

Домашнее задание по эконометрике для группы Э_Б2017_Э_3 №3 (часть 1), 2019-2020

I. В этом задании необходимо будет использовать данные с сайта Росстата www.gks.ru, раздел «Публикации», Регионы России. Социально-экономические показатели. Единицами наблюдений будут служить регионы (начиная с Белгородской области и заканчивая Чукотским автономным округом) для определенного Вам года (он сохраняется с предыдущего задания). Если для какого-то региона данные отсутствуют, то соответствующий регион не включается в выборку)

Вам понадобятся данные для следующих региональных показателей:

- 1) **Уровень безработицы (в %)**, можно найти в разделе «Труд»
- 2) Валовой региональный продукт на душу населения (в рублях, можно найти в разделе «Валовой региональный продукт»)
- 3) Стоимость фиксированного набора товаров и услуг (в % к среднероссийской, можно найти в разделе «Цены и тарифы»)
- 4) Удельный вес городского населения (в %, можно найти в разделе «Население»)
- 5) Доля занятого населения, имеющего высшее образование услуг (в %, можно найти в разделе «Труд»)

В качестве зависимой переменной в Ваших моделях будет использоваться переменная $Unemployment = \text{Уровень безработицы (в \%)} ,$

В качестве объясняющих переменных в Ваших моделях будут использоваться переменные:

1) $GDPpcarrrrr$ (Валовой региональный продукт на душу населения с учетом паритета покупательной способности), рассчитывается по формуле:

$(\text{Валовой региональный продукт на душу населения} * 100) / \text{Стоимость фиксированного набора товаров и услуг}$

2) $Urbanshare = \text{Удельный вес городского населения}$

3) $Higheduc = \text{Доля занятого населения, имеющего высшее образование}$

4) $WEST = Dummy$ – переменная, равная 1 для западных регионов и 0 для восточных (уже была создана при выполнении задания 2.1).

5) Дополнительные переменные, которые Вы выберете самостоятельно, например,

1) из сборника Регионы России,

2) Можно ввести дамми переменную, равную 1 для регионов, богатых природными ресурсами

3) Произведения некоторых непрерывных объясняющих переменных на дамми, например, Urbanshare*WEST и т.п.

II. А) Найдите дескриптивные статистики (min, max, выборочное среднее, среднеквадратичное отклонение) для всех выбранных переменных.

В) постройте диаграммы рассеяния зависимой переменной от всех независимых.

III. Оцените Линейную модель. Сохраните остатки регрессии.

IV. Проведите тесты Уайта, Бройша – Пагана, Голдфелда-Квандта, Глейзера на выявление гетероскедастичности возмущений. Если гетероскедастичность будет выявлена, то проведите коррекцию.

V. Постройте гистограмму распределения остатков регрессии и qqplot.

С помощью тестов Колмогорова-Смирнова, Харке-Бера, Шапиро-Уилка проверьте гипотезу о нормальности распределения случайных ошибок.

VI. Рассчитайте Стьюдентизированные остатки регрессии, постройте их график и с его помощью определите возможные наблюдения, которые являются выбросами. Если такие наблюдения будут выявлены, то оцените регрессию

а) без наблюдений, являющихся выбросами,

б) введя одну или несколько дамми переменных на выбросы,

б) R-регрессию Хубера.

VII Сравните результаты оценки регрессий, оцененных с помощью МНК (с выбросами и без), с коррекцией гетероскедастичности, R-регрессии Хубера.

Оформите результаты в виде связного текста (не забыв привести результаты оценок всех оцененных регрессий и проведенных тестов)

и отправьте учебному ассистенту Ладе Кулецкой по адресу lada.kuletskaya@gmail.com не позднее 23-59 (время) 25 января 2020. Это должны быть файлы в формате doc и pdf (оба) с названием: год.3.1.doc (или год.3.1.docx) и год.3.1.pdf, например, 2003.3.1.doc + 2003.3.1.pdf.

Время для личной беседы:

28 января в 13-00 – 14-50 или 31 января (до 17-45) или 4 февраля (с 13-00).

Записаться на конкретное время можно на семинаре 14 января или 21 января или по почте demidova@hse.ru.