

Визуализация внешних данных 2016 US Election

Николаев Сергей, 517 группа

1 октября 2018 г.

Что за данные?

- Таблица `county_facts` - демографические и экономические показатели по различным округам США

	fips	area_name	state_abbreviation	PST045214	PST040210	PST120214
0	0	United States	NaN	318857056	308758105	3.3
1	1000	Alabama	NaN	4849377	4780127	1.4
2	1001	Autauga County	AL	55395	54571	1.5
3	1003	Baldwin County	AL	200111	182265	9.8
4	1005	Barbour County	AL	26887	27457	-2.1

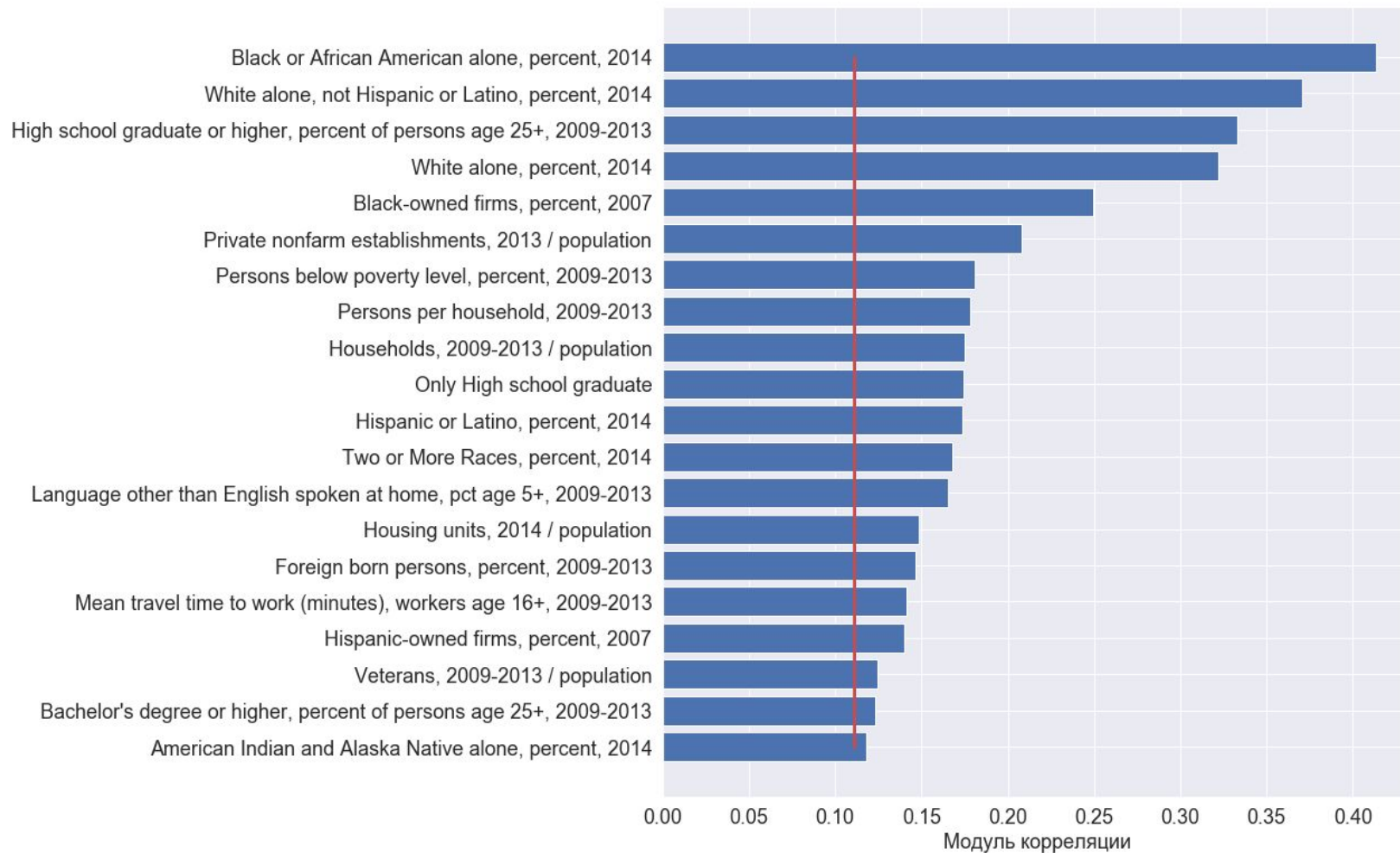
Что за данные?

- Таблица `primary_results` - результаты праймериз в Демократической и Республиканской партиях

