# Final Project Proposal

*Max Wagner*

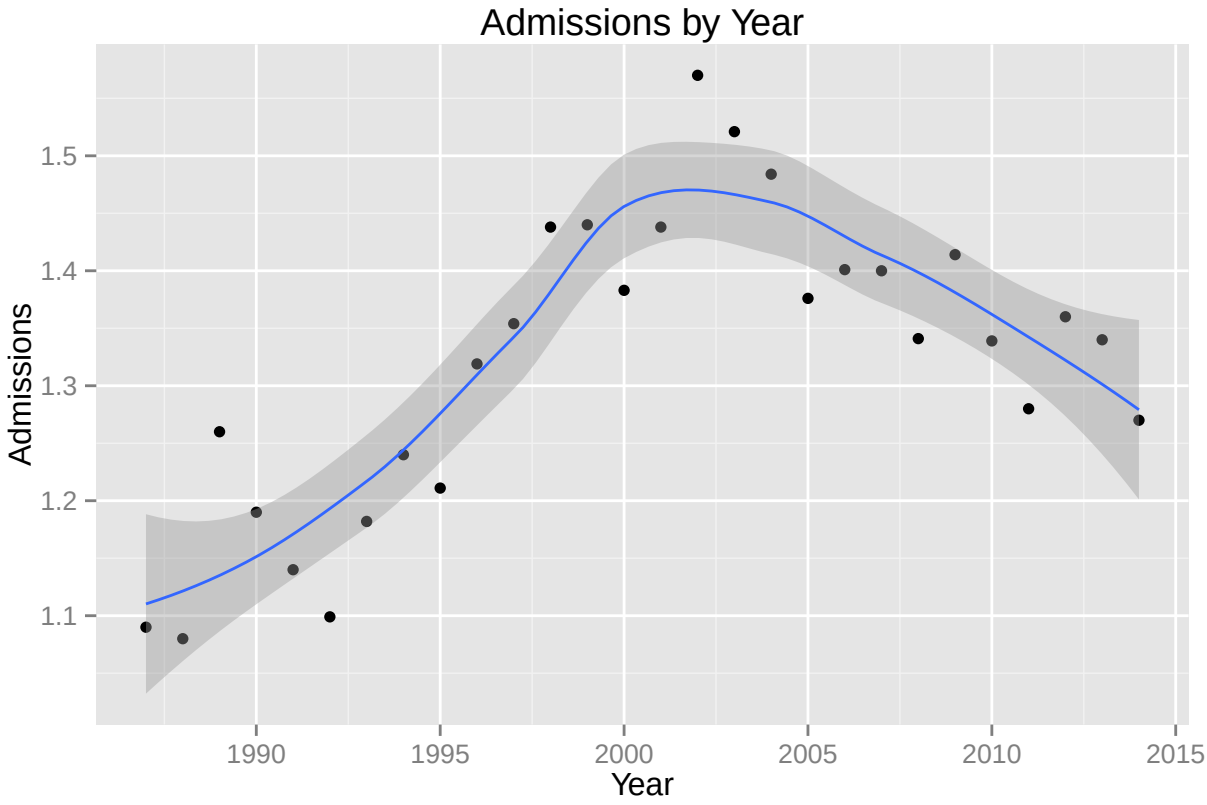
*November 18, 2015*

## Introduction

## Movie Admissions

Taking a table from NATO Online, we can see the total admissions by year in the US and Canada combined. The `Admissions` column is in billions. There is a clear trend of increasing attendance from 1987 to the early 2000's, and a definite decline from then until present day. White Hutchinson Leisure and Learning Group reported a 27% drop in revenue since 2002 to acompany the decline in admissions. There are number of possible reasons for the trend of rise and decline. We'll take a look at a few of the possible factors in the next section.

