

Семинары №4-6.

Все задания данного семинара требуется выполнять в пакете Stata.

Задание 1. Продолжение задания из семинара 3.

В совокупности данных HOUSING содержится информация о ценах и характеристиках домов, проданных в течение июля, августа и сентября 1987 года в канадском городе Виндзоре. Цель исследования состоит в получении **гедонистической ценовой функции**.

Гедонистическая цена приписывается неявной цене определенного характерного признака оцениваемого объекта (для дома, например, таким характерным признаком может быть число спален) и влияет на его продажную цену. Типичными товарами, для которых оцениваются гедонистические ценовые функции, являются компьютеры, автомобили и дома.

Гедонистическая ценовая функция описывает саму ожидаемую цену дома (или ее логарифма) как функцию множества характеристик:

- размера участка земли в собственности,
- числа спален,
- числа полностью оборудованных ванных комнат,
- числа мест в гараже,
- числа этажей,
- наличия подъездной дороги,
- наличия комнаты отдыха,
- наличия обустроенного подвального помещения,
- наличия центрального кондиционирования воздуха,
- наличия водяного отопления на газе и
- расположения в привилегированном районе.

1. Загрузите файл с данными в программу Стата.

Переменные

- Price – цена продажи дома (в канадских долларах),
- Lotsize - размер участка земли в собственности (в квадратных футах),
- Bedrooms – число спален,
- Bathrooms – число ванных,

-
- Stories – количество этажей,
 - Driveway – dummy – переменная, равная 1 при наличии подъездного пути,
 - Recroom - dummy – переменная, равная 1 при наличии комнаты отдыха,
 - Fullbase - dummy – переменная, равная 1 при наличии подвала,
 - Gashw - dummy – переменная, равная 1 при наличии газа для подогрева воды,
 - Airco - dummy – переменная, равная 1 при наличии централизованной системы кондиционирования воздуха,
 - Garager – число машино-мест в гараже,
 - Prefarea - dummy – переменная, равная 1 при расположении в престижном пригороде.
2. Оцените линейную модель (модель А), которая объясняет переменную *price* размером участка земли, числом спален, числом ванных комнат, наличием кондиционирования воздуха и константой. Интерпретируйте результаты.
 3. Протестируйте наличие мультиколлинеарности в модели. Примите соответствующие меры.
 4. Протестируйте наличие гетероскедастичности в модели. Проведите тесты Уайта, Бройша-Пагана, Голфелда-Квандта и Глейзера (предположите зависимость остатков от размера участка земли). Примите соответствующие меры.
 5. Проведите тест Рамсея для модели А с тремя вспомогательными регрессорами. Сделайте выводы.
 6. Протестируйте, улучшает ли модель включение дополнительных регрессоров: наличия подъездной дороги, комнаты отдыха, обустроенного подвального помещения, водяного отопления на газе, числа мест в гараже, расположения в привилегированном районе и числа этажей. Выполните известные вам тесты на адекватность этой модели и интерпретируйте полученные результаты. Проверьте тест на значимость группы коэффициентов.
 7. Выберите факторы, которые должны быть исключены из уравнения регрессии, используя метод пошагового исключения незначимых переменных.
 8. Выберите факторы, которые должны быть включены в уравнение регрессии, используя метод пошагового включения переменных.
 9. Представьте себе, что Вы имеете собственный двухэтажный дом на земельном участке 10 000 кв. футов, расположенный в привилегированной окрестности города, с четырьмя спальнями, одной ванной, с двухместным гаражом, подъездной дорогой, с комнатой отдыха, кондиционированием

воздуха, оборудованным подвальным помещением и водяным отоплением на газе. Чему равнялась бы ожидаемая цена продажи Вашего дома?

10. Оцените линейную модель (модель В), которая объясняет переменную *price* всеми оставшимися факторами в файле с данными.
11. Оцените модель С, которая объясняет логарифм переменной *price* логарифмом размера участка земли, числом спален, числом ванных комнат, наличия кондиционирования воздуха, константой и другими факторами из файла. Интерпретируйте результаты.
12. Проведите тест Бокса-Кокса. Какая функциональная спецификация в итоге оказалась более адекватна данным?
13. Создайте 4 фиктивных переменных для числа спален: 2 или меньше, 3, 4 и 5 или больше. Оцените модель (модель D) для логарифма цены, которая включает логарифм размера участка земли, число ванных комнат, наличие кондиционирования воздуха, три из четырех упомянутых фиктивных переменных и константу. Интерпретируйте результаты.
14. Включите все 4 фиктивные переменные в модель и повторно оцените ее. Что получилось? Почему?
15. Выполните тест на адекватность выбранной модели и сделайте выводы.
16. Постройте график остатков и предсказанные значения цены по модели.
17. Протестируйте нормальность остатков.
18. Протестируйте наличие мультиколлинеарности в выбранной модели. Примите соответствующие меры.
19. Протестируйте наличие гетероскедастичности в модели. Прделайте тесты Уайта, Бройша-Пагана. Примите соответствующие меры.
20. Проведите тест Рамсея для модели D. Сделайте выводы.
21. Прделайте тест Чоу двумя способами для выбора между моделями с числом спален для 2 или меньше и больше 2-х.
22. Как вы думаете по каким переменным выборка может быть неоднородной? Проверьте свое предположение с помощью теста Чоу или техники дамми переменных.
23. Какую модель Вы выберете в качестве итоговой?

Задание 2. Для самостоятельной работы.

В файле *clothing* (STATA) содержатся данные о продажах одежды в 400 немецких магазинах одежды.

Переменные:

tsales – среднегодовые продажи в гильденах,
sales - продажи в расчете на квадратный метр,
margin – маржинальная валовая прибыль,
nown – количество собственников (менеджеров),
nfull – количество полностью занятых,
npart - количество частично занятых,
nauх – количество временно работающих,
hoursw – общее число отработанных часов,
hourspw – количество отработанных часов в расчете на одного работающего,
inv1 – капиталовложения в помещения,
inv2 - капиталовложения в автоматизацию,
ssize – размер магазина в м²,
start – год открытия магазина.

- 1) Оцените зависимость среднегодовых продаж (переменная tsales) или продаж в расчете на квадратный метр (sales) или маржинальной валовой прибыли (margin) от всех остальных переменных.
- 2) Проверьте адекватность регрессии. Если регрессия адекватна, то переходите к следующим пунктам.
- 3) Прodelайте те же упражнения, что и в предыдущем задании.