



GAME PRESENTATION TEMPLATE

CS:GO

Game Presentation

Press Any Key



# DỰ ĐOÁN ĐIỂM SỐ TRẬN CHIẾN



## About Data:

Data & Source code: [click](#)

- Tập dữ liệu này được thu thập nhằm phục vụ cho bài toán dự đoán điểm số (points) của người chơi trong trò chơi Counter-Strike: Global Offensive (CSGO) – một trong những tựa game FPS phổ biến nhất hiện nay.
- Dữ liệu phản ánh hiệu suất cá nhân của người chơi thông qua nhiều đặc trưng kỹ thuật và hành vi trong mỗi trận đấu, bao gồm:
  - Thông số kỹ thuật: số lượng mạng hạ gục (kills), hỗ trợ (assists), số lần chết (deaths), MVPs, ping,...
  - Hiệu suất: tỷ lệ headshot (hs\_percent), thời gian chờ trận (wait\_time\_s), thời gian trận đấu (match\_time\_s)...
  - Thông tin trận đấu: bản đồ (map), tỉ số giữa hai đội, kết quả (result: Win/Lost/Tie) và thời điểm thi đấu (date, day, month, year).

# DỰ ĐOÁN ĐIỂM SỐ TRẬN CHIẾN

## About Data:

-Điểm số (points) là chỉ số phản ánh tổng hợp chất lượng thi đấu của người chơi, được sử dụng làm biến mục tiêu trong mô hình dự đoán.

-Mục tiêu cuối cùng là xây dựng một mô hình học máy có thể dự đoán điểm số người chơi dựa trên các yếu tố đầu vào từ đó có thể ứng dụng vào:

- Đánh giá nhanh trình độ hoặc phong độ người chơi.
- Gợi ý chiến lược đội hình phù hợp.
- Tự động xếp hạng hoặc đề xuất cải thiện kỹ năng.

