An Analysis of the effects of housing characteristics on residential density

Daniel Hartig

Arlington, Virginia

Cambridge, Massachusetts

Housing characteristics

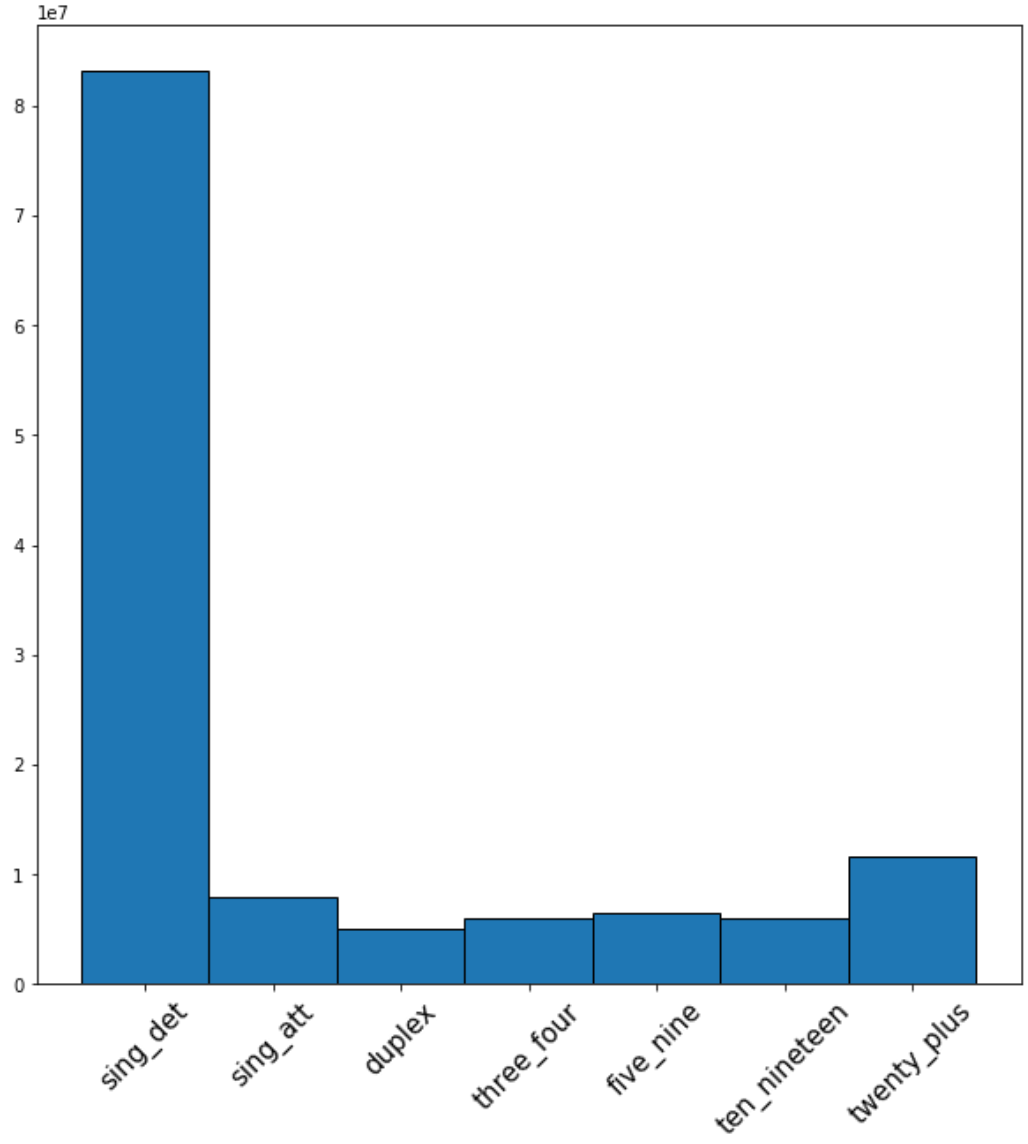
	Arlington	Cambridge
Area (km^2)	63	52
Population	223446	261760
Pop Density	3553	5068
Housing Units	109647	114074

	Arlington	Cambridge
Single Family, Detached	29115	16087
Single Family, Attached	10489	8601
Duplex	1044	23516
Triplex/Four-plex	4138	22338
5-9 Apartment	6231	10192
10-19 Apartment	8816	6062
20+ Apartment	49528	27116

