



# **Estatísticas Agregadas**



# 1 Recrutamento Estatístico

# Ranqueando Jogadores

- Temos diversas estatísticas do jogador, como gols marcados, assistências, minutos jogados, distância percorrida e porcentagem de passes concluídos;
- Isoladamente, elas não nos dizem quase nada sobre as habilidades de um jogador de futebol.

 **PREMIER LEAGUE**  
Round 9


 **6 - 3** 



 **Phil Foden**  
 **RIGHT WINGER**



**9.7**  
RATING

Minutes played	75
Goals	3
Shots inside box	5
xG on target (xGOT)	1.76
Accurate passes	21/22 (95%)




 **PREMIER LEAGUE**  
Round 11

 **1 - 0** 

 **Alisson Becker**  
 **KEEPER**

**8.4**  
RATING

Chances created	1
Assists	1
Recoveries	13
Saves	6
Accurate passes	30/37 (81%)

# Ranqueando Jogadores

- Mesmo os números disponíveis nas plataformas de aferição são difíceis de contextualizar.

Match	Position	Minutes played	Total actions / successful	Goals	Assists	Shots / on target	xG	Passes / accurate	Long passes / accurate	Crosses / accurate	Dribbles / successful	Duels / won	Aerial duels / won
<div>AVERAGE / 90</div> <div>TOTAL</div>		90	75 70%	0.06	0.04	0.43 31.4%	0.05	46.11 80.2%	8.49 49.9%	0.17 46.4%	0.92 73.3%	20.41 64.4%	8.61 64.6%
Everton 4:2 Brighton & Hove Albion England, Premier League, 03.10.2020	LCB	96	89/66 74%	0	0	1/0 0%	0.01	69/58 84%	8/1 13%	0	2/1 50%	12/7 58%	2/1 50%
Brighton & Hove Albion 2:3 Manchester United England, Premier League, 26.09.2020	LCB	102	70/53 76%	0	0	1/0 0%	0.11	52/45 87%	4/1 25%	0	0	12/6 50%	4/1 25%

- Existem dois grandes problemas com esses números:
  - **Contexto:** não sabemos, por exemplo, se o jogador estava completando passes simples entre zagueiros na linha de fundo, ou passes de ataque mais difíceis.
  - **Comparação:** não podemos ver como esse jogador em particular se compara a outros da mesma liga.



“Se eu tiver que dar um bote,  
eu já cometi um erro”

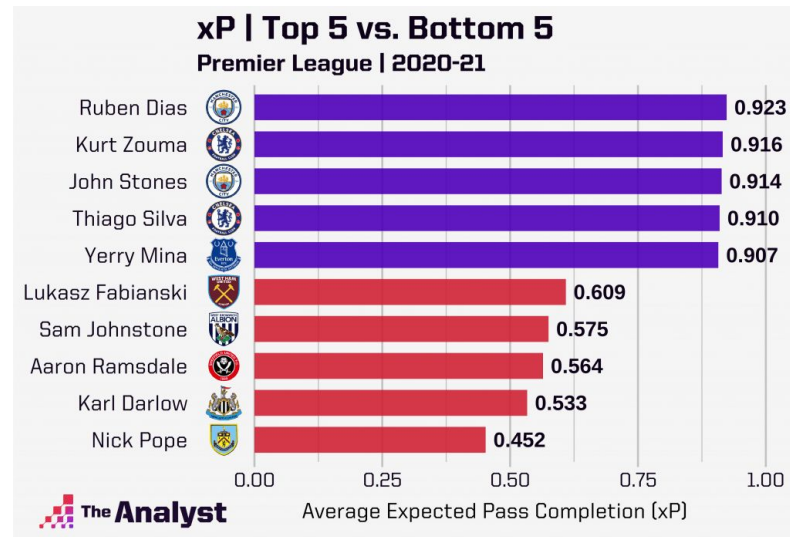
**Paolo Maldini**

# Contexto: o que medir

- Um bom zagueiro deve estar bem posicionado para:
  - Evitar que o adversário pegue a bola,
  - Para desencorajar um passe sendo feito,
  - ou pronto para interceptar um passe quando for feito.
- Maldini nos diz que o desarme é o último recurso de uma defesa ruim.
- Então, se contarmos os botes, estamos medindo uma defesa ruim!

