



7/15/2016

Analysis of Vivendi and Competitors' Sales Performance

Presented to:

CEO, Vivendi Universal

Prepared by:

Margaret Kapitany

Nikita Solovyev

William Fong

Danny Chahal

Itzel Vargas

Anthony Siu

Executive Summary

This report was generated to investigate the causes of declining average revenue per game by launch year of Vivendi Universal (Vivendi). Utilizing the framework of decision process provided, a gap analysis and root-cause analysis are conducted to assess whether a sales gap exists, between Vivendi and its competitors, and if so, why.

Revenue per game is the accepted performance metric utilized in this analysis as it is relevant to measuring the objective and readily available for use. Specifically, life-to-date (LTD) sales is chosen as the response variable to indicate the sales performance of games.

The results of the gap analysis indicate that a performance (LTD sales) gap exists between Vivendi's games and its competitors' games. Vivendi is outperformed by its competitors and thus, a root-cause analysis is conducted to identify the causes of the gap.

Through statistical analysis and data-mining of the data set, we reject Rudy's first theory that Vivendi produces fewer violent video games than its peer group. However, we agree with his second theory that "what we need is a Madden" as its license is worth up to \$16.72 million to Vivendi.

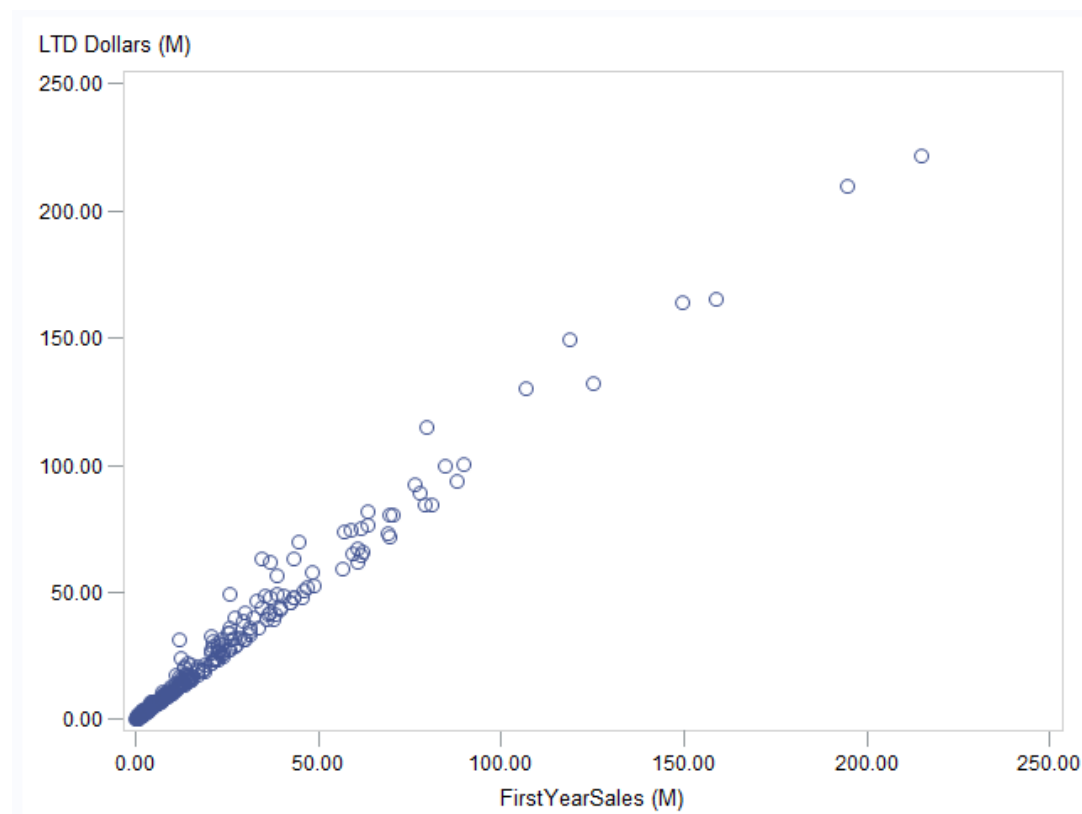
Objectives

First Year Sales vs. LTD Sales

The first-year sales include all the sales of the video games during the first year of release and the Life-to-Date (LTD) sales include the sales from the first year and the years afterwards. A scatter plot and a correlation matrix was created to demonstrate the relationship between first-year sales and LTD sales.

