

Lego Analysis

Analysis of Lego

Methodology:

In order to investigate LEGO set trends and theme popularity, we used the LEGO data on TidyTuesday provided by Rebrickable.

This dataset contained all the information about officially released LEGO pieces where we focus on theme names, number of parts, colors and release years. The dataset was filtered to exclude sets with missing or zero part counts, as these entries would not contribute meaningfully to part-based analyses.

We first grouped sets by theme to calculate the total number of sets and the median number of parts per theme Figure 1. The top 20 most and least popular themes were identified and visualized using a horizontal bar chart Figure 2 & Figure 3. Each bar's length represents the number of sets per theme and fill color indicates the median number of parts. This allowed us to compare popularity with complexity as measured by part count.

In addition to thematic analysis, we tracked the evolution of LEGO set complexity over time measured by median part count and color diversity. These trends are visualized in Figure 4 & Figure 5, where x-axis shows the year and y-axis showing the corresponding median part or unique colour count. A smooth trend line over the years highlights a general increase in part and colour count reflecting growing set intricacy over time.

Finally, to explore the largest sets ever produced by LEGO, a bar chart Figure 6 was created, where each bar represents a set and is segmented by the proportion of colours used. The bars are filled with the actual hex colours replacing fill in each set, allowing for a really clear visualisation of both the total number of parts and the colour distribution, and provides insights into which colours are most prevalent in large complex builds and highlight some of the design choices behind some of LEGO's most detailed sets.

Theme popularity - based on the number of sets released under each theme. More sets = more popular

Results:

1. Most Popular and least Popular LEGO Themes (Figure 2 & Figure 3)

The most successful LEGO themes share common features: strong brand identity, wide audience appeal, and ongoing product development. The least popular themes were often short-lived, targeted at small audiences, or limited by licensing. Understanding these patterns helps LEGO and similar brands plan better product strategies and maximize market impact.

2. Complexity of LEGO Sets Over the Years (Figure 4)

LEGO sets became more complex up to the late 1970s. Since then, the median complexity has plateaued, likely due to diversification of set sizes to meet the needs of different age groups and skill levels.

3. Diversity of Colors in LEGO Over the Years (Figure 5)

The growing number of unique colors used in LEGO sets is a strong indicator of increased design sophistication and build complexity. Color diversity not only enhances the visual experience, but also plays a critical role in enabling LEGO's modern design capabilities and supporting more engaging, realistic, and intricate builds.

4. LEGO Sets with the Most Parts and Color Proportions (Figure 6)

The number of parts in a LEGO set often correlates with either extensive design detail (e.g., maps, buildings) or technical function (e.g., vehicles, machinery). Color proportions reveal the thematic purpose of each set—some aim for vibrant storytelling, while others reflect precision and realism. Together, these sets showcase LEGO's versatility across both form and function.

Table 1: Theme ordered by sets with their median parts

theme_name	n_sets	median_parts
Technic	810	64
Star Wars	787	148
Friends	494	87
Ninjago	444	58
Basic Set	439	52
Supplemental	431	12
Gear	409	15
Classic Town	394	66
Town	374	20
Bionicle	357	43
Service Packs	347	16
Books	316	8
Duplo and Explore	239	80
Racers	226	54
City	207	169
Creator	201	59
Creator 3-in-1	193	256
LEGO Brand Store	156	31
Monthly Mini Model Build	148	41
Police	148	98
Basic	138	194
Duplo	138	15
Christmas	135	50
Brickheadz	130	160
Harry Potter	129	263
Bulk Bricks	124	50
Promotional	117	68
Nexo Knights	113	47
Fabuland	105	18
Soccer	100	16

