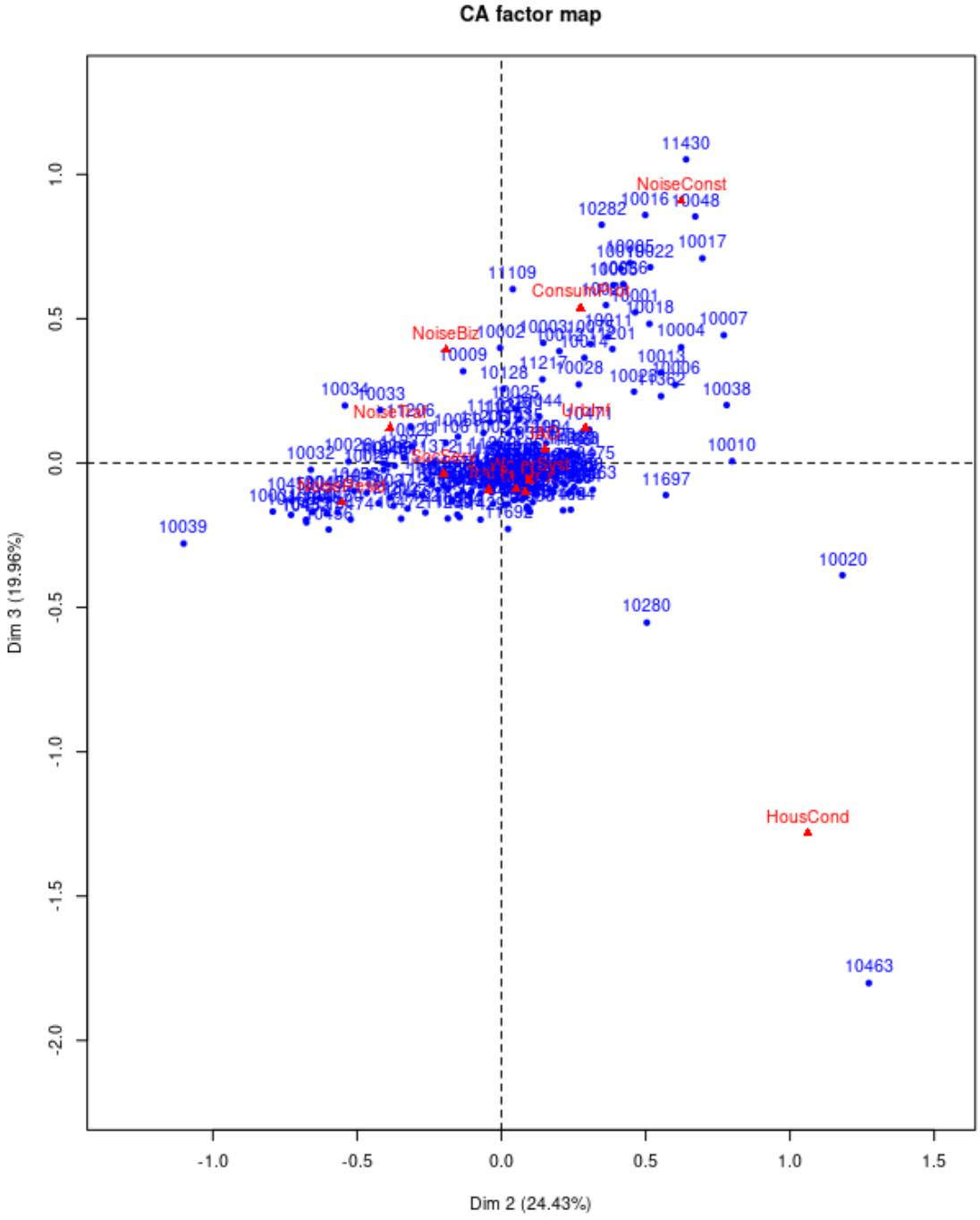
- Observe how 26 row marginals < 5 / nbr_calls
  → Can we suppress them ? *(… χ^2-test of independence)*

