TABLEAU CONFERENCE

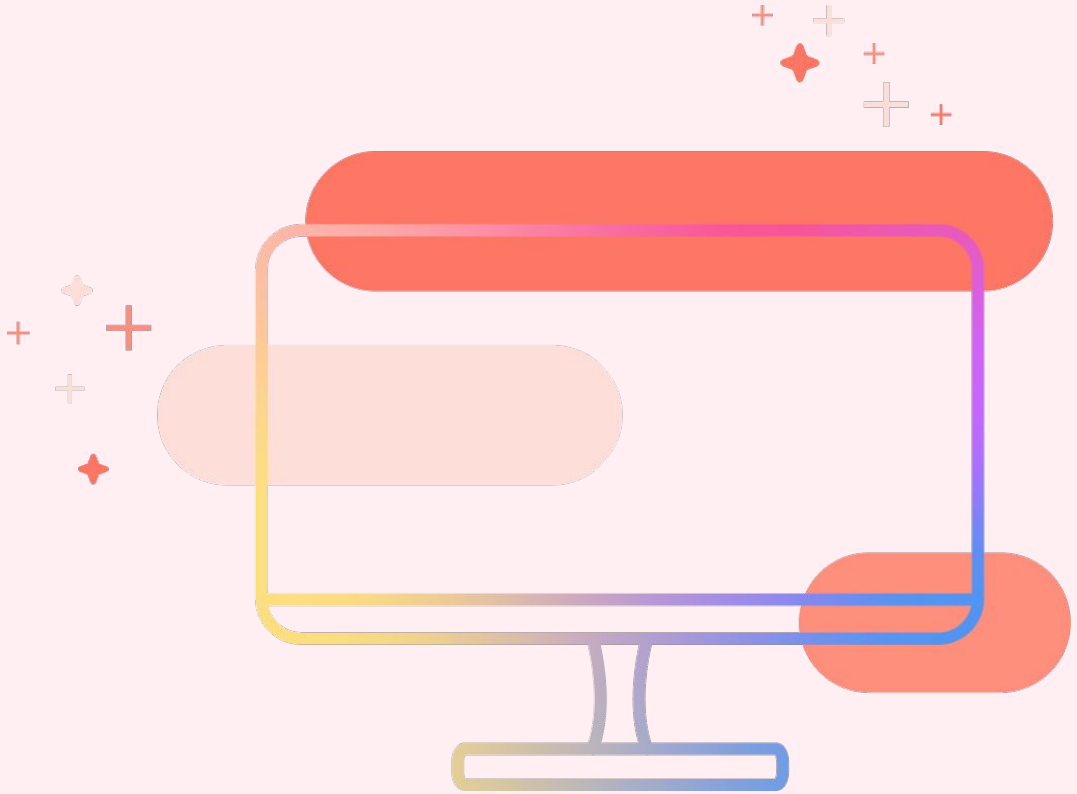
Tune in using Live Session Audio.

- Scan the QR code or navigate to the Live Session Audio Player.
- Connect headphones to your mobile device.
- Choose this location.
- Begin listening or view captions on your phone.

Salesforce Events Mobile app required.

TABLEAU CONFERENCE

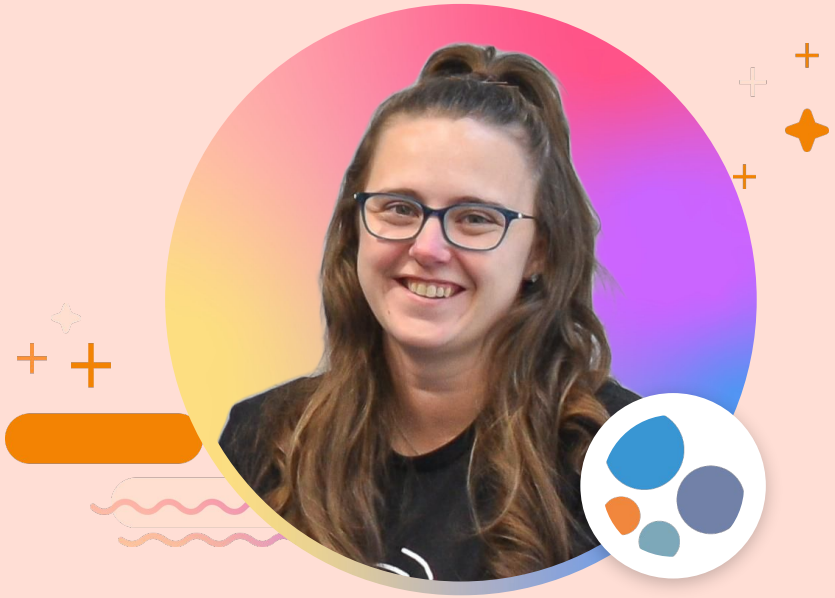
Taming the Data Beast: Normalization Strategies in Tableau

