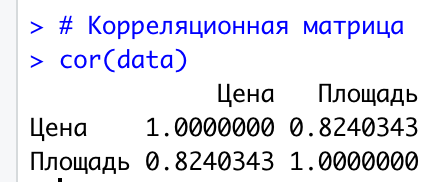
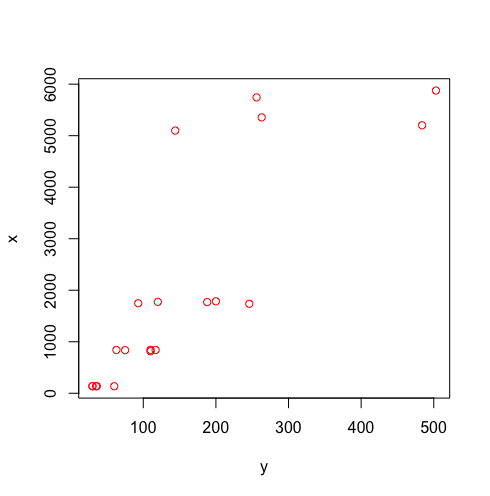
Данные

|  |  |
| --- | --- |
| **Цена** | **Площадь** |
| **5876** | 503 |
| **5743** | 256 |
| **5355** | 263 |
| **5202** | 484 |
| **5099** | 144 |
| **1785** | 200 |
| **1774** | 120 |
| **1770** | 188 |
| **1747** | 93 |
| **1737** | 246 |
| **842** | 110 |
| **840** | 63 |
| **840** | 117 |
| **836** | 75 |
| **816** | 110 |
| **137** | 30 |
| **137** | 36 |
| **137** | 30 |
| **136** | 35 |
| **136** | 60 |

А) Исследовать зависимость цены квартиры (в тыс. у.е.) в г. Москве от общей площади (м2). Выборка сделана таким образом, чтобы учесть элитное, среднее и рядовое жилье. Именно в силу данного разнообразия следует ожидать наличия свойства гетероскедастичности.



Коэфф 0.82 – говорит о том, что связь есть, причем достаточно сильная (что подтверждается здравой логикой)



Б) Построите 95%-й доверительный интервал для коэффициентов модели.

2.5 % 97.5 %

(Intercept) -852.64222 935.18689

y 8.36569 17.00235

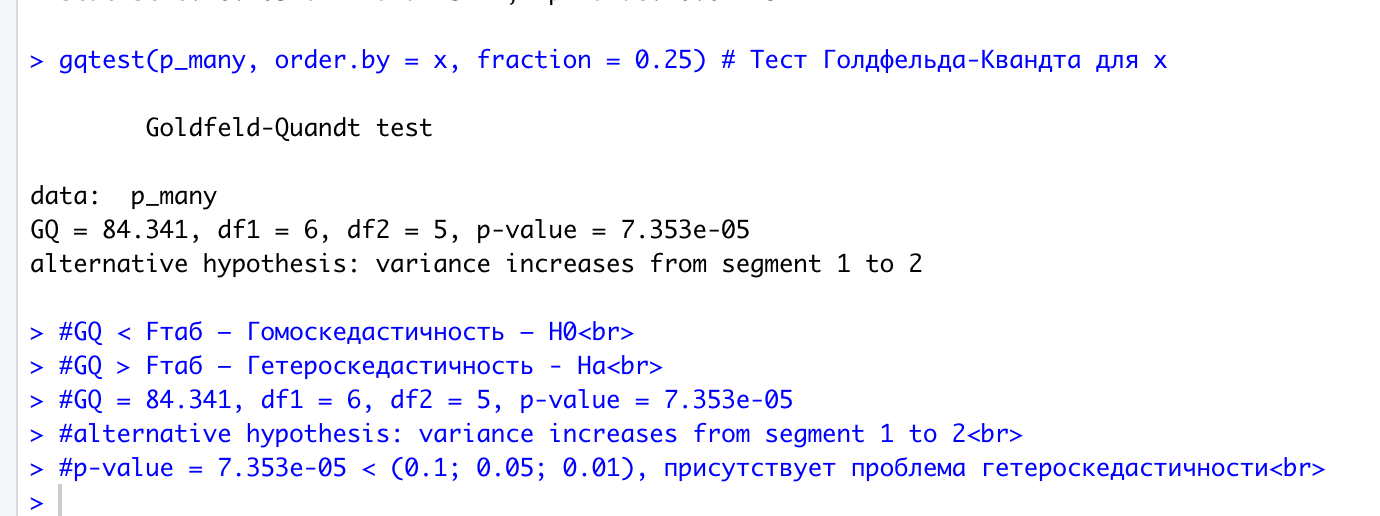
В) Проверьте значимость модели регрессии в целом и каждого коэффициента модели по отдельности.  
Я получил следующие данные и вывод

