



## Building Lego's Marketing Strategy

Rachel Burns, Alexandra Iov, Adara Liao

### Introduction

Lego sets have captured the imaginations of generations. Since inception 80 years ago, the company has expanded from simple bricks, to theme parks, movie and video game tie-ins and even its own movies. Today, it employs 17,000 employees across the world.

We will investigate Lego's rich history of its beloved Lego bricks and delve into how its product's themes and colors, movie strategy have changed over time, and how those factors created financial success for Lego company. We will conclude with our take on Lego's marketing strategy.

### Data Sets

- Kaggle Dataset with 9 separate files
  - Source: <https://www.kaggle.com/ratman/lego-database>
  - Consolidated dataset has 21 columns and 580,251 entries. Some values are missing that did not match between tables, and some part names we chose to concatenate because one part id corresponded to multiple rows with a, b, c etc appended to the part id in the part lookup table.
- Financial data contains 27 rows with 15 entries
  - <https://www.lego.com/en-us/aboutus/lego-group/annual-report/>
- Movie data from boxofficemojo.com with 100 entries for each year since 1980 with 2 crucial columns

- <http://www.boxofficemojo.com/yearly/chart/?yr=2017&view=releasedate&view2=domestic&sort=gross&order=DESC&&p=.htm>

## Lego Financials

### Methodology:

#### Data Cleaning

As the financial data was buried in PDFs, we opened individual years' financial reports and stored select information in a CSV file. There was not much data cleaning in the Pythonic sense as we pre-selected relevant financial data to store in CSV file.

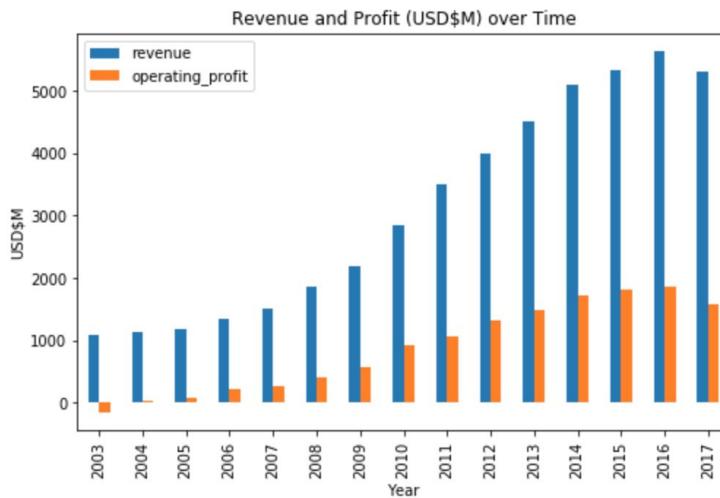
#### Data Analysis

We joined the financial dataset and Kaggle's dataset by year (with Kaggle dataset being aggregated by year, thus losing granular details on theme, set names, part ID etc). This will be used to examine relationship between operating expenses and number of sets maintained by the Lego production team on any year, which is a proxy of how complicated manufacturing operations are. Additionally, we created additional columns in financial dataset to measure year-over-year change in revenue and profitability.

### Findings:

#### Industry-leading revenue growth and profitability

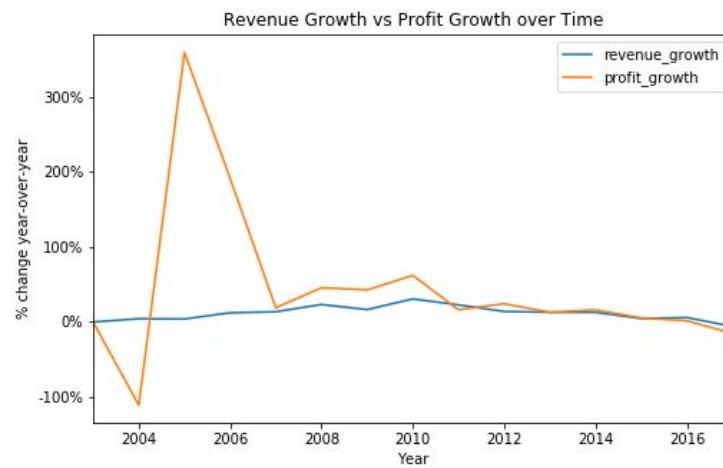
First we looked at overall performance of Lego as a business by plotting revenue and operating margin over time (2003-2017).



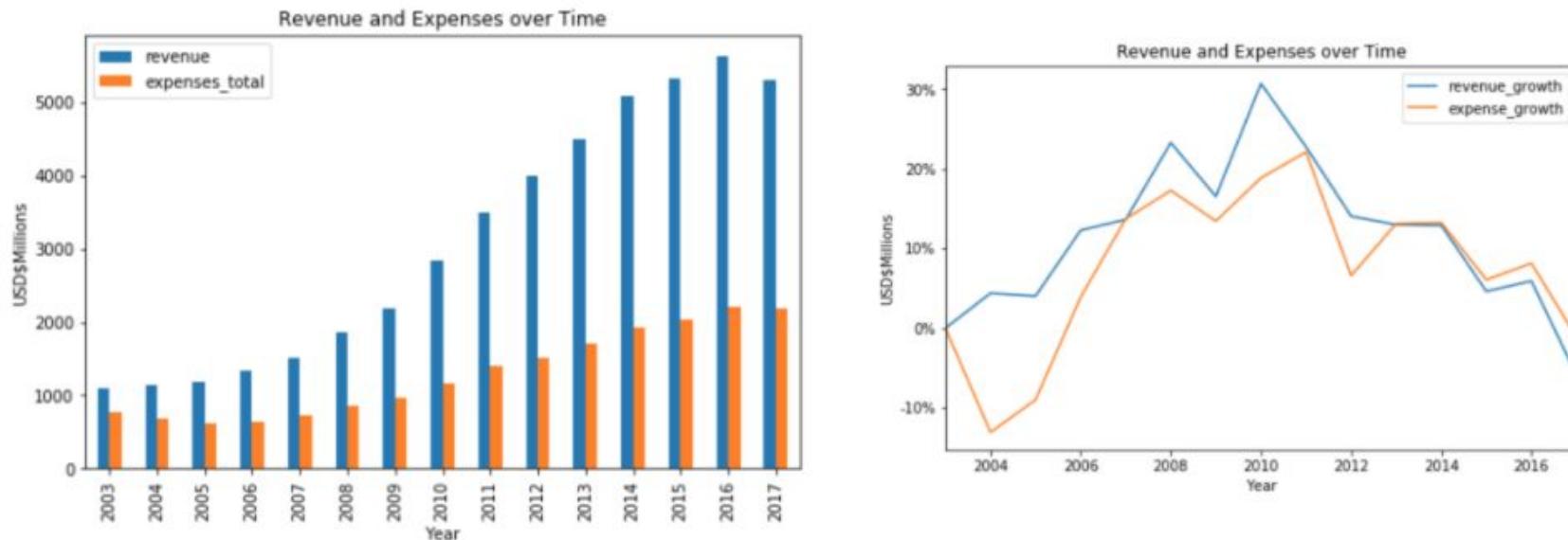
The revenue growth rate in the past 15 years was impressive at average of 12% every year and average profitability of 24% of revenue. This is better than leading global toy companies with similar revenue of \$5billion per year, such as Mattel and Hasbro. Mattel average 10% profitability, while Hasbro average profitability was 15%. Mattel and Hasbro, which sells just toys, are 40%-50% less profitable than Lego.