Learn More



# QUY MÔ DỮ LIỆU

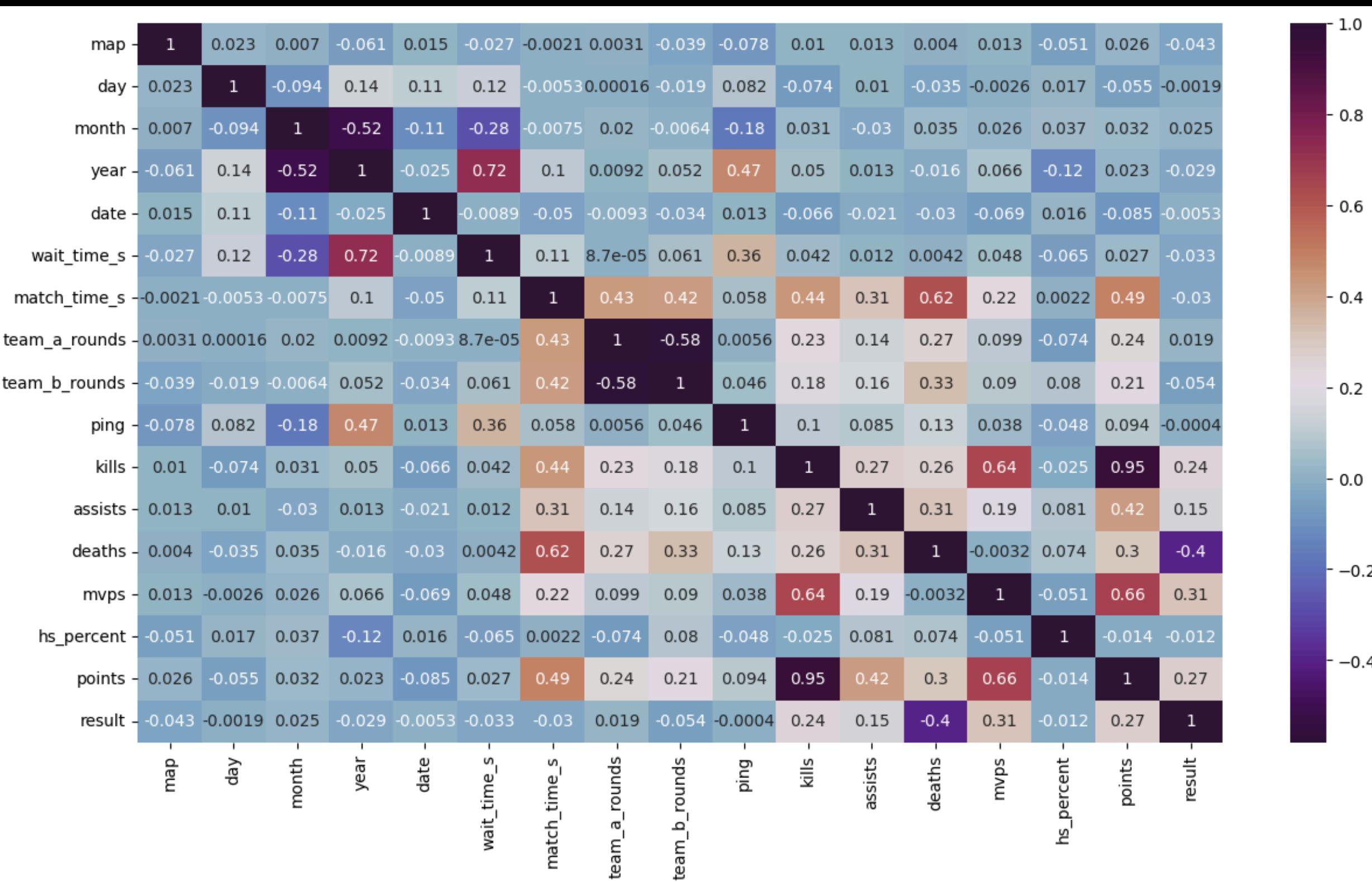
## Thông tin về Data

-Data gồm 1133 dòng và 17 biến dữ liệu

## VARIABLE DESCRIPTION

	Tên biến	Kiểu dữ liệu	Mô tả
1	map	object	Tên bản đồ trong trận đấu (ví dụ: Mirage, Dust2...)
2	day	int64	Ngày diễn ra trận đấu
3	month	int64	Tháng diễn ra trận đấu
4	year	int64	Năm diễn ra trận đấu
5	date	object	Ngày tháng đầy đủ dạng chuỗi (ví dụ: 30/7/2018)
6	wait_time_s	int64	Thời gian chờ ghép trận, tính bằng giây
7	match_time_s	int64	Thời lượng trận đấu, tính bằng giây
8	team_a_rounds	int64	Số round (ván) thắng của đội A
9	team_b_rounds	int64	Số round thắng của đội B
10	ping	int64	Dộ trễ mạng (ping) của người chơi trong trận đấu
11	kills	int64	Số mạng người chơi hạ gục trong trận
12	assists	int64	Số lần hỗ trợ hạ gục đối thủ
13	deaths	int64	Số lần người chơi bị hạ gục
14	mvp	int64	Số lần người chơi được MVP (người chơi xuất sắc nhất round)
15	hs_percent	int64	Tỉ lệ headshot (bắn vào đầu) trong tổng số kill, tính theo %
16	points	int64	Biến mục tiêu – điểm số tổng kết người chơi đạt được trong trận
17	result	object	Kết quả trận đấu: Win, Lost, hoặc Tie

