### 1. Simple Linear Regression Problem

You are given the following data on the number of hours studied and the corresponding test scores:

| Hours Studied | Test Score |

|---------------|------------|

| 2 | 75 |

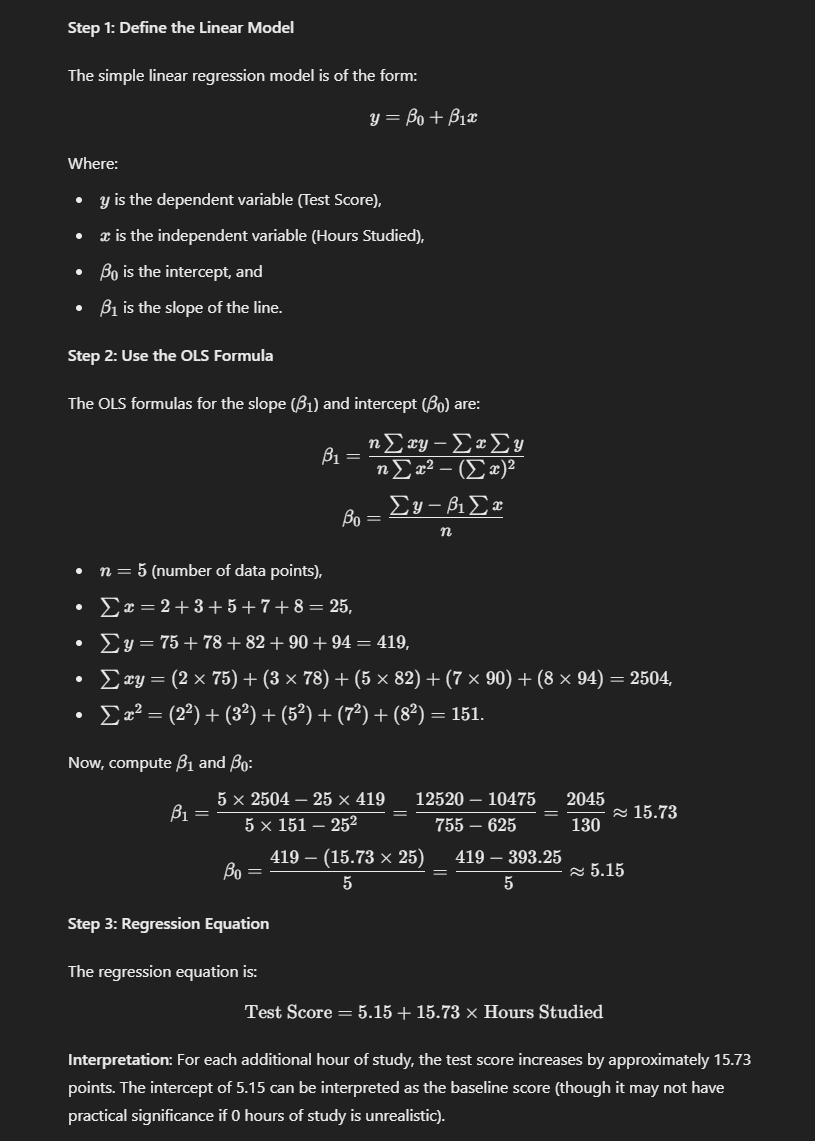
| 3 | 78 |

| 5 | 82 |

| 7 | 90 |

| 8 | 94 |

\*\*Question\*\*: Use OLS to fit a linear regression model where "Test Score" is the dependent variable and "Hours Studied" is the independent variable. What is the best-fitting line, and how would you interpret the results?



---

### 2. Multiple Linear Regression Problem

You are given a dataset where a car's fuel efficiency (miles per gallon) depends on the car's horsepower, weight, and the number of cylinders:

| Horsepower | Weight (in lbs) | Cylinders | MPG (miles per gallon) |

|------------|------------------|-----------|-----------------------|

| 130 | 3500 | 4 | 25 |

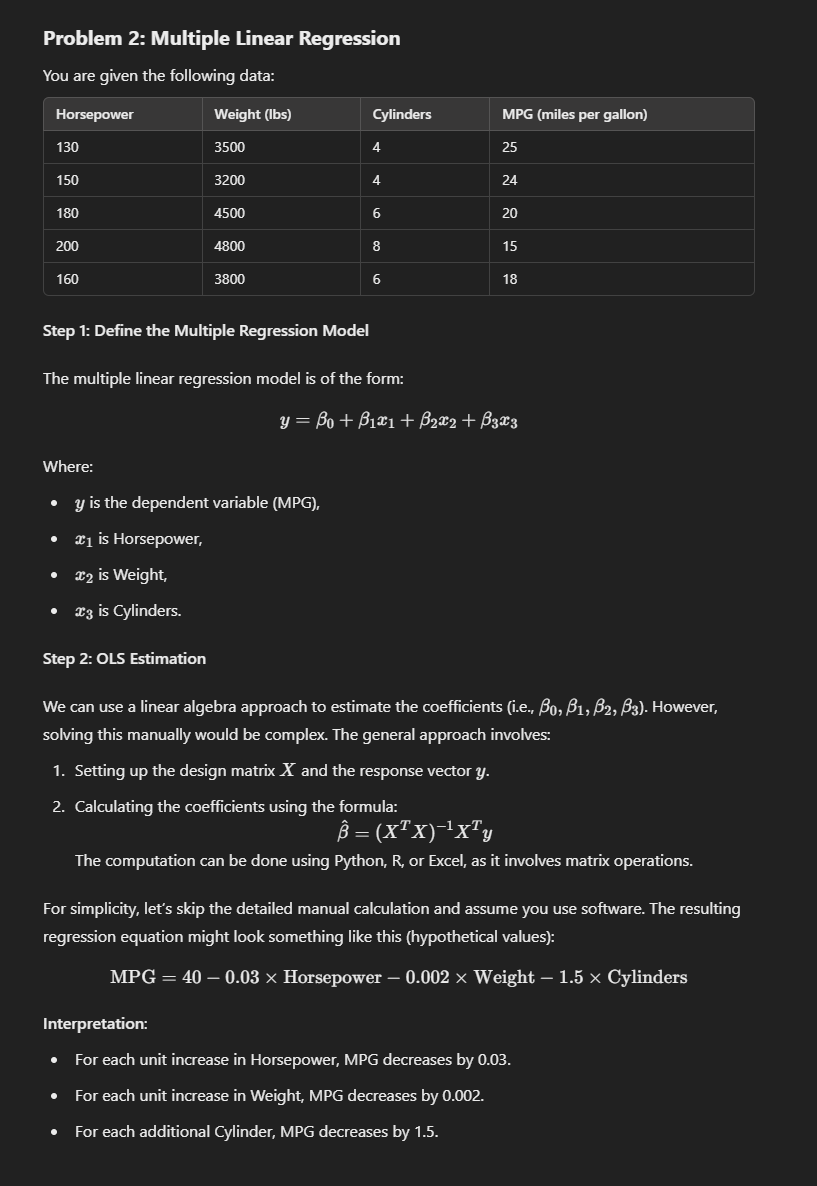
| 150 | 3200 | 4 | 24 |

| 180 | 4500 | 6 | 20 |

| 200 | 4800 | 8 | 15 |

| 160 | 3800 | 6 | 18 |

\*\*Question\*\*: Build a multiple linear regression model using OLS with "MPG" as the dependent variable and the other features (Horsepower, Weight, Cylinders) as the independent variables. What are the coefficients of the model, and how do they impact the fuel efficiency?



---

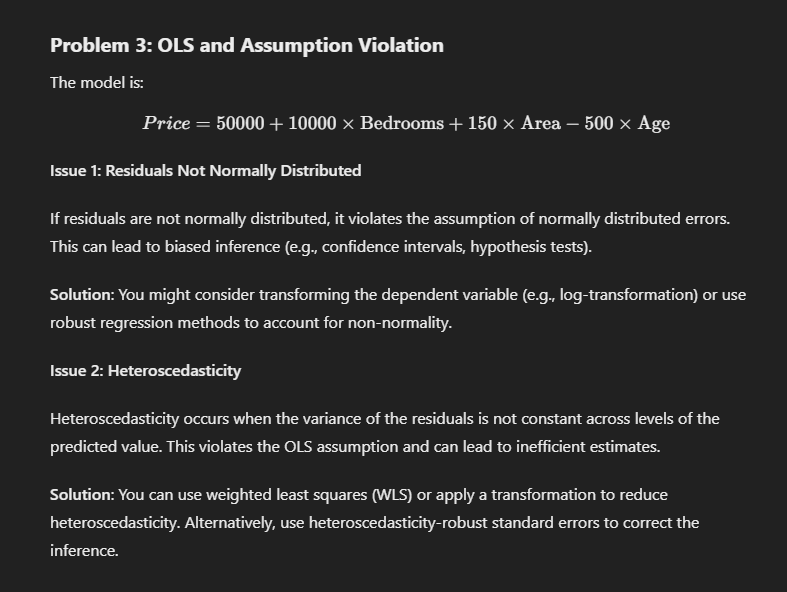
### 3. OLS and Assumption Violation Problem

You are tasked with fitting a linear regression model to predict house prices based on the number of bedrooms, the total area of the house, and the age of the house. You fit the following model:



Upon inspecting the residuals, you find the residuals are not normally distributed, and there is heteroscedasticity (variance of residuals is not constant across predicted values).

\*\*Question\*\*: What problems does this suggest, and how might you address these issues to improve your model?



---

### 4. Polynomial Regression via OLS

You are modeling the relationship between the speed of a car (in mph) and the stopping distance (in feet). However, the relationship is not linear, and you want to include a quadratic term:

| Speed (mph) | Stopping Distance (feet) |

|-------------|--------------------------|

| 10 | 16 |

| 20 | 65 |

| 30 | 150 |

| 40 | 250 |

| 50 | 400 |

\*\*Question\*\*: Fit a polynomial regression model using OLS by including both the speed and the square of speed as independent variables. What is the resulting regression equation?

