

Наша гениальная
цитата

DataSad

УЧАСТНИКИ

Макаров Никита
Программист, капитан

Смирнов Максим
Программист

Разин Денис
Экономист

Биккинина Ксения
Программист

Цыганова Елена
Дизайнер

Пухватуллин Марсель
Программист

СТРУКТУРА ДАННЫХ

3871 строка

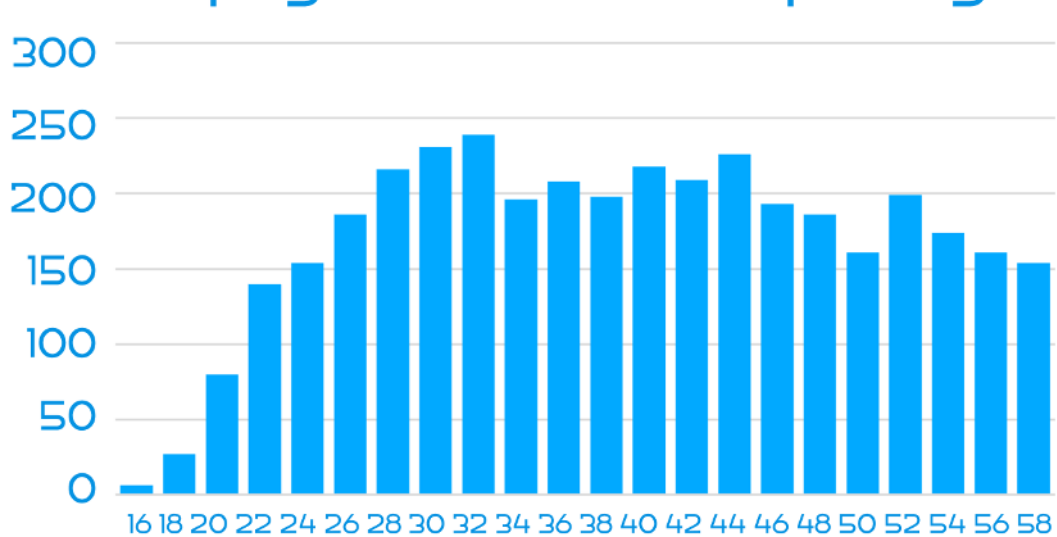
4 образования

2016 год

38 регионов

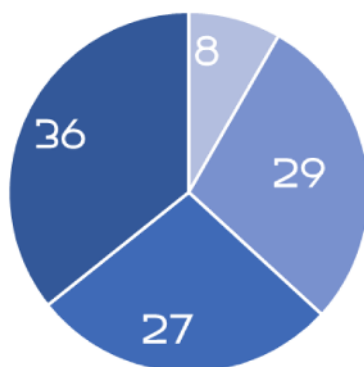
ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ

Распределение по возрасту



РАСПРЕДЕЛЕНИЕ ОБРАЗОВАНИЯ

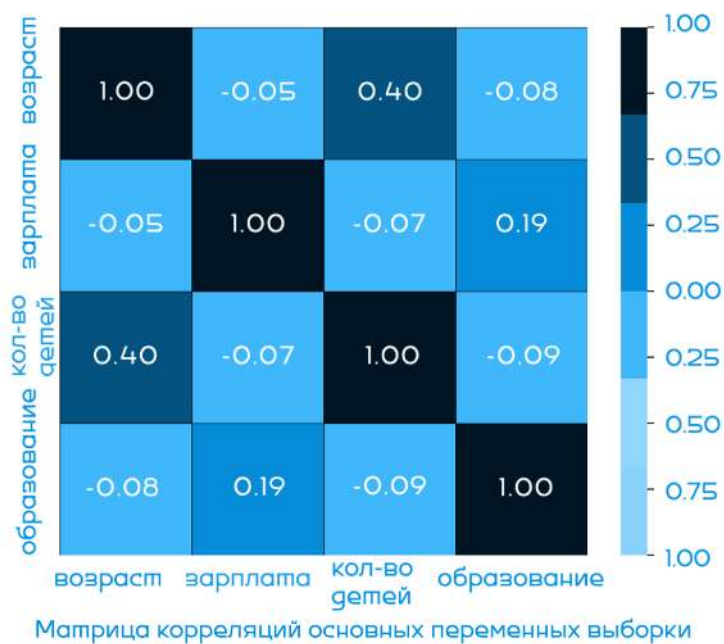
Уровень образованности населения (%)