# ĐÁNH GIÁ HỆ SỐ TƯƠNG QUAN VỚI POINT



## I/ Đánh Giá

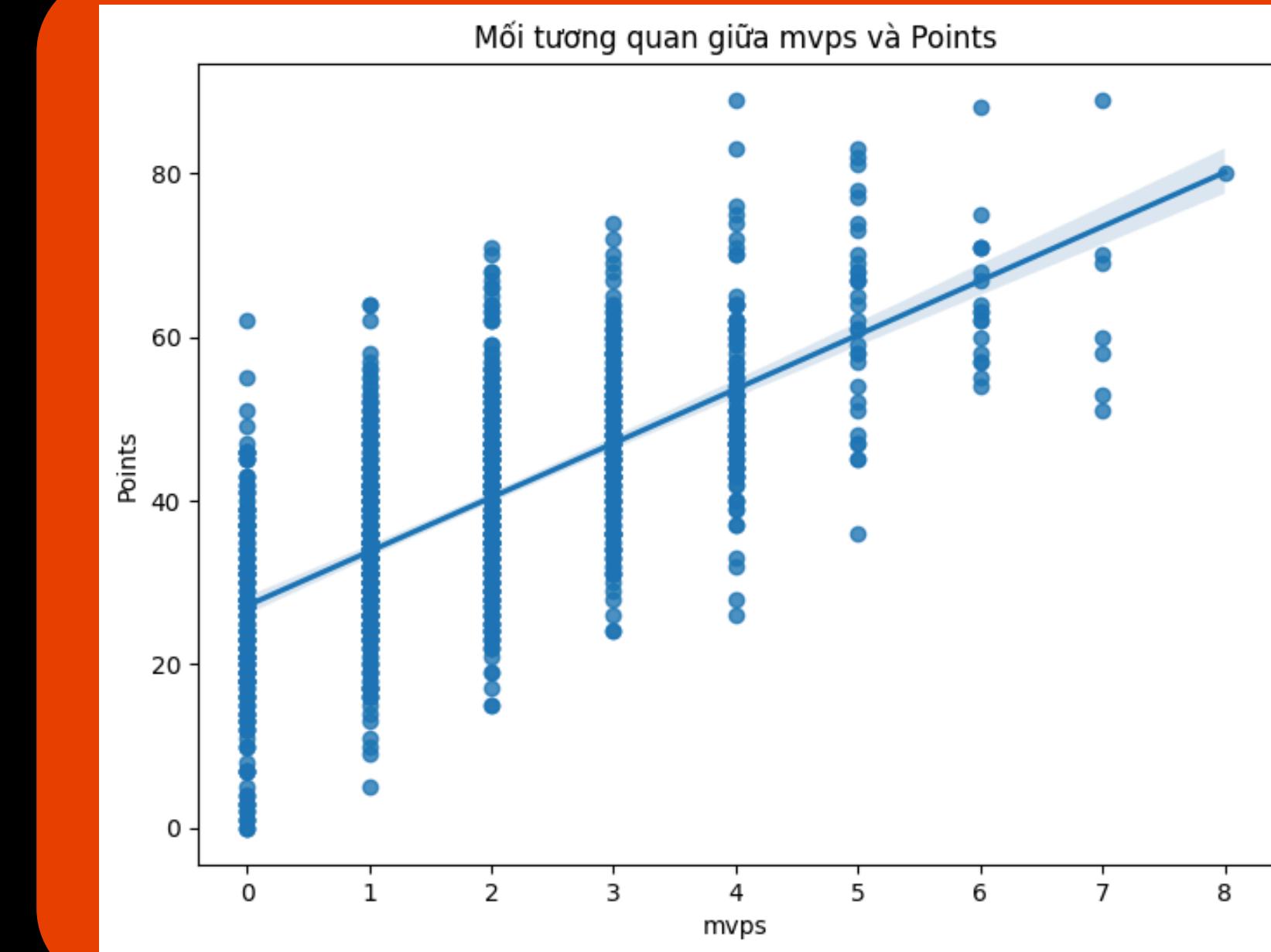
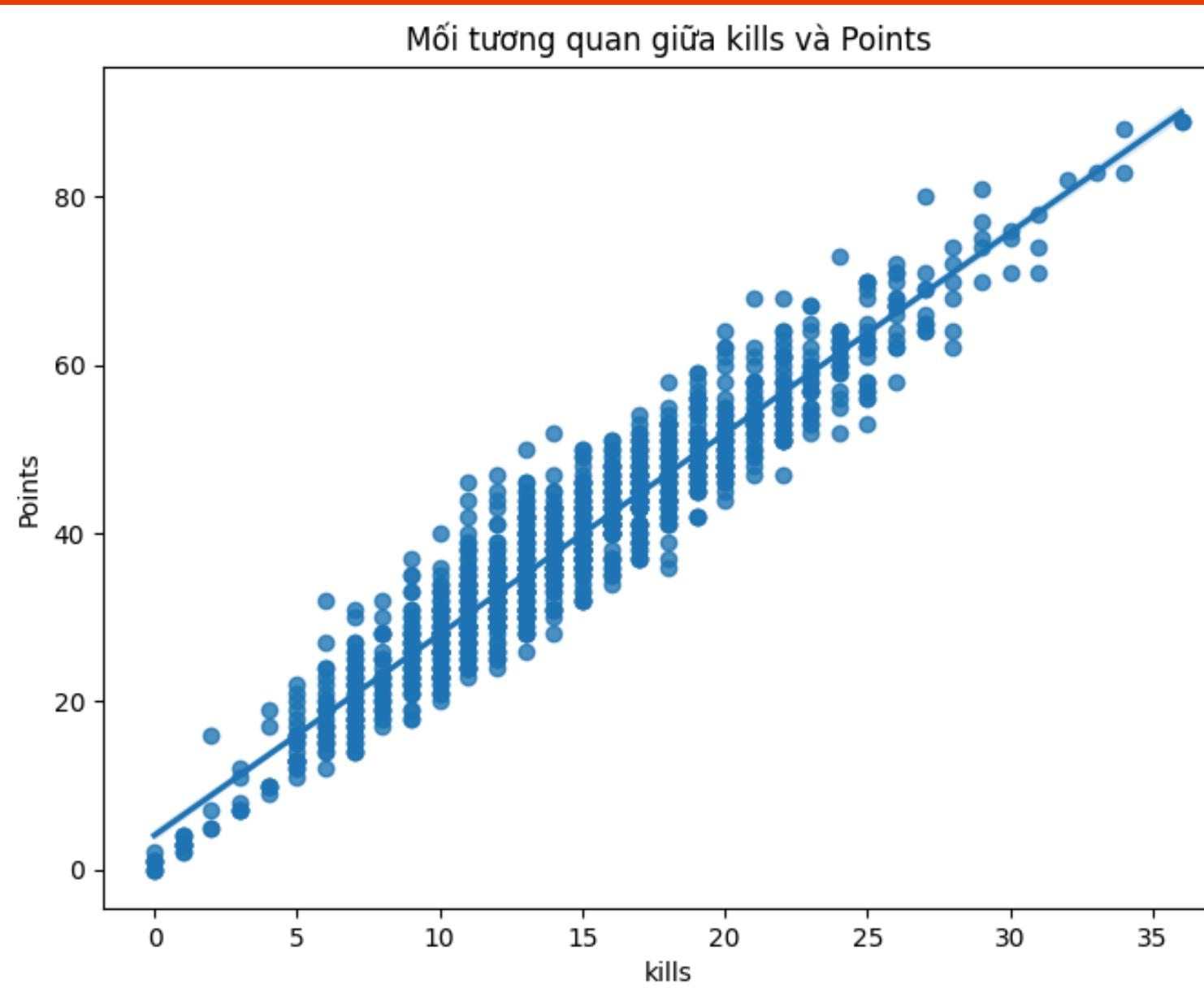
-Kills: (0.95) rất mạnh - số mạng hạ gục là yếu tố chính quyết định số điểm  
 -Mvps: (0.66) mạnh - người chơi được MVP càng nhiều thường điểm sẽ càng cao  
 -Assists: (0.42) trung bình - hỗ trợ diệt địch góp phần tăng điểm  
 -Death: (0.3) yếu - số lần chết tăng lên thì điểm số có xu hướng tăng nhẹ  
 -match\_time\_s: (0.49) trung bình - trận kéo dài thường có điểm cao hơn  
 -result: (0.27) yếu - thắng/ thua có liên hệ đến điểm nhưng không mạnh

-Các biến còn lại ít ảnh hưởng tới điểm số

## II/ Kết Luận

-Kills là yếu tố quyết định nhất đến điểm số, theo sau là mvps, assists, và match\_time\_s.  
 -Một số biến như deaths có tương quan dương nhẹ, cho thấy người chơi càng tham gia sâu (kể cả hy sinh nhiều), vẫn có điểm cao nếu thi đấu hiệu quả.  
 -Không nên đưa những biến ít ảnh hưởng tới điểm số vào mô hình học máy.

