



Báo cáo thực hành phân tích dữ liệu trang facebook Cartoon Network Asia

A. Thực hiện cào dữ liệu:

- Lưu file dưới tên: "CartoonNetworkAsia_81Post_Raw.csv"

B. Làm sạch và tiền xử lý dữ liệu:

1. Làm sạch dữ liệu:

- Đọc file "CartoonNetworkAsia_81Post_Raw.csv"

- Lựa chọn các thuộc tính cần phân tích:

post_id	post_text	time	likes	comments	shares	reactions	reaction_count
---------	-----------	------	-------	----------	--------	-----------	----------------

- Loại bỏ các cột có dữ liệu NULL, NaN

- Lưu file dưới tên: "CartoonNetworkAsia_81Post_Processing_01.csv"

2. Tiền xử lí dữ liệu:

- Tách thuộc tính reactions thành từng loại reaction:

- Đọc file "CartoonNetworkAsia_81Post_Processing_01.csv"

- Tách thuộc tính reactions thành từng loại reaction

post_id	post_text	time	likes	comments	shares	reaction_count
---------	-----------	------	-------	----------	--------	----------------

- Lưu file dưới tên: "CartoonNetworkAsia_81Post_Processing_02_SeperatedReactions.csv"

- Tách lấy giá trị hour, month, day từ thuộc tính time:

- Đọc file: "CartoonNetworkAsia_81Post_Processing_02_SeperatedReactions.csv"

- Tách giờ, ngày, tháng đăng bài từ thuộc tính time

post_id	post_text	likes	comments	shares	reaction_count	like
---------	-----------	-------	----------	--------	----------------	------

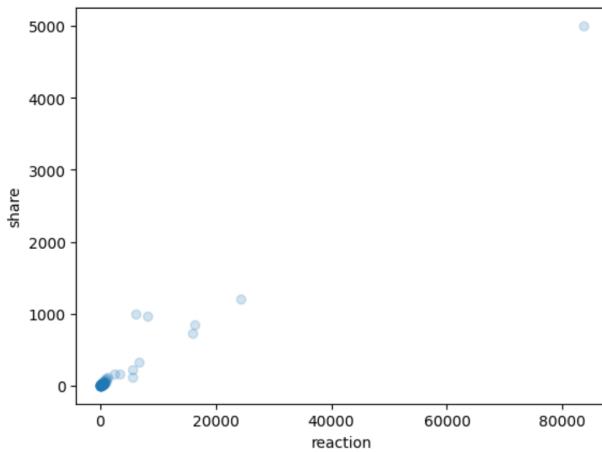
- Lưu file dưới tên: "CartoonNetworkAsia_81Post_Processing_03_SeperatedHourDate.csv"

- Lưu file dữ liệu đã hoàn thành xử lí dưới tên: "CartoonNetworkAsia_81Post_ProCESSED.csv"

C. Phân tích dữ liệu:

1. Tương quan giữa số lượng reaction và comment, reaction và share, tỉ trọng giữa các loại reaction:

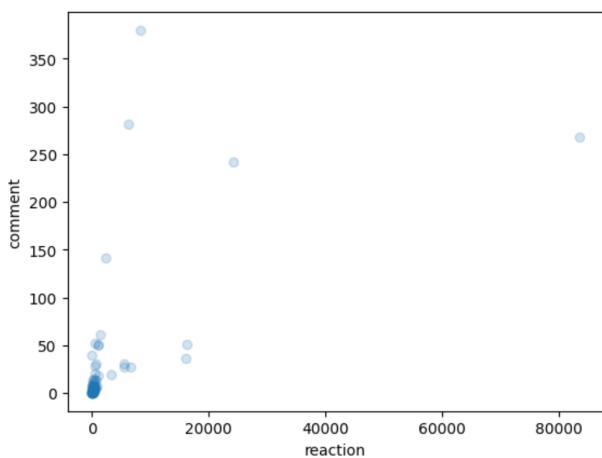
- Tương quan giữa reaction và share



- Nhận xét:

- số lượng reaction và share tỉ lệ thuận

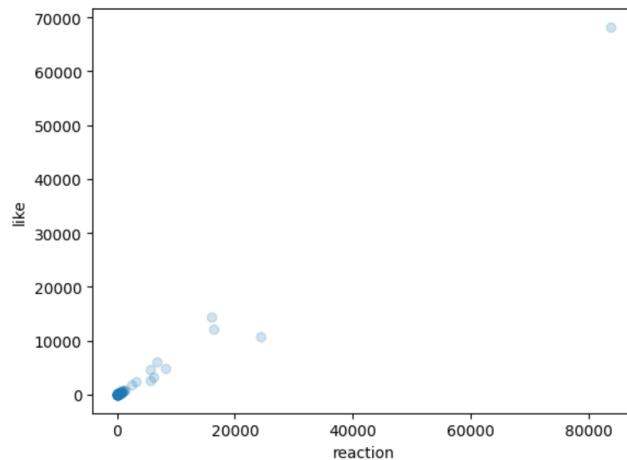
b. Tương quan giữa reaction và comment



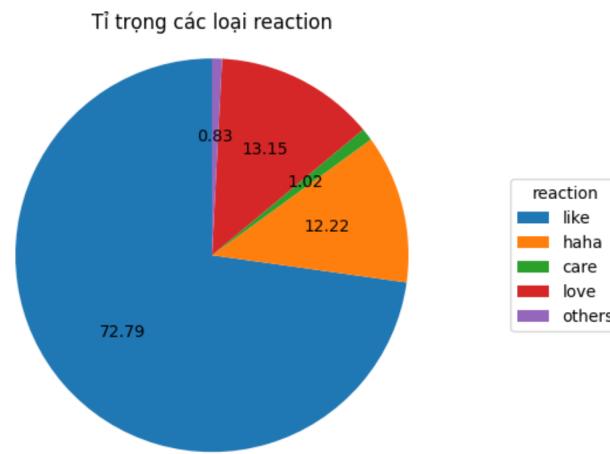
- Nhận xét:

- số lượng reaction và comment không quá phụ thuộc vào nhau
- số lượng comment ít hơn rất nhiều so với reaction

c. Tương quan giữa tổng số reaction và like

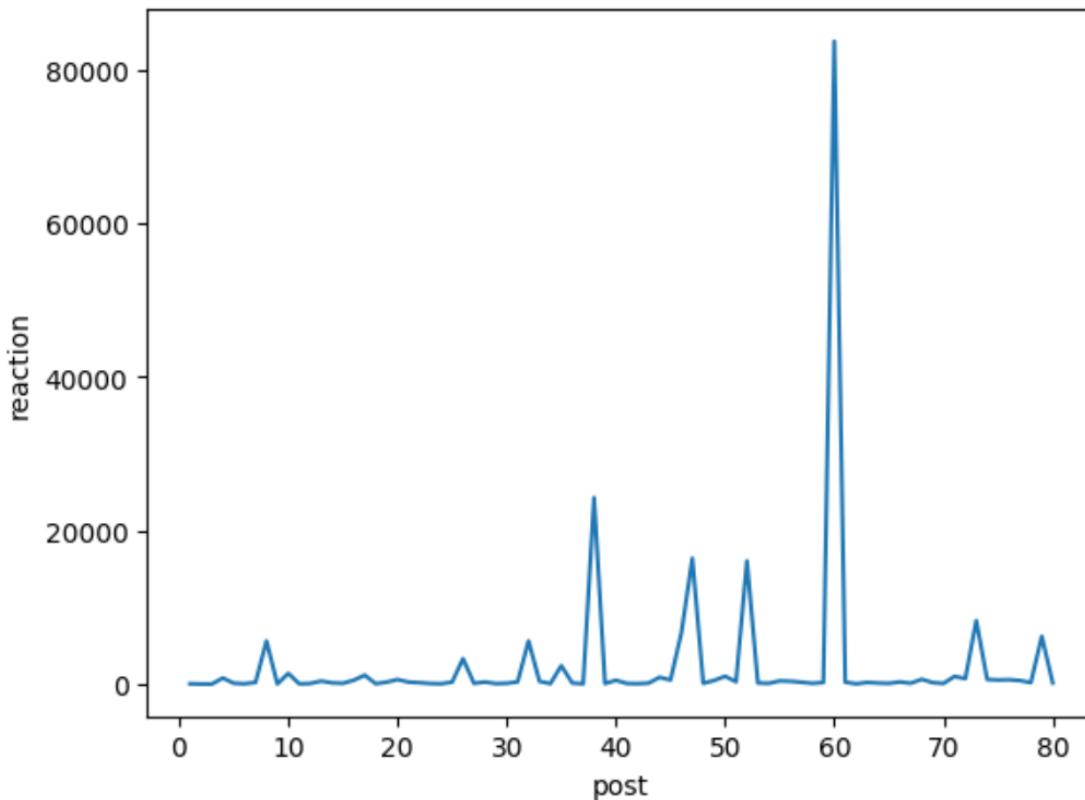


- Nhận xét:
 - Số lượng reaction và like tỉ lệ thuận, cho thấy rằng like là loại reaction chủ yếu của các bài đăng
- d. Tỉ trọng giữa các loại reaction



- Nhận xét:
 - Like là loại reaction chủ yếu, chiếm 72% tổng số lượng reaction của các bài đăng, tiếp theo đó là love với 13% và haha với 12,22%, care là loại reaction có tỉ trọng thấp hơn với 0.02%
 - Các loại reaction khác như wow, sad, angry có tỉ trọng dưới 1%, trong đó angry chiếm tỉ trọng thấp nhất với 0.1%

e. Biến động số lượng reaction:



- Nhận xét:

- Số lượng reaction qua từng bài đăng trong khoảng bài đăng số 10 đến 37 khá ổn định, từ bài đăng số 38 trở đi số lượng reaction biến động rõ rệt, đặc biệt là các bài đăng số 38, 46, 47, 52, 60, 73, 79 có lượng reaction khá cao, bài đăng số 60 có số lượng reaction cao bất thường

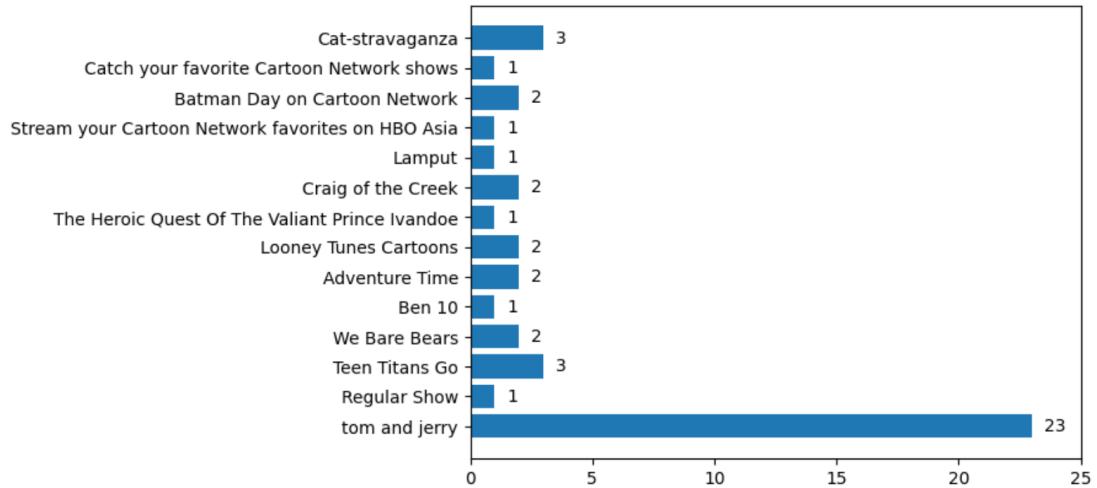
f. Nhận xét về mối quan hệ giữa reaction, share, comment, giữa tỉ trọng các loại reaction:

- Người xem chủ yếu chỉ thà like, số lượng ít hơn sẽ share khi thấy có bài đăng của Cartoon Network Asia chứ không thật sự đọc nội dung, có thể thấy được từ số lượng reaction nhiều nhưng chủ yếu là like nhưng rất ít comment
- Số lượng reaction của trang không ổn định, điều này phụ thuộc vào nội dung bài đăng, thời gian đăng bài và sự kiện xung quanh thời gian đăng bài

2. Tương quan giữa nội dung bài đăng và số lượng reaction:

a. Về các bài đăng về 1 chương trình cụ thể:

- i. Số lượng bài đăng của từng chương trình:



- Tom and Jerry: 23 bài
- Regular Show 1 bài
- Teen Titans Go 3 bài
- We Bare Bears 2 bài
- Ben 10 1 bài
- Adventure Time 2 bài
- Looney Tunes Cartoons 2 bài
- The Heroic Quest Of The Valiant Prince Ivandoe 1 bài
- Craig of the Creek 2 bài
- Lamput 1 bài
- Stream your Cartoon Network favorites on HBO Asia 1 bài
- Batman Day on Cartoon Network 2 bài
- Catch your favorite Cartoon Network shows 1 bài
- Cat-stravaganza 3 bài

ii. Tỉ trọng số lượng bài đăng về các chương trình:

Tỉ trọng số lượng bài đăng về các chương trình

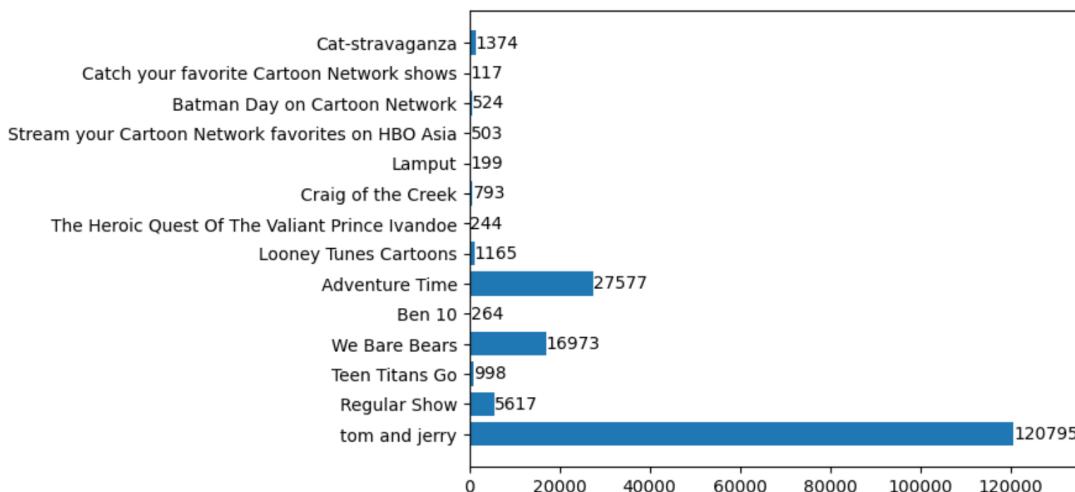


iii. Nhận xét:

- Tom and Jerry là chương trình được đăng nhiều bài nhất với tổng cộng 23 bài đăng, chiếm 51.11% số lượng bài đăng về các chương trình trên CartoonNetworkAsia
- Các chương trình khác có số lượng bài đăng dao động trong khoảng 1 đến 3 bài đăng

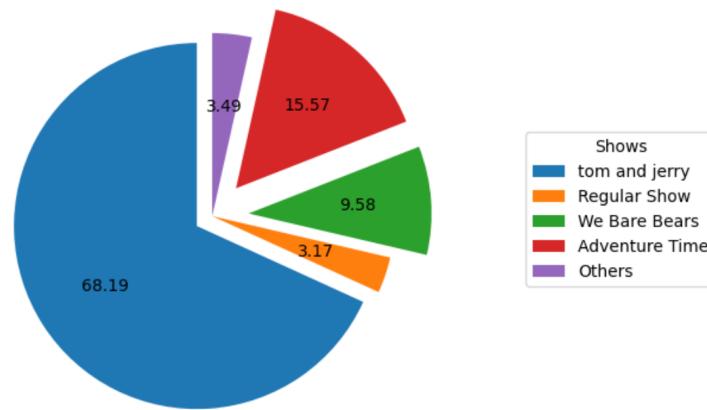
b. Về số lượng reaction của từng bài đăng về 1 chương trình cụ thể:

i. số lượng reaction của từng chương trình



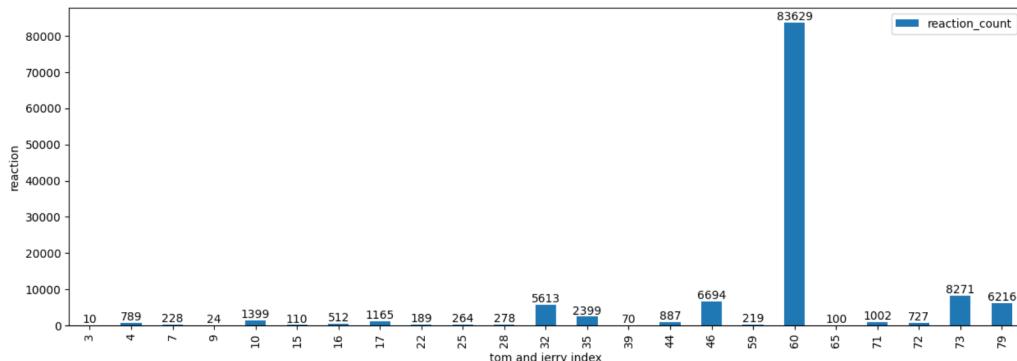
ii. Tỉ trọng về reaction giữa các chương trình

