Министерство образования и науки Российской федерации

Алтайский государственный технический университет

им. И.И. Ползунова

Отчет

о лабораторной работе №3

по теме: «Построение и оценка регрессионных моделей»

предмет: Компьютерные методы анализа и прогнозирования в экономических системах

Выполнил Студент группы 8ПИЭ-41

Хартов А.Е.

Проверил Жуковский М.С.

Барнаул 2024

**Цель работы**

Цель данной лабораторной работы — освоить методы построения и анализа регрессионных моделей для экономических данных с использованием языка программирования R. В рамках работы вы научитесь:

- Разрабатывать линейные модели для экономических данных.

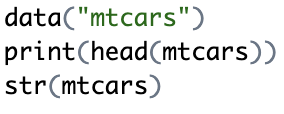
- Анализировать остатки и проводить диагностику моделей.

- Проверять гипотезы о значимости коэффициентов.

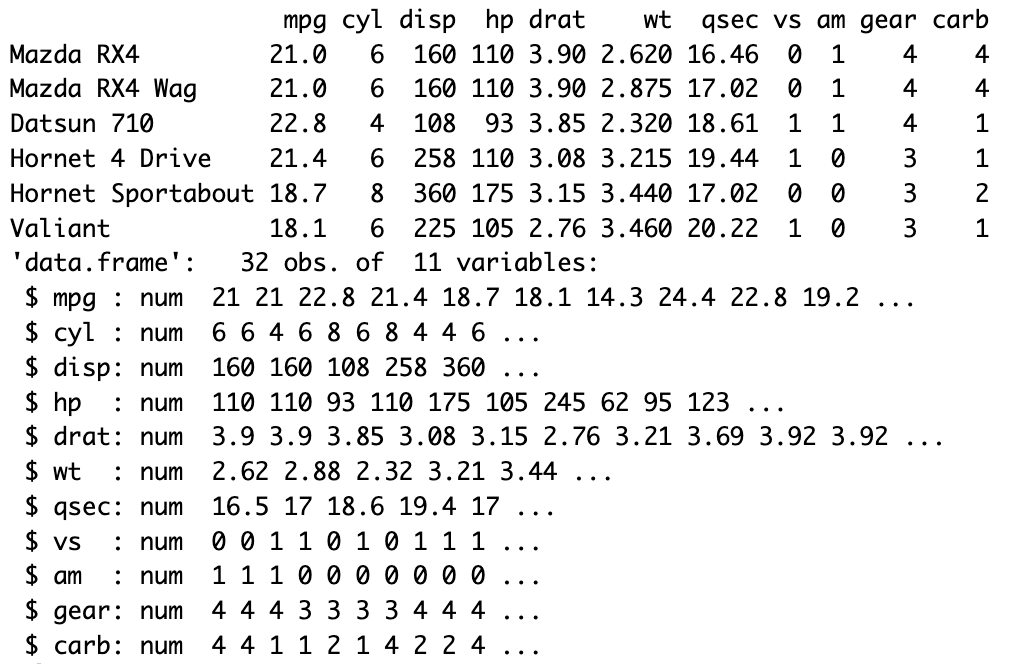
- Использовать регрессионные модели для прогнозирования.

Задание 1: Подготовка данных

1. Импорт данных:



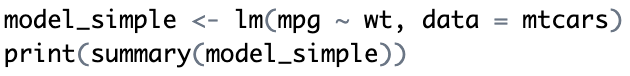
Результат:



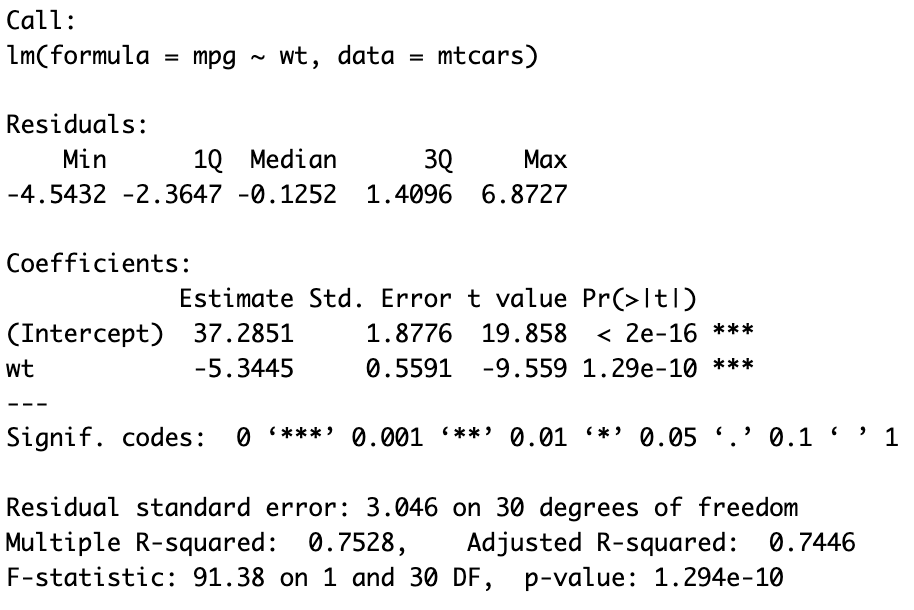
Задание 2: Построение линейной регрессии

1. Построение простой линейной регрессии:

Модель зависимости `mpg` от `wt`:



Результаты:



Анализ результатов:

- Коэффициенты модели:

Intercept = 37.2851 - Среднее значение mpg при нулевом wt. Хотя значение "вес = 0" не имеет физического смысла, оно важно для модели как точка отсчета.

wt = -5.3445 - при увеличении веса автомобиля на 1000 фунтов, расход топлива уменьшается в среднем на 5.3445 mpg. Это отрицательная связь.

- Значимость коэффициентов (p-значения):

|  |  |
| --- | --- |
|  | p-value |
| Intercept | <2e-16 |
| wt | 1.29e-10 |

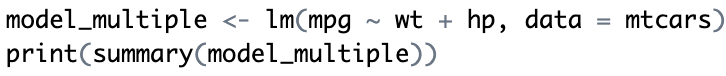
Оба коэффициента статистически значимы: p-значения значительно меньше 0.05. Вес автомобиля сильно влияет на расход топлива.

- Коэффициент детерминации (R^2):

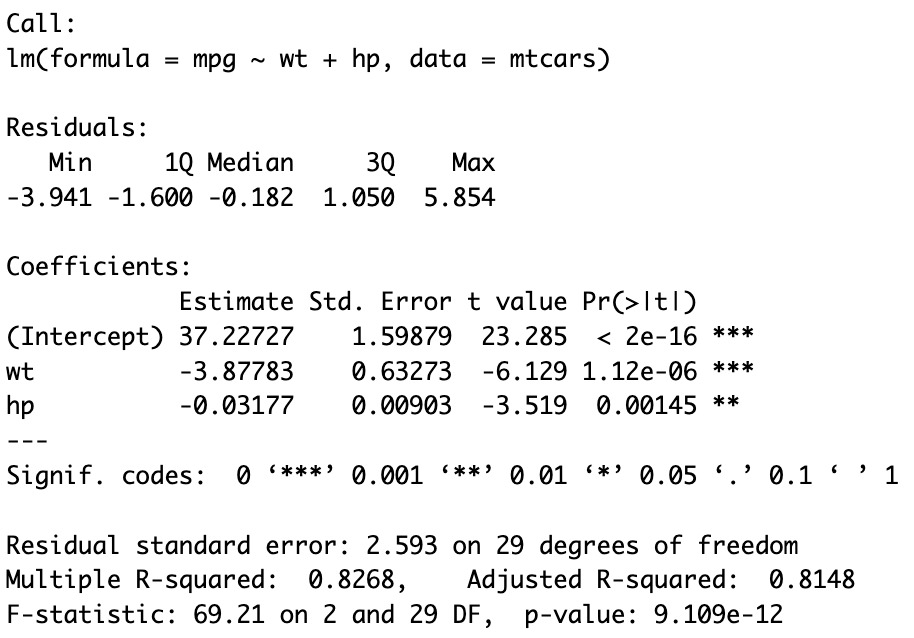
R^2 = 0.7528 - Примерно 75% вариации в значениях mpg объясняется весом автомобиля. Это говорит о сильной линейной зависимости между переменными.

