

**Домашнее задание №1 (сдаваемое)**  
**Дедлайн: 23:59 19 февраля**

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, исключенного 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.

1. Оцените линейную регрессионную модель (модель *m1*), в которой изменение в количестве сельских школ на душу населения является зависимой переменной, все остальные представленные в таблице переменные – предикторы. Здесь и далее при оценивании используйте робастные стандартные ошибки, состоятельные в условиях гетероскедастичности (тип ошибок – НСЗ). Проинтерпретируйте полученные результаты: какой характер взаимосвязи объясняющих переменных и отклика, также прокомментируйте значимость
2. Проинтерпретируйте значение  $R^2$ . Чем отличается мера скорректированного  $R^2$  ( $R^2$  adjusted) от  $R^2$ ?
3. Протестируйте, есть ли мультиколлинеарность в модели *m1*, насколько сильно мультиколлинеарность отражается на результатах. Используйте как визуальные диагностики, коэффициенты корреляции, так и значения VIF. Проинтерпретируйте полученные результаты и сделайте вывод
4. Приведите свидетельства гетероскедастичности для модели *m1*:
  - Используйте визуальные способы определения гетероскедастичности. Сделайте вывод
  - Используйте формальный тест Бреуша—Пагана. Сделайте вывод
5. Согласно гипотезе авторов, 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. Для этого выполните последующие шаги:
  - 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. Не забудьте добавить в модель контрольные переменные (все остальные переменные в таблице с описанием данных). 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)
6. 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 изменится 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 график: как изменения предельного эффекта, так и значимость предельного эффекта