Scatterplot Analysis

The scatterplot indicates a positive correlation between first year sales and LTD sales, as demonstrated by the positive slope.



Correlation Matrix Analysis

The correlation matrix reveals a correlation value of 0.99030428 between first year sales and LTD sales, indicating that a strong positive correlation exists.

	FirstYearSales (M)	LTD Dollars (M)
FirstYearSales (M)	1	
LTD Dollars (M)	0.99030428	1

Statistics of Video Game First Year Sales

For all video games (with the data we have available), the minimum first year sales has been \$0, with a maximum first year sales of \$265.5 million. With an average first year sales of \$11.2 million and a median of \$0.34 million, we can conclude that the distribution of video game sales is positively skewed and there are only a few video games that performed significantly well.

Determining the Response Variable

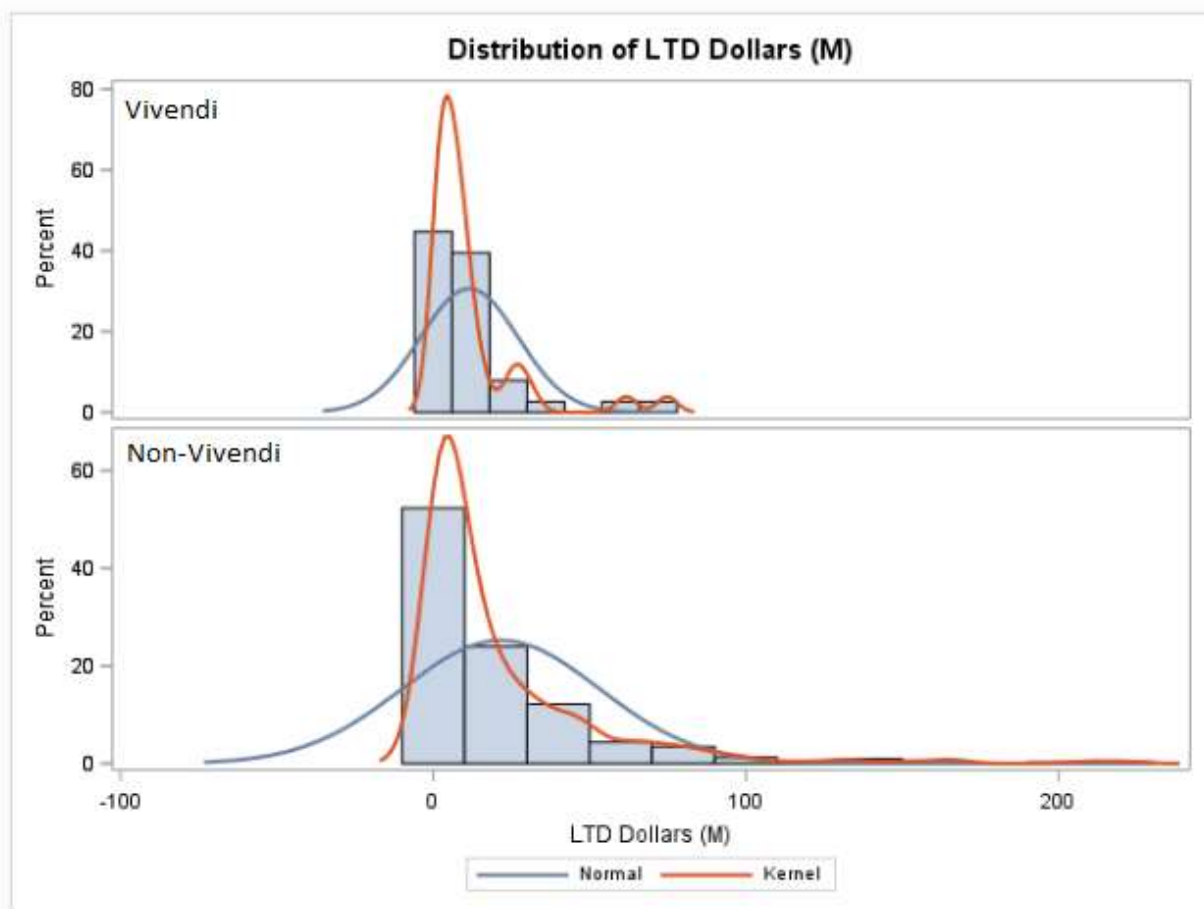
Revenue per game is the accepted performance metric utilized in this analysis as it is relevant to measuring the objective and readily available for use. LTD sales takes into account of the following factors which may contribute to the increasing sales of games in subsequent years:

- Strong performance from sequels leading owners to buying previous titles
- Exponential growth in player base leading to sales in later years
- Increased uptake with players reduction in prices on games in later years may lead to
- Owner with older consoles may continue to buy older games

Given Vivendi continues to earn revenue over the life-to-date sales of a video game, using first year sales could lead to errors in analysis. Therefore, life-to-date (LTD) sales is chosen as the response variable to indicate the sales performance of games.