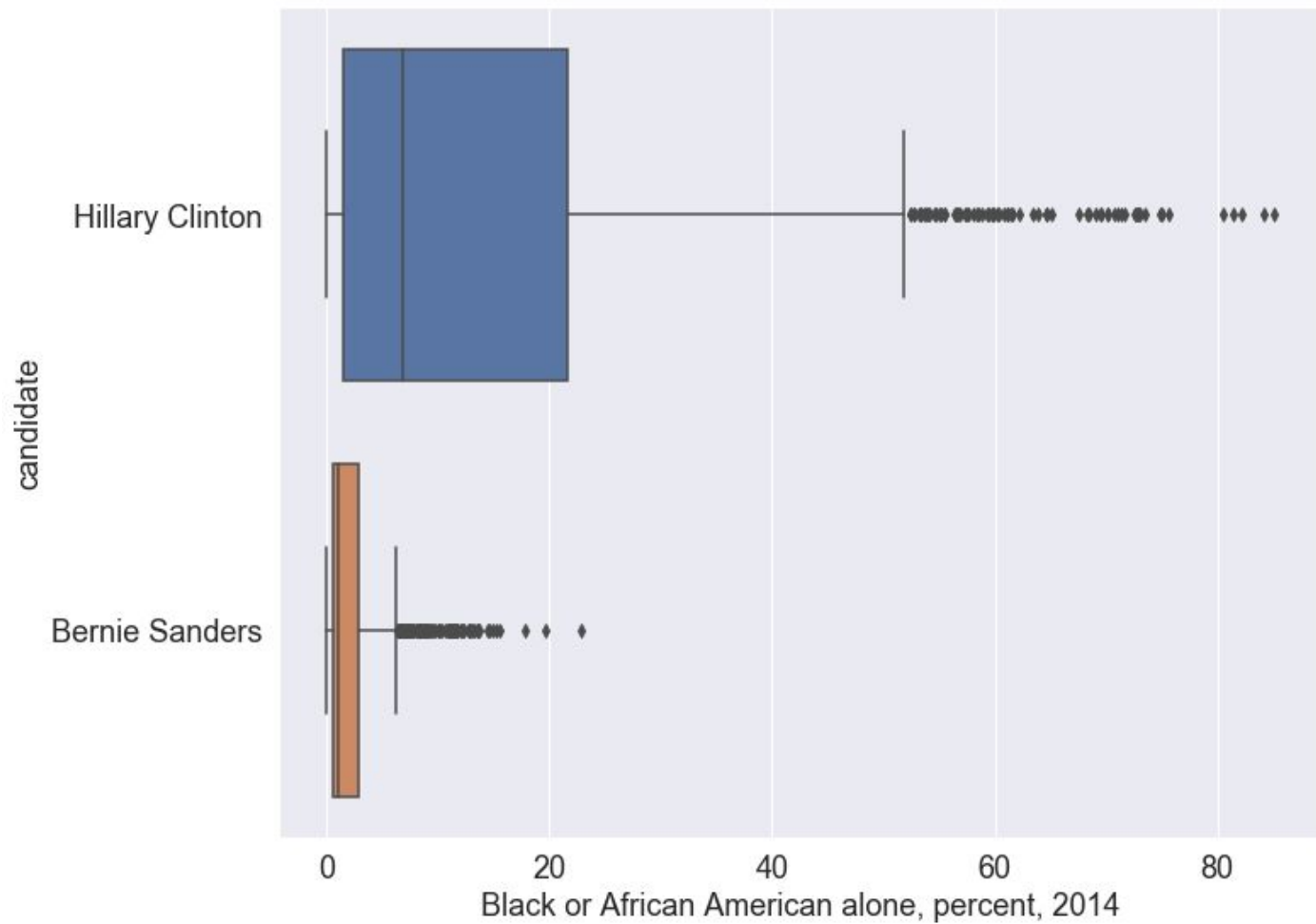
	state	state_abbreviation	county	fips	party	candidate	votes	fraction_votes
0	Alabama	AL	Autauga	1001.0	Democrat	Bernie Sanders	544	0.182
1	Alabama	AL	Autauga	1001.0	Democrat	Hillary Clinton	2387	0.800
2	Alabama	AL	Baldwin	1003.0	Democrat	Bernie Sanders	2694	0.329
3	Alabama	AL	Baldwin	1003.0	Democrat	Hillary Clinton	5290	0.647
4	Alabama	AL	Barbour	1005.0	Democrat	Bernie Sanders	222	0.078

- Будем рассматривать признаки в контексте праймериз Демократической партии
- Два претендента - Хиллари Клинтон и Берни Сандерс
- Признаков много (>60), поэтому визуализировать все не получится, надо производить отбор признаков

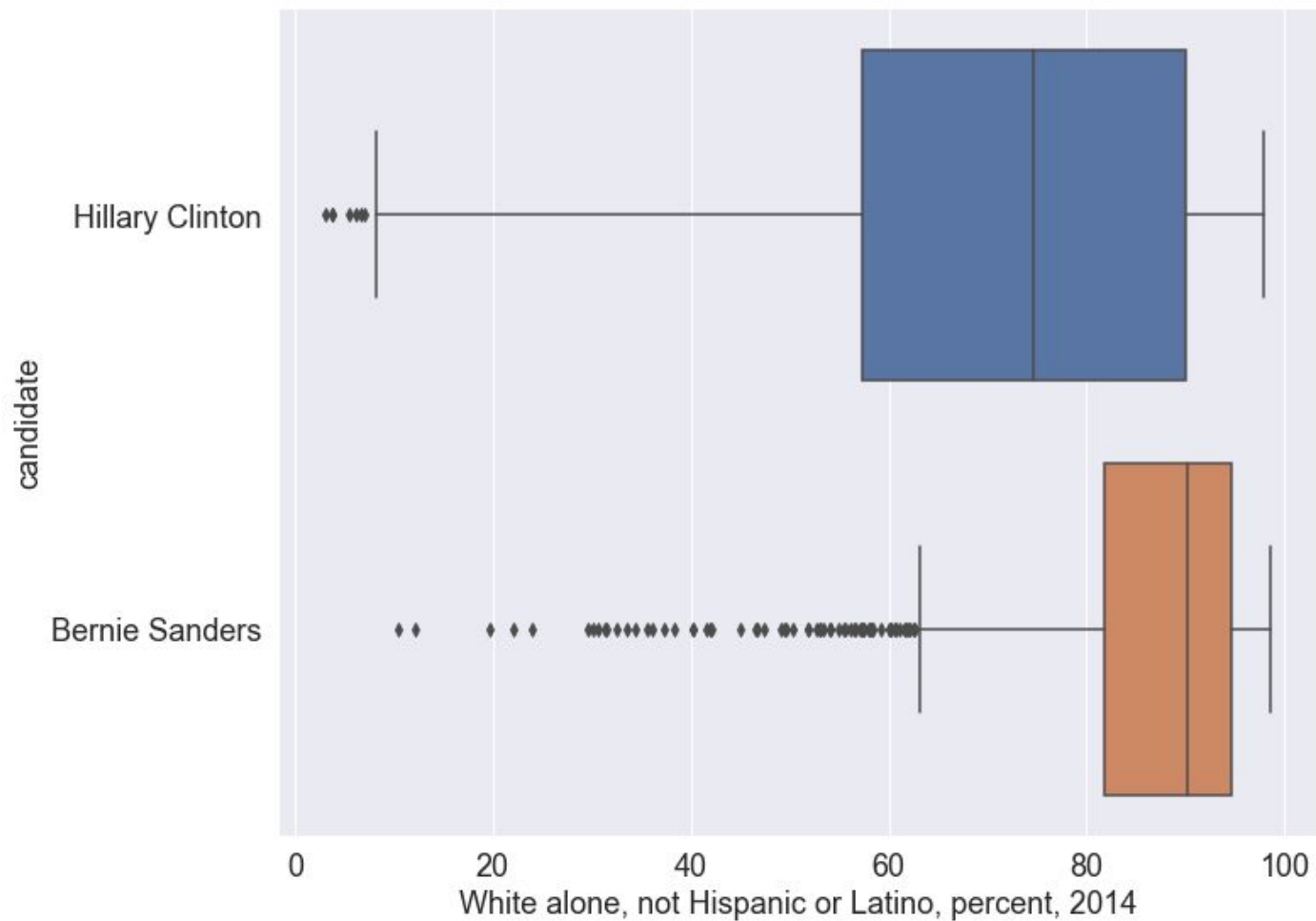
Топ признаков по модулю корреляции с целевой переменной