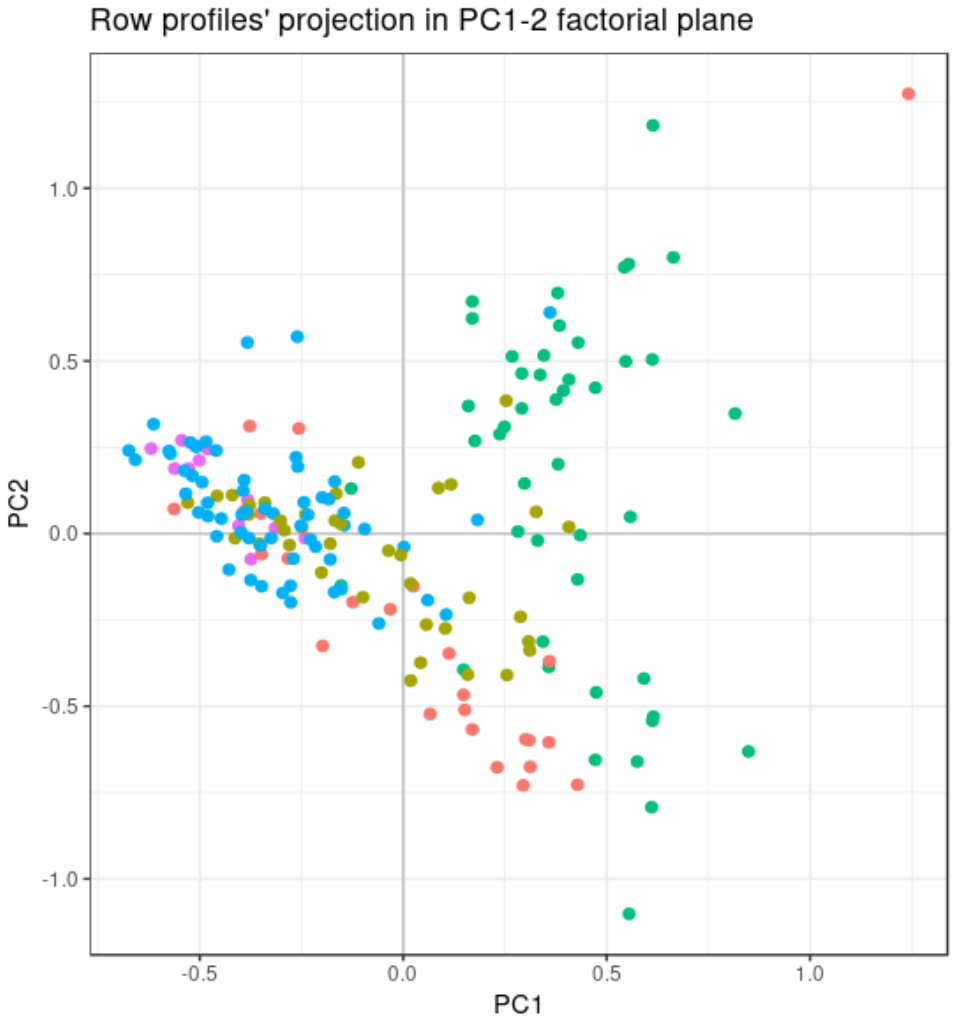
- Run CA with row marginals as row profile weights
  (*χ^2-metric)*