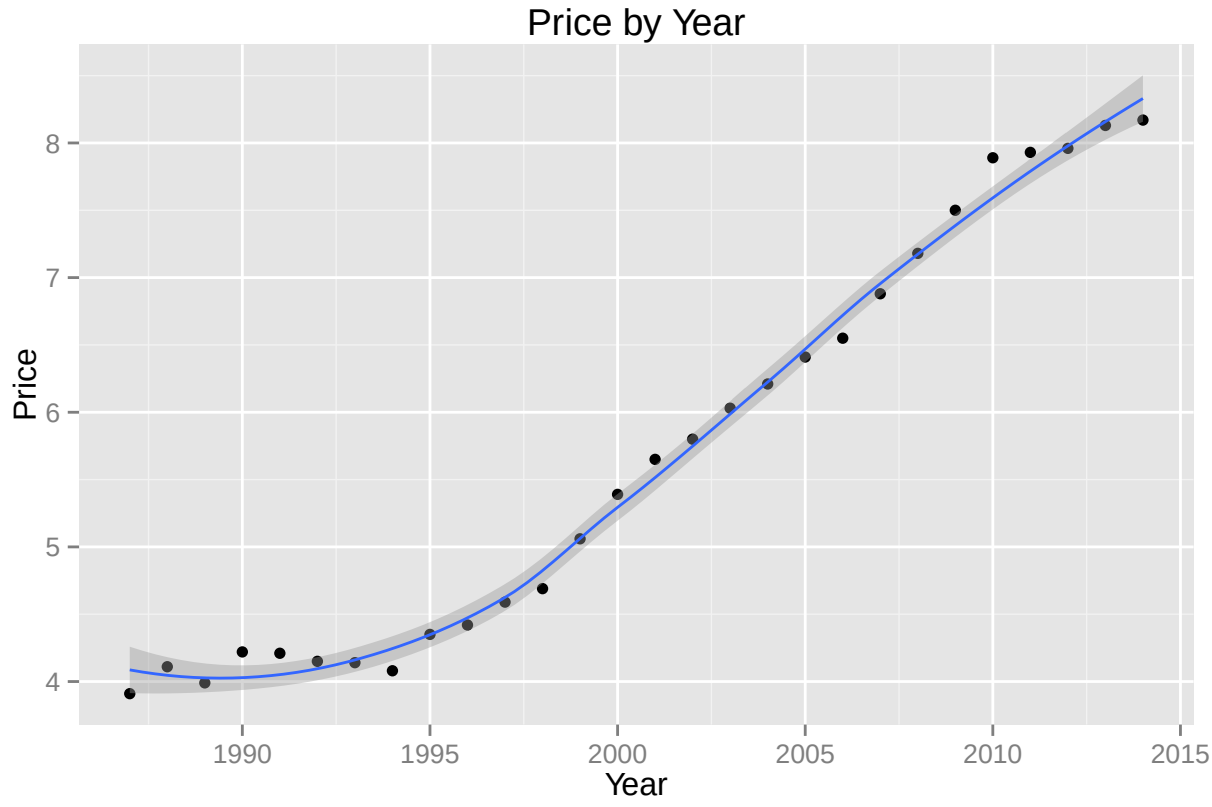
| Year | Admissions | Year | Admissions |
|---|---|---|---|
| 1987 | 1.090 | 2001 | 1.438 |
| 1988 | 1.080 | 2002 | 1.570 |
| 1989 | 1.260 | 2003 | 1.521 |
| 1990 | 1.190 | 2004 | 1.484 |
| 1991 | 1.140 | 2005 | 1.376 |
| 1992 | 1.099 | 2006 | 1.401 |
| 1993 | 1.182 | 2007 | 1.400 |
| 1994 | 1.240 | 2008 | 1.341 |
| 1995 | 1.211 | 2009 | 1.414 |
| 1996 | 1.319 | 2010 | 1.339 |
| 1997 | 1.354 | 2011 | 1.280 |
| 1998 | 1.438 | 2012 | 1.360 |
| 1999 | 1.440 | 2013 | 1.340 |
| 2000 | 1.383 | 2014 | 1.270 |




