



Data Literacy & Story Telling in Oracle

Abi Giles-Haigh

Innovation and Analytics Director



- PhD in ML – Computation Modelling of the human heart
- Oracle DBA for 5 years
- Worked public health care, private banking





Tech Superstars Unite

Get worldwide recognition as an Oracle ACE



Oracle.com profile page



Swag, certification exam credit & event passes



Exclusive content



Networking events



Your own Oracle cloud account



Travel support

✕ @oracleace

in [Linkedin.com/groups/72183](https://www.linkedin.com/groups/72183)

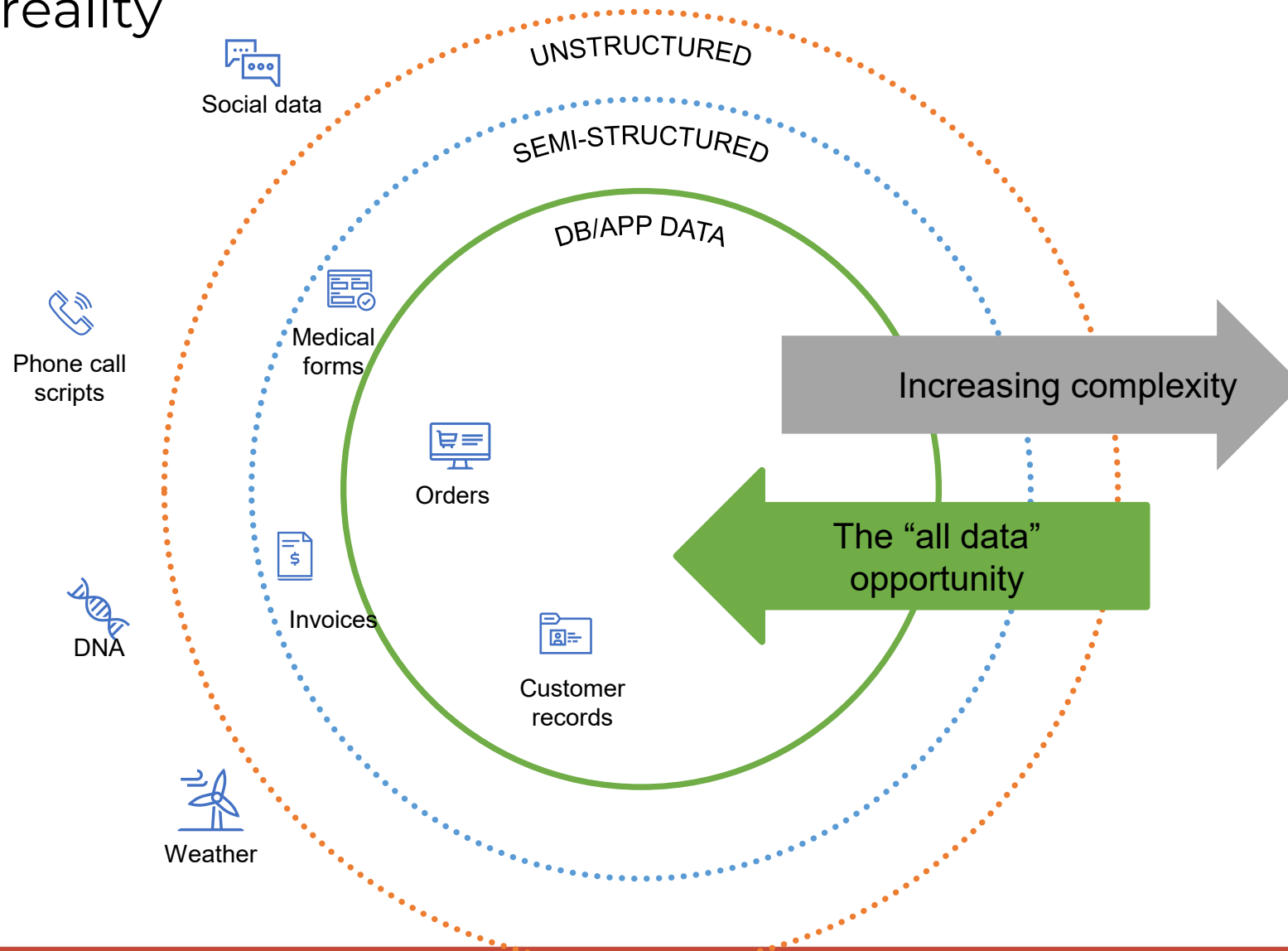
🦋 @oracleace.bsky.social

Learn more at:
ace.oracle.com





The business reality





Why we're here

42.2% annual growth (over two years) for enterprise data. Seagate: Rethink Data 2024

- Organisations collect massive amounts of data
- Dashboards don't always equal understanding
- People are overwhelmed, not empowered
- Insights aren't reaching decision makers

*77% of business leaders say the dashboards/charts they get do **not** always relate directly to the decisions they need to make*

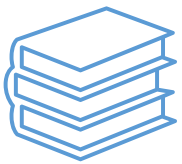
https://www.prnewswire.com/news-releases/global-study-70-of-business-leaders-would-prefer-a-robot-to-make-their-decisions-301799591.html?utm_source=chatgpt.com



What is data literacy?

“the ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case, application and resulting value.”

- Gartner



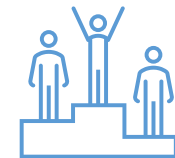
Read...



Analyse...



Communicate...



Results!

Data-Driven Culture vs. Data-Collected Culture

- Most organizations **collect data**
- Few organizations **change behaviour** because of data
- Reports $\neq$ decisions
- Insights only matter when they drive action

