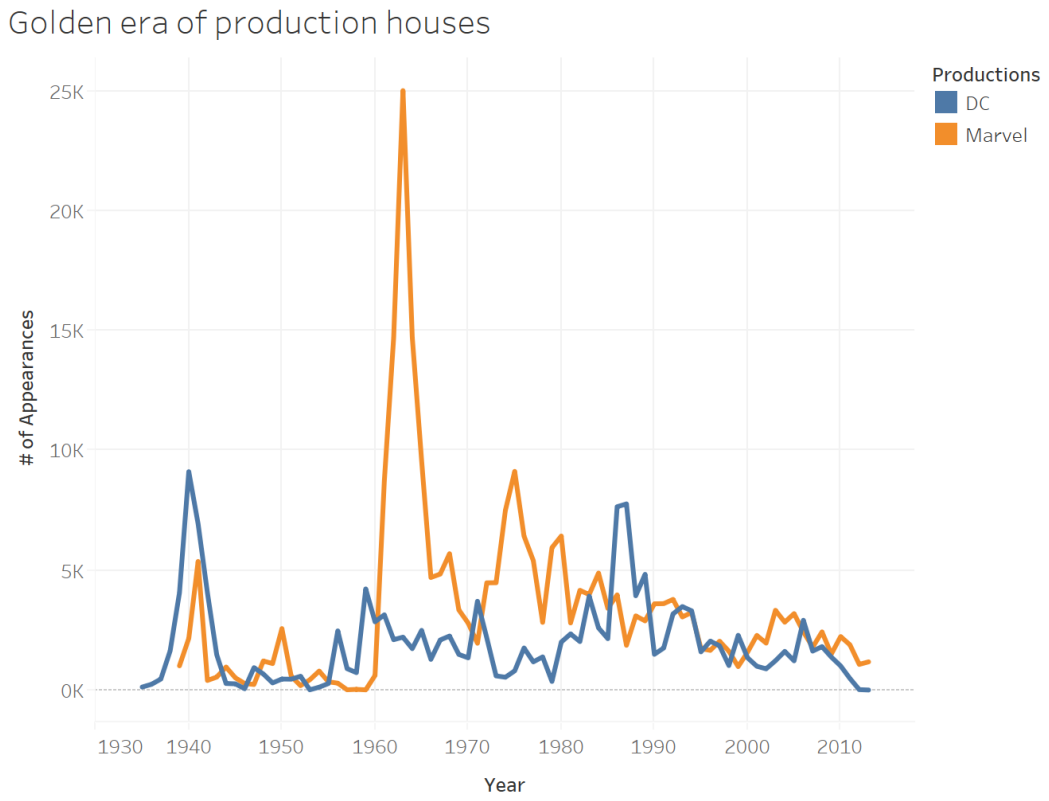
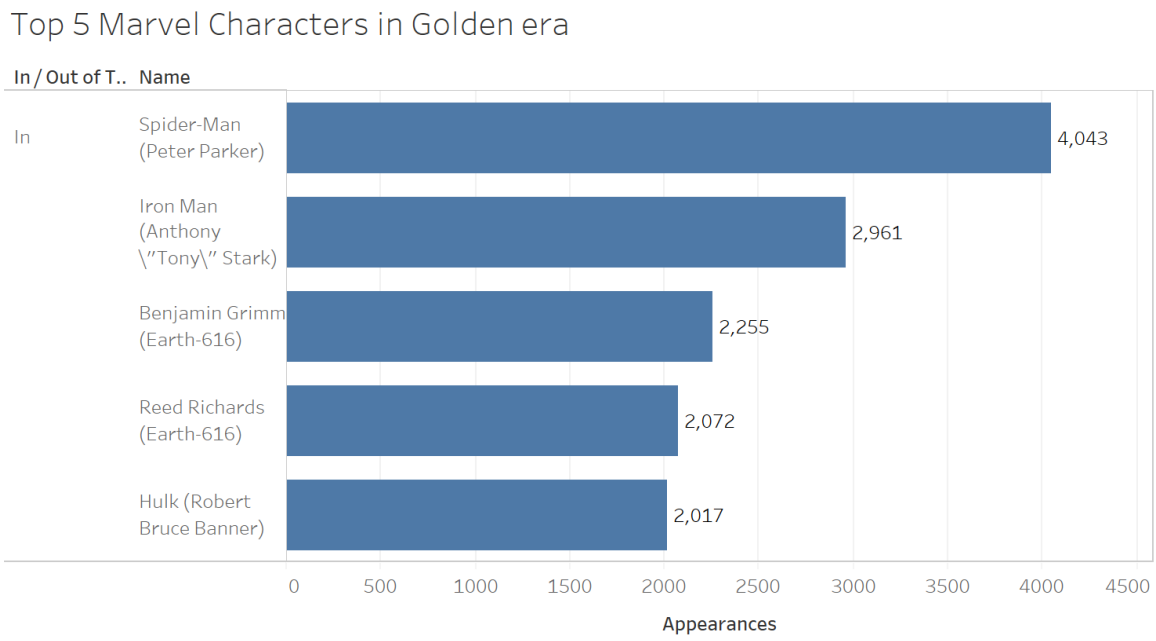
I have selected the superheroes dataset that has their Wikipedia search appearances. It also had the character descriptions like hair color, eyes color, identity, year appeared, sex etc.,
Data Source: [https://github.com/fivethirtyeight/data/tree/master/comic-characters](https://github.com/fivethirtyeight/data/tree/master/comic-characters)