# Contexto: o que medir

- A porcentagem de passes concluídos pode ser interpretada como uma habilidade de manter a bola.
- Mas também pode medir um estilo de posse de futebol onde são realizados passes simples.



# Contexto: o que medir

- A porcentagem de passes concluídos pode ser interpretada como uma habilidade de manter a bola. Mas também pode medir um estilo de posse de futebol onde são realizados passes simples.
- Um bom exemplo disso foi um estudo da Transfermarkt sobre as maiores taxas de conclusão de passes.
  - Eles descobriram que as maiores taxas de conclusão de passes nas cinco grandes ligas europeias eram de:
    - Thiago Silva (PSG), 95,5%
    - Presnel Kimpembe (PSG), 95,0%
  - Os dois jogaram juntos como zagueiros centrais em um time que dominava o campeonato.
  - Tudo o que essa estatística mostra é que eles estavam rolando para a frente e para trás um para o outro!



# Contexto: o que medir

- Mais para frente veremos o exemplo do gráfico de radar para atacantes.
- Nestes gráficos veremos:
  - Passes e recepções no terço final;
  - Passes inteligentes;
  - Passes importantes;

# Contexto: o que medir

- Mais para frente veremos o exemplo do gráfico de radar para atacantes.
- Nestes gráficos veremos:
  - Passes e recepções no terço final;
    - Adiciona contexto aos passes recebidos por atacantes: eles estão em áreas perigosas.
      - É muito mais difícil completar passes perto do gol adversário.
  - Passes inteligentes;
  - Passes importantes;

# Contexto: o que medir

- Mais para frente veremos o exemplo do gráfico de radar para atacantes.
- Nestes gráficos veremos:
  - Passes e recepções no terço final;
  - Passes inteligentes;
    - WyScout categoriza como: “Um passe criativo e penetrante que tenta quebrar as linhas defensivas do adversário para obter uma vantagem significativa no ataque”
  - Passes importantes;

# Contexto: o que medir

- Mais para frente veremos o exemplo do gráfico de radar para atacantes.
- Nestes gráficos veremos:
  - Passes e recepções no terço final;
  - Passes inteligentes;
  - Passes importantes;
    - “Criam imediatamente uma oportunidade clara de gol para um companheiro de equipe”

# Contexto: o que medir

- Mais para frente veremos o exemplo do gráfico de radar para atacantes.
- Nestes gráficos veremos:
  - Passes e recepções no terço final;
  - **Passes inteligentes;**
  - **Passes importantes;**
- **Ambos** são um tanto subjetivos - é a pessoa que coleta os dados que decide se um passe é inteligente ou chave - mas acrescentam um contexto importante.
- Juntas (com assistências), essas cinco estatísticas fornecem uma boa medida geral de quão bem um jogador pode passar a bola ao atacar.

# Comparação: percentis de classificações

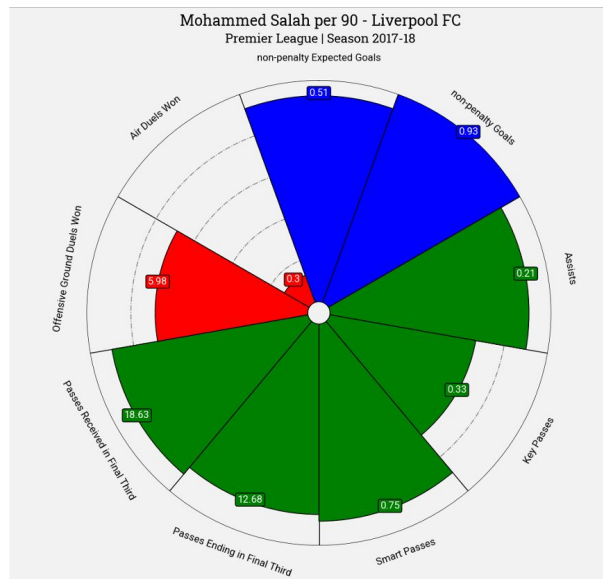
- O exemplo abaixo mostra como as contagens são convertidas em percentis de classificação.