Gap-Analysis

Based on the t-test's values, in order to assess variance of the results, we found that the F-test's p-value of <0.0001 is lower 0.05 and determined that there was unequal variance. Therefore, the Satterthwaite t-test was used to reference the p-value of 0.0016 which is lower than 0.05. This comparison indicates that the null hypothesis is rejected and concludes that there is a significant difference between the performance of Vivendi and Non-Vivendi Peers, which include LucasArts, THQ, Electronic Arts, Ubisoft, and Sony, with a difference of \$10.07 million in dollar LTD sales.



t Test

The TTEST Procedure

Variable: LTD Dollars (M)

PublisherCompare	N	Mean	Std Dev	Std Err	Minimum	Maximum
Vivendi	38	11.4995	15.6404	2.5372	0.8276	74.8384
Non-Vivendi	321	21.5717	31.6295	1.7654	0.1609	221.5
Diff (1-2)		-10.0722	30.3660	5.2094		

PublisherCompare	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
Vivendi		11.4995	6.3587 16.6404	15.6404	12.7510 20.2347
Non-Vivendi		21.5717	18.0985 25.0450	31.6295	29.3574 34.2858
Diff (1-2)	Pooled	-10.0722	-20.3172 0.1728	30.3660	28.2927 32.7697
Diff (1-2)	Satterthwaite	-10.0722	-16.2242 -3.9202		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	357	-1.93	0.0540
Satterthwaite	Unequal	79.348	-3.26	0.0016

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	320	37	4.09	<.0001

Root-cause Analysis

Rudy's Theory

Rudy claims that Vivendi's profits are lacking behind its peer group because he believes Vivendi produces less violent video games, which he believes generates greater sales. To verify if Rudy's claim is valid or not, a contingency table is created to compare violent vs. non-violent game production between Vivendi and its peer group. Violent video games are defined as games that fall in the genre of action, fighting, or shooting.

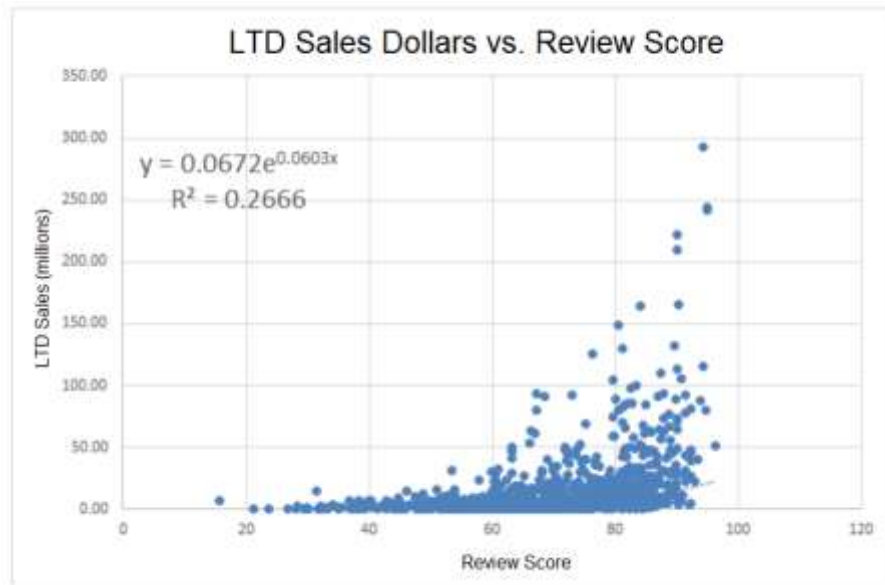
Based on the contingency table below, it is revealed that Vivendi actually produces a greater percentage of violent video games (71.05%) than its peer group (31.78%). Therefore, it is not true that Vivendi produces fewer violent video games than its peer group. Thus, Rudy's claim is incorrect.