## Ticket Prices

A table from NATO Online has information on the yearly average US movie ticket price. Ticket prices, as expected, increased linearly for each year. The `Price` column is in USD. The plot below for prices do not indicate any sort of trend that would indicate a reason for the early 2000's peak in attendance. A potential reason for decline directly following could be that customers did not feel movies were worth the increased price.

| | Year | Price | Year | Price |
|---|---|---|---|---|
| 20 | 1987 | 3.91 | 2001 | 5.65 |
| 21 | 1988 | 4.11 | 2002 | 5.80 |
| 22 | 1989 | 3.99 | 2003 | 6.03 |
| 23 | 1990 | 4.22 | 2004 | 6.21 |
| 24 | 1991 | 4.21 | 2005 | 6.41 |
| 25 | 1992 | 4.15 | 2006 | 6.55 |
| 26 | 1993 | 4.14 | 2007 | 6.88 |
| 27 | 1994 | 4.08 | 2008 | 7.18 |
| 28 | 1995 | 4.35 | 2009 | 7.50 |
| 29 | 1996 | 4.42 | 2010 | 7.89 |
| 30 | 1997 | 4.59 | 2011 | 7.93 |
| 31 | 1998 | 4.69 | 2012 | 7.96 |
| 32 | 1999 | 5.06 | 2013 | 8.13 |
| 33 | 2000 | 5.39 | 2014 | 8.17 |




## Internet Traffic

The following table from Cisco shows a relative estimate of internet bandwidth used. The first row is per day, the second is per hour, and the rest are per second. The bandwidth used from 1992 to 1997 is relatively small. The amount begins to be substantial in 2002, and only increases from there. The estimated usage in 2019 is a staggering 51,694% increase over the usage in 2002.

| Year | Bandwidth |
|---|---|
| 1992 | 100 GBpd |
| 1997 | 100 GBph |
| 2002 | 100 GBps |
| 2007 | 2000 GBps |
| 2014 | 16144 GBps |
| 2019 | 51794 GBps |

The following is a table from Sandvine, outlining the overall bandwidth distribution from 2013. Netflix, Youtube, Amazon Video, and Hulu are all inside the top 10 for total bandwidth, with Netflix accounting for almost 32% of all data transferred. What this means is that people are watching a substantial amount of video every day, and aren't going to the theatre to do it. BitTorrent and HTTP are important to the overall view as well. Not all file sharing and BitTorrent traffic are media files, but the majority are, and together they account for 14% of all traffic.