# ĐÁNH GIÁ HỆ SỐ TƯƠNG QUAN VỚI POINT



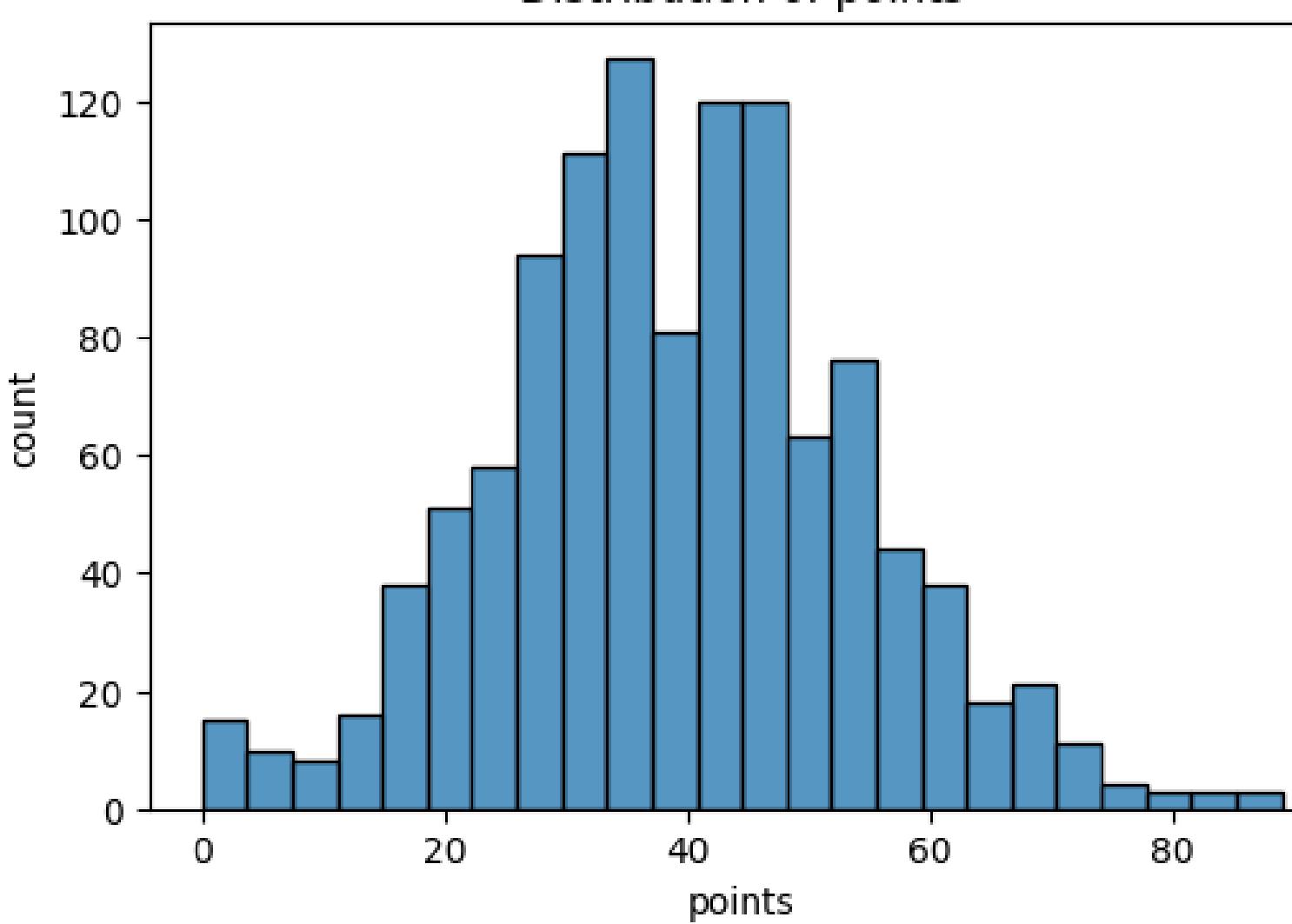
Kills có tương quan mạnh với Points hơn mvps.