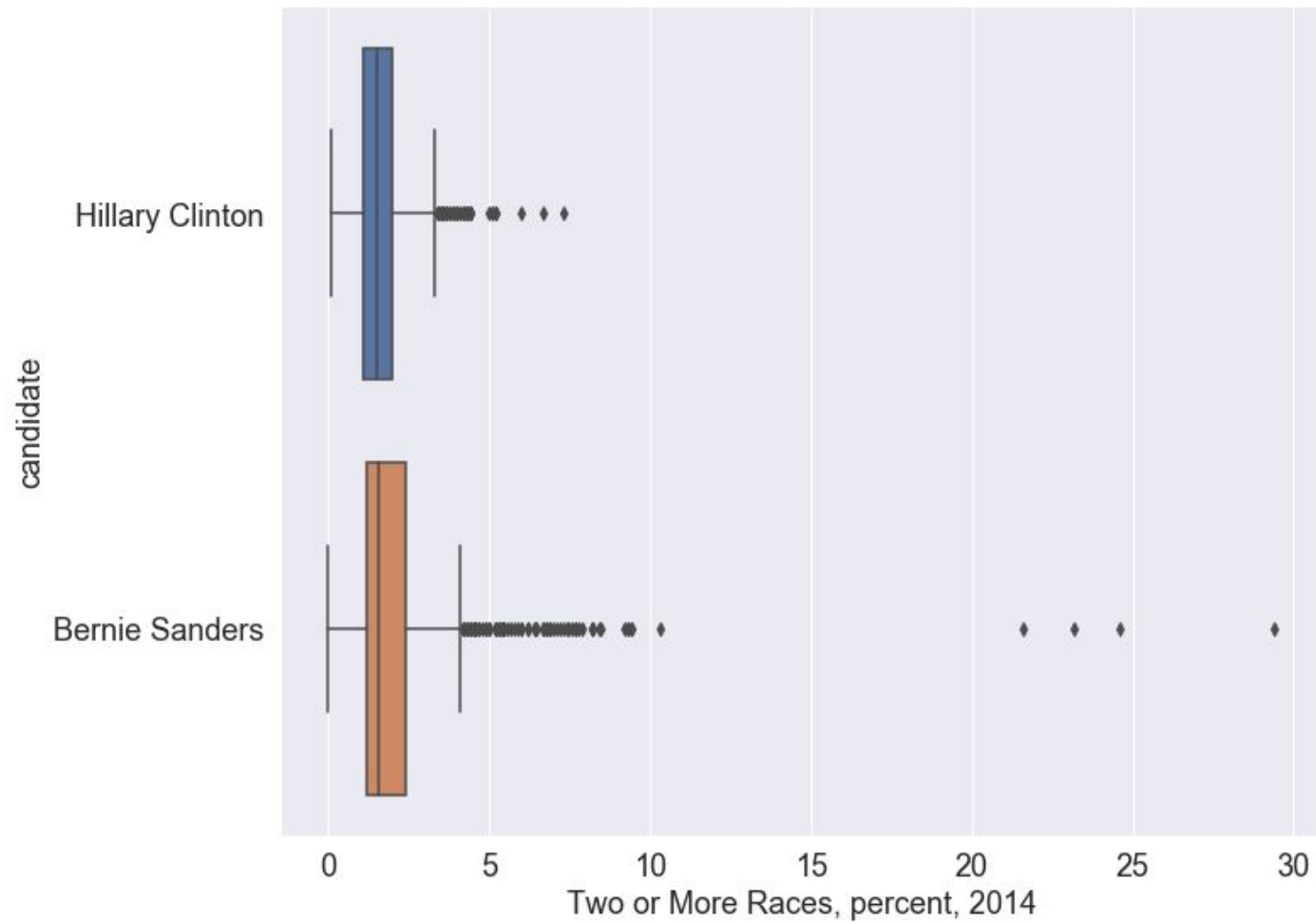
Зависимость от расы



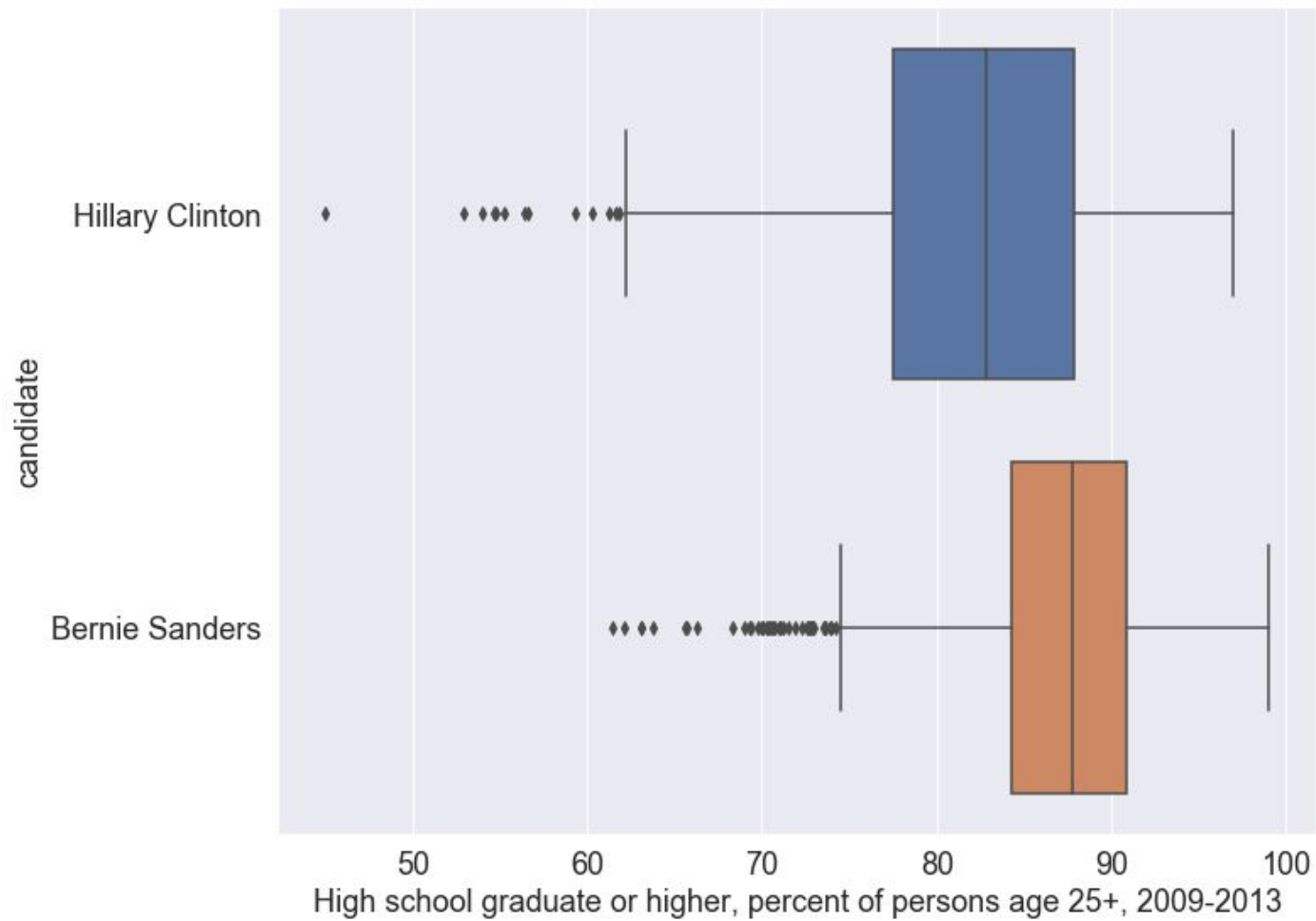