My first intention was to compare the DC and Marvel (Production companies). The trend of both the production companies by the year in which the characters were introduced. It will give me the golden era of the production houses.



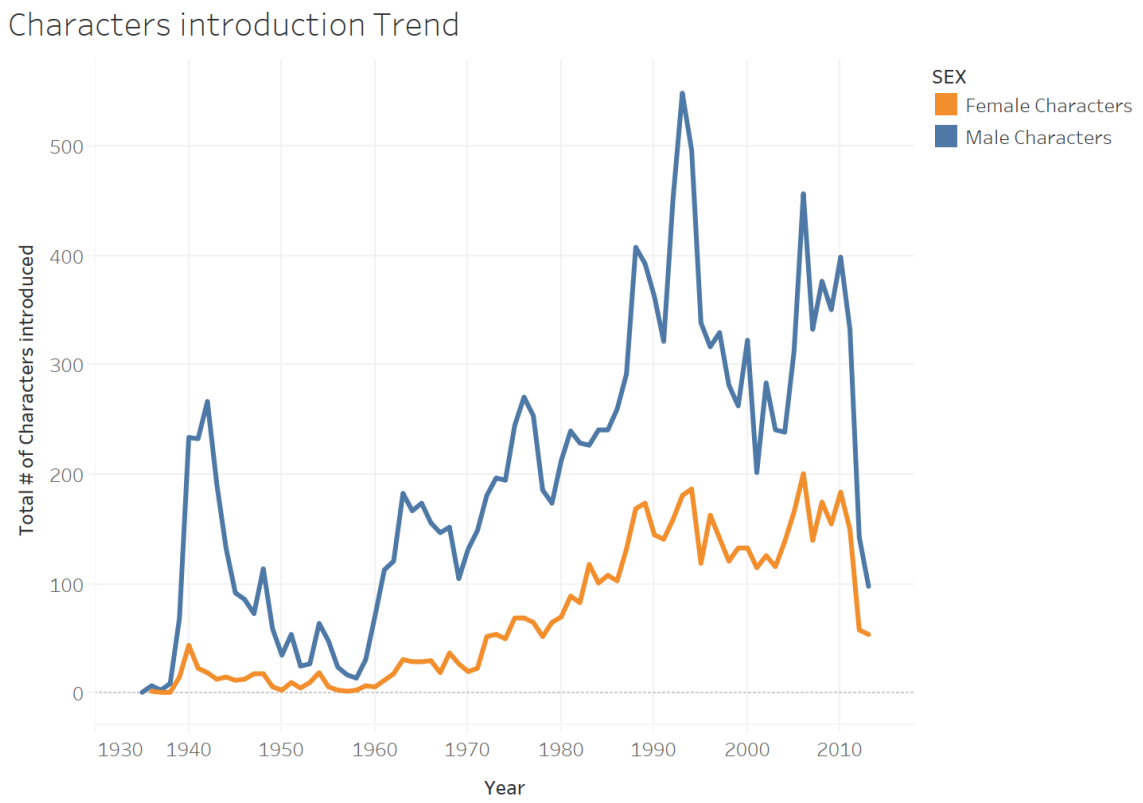

The trend of sum of Appearances for Year. Color shows details about Productions.

The Marvel production had highest peak between1960 to 1970. Hence I drilled down to find which superheroes were popular at that period. I found the top 5 superheroes and noticed that all the 5 were male superheroes.




Sum of Appearances for each Name broken down by In / Out of Top N superheros-Marvel. The marks are labeled by sum of Appearances. The data is filtered on Year and Productions. The Year filter ranges from 1960 to 1970. The Productions filter keeps Marvel. The view is filtered on In / Out of Top N superheros-Marvel, which has multiple members selected.

