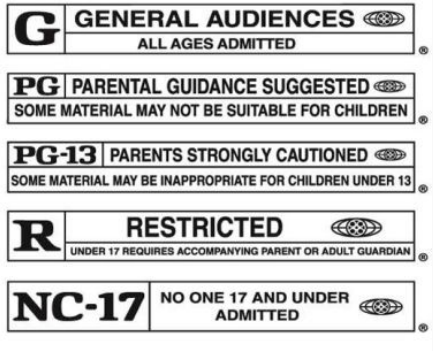
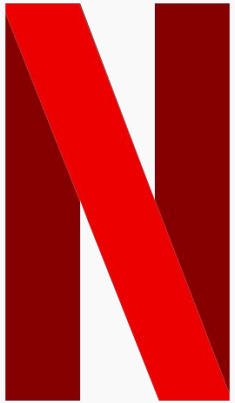
YouTube Video Link

https://youtu.be/OfeyO59Vy0M

CS108 Final Project

Chase Lenhart





Introduction

1. Netflix stock prices
2. In relation to movie/tv show ratings

Question being asked:

Do movie/tv show ratings released on Netflix have an affect on Netflix stock prices?

Stock Dataset Introduction:

This dataset includes the stock prices for Netflix (Ticker NFLX). The dataset includes indices of days (daily stock prices). The various columns included are High (high of the stock on the day), Volume (how many shares traded on the day), and more. This is interesting to me due to my interest in business, and further stocks.

Rating Dataset Introduction:

This dataset includes various movie and tv show releases on Netflix, over 7000 of them to be exact. The indices are in days as well making it comparable to the stock dataset. The dataset includes columns such as the director of the movie or tv show, the rating of the movie or tv show, and more. This dataset is interesting because it can be used to compare the ratings of movies and tv shows on Netflix to stock prices.

1

Netflix Stock Prices

yahoo! finance

Netflix Stock Price Dataframe

Current

- High
- Low
- Volume
- Open
- Close

Added

- High_chg
- High_rm
- High_rm_pct
- High_pct
- Stock_Return
- Stock_Return_pct
- Volume_pct

Stock Dataframe .describe()

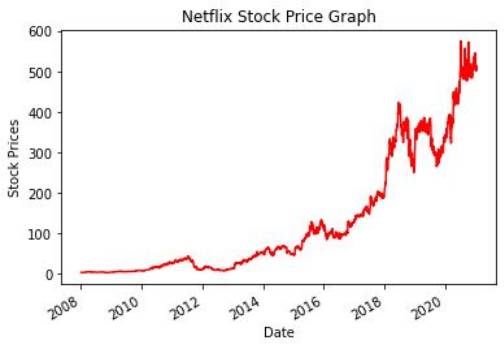
	High	Low	Volume
count	3255.000000	3255.000000	3.255000e+03
mean	128.397575	124.296519	1.899124e+07
std	148.227791	143.438868	2.022961e+07
min	2.817143	2.557143	1.144000e+06
25%	14.844285	14.090000	7.395650e+06
50%	60.447144	58.861427	1.265320e+07
75%	190.580002	186.394997	2.315335e+07
max	575.369995	541.000000	3.155418e+08

```
In [365]: df_stock.columns
Out[365]:
Index(['High', 'Low', 'Volume', 'Open', 'Close', 'Dividends', 'High_chg',
       'High_rm', 'High_pct', 'Volume_pct', 'High_rm_pct'],
      dtype='object')
```