2. Построение множественной линейной регрессии:

Модель зависимости `mpg` от `wt` и `hp`:



Результат:



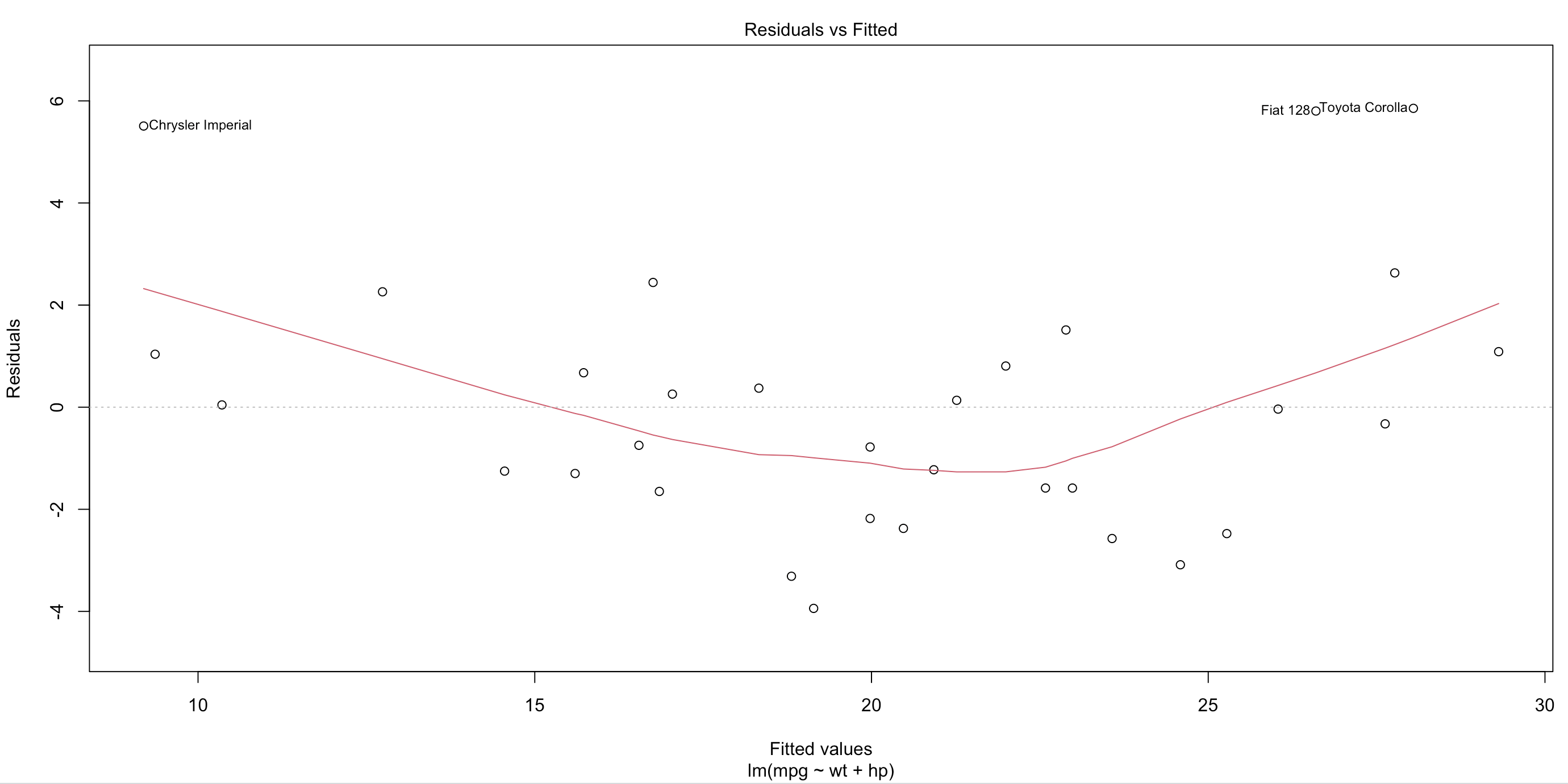
Сравнение результатов с простой регрессией:

Значение Intercept осталось таким же, значение wt уменьшилось по модулю. После учёта мощности (hp) влияние веса на расход ослабло, так как часть вариации в mpg объясняется теперь hp.

Коэффициент R^2 увеличился до 0.8268, видно что модель улучшилась. Модель теперь объясняет больше вариации в mpg.

