

## Домашнее задание

### Deadline: 23:59 9 марта

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.*

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:

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

**Задание 1.** Оцените регрессионную модель, в которой изменение в количестве сельских школ на душу населения является зависимой переменной, все остальные представленные в таблице переменные – предикторы.

Проинтерпретируйте полученные результаты: какой характер взаимосвязи объясняющих переменных и отклика, прокомментируйте значимость. Есть ли теоретические основания полагать, что данная модель страдает от гетероскедастичности? Опишите эти возможные основания (пока без статистических тестов, визуализаций, только на основании Ваших теоретических ожиданий).

**Задание 2.** Протестируйте, есть ли гетероскедастичность в модели.

1. Используйте визуальные способы определения гетероскедастичности. Сделайте вывод.
2. Используйте формальный тест Бреуша—Пагана. Сделайте вывод.
3. Есть ли основания полагать, что вариация ошибок зависит от одной из объясняющих переменных? Протестируйте такую монотонную зависимость посредством теста Goldfeld-Quandt.
4. Если Вы обнаружили гетероскедастичность, переоцените стандартные ошибки (используйте тип ошибок *НСЗ*). Изменились ли значимо результаты?

**Задание 3.** Протестируйте, есть ли мультиколлинеарность. Используйте как визуальные диагностики, так и значения VIF. Проинтерпретируйте полученные результаты и сделайте вывод.