### **Profitability growth rate varies more than revenue growth rate**

Profit growth rate swings more widely than revenue growth rate. The large changes in profit growth rate in 2004 and 2005 indicate that there is a sensitive break-even revenue level (\$1.2 billion). After this level, profitability moves at same rate as revenue increase.



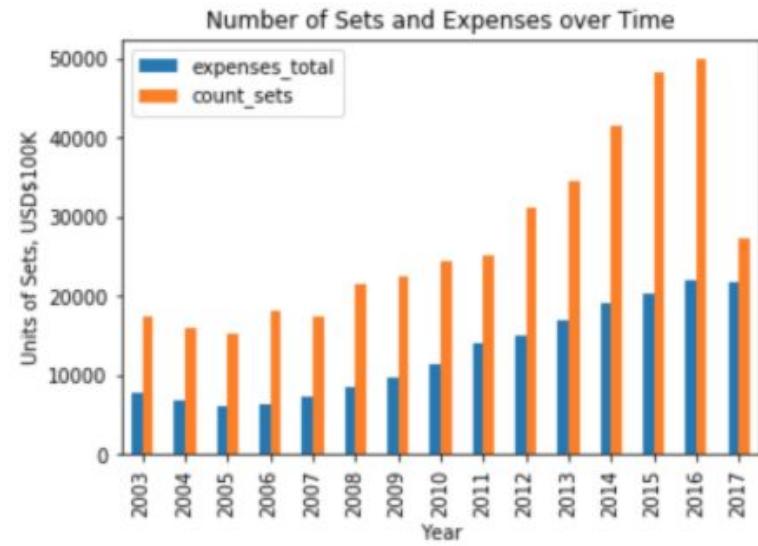
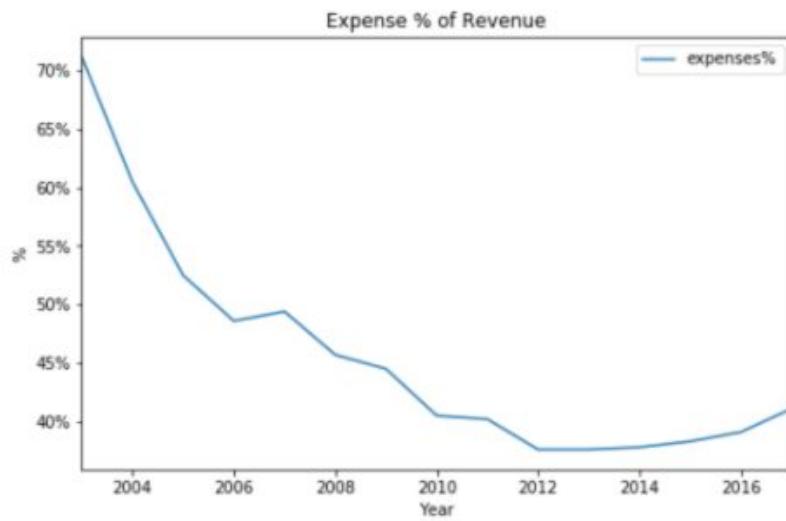
**Potential reason for profitability margins that are higher than peers and keep growing**



Examining why Lego has been consistently able to increase profitability, from 6% in 2003 to 30% in 2017. One reason is the disciplined expenses. The expenses grow much less than revenue. To achieve a revenue growth, Lego needs to spend less in additional resources.

A strong advantage that Lego has over traditional toy companies such as Mattel and Hasbro is that Lego has a profitable and strong movie franchise for its products and produces chart-topping movies. For the years that financial data is available, the total revenue divided by number of lego sets produced is \$33,866, implying that there is a high percentage of movie-related revenue in the sales numbers reported.

**Expenses are kept under control even as number of sets that needed to be produced increase every year.**

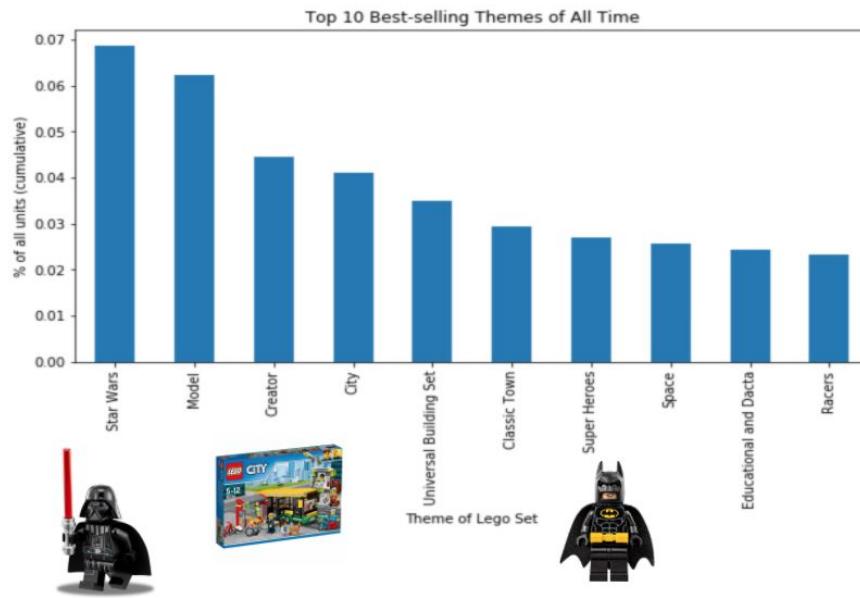


The economies of scale can be seen by the sharp drop in expenses as % of revenue. The discipline in operational efficiency can be contrasted with the growing number of sets that Lego produces every year. The number increases in red bar, while expenses do not grow as much.

The number of sets that Lego produces at any one time has grown. Despite the complexity of operations that arises, Lego's operations team has done a good job.

### Among best-selling themes of all time, Star Wars and SuperHeroes appeared, highlighting importance of movie franchise tie-ins

Since 1950, Lego has produced 2M sets. A prominent strategy of Lego is to differentiate their products with movie franchise tie-ins. We looked at the cumulative number of sets sold by theme, but since Lego does not publish number of sets sold, we used number of sets manufactured (in Kaggle dataset) as a proxy.



The top 10 best sellers account for 40% of all units ever sold, indicating the importance of having a few best hitting themes. From analysis, Star Wars is the best-selling theme of all time, with 132K units produced and 7% of all-time sales. A relatively newcomer (Star Wars first appeared since 2009), Star Wars has overtaken classic themes such as Model (since 1983) and Creator (since 2001).