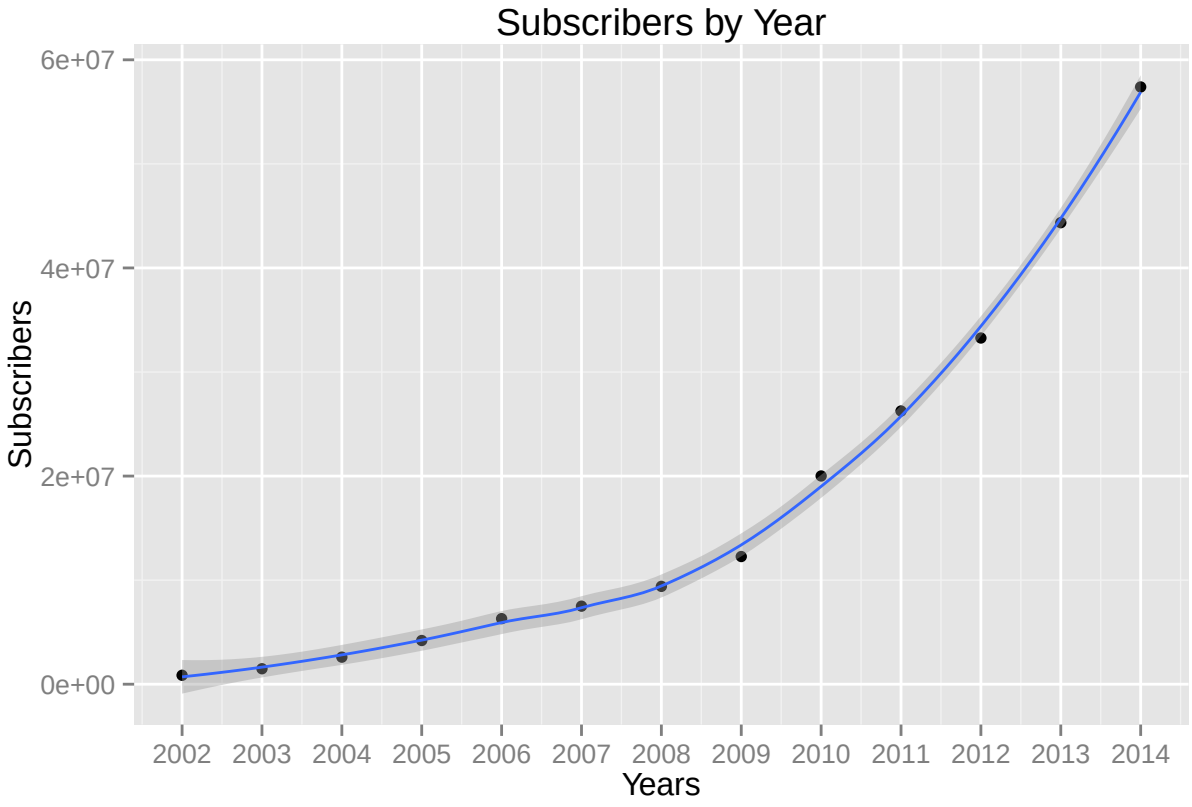
| | Upstream | | Downstream | | Aggregate | |
|---|---|---|---|---|---|---|
| Rank | Application | Share | Application | Share | Application | Share |
| 1 | BitTorrent | 36.35% | Netflix | 31.62% | Netflix | 28.18% |
| 2 | HTTP | 6.03% | YouTube | 18.69% | YouTube | 16.78% |
| 3 | SSL | 5.87% | HTTP | 9.74% | HTTP | 9.26% |
| 4 | Netflix | 4.44% | BitTorrent | 4.05% | BitTorrent | 7.39% |
| 5 | YouTube | 3.63% | iTunes | 3.27% | iTunes | 2.91% |
| 6 | Skype | 2.76% | MPEG - Other | 2.60% | SSL | 2.54% |
| 7 | QVoD | 2.55% | SSL | 2.05% | MPEG - Other | 2.32% |
| 8 | Facebook | 1.54% | Amazon Video | 1.61% | Amazon Video | 1.48% |
| 9 | FaceTime | 1.44% | Facebook | 1.31% | Facebook | 1.34% |
| 10 | Dropbox | 1.39% | Hulu | 1.29% | Hulu | 1.15% |
| | | 66.00% | | 76.23% | | 73.35% |