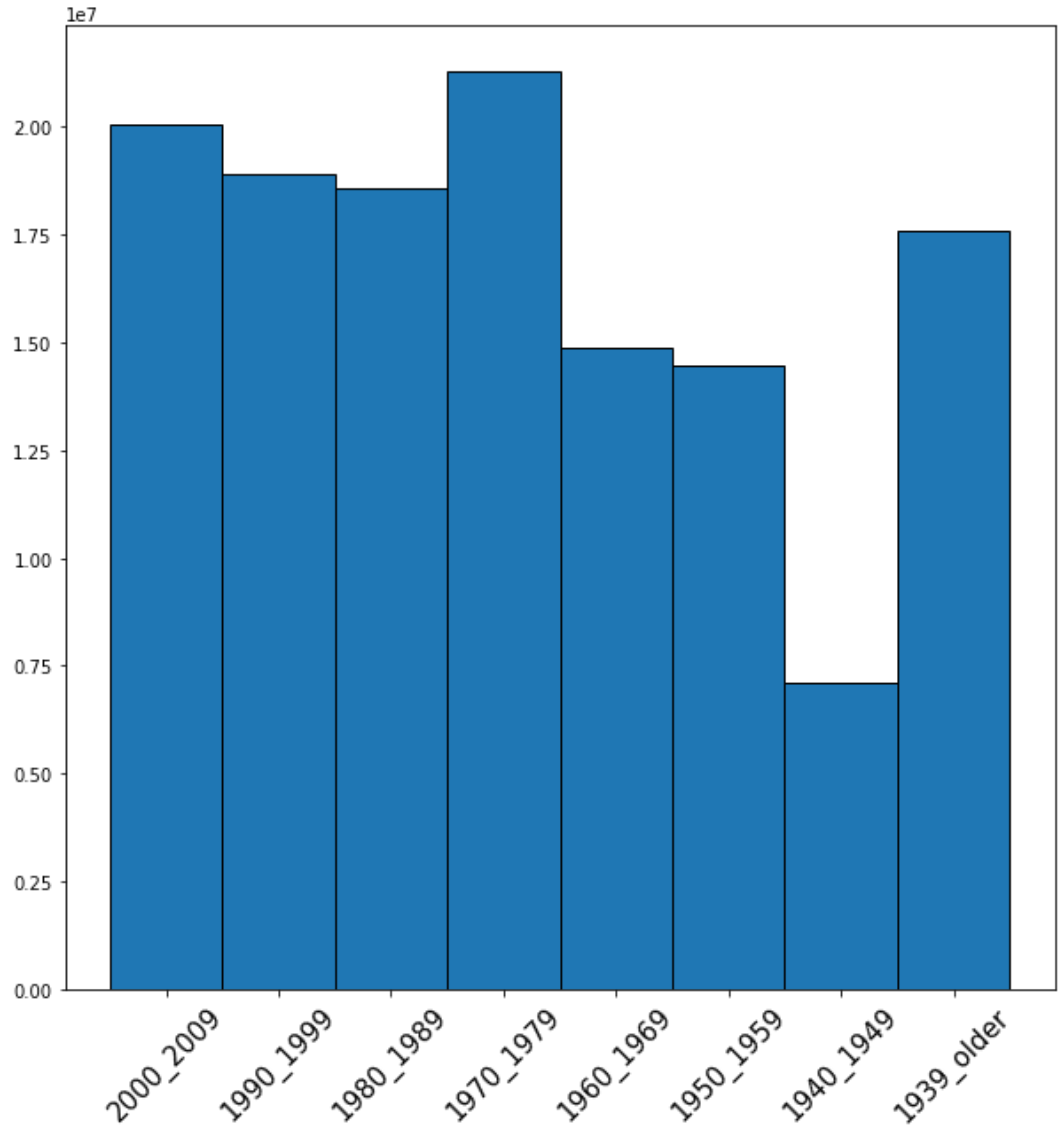
	Arlington	Cambridge
2000-2009	17044	7678
1990-1999	10335	3751
1980-1989	12071	7478
1970-1979	11861	7909
1960-1969	12339	6989
1950-1959	16698	6928
1940-1949	16145	6681
1939 or earlier	10227	65678