Зависимость от расы



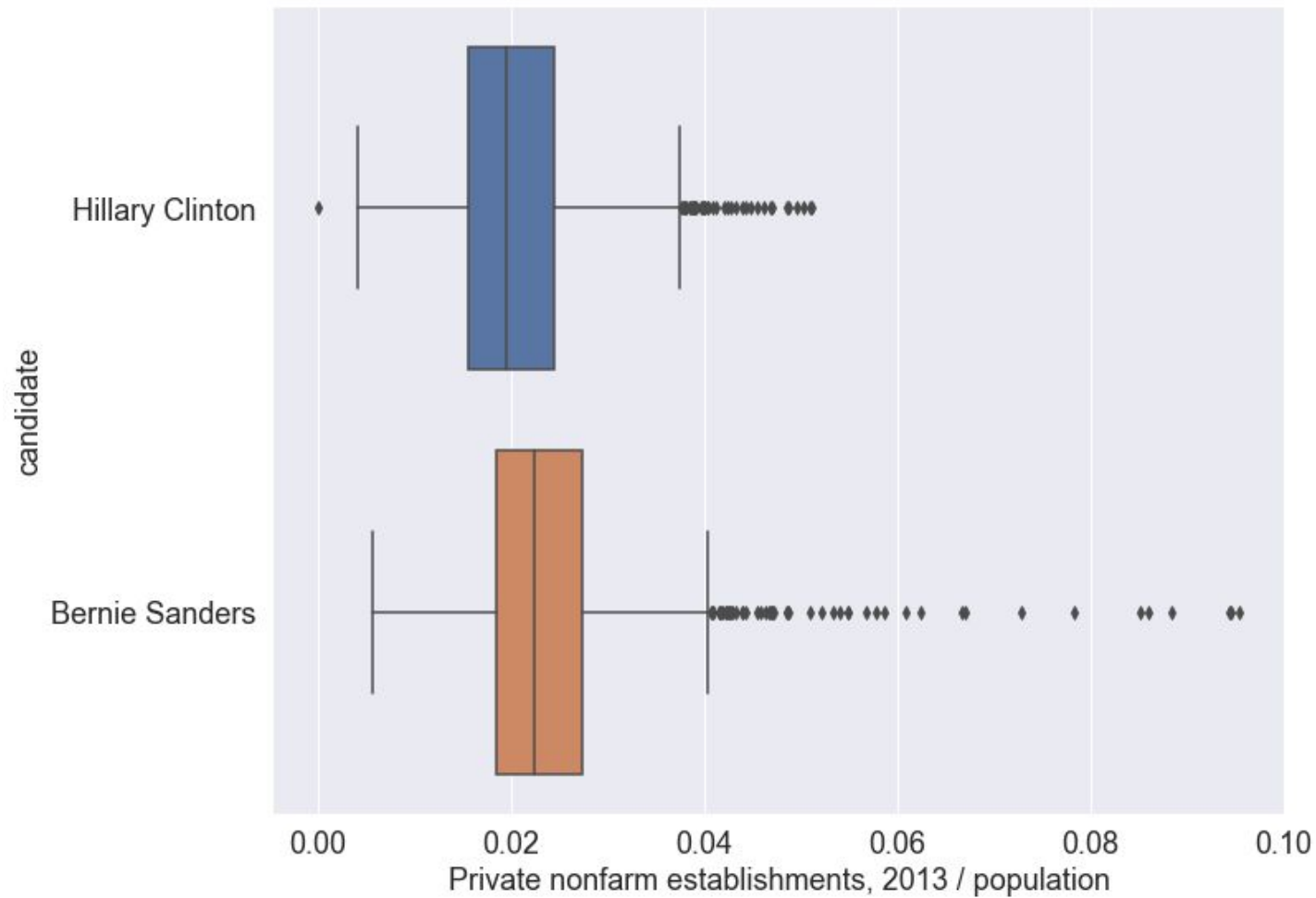
Однако...