## Colors and Themes - Rachael Burns

### Methodology:

#### Data Cleaning

I denormalized the csv files from the Kaggle database to the part level using json manipulations- I wanted to be able to see exactly

how the data related or not, and I was not yet clear on how easy it would have been in pandas. One thing I found was that the part number did not match between the associative table for parts to sets and the lookup table for part names and colors- some matched to the id plus multiple letters, for example two parts of a mast. We decided to match with regular expressions and concatenate the part names, and in every case the parts had the same color. There were some values with missing years and some with [no color] as the color, which I stripped out for the analysis. A final data quality issue is that the data was last updated in the middle of 2017. For this reason I cut 2017 out of the data used to show growth trends but kept it in the plots of colors per year and per-set complexity.

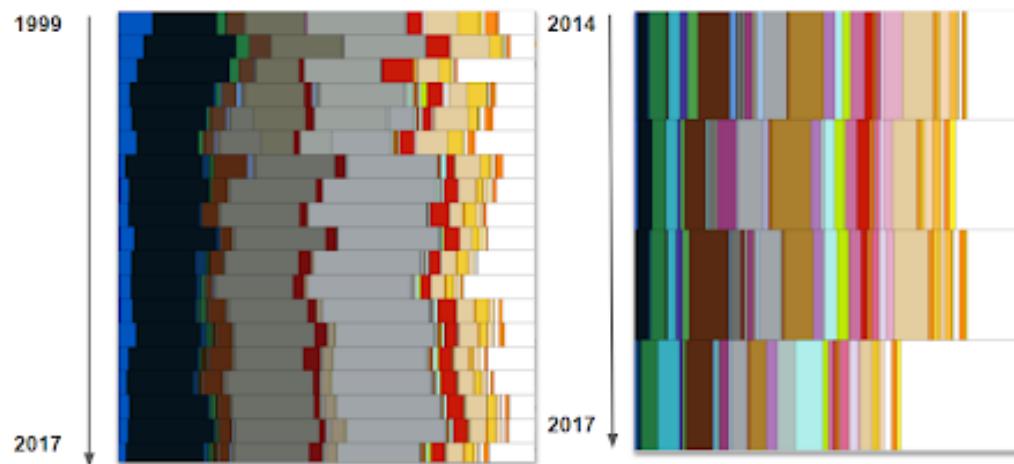
## Data Analysis

I wrote a quick function with the Graphics module to take any slice of the Legos data frame and plot out the proportion of each color over time, using the real rgb values in the data. I used this as a tool for exploratory analysis, and I sliced, diced, aggregated and plotted to validate what I saw.

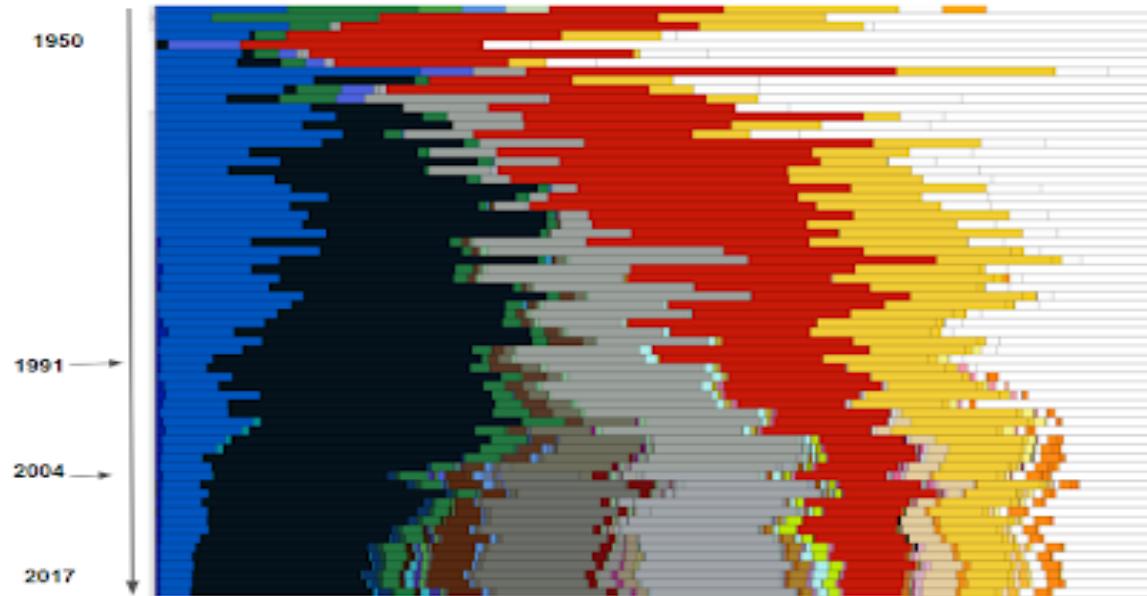
My unit of analysis was unique parts released as part of a set in a given year. I looked at these parts across themes, which are brands that defines multiple sets over multiple years. Besides looking at the Legos in various themes, key sub-populations I looked at were Legos with "Star Wars" in the theme name and Legos with "Pink" in the color name.

## Findings:

Lego themes usually have a consistent color profile, which creates a mood or evokes a memory. See if you can guess these themes:



If you guessed Star Wars and Disney Princess, you're right! These examples show how much color defines a brand. From this we can infer that looking at aggregated colors will show Lego's overall brand strategy over time.

