

Домашнее задание 2
Дедлайн: 23.59 21 октября 2025

Для проверки решение ДЗ в файле формата .ipynb
необходимо отправить на следующие почты:
dskuznetsova_2@edu.hse.ru; darysalnikova@yandex.ru

Тема письма в формате: Фамилия ИО студента, ДЗ2

Задание выполняется на данных [hwdata.dta](#). Это данные из исследования *Dower, Paul Castaneda, Evgeny Finkel, Scott Gehlbach, and Steven Nafziger. "Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms." American Political Science Review 112, no. 1 (2018): 125–47.* Статья доступна по [ссылке](#).

Авторы обращались к вопросу, как вероятность коллективного действия народа, исключительно из процесса принятия решений, влияет на склонность элит к перераспределению благ в пользу народа (финансирования публичных благ)? Показателем «изменение в количестве сельских школ с 1860 до 1880 гг. на душу сельского населения уезда» используется для измерения объема финансирования публичных благ. Частота крестьянских выступлений в годы, предшествующие Земской реформе, – показатель способности к коллективному действию. Ниже представлено краткое описание переменных:

ch_schools_pc	Изменение в количестве сельских школ с 1860 до 1880 гг. на душу сельского населения уезда
afreq	Доля лет между 1851 и 1863 гг., в которые были зафиксированы крестьянские выступления
nozemstvo	Бинарная переменная: Единицей закодированы уезды тех губерний, в которых в результате реформы 1864 года земства созданы не были, 0 – в противном случае.
distance_moscow	Расстояние от Москвы до центра уезда
goodsoil	Показатель плодородности почвы
lnurban	Логарифм городского населения уезда на 1863 г.
lnpopn	Логарифм населения уезда на 1863 г.
province_capital	Бинарная переменная: принимает значение 1, если в уезде находился «столичный» город губернии, 0 – в противном случае.

Задание 1. В этом домашнем задании нас будет интересовать гипотеза авторов, согласно которой, там, где элиты не пошли на уступки демократизации, большая способность народа к коллективному действию приведет к большему перераспределению. Ожидается, что большее финансирование общественных благ в результате увеличивающейся опасности крестьянских выступлений будет осуществляться в тех губерниях, на которые не была распространена Земская реформа. Это предположение необходимо проверить.

- Создайте переменную *zemstvo* посредством перекодирования *nozemstvo* таким образом, чтобы единицей обозначались уезды тех губерний, в которых в результате реформы 1864 г.

земства были созданы, нулем – наоборот, уезды тех губерний, в которых земства не были созданы.

- Визуализируйте взаимосвязь изменения в количестве сельских школ на душу населения (*ch_schools_pc* – по оси ординат) и доли лет, в которые были зафиксированы крестьянские выступления (*afreq* – по оси абсцисс), в разных группах, а именно сформированные по признаку «наличие земства» и «отсутствие земства». Сделайте предварительный вывод на основе указанного графика.
- Далее оцените соответствующую спецификацию регрессионной модели, которая позволяет протестировать различия во взаимосвязи изменения в количестве сельских школ на душу населения (зависимая переменная) и доли лет, в которые были зафиксированы крестьянские выступления, в зависимости от наличия земства, для начала без контрольных переменных. До оценивания регрессионной модели приведите описательные статистики по необходимым переменным и объясните, почему предварительные описательные статистики имеют особую важность в контексте регрессионного анализа с переменными взаимодействия. Если считаете нужным, предложите возможные преобразования предикторов для более удобной интерпретации и примените их при оценивании модели. Проинтерпретируйте все оценки коэффициентов в регрессионной модели (как с точки зрения характера связи, так и с точки зрения значимости).
- Добавьте в регрессионную модель контрольные переменные (добавьте оставшиеся переменные из таблицы с описанием показателей). Изменились ли значимым образом результаты? Соотносятся ли результаты с предположением авторов о характере взаимного влияния Земской реформы и крестьянских выступлений?

Задание 2. Оцените регрессионную модель, в которой тестируется взаимный эффект показателя плодородности почвы (*goodsoil*) и логарифма городского населения уезда на 1863 г. (*lnurban*) на изменение в количестве сельских школ на душу населения. Остальные показатели из таблицы с описанием переменных используйте как контрольные. Как изменяется предельный эффект плодородности почвы и его значимость в зависимости от «условия» (показателя городского населения)? Приведите необходимые расчеты и представьте соответствующий график зависимости предельного эффекта *goodsoil* на *ch_school_pc*. Проинтерпретируйте как предельный эффект, так и его значимость.

Задание 3. Бонусное задание (выполняется по желанию): можно получить до 2-ух дополнительных баллов к семинарской активности

В статье Hainmueller et al. (2018) «How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice» приводятся ошибки, допускаемые довольно часто исследователями при оценивании линейных регрессионных моделей с переменными взаимодействия. Ознакомьтесь с данной [статьей](#) и порассуждайте, какие слабые места можно выделить в оцененных в рамках этого домашнего задания моделях. Свой ответ поясните и подкрепите свой ответ эмпирическими свидетельствами.