Econ 3040 - Assignment 2: Sales from different video game consoles

Ryan T. Godwin

The big three 7th generation video game consoles released in 2005-2006 were the Nintendo Wii, Playstation 3, and Xbox 360. Use the data below to determine if the critic score of a video game is related to its sales. The data for your assignment is the same as from Computer Lab 1. The data was obtained by Abdulshaheed Alqunber.

Due date: February 20th, 2026. Worth 3% of your final grade.

Instructions: Submit your assignment in the "Assignment 2" drop box on UM Learn. *You must complete your assignment individually.* Submit both your answers to the assignment questions, and the R code that you use. Make sure all graphs and tables are appropriately labelled. Load the data using:

```
mydata <- read.csv("https://rtgodwin.com/data/vidsales.csv")
```

---

1. Make a subset that contains only `Wii`, `PS3`, and `X360` games. **Use this subset for the remainder of the assignment**. Please see the hint in Computer Lab 1 for making subsets.
2. Draw a scatterplot of `Score` (on the x-axis) vs. `Sales` (on the y-axis). Include this scatterplot in your assignment.
3. How much does it pay to make a *good* game? Estimate a model by least squares (LS) to answer the question: how much more does a video game sell on average, when it has a higher critic score rating?
4. Interpret the R-square from the estimated model in question 3.
5. Test the hypothesis that `Score` has no effect on `Sales`.
6. Estimate a new model. Use dummy variables to differentiate between `Wii`, `PS3`, and `X360` games. What is the average difference in sales for these three consoles?
7. In the above question, there are three categories for `Platform`. Why were only *two* dummy variables used in the model?
8. Using the entire sample (all 4706 observations) estimate a model with `Sales` as the dependent variable and `Score, Publisher, Platform` and `Genre` as the regressors. (Do not include the entire table of output results in your assignment, it is too big). What is the effect of `Score` on `Sales`? What is the reason for including these additional variables in the model, even though we only care about the effect of `Score`?