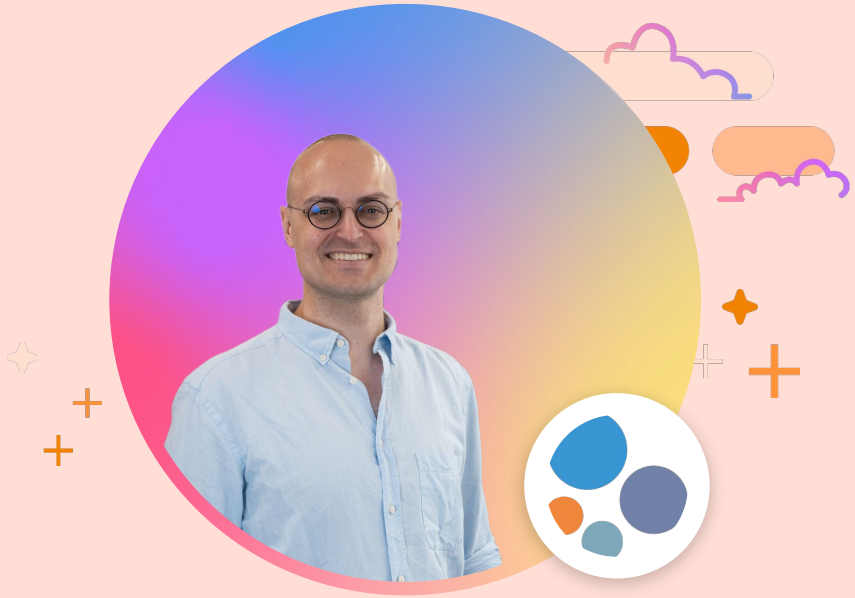


Lorna Brown, Data School Coach
@_Lorna_Brown
Lorna.Brown@theinformationlab.co.uk
She/Her

Robbin Vernooij, Data School Coach
@RobbinVernooij
Robbin.Vernooij@theinformationlab.co.uk
He/Him

TABLEAU CONFERENCE





Lorna Brown

Data School Coach,

The Information Lab

Robbin Vernooij

Data School Coach,

The Information Lab

TABLEAU CONFERENCE

Agenda

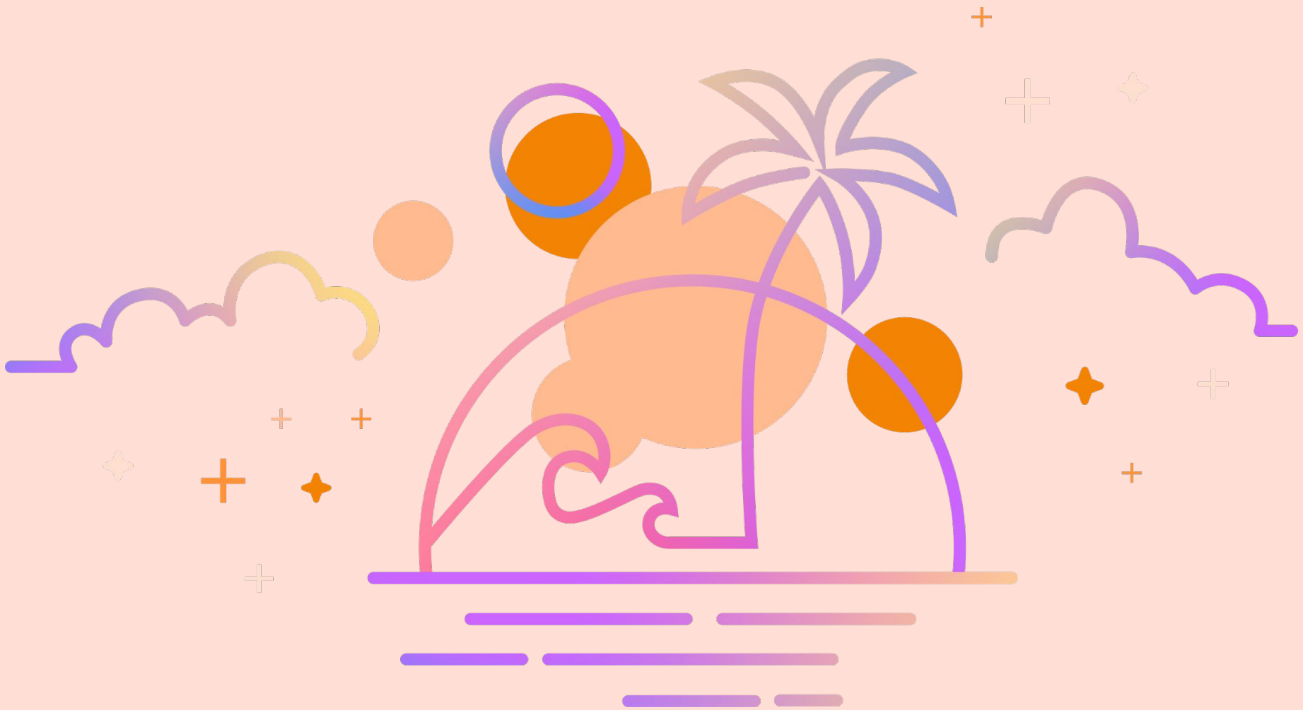
01 What is Data Normalization?

