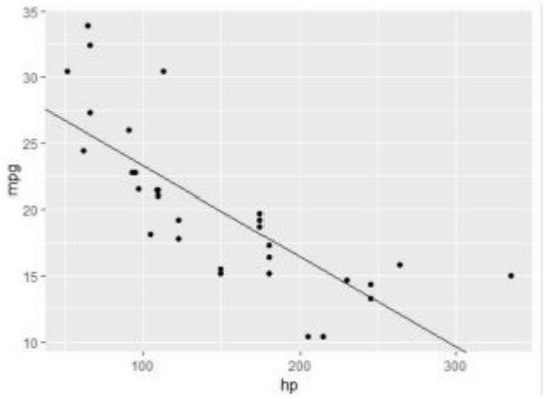
COP2073C Practice Exercise 10

For this exercise you will need to install and load the tidyverse.

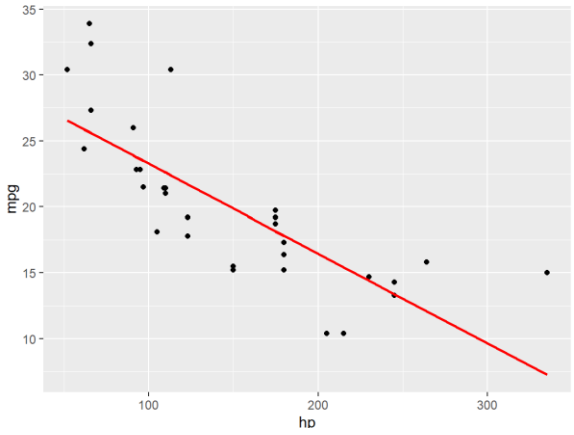
In this exercise we will use modelr and base R functions to practice analyzing a linear relationship.

Instructions:

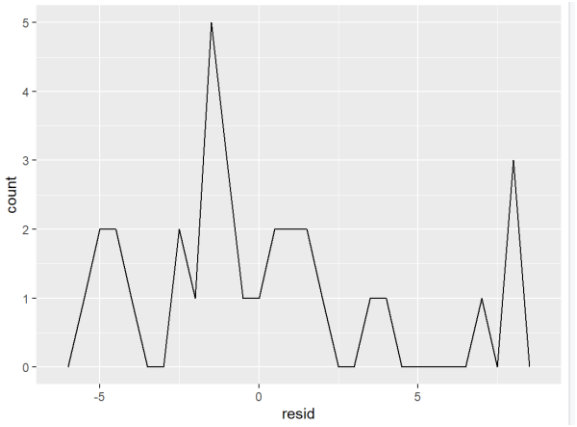
1. Create a tibble using the hp (horsepower) and mpg (miles per gallon) columns of the mtcars dataset.
2. Use the lm function to determine the coefficients of the linear relationship between the two variables (mpg ~ hp).
3. Print the coefficients (hint: use the coef() function).
4. Plot the points and the fitted line as shown here:



5. Create a data grid from the tibble's horsepower (hp) column, then add predictions using the grid.
6. Plot the resulting values as shown below and compare to the plot created from the lm function coefficients.



7. Calculate the residuals and print them.
8. Plot a frequency polygon for the residuals as shown here:



Expected Output:

```
coef: 30.09886 -0.06822828
[1] "residuals:"
# A tibble: 32 × 3
      hp   mpg  resid
   <dbl> <dbl>  <dbl>
 1   110  21   -1.59
 2   110  21   -1.59
 3    93  22.8 -0.954
 4   110  21.4 -1.19
 5   175  18.7  0.541
 6   105  18.1 -4.83
 7   245  14.3  0.917
 8    62  24.4 -1.47
 9    95  22.8 -0.817
10   123  19.2 -2.51
# ℹ 22 more rows
```

Analysis:

This program analyzes the relationship between horsepower (hp) and miles per gallon (mpg) in the mtcars dataset using linear regression. The linear model fitted with lm reveals a negative relationship between horsepower and fuel efficiency, as the coefficients indicate that mpg decreases as hp increases. The fitted line, plotted alongside the data points, visually confirms this trend. A second plot, showing predicted mpg values based on the data grid, aligns closely with the regression line, illustrating that the trend continues in the predicted values. The residuals plot (showing the differences between observed and predicted mpg values) suggests that most predictions are close to the actual values, though some variation exists, indicating that while the model captures the general trend, it may not explain all the variability in mpg.