Table of Super Genre Violence by Compare				
		Compare(0=Vivendi,1=Peer,2=Non-Peer)		
		Non-Vivendi	Vivendi Universal	Total
Super Genre Violence				
Non-Violent	Frequency	219	11	230
	Col Pct	68.22	28.95	
Violent	Frequency	102	27	129
	Col Pct	31.78	71.05	
Total	Frequency	321	38	359

Regression Analysis

Scatterplot Analysis

The scatterplot below indicates a positive exponential correlation between LTD sales and review score. Assessing the accuracy of the model, we conclude that only approximately 26.67% of the variability in LTD sales can be explained by review score. However, it can be concluded that the higher the review score, the higher LTD sales can be expected for a game.



Simple Regression Analysis

Since there is an exponential relationship between LTD Sales and Review, we transformed LTD Sales by taking the natural logarithm. It was assumed that Vivendi was interested in the entire dataset and, thus, no game publishers or developers were filtered out. For every 1 unit change in review score, on average, there is a 6% change in the LTD Sales. Again, it is shown that Review Score by itself can only explain 26.67% of the variability in LTD Sales. Therefore, Review is not a good predictor of LTD Sales.

Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: LTDSalesLN

Number of Observations Read	1256
Number of Observations Used	1200
Number of Observations with Missing Values	56

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	747.11732	747.11732	435.60	<.0001
Error	1198	2054.75407	1.71515		
Corrected Total	1199	2801.87140			

Root MSE	1.30964	R-Square	0.2666
Dependent Mean	1.48834	Adj R-Sq	0.2660
Coeff Var	87.99338		

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-2.70077	0.20424	-13.22	<.0001
Review	1	0.06034	0.00289	20.87	<.0001

Normally Distributed Residuals Assumption

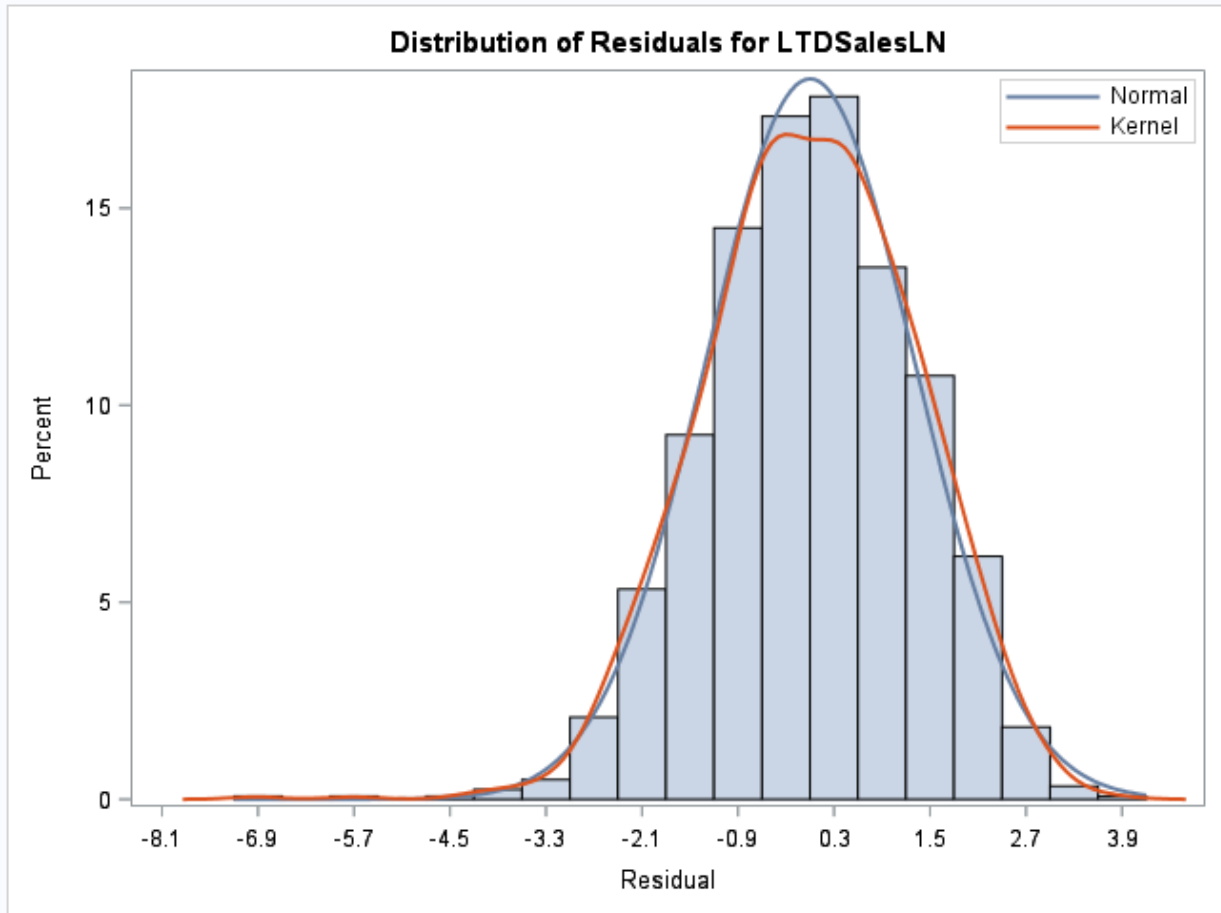
Yes, the residuals are normally distributed as LTD Sales was transformed logarithmically to account for the non-linear relationship.