|  |  |  |
| --- | --- | --- |
| Коэффициент | Значение | Вывод |
| R^2 | 0.6790325 < 0.8 | Модель приемлема (R не очень близка к 1) |
| R^2 скорректированный | 0.661201 | Модель приемлема (R не очень близка к 1) |
| Стандартная ошибка Se | 1227.757 | Слишком сильная ошибка, плохо |
| Ошибка аппроксимации | 662.0046 | качество среднее |
| F-статистика | 38.08045 |  |
| P-value | 0.02275 |  |

Можем сделать вывод о том, что модель достаточно среднего качества и внушает скепсис

# Проверка на существование проблемы гетероскедастичности

## Тест Гольдфельда-Квандта



GQ < Fтаб – Гомоскедастичность – H0<br>

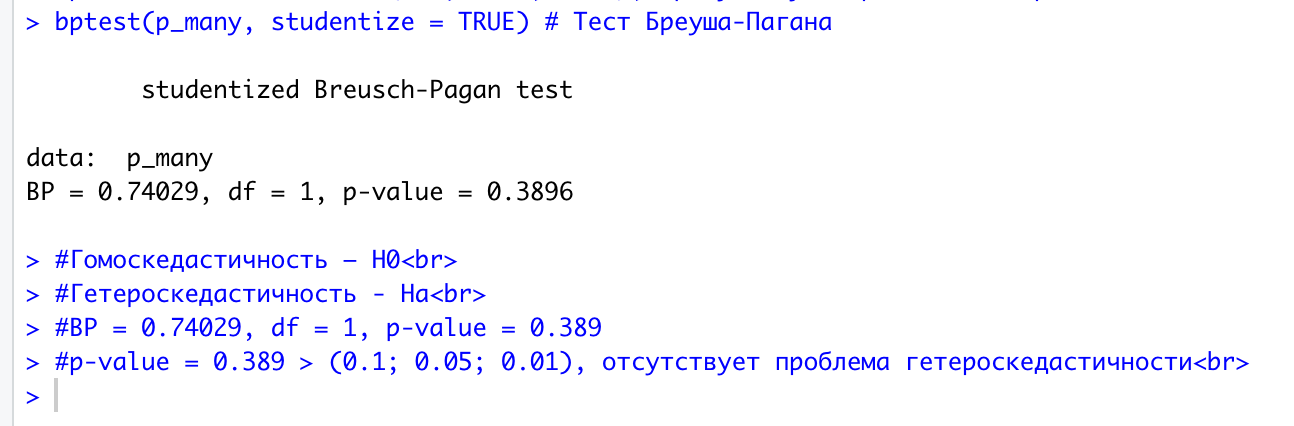
GQ > Fтаб – Гетероскедастичность - Hа<br>

GQ = 84.341, df1 = 6, df2 = 5, p-value = 7.353e-05

alternative hypothesis: variance increases from segment 1 to 2

p-value = 7.353e-05 < (0.1; 0.05; 0.01), присутствует проблема гетероскедастичности