Unsorted		Sorted		
Scores		Name	Rank	Percentile
David	31	Andrew	1	100
Andrew	45	William	2	87.5
George	38	Charles	3	75
Henry	41	Henry	4	62.5
Jack	37	John	5.5	43.75
John	38	<u>George</u>	5.5	43.75
Charles	42	Jack	7	25
William	43	David	8	12.5

- Classificamos os indivíduos de acordo com suas pontuações e, em seguida, atribuímos uma pontuação de percentil do mais alto (100%) ao mais baixo (0%).

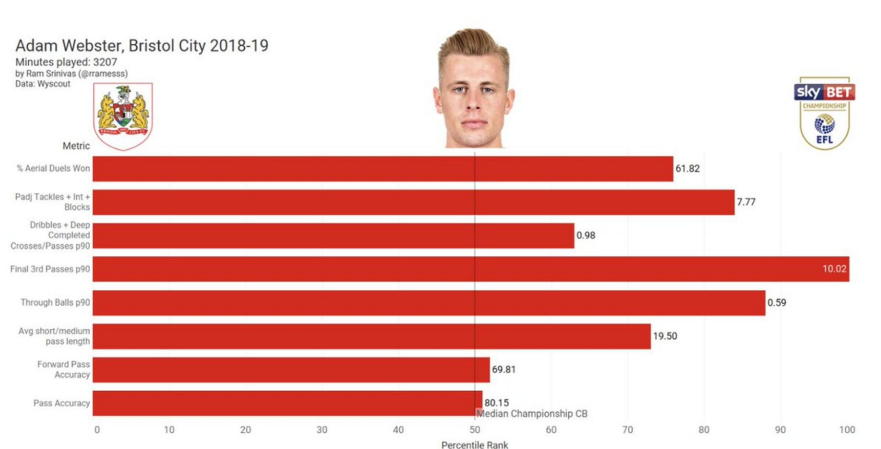
# Comparação: percentis de classificações

- Ao fazer esse ranking, é importante comparar com jogadores em **posição semelhante** na **mesma liga**.
- Esta classificação é mostrada abaixo para Mohammed Salah em forma de radar.



# Comparação: percentis de classificações

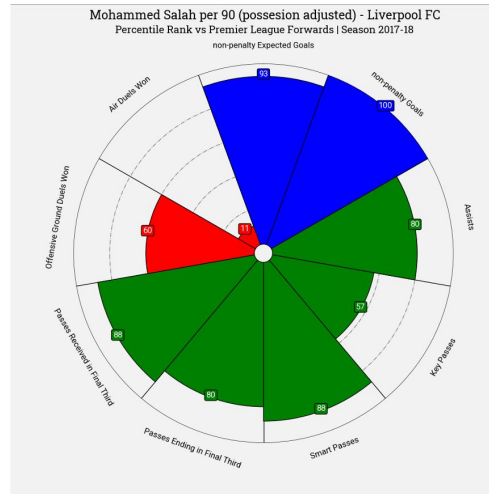
- Pode-se argumentar que os radares distorcem os dados, porque a área do radar não é proporcional ao percentil.
  - Jogadores nos percentis superiores preenchem muito mais o radar.
- Uma alternativa é fazer gráficos de barras, como no exemplo a seguir, criado para procurar zagueiros centrais.



