new file

2024-04-17

THEMES OF THE TOP 1000 ANIME IN 2023

Rationale:

Japanese animation, also known as **Anime**, has gained immense popularity over the years. As I grew up watching anime, the topic of which Genre or Theme was more interesting always sparked a debate between me and my friends. And thus, for my final project for PSY6422, I will try to visualise common recurring themes in the top 1000 highly rated anime of the year 2023 on My Anime List.

Action, adventure, comedy, drama, romance, fantasy, sci-fi, and many more genres are covered in anime. Better recommendation systems can be created by having a better understanding of which genres perform better. The reason why I choose themes over genre is because one anime could have multiple themes thus making our criteria more inclusive and relatively comprehensive.

The data will be visualised in a bar graph as it is approparite to convery the information I want to visualise (Ranking the count of themes & sources). I am using Plotly to create an interactive version of my plot as the graph as multiple columns which can be hard to follow & the reader might benefit from the interactive nature of the graph by simply hovering over the column to see its description i.e, The name of the theme and the count.

Questions I aim to visualise:

- The top animes had which themes?
- The top animes were derived from which source?

A still from Tenki no ko by Makato Shinkai

Source of the Dataset :

The dataset was acquired from Md Kazi Sajiduddin on kaggle. It was created around July 2023. Jikan Application Programming Interface (4.0.0) was used to extract the anime dataset via the My Anime list. The original dataset retrived anime-related data, inclduing the original title, the english title, Demographics, Start season, Airing date,Format, Studios, Synopsis, Production house, The User ID and the scores given by the users. MyAnimeList.

What is My Anime List?

Frequently shortened as MAL, MyAnimeList is a volunteer-run website that provides social networking and social cataloging services for fans of anime and manga. Users of the website can score and arrange anime and manga using a system similar to a list. It offers a comprehensive database on anime and manga and makes it easier to find users with similar interests.

What will my project include?

The data included 24,985 anime titles that were rated by users on My Anime List. The original dataset had a plethora of information including The original title, english title, Demographics, Start season, Airing date,Format, Studios, Synopsis, Production house, The User ID and the scores given by the users. For my project, I will examine the top 1000 anime titles in the dataset to identify recurring themes. Additionally, I will also visualise if the *Source* of the anime, taking a look at if the anime was derived from a Manga *(Comic book)*, Web novel, Light novel or was an original creation and so on. Thus, I make sure only these columns are retrived from the rawdata.

Folders in my project:

The /Data consists of the raw data acquired from kaggle, /figures consist of the Plots generated in the project and /images consist of the image used in the project.

A table of the total theme count from top rated 1000 anime titles on My Anime List:

```
kable(theme_counts, format = "markdown")
```

Themes	Count
School	251
Adult Cast	98
Historical	80
Psychological	79
Super Power	73
Mythology	63
Military	62
Isekai	60
Gore	48
Mecha	48
Gag Humor	44
Iyashikei	39
Parody	39
Music	36
Love Polygon	35
Team Sports	32
Reincarnation	27
Time Travel	26
Workplace	26
CGDCT	25
Harem	25
Organized Crime	25
Space	25
Otaku Culture	24
Survival	23
Detective	22
Vampire	22
Romantic Subtext	20
Childcare	19
Martial Arts	19
Samurai	19
Video Game	17
Mahou Shoujo	16
Strategy Game	13
Anthropomorphic	12
Performing Arts	11
Visual Arts	11
Racing	10
Combat Sports	9
Delinquents	7
High Stakes Game	7

Themes	Count
Idols (Female)	6
Showbiz	6
Reverse Harem	4
Crossdressing	2
Educational	1
Magical Sex Shift	1
Medical	1
Pets	1