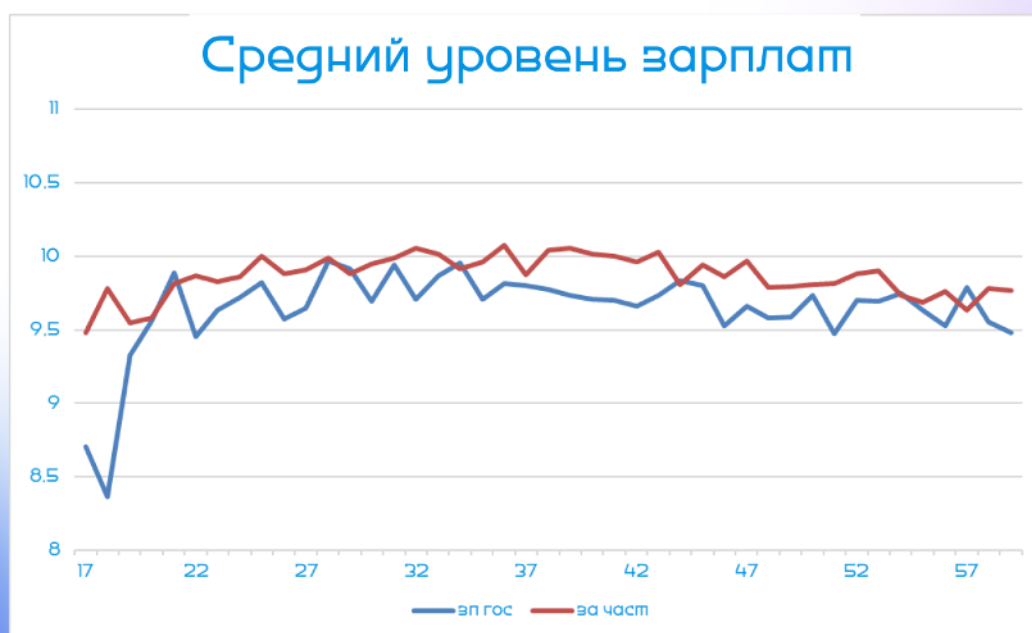
■ без образования
■ ПТУ

■ среднее образования
■ Высшее образования

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ



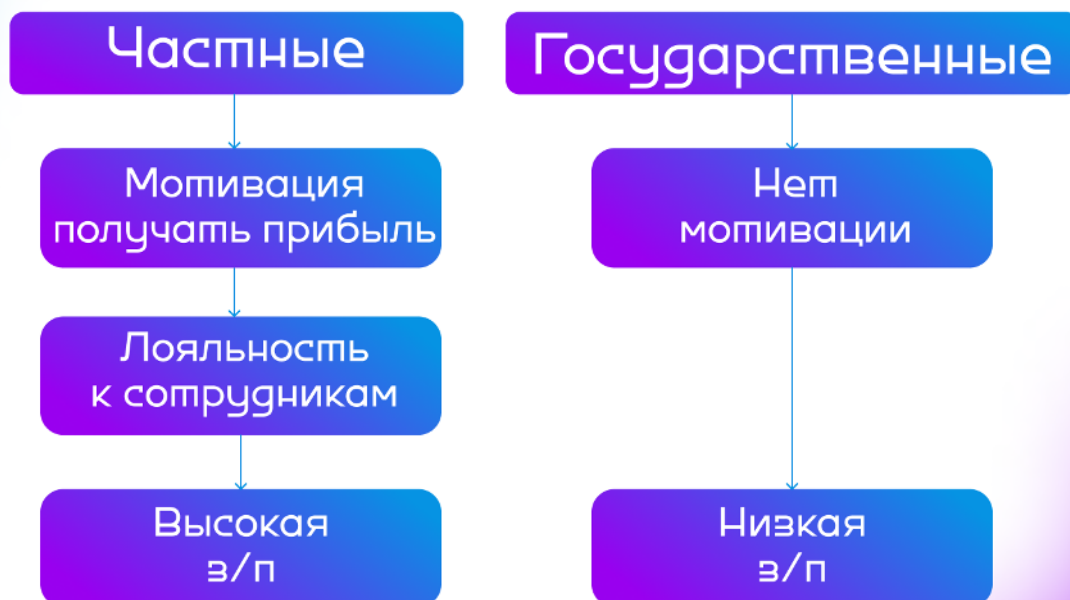
ЗАРПЛАТЫ ЧАСТНЫХ И ГОСУДАРСТВЕННЫХ РАБОТНИКОВ



ГИПОТЕЗА

Люди, работающие в частных компаниях получают в среднем больше, чем люди, работающие в госструктурах

МЕХАНИЗМ



РАБОТА С ДАННЫМИ

	МИН.	МАКС.	КОЛИЧЕСТВО
Группа 1	9.127	9.561	11
Группа 2	9.586	9.811	12
Группа 3	9.880	10.193	12
Группа 4	10.344	10.454	2

Группировка по зарплатам

РАБОТА С ДАННЫМИ

Результат удаления
выбросов:

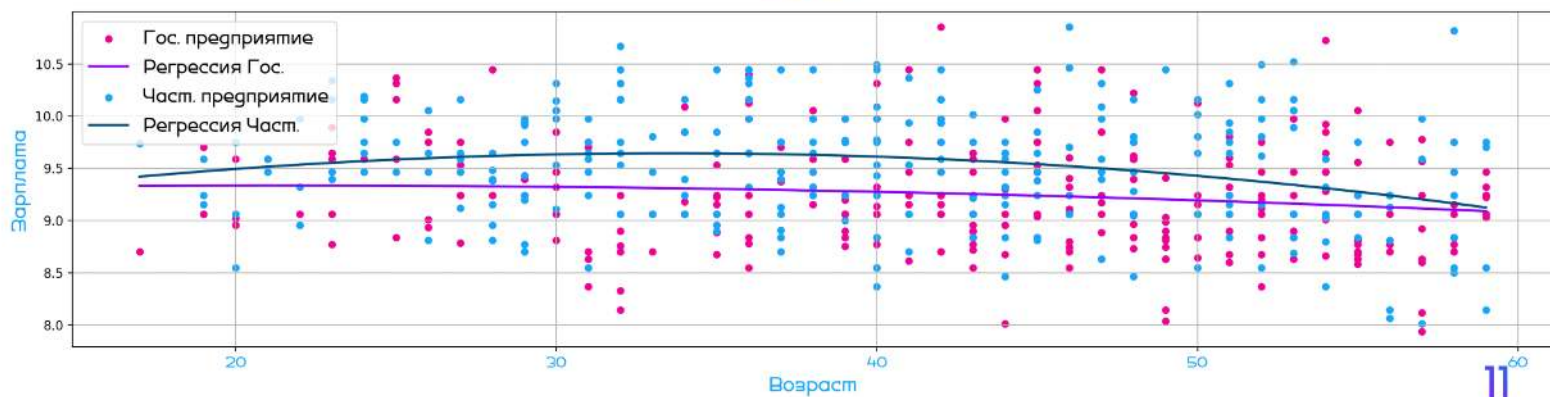
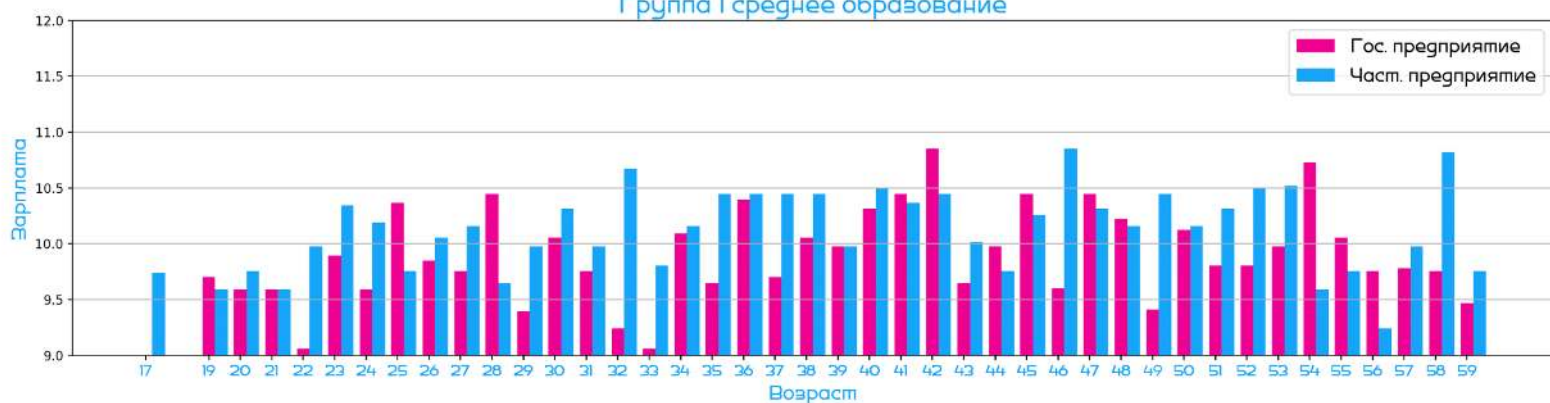
1 группа: 836 → 810

2 группа: 988 → 951

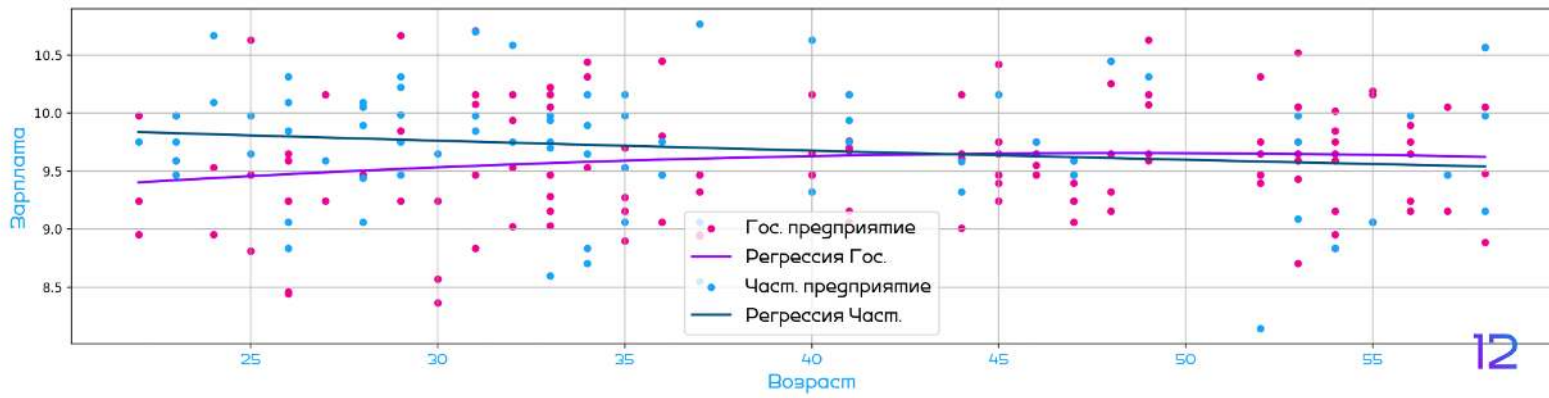
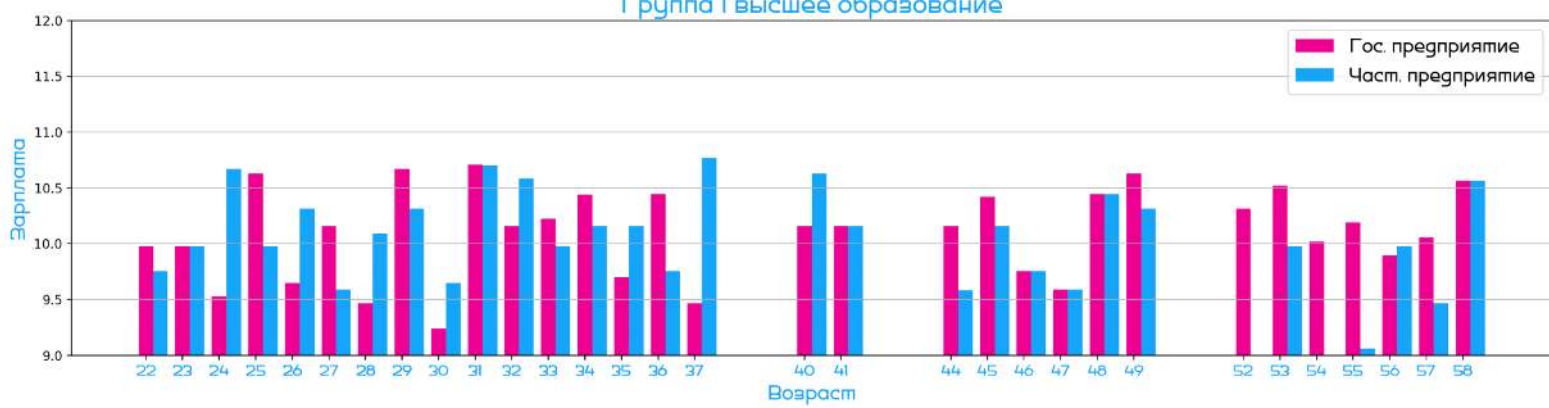
3 группа: 1163 → 1124