Kills nên được ưu tiên làm đặc trưng chính trong mô hình dự đoán điểm số

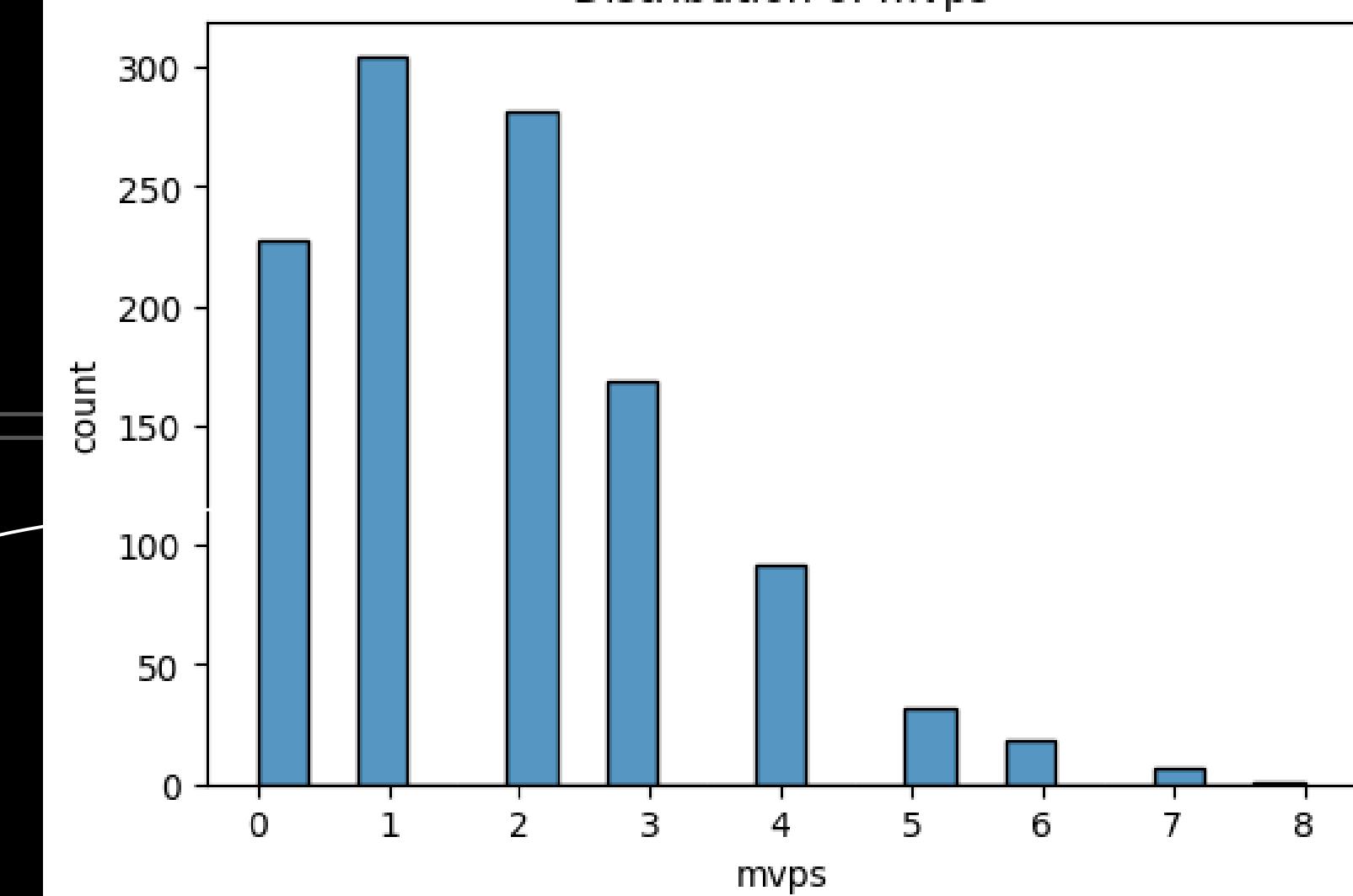
- Tương quan tuyến tính dương mạnh mẽ: Dữ liệu bám sát đường hồi quy, thể hiện mối quan hệ gần như tuyến tính.
- Phân tán thấp: Ít nhiễu, độ chính xác cao hơn khi dự đoán điểm số từ kills.

# Độ PHÂN BỐ DỮ LIỆU

Distribution of points



Distribution of mvps



## 1. points

Phân bố: Gần giống chuẩn, hình chuông.

Độ phân tán: Vừa phải, trải rộng từ 0 đến hơn 80, tập trung quanh trung tâm 30–50.

Nhận xét: Điểm số có sự phân bố đa dạng giữa các cá nhân, cho thấy sự khác biệt khá rõ về hiệu suất.