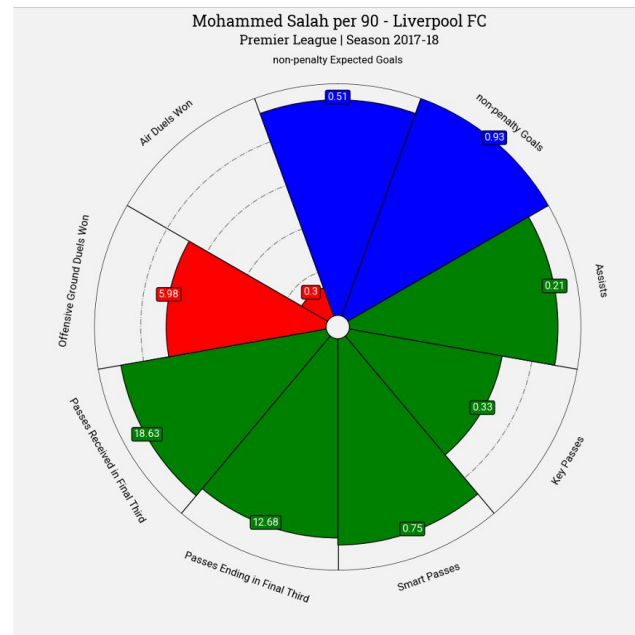
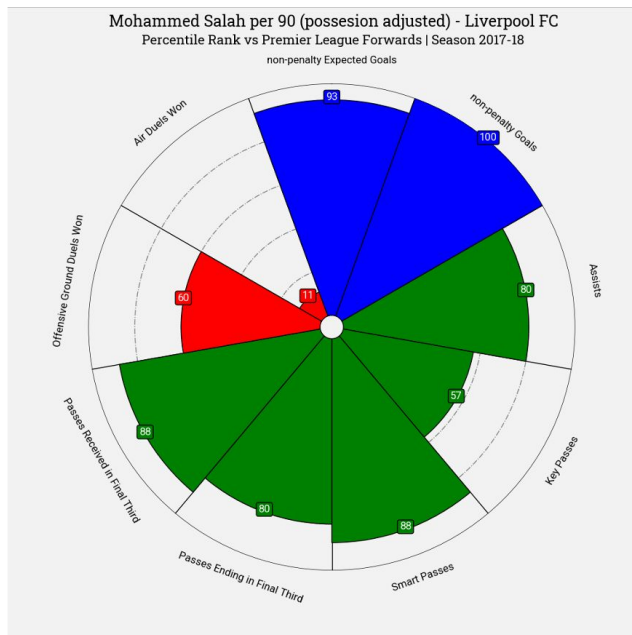


# Normalização

- Salah joga pelo Liverpool, que costuma ter mais posse de bola que o adversário.
  - Ele tem mais oportunidades de fazer passes e chutar do que um atacante jogando em um time que tem menos posse de bola.
- No radar a seguir, calculamos os minutos jogados para cada jogador e depois normalizamos pela posse do time: obtemos chutes por minuto jogados onde o time está com a posse de bola. Isso reduz ligeiramente a classificação de Salah.