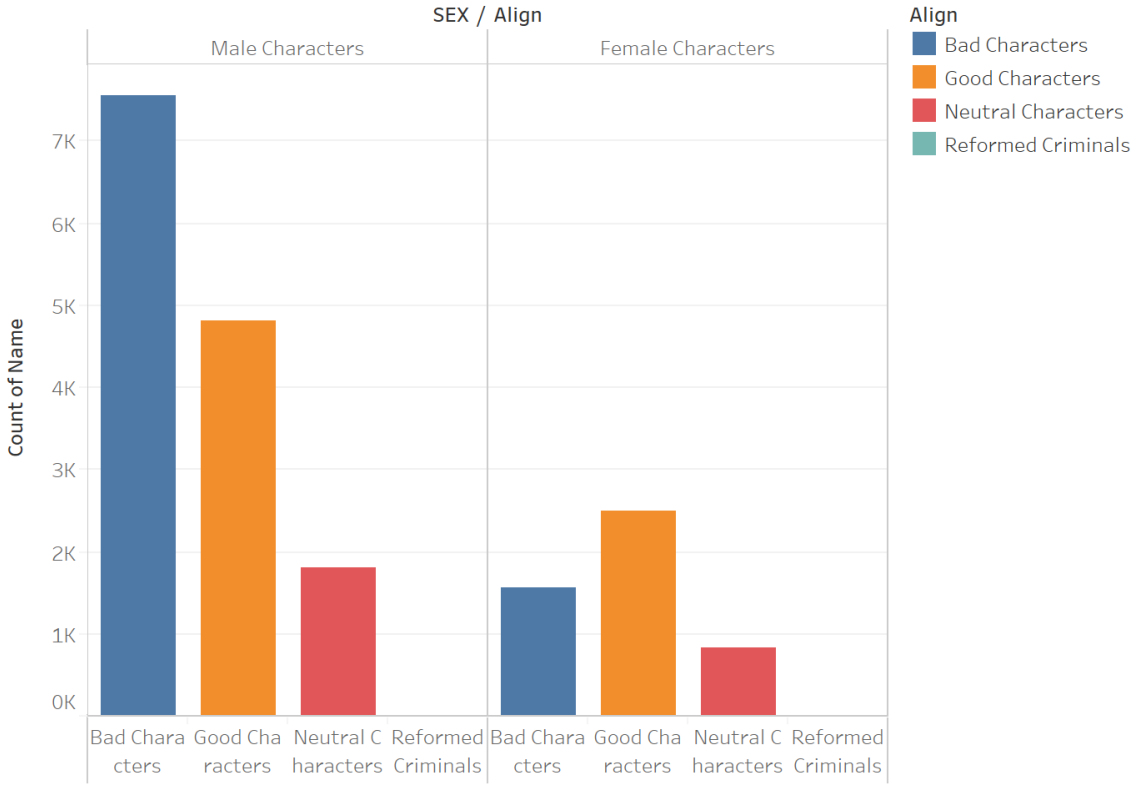
This probed me a question of whether female characters were popular or not. Hence I created a visualization to compare the male vs female characters. I excluded the Agender, transgender, gender less characters and compared only the male and female gender.




The trend of count of Name for Year. Color shows details about SEX. The view is filtered on SEX, which keeps Female Characters and Male Characters.

The graph showed me that both the female and male characters were popular at same period. Whenever there is a rise in the popularity of the male characters, the popularity of female characters also increased. But the popularity of female was less compared to male characters in all the period. My assumption was female characters was introduced at a certain period of time. But the visualization showed that the female characters were from the beginning. The graph was not rising continuously and it was fluctuating. I wanted to find the reason behind it. So I wanted to know about the characterization of both genders (whether good or bad).

## Character alignment by Gender




Count of Name for each Align broken down by SEX. Color shows details about Align. The view is filtered on SEX and Align. The SEX filter keeps Female Characters and Male Characters. The Align filter keeps Bad Characters, Good Characters, Neutral Characters and Reformed Criminals.

Most of the bad characters were male. The production companies introduced most of the female characters as good. Based on this, I wanted to know how the companies designed the characters i.e., their identity, hair color and eyes color. The visualization helped me to conclude that in the comic world the bad characters are identified secretly and good characters are characterized as being a public figure.