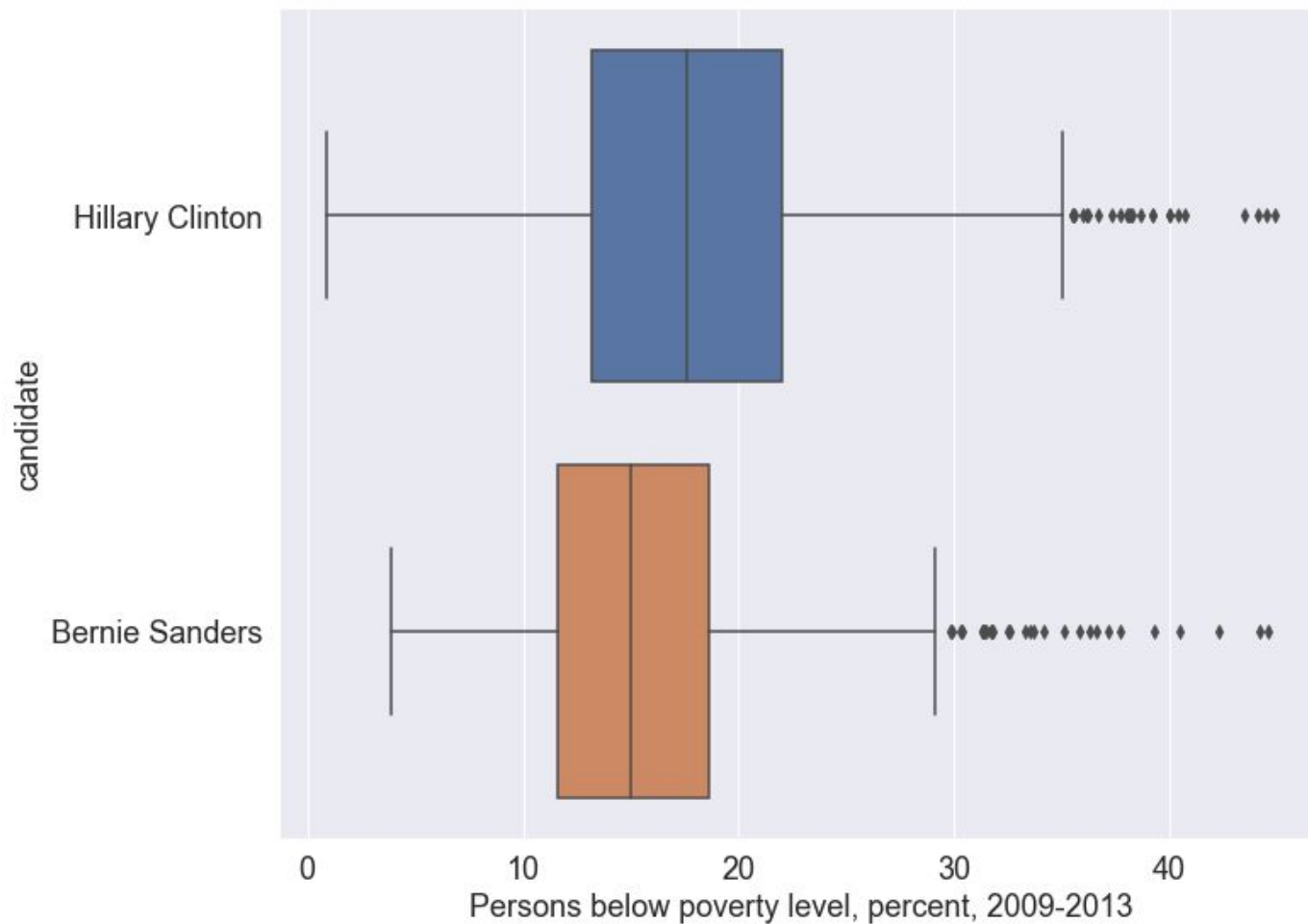
Образование



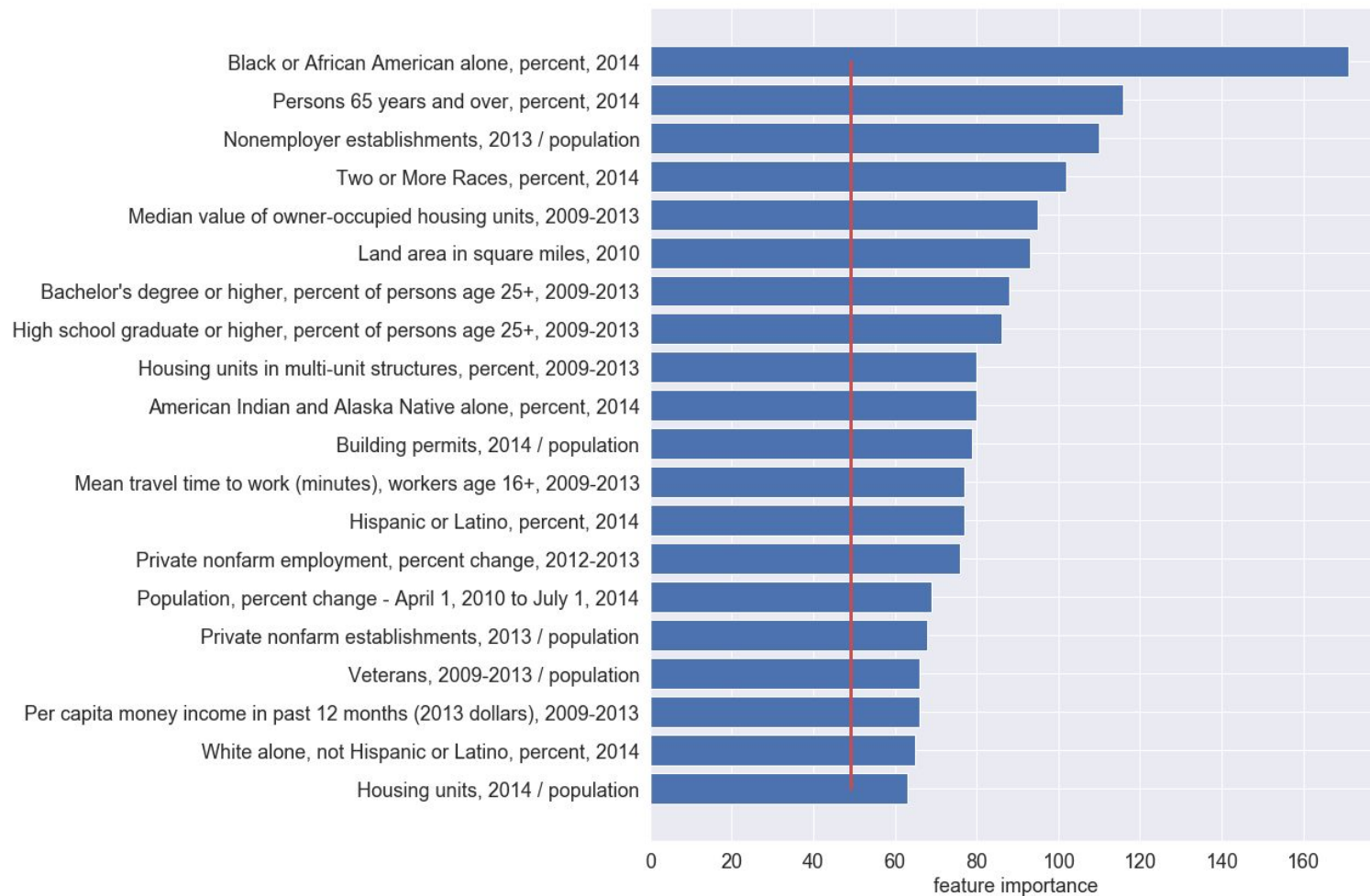
Несельскохозяйственные организации



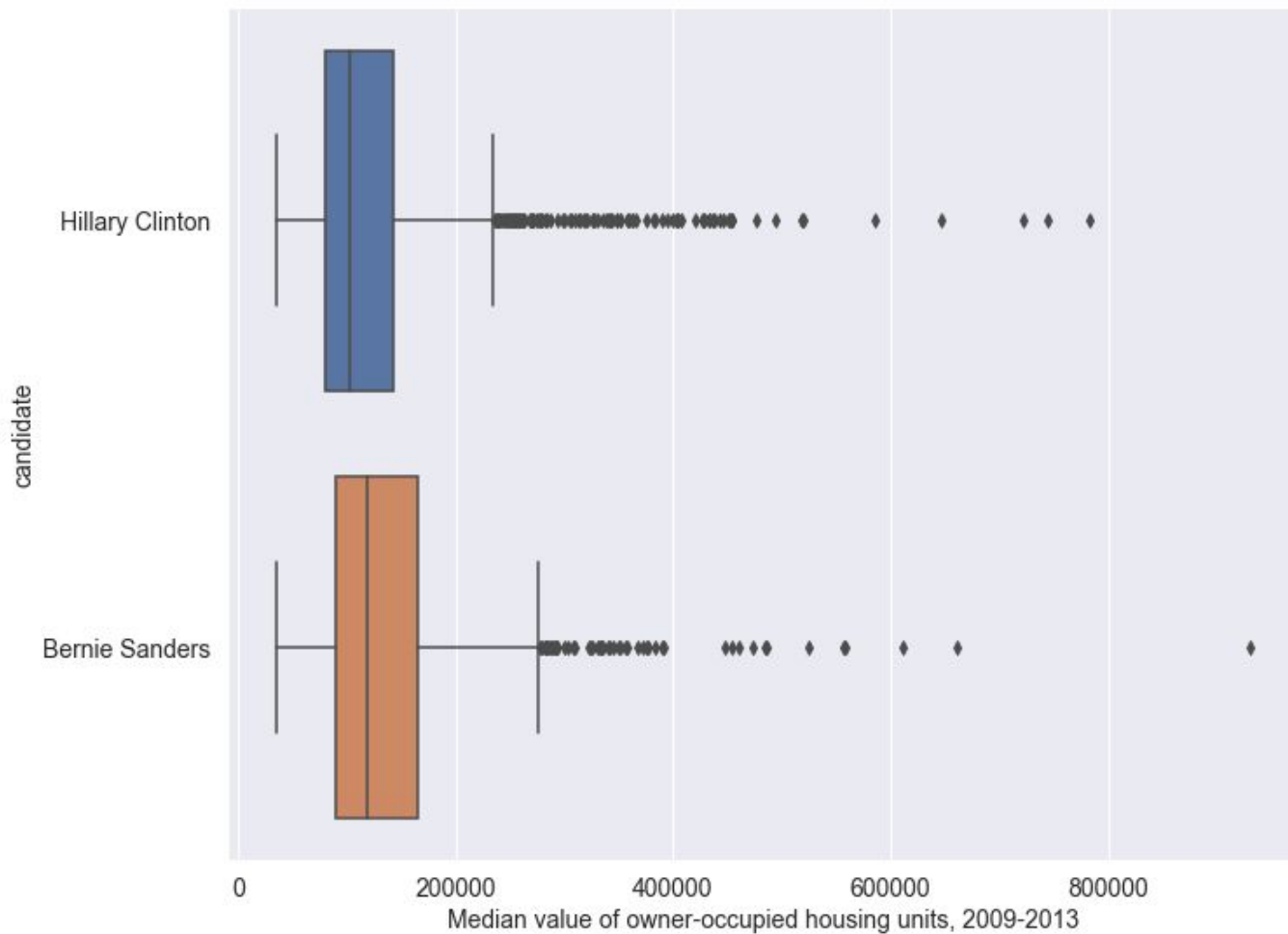
Население ниже уровня бедности