# Comparando os radares

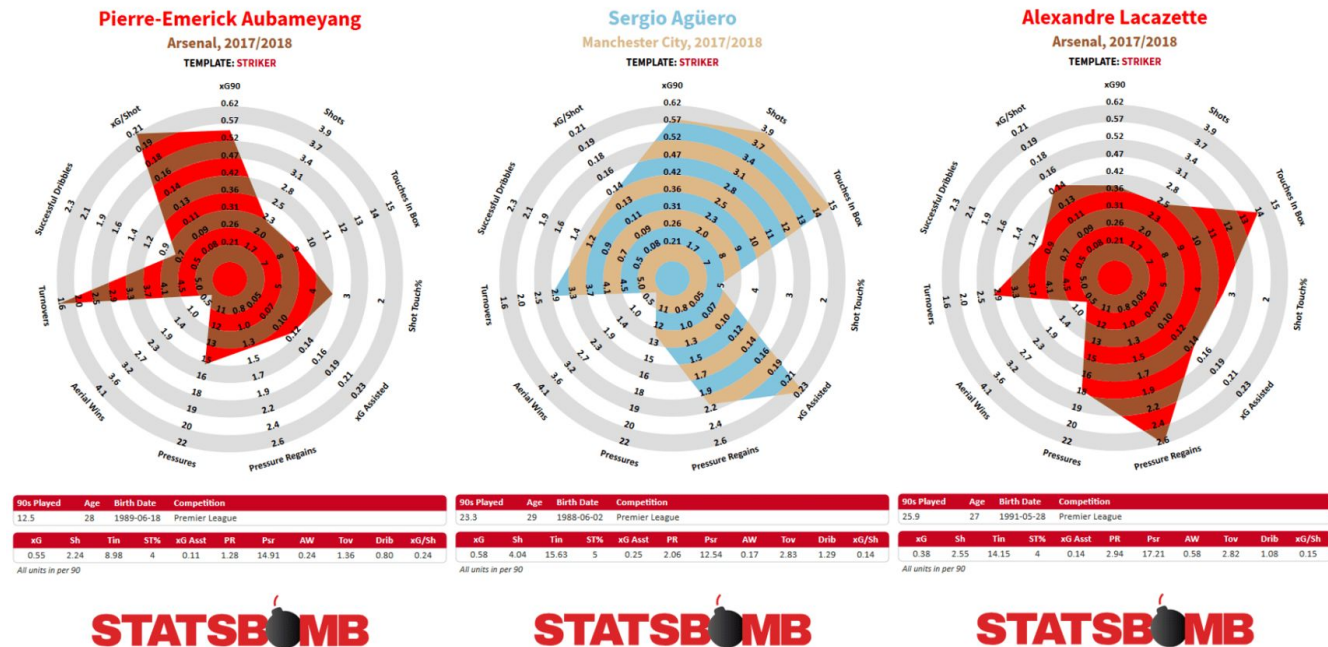


# Normalização

- Corrigir a posse de bola não é a única maneira de normalizar os dados do futebol.
- Métricas, como:
  - **Shot Touch:** A quantidade de chutes de um jogador em relação ao número de vezes que ele toca na bola.
  - **Pressure Regains:** Vezes que a equipe recuperou a bola 5 segundos após o jogador pressionar um adversário, por 90 minutos.
  - **Turnovers:** a taxa de perda de bola de um jogador devido a perda de controle ou drible mal sucedido a cada 90 minutos de jogo.
- são ajustadas não apenas para minutos jogados e posse de bola, mas também para o número de toques que um jogador dá.

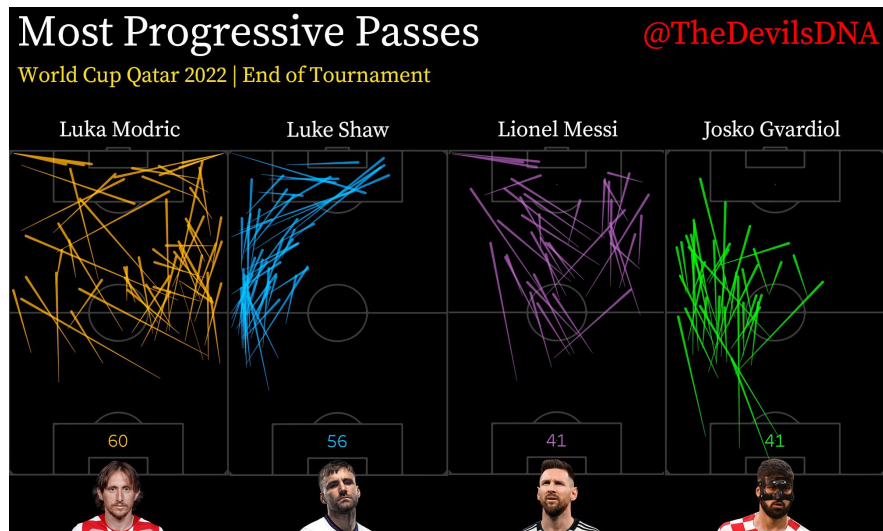
# Normalização

- Gráficos de radar de atacantes do Statsbomb.



# Normalização

- Mais sobre normalização de dados de futebol:
  - [Postagens originais](#) de Ted Knutson do Statsbomb sobre o assunto;
  - [Artigo](#) de Mark Carey.



# Modéstia

- Nesta seção, nos concentramos no contexto e na comparação.
- Restam dois desafios:
  - Um que podemos abordar com uma modelagem melhor;
    - Atribuir valores às ações;
    - Embora rótulos como passes inteligentes nos ajudem, ainda não temos uma estrutura para avaliar outras ações da mesma forma que valorizamos chutes com gols esperados.
  - Outro que não podemos.
    - Desafio de combinar análises com o tradicional levantamento qualitativo.
    - No [vídeo](#), há algumas dicas nesse sentido.
    - O melhor é a modéstia.
      - “This modesty is, in my experience, a characteristic of many of the best applied mathematicians and statisticians.”