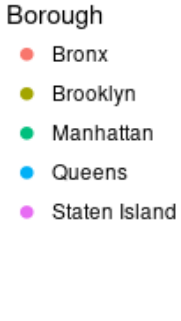
MVA – CA

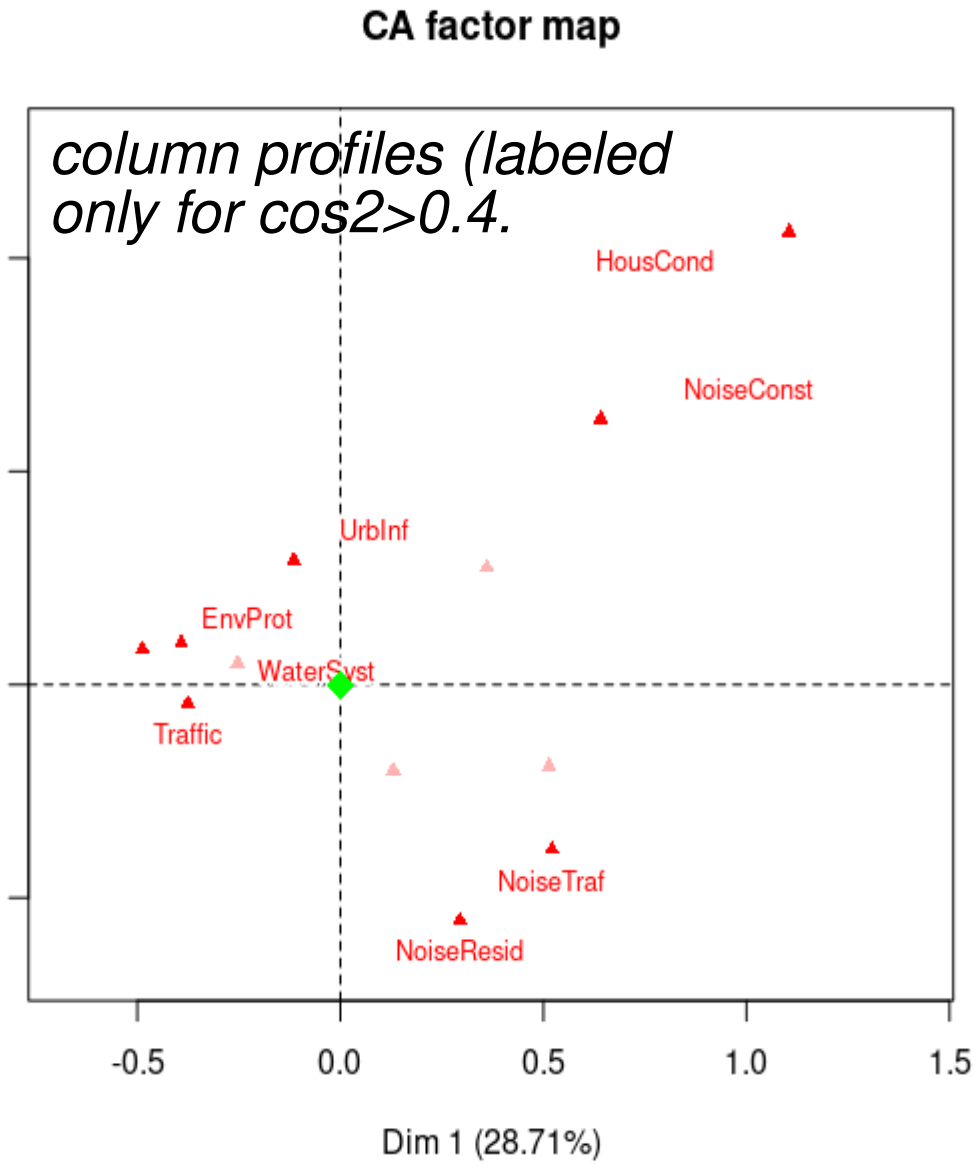
(April 2014)