DATA



SORTED



ARRANGED



PRESENTED
VISUALLY

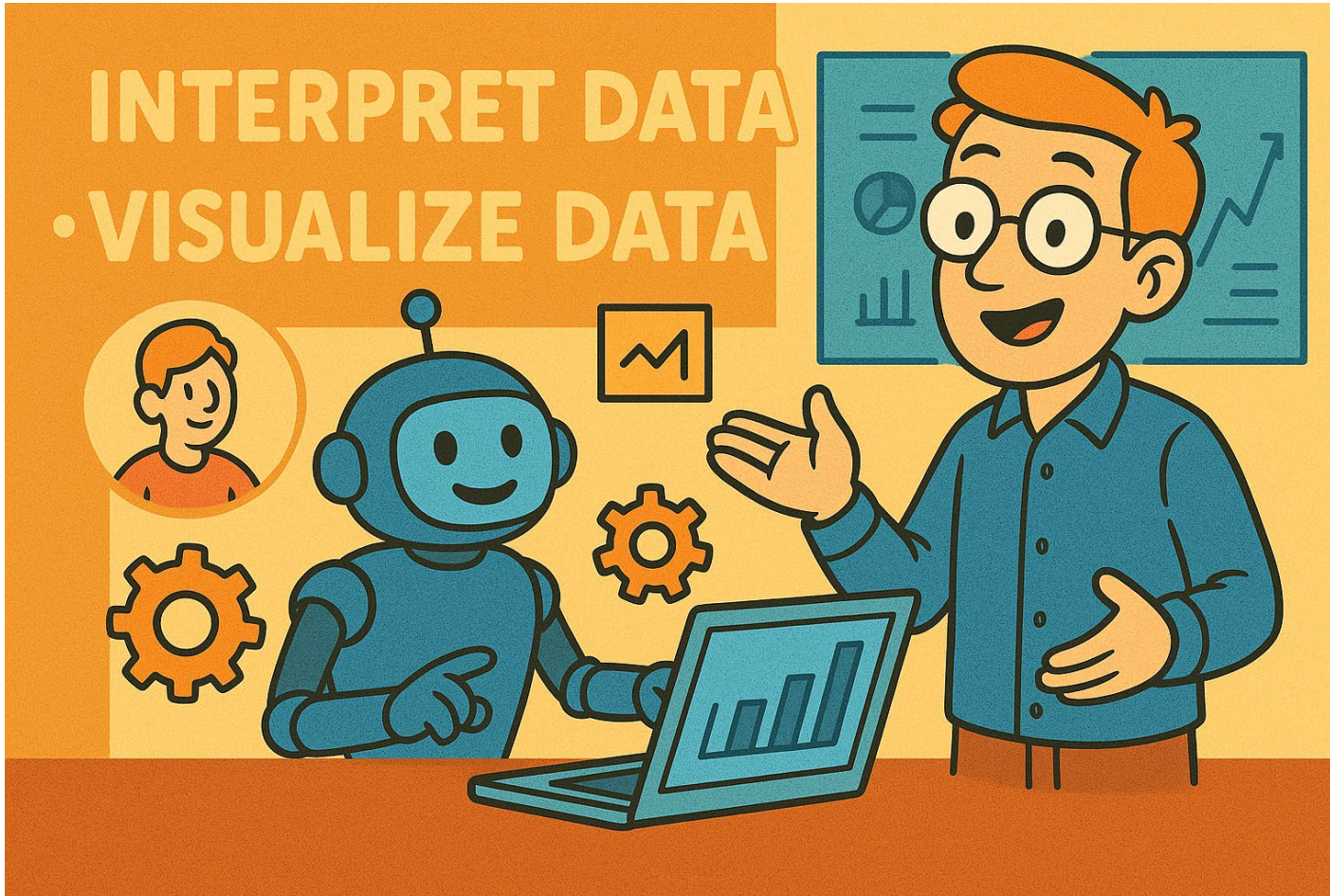


EXPLAINED
WITH A STORY



by B. Rossen & K. Lurie

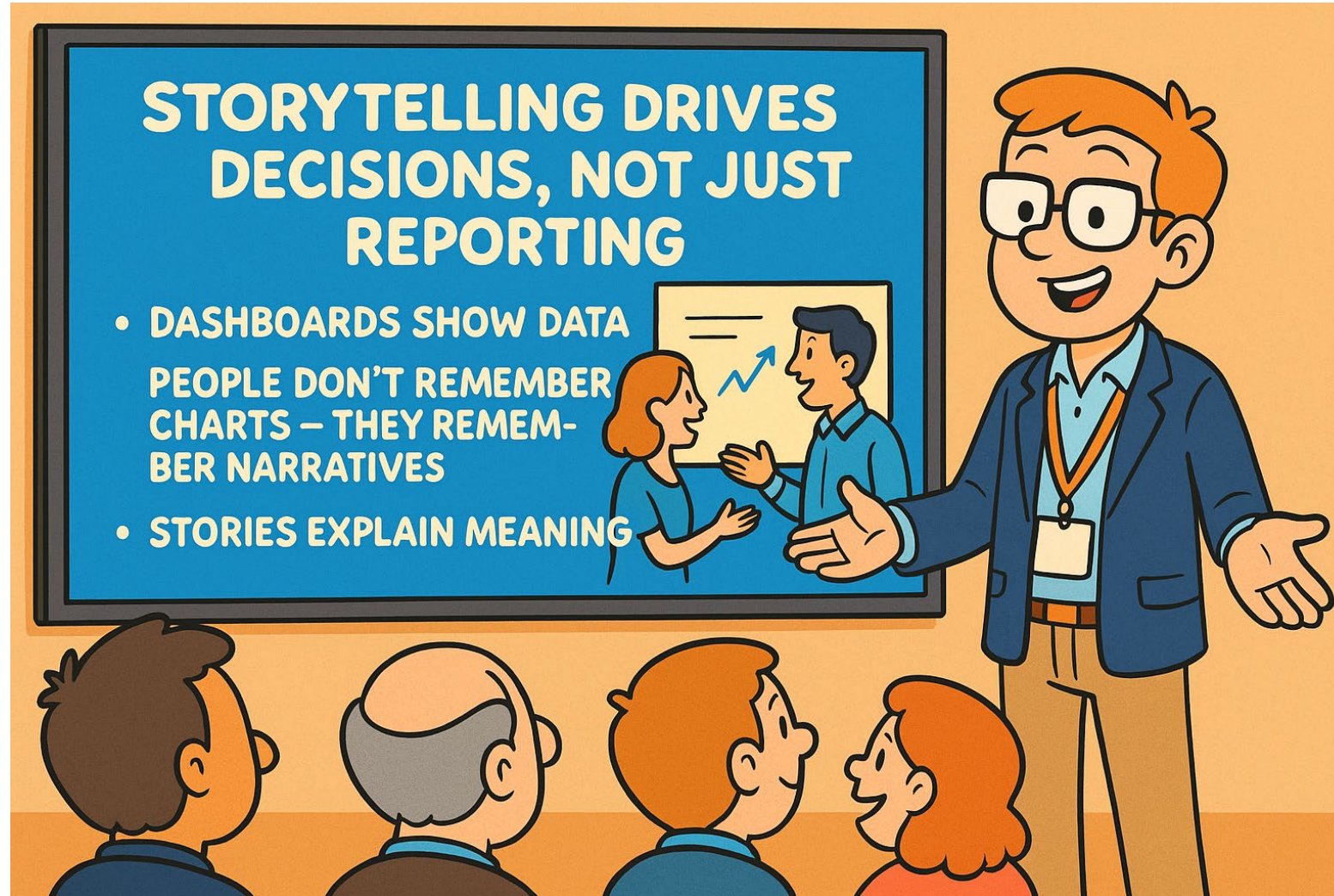
Core Skills of Data Literacy



To be data literate, people need to:

- Interpret data
- Visualise data
- Communicate insights
- Stay curious and ask better questions

The Rise of Data Storytelling





Common Barriers to Data Understanding

Generative AI
NLP
Data Science
IoT
Big Data

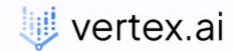


Too much jargon

Tools are complex

Too many dashboards, not enough clarity

Analysis paralysis: overwhelmed, not empowered





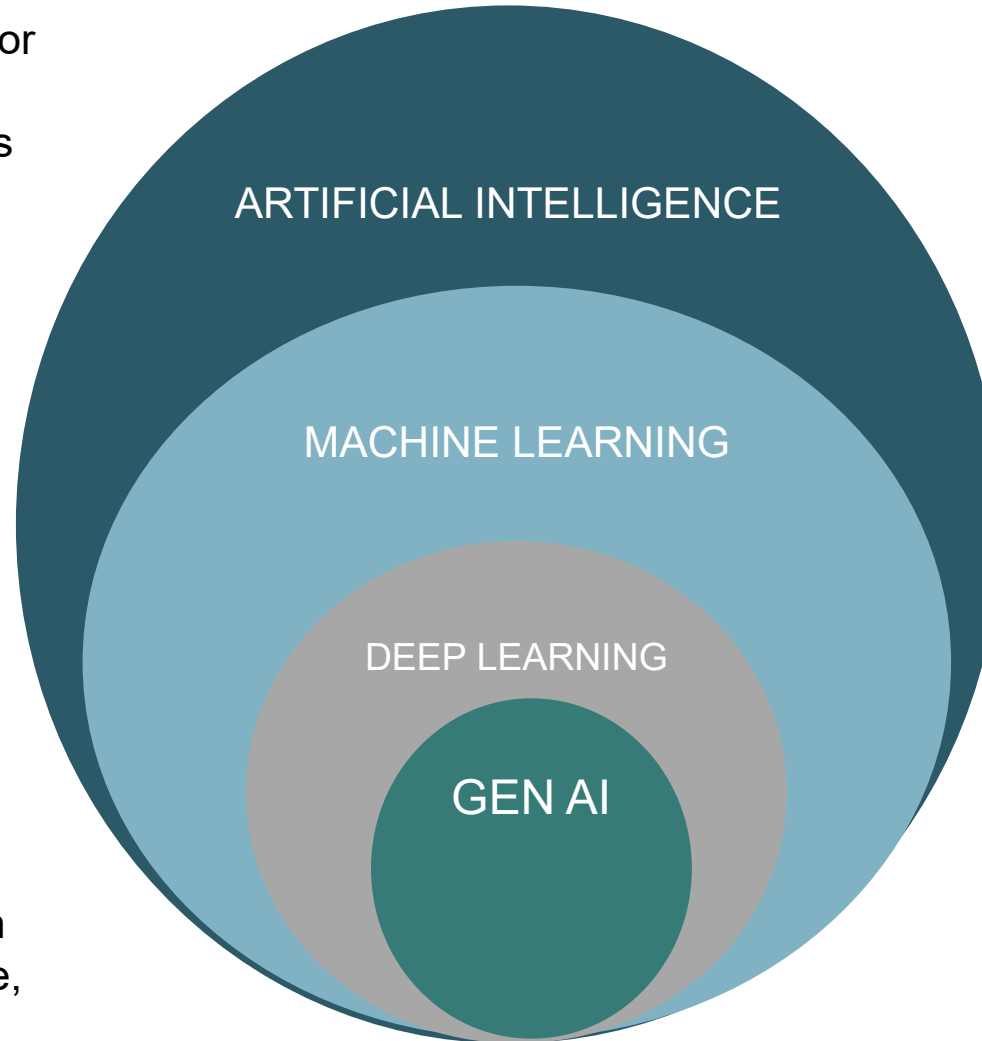
How is it connected to machine learning and data science?

Artificial Intelligence refers to systems or programs that can perform tasks that typically require human intelligence. It is a multidisciplinary field that combines various branches.

Machine Learning is a subset of AI that focuses on the development of algorithms that enable computer systems to learn from experience without being explicitly programmed.

Deep Learning is a subfield of machine learning that uses a structure of algorithms called neural networks.