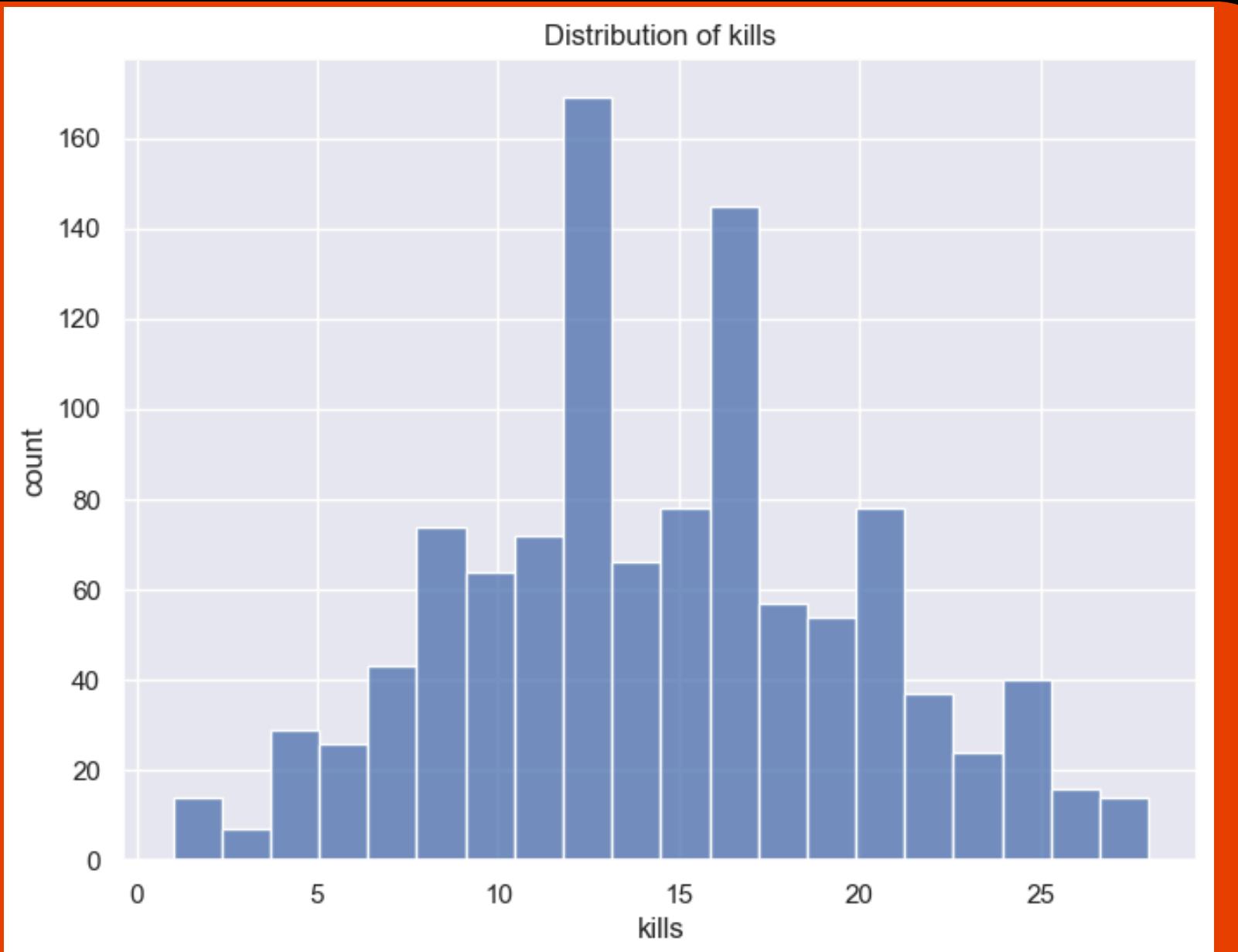
## 2. mvps

Phân bố: Rất lệch phải (skewed right).

Độ phân tán: Nhỏ, giá trị chủ yếu tập trung ở 0–3 MVPs, rất ít người đạt trên 4.

Nhận xét: Chỉ một số ít người chơi có số MVP nổi bật, trong khi phần lớn có rất ít hoặc không có MVP.