Distribution of characteristics

Type of Housing



Age of Housing



Two way contingency table

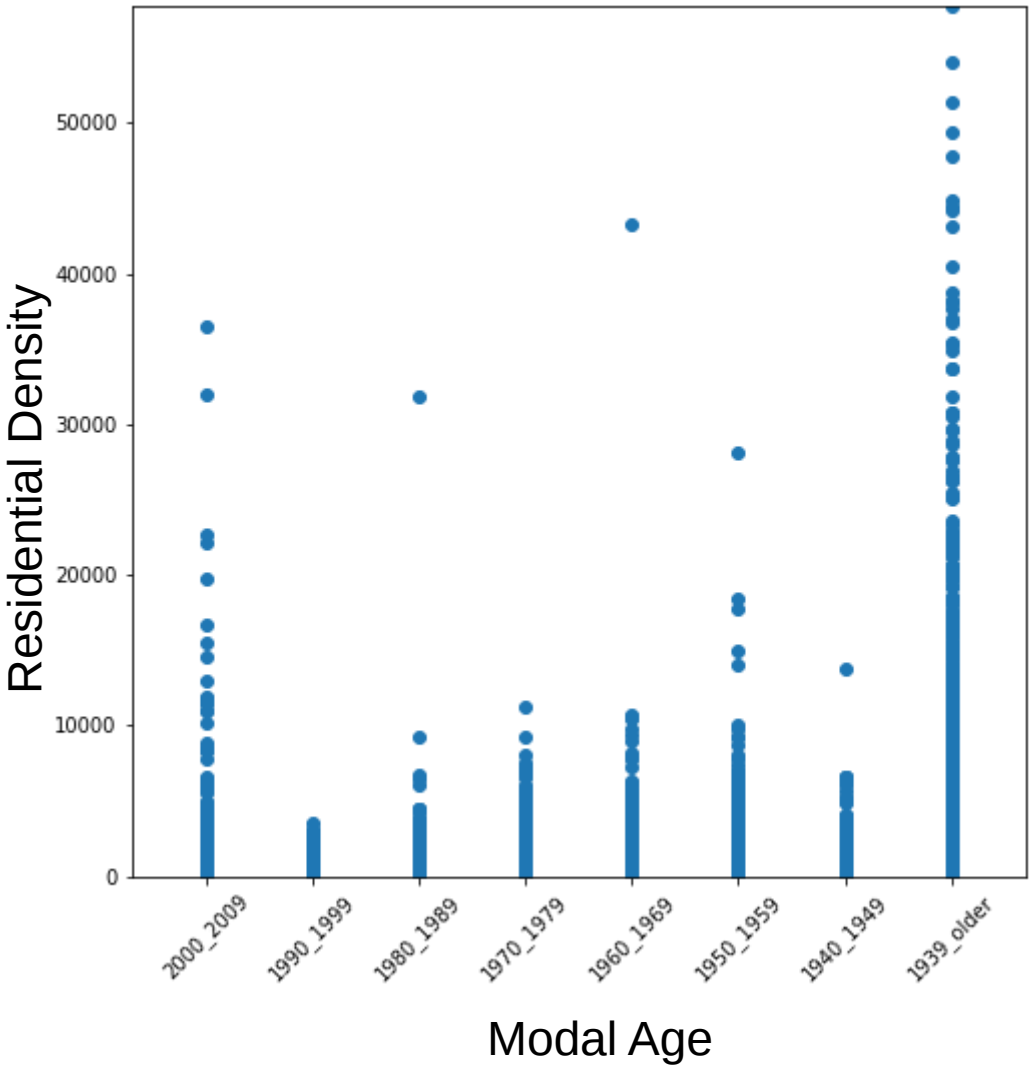
	2000_2009	1990_1999	1980_1989	1970_1979	1960_1969	1950_1959	1940_1949	1939_older
sing_det	13231648	12383720	11622307	13186186	8981531	8881509	4171739	9434389
sing_att	1080944	1057279	1117296	1219262	876154	868056	460750	1147528
duplex	481991	451096	493335	639427	569156	650298	396436	1294500
three_four	672204	6457310	725539	887440	686525	689152	380892	1204765
five_nine	825655	819087	911708	1061701	763556	704955	345267	9250555
ten_nineteen	829216	805582	907553	102648	719441	644608	296961	730331
twenty_plus	1456613	1206455	1436357	1784572	1436910	1335572	694701	2119241