# 2 Gráficos de Radar

# Gráficos de Radar

- Nesta parte, vamos realizar o passo a passo de fazer radares de jogador para um atacante.
- Calculamos as seguintes métricas diretamente de uma contagem de ações nos dados de evento Wyscout:
  - Gols sem ser de pênalti;
  - Assistências;
  - Passes-chave;
  - Passes inteligentes;
  - Duelos de aéreos vencidos;
  - Duelos de ataque (terrestres) vencidos.
- Acrescentamos a estes nossos próprios cálculos de,
  - xG sem pênalti;
  - Passes terminando no terço final do campo;
  - Recepções no terço final do campo.



# Gráficos de Radar

***Acho que toda essa parte deve ser prática***



# 3 Estudo de caso

Recrutamento de planilhas

# Jogador recrutado por dados

- Na temporada 2014-15, N'Golo Kanté liderou a tabela de desarmes bem-sucedidos por jogo em toda a Europa.
- Na época, ele jogava pelo Caen, da Ligue 1 francesa . Enviar um olheiro para checá-lo foi simples, e muitos clubes fizeram exatamente isso.
- Foi Steve Walsh, do Leicester City, quem mais insistiu com seu chefe sobre a contratação de Kanté.
- O Leicester o comprou por £ 5,6 milhões, venceu a Premier League e o vendeu para o Chelsea por £ 32 milhões no ano seguinte.
- A estatística de desarmes era simples, mas não mentiu.



# Recrutamento por planilhas

- Os olheiros têm grandes planilhas onde as colunas incluem desarmes feitos, interceptações, passes e dribles.
  - Eles classificam as planilhas pela coluna em que estão mais interessados.
  - A linha A mostra o melhor jogador, a linha B o segundo melhor e assim por diante.
  - É a partir daí que iniciam a busca.
- Quando Steve Walsh assumiu o cargo de diretor de futebol do Everton, ele assinou a Linha B em sua planilha - Idrissa Gueye, do Aston Villa.
  - Como Kanté, Gueye também veio da França para a Premiership em 2015, e durante sua primeira temporada ficou em segundo lugar (atrás apenas para Kanté) em desarmes e interceptações.

# Recrutamento por planilhas

- As equipes não dependem apenas de planilhas para comprar jogadores.
- Todos os olheiros profissionais concordam que é importante ver um jogador em ação antes de contratá-lo.
  - A filmagem das partidas é valiosa, mas não há substituto para sentar na beira do campo e observar como um jogador reage ao movimento do jogo ao seu redor e aos companheiros de equipe.
- Uma avaliação adequada envolve visitas a vários jogos e, se possível, a oportunidade de falar com o jogador e vê-lo treinar.
- A verdadeira questão para um clube é obter um bom equilíbrio entre a triagem estatística inicial, o uso de redes de observação, análises de vídeo, assistir partidas ao vivo e usar dados para checar a tomada de decisões.
  - A contabilização de tudo isso dificulta o trabalho de escolher jogadores.

# Recrutamento por planilhas

A diferença de resultado de Aston Villa e Leicester City durante a temporada 2015-16:

- Ambas as equipes seguiram os mesmos princípios de iniciar uma busca usando estatísticas.
  - Ambos encontraram jogadores subvalorizados na França durante 2014–15.
    - O Leicester contratou Kanté e Riyad Mahrez e venceu a Premier League pela primeira vez em sua história.
    - Villa contratou Gueye e três outros jogadores do campeonato francês, mas foi rebaixado com o menor total de pontos de sua história.
- A maior diferença entre Leicester e Villa parece estar na confiança mútua entre a comissão técnica e os analistas.
  - O Leicester encontrou uma maneira de integrar as estatísticas em todos os aspectos de suas operações;
  - A equipe de recrutamento do Aston Villa e o gerente Tim Sherwood não chegaram a um acordo sobre como as estatísticas deveriam ser usadas.