Linear Regression Results

The REG Procedure

Model: Linear_Regression_Model

Dependent Variable: LTDSalesLN



Root Causes of Game Success

First, we created a correlation matrix (see below) to investigate if there was multicollinearity between explanatory variables in the dataset. Variables with a correlation above 0.6 were removed and, as a result, the following variables were used in multiple linear regression:

- GCN
- XBX
- PS2
- Launch
- Review
- PriorSaleBrand (M)
- ConcurrentReleaseBrand
- ConcurrentReleaseGenre
- PriorReleaseBrand
- PriorReleaseGenre
- FirstYearRatio
- LTD ASP

	GCN	XBX	PS2	Launch	Review	PriorSaleBrand (M)	ConcurrentReleaseBrand	ConcurrentReleaseGenre	PriorReleaseBrand	PriorReleaseGenre	FirstYearRatio	LTD ASP
GCN	1											
XBX	0.13	1										
PS2	-0.15	-0.23	1									
Launch	0.08	0.23	-0.1	1								
Review	0.01	0.06	0.07	-0.01	1							
PriorSaleBrand (M)	0.16	0.14	0.1	0.2	0.24	1						
ConcurrentReleaseBrand	0.06	0.13	0.1	0.08	0.13	0.25	1					
ConcurrentReleaseGenre	0.27	0.23	0.23	0.19	-0.07	0.05	0.05	1				
PriorReleaseBrand	0.09	0.18	0.12	0.29	0.21	0.46	0.59	0.03	1			
PriorReleaseGenre	0.16	0.2	0.24	0.48	-0.03	0.15	0.02	0.76	0.19	1		
FirstYearRatio	-0.01	0	0.05	-0.23	0.15	0.08	0.04	-0.02	0	-0.11	1	
LTD ASP	-0.06	-0.02	0.08	-0.08	0.39	0.2	0.04	-0.07	0.09	-0.04	0.27	1

The only exception was granted to keep both ConcurrentReleaseGenre and PriorReleaseGenre, despite the fact that their correlation coefficient exceeded our set benchmark. These two variables provide different data and even though they are not fully independent (preceding genres sold influence sale of competing genres now), it is reasonable to keep both of them in the model.

Next SAS was used to run a linear regression using the stepwise selection method with a significance level of 0.10 set for variable to enter and remain in the model.

Linear Regression Results

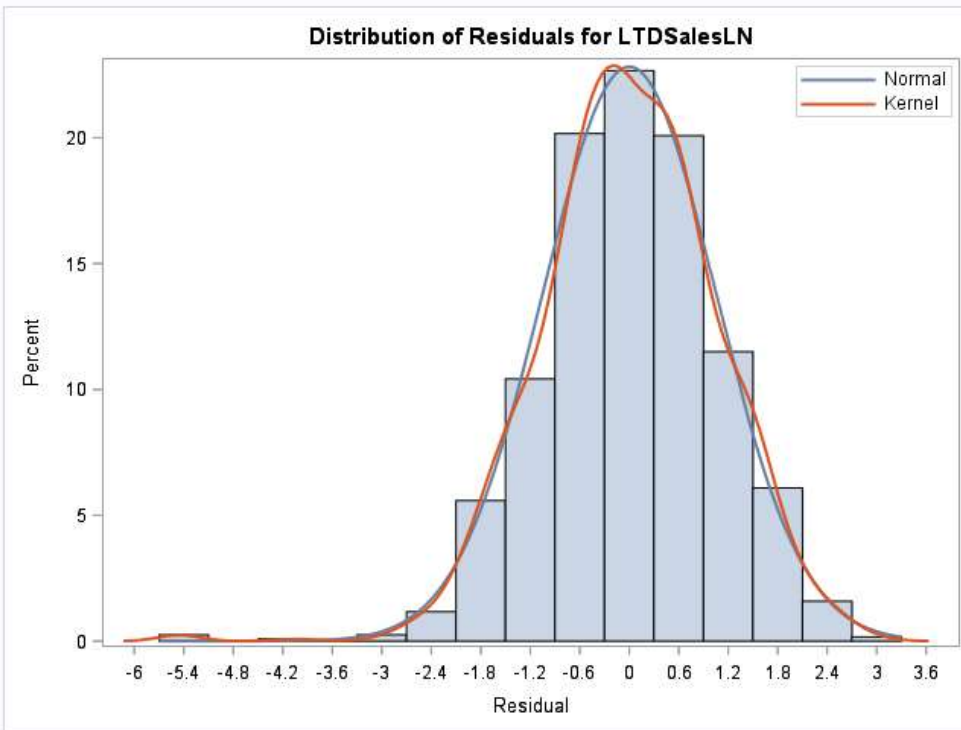
The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: LTDSalesLN