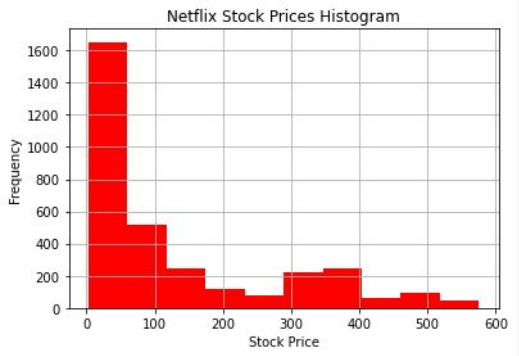
Netflix Stock Price Plots

Plots

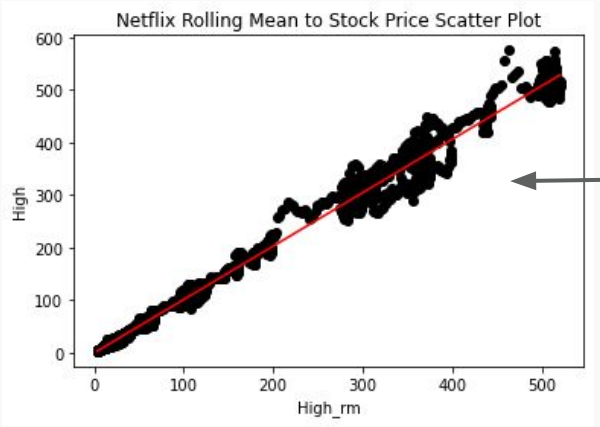
- Line graph
- Histogram
- Regression of stock prices and rolling mean
- **KDE curve**
- **Hexbin plot**



This graph shows the stock price of Netflix stock since 2008. The stock experienced overall growth with the biggest rise being during the pandemic.



This graph shows the histogram (frequencies) of stock prices of Netflix stock since 2008. The stock price fell in the range of $0 to $100 for the majority of the range of dates.



This graph shows the scatter plot of the rolling mean of Netflix stock prices to the Netflix stock price, as well as the regression line. The rolling mean is a good indicator of Netflix stock price.

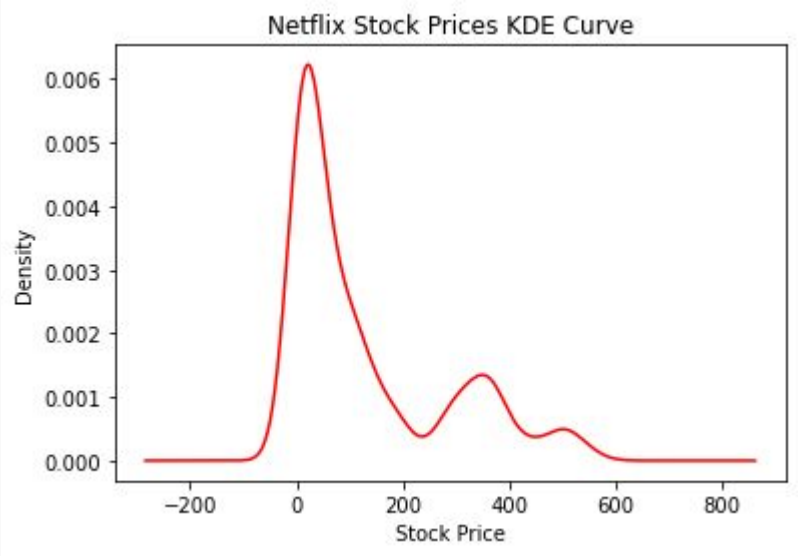
OLS Regression Results

Dep. Variable:	High	R-squared (uncentered):	0.996
Model:	OLS	Adj. R-squared (uncentered):	0.996
Method:	Least Squares	F-statistic:	7.279e+05
Date:	Sat, 09 Dec 2023	Prob (F-statistic):	0.00
Time:	21:23:21	Log-Likelihood:	-12988.
No. Observations:	3255	AIC:	2.598e+04
Df Residuals:	3254	BIC:	2.598e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
High_rm	1.0152	0.001	853.199	0.000	1.013	1.018

Omnibus:	528.929	Durbin-Watson:	0.111
Prob(Omnibus):	0.000	Jarque-Bera (JB):	7637.229
Skew:	0.292	Prob(JB):	0.00
Kurtosis:	10.481	Cond. No.	1.00

Netflix Stock Price Plots



KDE Curve

A KDE curve shows the density (occurence weight) of a column of values in a dataframe. In this case, the Netflix stock density appears to be between 0 and 200 for the most part, with a spike in between 200 and 400.

Hexbin Plot

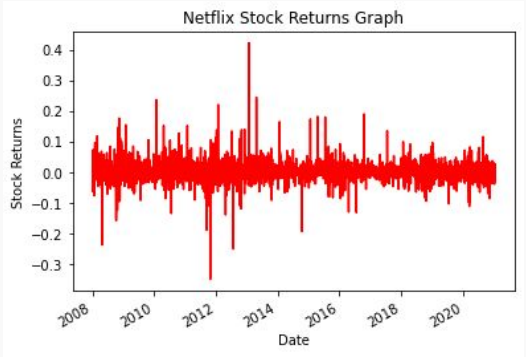
A hexbin plot shows the density of data using hexagons in a 2D space. In this case, the Netflix stock price is shown against the volume of stock shares. With high density in the bottom left corner and going slightly rightward, the stock price seems to be highly dense in the lower stock price and medium volume range.

Netflix Stock Price Operations

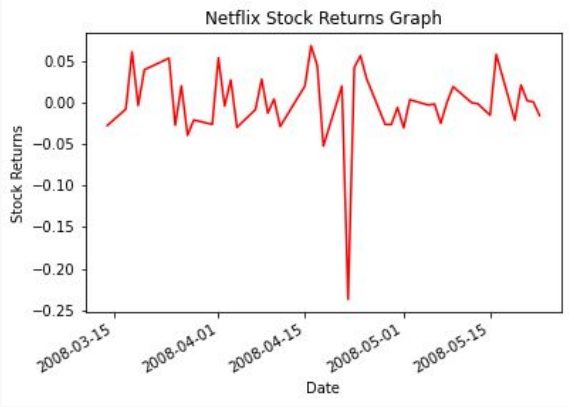
Operations

- Probabilities of stock rising or falling
- Return on stock prices

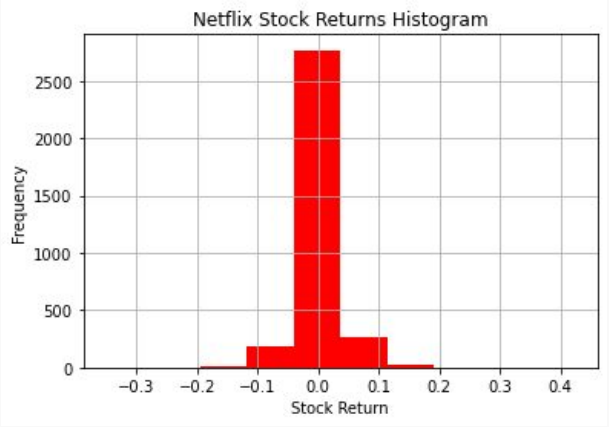
```
In [68]: probabilities(df_combo)
Probability of Stock Price Rising = 0.5378048780487805
Probability of Stock Price Falling = 0.4621951219512195
```



This line graph shows the various returns of Netflix stocks over time. It seems the returns were most volatile from 2012 - 2014.



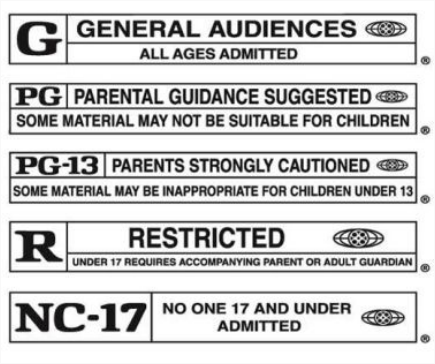
This is the line graph of stock returns for Netflix stock for the date range March 15, 2008 - May 15, 2008. This shows the returns much more clearly.



This histogram shows the stock returns of the Netflix stock. According to the histogram, the stock returns sway slightly leftward of 0, having most returns being negative.

2

Netflix TV/Movie Ratings



TV/Movie Rating Dataframe

Current

- Title
- Type
- Director
- Date Added
- Rating
- Duration

Added

- Rating_Count
- Rating_Cum
- Rating_Cum_pct
- Rating_chg

Rating Dataframe .describe()

	title	rating	director
count	7787	7780	5398
unique	7787	14	4049
top	To and From New York	TV-MA	Raúl Campos, Jan Suter
freq	1	2863	18

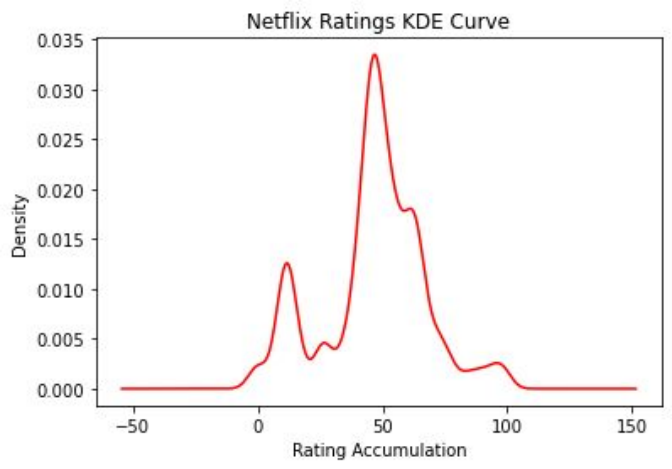
```
good_movies = ['TV-PG', 'TV-14', 'PG', 'TV-G', 'TV-Y']
bad_movies = ['TV-MA', 'NR', 'R']
```

```
In [391]: df_data.columns
Out[391]:
Index(['title', 'type', 'director', 'date_added', 'rating', 'duration',
       'Rating_Count', 'Rating_Cum', 'Rating_Cum_pct', 'Rating_chg'],
      dtype='object')
```

TV/Movie Rating Plots

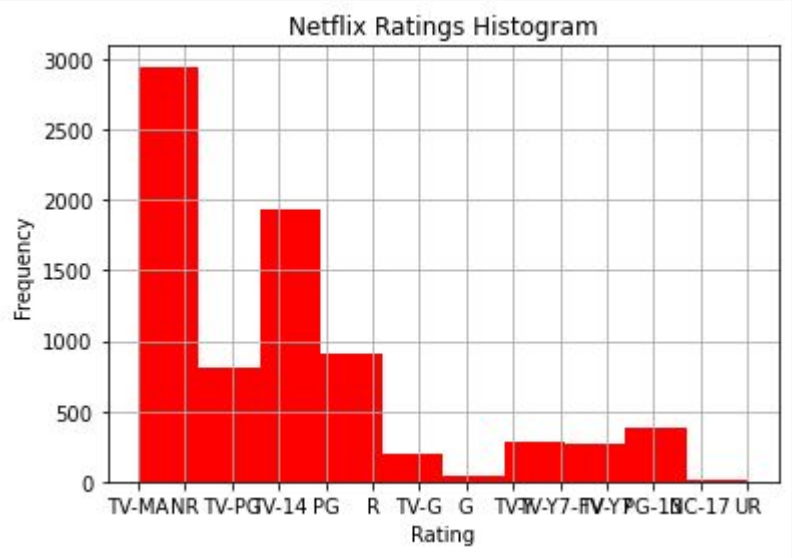
Plots

- Histogram
- **KDE curve**



KDE Curve

This KDE curve shows the rating accumulation percentage of Netflix tv show and movie ratings. With the percentage having spikes at 50% and below, the ratings released are frequently under the highest rating accumulation (good ratings).



This histogram shows the frequencies of Netflix tv show and movie ratings. The frequencies are high for TV-MA and TV-14 ratings.

Rating Data Operations

Operations

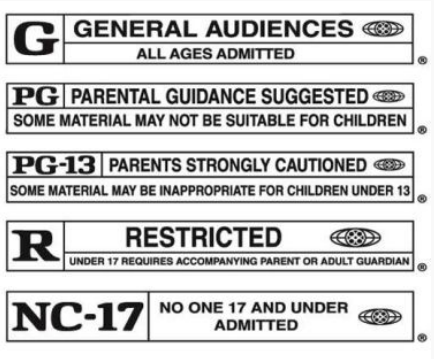
- Removing duplicate dates
- Probabilities of movie/tv ratings
- Best director

```
Probability of Rating being Positive = 0.365853658536585 36
Probability of Rating being Negative = 0.6341463414634146
```

```
In [203]: best_director(df_combo)
Best Director = Sorin Dan Mihalcescu
Stock Changes for Sorin Dan Mihalcescu = $0.24955573485257287
```

3

Combination of Both Dataframes



Stock Price & Rating Plots/Operations

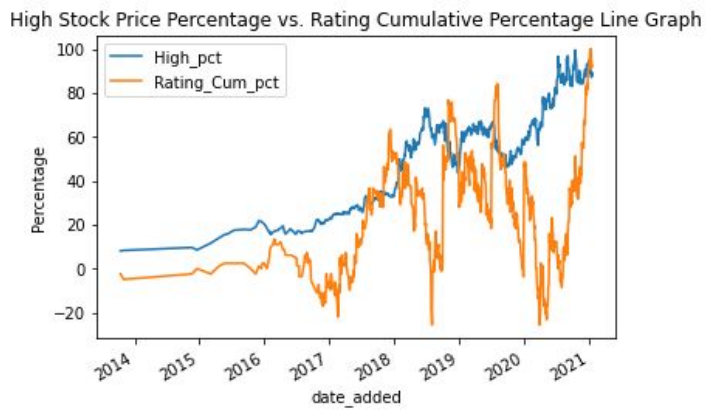
Plots/Operations

- New dataframe combining the two
- Conditional probabilities
- Plot stock prices to cumulative data
- **KDE curve**
- **Hexbin plot**
- Scatter plot for ratings and stock data

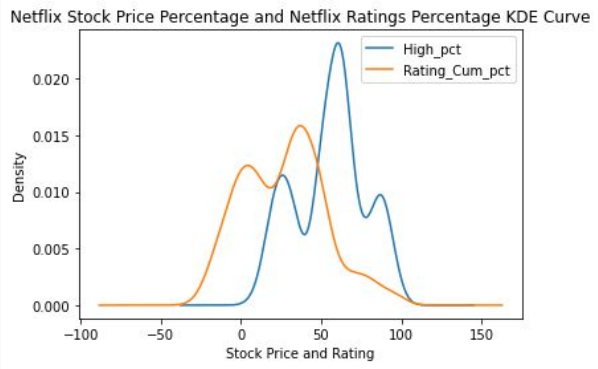
```
In [123]: df_combo.columns
Out[123]:
Index(['title', 'type', 'director', 'date_added', 'rating', 'duration',
       'Rating_Count', 'Rating_Cum', 'Rating_Cum_pct', 'Rating_chg',
       'Rating Count', 'High', 'High_pct', 'High_chg', 'Stock_Return',
       'Stock_Return_pct'],
      dtype='object')
```

```
Probability of Rating Rises Given Stock Rises = 0.36507936507936506
Probability of Stock Rises Given Rating Rises = 0.5366666666666667
```

This line graph shows both the stock price percentage and the rating cumulative percentage. When the rating percentage falls (worse rated movies) the stock price rises.

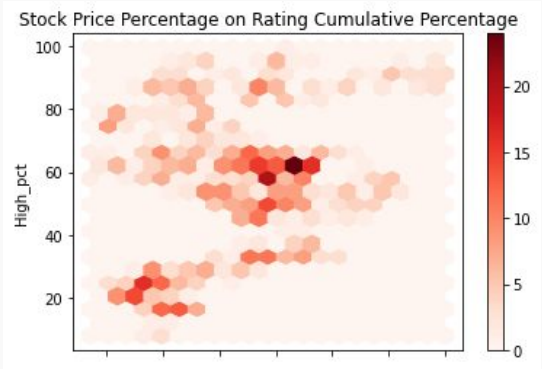


Stock Price & Rating Plots/Operations



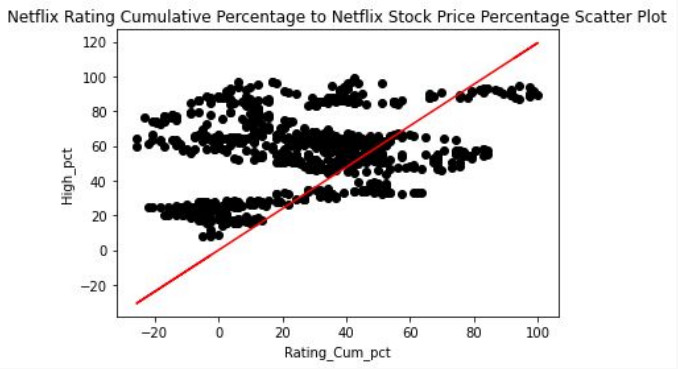
KDE Curve

This KDE curve shows both the KDE curves for stock percentage and rating cumulative percentage. It shows that the stock price density are higher than the rating cumulative percentage.



Hexbin Plot

This hexbin plot shows the rating cumulative percentage on the x axis and the stock price percentage on the y axis. The majority of density fall in the middle and left side of the graph showing that with lower rated movies come higher stock prices.