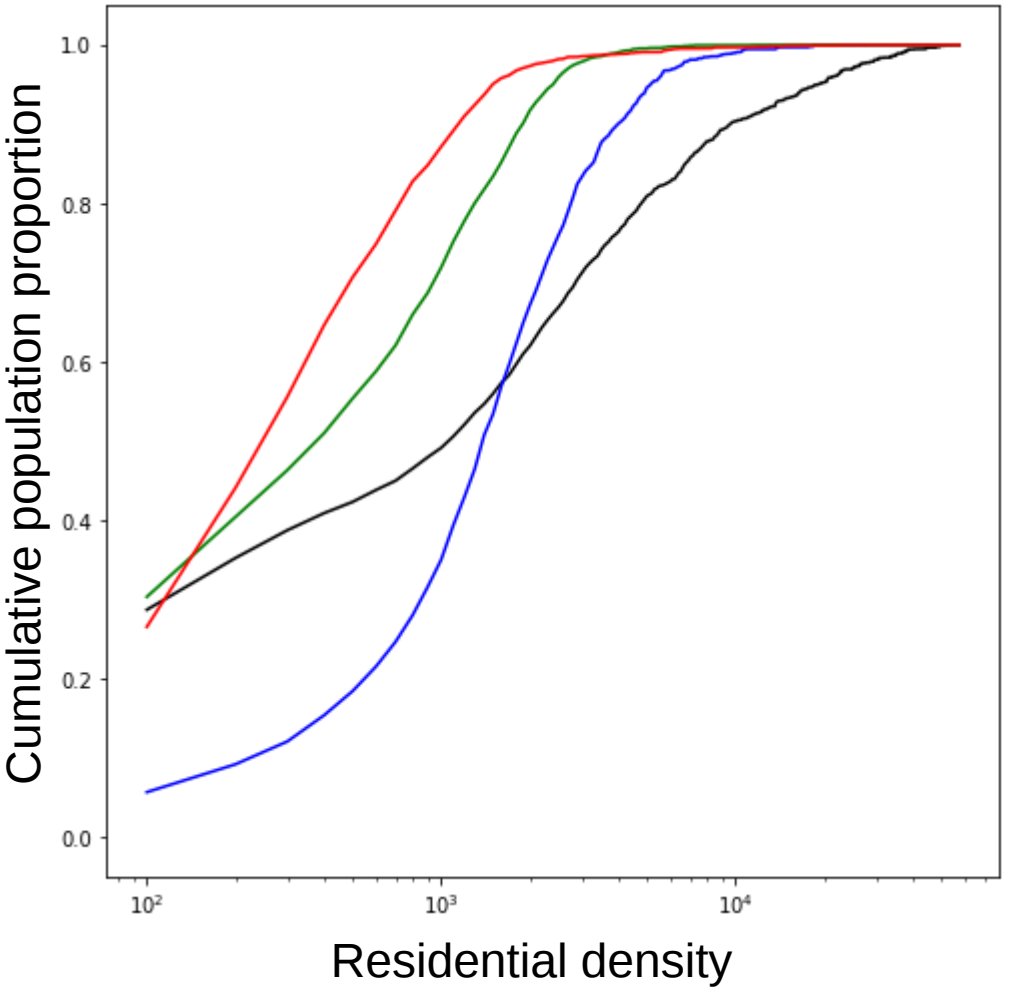
Chi-square score of 2146451, with p-value = 0.0

Fractional difference from expected count

	2000_2009	1990_1999	1980_1989	1970_1979	1960_1969	1950_1959	1940_1949	1939_older
sing_det	0.082	0.082	0.025	0.011	-0.028	-0.021	-0.061	-0.150
sing_att	-0.075	-0.033	0.031	-0.022	-0.008	0.001	0.085	0.082
duplex	-0.352	-0.351	-0.284	-0.193	0.013	0.180	0.469	0.920
three_four	-0.236	-0.215	-0.110	-0.054	0.033	0.056	0.197	0.508
five_nine	-0.131	-0.077	0.036	0.049	0.065	0.001	0.001	0.074
ten_nineteen	-0.069	-0.032	0.100	0.082	0.070	-0.023	-0.081	-0.096
twenty_plus	-0.150	-0.247	-0.096	-0.023	0.110	0.051	0.117	0.363