Задание 3: Анализ остатков и диагностика модели

1. График остатков:

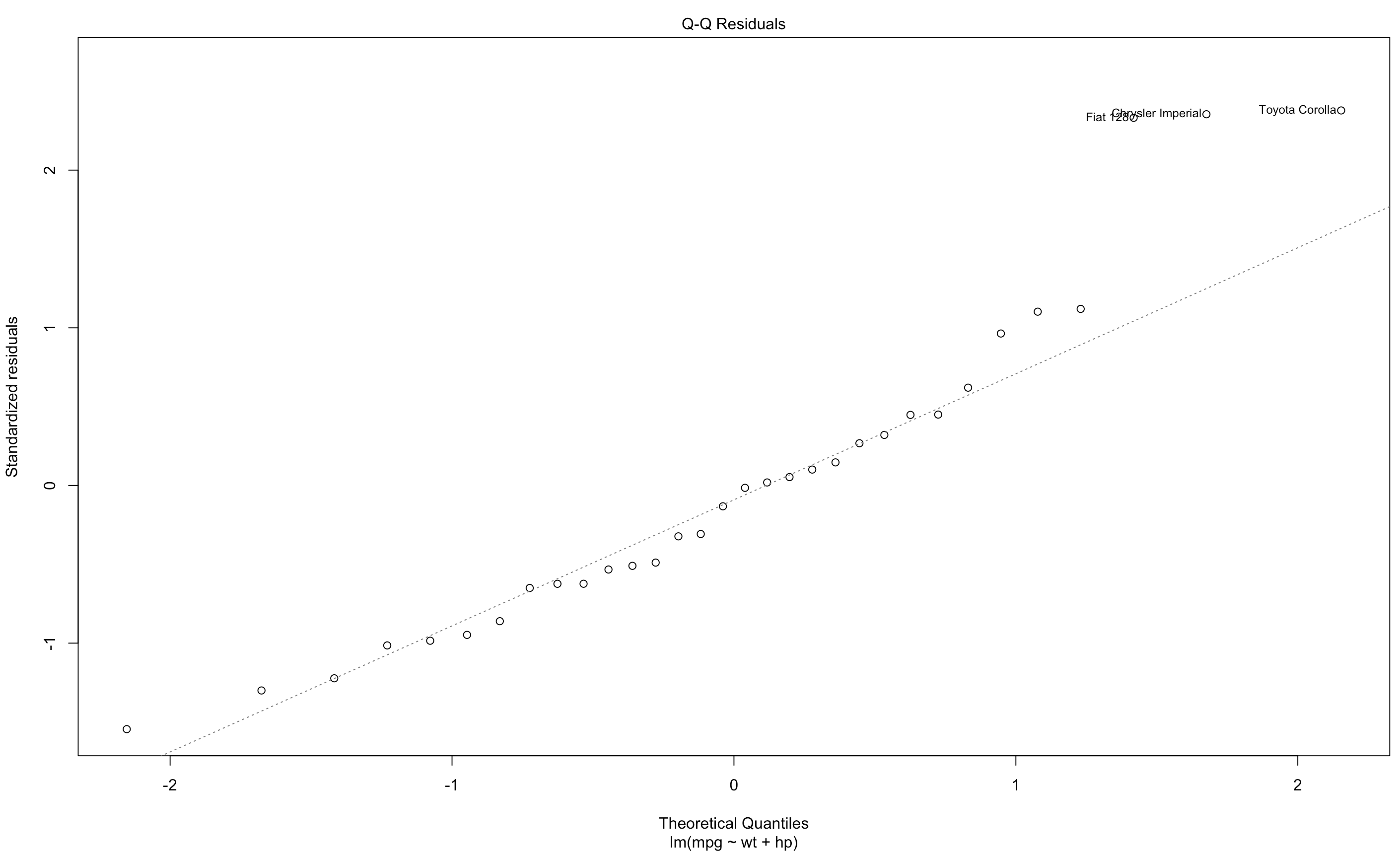


- Есть ли систематические отклонения?

Да, систематические отклонения есть, так как распределение остатков формируют U-образную форму

2. Q-Q график:





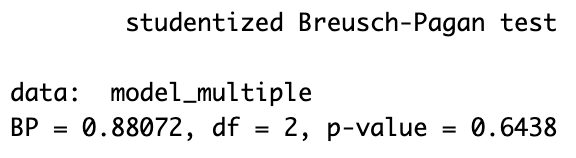
- Выполняется ли предположение о нормальности остатков?

