

# CS 4525 Project Topic, Data

Eric Foote

October 2018

My data is from <https://www.kaggle.com/gregorut/videogamesales> and it is video game sales as of October 2016. It is a single file containing 16600 tuples and 11 attributes (Rank, Name, Platform, Year, Genre, Publisher, NA Sales, EU Sales, JP Sales, Other Sales, Global Sales) so in theory it should be large enough to be interesting since the file size is 1.4 MB but should hopefully be able to fit onto the VM's setup in the lab. There is a lot of possibilities for this data set; therefore, my topic is to compare the purchasing habits of North Americans vs Europeans vs Japanese. To achieve this I want to do the following, group franchises together and compare sales, group platforms (Playstation, Nintendo, etc...) together and compare sales, group genre together and compare sales and finally group publishers (Electronic Arts, Ubisoft, etc...) together and compare sales. I don't have a team for this topic.