Analysis of age distribution





Black – 1939 or before
Blue – 1940-1969
Green – 1970-1999
Red – 2000 or later

Housing type versus density regression

Using a flag for at least 20% of housing in a type category as feature

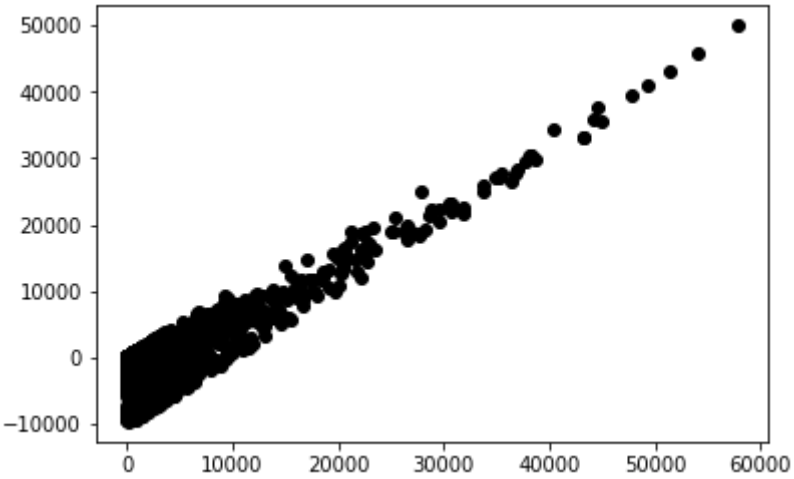
OLS Regression Results

Dep. Variable:	density	R-squared:	0.244
Model:	OLS	Adj. R-squared:	0.244
Method:	Least Squares	F-statistic:	1757.
Date:	Sun, 03 Dec 2017	Prob (F-statistic):	0.00
Time:	13:33:57	Log-Likelihood:	-2.8745e+05
No. Observations:	32633	AIC:	5.749e+05
Df Residuals:	32626	BIC:	5.750e+05
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	252.1625	9.315	27.069	0.000	233.904	270.421
sing_att[T.True]	743.6597	57.598	12.911	0.000	630.765	856.555
duplex[T.True]	2172.6819	83.310	26.080	0.000	2009.392	2335.972
three_four[T.True]	2378.1527	101.683	23.388	0.000	2178.851	2577.455
five_nine[T.True]	858.0827	122.476	7.006	0.000	618.026	1098.139
ten_nineteen[T.True]	384.4179	122.999	3.125	0.002	143.336	625.500
twenty_plus[T.True]	4336.1339	48.463	89.473	0.000	4241.145	4431.123

Omnibus:	58417.086	Durbin-Watson:	0.758
Prob(Omnibus):	0.000	Jarque-Bera (JB):	105123996.503
Skew:	12.922	Prob(JB):	0.00
Kurtosis:	279.850	Cond. No.	14.5

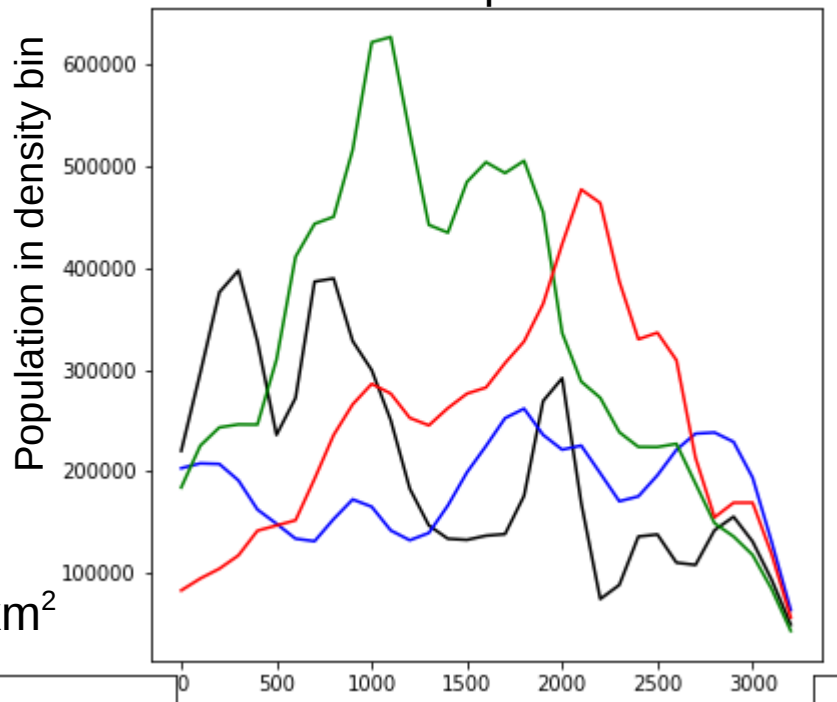
- Lowest condition number, but R^2 is almost garbage
- Problems: estimate of 100 % large apartments is lower than estimate of 50/50 large and small apartments
- Look at that residual plot!
- At least the coefficients all have a real meaning
- Look a the groupings of coefficients