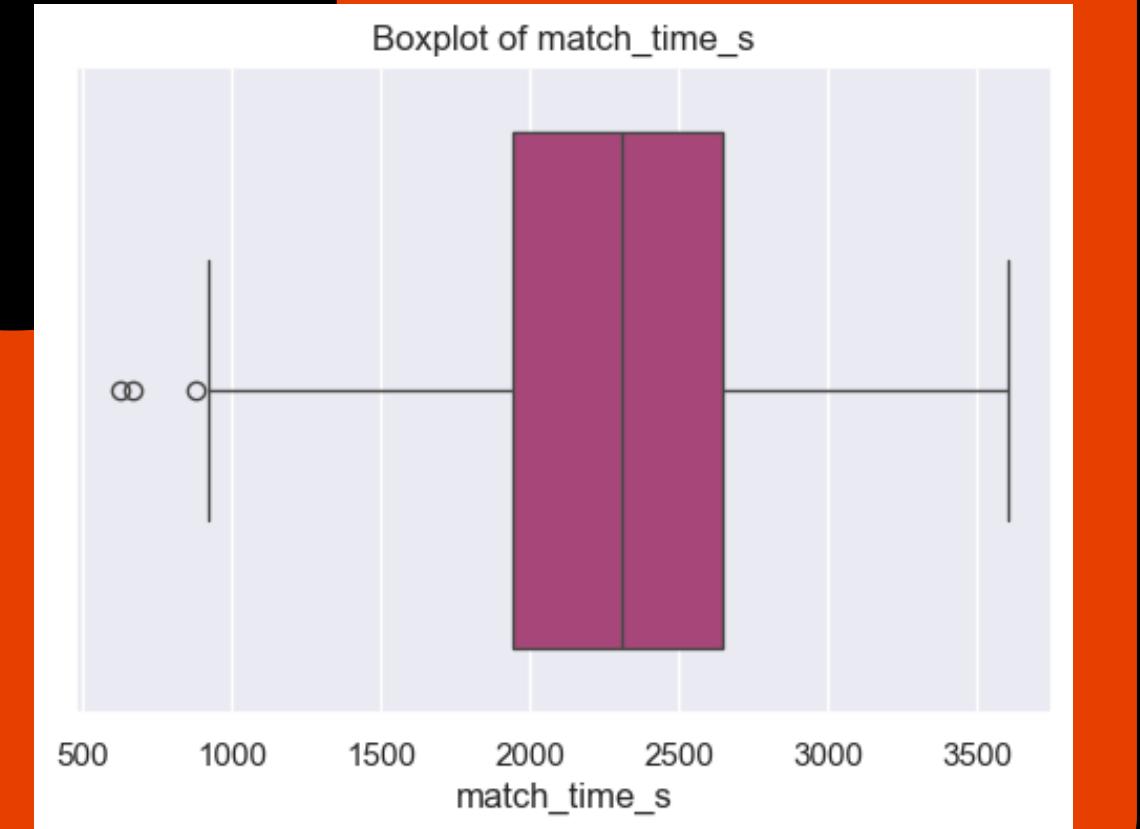
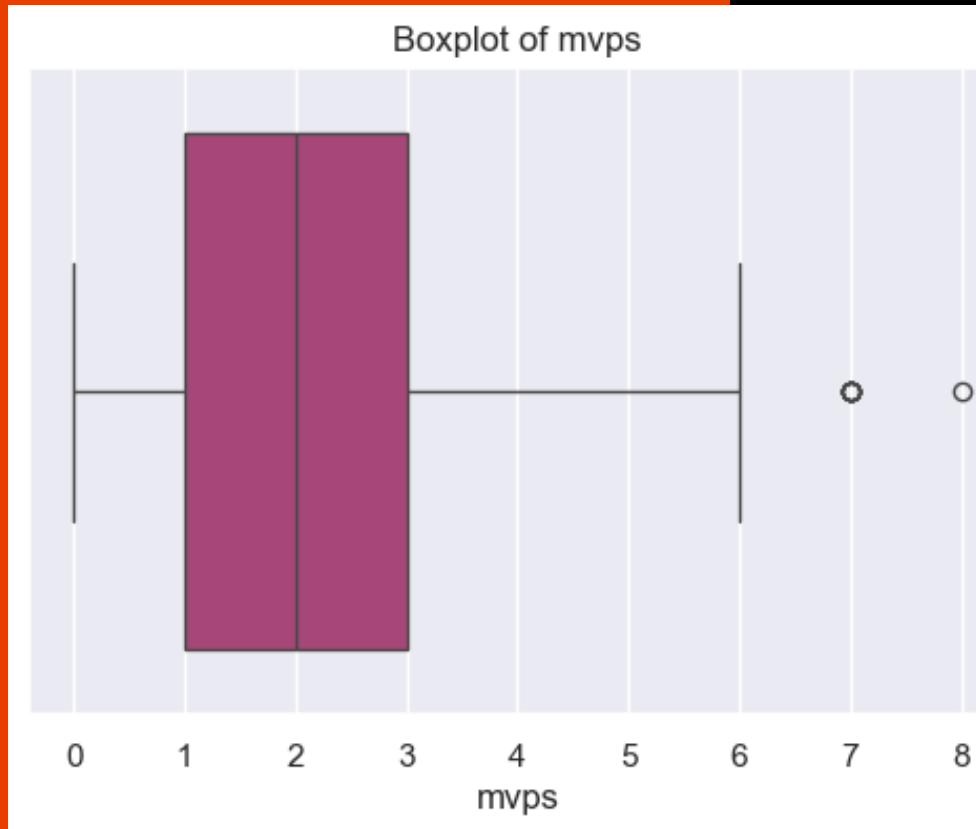
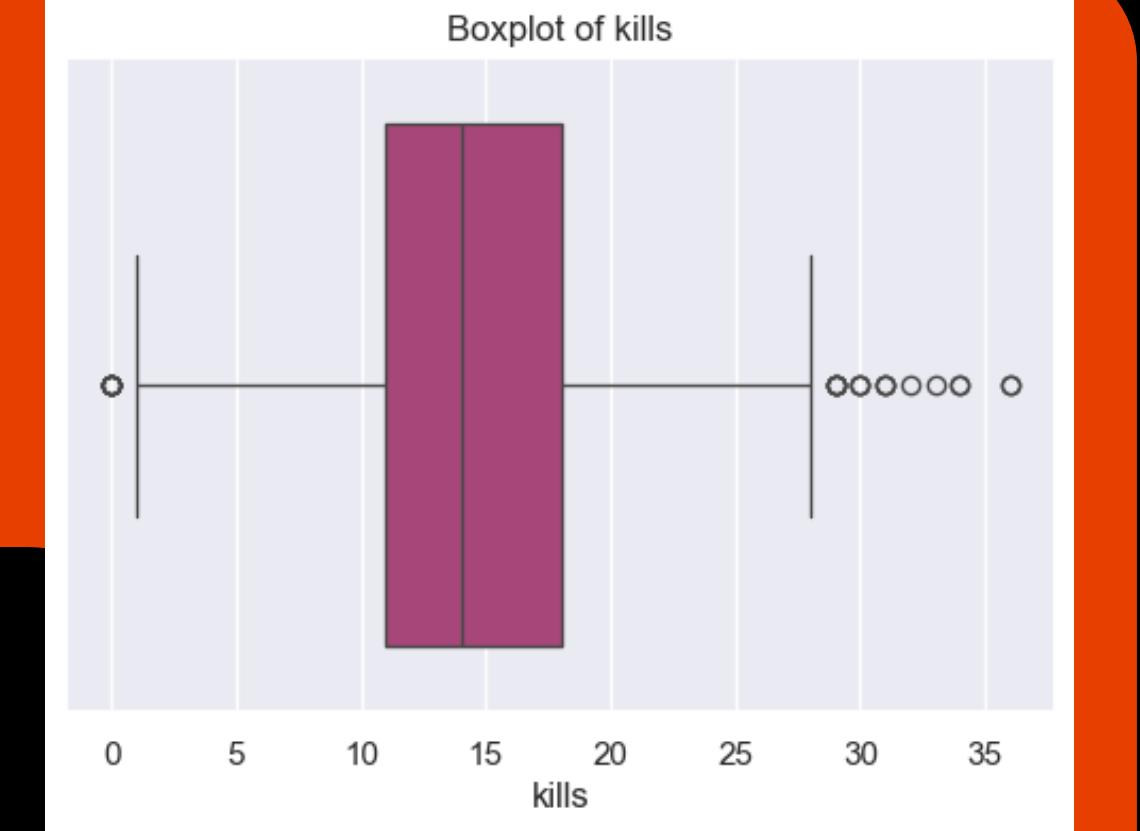
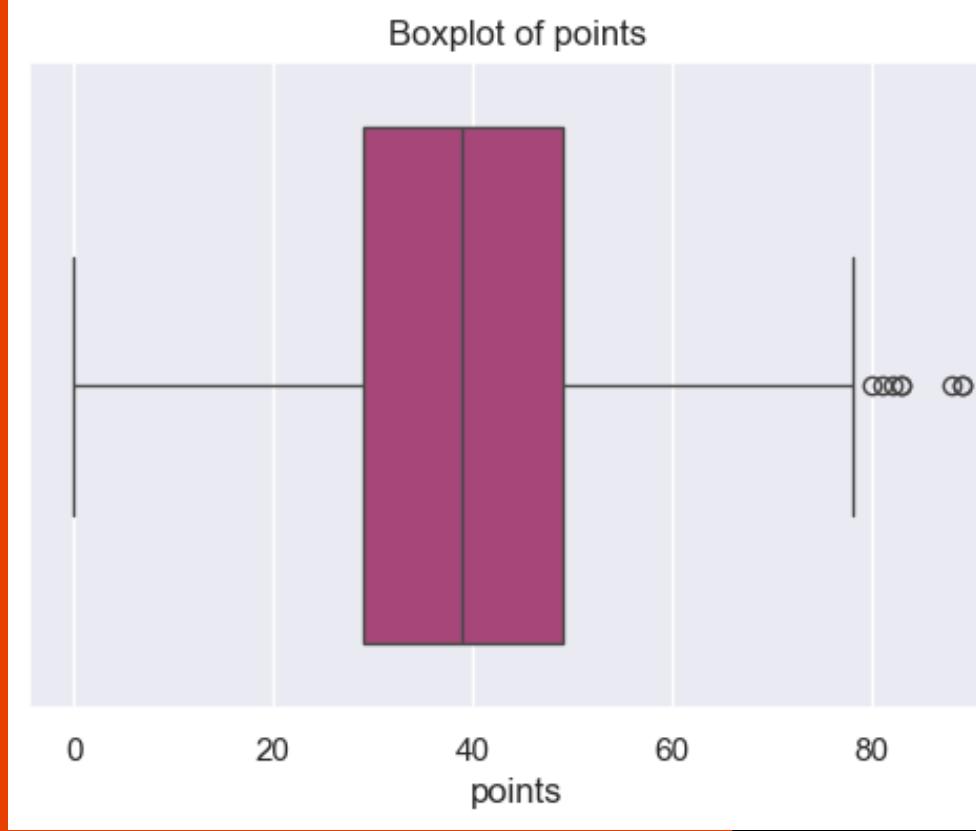
# Độ PHÂN BỐ DỮ LIỆU