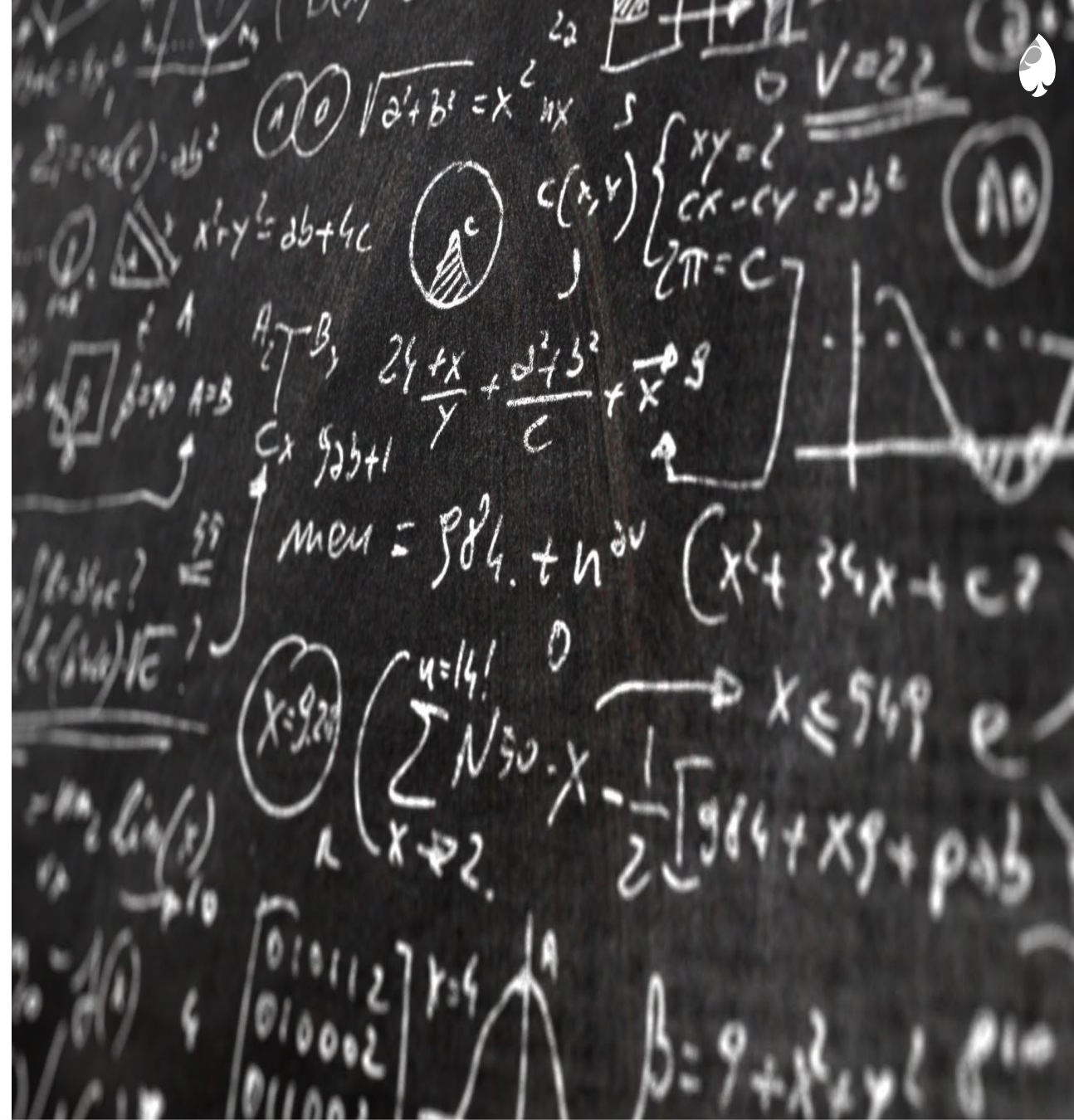
Generative AI is a subset of AI that can create new content such as text, image, code, synthetic data.



Data Science
Data Science = Data + Science

An interdisciplinary field that combines statistics, computer science and domain knowledge to extract insights from data. It covers various techniques such as statistics, prescriptive analytics, predictive analytics, machine learning, deep learning.

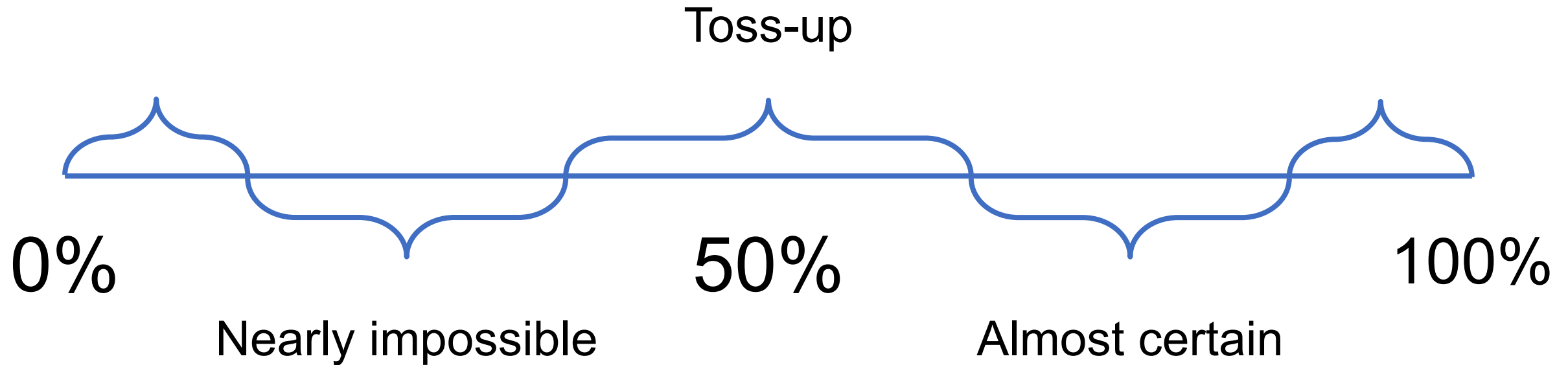
Why AI/ML Makes Data Literacy **MORE** Important





Probability

% chance of an event



Evaluation of Accuracy

Correct predictions/all cases

A value between 0-1

Works quite well when the data is evenly balanced

We always say no -> accuracy is very good



https://gking.harvard.edu/publications/google-flu-trends-still-appears-sick%C2%A0-evaluation-2013%E2%80%902014-flu-season?utm_source=chatgpt.com



Evaluation : Confusion Matrix

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	True positive	False positive
	Negative (0)	False negative	True negative



Evaluation : Confusion Matrix

		Actual Values	
		Fraud Positive (1)	Not Fraud Negative (0)
Predicted Values	Predicted Fraud Positive (1)	200 (True positive)	200 (False positive)
	Predicted Not Fraud Negative (0)	800 (False negative)	98,800 (True negative)



Evaluation : Confusion Matrix

		Actual Values	
		Fraud Positive (1)	Not Fraud Negative (0)
Predicted Values	Predicted Fraud Positive (1)	200 (True positive)	200 (False positive)
	Predicted Not Fraud Negative (0)	800 (False negative)	98,800 (True negative)



Evaluation : Confusion Matrix

		Actual Values	
		Fraud Positive (1)	Not Fraud Negative (0)
Predicted Values	Predicted Fraud Positive (1)	700 (True positive)	700 (False positive)
	Predicted Not Fraud Negative (0)	300 (False negative)	98,300 (True negative)

Context (additional domain knowledge)

Context means the setting of a word or event.....

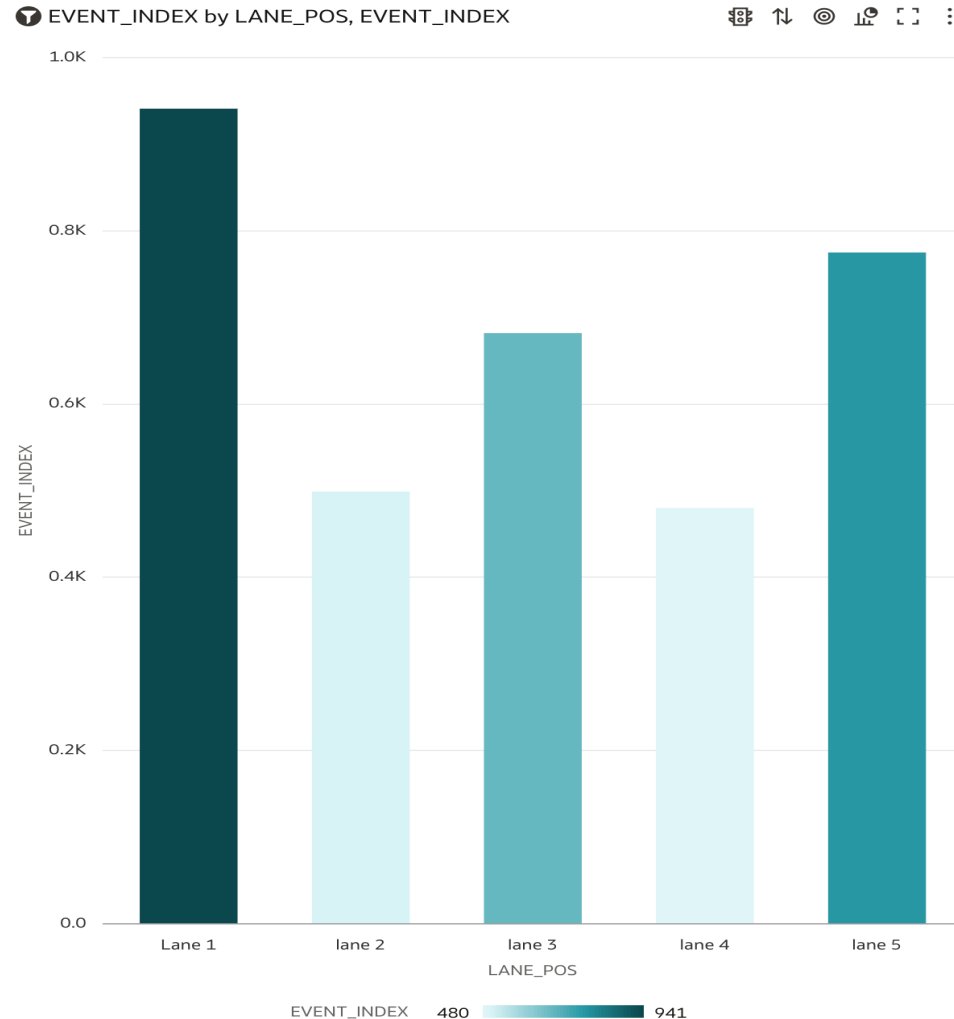
9 types of Context in contextual search

- Individual
- Demographic profile
- Interest profile
- Location
- Device
- Date
- Time
- Weather
- Mood



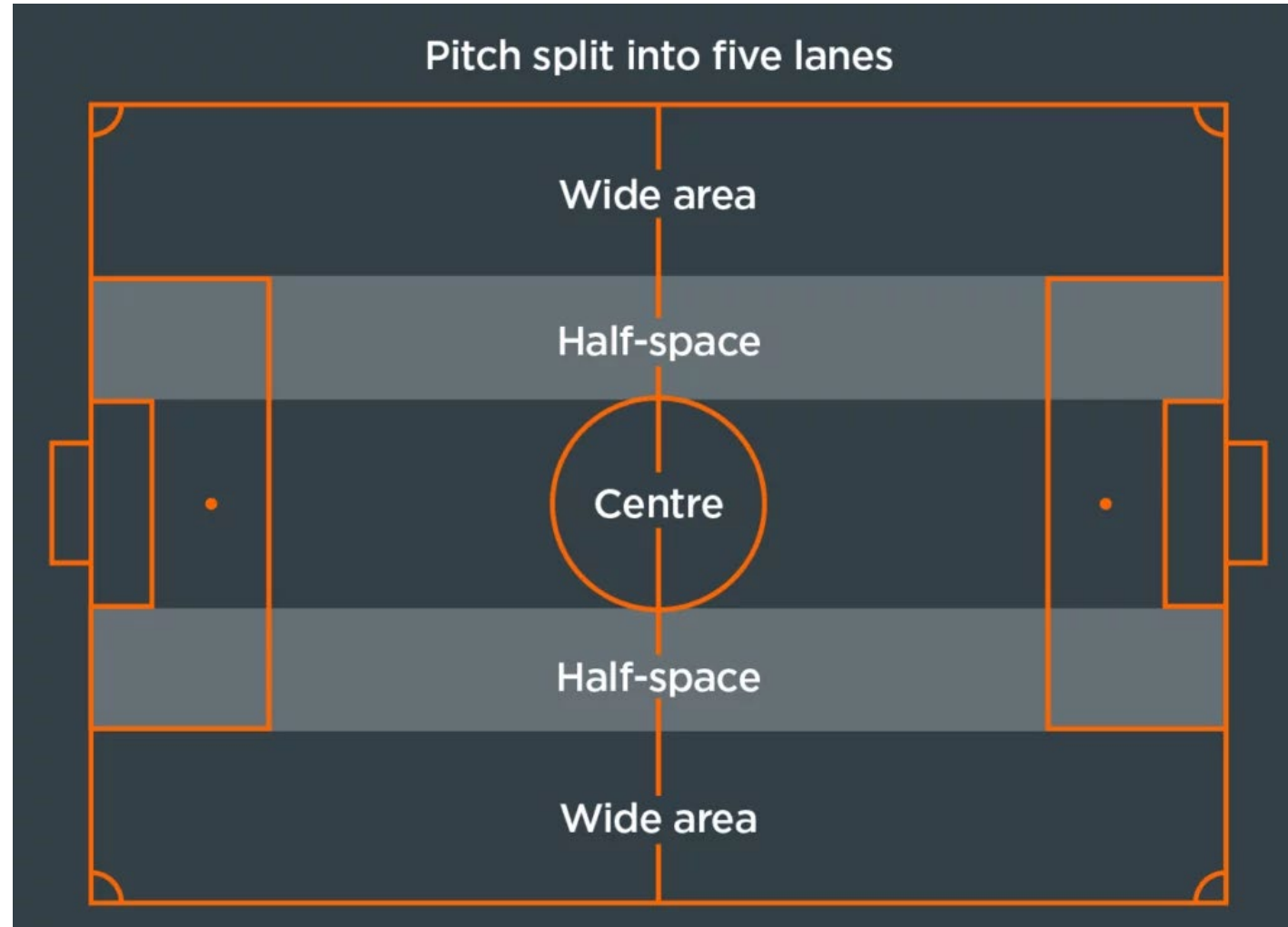
[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)

Context



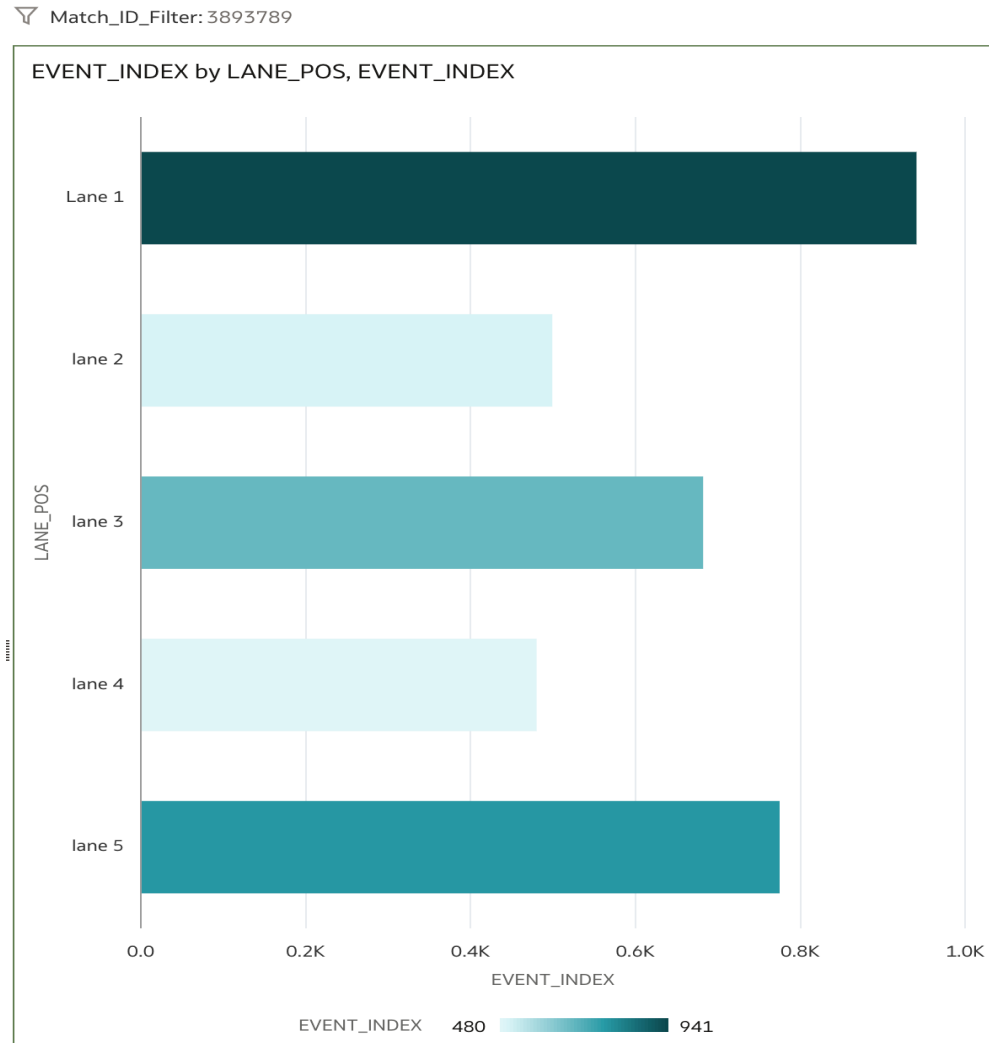
- Analysis of the number of passes and where they come from
- Measure (number of passes), Attribute (location - lane 1,2,3,4,5)
- So, a bar chart.....

Context





Context

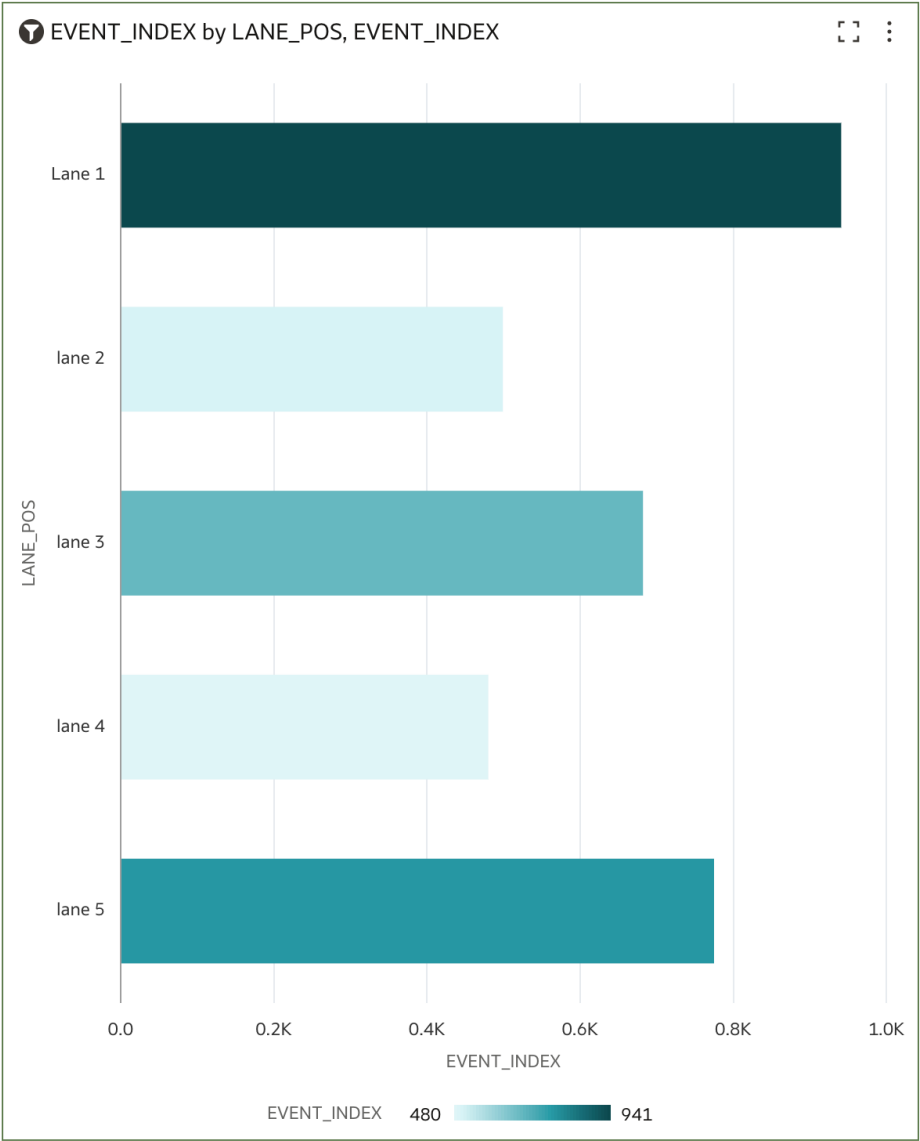


- Analysis of the number of passes and where they come from
- Measure (number of passes), Attribute (location - lane 1,2,3,4,5)
- So, a **Sideways** bar chart.....

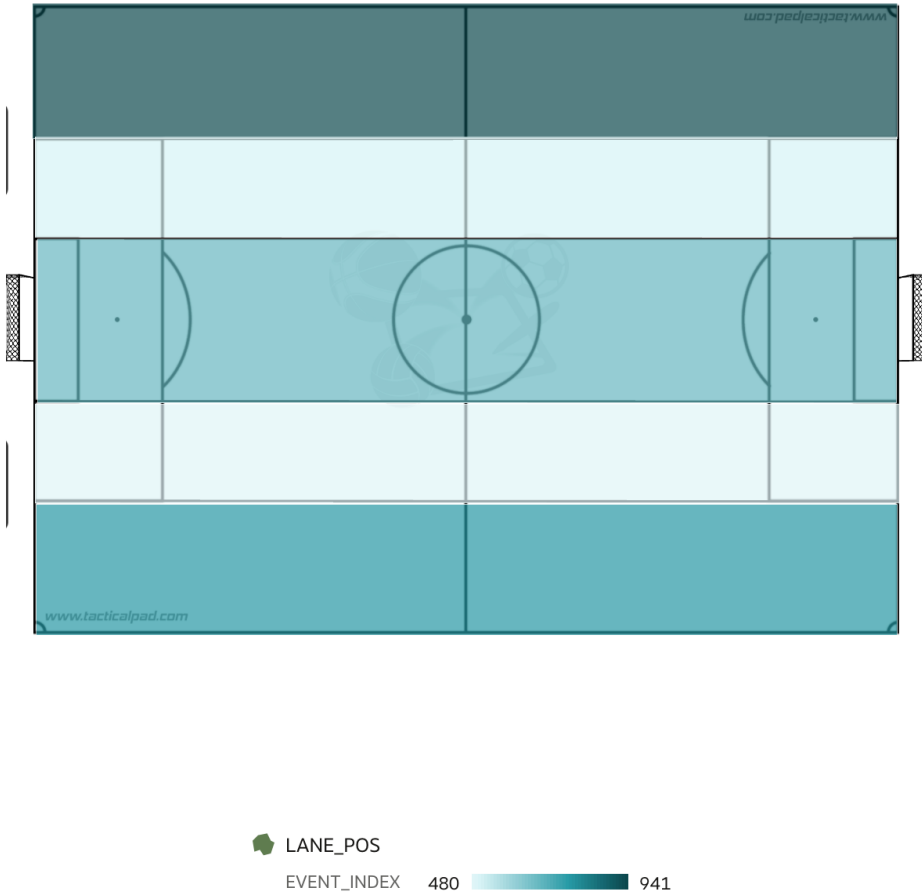


Match_ID_Filter: 3893789

Context



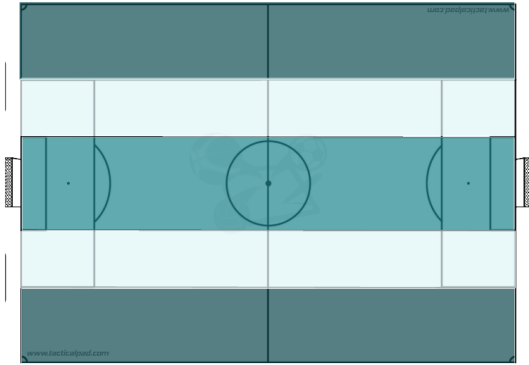
LANE_POS, EVENT_INDEX





Context

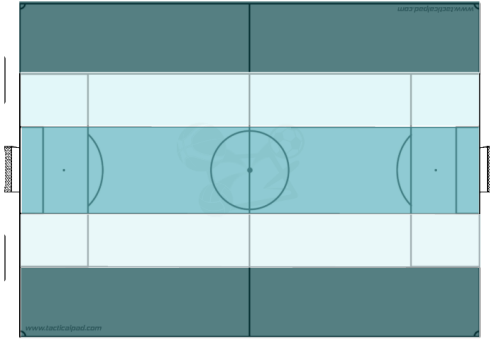
LANE_POS, EVENT_INDEX



LANE_POS
EVENT_INDEX 28K 47K

LANE_POS, EVENT_INDEX

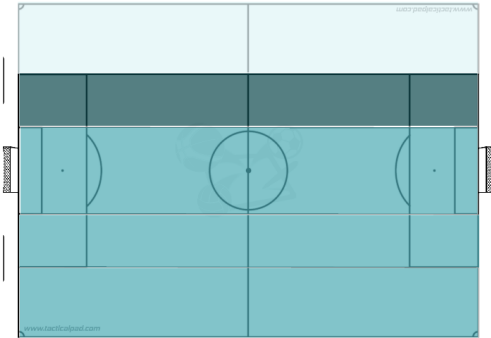
UNDER_PRESSURE: True



LANE_POS
EVENT_INDEX 6K 12K

LANE_POS, EVENT_INDEX

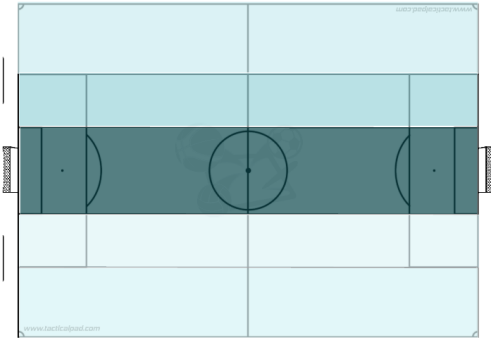
UNDER_PRESSURE: True PASS_THROUGH_BALL: True



LANE_POS
EVENT_INDEX 4 6

LANE_POS, EVENT_INDEX

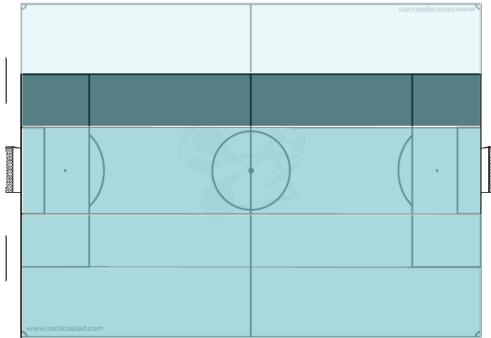
UNDER_PRESSURE: True PASS_SHOT_ASSIST: True



LANE_POS
EVENT_INDEX 17 45

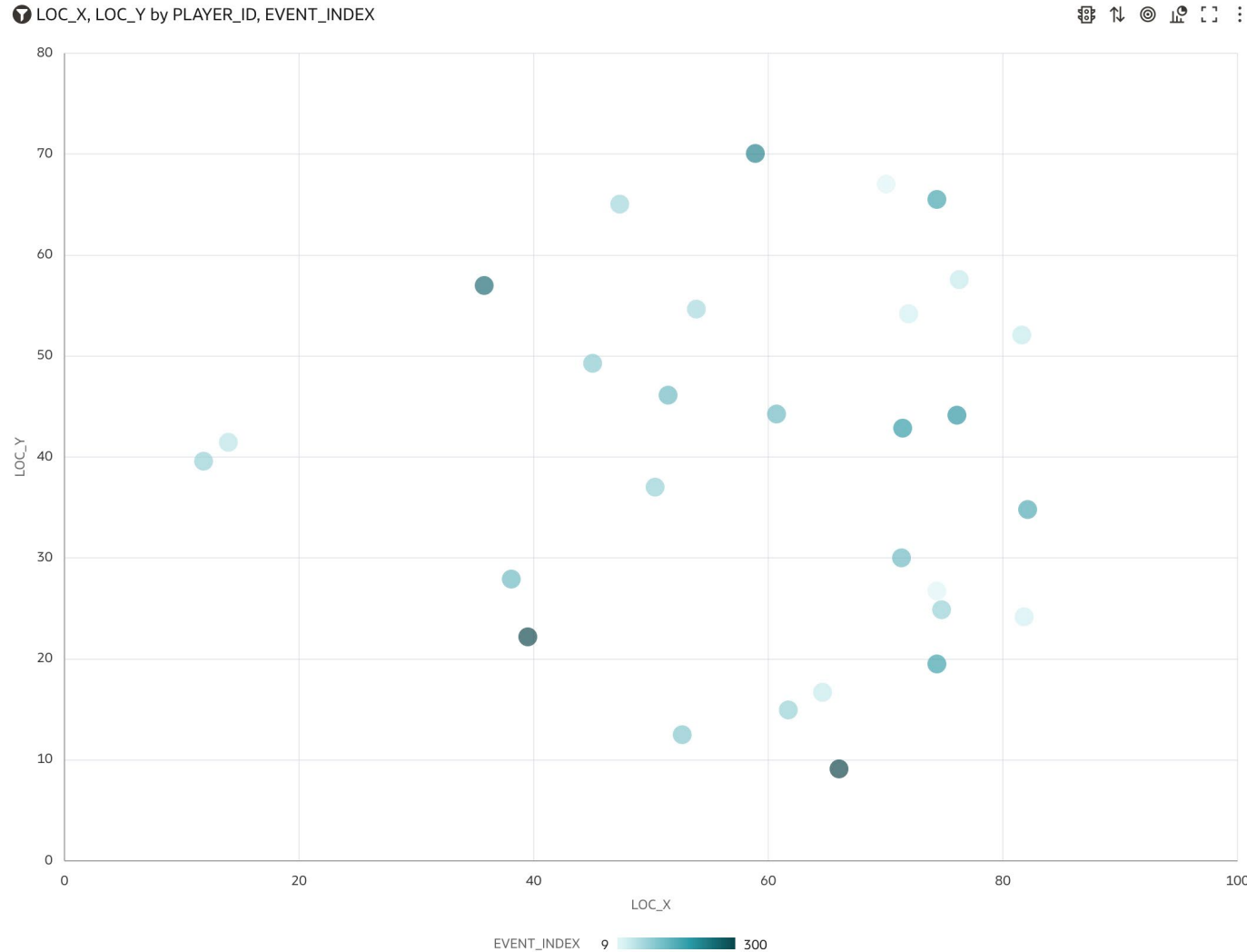
LANE_POS, EVENT_INDEX

UNDER_PRESSURE: True PASS_GOAL_ASSIST: True



LANE_POS
EVENT_INDEX 1 4

Context: What about X and Y coordinates....

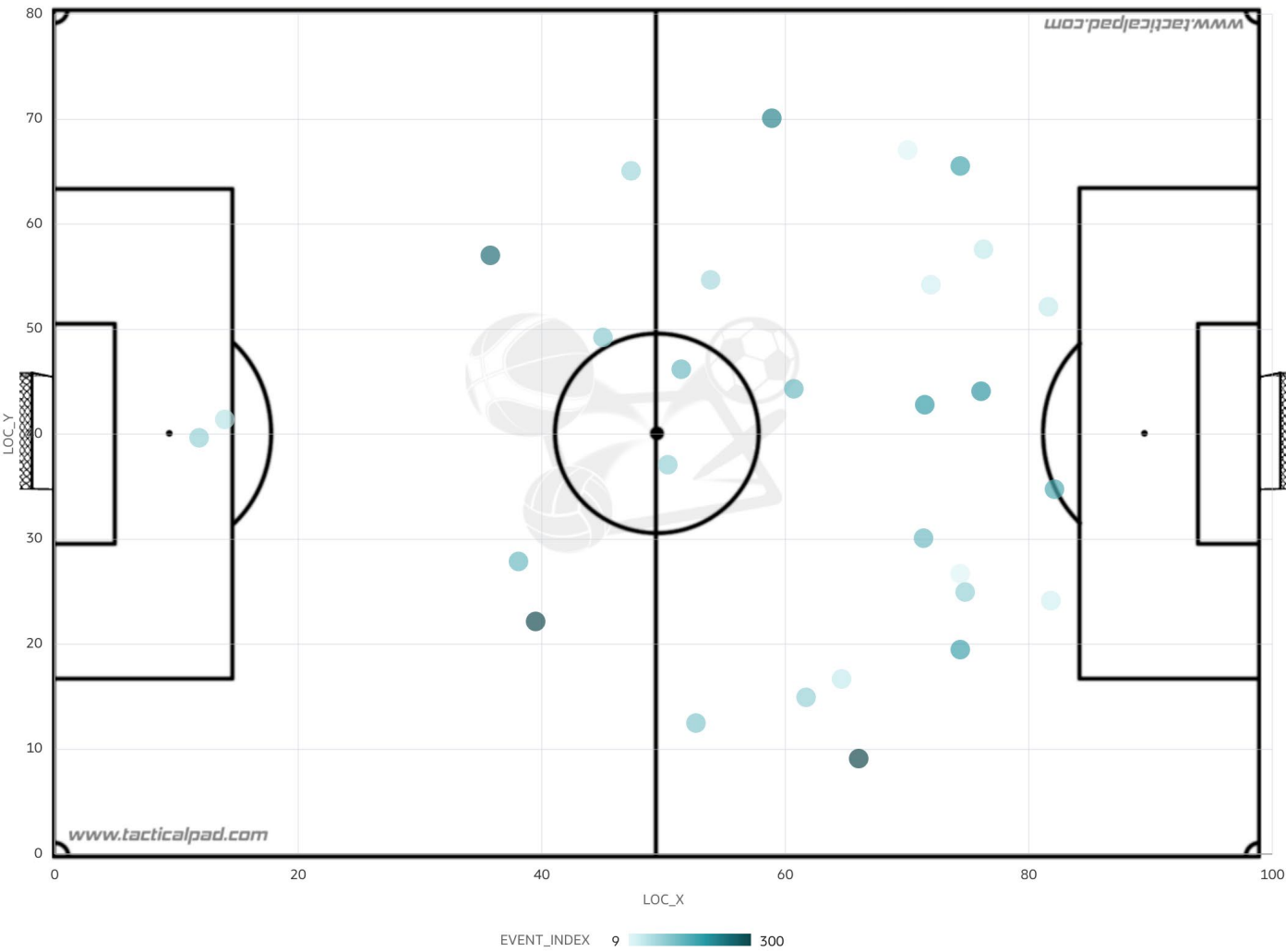


- Analysis on the average position each player was on the pitch when completing passes
- X and Y coordinates per passing event
- So, that's a scatter plot?



Context: What about X and Y coordinates....

LOC_X, LOC_Y by PLAYER_ID, EVENT_INDEX



▲ Style

Background

Custom

Fill Color

Color Transparency

0%

Image Source

Pitch_View_JPG3.jpg

↑

Image Align

Image Transparency

0%

Image Aspect Ratio

Unlocked

Image Width

1115

Image Height

780

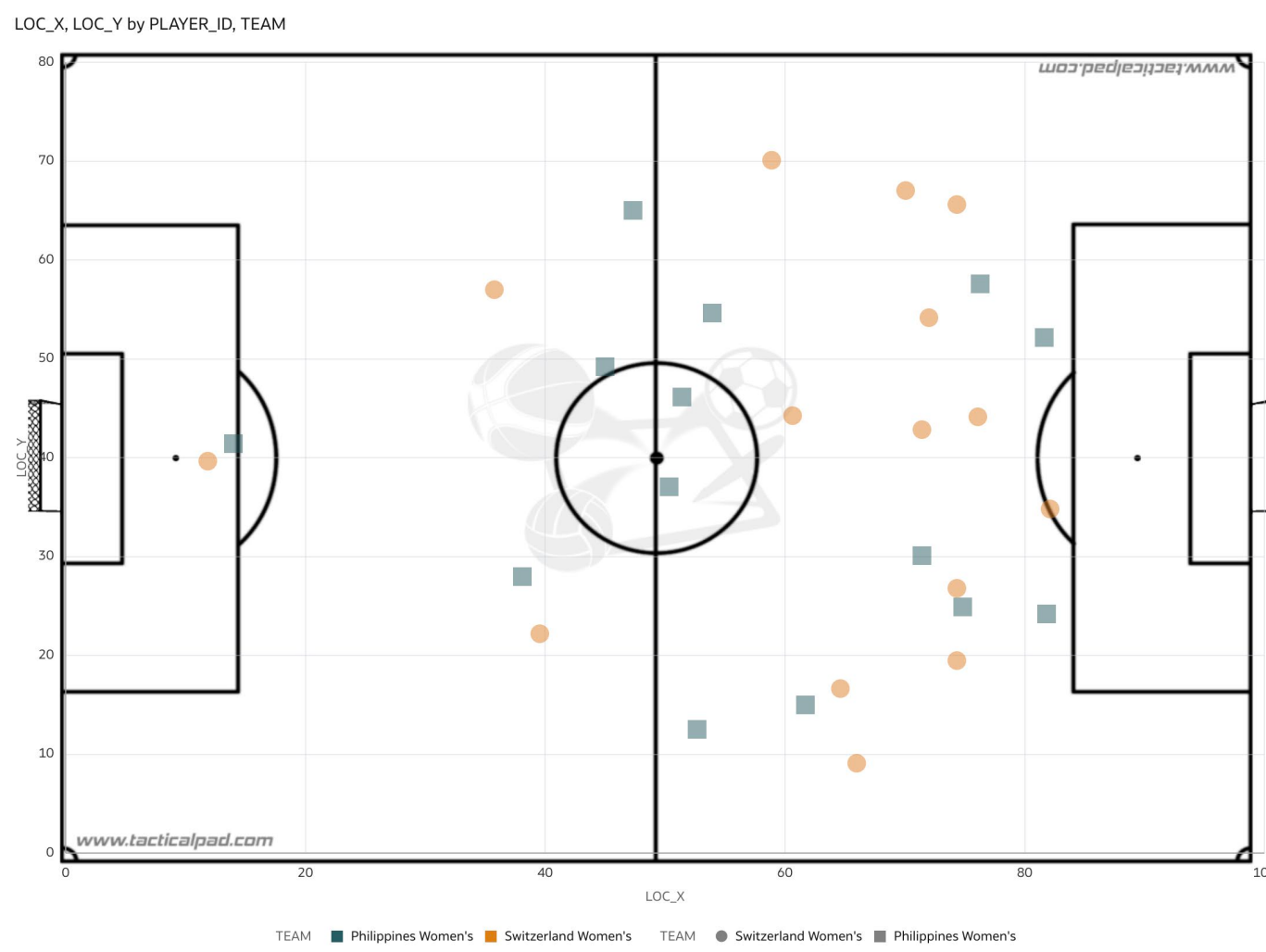
Border

None

Shadow

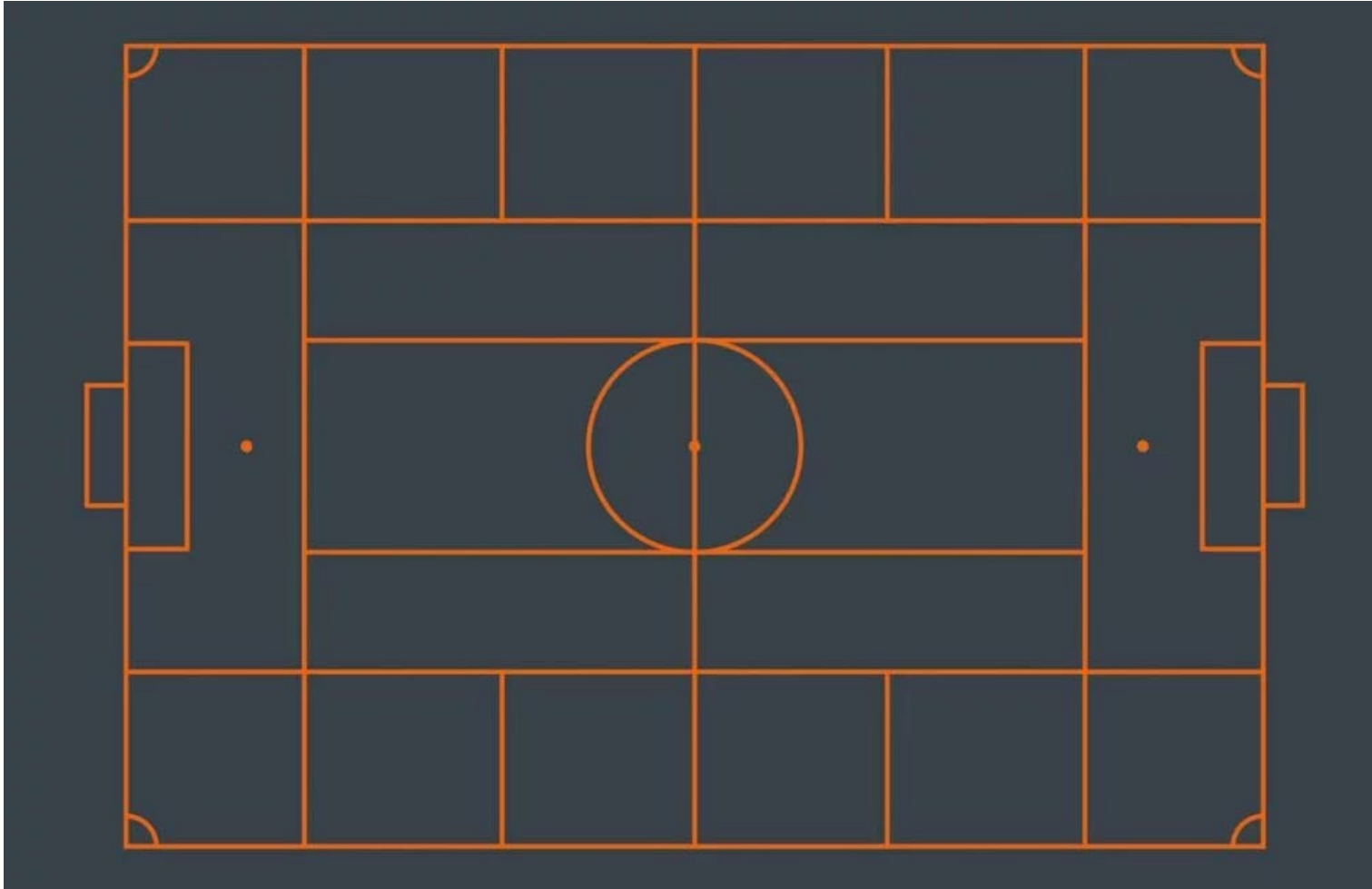
None

Context: What about X and Y coordinates....



To colour or not
to colour....

Context: Pitch orientation



The playing field is divided into zones with four vertical lines and some horizontal lines.

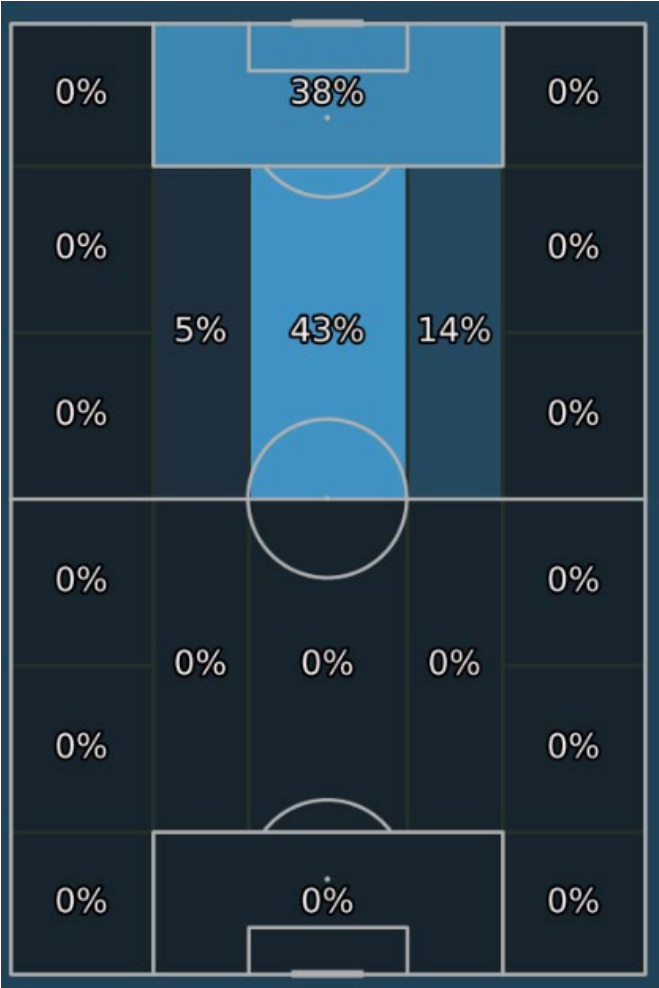


Where do we have shots from: Context + Data

Homea Team Goals

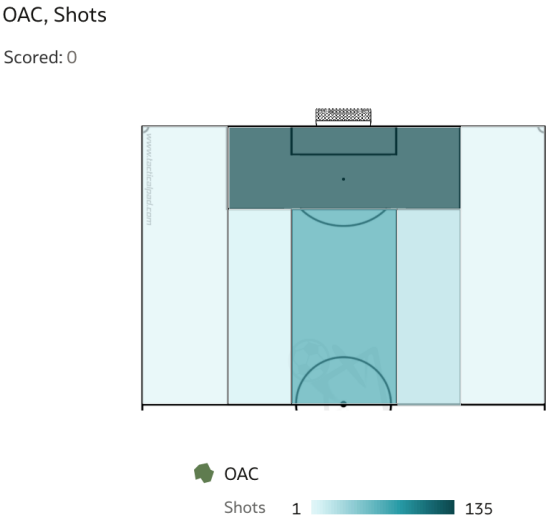
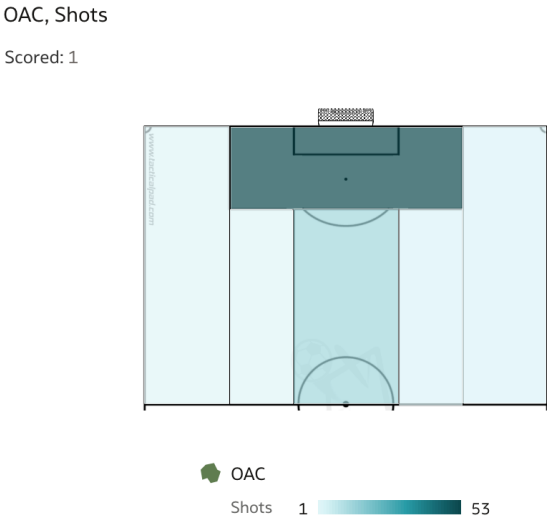
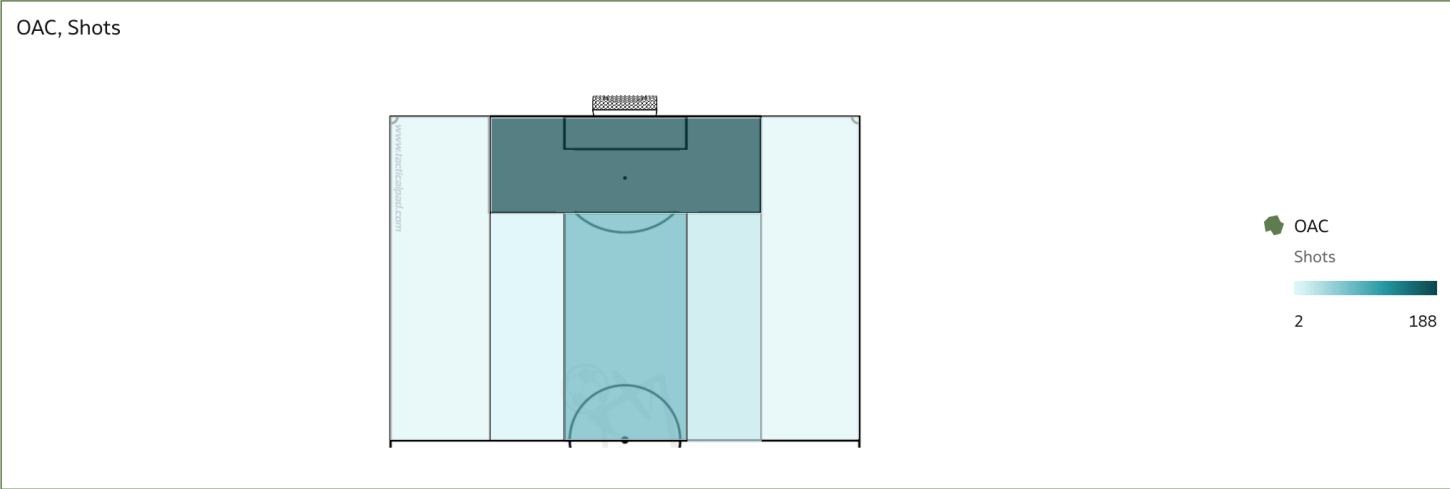


Home Team No Goal





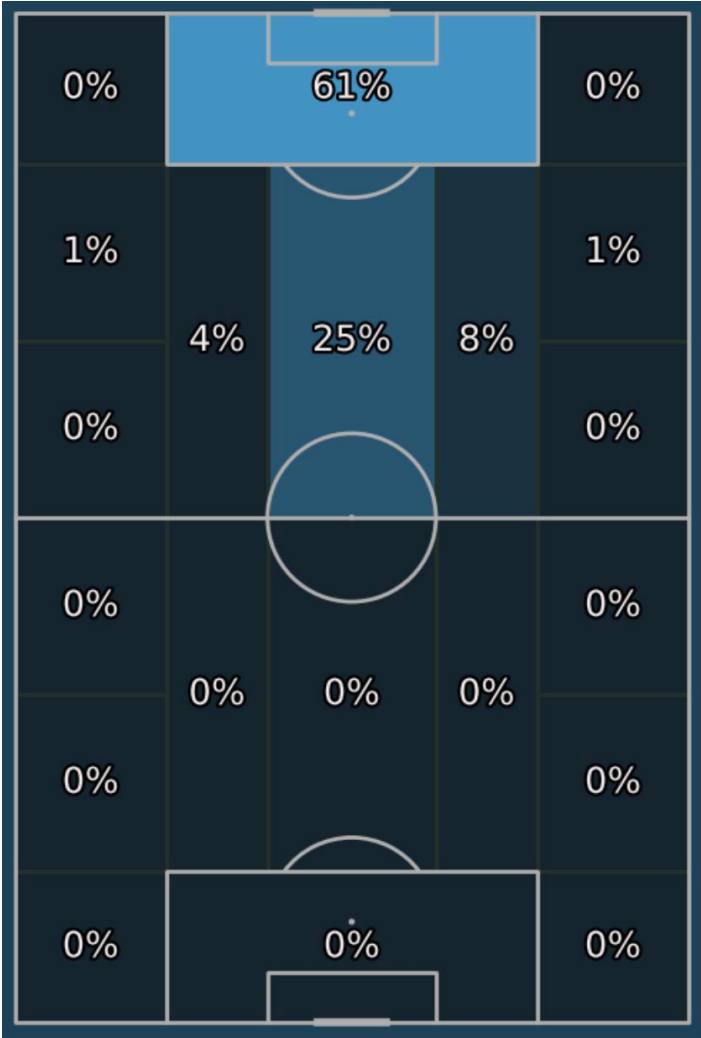
Where do we have shots from: Context + Data



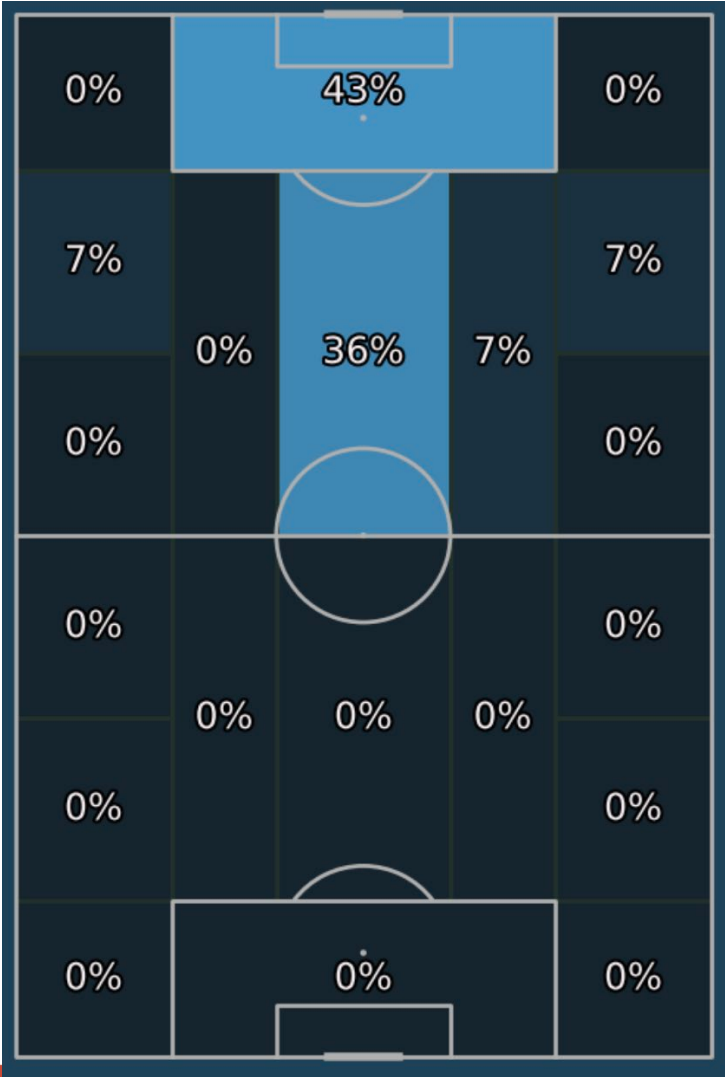


Look back at the season....

Shots throughout the season



Goals scored throughout the season





Let's build our story

Chapter 1 – Introduction

Thinking and Understanding

Chapter 2 –
Context
Who, where and
when

Chapter 3 –
Purpose
Why?

Chapter 4 –
Design
How?

Don't Assume knowledge

Serve the setting, as well as
the audience



What top tips would I give....

Data: The Foundation of Every Data Story

- Gather high-quality and relevant data.
- Understand the data's source and credibility.
- Ensure data is processed and cleaned for accuracy.
- Highlight key data points and metrics.

Analysis: Discovering Insights and Patterns

- Choose appropriate analysis techniques.
- Identify correlations and causations.
- Incorporate statistical significance and confidence levels.



What top tips would I give.....

Narrative: Crafting a Compelling Story

- Define the purpose and audience of the data story
- Establish a clear and engaging storyline
- Identify the central message or takeaway
- Incorporate real-world examples and anecdotes

Visualisation: Enhancing Data Communication

- Select suitable visualisation types for the data
- Emphasize clarity, simplicity, and aesthetics
- Add interactive elements for user engagement
- Use visualisations to support the narrative



What top tips would I give.....

Context: Providing a Holistic Perspective

- Place the data within context
- Address potential biases and limitations
- Compare data with history or benchmarks
- Explain external factors influencing the data

Delivery: Matching the Presentation to the Audience

- Choose the suitable medium for delivery (e.g., blog, presentation, infographic)
- Adapt the data story to the audience's knowledge level
- Incorporate storytelling techniques
- Encourage audience interaction and feedback



Ideas and Questions always welcome....

Dr. Abi Giles-Haigh

Oracle Analytics and Innovation Director,
Capgemini

Abi.Giles-Haigh@Capgemini.com

AbiHaigh@Hotmail.com

@Abi_Giles_Haigh

