

# Báo cáo cuối kỳ

Nhóm 9

# Thành viên nhóm

- Nguyễn Thanh Bình – 20133025
- Nguyễn Nhật Triều – 20133102
- Đoàn Quốc Trung – 20133104
- Mai Thành Trung - 20133105

# Đề Xuất

- Tên đề tài : Phân tích dữ liệu các kênh stream trên nền tảng Twitch
- Nguồn tải : <https://www.kaggle.com/datasets/aayushmishra1512/twitchdata>
- Nguồn thu thập dữ liệu : <https://www.twitch.tv/>
- Cách thu thập dữ liệu : Khảo sát các kênh trên Twitch

# Ý tưởng

- Ngành streamer khá linh hoạt về thời gian làm việc, công cụ và thu nhập
- Mức lương cao làm cho ngành này trở nên khá “hot” đối với các bạn trẻ.
- Nhiều nền tảng livestream ra đời để đáp ứng mọi nhu cầu. Trong đó có Twitch
  - Là nơi mà các streamer hoạt động trong một cộng đồng.
  - Một trong những nền tảng livestream lớn nhất.
  - Bao gồm nhiều streamer nổi tiếng của thế giới.
  - Streamer thu hút được nhiều người xem hơn, đồng thời làm tăng việc thu nhập của họ .
- Lượng người theo dõi rất quan trọng đối với 1 kênh stream . Do đó , hãy thử dự đoán lượng người theo dõi của 1 kênh đạt được sau 1 năm là bao nhiêu và phân tích mối quan hệ giữa lượng người theo dõi với các yếu tố còn lại khác của 1 kênh stream điển hình

# Mô tả dữ liệu

- Dữ liệu gồm 11 cột, 1000 dòng.
- Mô tả biến :
  - Channel :Tên kênh của streamer
  - Watch time (minutes): tổng thời gian xem của viewers cho streamer
  - Stream time (minutes): Tổng thời gian streamer phát trực tiếp
  - Peak views : Lượt xem cao nhất của streamer
  - Average viewers : Lượt xem trung bình của viewer
  - Followers : số người theo dõi kênh.
  - Followers gained: số người theo dõi kênh đạt được sau 1 năm
  - Views gained: Tổng số lượt xem đã đạt được sau 1 năm
  - Partnered : Hợp tác với Twitch
  - Mature: Độ tuổi người xem kênh
  - Language : Ngôn ngữ mà streamer sử dụng khi stream

## Mô tả dữ liệu

Số quan sát : 1000

## Kiểu dữ liệu của từng cột

```

Rows: 1,000
Columns: 11
$ Channel      <chr> "xQcow", "summit1g", "Gaulés", "ESL_CSGO", "Tfue", "Asmongold", "NICKMERCs", "Fextralife"...
$ Watch.time.Minutes. <dbl> 6196161750, 6091677300, 5644590915, 3970318140, 3671000070, 3668799075, 3360675195, 33018...
$ Stream.time.minutes. <int> 215250, 211845, 515280, 517740, 123660, 82260, 136275, 147885, 122490, 92880, 108780, 128...
$ Peak.viewers  <int> 222720, 310998, 387315, 300575, 285644, 263720, 115633, 68795, 89387, 125408, 142067, 891...
$ Average.viewers <int> 27716, 25610, 10976, 7714, 29602, 42414, 24181, 18985, 22381, 12377, 25664, 21739, 20960,...
$ Followers     <int> 3246298, 5310163, 1767635, 3944850, 8938903, 1563438, 4074287, 508816, 3530767, 2607076, ...
$ Followers.gained <int> 1734810, 1370184, 1023779, 703986, 2068424, 554201, 1089824, 425468, 951730, 1532689, 124...
$ Views.gained  <int> 93036735, 89705964, 102611607, 106546942, 78998587, 61715781, 46084211, 670137548, 513499...
$ Partnered    <chr> "True", "True", "True", "True", "True", "True", "True", "True", "True", "True", "True", "True", "...
$ Mature       <chr> "False", "False", "True", "False", "False", "False", "False", "False", "False", "False", "False", "...
$ Language     <chr> "English", "English", "Portuguese", "English", "English", "English", "English", "English", "English"

```

# Câu hỏi phân tích

- Dự đoán lượng người theo dõi đạt được sau 1 năm
- Lượng người theo dõi có ảnh hưởng đến lượt average viewers hay không ?
- Thời gian stream có ảnh hưởng đến lượng người theo dõi một của streamer hay không ?
- Độ tuổi của người xem có ảnh hưởng đến lượt xem không ?

# Phân công công việc

- Từ ngày 1/11 đến ngày 27/11 , hoàn thành phần Milestone và phân chia công việc , chuẩn bị nội dung phần Presentation
- Từ ngày 1/12 đến ngày 11/12 hoàn thành phần Presentation , phân chia công việc , chuẩn bị nội dung phần Report
- Từ ngày 12/12 đến ngày 18/12 hoàn thành phần Report và phần Peer Assessment