Number of Observations Read	1256
Number of Observations Used	1200
Number of Observations with Missing Values	56

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	1483.39146	247.23191	223.70	<.0001
Error	1193	1318.47994	1.10518		
Corrected Total	1199	2801.87140			

Root MSE	1.05128	R-Square	0.5294
Dependent Mean	1.48834	Adj R-Sq	0.5271
Coeff Var	70.63420		

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
Intercept	1	458.70938	55.31420	8.29	<.0001	0
NumberOfConsoles	1	0.74826	0.04122	18.15	<.0001	1.09070
Review	1	0.04047	0.00257	15.73	<.0001	1.22946
Launch Year	1	-0.23104	0.02762	-8.37	<.0001	1.05745
Launch Month	1	0.02736	0.00897	3.05	0.0023	1.02084
PriorSaleBrand (M)	1	0.01118	0.00143	7.82	<.0001	1.19135
LTD ASP	1	0.03610	0.00315	11.47	<.0001	1.22882



Our model has an adjusted r-square of 0.5281 which means it can explain 52.81% of the variability in the natural logarithm of LTD Sales. Our model is valid as the VIF of the variables are all under 10 which means there is no multicollinearity in the model. As well, the residuals of this model are normally distributed as seen below. Within the model, PriorSaleBrand (M) has the lowest explanatory power as a 1 unit change in the PriorSaleBrand results in a 1% change in the LTD Sales.

Thus, we conclude that the variables that represent the root causes of game success are:

- Number of Consoles
- Review
- Launch Year
- Launch Month
- PriorSaleBrand (M)
- LTD ASP

Data Mining

Confusion Matrix

Using the set of variables left after the correlation analysis, we were able to build a quite accurate model achieving approximately 86% accuracy. The top three determinants of the market performance of video games based on our classification tree analysis (see below) are the following:

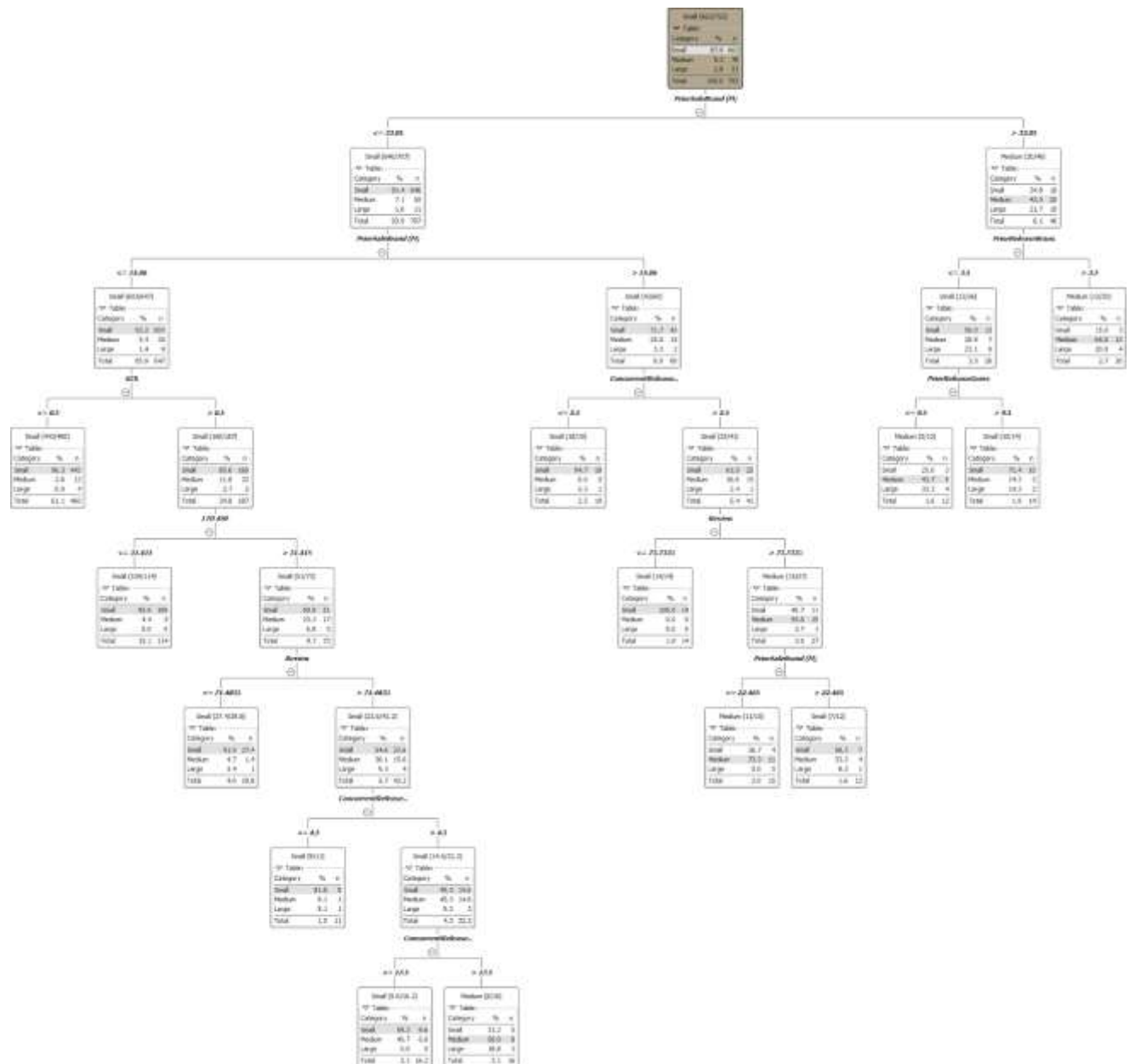
- Prior Sale Brand
- Type of Console
- Concurrent Release Brand

LTD Dollars...	Small	Medium	Large
Small	420	23	0
Medium	41	13	0
Large	4	2	0

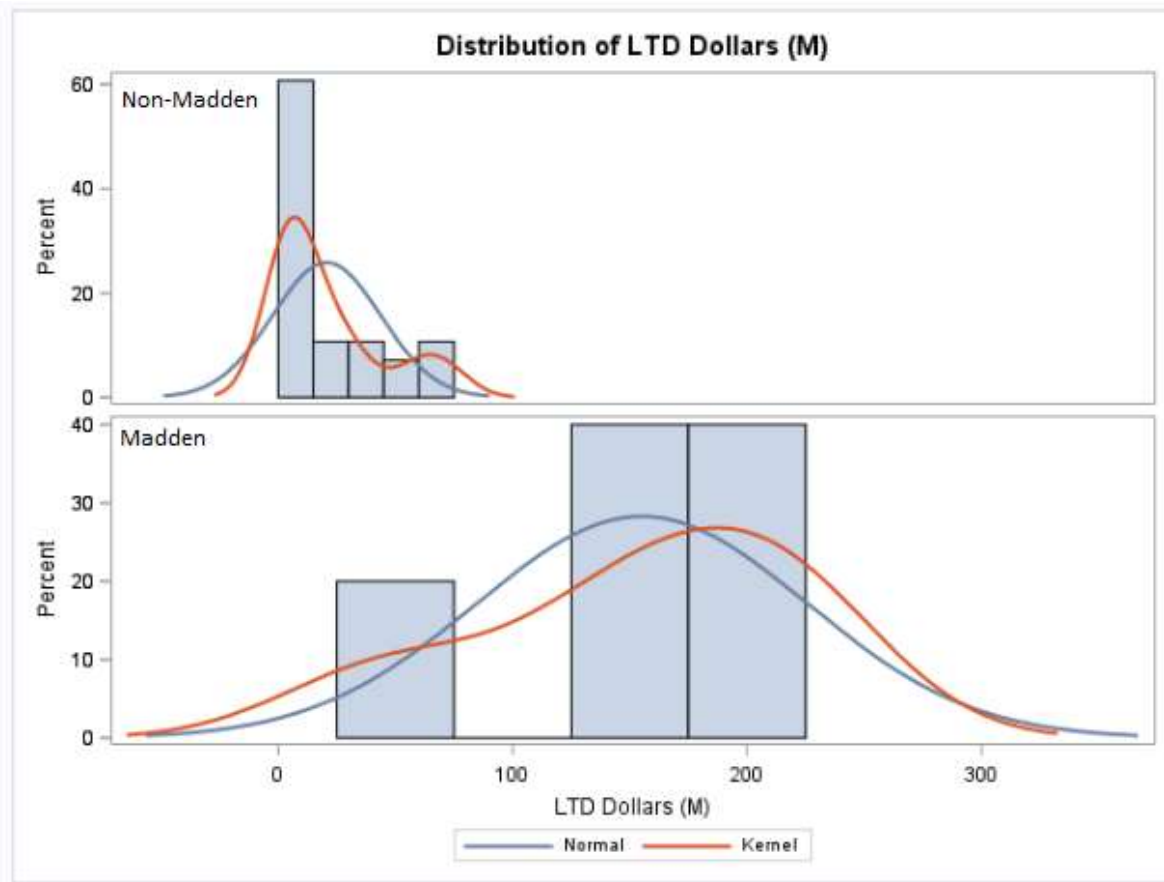
Correct classified: 433	Wrong classified: 70
Accuracy: 86.083 %	Error: 13.917 %
Cohen's kappa (κ) 0.217	