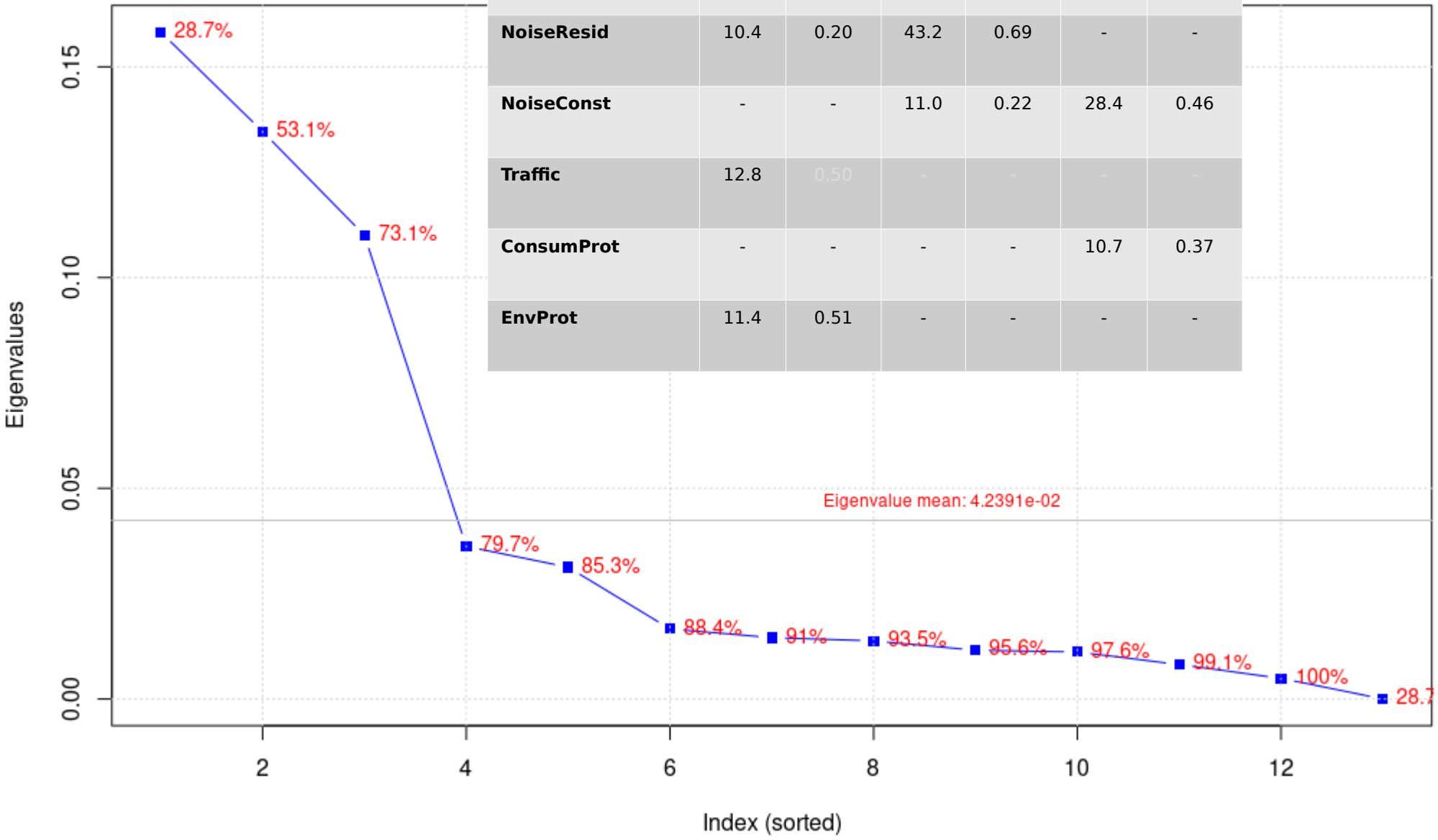

Row profiles' projection in PC1-2 factorial plane
PC2
PC1
1.0
0.5
0.0
-0.5
-1.0
-0.5
0.0
0.5
1.0