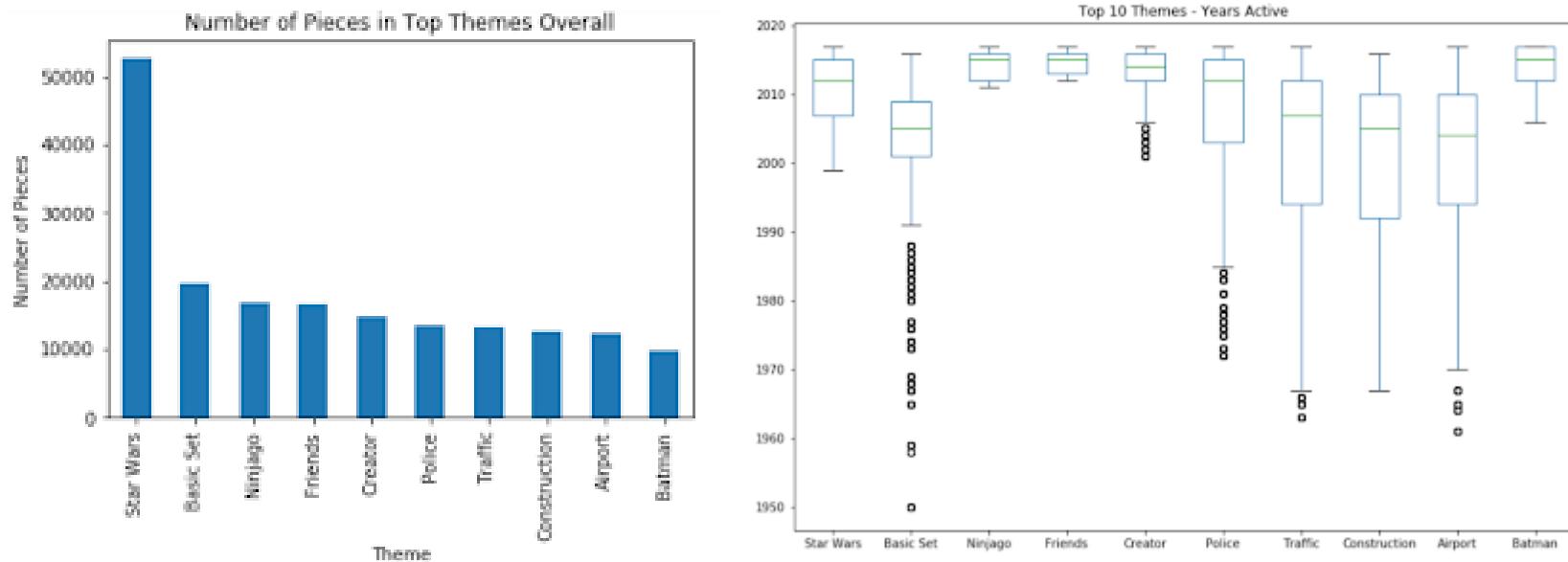


Looking at the colors Lego has used for every year's inventory since 1950 tells a story about their marketing strategy. Note that this plot is somewhat distorted because the number of parts released per year has grown significantly over time, so there are not actually more red pieces over all in 1950 than 2017, although the proportion of red pieces in 1950 is higher. Also note that there are some textures like glitter, chrome, transparent and milky that are not visible in this style of plot.

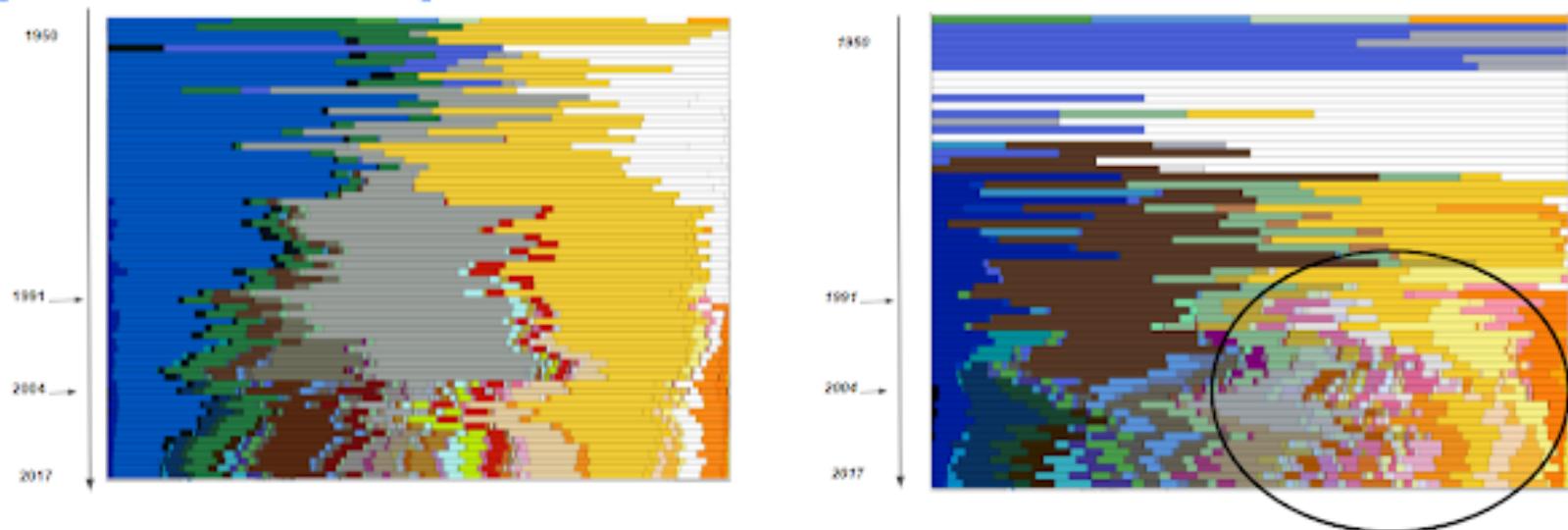
Lego's use of white has persisted but somewhat diminished over time. Black is the most common color overall, making up almost 20% of all Lego pieces ever manufactured, but for the first 14 years of Lego it was not in the top 3 colors.

Overall, you see that White and Red are constants throughout Lego's history, and are in the top 5 for every year. But the last year Red was in the top three was 1991 - since then, you see grey colors start to overtake the share once held by red, yellow, and blue.

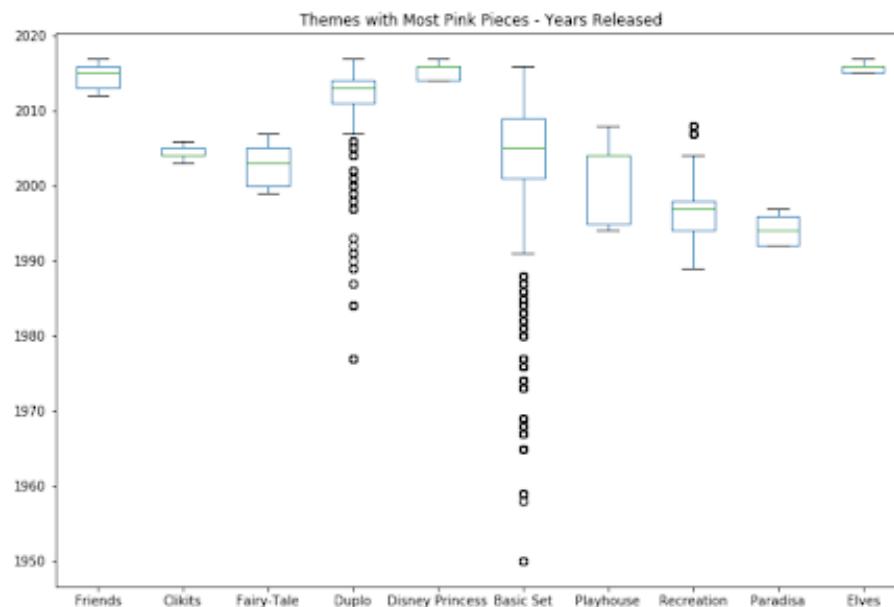
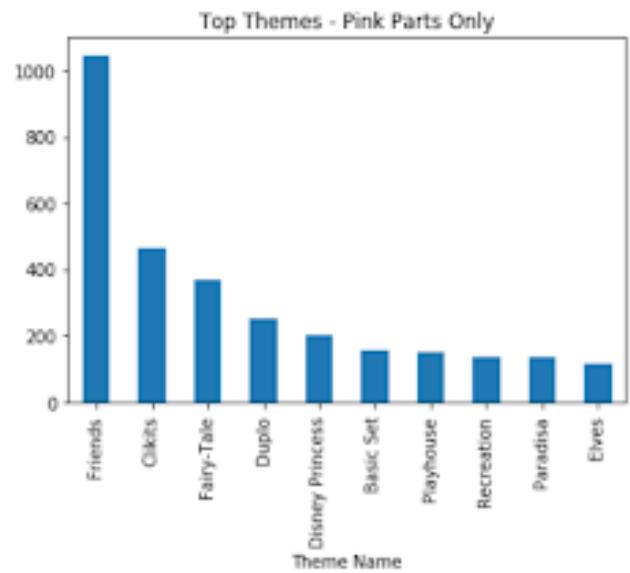
Light Bluish Grey was never a top 3 color until 2004, but it has been in the top 3 every single year since then. Dark Bluish Grey was similarly absent from the top 5 every year until 2004, but since then it was in the top 5 every single year.



