

Домашнее задание 3. Знакомство с RLMS

14 ноября 2012 г.

1. Прочитайте про RLMS, <http://www.hse.ru/rlms/>

Посмотрите описание проекта. Прочитайте вестник RLMS, чтобы иметь представление о том, какие исследования можно строить на основе RLMS.

2. Скачайте любую волну RLMS по своему выбору. Скачайте описание переменных.

Прочитайте описание переменных. Там их больше тысячи. Попадают довольно прикольные. Мне нравится pc9.6.5a, «У Вас есть GPRS навигатор?»

3. Загрузите данные в R.

Данные RLMS выложены на сайте в формате SPSS. SPSS это потихоньку погибающий статистический пакет для домохозяек. Для чтения формата .sav в таблицу данных R можно сделать так

```
library(foreign)
options(width = 20)
file.name <- "/home/boris/downloads/r20hall23c.sav"
h <- read.spss(file.name, to.data.frame = TRUE)

## Warning: /home/boris/downloads/r20hall23c.sav: Unrecognized record type 7, subtype 18 encountered
## in system file
## re-encoding from CP1251

cat("ffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff\n",
    fill = TRUE)

## ffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff
```

Первая команда, library(foreign), подгружает библиотеку R, в которой содержатся команды для чтения вражеских форматов, spss, stata, etc

Описания переменных при этом также загружаются в таблицу данных. Можно их выделить в отдельный вектор и прочитать, например, про переменную pc9.631a.

```
var.labels <- attributes(h)$variable.labels
var.labels["pc9.631a"]

## NULL
```

4. Выберите любую количественную переменную в качестве зависимой и несколько переменных в качестве объясняющей.

Цель этой домашки скорее ознакомится с наличием мониторинга RLMS, поэтому можно не сильно заморачиваться с этим этапом. Хотя в реальности тут-то всё самое интересное и начинается. За оригинальные гипотезы будут плюшки.

5. Опишите выбранные переменные.

Постройте симпатичные графики. Посчитайте описательные статистики. Много ли пропущенных наблюдений? Есть ли что-нибудь интересенькое?

6. Постройте регрессию зависимой переменной на объясняющие.

Проверьте гипотезу о значимости каждого полученного коэффициента. Проверьте гипотезу о значимости регрессии в целом. Для нескольких коэффициентов (двух достаточно) постройте 95%-ый доверительный интервал.

7. Напишите свои пожелания и комментарии.

Какие домашки хочется сделать? Что не ясно в курсе эконометрики? Содержательные комментарии позволяют получить бонус. Искусная лесть оценивается :)