Borough
Bronx
Brooklyn
Manhattan
Queens
Staten Island

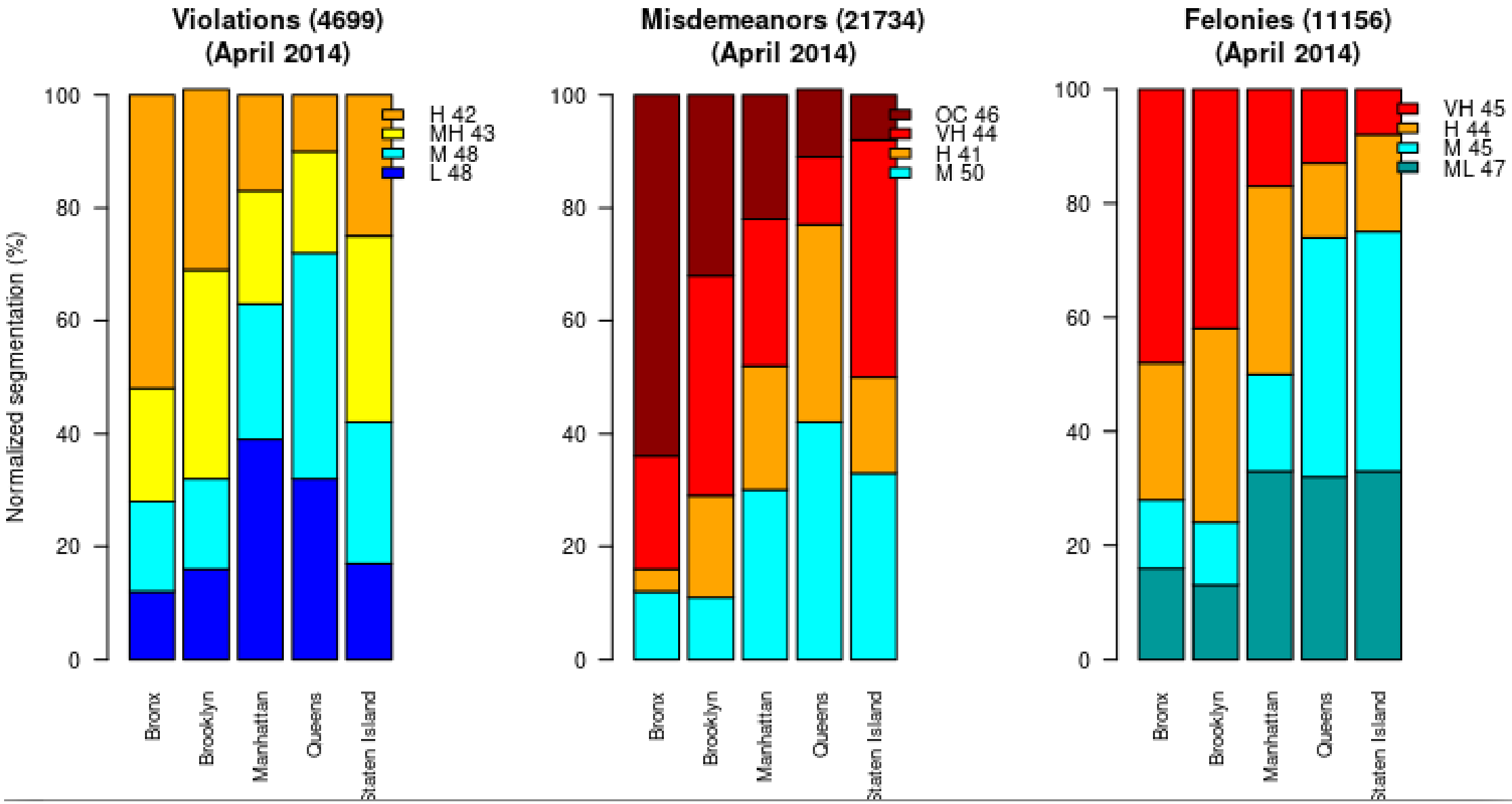

CA factor map
column profiles (labeled only for cos2>0.4.
HousCond
NoiseConst
UrbInf
EnvProt
WaterSyst
Traffic
NoiseTraf
NoiseResid
-0.5
0.0
0.5
1.0
1.5
Dim 1 (28.71%)