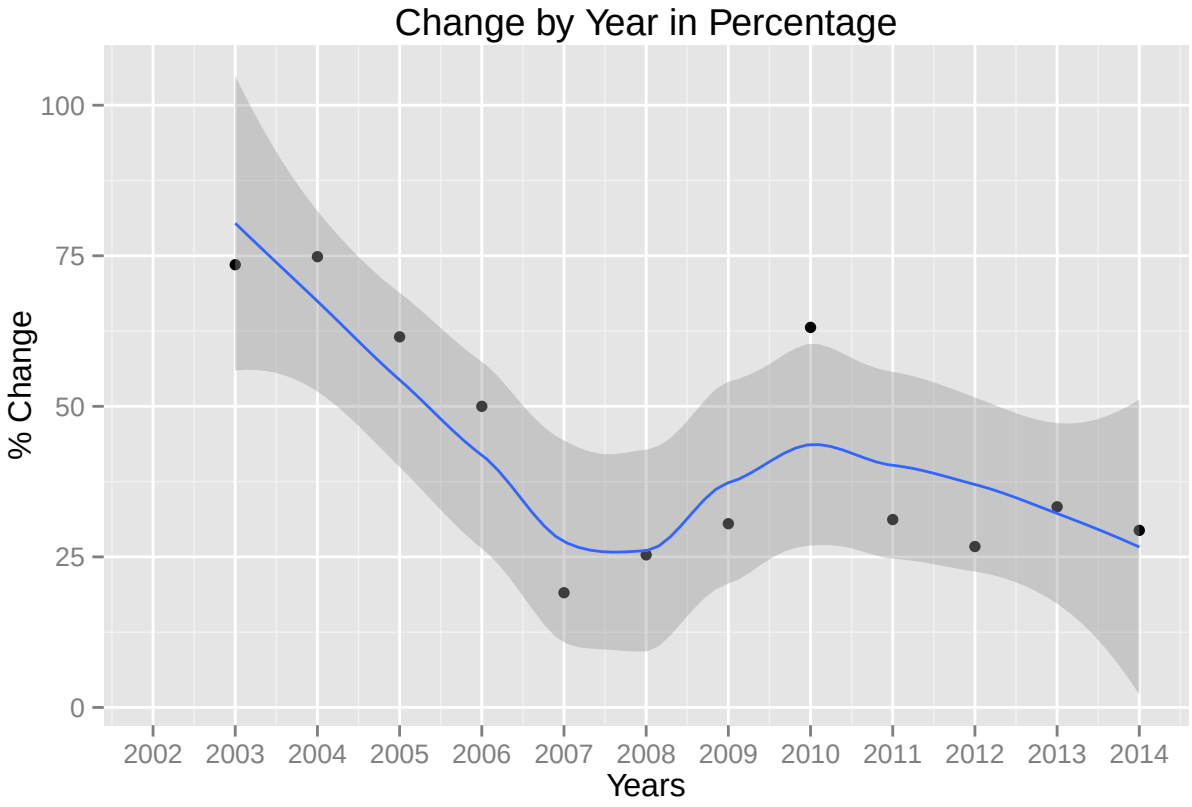
sandvine

## Netflix

Taking a closer look at Netflix in particular, and grabbing their subscriber data directly from their revenue excel documents, we can see that their userbase has substantially increased each year following 2002.