Tỉ trọng số lượng reaction giữa các chương trình

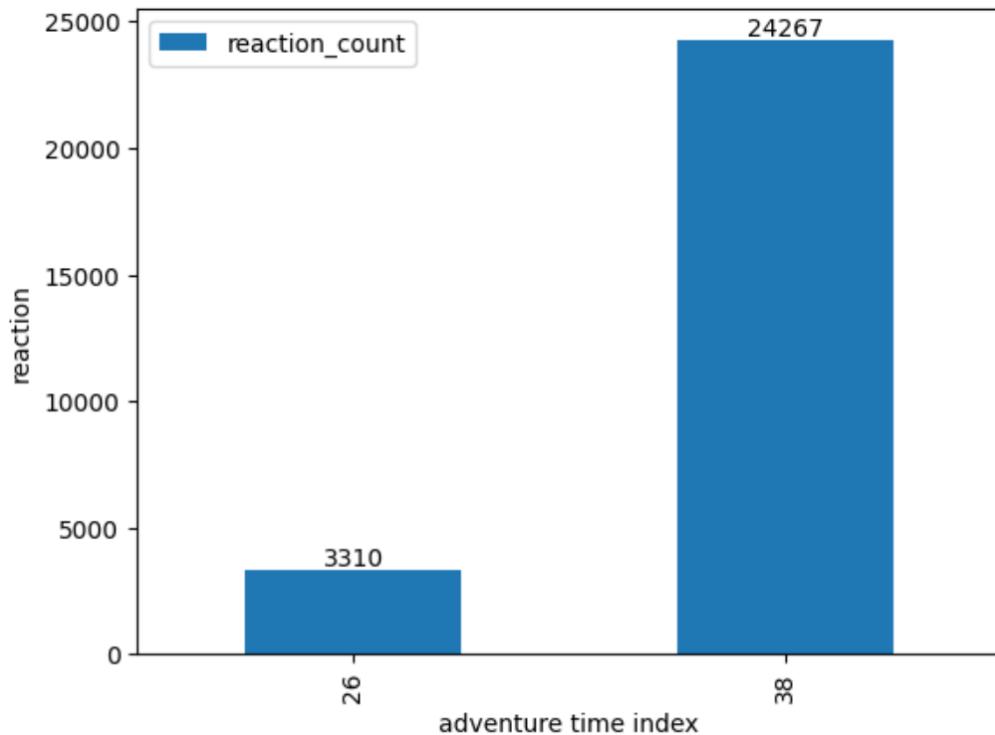


iii. Nhận xét:

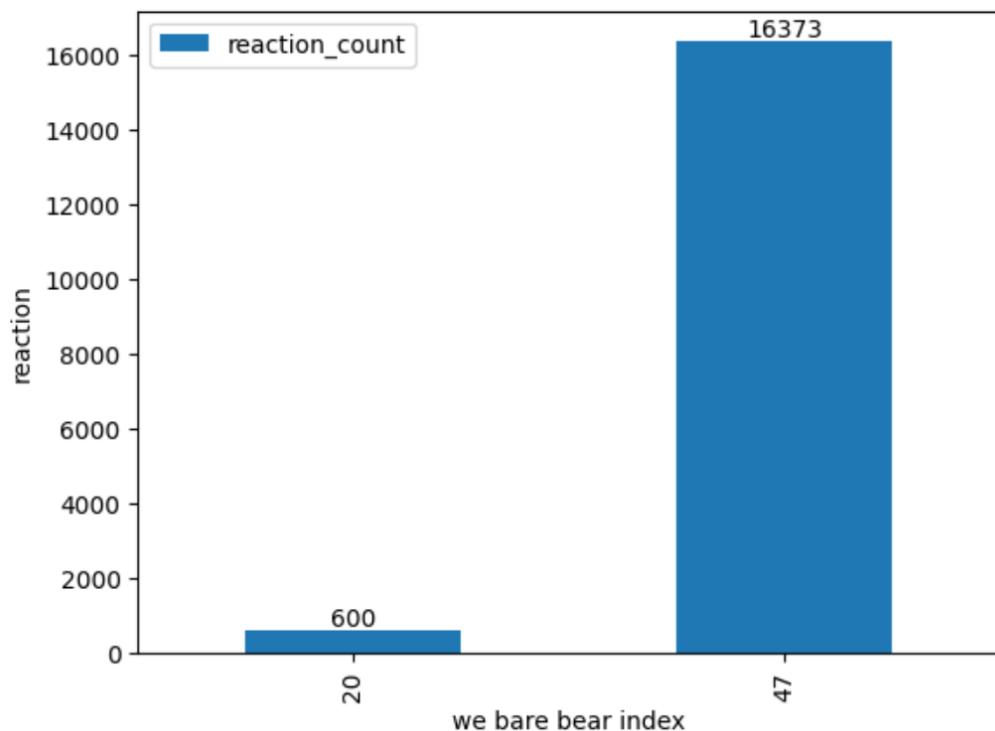
- Tom and Jerry là chương trình có số lượng reaction cao nhất với tổng cộng 120795 reaction, chiếm 68.19% số lượng reaction về các chương trình trên CartoonNetworkAsia
- Adventure Time có lượng reaction cao thứ 2 với tổng cộng 27577 reaction, chiếm 15.57%
- We Bare Bear có lượng reaction cao thứ 3 với tổng cộng 16973 reaction, chiếm 9.58%
- Regular Show có lượng reaction cao thứ 4 với tổng cộng 5617 reaction, chiếm 3.49%
- Các chương trình khác có số lượng reaction dưới 1%
- So với số lượng reaction của các bài đăng về Tom and Jerry, các chương trình Adventure Time và We Bare Bear chỉ có 2 bài đăng nhưng đều có lượng reaction rất cao (Adventure Time chiếm 22,83%, We Bare Bear chiếm 14,05%). Ta đặt ra giả thuyết rằng 2 chương trình Adventure Time và We Bare Bear có sự hấp dẫn hơn Tom and Jerry
 - Số lượng reaction của các bài đăng về Tom and Jerry



