

RLMS — это набор данных по индивидам и домохозяйствам России. В этом домашнем задании вам предстоит провести своё маленькое и гордое первое исследование.

1. Запишитесь на волну RLMS по ссылке [http://tiny.cc/hse\\_master\\_ea\\_ha\\_rlms](http://tiny.cc/hse_master_ea_ha_rlms).

На каждую волну может зарегистрироваться не более трёх человек. При этом разрешается выполнять работу как индивидуально, так и в группе. Скажем, на волну 22 может записаться Андрей, выполняющий работу в одиночку, и команда Бориса и Владимира, выполняющих работу вместе.

2. Скачайте и импортируйте данные выбранной вами волны.

Данные RLMS доступны по ссылке <https://www.hse.ru/rlms/spss>.

3. Ознакомьтесь с доступными переменными.

Разумно посмотреть документацию на официальном сайте. Будьте осторожны, в наблюдениях много пропусков!

4. Сформулируйте исследовательский вопрос, который можно изучить с помощью этих данных.

Достаточно сформулировать вопрос о наличии статистической связи между переменными. Стоит осознавать, что обнаружение причинно-следственных связей — очень трудная задача, браться за неё мы не требуем, хотя и не запрещаем.

В постановке вопроса в качестве зависимой переменной должна использоваться количественная переменная. Среди предикторов должно быть минимум две количественных переменных и минимум одна факторная.

5. Визуализируйте используемые переменные и приведите описательные статистики.

Дайте краткую характеристику используемым переменным. Используйте гистограммы и другие способы визуального представления информации.

6. Оцените две регрессионных модели: простую и более общую сложную.

Для сложной модели проинтерпретируйте коэффициенты и постройте доверительные интервалы для них. С помощью теста сделайте выбор между простой и сложной моделью.

7. Для отобранной модели и выдуманного нового наблюдения постройте прогноз и предиктивный интервал для прогноза.

8. Описанные выше пункты являются минимальным требованием. Можно сделать больше! Удачи :)

9. Работу следует представить в виде отчёта в pdf формате. В начале работы должен идти текст с графиками, в конце работы в качестве приложения должен идти код. Общий объем текста (без приложений) должен составлять **не более 10 страниц**.

Дедлайн сдачи - 27 октября 2021, 20:59. До указанного времени файл в формате pdf должен быть загружен по ссылке <https://www.dropbox.com/request/qvCEbydgb2UD8m1FwUuX>.