**Homework 4 – Regression Analysis**

The dataset for this exercise is available in VideoGamesSales\_Main.7bdat. This dataset contains information on the global sales and critic and user review ratings for videogames launched between 2001 and 2012 (from [www.vgchartz.com](http://www.vgchartz.com)). The variables are:

* Name of the game
* Videogame platform on which it was released.

|  |  |
| --- | --- |
| **Platform** |  |
| **DS** | Nintendo DS |
| **GBA** | Nintendo Game Boy Advance |
| **GC** | Nintendo Game Cube |
| **PC** | Personal Computer |
| **PS2** | Sony PlayStation 2 |
| **PS3** | Sony PlayStation 2 |
| **PSP** | Sony PlayStation Portable |
| **Wii** | Nintendo Wii |
| **X360** | Microsoft XBOX 360 |
| **XB** | Microsoft XBOX |

* Videogame Genre (e.g., Action, Sports, Shooter etc.)
* Publisher
* Developer
* Rating: E = Everyone, E10+ = Everyone 10+, T = Teen, M = Mature
* Global Sales (Millions of units)
* Year of release
* Critic Score (0 – 100): Average critic rating
* Critic Count : Number of critic ratings
* User Score (0 – 10): Average user rating
* User Count: Number of user ratings

1. Develop a regression model that links global sales to video game reviews. Explore ways in which the model fit could be improved through suitable changes to the model specification and variables.
   1. Present the final model and results.
   2. Explain how you developed your model (what was your initial model, what were the key variations you tried and how did you arrive at the final model – and the thought process behind these steps).
   3. Interpret the model results.
2. For the model you constructed, verify whether the various regression assumptions discussed in class are satisfied. If an assumption is violated, discuss how it can be handled, and implement the same. Discuss whether this change had a practically significant impact on your model results.
3. Submit: (i) your answers to the above, (ii) the SAS code and (iii) SAS output – either as one word document, or as a zipped folder containing these different files clearly named.