Classification Tree

The main predictor of market performance of video games according to the classification tree analysis is PriorSaleBrand, which is the first year sales of the previously released game of the same brand. If the previous game of the same brand achieved market sales of less than 33.05 million, then the chances of the next game selling above \$30 million is below 8.6%. If the previous game of the same brand achieved market sales of above 33.05 million, then the chances of selling above \$30 million is 65.2%.



“We Need a Madden”



The dummy variable, Madden, was created for all game titles that contain Madden. Note: Only five games contain Madden within the given dataset. The data set was filtered with only football as the sub-genre. From the histograms above, it appears that there is a LTD sales gap between non-Madden and Madden game titles.

t Test

The TTEST Procedure

Variable: LTD Dollars (M)

Madden	N	Mean	Std Dev	Std Err	Minimum	Maximum
Non-Madden	28	20.6366	23.1639	4.3776	0.3331	73.4558
Madden	5	155.1	70.5027	31.5298	46.2746	221.5
Diff (1-2)		-134.5	33.2972	16.1659		

Madden	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
Non-Madden		20.6366	11.6546 29.6186	23.1639	18.3139 31.5293
Madden		155.1	67.5562 242.6	70.5027	42.2405 202.6
Diff (1-2)	Pooled	-134.5	-167.4 -101.5	33.2972	26.6945 44.2680
Diff (1-2)	Satterthwaite	-134.5	-221.6 -47.3688		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	31	-8.32	<.0001
Satterthwaite	Unequal	4.1555	-4.22	0.0124

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	4	27	9.26	0.0002

On average, there is a \$134.5 million difference in LTD Sales between non-Madden and Madden game titles.

Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: LTDSalesLN

Number of Observations Read	33
Number of Observations Used	33

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	65.39989	16.34997	36.11	<.0001
Error	28	12.67721	0.45276		
Corrected Total	32	78.07710			

Root MSE	0.67287	R-Square	0.8376
Dependent Mean	2.71435	Adj R-Sq	0.8144
Coeff Var	24.78946		

Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
Intercept	Intercept	1	973.50243	420.29893	2.32	0.0281	0
Madden	0=Non-Madden, 1=Madden	1	0.81563	0.40547	2.01	0.0540	1.54053
Review		1	0.08166	0.01124	7.26	<.0001	1.52055
Launch Year		1	-0.48842	0.21010	-2.32	0.0276	4.50824
PriorReleaseGenre		1	0.04117	0.01301	3.16	0.0037	4.45337

SAS was used to run a linear regression using the stepwise selection method with a significance level of 0.10 set for variable to enter and remain in the model. As such, after accounting for the effects of review scores, launch year, and prior release genre, Madden games generate 81% more LTD Sales than their "Football" genre counterparts on average. Thus, the Madden license is worth up to \$16.72 million.