4 группа: 412 → 373

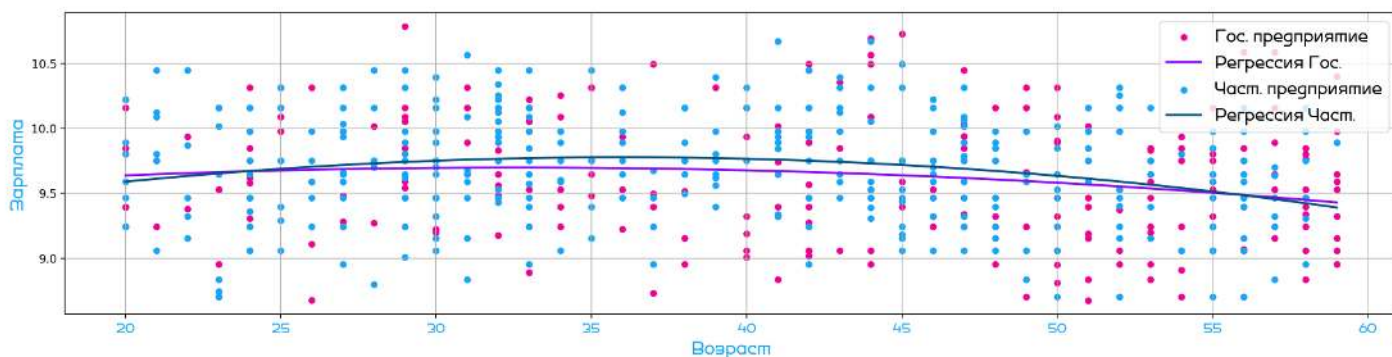
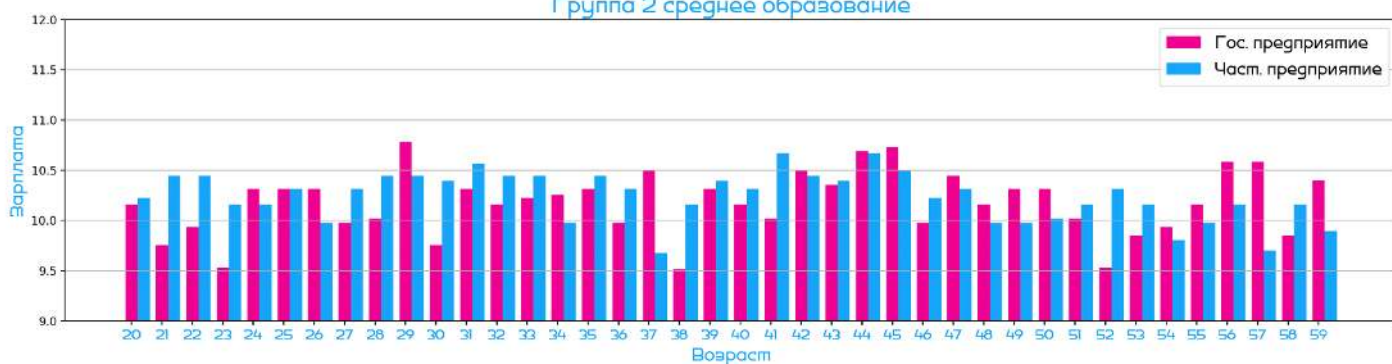
Группа 1 среднее образование