Distributions of dominant housing types by density

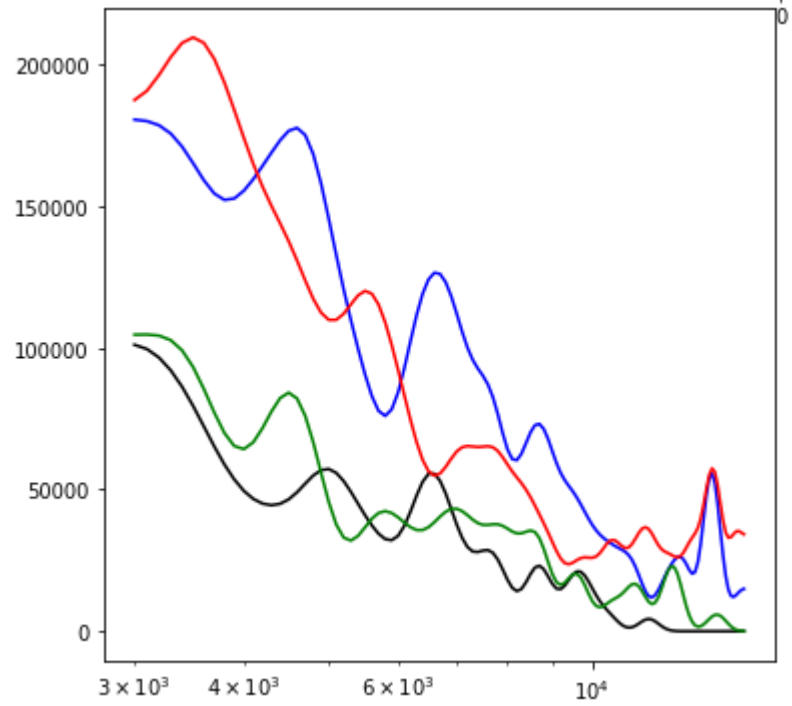
- Rowhouses: Black
- Duplex to Fourplex: Blue
- Smaller Apts: Green
- Large Apts: Red

0 – 3333 per km^2

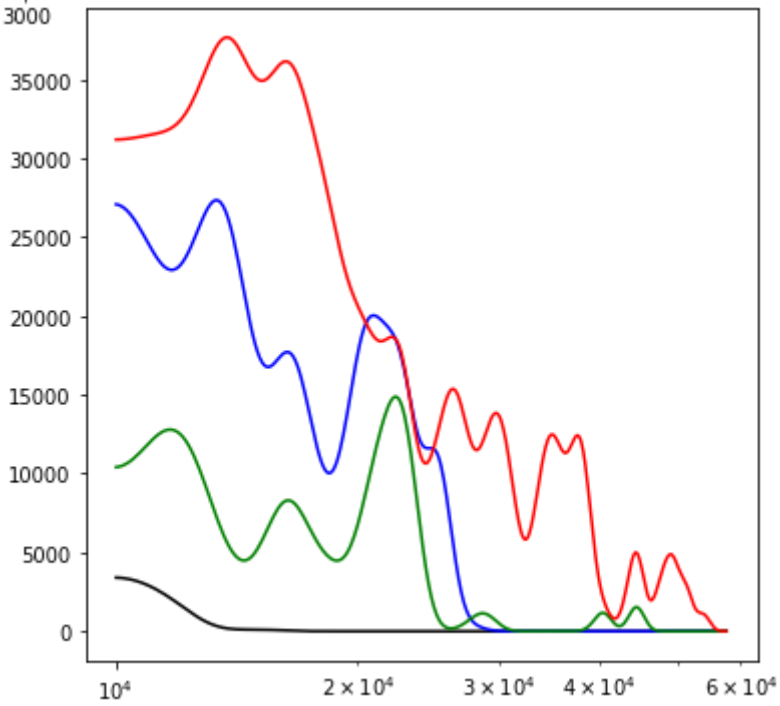


- A zip code is counted in a type category if at least 20% of the housing units in that zip code are in that type category.

3333 – 15000 per km^2



10000 – 60000 per km^2



Conclusions

- Zip codes built up before 1940 are the densest

- Densities above suburban (~3,000 people per km^2; like Old Town Alexandria) are dominated by either 2-4 unit housing or large apartments

- Densities in the range of big cities (~10,000 people per km^2 and higher; like Dupont Circle) are dominated by large apartments

- These two types of housing are heavily overrepresented in zip codes built up before 1940

- We haven’t sufficiently disentagled whether it is the housing age or housing type that is responsible for higher density.

Any Questions?