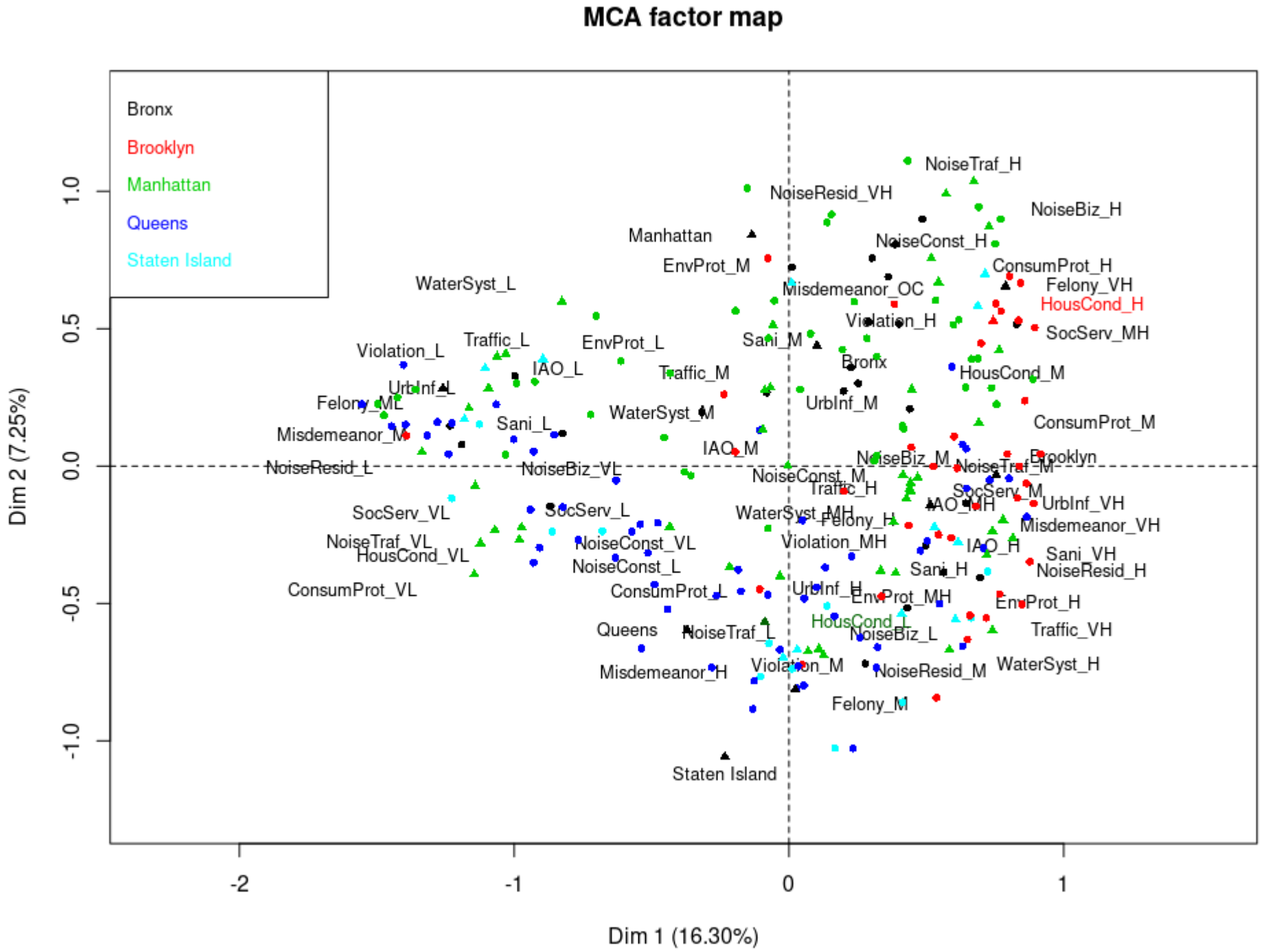
MVA – PCA

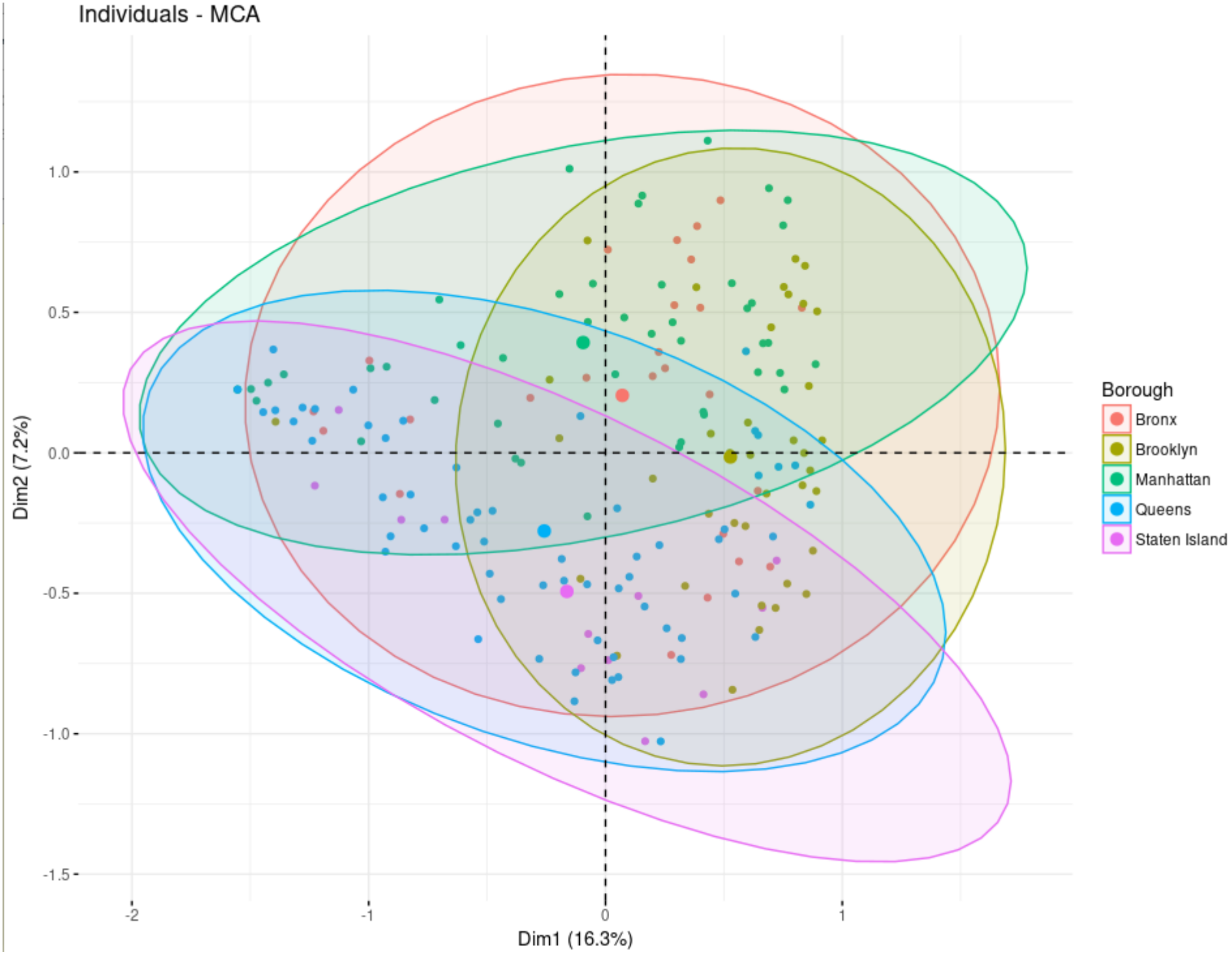
	Dim 1		Dim 2		Dim 3	
	ctr	*cos2*	*ctr*	*cos2*	*ctr*	*cos2*
HousCond	23.7	0.30	25.7	0.29	45.9	0.41
NoiseResid	10.4	0.20	43.2	0.69	-	-
NoiseConst	-	-	11.0	0.22	28.4	0.46
Traffic	12.8	0.50	-	-	-	-
ConsumProt	-	-	-	-	10.7	0.37
EnvProt	11.4	0.51	-	-	-	-