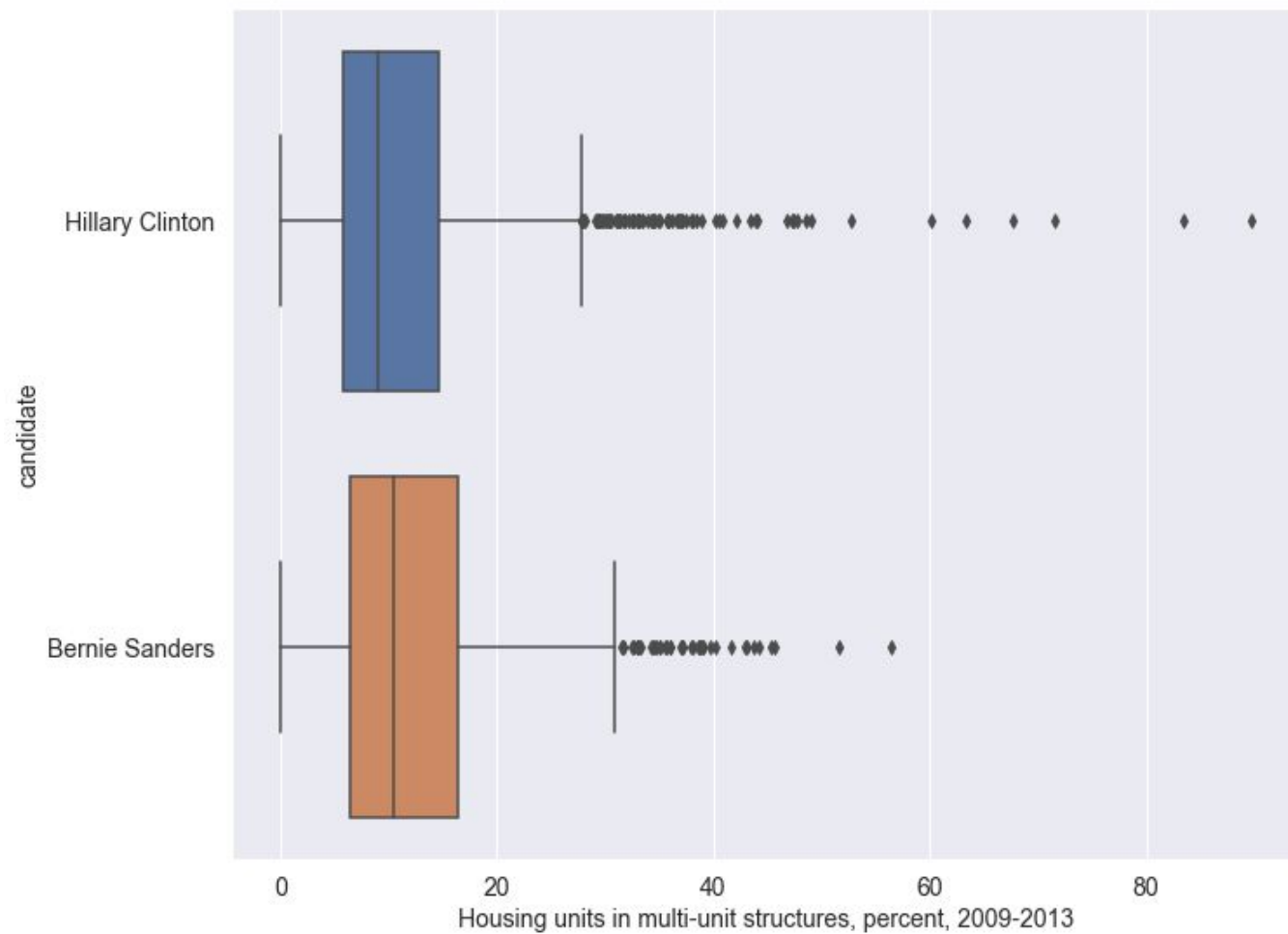
На что обращает внимание lightgbm



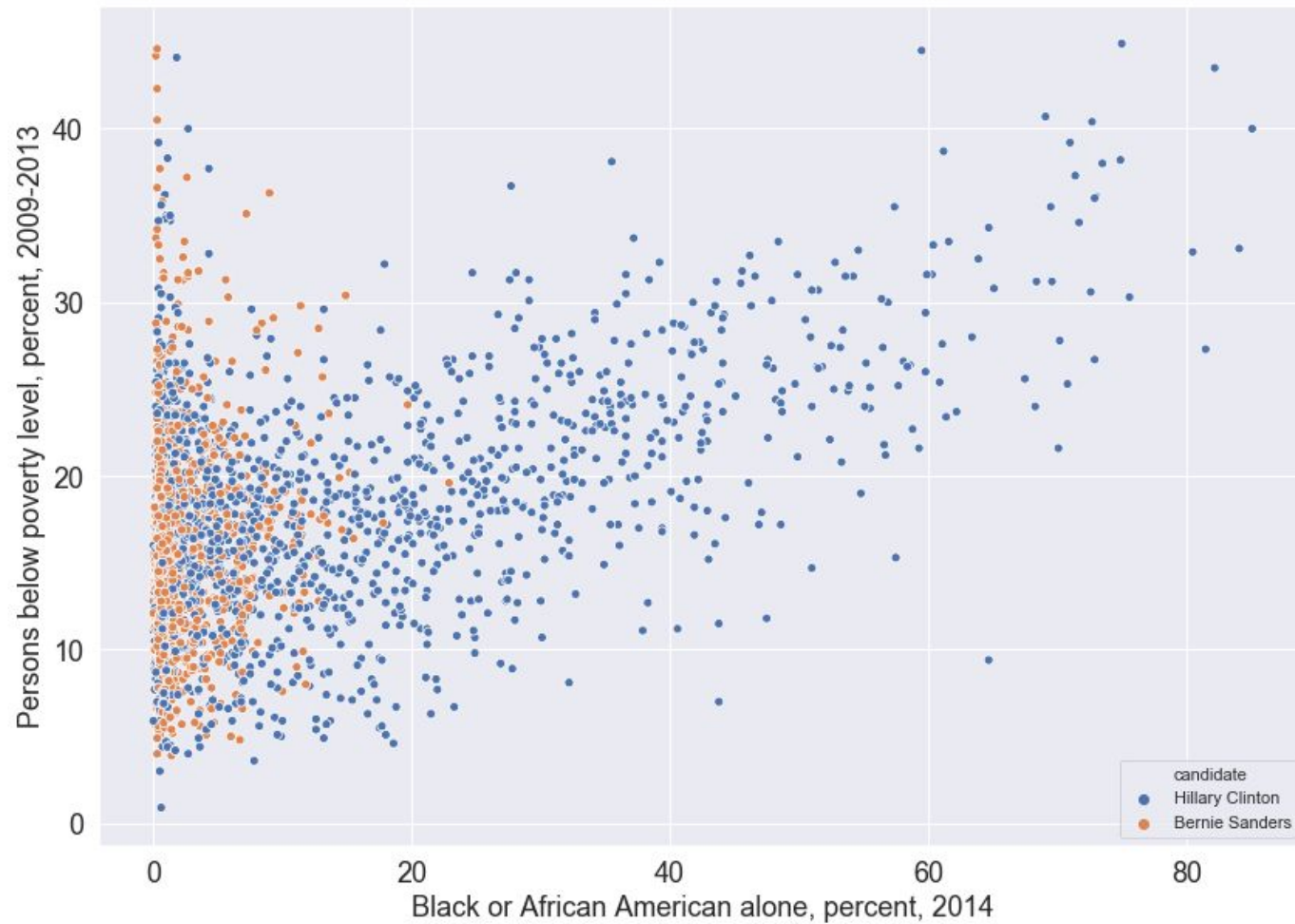
Медианная стоимость жилья



Доля многоквартирных домов