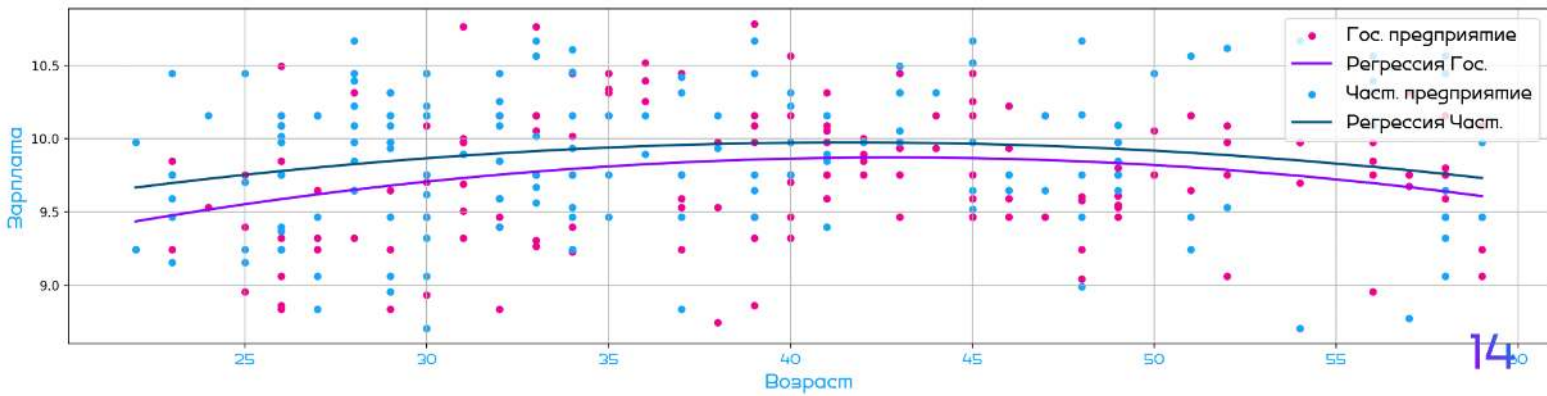
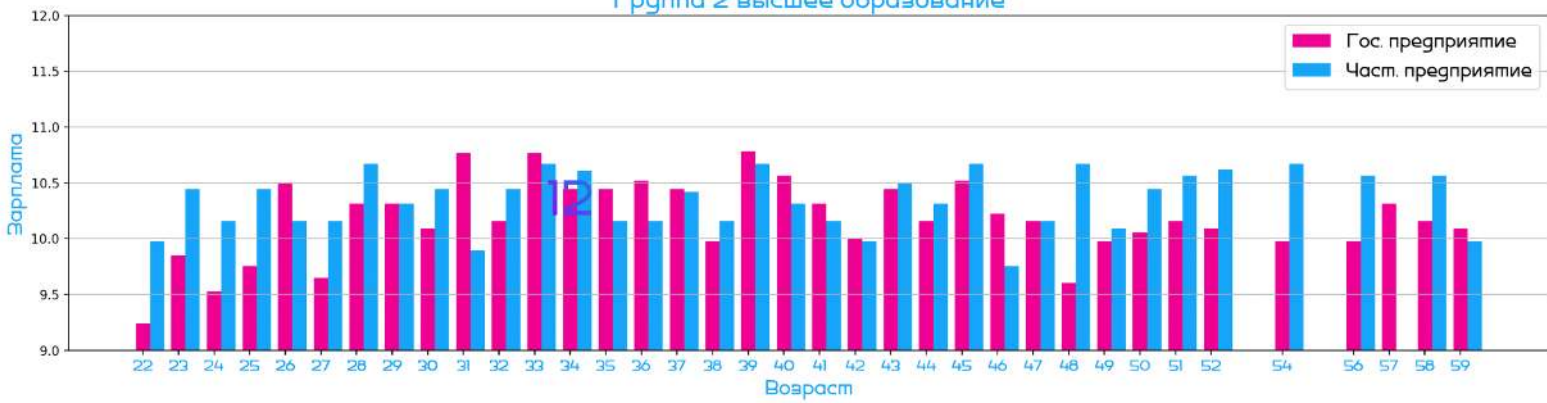
Группа 1 высшее образование



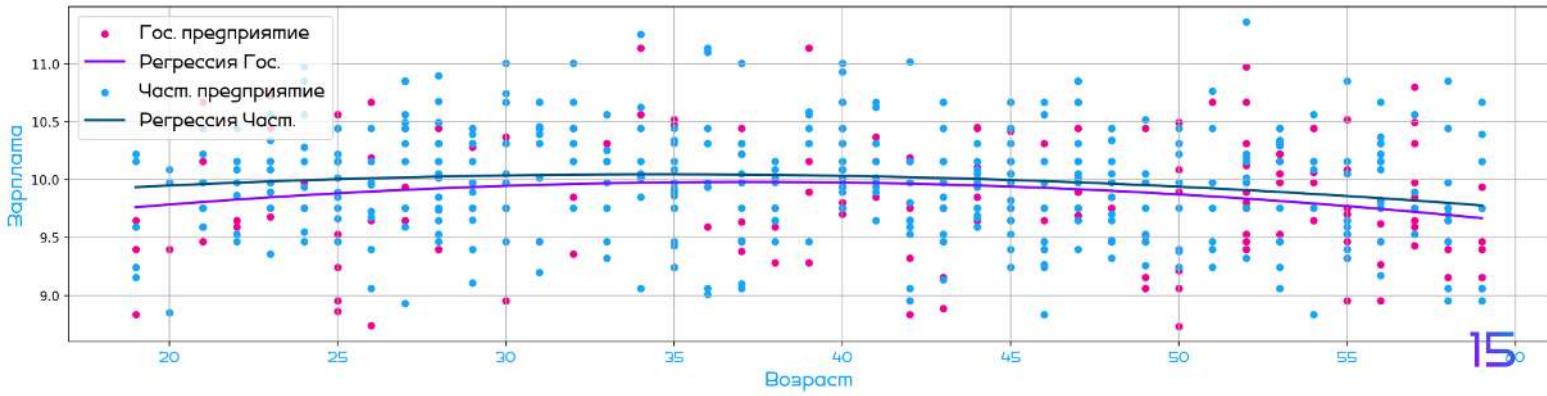
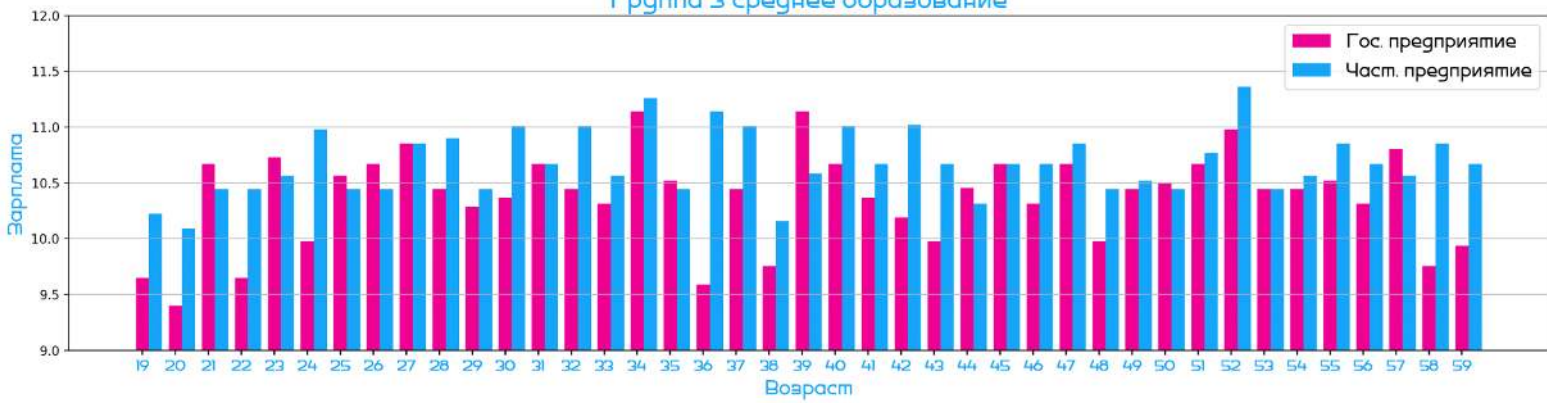
Группа 2 среднее образование