This scatter plot shows rating cumulative percentage and its effect on stock percentage. The regression line also shows an upward trend, but the data points do not show this correlation. The r^2 for this scatter plot is low giving a lack of correlation.

New Functions and Data Analytics Learned

Function	Purpose
df.drop	Drops a row from a dataframe given an index
df.sort_index	Sorts a dataframe in ascending or descending order based on parameters given
df.apply	Applies a function (a parameter) to each row in dataframe
df.dropna	Drops all rows that have any na values in specified column
df.pct_change	Finds the percent change from row to row for a column
df.cumsum	Finds the cumulative sum for the column and stores it into each row for the sum leading up to that row

New Plots/Pandas Used

Stock Returns

Plotting the returns of a stock using the pct_change function

Hexbin Plot

Plotting two columns using hexagons showing density

KDE Curve

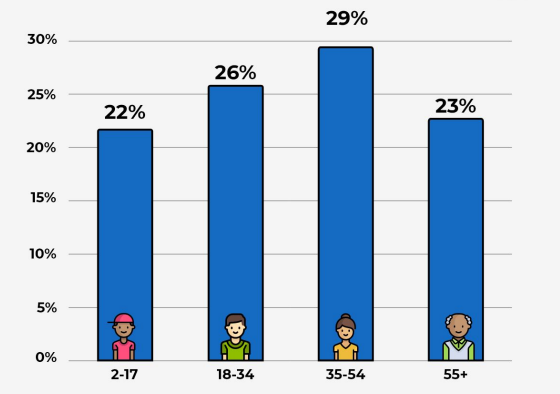
Plotting various columns' densities on a line plot

Optimization of Dataframe

Comparing each row of dataframe to find best director on Netflix based on stock price changes

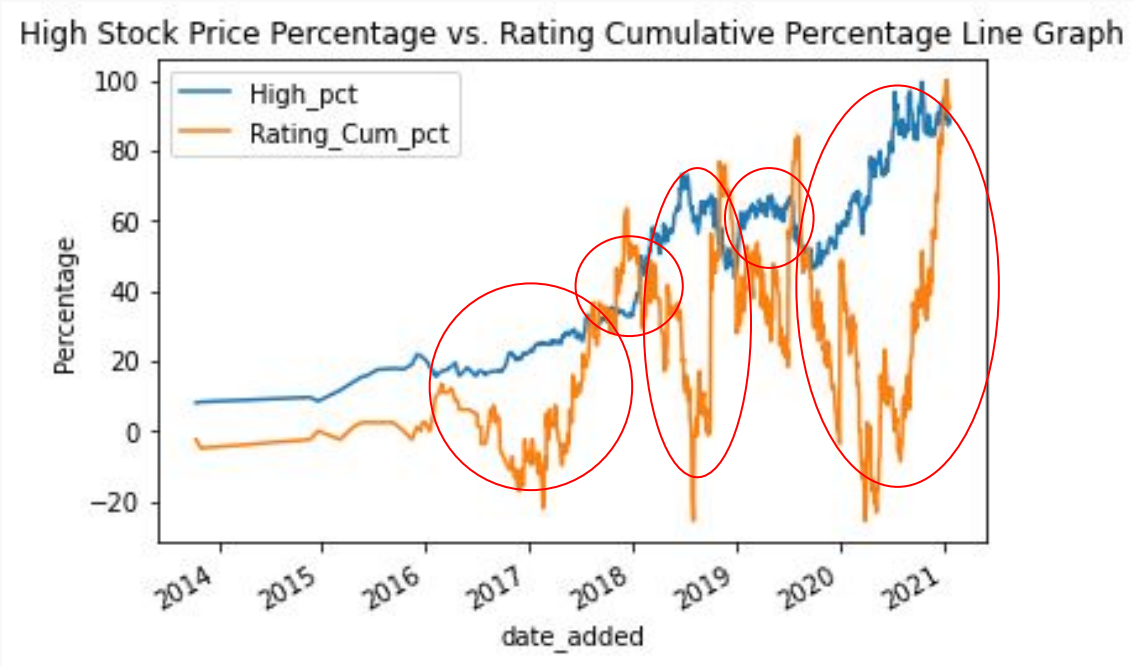
Answering the Question:

Do movie/tv show ratings released on Netflix have an affect on Netflix stock prices?