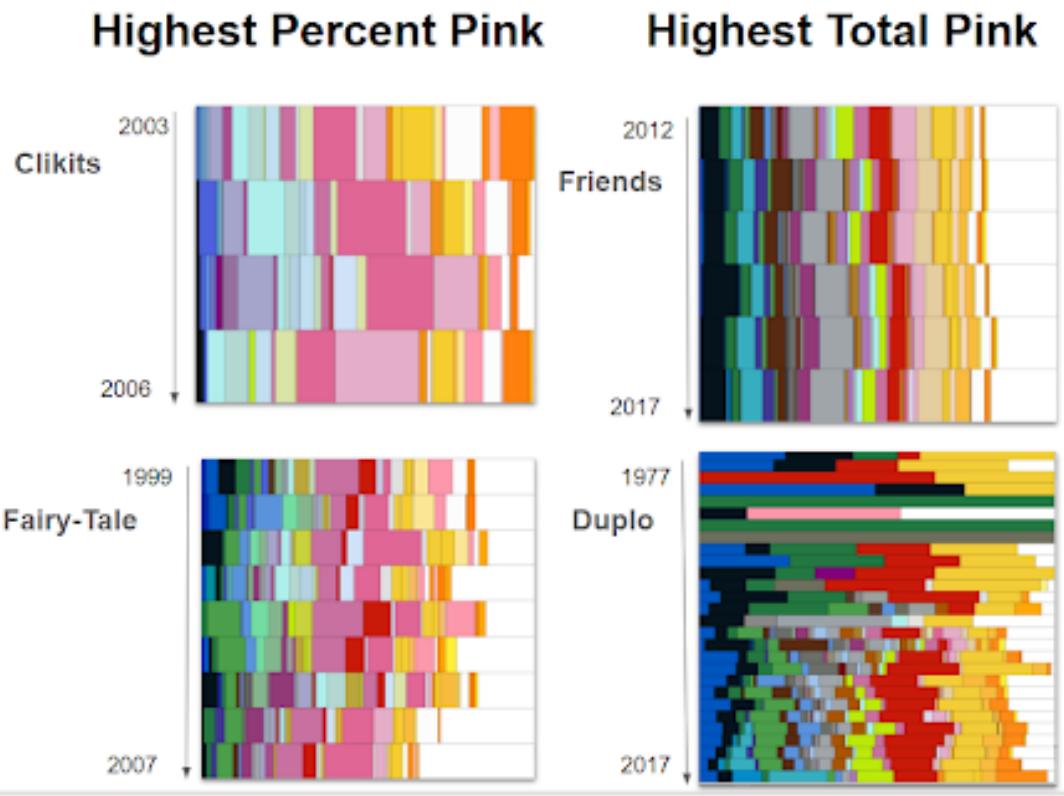
Light Bluish Grey and Dark Bluish Grey lead us into the story of Star Wars legos. Along with Black they are the most common colors in themes with "Star Wars" in the name, of which there are 12 distinct sub-themes since 1999. In all, Star Wars legos make up about 9% of all pieces ever produced by Lego, over 50,000 unique pieces across 18 years. Even before I grouped them together, two distinct Star Wars themes were in this top 10 chart. This is especially notable because compared to other top 10 themes like Basic Set, Traffic, Police, Construction and Airport, Star Wars has been in the market for only a fraction of Lego's life span.



These plots show all but the top 5 colors, in terms of number of pieces, on the left, and all but the top 10 colors on the right. You see that the primary colors blue and yellow have not diminished much over Lego's entire history - probably because of the continued importance of the early childhood market. The white you see is a different white called 'Milky White' and was big between 1963 and 1977 but has not been in the top 10 ever since. A popular shade of grey almost disappears after 2004 - probably because it was replaced by the blue grays of the Star Wars themes. The same thing happened to brown - Star Wars' richer brown after 2004 pushes out the greyer brown. You see this other brown magnified from left to right. What stands out here is that in 1991, you start to see the emergence of pink and purple - in fact, there is not a single Lego in any shade of pink before 1991. This indicates a shift in Lego's strategy from unisex marketing to market segmentation based on gender.



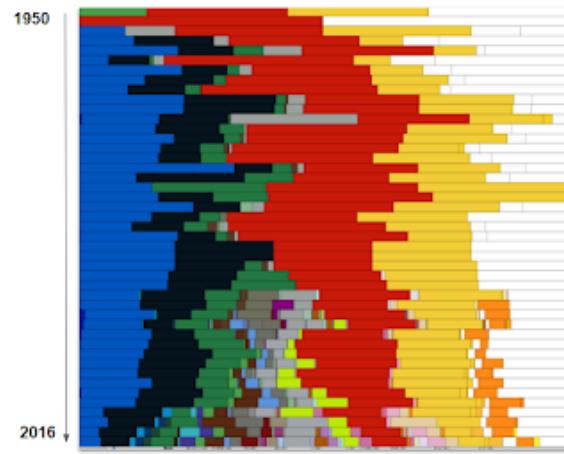
Subsetting to show only the 4,296 lego pieces in shades of pink reveals an almost entirely different set of top 10 themes - only Basic Set and Friends are also in the top 10 overall. Duplo and Basic Set both introduced Pink in the early 90s, but they have a much longer history at Lego. Friends has the most pink pieces overall, but it is not the highest percent pink.



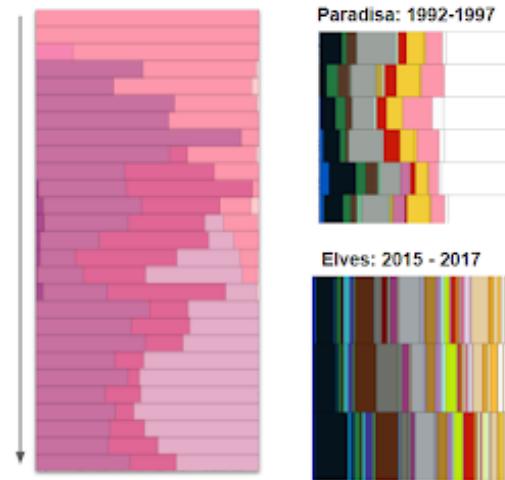
This shows the color profiles of the most pink sets, by both the highest percent of pieces in shades of pink, and highest volume of pieces in shades of pink.

Clikits only lasted 4 years but released over 400 pink parts. In all, Clikits pieces are 36% pink!