- Phân bố gần chuẩn (normal-like), nhưng hơi lệch phải nhẹ (positive skew).**
- Đỉnh phân bố nằm quanh 13–17 kills.**
- Có đuôi phải kéo dài đến khoảng 38 kills**
- Không hoàn toàn đối xứng: lệch phải nhẹ.**
- Dài hơn đuôi trái → tồn tại những người chơi rất xuất sắc >30 kills là hiếm nhưng có tồn tại (~1–2%).**
- Vẫn có một số trận người chơi đạt 0–3 kills.**
- Đánh giá**
  - **Trình độ phổ biến: 12–22 kills → người chơi trung bình và khá.**
  - **Người chơi vượt trội: 30–38 kills → rất ít, có thể là pro hoặc trận kéo dài.**
  - **Giá trị cực tiểu (0 kills) → những người chơi yếu, AFK hoặc bị ép.**
  - **Vì kills phân bố gần chuẩn, lệch phải nhẹ → points cũng sẽ phân bố tương tự, nhưng giá trị cao hơn.**
  - **Các điểm cao (trên 70+) thường gắn với kills ≥ 25.**



# KIỂM TRA VÀ REMOVE OUTLIERS



Kết luận:

Dữ liệu ghi nhận nhiều outlier ở cả hai đầu phân phối của các biến points, kills, mvps và match time. Những giá trị ngoại lệ cao phản ánh người chơi có hiệu suất vượt trội hoặc các trận đấu kéo dài. Trong khi đó, giá trị quá thấp có thể do thoát trận hoặc đóng góp thấp. Việc xử lý các outlier này là cần thiết để giảm ảnh hưởng đến mô hình học máy và đảm bảo kết quả phân tích chính xác.

# LỰA CHỌN VÀ XÂY DỰNG MODEL



- Biến points là biến số thực liên tục → phù hợp với các mô hình hồi quy để mô hình hóa mối quan hệ tuyến tính giữa points và các biến như kills, mvps, match time...
- Qua phân tích biểu đồ phân tán và hệ số tương quan, các biến kills, mvps, và assists có tương quan dương mạnh với points.
- Nên sử dụng mô hình **LinearRegression** vì đã đáp ứng điều kiện cốt lõi là biến mục tiêu phải tuyến tính với biến đầu vào



# ĐÁNH GIÁ KẾT QUẢ

Model	justed R-Square	R-Squared	RMSE	Time Taken
LinearRegression	0.946099	0.948538	3.02823	0.00799894

Mean Absolute Error (MAE): 2.5206

Mean Squared Error (MSE): 9.1591

Mô hình Linear Regression:  $R^2 = 0.9485$  cho thấy khả năng mô hình dự đoán khá chuẩn. Các chỉ số lỗi thấp (RMSE = 3.03, MAE = 2.52) chứng minh mô hình có độ chính xác tốt và phù hợp với mục tiêu dự đoán points trong bộ dữ liệu CSGO



# KẾT LUẬN

- Hỗ trợ HLV phát hiện điểm mạnh, điểm yếu của từng game thủ để đưa ra trình lô trình huấn luyện.
- Tối ưu đội hình thi đấu dựa vào phân tích dữ liệu hiệu suất.
- Mô hình giúp dự đoán tiềm năng phát triển của người chơi.
- Đánh giá tuyển dụng game thủ dựa trên hiệu suất chứ không chỉ thành tích đơn lẻ.
- Đưa ra trình độ ghép trận phù hợp dựa trên kỹ năng chơi.
- Tạo sự cân bằng sức mạnh game



G A M E   P R E S E N T A T I O N   T E M P L A T E

THANK YOU

THANK YOU