Большинство точек расположено вдоль прямой линии, что соответствует нормальному распределению остатков. Предположение о нормальности остатков выполняется

3. Тест на гетероскедастичность:

тест Бройша-Пагана:





4. Мультиколлинеарность:

- Проверьте корреляцию между предикторами:

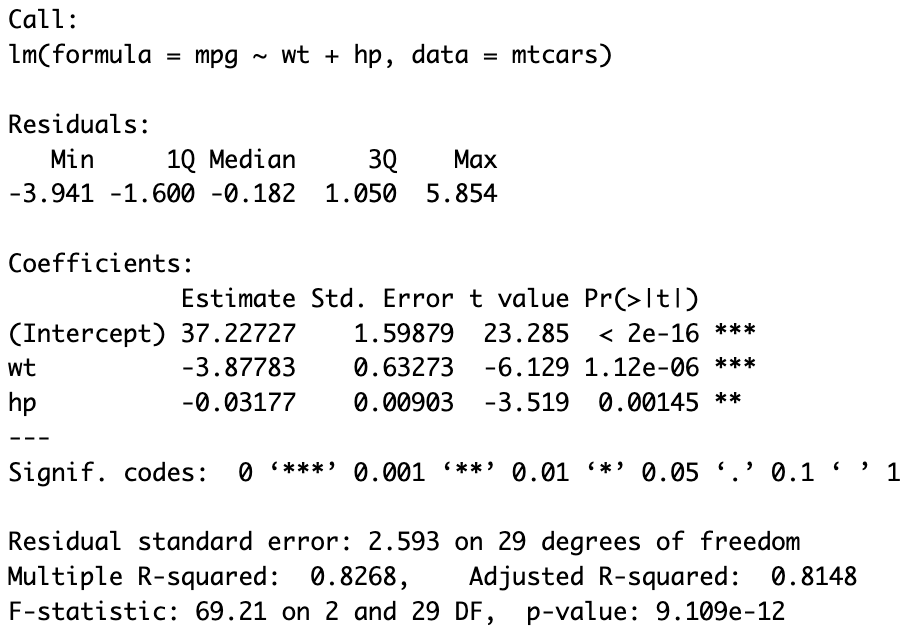


Корреляция равна 0.6587479. Корреляция меньше 0.7 значит она умеренно высокая.

Задание 4: Проверка гипотез о значимости коэффициентов

1. Гипотеза о значимости коэффициентов:

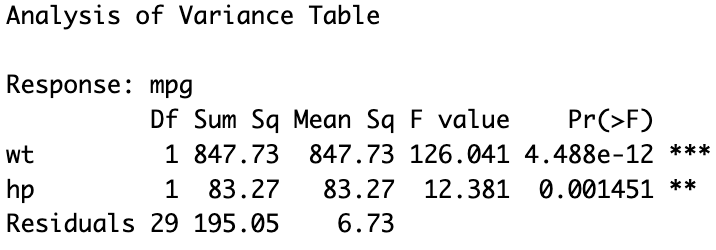




p-value для всех коэффициентов (Intercept, wt, hp) меньше 0.05. Это означает, что мы отвергаем нулевую гипотезу для каждого коэффициента. Следовательно, вес машины (wt) и мощность двигателя (hp) оказывают статистически значимое влияние на расход топлива (mpg).

2. Общая значимость модели:

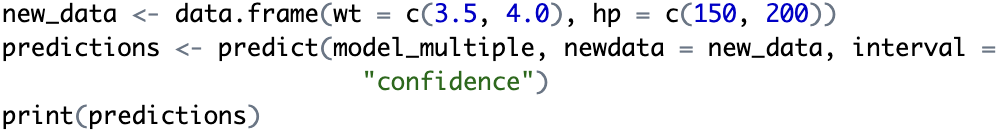




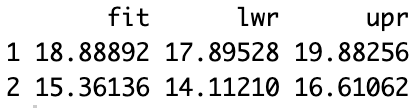
Видно, что для wt и для hp очень маленькие p-value (< 0.05).

Так как p-value для всей модели меньше 0.05, мы отвергаем нулевую гипотезу. Следовательно, хотя бы один предиктор (wt или hp) существенно связан с расходом топлива (mpg).

Задание 5: Прогнозирование на основе регрессионной модели 1. Прогнозирование новых значений:



Результат:



2. Визуализация прогнозов:

График фактических и прогнозируемых значений:

