



footure  
academy

ESPECIALIZAÇÃO

CIÊNCIA NOVO  
DE DADOS

NO FUTEBOL

# Aula 12 - Como analisar gráficos

Caio Batatinha

## AGENDA

- Motivação
- Boas práticas de visualização de dados
- Galeria com tipos de gráficos

## OBJETIVOS DE APRENDIZAGEM

- Entender os diferentes de gráfico, como e quando utilizar.

## MOTIVAÇÃO

Como revisado na última aula, gráficos e técnicas de visualização de dados são ferramentas mais efetivas na comunicação da análise de dados. Gráficos conseguem traduzir a complexidade dos dados crus em informações valiosas para tomadores de decisão. A missão de identificar padrões, relacionamentos e tendências é facilitada pelos elementos visuais.

## BOAS PRÁTICAS DE VISUALIZAÇÃO DE DADOS

Ao trabalhar com gráficos e visualizações de dados, alguns tópicos devem ser levados em consideração para auxiliar no tipo de seleção de ferramenta gráfica e tradução desses dados em uma informação clara.

1. Propósito
2. Audiência
3. Tipo de dado

## PROPÓSITO

O primeiro passo a ser tomado no trabalho com gráficos é pensar na seguinte pergunta:

“ Qual a pergunta que essa visualização busca solucionar?”

Esse questionamento inicial poupa o tempo de selecionar dados e gráficos que no fim não vão ser capazes de trazer uma informação. Após realizar suas análises coletar informações, eleger quais histórias a destacar é a tarefa que norteia o escopo do trabalho.

## AUDIÊNCIA

A função do gráfico é comunicar informações e histórias observadas na análise dos dados. Em qualquer negócio, pensar na audiência é um passo chave na confecção dos visuais:

- É uma audiência técnica ou não técnica?
- Seleção de linguagem e que se adapte ao contexto.
- Forma de apresentação, uma entrega estática como um relatório a ser analisado periodicamente, ou uma apresentação oral?
- Menos é mais. Combinar vários elementos podem ser prejudiciais e acabar desviando a mensagem principal do gráfico.

## TIPO DE DADO

TIPO DE DADO	NATUREZA	GRÁFICO
Categórico	Discreta (Qualitativo)	Gráfico de Barras, Tabelas
Numérico (Valor estático)	Contínua	Texto simples, Tabelas
Numérico (Ao longo do tempo)	Temporal	Gráfico de linha, Gráfico de inclinação
Numérico (Distribuição)	Contínua	Histograma, Box Plot
Relacional (Duas variáveis)	Quantitativa	Gráfico de dispersão (Scatterplot)
Espacial	Relaciona/Contínua	Campogramas
Proporção	Discreta (Quantitativa)	Gráficos empilhados, Waffle, Pizza

# **TIPOS DE GRÁFICOS**

## TEXTO SIMPLES

### Função:

Comunicar diretamente informações relevantes.

### Quando utilizar:

Quando um ou mais números são suficientes para resumir a história principal.

### Exemplo:

Ao lado compreendemos de maneira rápida a idade de Pelé e algumas métricas na copa de 70 como partidas jogadas, gols e assistências.

### Brasil 70

by Leo M. Sa Freire

[Back to main page](#)

### Pelé



Age	29	Matches	6
xG	4.33	Goals	4
xA	5.03	Assists	5
Pass xT	1.76	Carry xT	0.73

## TABELA

### Função:

Disponibiliza diferentes números e informações de uma forma organizada.

### Quando utilizar:

Quando o propósito é comparar diferentes jogadores, times, jogos em diferentes métricas. Elementos visuais como negrito, cores diferentes são importantes para direcionar informações à um público.

**EVITE** utilizar tabelas em apresentações orais, por conter vários elementos dispersa facilmente a audiência.

### Exemplo:

A tabela ao lado disponibiliza todos os jogos que ocorreram na segunda rodada da Bundesliga 22/23. Elementos como o mapa de calor direciona a audiência a perceber times mais efetivos nas finalizações.

### Bundesliga Round 2 Summary

Viz by @sonofacorner | Season 2022/2023

MATCH	XG		XGOT		SHOTS		SHOT ACCURACY	
	H	A	H	A	H	A	H	A
 Freiburg (1) vs. Borussia Dortmund (3) 	1.0 (-0.0)	1.1 (+1.9)	1.4 (-0.4)	1.2 (+1.8)	16	13	44%	46%
 RB Leipzig (2) vs. 1. FC Köln (2) 	1.0 (+0.0)	1.5 (+0.5)	1.8 (+0.2)	2.0 (-0.0)	9	13	56%	31%
 Hertha BSC (1) vs. Eintracht Frankfurt (1) 	1.7 (-0.7)	1.6 (-0.6)	1.0 (+0.0)	0.6 (+0.4)	12	11	42%	9%
 Werder Bremen (2) vs. VfB Stuttgart (2) 	0.5 (+1.5)	1.4 (+0.6)	1.1 (+0.9)	2.4 (-0.4)	5	14	60%	50%
 Bayer Leverkusen (1) vs. Augsburg (2) 	2.6 (+1.6)	0.4 (+1.6)	2.4 (+1.4)	1.4 (+0.6)	24	6	42%	67%
 TSG Hoffenheim (3) vs. Bochum (2) 	2.6 (+0.4)	1.5 (+0.5)	3.5 (-0.5)	2.5 (-0.5)	23	11	52%	45%
 Schalke 04 (2) vs. Borussia M'Gladbach (2) 	1.4 (+0.6)	1.3 (+0.7)	1.8 (+0.2)	1.6 (+0.4)	12	14	33%	57%
 Mainz 05 (0) vs. Union Berlin (0) 	0.6 (-0.6)	0.4 (-0.4)	0.5 (-0.5)	0.5 (-0.5)	11	10	45%	20%
 Bayern München (2) vs. Wolfsburg (0) 	2.1 (-0.1)	0.4 (-0.4)	1.7 (+0.3)	0.7 (-0.7)	24	8	33%	25%

## GRÁFICO DE BARRAS

### Função:

Versátil na comparação de diferentes dados categóricos. Divide-se em alguns tipos: Vertical, horizontal, empilhado e 100% empilhado.

### Quando utilizar:

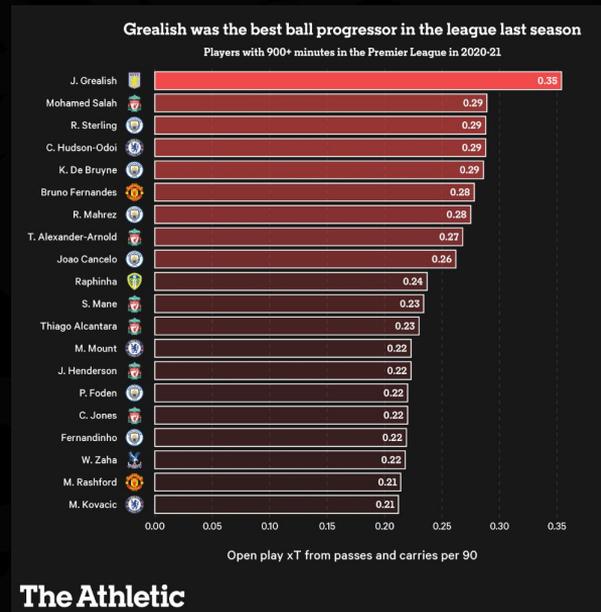
Vertical: Comparar diferentes magnitudes de forma lateral.

Horizontal: Recomendando utilizar quando os nomes das categorias são longos.

**CUIDADO** ao utilizar mais de um gráfico de barra na mesma tela, a escala é importante e pode causar falsas impressões.

### Exemplo:

O gráfico de barras ao lado horizontal e como as categorias são nome de jogadores, a orientação se torna mais agradável na horizontal.



## GRÁFICO DE BARRAS EMPILHADO

### Função:

Versátil na comparação de diferentes dados categóricos. Divide-se em alguns tipos: Vertical, horizontal, empilhado e 100% empilhado.

### Quando utilizar:

Empilhado: Comparar diferentes proporções que compõem o mesmo dado

**CUIDADO** com a cor que foi utilizado para não sobrepor.

### Exemplo:

O gráfico ao lado traz a informação sobre o xT, mas diferenciando quanto é proveniente de passe e quanto de drible.



## GRÁFICO DE LINHA

### Função:

Mostrar dados ou tendências que variam de forma contínua.

### Quando utilizar:

Dados de natureza temporal como uma média móvel ao longo da competição ou o xG acumulado ao longo de uma partida.

### Exemplo:

O gráfico ao lado traz a evolução de pontos reais e pontos esperados ao longo das rodadas de uma temporada europeia. A conexão entre os pontos evoca ideia de continuidade.

### Which Teams are Failing to Meet Expectations?

Expected points (in probabilistic terms) vs. actual points | Probabilities taken from FiveThirtyEight | viz by @sonofacorner



## GRÁFICO DE INCLINAÇÃO

### Função:

Enfatizar a diferença de estado entre dois pontos.

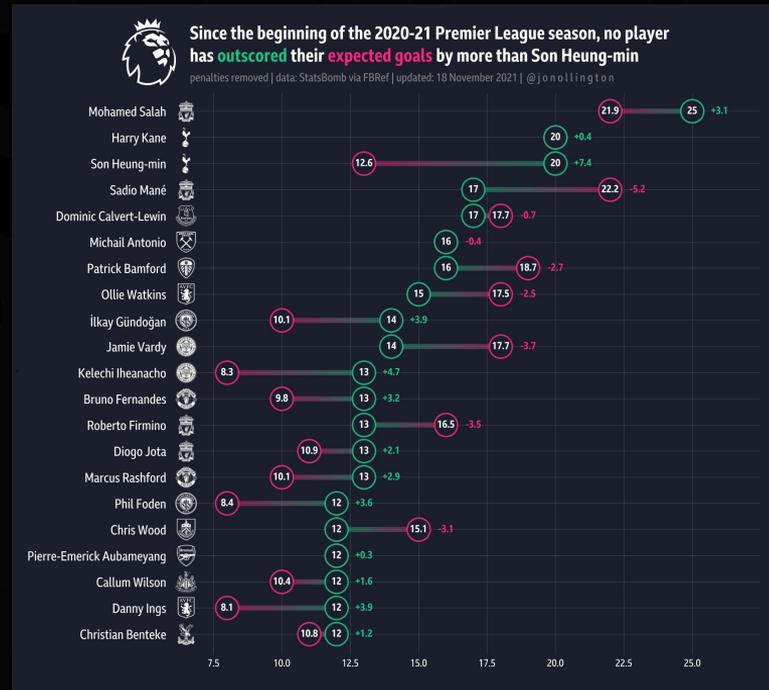
### Quando utilizar:

Comparações de natureza binária, como diferença de pontos na temporada, ou dados de fornecedores diferentes.

### Exemplo:

O gráfico ao lado traz a comparação entre gols marcados e expected goals dos artilheiros da Premier League 20-21. Fica explícito a diferença de gols esperados e gols reais marcados por Son.

## COMO ANALISAR GRÁFICOS



## GRÁFICO DE DISPERSÃO (SCATTER)

### Função:

Mostrar o relacionamento entre duas variáveis.

### Quando utilizar:

Excelente para identificar correlação entre variáveis, possíveis tendências e outliers.

### Exemplo:

O gráfico ao lado traz o xT por passes e xT drible para entender quais jogadores criam mais via apenas dribles, outros via passes e quais se complementam. Fica notório a identificação de outliers como Messi e Neymar que adicionam valor em ambas métricas.

### xT pass vs. xT dribble

From 2017/18 season | Top 5 leagues | Minimum of 3600 minutes played



## GRÁFICO WAFFLE

### Função:

Mostrar proporções diferentes em termos de área, associando a cor a proporção.

### Quando utilizar:

Mostrar diferentes jogadores/times e como a proporção entre eles se altera para o mesmo tipo de dado. Exemplo, mostrar com quais partes do corpo um jogador marcou numa competição: Cabeça, pé dominante, pé fraco e outros.

### Exemplo:

O gráfico ao lado traz a comparação entre os criadores na Europa e como eles geram os passes chaves, via escanteio, cruzamentos, faltas indiretas e jogadas de open play.



## GRÁFICO RADAR

### Função:

Comparar diferentes métricas de forma simultânea.

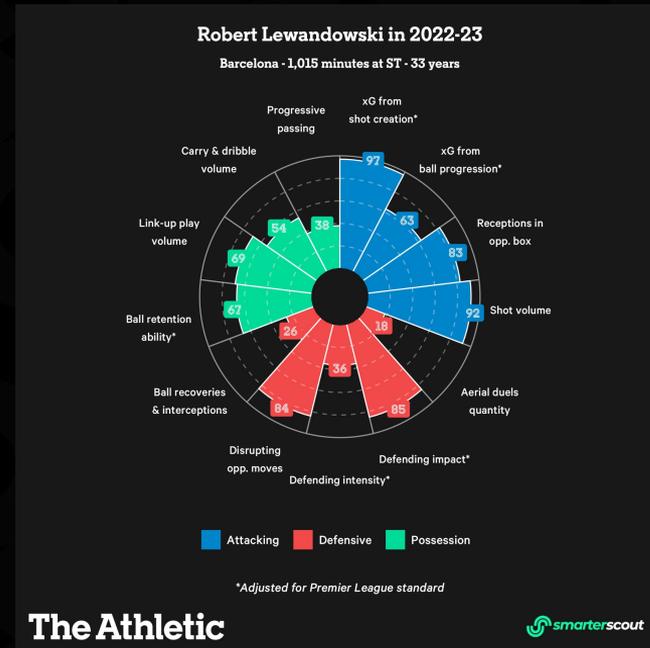
### Quando utilizar:

Destacar as valências de um jogador ou time de forma visual, excelente ao apresentar em formas de percentil cada métrica.

### Exemplo:

O gráfico ao lado traz o radar de Lewandowski em 2022-2023, nota-se de imediato o seu volume de finalizações e xG a partir de chutes, recepções na área adversária, movimentos de desmarque e impacto defensivo.

## COMO ANALISAR GRÁFICOS



## CAMPOGRAMAS

**Função:**

Exibir de maneira oficial comportamento observado no campo.

**Quando utilizar:**

Para dados espaciais, (x,y,z), a melhor forma de traduzir os dados é através de campoogramas. A ideia de movimentação e ações chave do jogador fica mais explícito.

**Exemplo:**

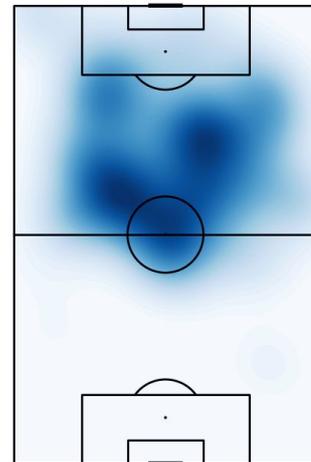
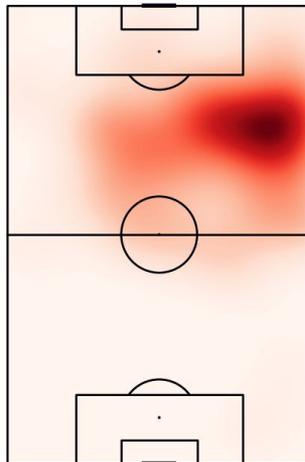
O mapa de calor ao lado destaca como a área de atuação de Messi se transforma a partir de uma mudança de função.

## COMO ANALISAR GRÁFICOS



### The Evolution of Lionel Messi

Actions in the match **before** and  
**after** becoming a False-9



## SIMPLE TEXT



Just because you have numbers doesn't mean you need a graph!

## TABLE

What is the main point I want to make?

OFTEN THERE ARE MORE EFFICIENT WAYS

	A	B	C
CATEGORY 1	25	125	500
CATEGORY 2	400	342	300
CATEGORY 3	35	174	100
CATEGORY 4	300	294	180

Avoid using tables in live presentations because people stop listening & start reading

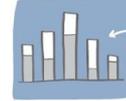
## BAR CHARTS

Great for categorical data

Easy for our eyes - comparing heights to a consistent baseline

Rule: Must have a zero baseline. No exceptions!

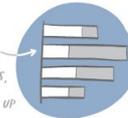
### VERTICAL



### HORIZONTAL



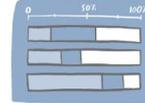
Good when category names are long



### STACKED

OFTEN MISUSED... EASIER TO COMPARE TOTAL & FIRST SERIES, BUT SEGMENTS UP THE STACK DON'T LINE UP

### 100% STACKED



TWO BASELINES FOR COMPARISON

## HEAT MAP

	A	B	C
CATEGORY 1	25	125	500
CATEGORY 2	400	342	300
CATEGORY 3	35	174	100
CATEGORY 4	300	294	180

EYES CAN EASILY PICK OUT BIG DIFFERENCES IN COLOR INTENSITY, but smaller ones don't stand out

Can work well when beginning to explore data and deciding where to dig further

## SCATTER PLOT



Good for encoding data simultaneously on two axes to identify what relationships exist

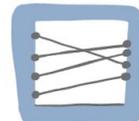
## LINE



Rule: The lines that connect the dots have to make sense! Most effective with continuous data, often time

## SLOPE GRAPH

A FANCY WORD FOR A LINE GRAPH WITH ONLY 2 POINTS



Useful to focus on change between two points in time or difference between groups

## WATERFALL

ADDITIONS & DEDUCTIONS



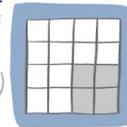
BEGINNING QUANTITY

ENDING QUANTITY

Often used in finance to show variance to budget

## SQUARE AREA

(AKA WAFFLE CHART)



THE GRID IS IMPORTANT BECAUSE WE TEND TO OVERESTIMATE AREAS

Good for showing numbers of very different magnitudes, or as an alternative to a pie chart

**Dúvidas?**

## REFERÊNCIAS UTILIZADAS

<https://x.com/petermckeeever>

<https://www.storytellingwithdata.com/blog/2021/4/6/how-do-i-know-which-graph-to-use>

<https://x.com/TonyElHabr>

<https://soccerment.com/expected-threat/>

<https://brasil70.streamlit.app/>

[https://github.com/sonofacorner/soc-viz-of-the-week/blob/main/10032022/figures/10032022\\_underperformers.png](https://github.com/sonofacorner/soc-viz-of-the-week/blob/main/10032022/figures/10032022_underperformers.png)

**OBRIGADO!**



@CcBatatinha



[linkedin.com/in/caio-batatinha/](https://www.linkedin.com/in/caio-batatinha/)



[caio.batatinha@footure.com.br](mailto:caio.batatinha@footure.com.br)  
[caiocbatatinha@gmail.com](mailto:caiocbatatinha@gmail.com)