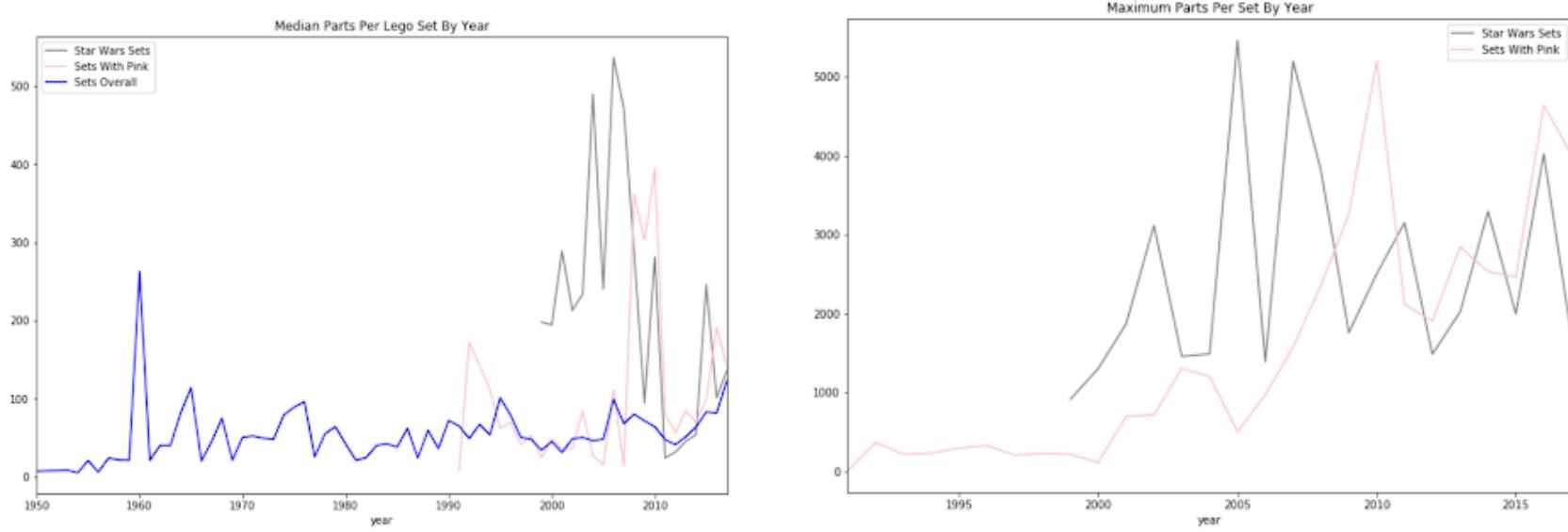
The volume of Duplos released grew considerably from only one unique piece in 1990 and 1992, which were cars, to 896 unique pieces in 2014. Duplos and Basic Set are two brands that we see change their strategy significantly after many years on the market to include more feminine colors, showing that themes are not static and Lego's marketers are willing to take risks with some of their more established brands to drive growth.



This plot shows how Lego preserves the fundamentals of the theme “Basic Set” even as it evolves.



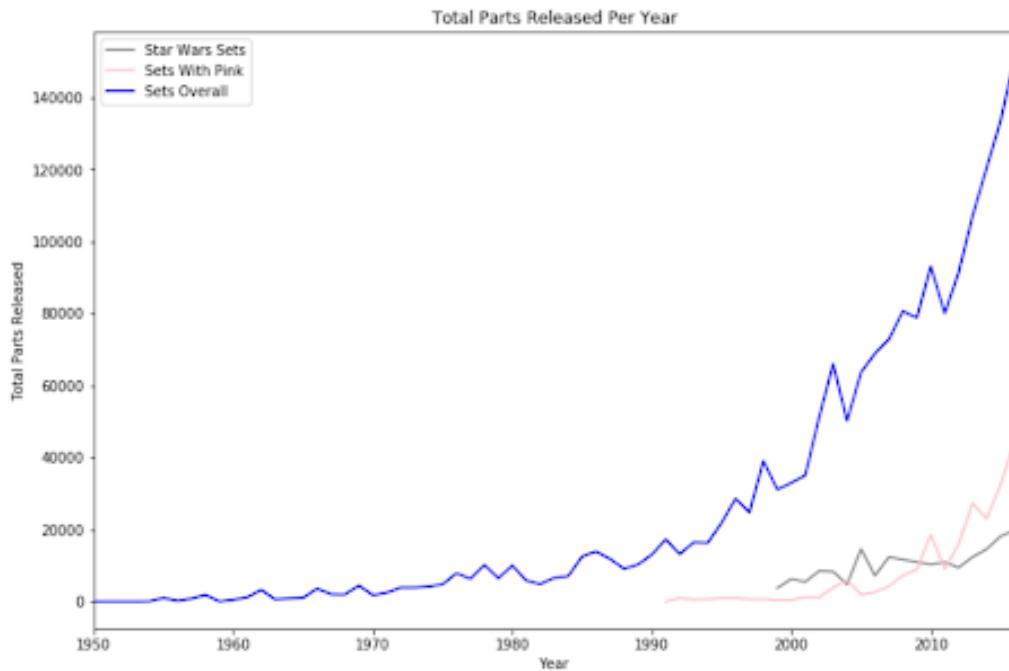
The color palette of pink Legos has grown more sophisticated over time - this shows the evolution from 1991 to 2017, along with 2 pink themes from the early 90s and the late 2000s. The tones are cooler and the rate of pink relative to the overall color palette is lower.



This growth in sophistication for pink sets shows up in the trends in parts per set as well. Star Wars is generally more complex, but in recent years sets with pink have gotten more complex

In the second plot, the most complex set with pink in 2010, more complex than any Star Wars set released that year, was a Toy Story Set. Toy Story was a one-year theme, and a second possible explanation for the low amount of pink released in 2010 is Toy Story's popularity with both boys and girls, leading Lego to reduce the major releases for their girl-centered brands.

I thought about whether it made sense to use all sets with any pink as my subset, but I think by now we understand that Lego does not use color by accident - every use of color is intentional and fits into Lego's overarching strategy. What we can see here from the plots of complexity is that sets with pink are comparable with the general population, and most recently challenge Star Wars in terms of sophistication. This shows that Lego has intentionally directed its strategy first towards capturing the market of young girls, and then towards providing more sophisticated toys for girls.



The number of pieces released per year has grown almost exponentially since 1950. We can see from the gray line that Lego used a lot of resources to scale their Star Wars brand fast. But, you notice that the pink line outpaces the grey one after 2010 - this represents the population of sets using pink. We know every time Lego uses a color, it is an intentional aspect of their brand, so this pink line shows us a strategic initiative to target the female demographic. If this trend continues, we can expect that content marketed to girls could be as significant to Lego as Star Wars.



1981



2005

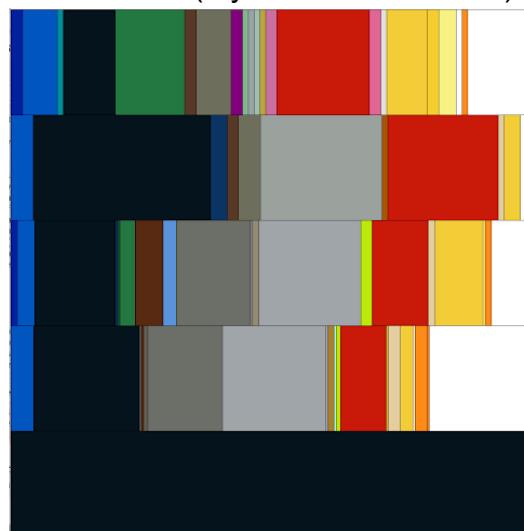


2017

Lego started out as a toy intended for organic creativity. Beginning in the early 1990s, their marketing strategy shifted to more branded and more gendered content, like the massive Lego Star Wars franchise and 2005's Clikits Pretty In Pink Beauty Set. These products sold, driving Lego's spectacular growth. In 2014, a letter written by a 7-year-old criticizing the lack of female career role models in Legos made headlines, and Lego reacted by again evolving their strategy. They launched an initiative to collect fan feedback and released the Women of NASA set last November under the new Lego Ideas theme. Within 24 hours it was the number 1 toy on Amazon.