# Recrutamento por planilhas

- Rory Campbell, olheiro técnico e analista do West Ham, enfatizou a necessidade de uma estratégia baseada em análises em todo o clube.
- Essa estratégia deve combinar todos os aspectos da avaliação do jogador: da análise estatística, passando pela compreensão da personalidade e atitude, até ver os pontos fortes e fracos dos jogadores em campo.
- A análise estatística é apenas uma parte do processo de avaliação de jogadores. Combinando as análises estatísticas com outros aspectos, é possível ter uma visão mais completa do jogador e tomar decisões mais bem informadas.
- Ter uma estratégia baseada em análises pode ajudar um clube a identificar jogadores promissores, otimizar a formação da equipe e melhorar o desempenho geral.



# 4 Outras abordagens

Outras abordagens de recrutamento



# Modelos Plus/Minus

- Uma forma bastante objetiva de avaliar jogadores, sem examinar minuciosamente suas ações, consiste em observar a taxa de gols marcados ou sofridos por sua equipe quando estão em campo e quando não estão.
- Essa é a ideia básica dos modelos Plus/Minus.

# Modelos Plus/Minus

# Avaliando ações

- Índice de Desempenho da Premier League:
  - Trabalho de Ian McHale e Philip Scarf
  - Analisaram, em detalhes, o que os jogadores contribuíram em um jogo.
  - Começaram construindo um modelo estatístico de como diferentes contribuições criam gols.
    - Em cada partida, eles contaram com que frequência os jogadores realizavam certas ações (passes, desarmes, cruzamentos, dribles, bloqueios, etc) e com que frequência recebiam cartões amarelos ou vermelhos.
    - Em seguida, eles empregaram um método estatístico de ajuste para avaliar o quão preciso o número dessas ações foi em prever o próprio número de chutes a gol ou o número de chutes a gol do time adversário.
      - Ajuste estatístico fornece uma medida do efeito, positivo ou negativo, na criação de chances.

# Avaliando ações

- Índice de Desempenho da Premier League:
  - Trabalho de Ian McHale e Philip Scarf
  - Analisaram, em detalhes, o que os jogadores contribuíram em um jogo.
  - Começaram construindo um modelo estatístico de como diferentes contribuições criam gols.
  - Descobriram que quanto mais um time passa a bola, mais chances cria.
    - Cruzamentos bem-sucedidos provaram-se uma das rotas mais prováveis para o gol.
  - Ao ajustar seu modelo aos dados da Premier League de 2003 a 2006, eles estimaram:
    - Cada cruzamento bem-sucedido valia cerca de 10 passes padrão em termos de criação de chutes.
    - Ser expulso equivalia a perder 41 intercepções.

# Índice de Desempenho da Prem League

- A maneira como eles fizeram isso foi realizar uma regressão linear sobre como o número de ações específicas impactava o número de gols.
- Abaixo está a saída de regressão para o modelo

Variables	Coef.	Std. err.	t-stat	p-value
Crosses	0.519	0.069	7.490	0.000
Dribbles	0.118	0.026	4.600	0.000
Passes	0.034	0.004	7.830	0.000
Opposition interceptions	-0.024	0.009	-2.520	0.012
Opposition yellows	0.253	0.134	1.890	0.059
Opposition reds	1.023	0.506	2.020	0.043
Opposition tackle win ratio	-0.170	0.090	-1.890	0.059
Opposition cleared	-0.017	0.009	-1.920	0.055
Constant	6.463	0.815	7.930	0.000

# Índice de Desempenho da Prem League

- Os coeficientes de regressão mostram a importância relativa de diferentes fatores na determinação da pontuação.
- Os jogadores são então recompensados por realizar ações, ponderadas com esses coeficientes.

Variables	Coef.	Std. err.	t-stat	p-value
Crosses	0.519	0.069	7.490	0.000
Dribbles	0.118	0.026	4.600	0.000
Passes	0.034	0.004	7.830	0.000
Opposition interceptions	-0.024	0.009	-2.520	0.012
Opposition yellows	0.253	0.134	1.890	0.059
Opposition reds	1.023	0.506	2.020	0.043
Opposition tackle win ratio	-0.170	0.090	-1.890	0.059
Opposition cleared	-0.017	0.009	-1.920	0.055
Constant	6.463	0.815	7.930	0.000






# Índice de Desempenho da Prem League

- Seu modelo estatístico permite quantificar o efeito da construção de um ataque que vem da defesa e do meio-campo.
- Ian e Phil adicionaram a probabilidade de chutes e defesas bem-sucedidos, o que lhes permitiu levar em consideração a influência de atacantes e goleiros na criação e supressão de gols.
- O método é o mais "objetivo" possível, no sentido de que se baseia inteiramente no número de várias ações realizadas em campo.
  - Ele pega as ações de cada jogador, relaciona-as com as chances de gol e descobre quais jogadores contribuem mais.