Figure 1: Theme by parts

All time most popular lego themes with median parts

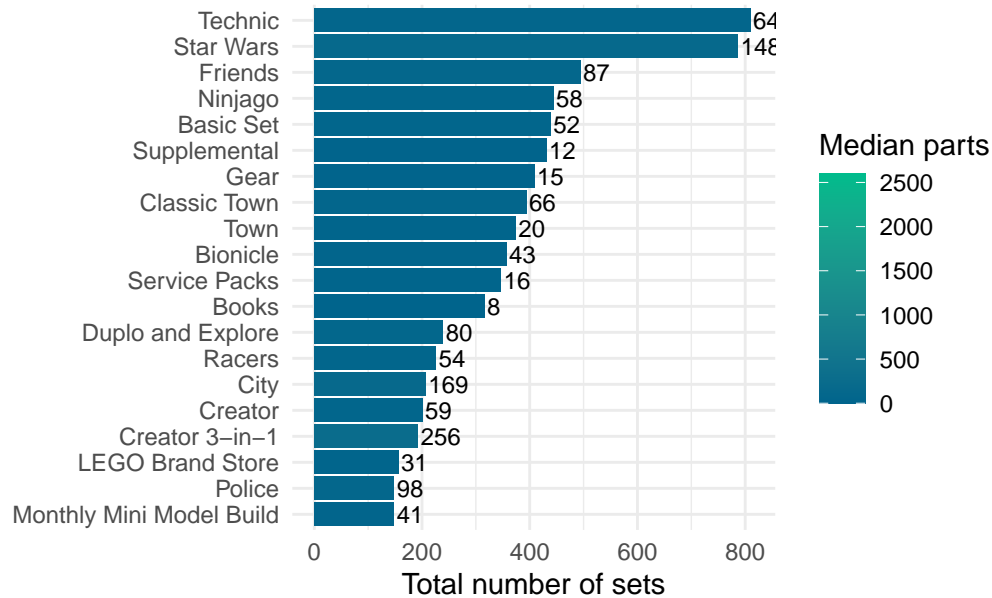


Figure 2: Most popular themes

All time least popular lego themes with median parts

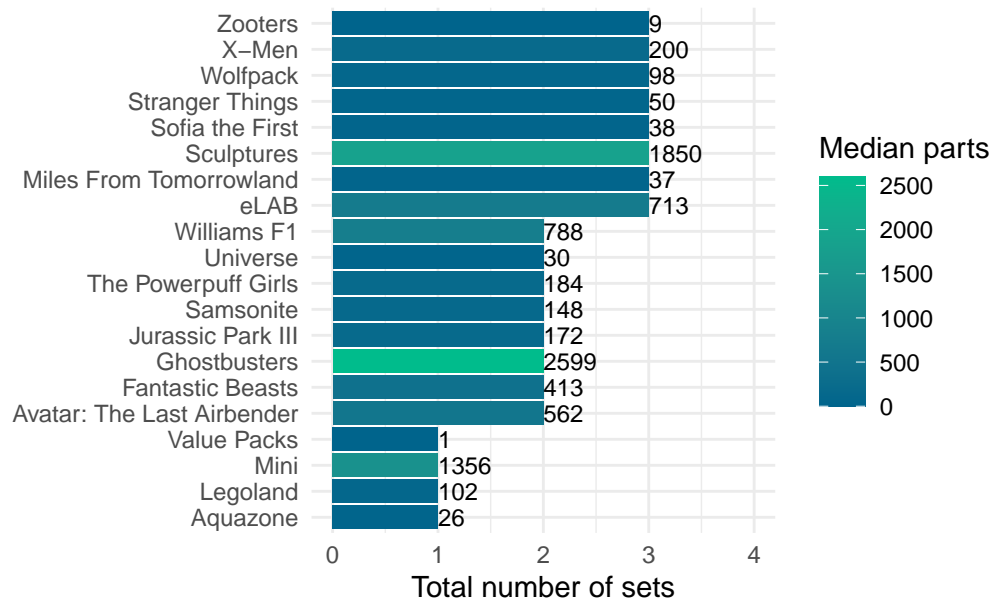


Figure 3: Least popular themes

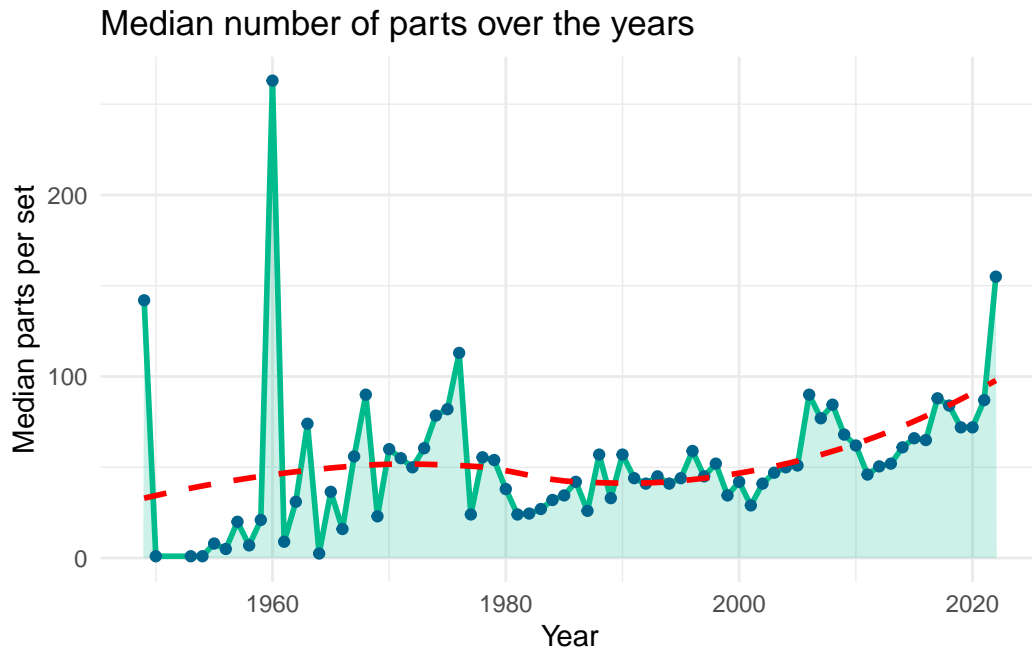


Figure 4: Median parts by year

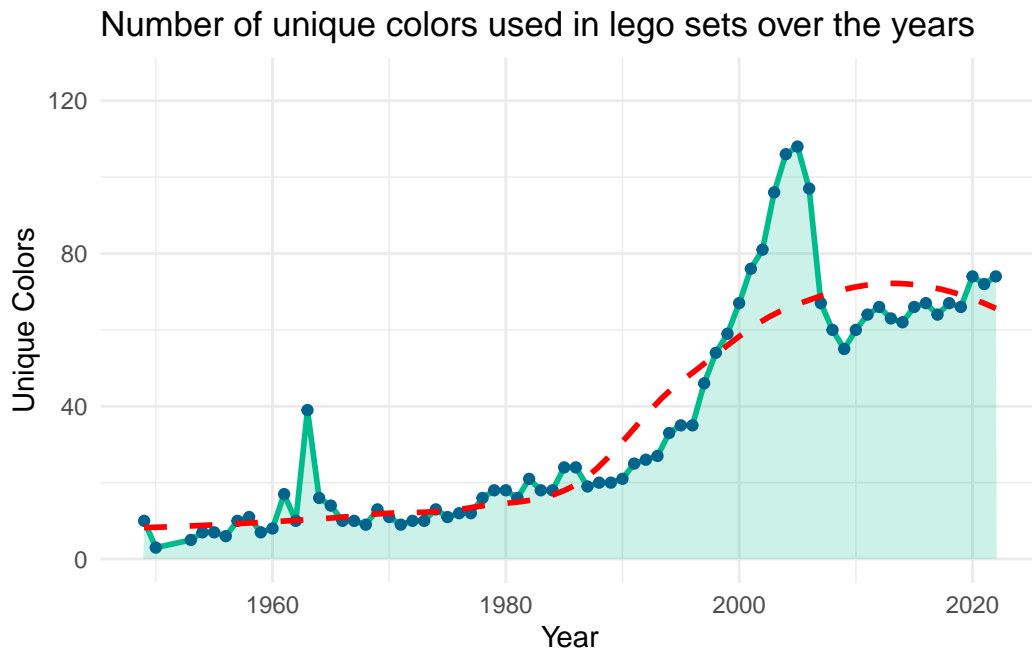


Figure 5: Unique colour count by year

Top 20 lego sets by # parts and colour proportion

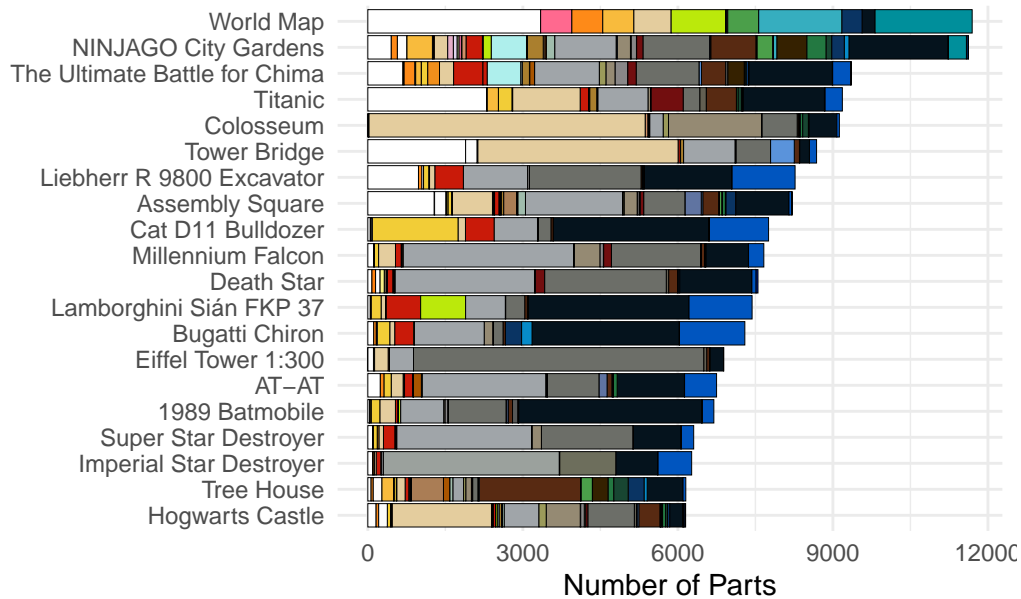


Figure 6: Top 20 sets by part count with color proportion

Discussion

Based on Figure 2 and Figure 3, it can be seen that the more popular themes have on average significantly fewer parts in comparison to the least popular themes. This indicates that accessibility, simplicity, and recognizability may play a more important role in determining a theme's success, rather than set complexity alone. Simpler sets can appeal to more customers, especially younger children and first time taster, making them more commercially successful.

However, the correlation analysis between the number of sets in each theme and the median number of parts per set produced a Pearson correlation coefficient of -0.087 . This indicates a very weak negative correlation, meaning there is little to no linear relationship between a theme's popularity (number of sets) and the complexity of its sets (median parts count). In other words, while popular themes often feature smaller sets, the number of sets in a theme does not strongly predict the typical complexity of those sets. In terms of inventory management, the production of simple sets within existing themes can help to boost sales.

Overall, the finding is aimed to emphasize that simplicity and accessibility are key factors of Lego theme popularity.