An interactive plot of the Theme count

```
# Assigning the rainbow theme to each unique theme in theme_count
theme_colors <- rainbow(length(unique(theme_counts$Themes)))
hover_text <-paste('Theme:', theme_counts$Themes, '<br>Count:' , theme_counts$Count) #setting the hover

# Creating the first graph with ggplot
fig1 <- ggplot(theme_counts, aes(x = reorder(Themes,-Count), y = Count, text = hover_text)) +
    geom_bar(stat = 'identity', fill = theme_colors) +
    labs(x = 'Themes', y = 'Count', title = "THEMES OF THE TOP 1000 ANIME IN 2023", caption = "Based on

    theme_minimal() +
    theme(
        plot.background = element_rect(fill = 'black'),  # To create a black background
        panel.background = element_rect(fill = 'black'), # To create a black panel background
        panel.grid.major = element_line(color = 'transparent'),  # To make major gridlines transparent
        axis.line = element_line(color = '#FFFFFF'),  # axis lines colour set as White
        axis.text = element_text(color = '#EEB4B4'),  # axis text colour set as rosybrown2
        axis.title = element_text(color = 'skyblue'), # axis title colour set as skynlue
        plot.title = element_text(color = 'skyblue', size = 18),  # Plot title colour set to blue & siz
        axis.text.x = element_text(angle = 45, hjust = 1, size = 7)  # x-axis text angle was adjusted t
    ) +

    guides(fill = FALSE)  # Removing the legend as the name of the column and count can be seen in the

#assigning the plot to plotly fr an interactive graph
fig1 <- ggplotly(fig1, tooltip= 'text')
fig1
```

```
# Saving the figure in the figures folder
ggsave(here('Figures', 'Themes_graph.png'))
```

A table of the total source count from top 1000 anime titles on My Anime List:

Source	Count
Manga	535
Original	160
Light Novel	154
Web Manga	39

Source	Count
Novel	35
Visual Novel	27
Yonkoma	24
Game	11
Other	9
Web Novel	3
Music	2
Card Game	1

Second Graph that plots the source of the top 1000 anime of 2023

```
# Creating the second bar graph in ggplot

source_colors <- rainbow(length(unique(source_counts$Source))) # Setting up rainbow themes for the grap
hover_text <-paste('<br>Count:' , source_counts$Count) #setting the hover text

fig2 <- ggplot(source_counts, aes(x =reorder(Source,Count), y = Count, fill = Source,text= hover_text))
    scale_y_continuous(breaks = seq(0, max(source_counts$Count), by = 100)) + # To make the intervals o
    geom_bar(stat= 'identity')+
    labs(x= 'Source', y= 'Count', title = "The source of the anime") +
    coord_flip() + # To create a horizontal chart
    theme_minimal() +

  theme(
        plot.background = element_rect(fill = 'black'),  # To create black background
        panel.background = element_rect(fill = 'black'), # To create a black panel background
        panel.grid.major = element_line(color = 'transparent'),  # To make major gridlines transparent
        axis.line = element_line(color = '#FFFFFF'),  # axis lines colour set as White
        axis.text = element_text(color = '#EEB4B4'),  # axis text colour set as rosybrown2
        axis.title = element_text(color = 'skyblue'), # axis title colour set as skyblue
        plot.title = element_text(color = 'skyblue', size = 14),  # Plot title colour set to blue & siz

    ) +

    scale_fill_manual(values = source_colors) +  # setting the colours in the plot
    guides(fill = FALSE) # removing the legend because the plot is interactive and the names and count

#assigning the plot to plotly for an interactive graph
fig2 <- ggplotly(fig2, tooltip = 'text')
fig2
```

```
# Saving the figure in the figures folder
ggsave(here('Figures', 'Source_graph.png'))
```

Insights: It is evident that the *School* theme is among the most highly regarded animes of 2023, making it one of the most popular recurring themes.Some of the top-rated themes in the 2023 anime ratings were Adult Cast, Historical, Psychological, Super Power, Mythology, Military, and Isekai. In the second graph we observe that a considerable amount of anime that were highly rated in 2023 were derived from Manga, followed by Light novels and original plots.

Closing remarks: I was able to pick up a new skill at my own pace with this module. I can say have relatively become capable of using R Studio and Github over time. I also used this chance to investigate various packages and themes that might improve my project in some way. I also explored managing project environments with *renv* to ensure the required packages are installed appropraitely over different devices. If I had more time to work on the project, I would have loved to plot all of the variables based on various criteria (for example, contrasting highly rated versus low rated anime titles) to have a comprehensive understanding of criteria that make an anime series hghly rated. I also had attempted to scrap the dataset for the cuurent year via My Anime List but was unsuccessful in doing so, therefore working on my web scraping skills would also be one of my future goals.

References :

- Anime Dataset 2023. (2023, July 28). Kaggle. https://www.kaggle.com/datasets/dbdmobile/myanimelist-dataset
- Grolemund, G., & Wickham, H. (2014). R for Data science. https://edtechbooks.org/r_data_science
- MyAnimeList.net: *anime and manga database and community.* (n.d.). MyAnimeList.net. https://myanimelist.net/
- R CODER. (n.d). *R colors [Full List, Color Converter and Color Picker]* | R CHARTS.