# Índice de Desempenho da Prem League

- Eles o testaram na temporada 2008/09 para determinar os melhores jogadores.
- Eles observaram quantos dos vários tipos de ação cada jogador realizou durante a temporada e calcularam o índice para cada um deles.



Clube	Pts	PJ	VIT	E	DER	GM	GC	SG
1  Manchester U...	90	38	28	6	4	68	24	44
2  Liverpool	86	38	25	11	2	77	27	50
3  Chelsea	83	38	25	8	5	68	24	44
4  Arsenal	72	38	20	12	6	68	37	31
5  Everton	63	38	17	12	9	55	37	18



# Índice de Desempenho da Prem League

- O melhor jogador segundo o modelo foi o goleiro do Fulham, Mark Schwarzer (com 7,29)
  - Ronaldo nem chegou ao top 20, nem nenhum jogador do Manchester United.
- Na verdade, a lista trazia apenas um atacante, Nicolas Anelka, do Chelsea, e um meio-campista, Gareth Barry. Todos os outros eram goleiros e zagueiros.
- Em seu artigo científico sobre o sistema, Ian e Phil defenderam seu modelo de contribuição de correspondência.
- Quando eles investigaram por que os atacantes e meio-campistas ofensivos não estavam em posições mais altas no ranking, descobriram que era porque estavam desperdiçando chances na frente do gol.
  - Frank Lampard, em particular, foi considerado "altamente variável" com seus chutes.
- Os defensores foram classificados como mais valiosos porque eram suas intercepções e bloqueios que impediam os gols.
- Os defensores evitaram muito mais gols em potencial do que os atacantes marcaram, tornando sua contribuição na partida mais significativa.

# Índice de Desempenho da Prem League

- Mais de dez anos depois, esses métodos são limitados e foram substituídos por outros.
- É um ponto de partida útil para métodos mais recentes.

Name	Team	Position	Subindex 1 points
Mark Schwarzer	Fulham	Goalkeeper	7.29
Gareth Barry	Aston Villa	Midfielder	7.06
Sol Campbell	Portsmouth	Defender	6.86
Gary Cahill	Bolton Wanderers	Defender	6.70
J. Lloyd Samuel	Bolton Wanderers	Defender	6.63
Robert Green	West Ham United	Goalkeeper	6.45
Heurelho Gomes	Tottenham Hotspur	Goalkeeper	6.43
David James	Portsmouth	Goalkeeper	6.40
Scott Carson	West Bromwich Albion	Goalkeeper	6.39
Nicolas Anelka	Chelsea	Striker	6.33
Richard Dunne	Manchester City	Defender	6.30
Andrew O'Brien	Bolton Wanderers	Defender	6.26
Jussi Jaaskelainen	Bolton Wanderers	Goalkeeper	6.20
Sylvain Distin	Portsmouth	Defender	6.12
Thomas Sorensen	Stoke City	Goalkeeper	6.10
Matthew Upson	West Ham United	Defender	6.03
Ryan Shawcross	Stoke City	Defender	5.85
Danny Collins	Sunderland	Defender	5.81
Michael Turner	Hull City	Defender	5.54
Anton Ferdinand	Sunderland	Defender	5.47

# Referências

- [Statistical scouting — Soccermatics documentation](#)
- [Statistical models — Soccermatics documentation](#)
- [PlusMinus - Soccermatics documentation](#)
- [Models for evaluating players part 1: Plus/minus and EA player ratings](#)
- D. Sumpter, Soccermatics: Mathematical Adventures in the Beautiful Game. Bloomsbury Publishing Plc, 2016.

# Obrigado!



luiza.chagas@dcc.ufmg.br



hugoriosneto@dcc.ufmg.br



adrianoc@dcc.ufmg.br



meira@dcc.ufmg.br