

Домашнее задание по эконометрике для группы Э_Б2016_Э_3 №3 (часть 1), 2018-2019

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. 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 (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) $GDP_{regcaprrrr}$ (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 и т.д.).
- 6) Произведения некоторых непрерывных объясняющих переменных на дамми, например, $Urbanshare * WEST$.

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

В) постройте диаграммы рассеяния зависимой переменной от всех независимых. Все ли зависимости похожи на линейные?

III. Оцените Линейную модель. Проведите тест Рамсея на правильную спецификацию модели. Если гипотеза о правильной спецификации модели отвергается, то попробуйте добавить в модель квадраты (и, возможно, кубы) некоторых переменных, зависимость от которых может быть нелинейной. Проведите тест Рамсея снова.

IV. Рассчитайте VIF-ы. Существует ли для построенной регрессии проблема мультиколлинеарности? Если да, то перейдите к следующим пунктам.

V. Выберите факторы, которые должны быть исключены из уравнения регрессии, используя метод пошагового исключения незначимых переменных.

VI. Выберите факторы, которые должны быть включены в уравнение регрессии, используя метод пошагового включения переменных.

VII. Сравните результаты, полученные в пунктах III, V, VI. В качестве показателя качества подгонки регрессии используйте коэффициент множественной детерминации, скорректированный на число степеней свободы. Дайте экономическую интерпретацию полученным результатам.

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

и отправьте учебному ассистенту Алленовой Анастасии по адресу nastallen97@mail.ru. не позднее 23-59 (время) 26 января 2019. 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 17-00 - 21-00.

Записаться на конкретное время можно на семинаре 14 января или по почте

demidova@hse.ru. Студенты с одним вариантом приходят на беседу вместе, но

отвечают на разные вопросы (заболевший может отсутствовать, но и баллов за устную беседу он/она не получает).