- Số lượng reaction của các bài đăng về Adventure Time



- Số lượng reaction của các bài đăng về We Bare Bear



- Kết luận:
 - Cả 2 chương trình Adventure Time và We Bare Bear đều có lượt reaction chênh lệch lớn giữa 2 bài đăng

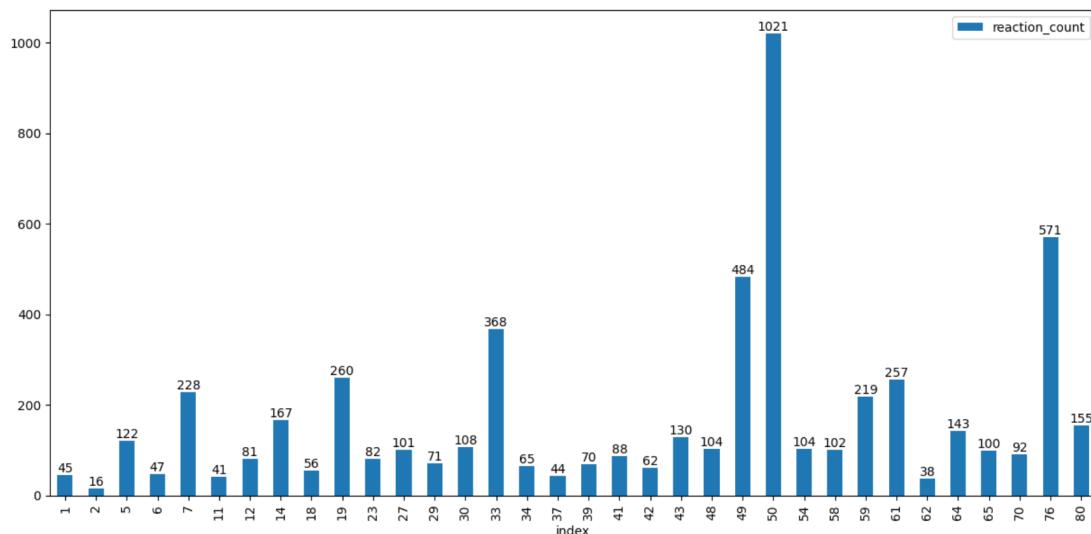
- Adventure Time có lượng reaction của bài đăng số 26 tuy thấp hơn bài đăng số 38 nhưng vẫn cao hơn mức trung bình của Tom and Jerry nên có thể kết luận rằng Adventure Time có sức hút hơn so với Tom and Jerry
- We Bare Bear tuy 2 bài đăng có lượng reaction cao nhưng bài đăng số 20 lại có số lượng reaction không quá cao so với Tom and Jerry nên không thể nói rằng We Bare Bear có sức hút hơn so với Tom and Jerry. Có thể lượt tương tác cao của bài đăng số 47 là do có sự kiện gì đó nổi bật

c. Về các bài đăng không phải các bài đăng về 1 chương trình cụ thể :

i. Các bài đăng ngoài chương trình:

- Cha-Ching
- SafeStepKids
- Lễ kỉ niệm các sự kiện của Warner Bros
- Các sự kiện giao lưu với người xem tại khu vui chơi

ii. Lượt tương tác của các bài đăng:



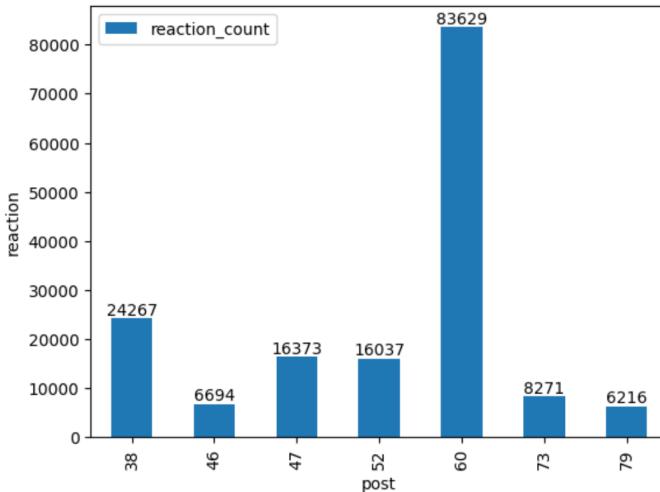
iii. Nhận xét:

- So với các bài đăng về các chương trình được phát sóng, các bài đăng về sự kiện, gameshow, lễ kỉ niệm có lượng reaction có phần ổn định hơn, duy trì ở mức 100 - 1000 reaction

d. Về những bài đăng có lượng reaction cao đột biến:

i. Về các bài đăng 38, 46, 47, 52, 60, 73, 79

- Nội dung:
 - 38: Finn cởi chiếc mũ mà cậu luôn đeo, cậu có 1 bộ tóc dài tuyệt đẹp
 - 46: Thông báo sắp chiếu 1 loạt movie về Tom và Jerry
 - 47: 1 buổi sáng bình thường của gấu trúc khi cậu sống trong trung tâm bảo tồn động vật quý hiếm
 - 52: Tom và Jerry trong bối cảnh truyện cổ tích khi Tom đóng vai quái vật
 - 60: Tom và Jerry đuổi bắt ở New York cùng 3 chú hồng hạc
 - 73: Jerry lần đầu ăn thử sâu riêng
 - 79: Thông báo sắp có phiên bản làm mới của Tom và Jerry, lấy bối cảnh tại Singapore
- Lượng reaction của các bài đăng trên:



- Nhận xét:

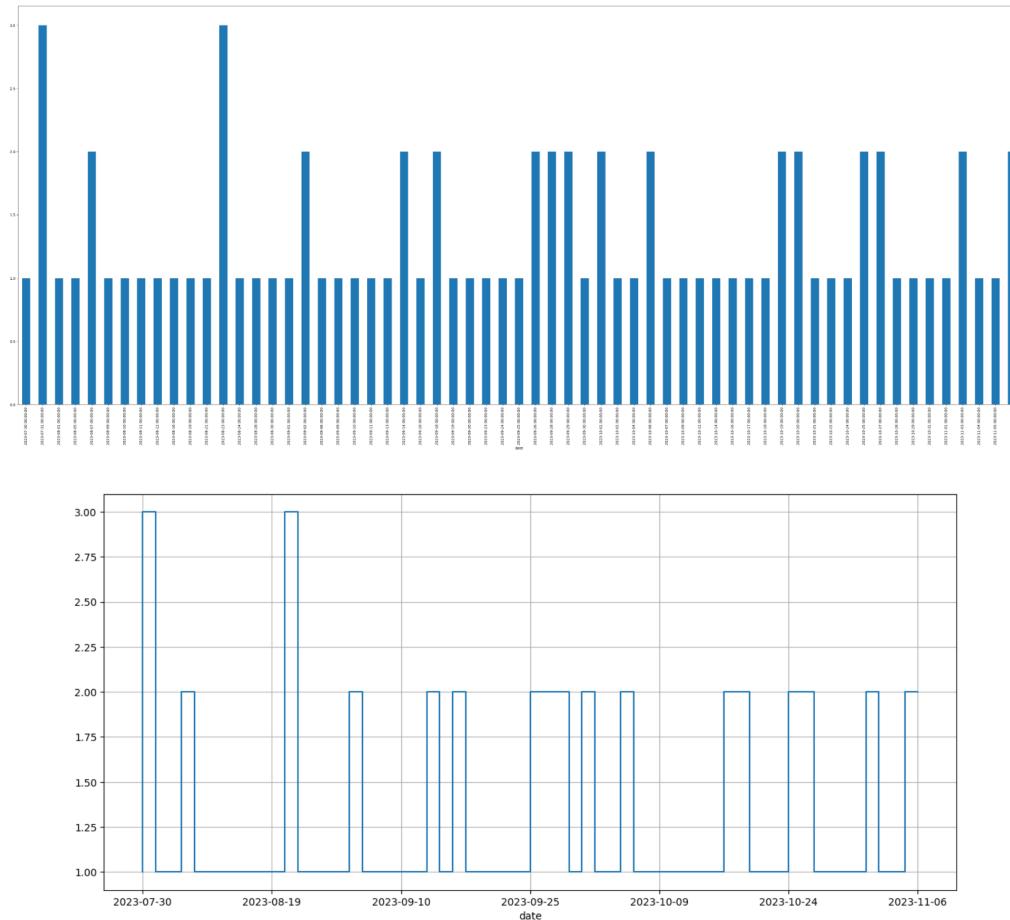
- Trong số 7 bài đăng trên có :
 - 5 bài về Tom and Jerry
 - 1 bài về Adventure Time
 - 1 bài về We Bare Bear

e. Nhận xét về mối quan hệ giữa nội dung và số lượng reaction:

- Tom and Jerry là chương trình được Cartoon Network Asia chú trọng quảng cáo nhất
- Adventure Time là chương trình được khán giả yêu thích nhất
- Đối tượng người theo dõi của Cartoon Network Asia chủ yếu là trẻ em yêu thích Tom and Jerry
- Đối tượng người xem của Adventure Time, và các chương trình lâu đài như We Bare Bear, Regular Show chủ yếu là thiếu niên, người xem cũ thấy bài đăng về chương trình gắn liền với tuổi thơ nên tương tác
- Số lượng bài đăng của trang Cartoon Network Asia về danh mục chương trình và danh mục ngoài chương trình khá cân đối (danh mục về các chương trình: 46 bài, danh mục ngoài chương trình: 34 bài)
- Các chương trình khác như Tom and Jerry, Ben 10, Teen Titans Go đều đã được làm lại phiên bản mới, làm phai một nội dung vốn có của phiên bản gốc, nhằm hướng tới tệp người xem là lứa tuổi trẻ em nên ít tương tác vì trẻ em thường theo dõi trên truyền hình, ít dùng facebook
- Các chuỗi chương trình về 1 sự kiện như 'Stream your Cartoon Network favorites on HBO Asia' hay 'Batman Day on Cartoon Network' chỉ nhằm mục đích thông báo lịch chiếu nên không quá thu hút người xem
- Số lượng reaction của các bài đăng về các chương trình dao động trong khoảng 895.5 reaction/bài, trong khi đó số lượng reaction của các bài đăng ngoài chương trình dao động trong khoảng 101.5 reaction/bài. Cho thấy rằng các bài đăng về 1 chương trình cụ thể sẽ hấp dẫn hơn các bài đăng về sự kiện hay lễ kỉ niệm

3. Cường độ đăng bài:

a. Số bài đăng theo ngày:

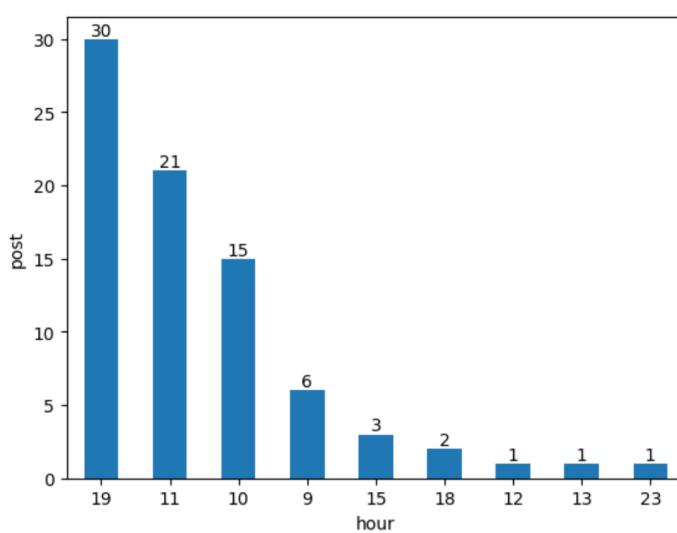


- Nhận xét:

- Cường độ đăng bài của Cartoon Network Asia là 1 bài 1 ngày

3. Tương quan giữa thời gian đăng bài và reaction:

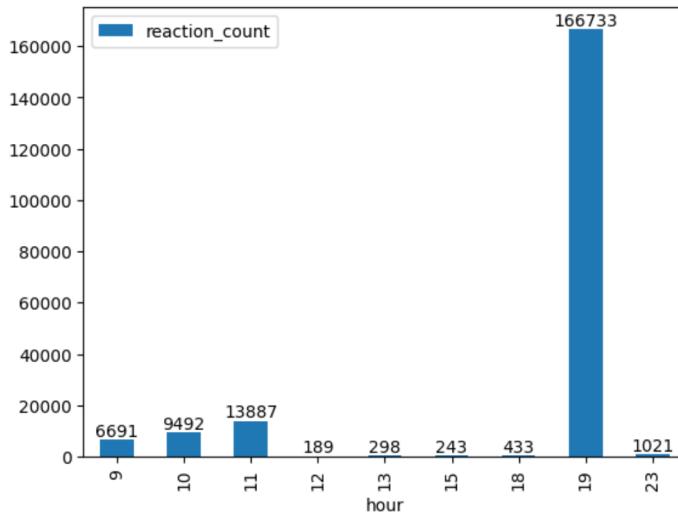
a. Số bài đăng theo giờ:



- Nhận xét:

- Cartoon Network Asia chủ yếu đăng bài vào lúc 19 giờ, tiếp theo đó là 11 giờ và 10 giờ
- Số lượng bài đăng vào lúc 9 giờ, 15 giờ, 18 giờ khá ít
- Số lượng bài đăng vào lúc 12 giờ, 13 giờ, 23 giờ rất ít
- Các thời gian còn lại trong ngày hầu như không đăng bài

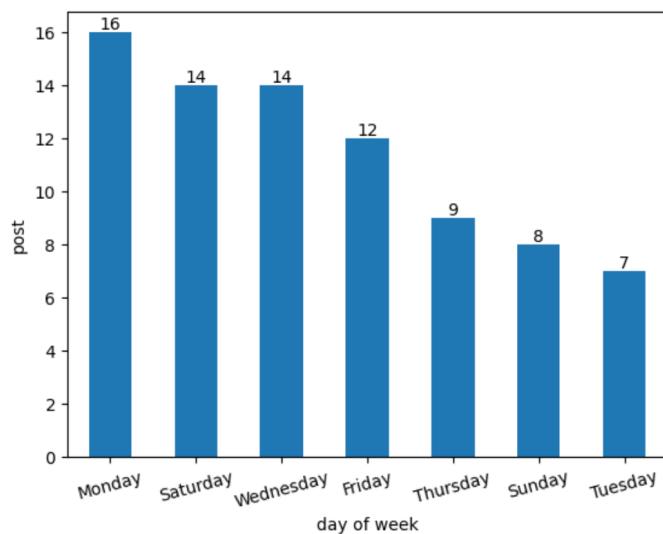
b. Reaction theo giờ:



- Nhận xét:

- Số lượng reaction của trang chủ yếu đến từ các bài được đăng vào lúc 19h
- Các bài đăng vào 11h, 10h, 9h có số lượng reaction ít hơn
- Các bài đăng vào 12h, 13h, 15h, 23h có lượng reaction rất ít

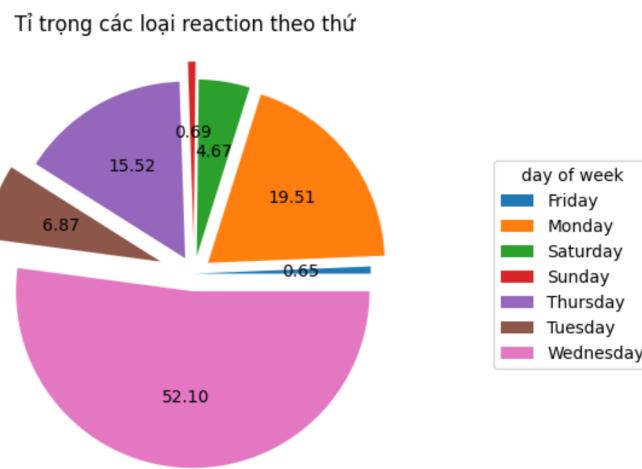
c. Số bài đăng theo thứ:



- Nhận xét:

- Số lượng bài đăng trên trang vào các ngày trong tuần khá đồng đều
- Số lượng bài đăng trên trang vào các ngày trong tuần dao động trong khoảng 12 - 14 bài

d. Reaction theo thứ:



- Nhận xét:
 - Các bài đăng vào thứ 4 có tổng số lượt reaction lớn nhất
 - Các bài đăng vào thứ 2, 3, 5 có tổng số lượt reaction khá cân bằng
 - Các bài đăng vào thứ 7 có tổng số lượt reaction thấp hơn
 - Các bài đăng vào thứ 6 và chủ nhật có tổng số lượt reaction rất thấp (dưới 1%)

e. Nhận xét về tương quan giữa thời gian đăng bài và số lượng reaction:

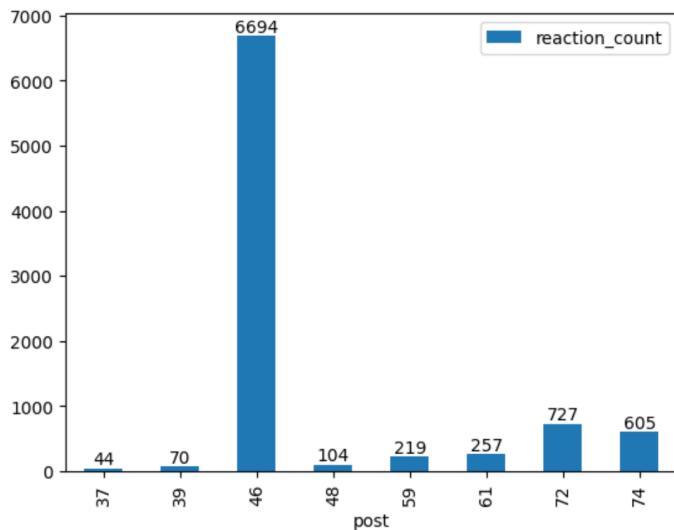
4. Các sự kiện nổi bật:

a. Các bài đăng 38, 46, 47, 52, 60, 73, 79:

- i. Về phần nội dung của các bài đăng trên, chỉ có 2 bài đăng số 46 và 79 là thông báo về 1 sự kiện cụ thể:
 - 46: Thông báo sắp chiếu 1 loạt movie về Tom và Jerry
 - 79: Thông báo sắp có phiên bản làm mới của Tom và Jerry, lấy bối cảnh tại Singapore
- ii. Về các bài đăng xung quanh các bài đăng 38, 47, 52, 60, 73:
 - 37: Sự kiện tham gia đi bộ đường dài tại sở thú Đài Bắc
 - 38: Finn cởi chiếc mũ mà cậu luôn đeo, cậu có 1 bộ tóc dài tuyệt đẹp
 - 39: Thông báo sắp tới lễ kỷ niệm 100 thành lập Warner Bros tại Singapore
 - 46: Thông báo phát sóng chuỗi phim về Tom and Jerry
 - 47: 1 buổi sáng bình thường của gấu trúc khi cậu sống trong trung tâm bảo tồn động vật quý hiếm
 - 48: Chào mừng ngày người dơi trên Cartoon Network, thông báo sắp chiếu 1 chuỗi chương trình liên quan tới batman
 - 59: Thông báo sắp tới lễ kỷ niệm 100 thành lập Warner Bros tại Singapore
 - 60: Tom và Jerry đuổi bắt ở New York cùng 3 chú hồng hạc
 - 61: Thông báo về trang web cung cấp kiến thức về biến đổi khí hậu cho trẻ em

- 72: Chúc mừng sinh nhật lần thứ 58 của Singapore
- [73: Jerry lần đầu ăn thử sầu riêng](#)
- 74: Thông báo sắp chiếu 1 chuỗi các chương trình về mèo trên kênh Cartoon Network

iii. Số lượng reaction của các sự kiện trên:



iv. Nhận xét về các sự kiện xoay quanh các bài đăng nhiều tương tác:

- Các bài đăng về chương trình có số lượt reaction cao và các bài đăng sự kiện xung quanh các bài đăng đó có nội dung không quá liên quan tới nhau
- Số lượng reaction cao đột biến của các bài đăng về các chương trình 38, 47, 52, [60](#), 73 không phải vì các sự kiện nổi bật xung quanh bài đăng đó mà vì sức hút của chương trình đó đối với khán giả (ngoại trừ bài đăng số 46, các bài đăng về sự kiện còn lại có số lượt reaction khá thấp, số lượt reaction của bài đăng số 46 cao có thể vì đó là chương trình Tom and Jerry phiên bản cũ)

D. Tổng kết về trang Cartoon Network Asia:

- Cartoon Network Asia là trang truyền thông của kênh truyền hình Cartoon Network khu vực châu Á, chủ yếu đăng tải các nội dung liên quan tới các bộ phim hoạt hình được chiếu trên kênh, lịch phát sóng, gameshow giao lưu với khán giả trong khu vực châu Á, các hoạt động kỉ niệm của Cartoon Network và Warner Bros.
- Các bài đăng trên trang Cartoon Network Asia phần lớn là phim hoạt hình hướng tới đối tượng thiếu nhi, thanh thiếu niên trong độ tuổi từ 14 đến 16 nhằm quảng bá cho kênh truyền hình Cartoon Network.
- Chương trình nổi bật là Tom and Jerry, với tổng số lượng bài đăng rất lớn, số lượt reaction tổng cộng, số lượt share và comment đều vượt trội hơn rất nhiều so với các chương trình khác được đăng bài trên trang. Tom and Jerry là chương trình được Cartoon Network Asia chú trọng khai thác nhiều nhất và là công cụ duy trì tương tác của trang
- Bên cạnh các bài đăng về chương trình chính trên trang, Cartoon Network Asia cũng thường xuyên đăng các bài về các sự kiện hay hoạt động giao lưu tại các khu vui chơi giải trí, nhằm mục đích tri ân, kỉ niệm các ngày lễ quan trọng của Cartoon Network và Warner Bros.