## Mindstorms Visualization for Gunnar

2002-2016 (5 years with releases)



The last row was when a set with a single piece called Transformer 10V DC was released in 2016. Pink shades disappear from Mindstorms after 2002, which is the opposite of the overall trend in Legos

## Lego and Film

### Introduction:

When taking a quick glance at the lego set themes, we found that movie names came up very often. It got us thinking; Lego could be targeting already existing movie audiences instead of trying to capture completely new customers. This would make sense if they are trying to cut down on marketing costs, since they could leverage existing brand awareness and love instead of investing in marketing Lego sets that they create. While this would increase licensing fees, it could be part of their overall successful marketing strategy. We decided to dive deeper into the relationship between Lego and films.

### Methodology:

We first needed to find movie titles by release year. Boxofficemojo had top grossing films for each year, so scraping the top 100

grossing films for the years 1980 - 2017 into a clean and concise csv was fairly straightforward.

	Year	Rank	Title
0	1980	1	Star Wars: The Empire Strikes Back
1	1980	3	Stir Crazy
2	1980	5	Any Which Way You Can
3	1980	7	Coal Miner's Daughter
4	1980	9	The Blue Lagoon

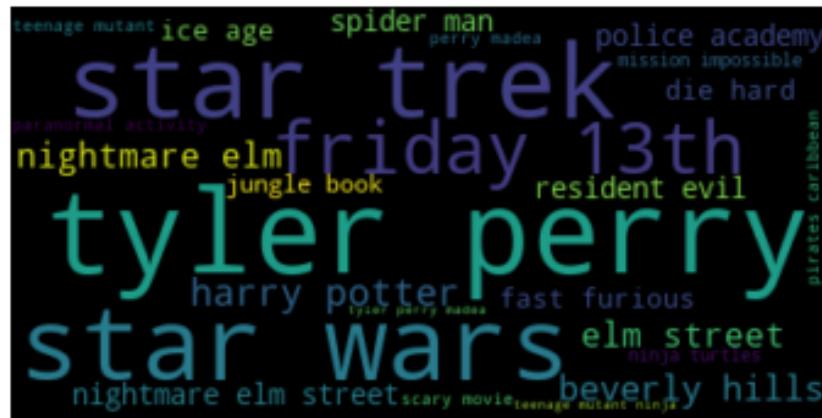
In terms of the Lego dataset, we had a master file, called lego\_out, that contained all of the files from kaggle joined into one csv. This was really helpful when looking at Lego movie information because it meant not having to join the set data with the theme data individually, it was all there in one file. While all of this information was useful, upon first glance, only a few columns were relevant to movies. Those were year, set\_name, theme\_name, and parent\_name. Year was important to see when these sets came out, but set\_name, theme\_name, and parent\_name all looked like they could have movie titles in them.

	year	set_name	theme_name	parent_name
0	2004.0	McDonald's Sports Set Number 6 - Orange Vest S...	Gravity Games	Sports
4	2012.0	Emma's Splash Pool	Friends	NaN
33	1999.0	Zo Weevil	Insectoids	Space
49	1990.0	T-Junction, Circle Plates	Town	Service Packs
51	1969.0	Train Wheels and Couplers	Supplemental	Classic

After further investigation, it was clear that set\_name did not have any movie titles, so theme\_name and parent\_name remained along with year. One issue was that either column theme\_name or parent\_name contained a movie title, so when searching for movie titles we would have to search through both columns. Our solution was to concatenate the contents of the columns into one column, aptly titled full\_name. We then grouped the data set to output only unique combinations of year and full\_name. Our data was then ready for exploration.

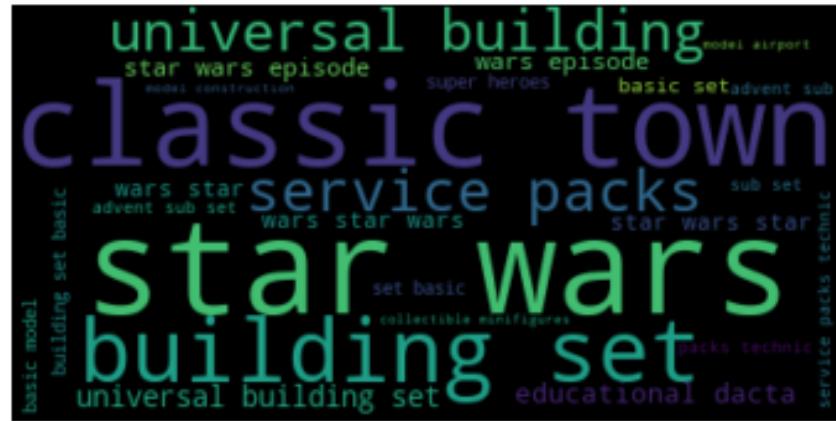
## **Titles Exploration:**

Our movie data had 3,800 rows of data, so there was a lot to explore. We decided it would be a good idea to focus on film franchises, since we know Star Wars was so popular in our Lego data set. One way to identify franchises was to measure the frequency of word pairs and triplets, because most films in a franchise contain the franchise name in their title. Using CountVectorizer and limiting to word pairs and word triplets (ngram\_range=(2, 3)), we identified the most common 25 phrases in film titles. We then visualized these titles by using a word cloud, so that we could easily see which titles occurred most frequently in the top 100 grossing movie lists in the past 37 years.



Immediately we saw some familiar names: Star Wars, Harry Potter, Spider Man. All of these film titles existed in our Lego data set. One interesting note; Tyler Perry was the most common word pair in our film data (15 films), meaning that there have been more Tyler Perry movies released than any other film franchise. And yet, Lego has not released any Tyler Perry themed sets. From this, we can assume that Lego Madea did not test well with target audiences.

