Part A - Lego Analysis

Analysis of Lego

```
# Loading libraries

library(tidyverse)
library(tidytuesdayR)
library(ggplot2)
library(scales)
```

Reading in dataset

```
# Read data from tidytuesday
tuesdata = tidytuesdayR::tt_load('2022-09-06')

for (name in names(tuesdata))
{var_name = gsub("\\.csv$", "", name)
    assign(var_name, tuesdata[[name]])}

rm(tuesdata)
```

Explore data

```
head(colors)
```

```
    4 2 Green 237841 FALSE
    5 3 Dark Turquoise 008F9B FALSE
    6 4 Red C91A09 FALSE
```

head(elements)

```
# A tibble: 6 x 3
  element_id part_num color_id
      <dbl> <chr>
                       <dbl>
    6300211 67906c01
                          14
1
2
    4566309 2564
                           0
                        1004
3 4275423 53657
4
    6194308 92926
                          71
5
    6229123 26561
                           4
    4241969 51035
                        1004
```

head(inventories)

```
# A tibble: 6 x 3
     id version set_num
  <dbl>
         <dbl> <chr>
1
    1
             1 7922-1
2
     3
             1 3931-1
3
     4
             1 6942-1
4 15
             1 5158-1
5
    16
             1 903-1
    17
             1 850950-1
```

head(inventory_minifigs)

```
# A tibble: 6 x 3
  inventory_id fig_num
                           quantity
         <dbl> <chr>
                              <dbl>
1
             3 fig-001549
                                  1
2
             4 fig-000764
                                  1
3
            19 fig-000555
                                  1
4
            25 fig-000574
                                  1
5
            26 fig-000842
                                  1
6
            26 fig-008641
                                  1
```

head(inventory_parts)

```
# A tibble: 6 x 6
  inventory_id part_num
                              color_id quantity is_spare img_url
         <dbl> <chr>
                                 <dbl>
                                           <dbl> <lgl>
                                                          <chr>
             1 48379c01
                                    72
1
                                               1 FALSE
                                                          https://cdn.rebrickabl~
2
             1 48395
                                     7
                                               1 FALSE
                                                          https://cdn.rebrickabl~
3
             1 stickerupn0077
                                  9999
                                               1 FALSE
                                                          <NA>
             1 upn0342
                                                          <NA>
4
                                     0
                                               1 FALSE
5
             1 upn0350
                                     25
                                               1 FALSE
                                                          <NA>
             3 2343
                                     47
                                               1 FALSE
                                                          https://cdn.rebrickabl~
```

head(inventory_sets)

```
# A tibble: 6 x 3
  inventory_id set_num quantity
        <dbl> <chr>
                          <dbl>
            35 75911-1
1
2
            35 75912-1
                               1
3
            39 75048-1
                               1
4
            39 75053-1
                               1
5
            50 4515-1
                               1
            50 4520-1
```

head(minifigs)

```
# A tibble: 6 x 4
 fig_num
          name
                                                      num_parts img_url
             <chr>
                                                           <dbl> <chr>
  <chr>
1 fig-000001 Toy Store Employee
                                                              4 https://cdn.reb~
2 fig-000002 Customer Kid
                                                               4 https://cdn.reb~
3 fig-000003 Assassin Droid, White
                                                               8 https://cdn.reb~
4 fig-000004 Man, White Torso, Black Legs, Brown Hair
                                                              4 https://cdn.reb~
5 fig-000005 Captain America with Short Legs
                                                               3 https://cdn.reb~
6 fig-000006 Lloyd Avatar
                                                               5 https://cdn.reb~
```

head(part_categories)

head(part_relationships)

```
# A tibble: 6 x 3
 rel_type child_part_num parent_part_num
  <chr>
           <chr>
                           <chr>
1 P
           3626cpr3662
                           3626c
           87079pr9974
2 P
                           87079
3 P
           3960pr9971
                           3960
4 R
           98653pr0003
                           98086pr0003
           98653pr0003
5 R
                           98088pat0003
6 R
           98653pr0003
                           98089pat0003
```

head(parts)

```
# A tibble: 6 x 4
 part_num name
                                                      part_cat_id part_material
 <chr>
           <chr>
                                                            <dbl> <chr>
1 003381
          Sticker Sheet for Set 663-1
                                                               58 Plastic
          Sticker Sheet for Sets 618-1, 628-2
2 003383
                                                               58 Plastic
3 003402
          Sticker Sheet for Sets 310-3, 311-1, 312-3
                                                               58 Plastic
          Sticker Sheet for Set 1550-1
4 003429
                                                               58 Plastic
5 003432
          Sticker Sheet for Sets 357-1, 355-1, 940-1
                                                               58 Plastic
6 003434
          Sticker Sheet for Set 575-2, 653-1, 460-1
                                                               58 Plastic
```

head(sets)

```
2 0011-2 Town Mini-Figures
                                      1979
                                                 67
                                                           12 https://cdn.rebri~
3 0011-3 Castle 2 for 1 Bonus Offer
                                                            0 https://cdn.rebri~
                                      1987
                                                199
4 0012-1 Space Mini-Figures
                                      1979
                                                143
                                                           12 https://cdn.rebri~
5 0013-1 Space Mini-Figures
                                                           12 https://cdn.rebri~
                                      1979
                                                143
6 0014-1 Space Mini-Figures
                                                           12 https://cdn.rebri~
                                      1979
                                                143
```

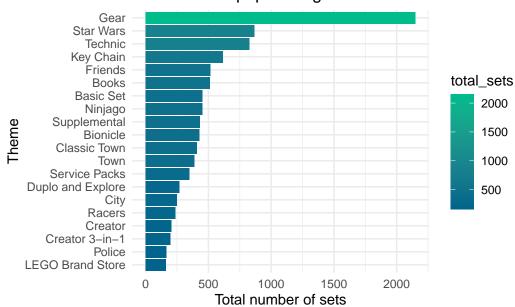
head(themes)