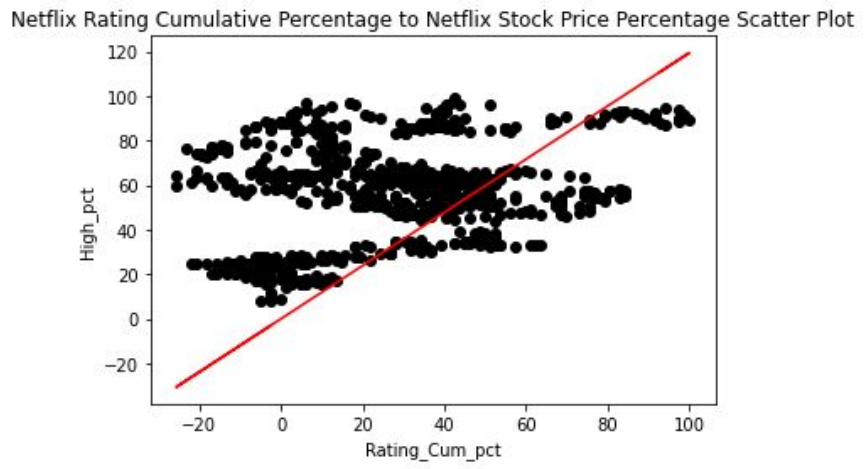
Segmentation = Age demographics

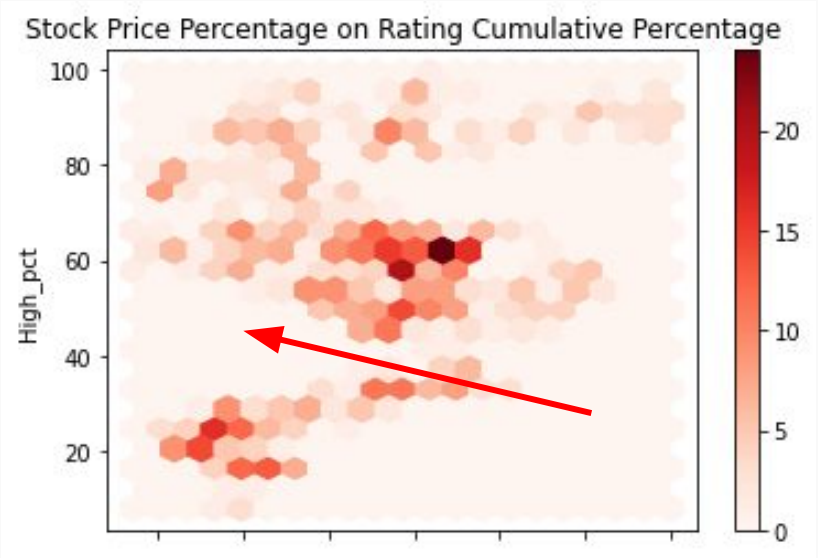
Target Market = Gen Z & Millennials



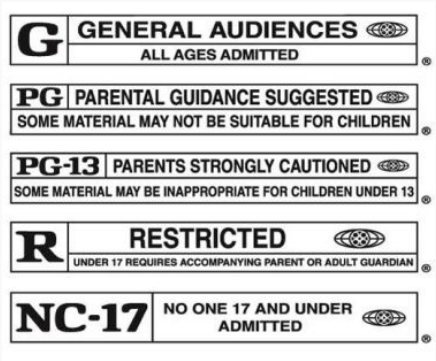
Answering the Question:

Do movie/tv show ratings released on Netflix have an affect on Netflix stock prices?





Thank you


G
GENERAL AUDIENCES
ALL AGES ADMITTED
PG
PARENTAL GUIDANCE SUGGESTED
SOME MATERIAL MAY NOT BE SUITABLE FOR CHILDREN
PG-13
PARENTS STRONGLY CAUTIONED
SOME MATERIAL MAY BE INAPPROPRIATE FOR CHILDREN UNDER 13
R
RESTRICTED
UNDER 17 REQUIRES ACCOMPANYING PARENT OR ADULT GUARDIAN
NC-17
NO ONE 17 AND UNDER ADMITTED