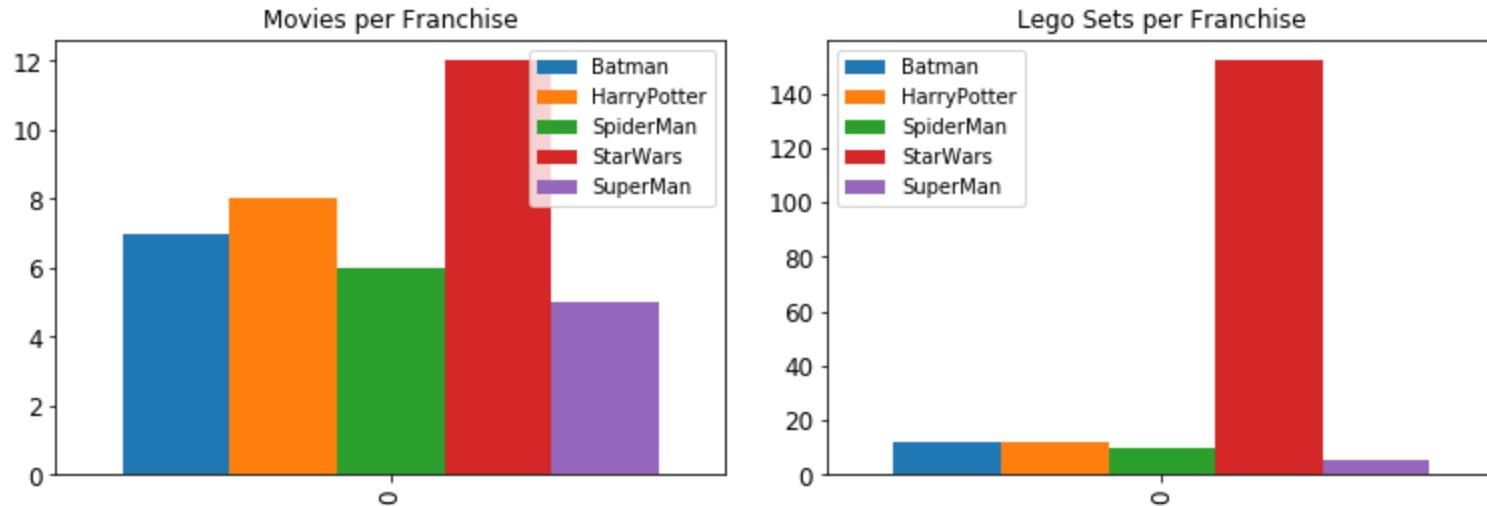
We decided to replicate this methodology using the lego data set, finding the top word pairs and triplets from our full\_name column.



At 226, the word pair with the highest frequency was Star Wars. There are multiple iterations of the “star wars” word pair occurring in the top 25 word pair and triplets frequencies, including “star wars episode”. What is interesting is that other than Star Wars, no other movie titles were in the most frequent 25 word pairs and triplets in the lego dataset. The next highest word pair that was even somewhat related to movies was “super heroes.” As previously mentioned, Star Wars has only existed as lego sets for the last 18 years, so the number of Star Wars sets in later years has been astounding.

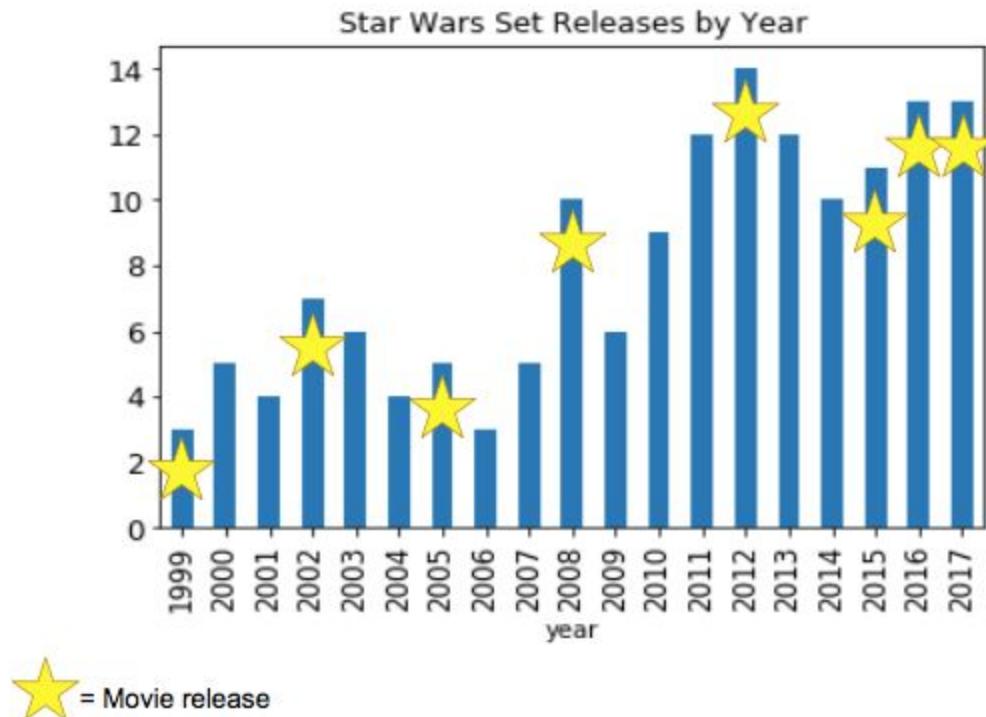
## **Super Heroes:**

We can see that Star Wars is clearly a popular set for Lego because they have produced more Lego themed sets than any other theme. However, Star Wars has released the most movies on any film franchise (except for Tyler Perry). Perhaps Lego was releasing an equal number of sets per film in each franchise. "Super Heroes" also appeared in our top 25 word pairs. We suspected that superheroes were just grouped together, and it was possible that when we looked at individual superhero movies, the number of sets released would reflect the number of movies in those franchises.



Our suspicions were not correct. Although the next largest film franchise, Harry Potter, also had the second highest number of Lego sets in this group, and the third largest film franchise, Batman, had the third highest number of Lego sets, Star Wars sets clearly over indexed. Directionally, this theory holds, but the magnitude of Star Wars sets proves that the number of films in a franchise do not determine how many Lego sets are created.

### **Star Wars Releases:**



In the early years of Star Wars sets, Lego did not have a fully thought-out plan for how many Star Wars themed sets to release and when. In fact, up until 2008 it looks like Lego experimented with the timing on their sets, releasing more sets the year after a Star Wars movie was released, like in 2000. By 2008, however, it is very clear that Lego had decided on a set-release strategy that coincided directly with movie launches. Star Wars movies were released in 2015, 2016, and 2017 and it looks as though Lego is ramping up their set production over these years. We anticipate a large amount of set releases in 2018 as there will be another Star Wars movie released.

## Conclusion

From 2003 onwards, Lego revenue and profits have steadily increased. Expenses have not kept up at the same rate, indicating Lego has had some major sales and efficiency wins. In an effort to understand how Lego has managed this impressive growth, we looked at different aspects of Lego's marketing strategy, namely how colors, branding, and film releases played into this growth.

We investigated 3 drivers of this growth: colors, themes, and film partnerships. We saw through colors and themes that Lego made an effort to target girls in new and inventive ways. First, Lego attempted to entice girls with very pink sets, such as clickits, while still targeting boys with Star Wars and more traditionally boy-themed sets. Lego is now transitioning to more united sets that appeal to both boys and girls, realizing that girls do not need purely bubble-gum pink sets to play with Legos. Lego also released sets related to film franchises, relying on the marketing of the films to attract customers. It is much cheaper to attract a customer who already loves a certain film than to bring in a customer who has no experience or knowledge of a new set Lego creates. Lastly, Lego began creating Star Wars sets in 1999. Since then the majority of Lego sets created over the years have been Star Wars. This partnership with Star Wars has clearly changed Lego's marketing strategy from one of releasing Lego-specific sets, like Classic Town, to branded sets. All of these changes to Lego's marketing strategy have grown Revenue and Profits in an impressive way.