## Теста Бройша-Пагана (Тест Уайта).



Гомоскедастичность – H0

Гетероскедастичность - Hа

BP = 0.74029, df = 1, p-value = 0.389

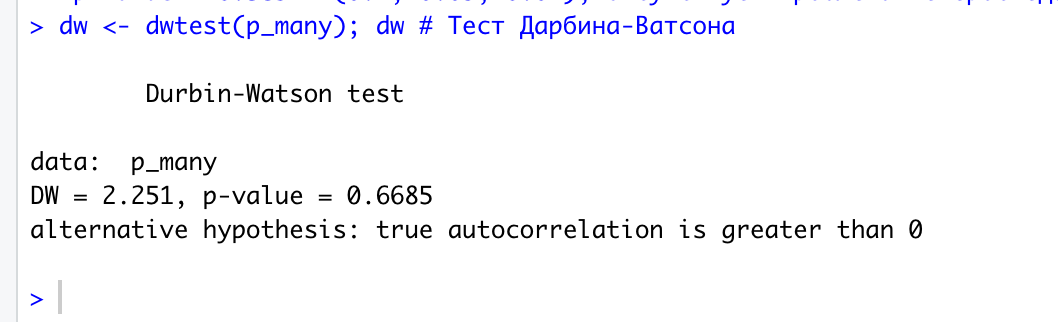
p-value = 0.389 > (0.1; 0.05; 0.01), отсутствует проблема гетероскедастичности

Тесты дали разные результаты

**Вывод: проблема гетероскедастичности существует**

# Проверка предпосылки об отсутствии автокорреляции

## Тест Дарбина-Уотсона



H0: нет автокорреляции<br>

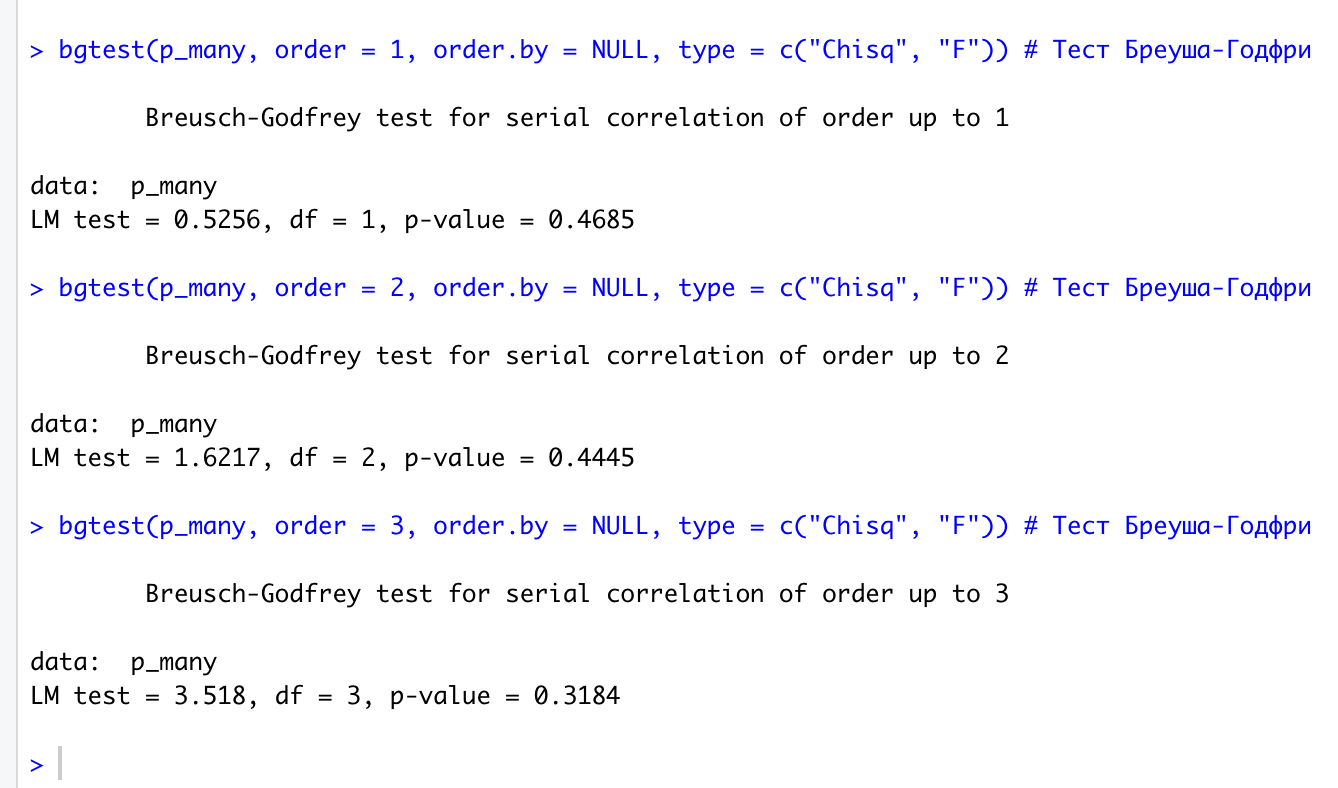
Ha: есть автокорреляция 1-го порядка<br>