## Character Alignment by Identity

| | ID | | | | |
|---|---|---|---|---|---|
| Align | Secret Identity | Public Identity | Null | No Dual Identity | Known to Authorit.. |
| Bad Characters | 4,337 | 2,097 | 2,243 | 447 | 3 |
| Good Characters | 2,416 | 2,901 | 1,335 | 637 | 10 |
| Neutral Characters | 902 | 943 | 413 | 374 | 1 |
| Reformed Criminals | 1 | 1 | 1 | | |

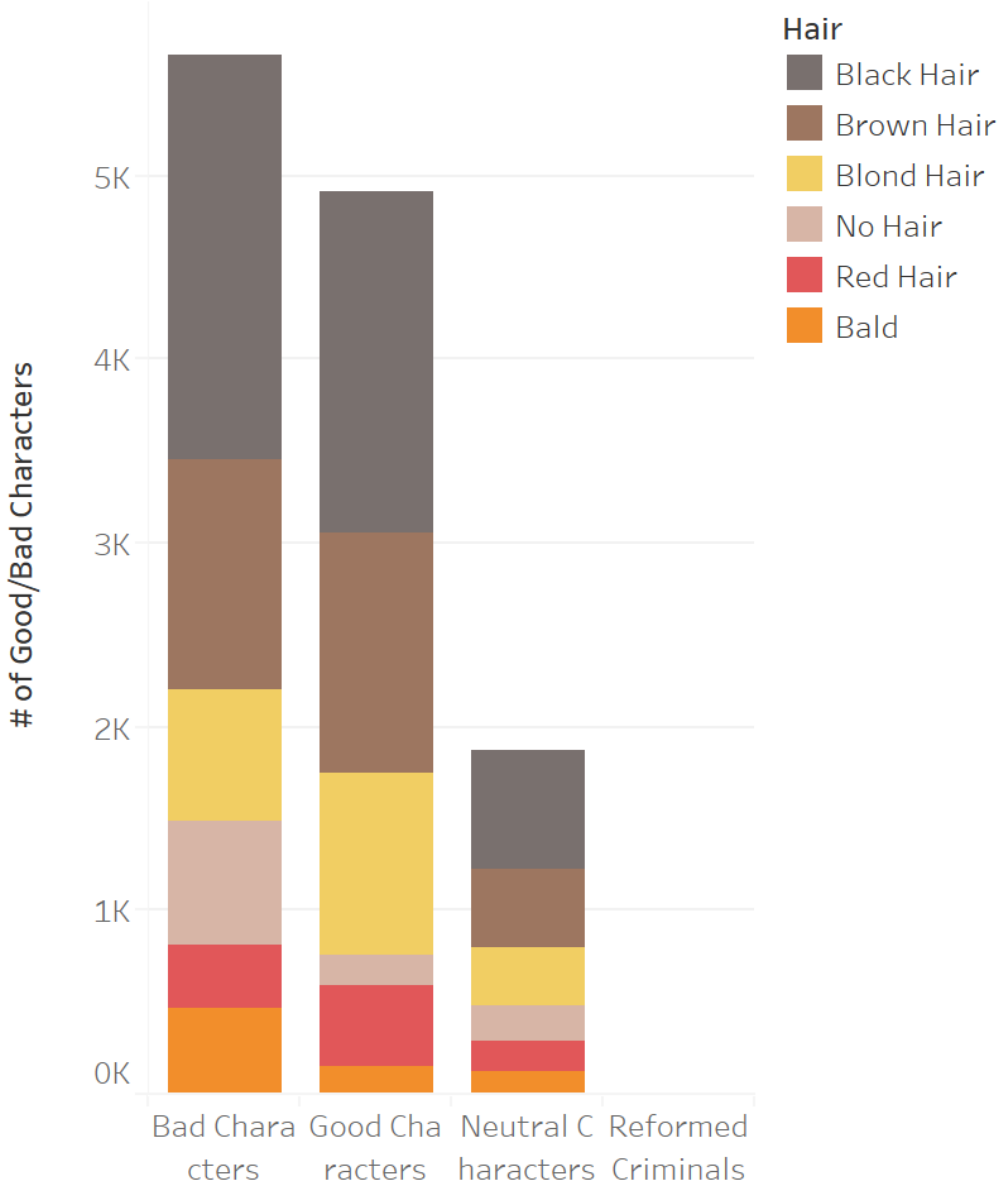
Count of Name

1 4,337

Count of Name broken down by ID vs. Align. Color shows count of Name. The marks are labeled by count of Name. The data is filtered on SEX, which keeps Female Characters and Male Characters. The view is filtered on Align and ID. The Align filter keeps Bad Characters, Good Characters, Neutral Characters and Reformed Criminals. The ID filter excludes Identity Unknown.

I tried to find if there is any association between hair color and alignment (good or bad) of characters. But there was no significant relation between them. Most of the hair colors used were black, brown and blond. The production companies should have used these colors for the printing and other extraneous reasons.

# Characterization based on Hair Color




Count of Align for each Align. Color shows details about Hair. The view is filtered on Align, Exclusions (Align,Hair) and Hair. The Align filter keeps Bad Characters, Good Characters, Neutral Characters and Reformed Criminals. The Exclusions (Align,Hair) filter keeps 96 members. The Hair filter keeps 6 of 29 members.