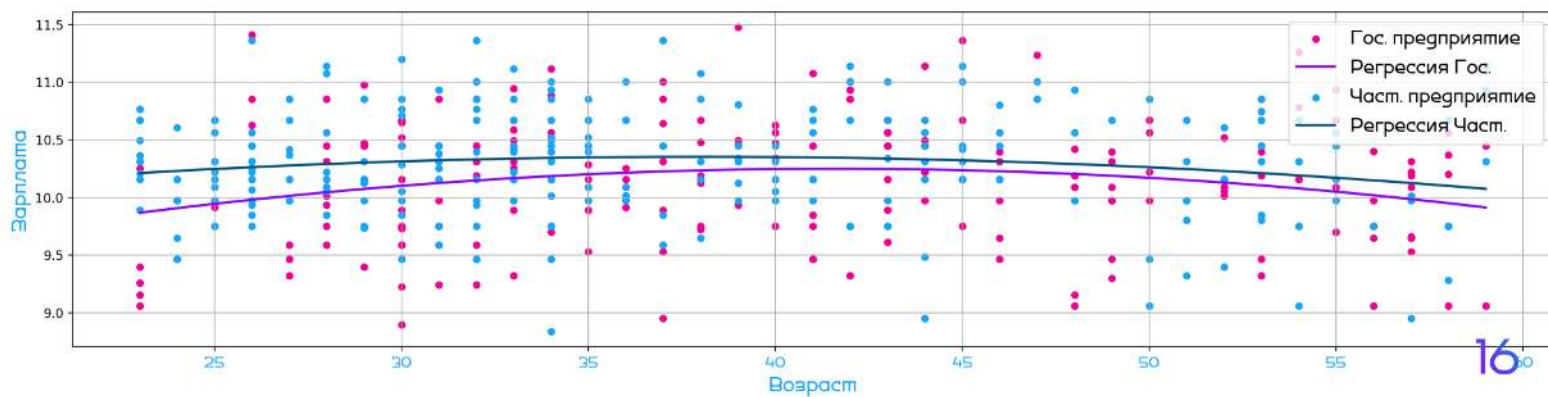
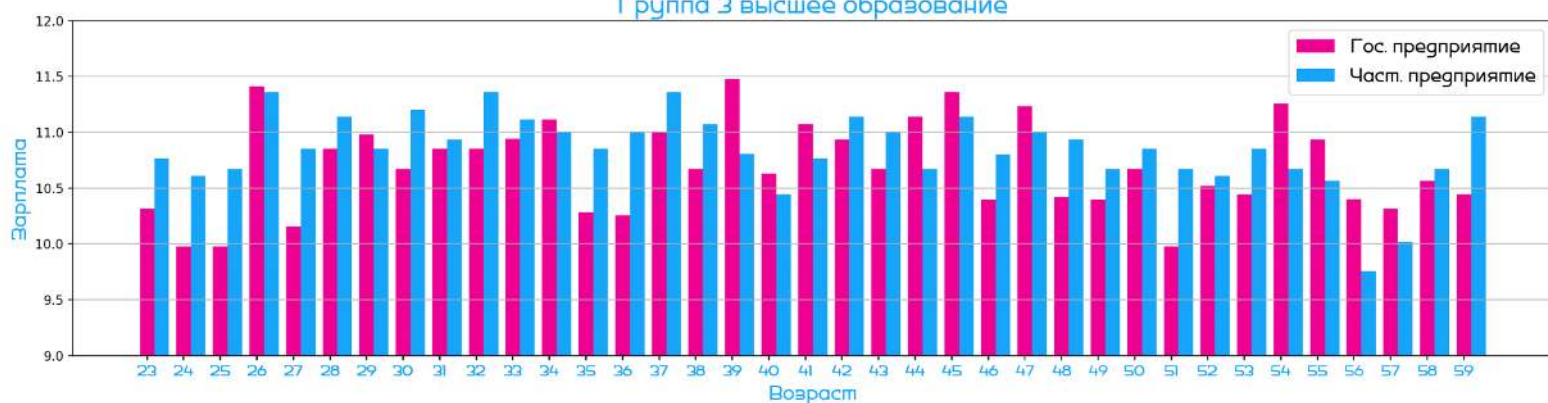
Группа 2 высшее образование