DW = 2.251, p-value = 0.6685

DW не стремится к 0, что говорит об отсутствии положительной автокорреляции

p-value = 0.6685 > (0.01, 0.05, 0.1) - принимаем гипотезу об отсутствии автокорреляции, отвергаем гипотезу о существовании автокоррелции

## Тест Бройша-Годфри



H0: нет автокорреляции

Ha: есть автокорреляция n порядка

LM test = 0.5256, df = 1, p-value = 0.4685

LM test = 1.6217, df = 2, p-value = 0.4445

LM test = 3.518, df = 3, p-value = 0.3184

LM test = 25.304, df = 3, p-value = 0.03291<br>

pv = 0.4685 > (0.01, 0.05, 0.1) => H0 принимается, автокорреляция 1 порядка отсутствует

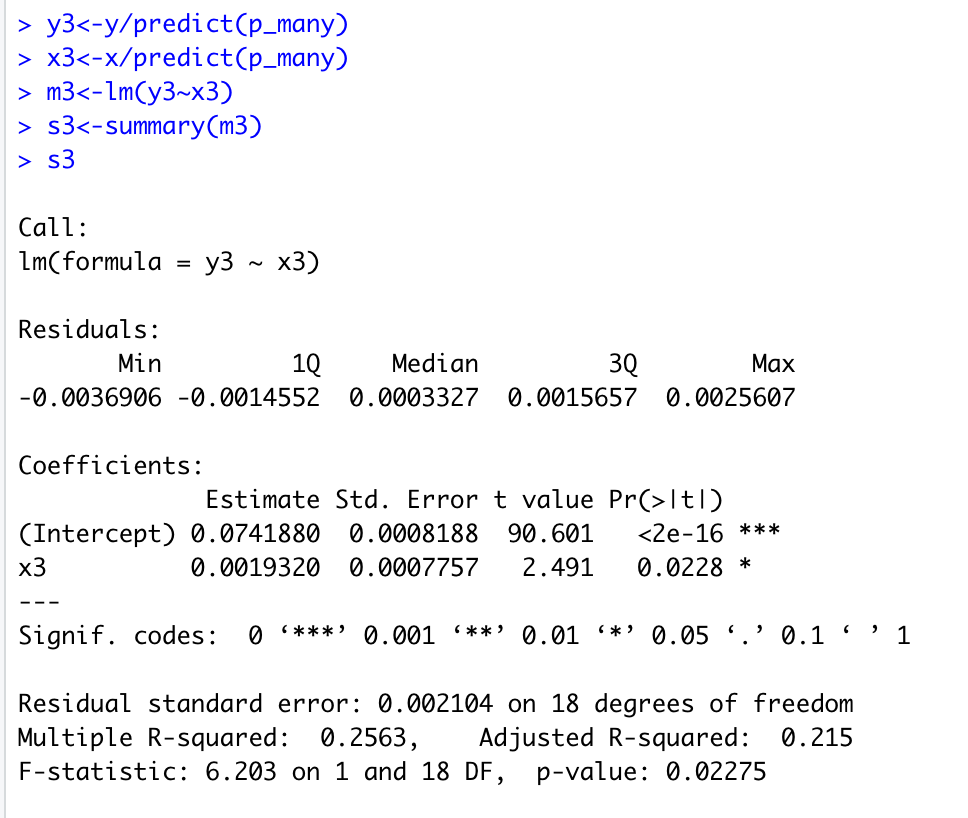
pv = 0.4445 > (0.01, 0.05, 0.1) => H0 принимается, автокорреляция 2 порядка отсутствует

pv = 0.3184 > (0.01, 0.05, 0.1) => H0 принимается, автокорреляция 3 порядка отсутствует

**Вывод: Автокорреляции отсутствует по 2 тестам**

Ж) При обнаружении проблем автокорреляции и(или) гетероскедастичности – предложить пути их устранения.

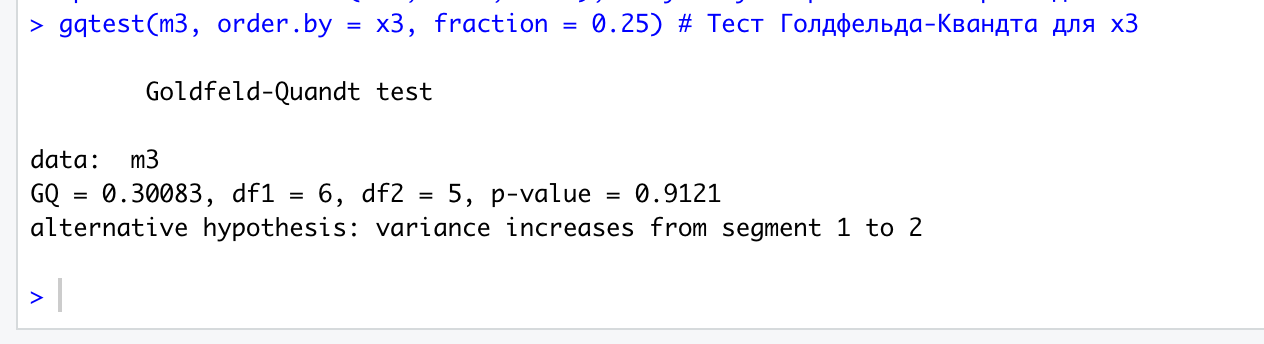
В модели имеется проблема гетероскедастичности, автокорреляции нет.  
Используем способ с ДВМНК (Доступный взвешенный метод наименьших квадратов)



Проводим тесты на гетероскедастичность еще раз ниже

# Проверка на существование проблемы гетероскедастичности (обновленная модель)

## Тест Гольдфельда-Квандта



GQ < Fтаб – Гомоскедастичность – H0

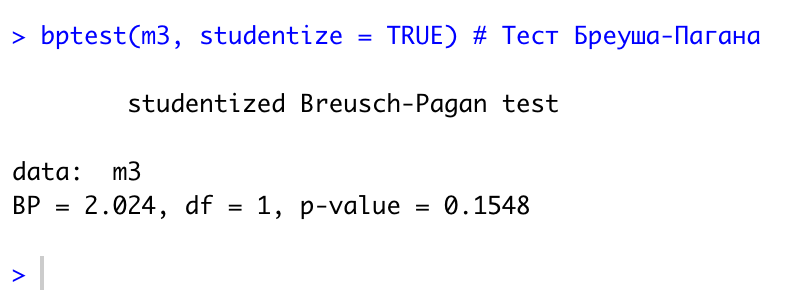
GQ > Fтаб – Гетероскедастичность - Hа

GQ = 0.30083, df1 = 6, df2 = 5, p-value = 0.9121

alternative hypothesis: variance increases from segment 1 to 2<br>

p-value = 0.9121 > (0.1; 0.05; 0.01), отсутствует проблема гетероскедастичности

## Теста Бройша-Пагана (Тест Уайта).



Гомоскедастичность – H0

Гетероскедастичность - Hа

BP = 2.024, df = 1, p-value = 0.1548

p-value = 0.1548 > (0.1; 0.05; 0.01), отсутствует проблема гетероскедастичности

**Вывод: проблемы гетероскедастичности в обновленной модели не существует**