Final Project Checkpoint 2: Model Construction and Visualization

Objective: Build and evaluate a multiple linear regression model for your dataset from checkpoint 1.

📌 Task 1: Build a Multiple Linear Regression Model (6 points)

1. Load your dataset into **RStudio** (if not already loaded from Checkpoint 1).
2. Train a multiple linear regression model using the `model_train()` function in R. Your model should include:
    - **One response variable (dependent variable).**
    - **At least two explanatory variables (independent variables).**
3. Use the following structure to train your model (add more explanatory variables as desired):

```
Model <- dataset |> model_train(response_variable ~ explanatory_variable1 + explanatory_variable2)
```

In the answer submission form, respond to the following prompt:

✍️Explain why you selected these explanatory variables.

📌 Task 2: Visualize the Model Predictions (6 points)

1. Create a scatter plot with the model's predictions using `point_plot()`.
2. Include a model annotation (`annot = "model"`) to visualize the predicted trend.
3. If needed, adjust the visualization by mapping additional variables to **color or facet** to improve clarity.

In the answer submission form, respond to the following question:

📸 Attach a screenshot of your visualization from RStudio.

✍️ Briefly describe the trend shown in your visualization.

📌 Task 3: Evaluate the Model Fit Using R^2 (6 points)

1. Calculate **R^2**, which measures how well the model explains the variation in the response variable.

```
Model |> R2()
```

In the answer submission form, respond to the following question:

✍️ Record your **R^2** value.

✍️ Interpret the **R^2** value in context**.** What does it tell you about the strength of the model?

📌 Task 4: Analyze Effect Sizes Using Model Coefficients (7 points)

1. Extract the model coefficients using the `conf_interval()` function:

```
Model |> conf_interval()
```

In the answer submission form, respond to the following question:

📸Attach a table of your model coefficient values. Column 1 is the explanatory variable name. Column 2 is the coefficient value.

✍️ Explain the significance of the coefficients in relation to your dataset.

After all tasks are complete:

📸 Attach the R Script from RStudio that contains your code from Tasks 1-4.

Submit answers on this Google Form