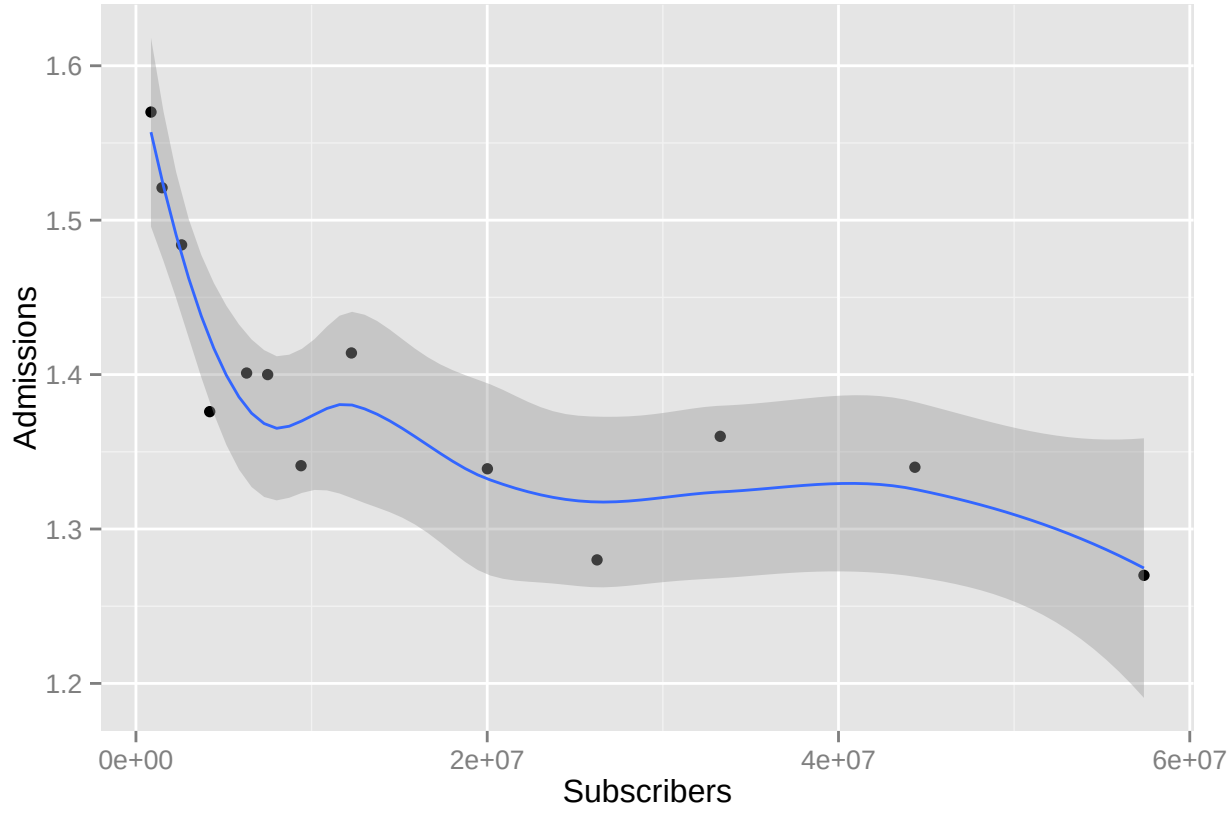
| Year | Subscribers | % Change |
|---|---|---|
| 2002 | 857000 | NA |
| 2003 | 1487000 | 73.51225 |
| 2004 | 2600000 | 74.84869 |
| 2005 | 4200000 | 61.53846 |
| 2006 | 6300000 | 50.00000 |
| 2007 | 7500000 | 19.04762 |
| 2008 | 9400000 | 25.33333 |
| 2009 | 12268000 | 30.51064 |
| 2010 | 20010000 | 63.10727 |
| 2011 | 26253000 | 31.19940 |
| 2012 | 33267000 | 26.71695 |
| 2013 | 44350000 | 33.31530 |
| 2014 | 57391000 | 29.40474 |


Subscribers by Year
6e+07
4e+07
2e+07
0e+00
Subscribers
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Years


Change by Year in Percentage
100
75
50
25
0
% Change
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Years

## Netflix and Theatres

Below is a plot, correlation, and summary stats on admissions numbers for theatres and netflix usage. The inital look is that as subscribers to netflix increase, the admissions to theatres go down.




```
##
##  Pearson's product-moment correlation
##
## data:  netflix_tickets$Subscribers and netflix_tickets$Admissions
## t = -3.5237, df = 11, p-value = 0.004768
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.9128949 -0.2959170
## sample estimates:
##        cor
## -0.7281761
```

The correlation value above of -0.7281761 indicates that that the correlation is significant, while can also be gained from the p-value being below 0.05.

```
##
## Call:
## lm(formula = netflix_tickets$Subscribers ~ netflix_tickets$Admissions)
##
## Residuals:
```

```
##       Min        1Q    Median        3Q       Max
## -15543993  -8691426  -1851095   9829591  21956502
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  223424510   58584966   3.814  0.00287 **
## netflix_tickets$Admissions -148023631   42008393  -3.524  0.00477 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12900000 on 11 degrees of freedom
## Multiple R-squared:  0.5302, Adjusted R-squared:  0.4875
## F-statistic: 12.42 on 1 and 11 DF,  p-value: 0.004768
```

The important note here is that the p value is much lower than the acceptable 0.05, and that the R-squared value is relatively high.

## Sources


- http://rpubs.com/catlin/rvest
- http://natoonline.org/data
- https://www.whitehutchinson.com/blog/2015/01/movie-attendance-continues-its-12-year-decline/
- http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html
- https://www.sandvine.com/downloads/general/global-internet-phenomena/2013/2h-2013-global-internet-phenomena-rep pdf
- http://ir.netflix.com/results.cfm