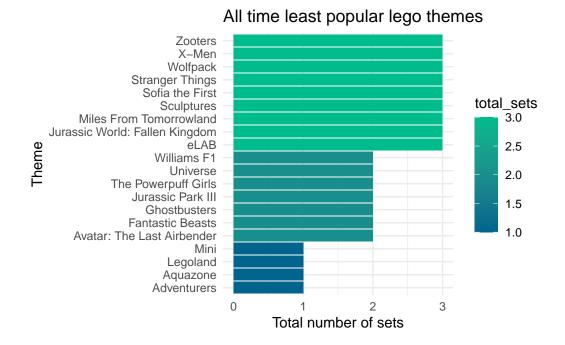
```
# A tibble: 6 x 3
     id name
                        parent_id
  <dbl> <chr>
                            <dbl>
1
     1 Technic
                               NA
2
      3 Competition
                                1
3
      4 Expert Builder
                                1
4
     16 RoboRiders
                                1
     17 Speed Slammers
                                1
5
     18 Star Wars
                                1
```

Considering the popularity of Lego, which is the most popular and the least popular themes of Lego.

```
# Joining sets with themes and filtering out the sets with no parts due to error in data col.
set_themes = sets |>
  left_join(themes |>
              select(id, theme_name = name),
              by = c(theme_id = "id")) |>
  mutate(numparts = na_if(num_parts, 0))
# Finding out the most popular themes
most_popular = set_themes |>
  group_by(theme_name) |>
  summarize(total_sets = n(), .groups = "drop") |>
  arrange(desc(total_sets)) |>
  head(20) |>
  ggplot(aes(x = total_sets, y = reorder(theme_name, total_sets))) +
  geom_bar(stat = "identity", aes(fill = total_sets)) +
  scale_fill_gradient(low = "#00648C", high = "#00BC8C")+
  labs(title = "All time most popular lego themes",
       x = "Total number of sets",
       y = "Theme") +
  theme_minimal()
```

All time most popular lego themes





Most Popular LEGO Themes

- 1. Gear is the most popular LEGO theme of all time, with over 2,000 sets released. This category includes accessories like keychains, stationery, and other merchandise, showing strong demand for LEGO-branded items beyond traditional building sets.
- 2. Star Wars comes in second with over 1,000 sets. Thanks to its powerful movie franchise, this theme appeals to both young fans and adult collectors and has remained a top performer for decades.
- 3. Technic also has over 1,000 sets and is popular among those who enjoy engineering-style builds. It's known for its complex mechanical features and realistic designs, attracting older children and adult hobbyists.
- 4. Other top themes include Key Chain, Friends, Books, and Basic Set, each with several hundred sets. This shows that LEGO successfully targets a variety of age groups and interests.

Least Popular LEGO Themes

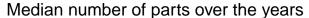
- 1. Many themes, such as Zooters, X-Men, Wolfpack, and Stranger Things, have only 1 to 3 sets. These were likely experimental or niche product lines that did not gain enough traction to expand.
- 2. Themes like Sculptures, eLAB, and Miles From Tomorrowland may have targeted specialized audiences, but limited appeal resulted in minimal set releases.
- 3. Some licensed themes, such as The Powerpuff Girls, Fantastic Beasts, and Avatar: The Last Airbender, may have faced market or licensing challenges, preventing long-term success.

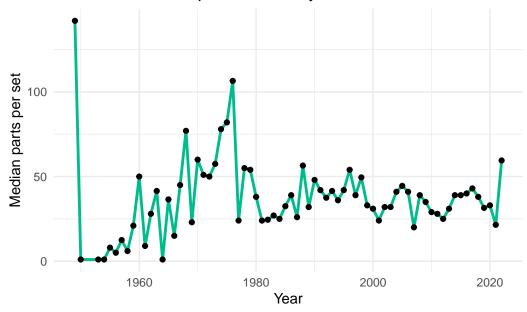
Conclusion

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.

Seeing the trends of Lego sets over the years, how hard or how complex have the Lego sets become.

```
theme_name_summary = set_themes |>
  group by(theme name) |>
  summarize(n_sets = n(),
            median_parts = median(num_parts)) |>
  arrange(desc(median_parts)) |>
  filter(n_sets >= 100)
part_year = set_themes |>
  group_by(year) |>
  summarise(median parts = median(num parts, na.rm = TRUE), .groups = "drop") |>
  ggplot(aes(x = year, y = median_parts)) +
  geom line(color = "#00BC8C", linewidth = 1) +
  geom point(color = "black") +
  labs(title = "Median number of parts over the years",
       x = "Year",
       y = "Median parts per set")+
  theme minimal()
part_year
```





Analysis: Complexity of LEGO Sets Over the Years

1. Early Years (1950s – early 1960s):

The chart shows a large spike in the 1950s, likely due to an outlier set with many parts. Aside from this, most sets had a low median part count, suggesting simple designs.

2. Growth Phase (1960s – late 1970s):

There was a clear upward trend in the median number of parts per set. This indicates that LEGO sets became more complex over time, possibly due to increased design innovation and consumer interest in more detailed builds.

3. Stabilization Period (1980s – 2020s):

After peaking in the late 1970s, the median number of parts stabilized around 40–60 parts per set. Despite some fluctuations, no significant long-term increase is observed.

4. Recent Years (2000s – 2020s):

Complexity remained steady, with occasional dips and rises. This suggests LEGO has balanced its product line with both simple and advanced sets.

Conclusion

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.