MVA – MCA


Violations (4699)
(April 2014)
Misdemeanors (21734)
(April 2014)
Felonies (11156)
(April 2014)
Normalized segmentation (%)
0
20
40
60
80
100
H 42
MH 43
M 48
L 48
OC 46
VH 44
H 41
M 50
VH 45
H 44
M 45
ML 47
Bronx
Brooklyn
Manhattan
Queens
Staten Island


MCA factor map
Bronx
Brooklyn
Manhattan
Queens
Staten Island
Dim 2 (7.25%)
Dim 1 (16.30%)
1.0
0.5
0.0
-0.5
-1.0
-2
-1
0
1
NoiseTraf_H
NoiseResid_VH
NoiseBiz_H
Manhattan
NoiseConst_H
EnvProt_M
ConsumProt_H
Felony_VH
WaterSyst_L
Misdemeanor_OC
HousCond_H
Violation_H
SocServ_MH
Traffic_L
EnvProt_L
Sani_M
Violation_L
IAO_L
Bronx
HousCond_M
Traffic_M
UrbInf_L
UrbInf_M
Felony_ML
WaterSyst_M
ConsumProt_M
Sani_L
Misdemeanor_M
IAO_M
NoiseBiz_M
Brooklyn
NoiseResid_L
NoiseBiz_VL
NoiseConst_M
NoiseTraf_M
Traffic_H
SocServ_M
IAO_VH
UrbInf_VH
SocServ_VL
SocServ_L
WaterSyst_MH
Felony_H
Misdemeanor_VH
NoiseTraf_VL
NoiseConst_VL
Violation_MH
IAO_H
Sani_VH
HousCond_VL
NoiseConst_L
Sani_H
NoiseResid_H
ConsumProt_VL
ConsumProt_L
UrbInf_H
EnvProt_MH
EnvProt_H
HousCond_L
Queens
NoiseTraf_L
NoiseBiz_L
Traffic_VH
Violation_M
NoiseResid_M
WaterSyst_H
Misdemeanor_H
Felony_M
Staten Island


Individuals - MCA
Dim2 (7.2%)
Dim1 (16.3%)
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2
-1
0
1
Borough
Bronx
Brooklyn
Manhattan
Queens
Staten Island