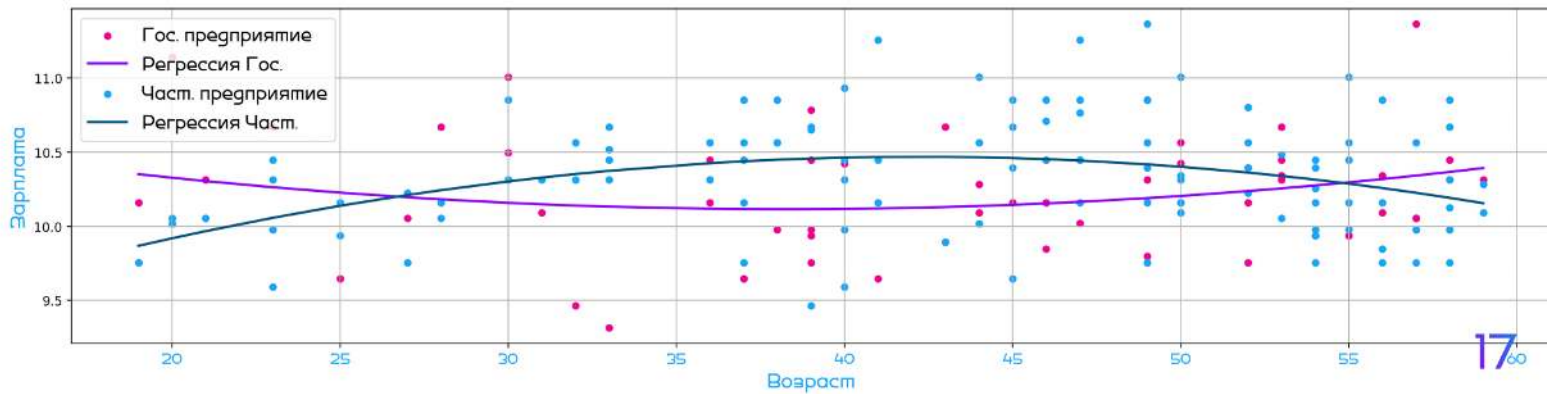
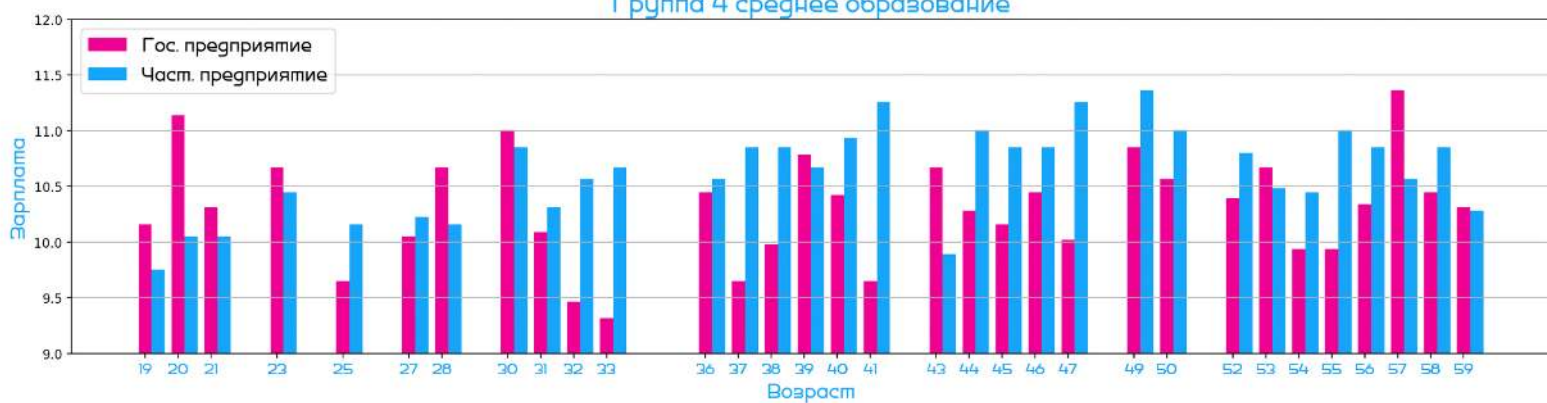
Группа 3 среднее образование