Уровень бедности от доли чернокожего населения



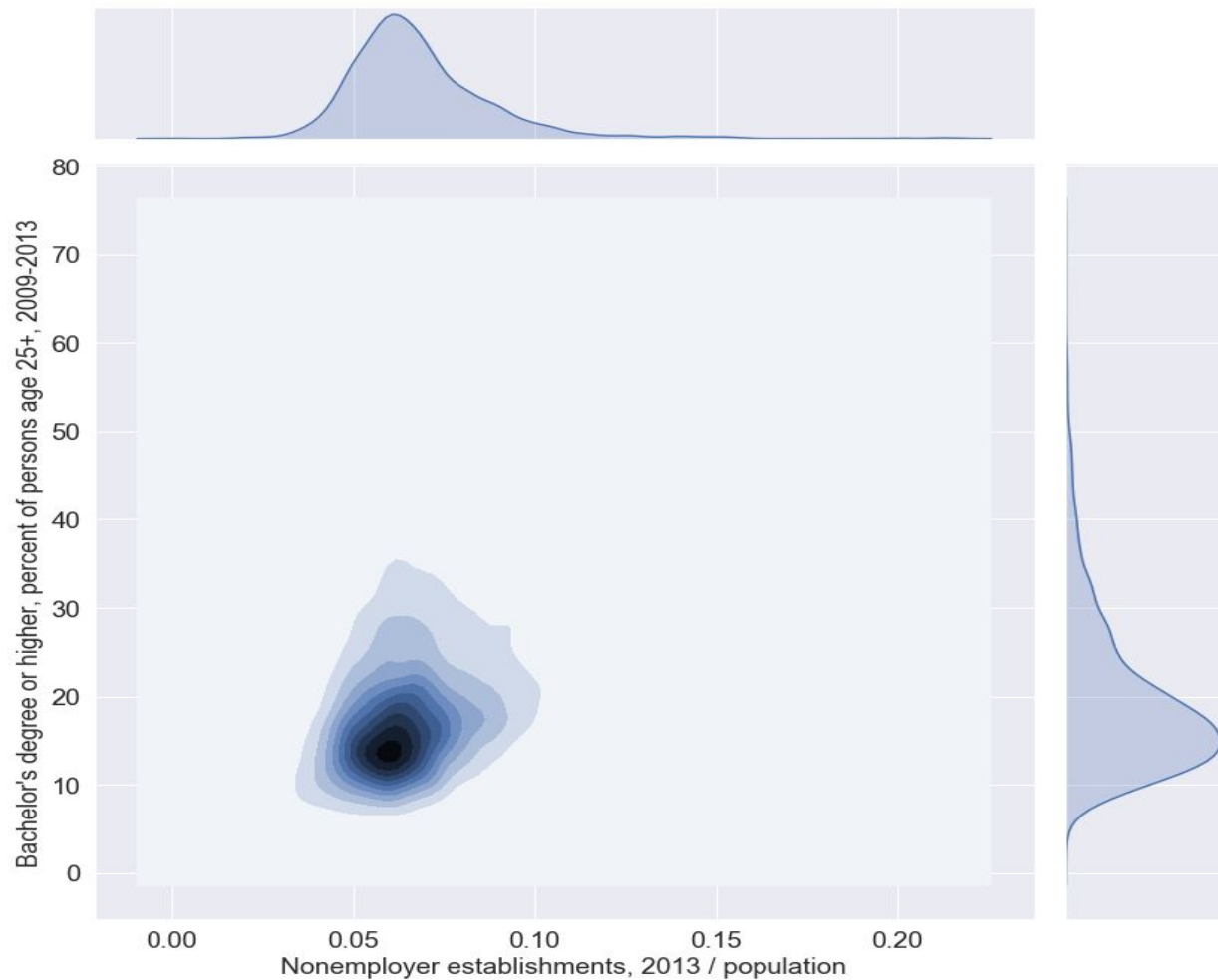
Зависимость среднего размера семьи от доли белых



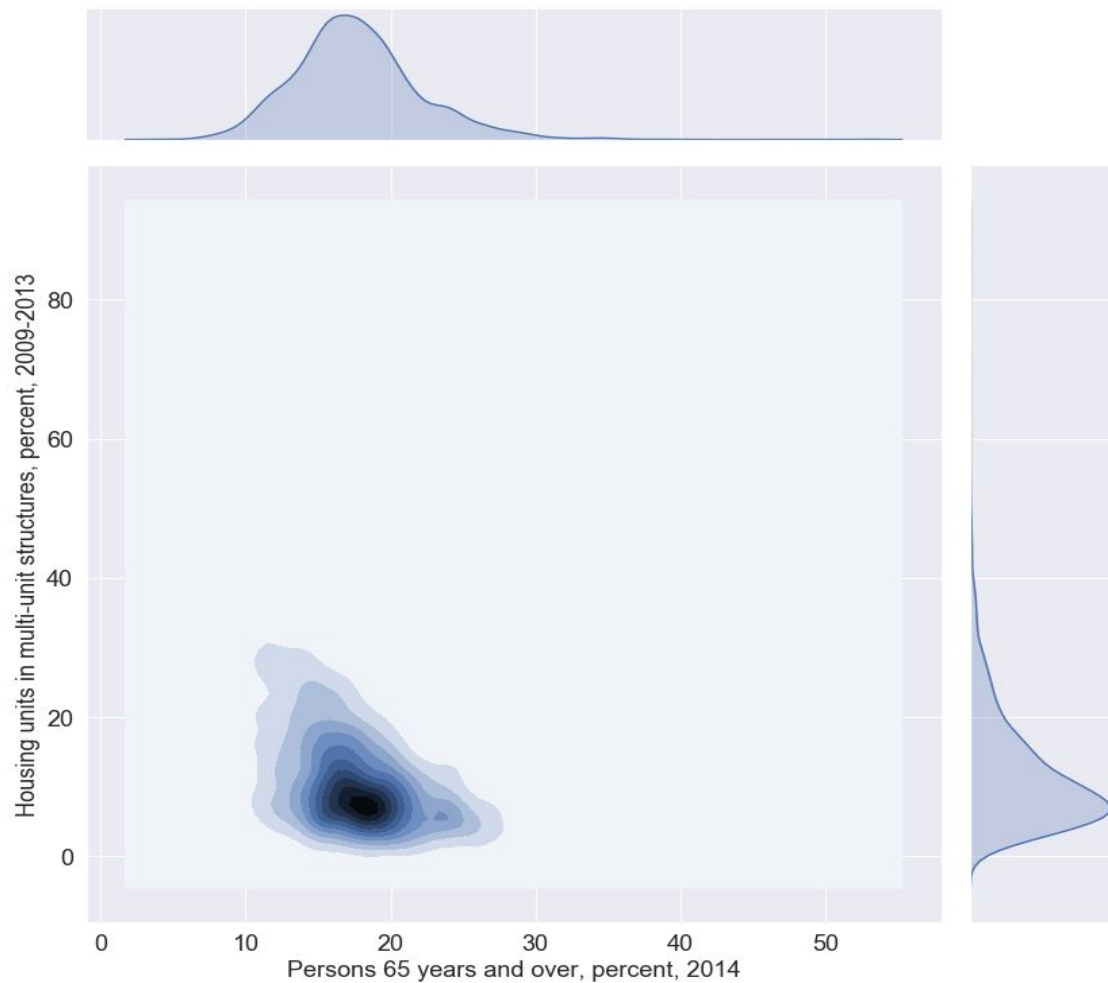
Зависимость уровня бедности от средней образованности



Зависимость доли людей с высшим образованием от количества организаций без наемных работников



Зависимость доли многоквартирных домов от доли пожилых людей



Спасибо за внимание