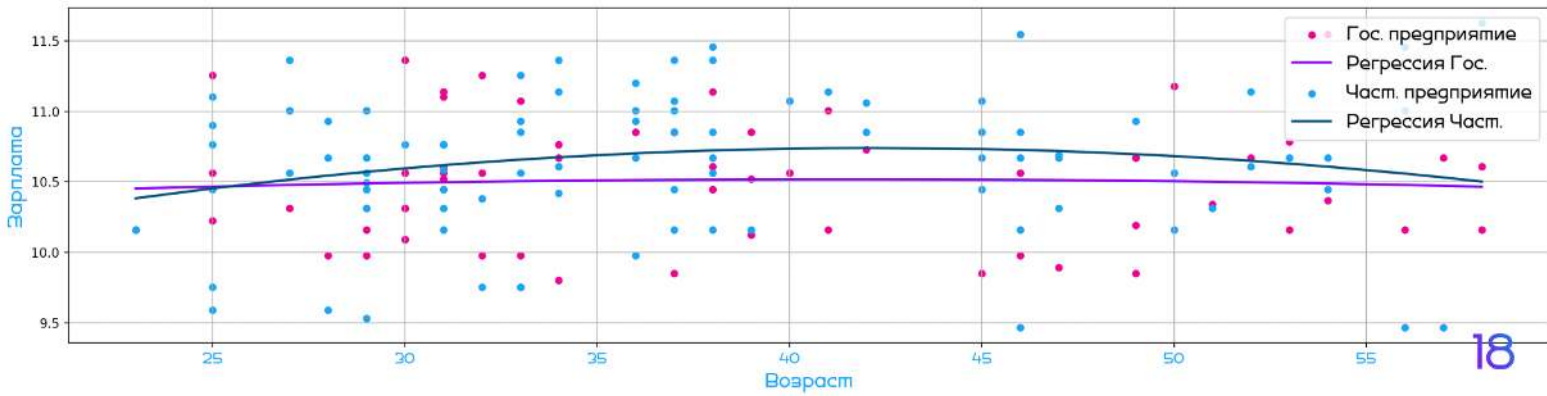
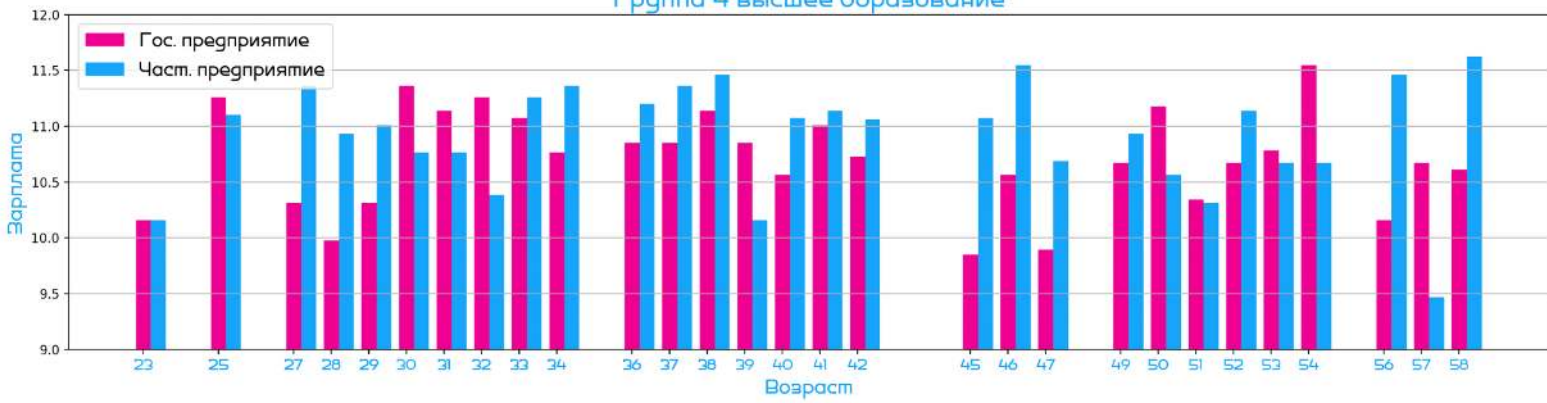
Группа 3 высшее образование



Группа 4 среднее образование



Группа 4 высшее образование



МАТЕМАТИЧЕСКАЯ ЧАСТЬ

	группа 1			группа 2			группа 3			группа 4		
	стат	p-value	кол-во	стат	p-value	кол-во	стат	p-value	кол-во	стат	p-value	кол-во
обр 1-2	5.94196	8.16996	450	2.36019	0.01904	250	2.73719	0.00672	300	3.30356	0.00151	350
обр 3	0.4682	0.64087	450	1.31178	0.19131	250	2.90255	0.00407	300	3.34222	0.00131	350

Математическая подтверждение теорем

Группы

1-2 уровень образования

3 уровень образования

частный

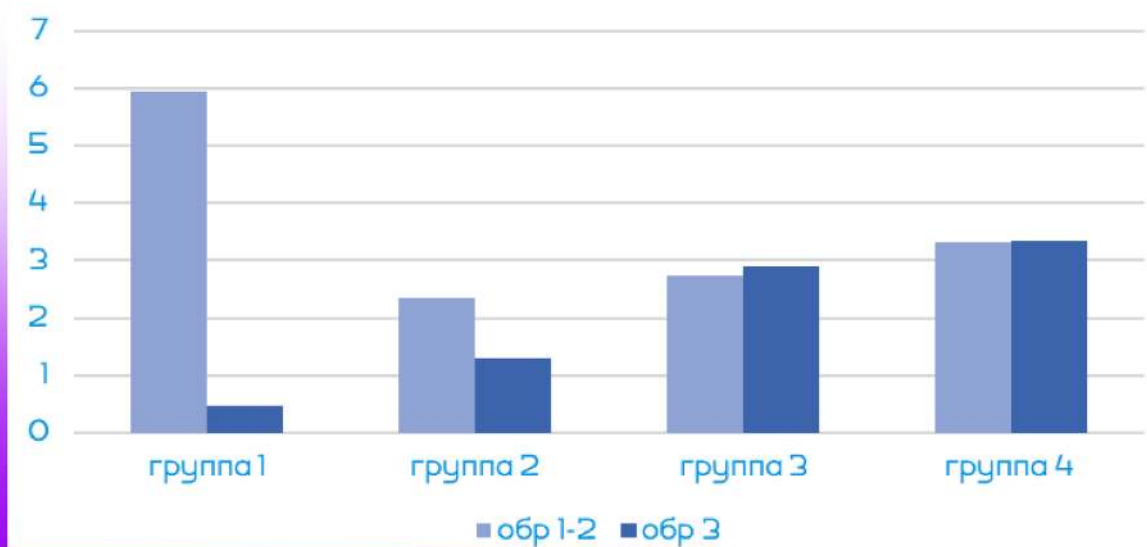
государственный

частный

государственный

МАТЕМАТИЧЕСКАЯ ЧАСТЬ

Дельта значений



ИНТЕРПРЕТАЦИЯ

✓ Гипотеза подтвердилась

Но не всех группах регионов

✓ Произведена проверка устойчивости

ОГРАНИЧЕНИЯ И ПЕРСПЕКТИВЫ

Ограничения

- Данные только по России
- 2016 год
- Небольшое кол-во переменных

Перспективы

- Для работы с молодежью
- Социальные проекты

Наша гениальная
цитата

DataSad