02 Examples in the Wild

03 Z Scores

04 Parallel Coordinates / Radar Charts

05 Preparing Data with Tableau Prep



What is Data Normalization?

And why do we need to do it?

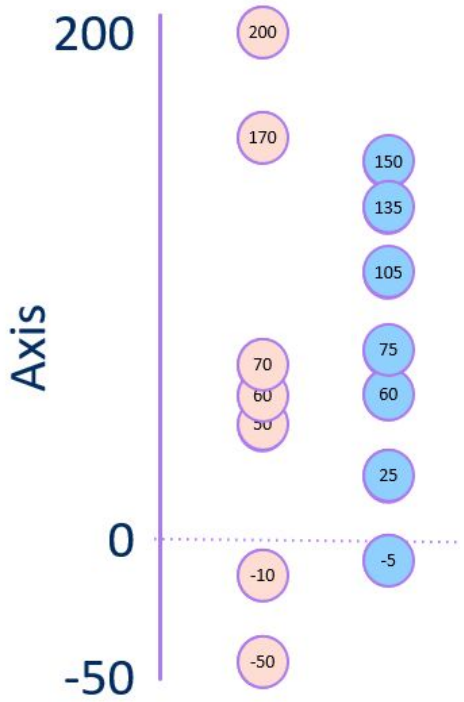
What is Data Normalization?

Why do we need to do it?

- The process of transforming data from various sources into a consistent format
- Allows for easier analysis and comparison by adjusting values to a common scale
- By normalizing the data you have the ability to compare like for like
- This can make your data easier to understand, give you more relevant analysis, or sometimes both!

What is Data Normalization?

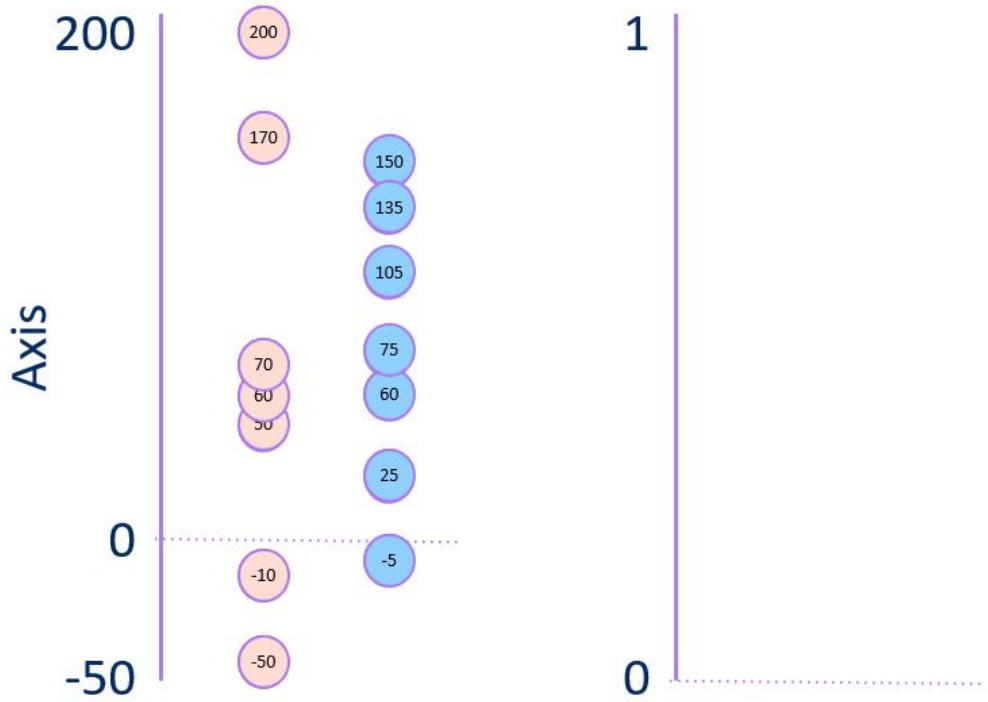
and why do we need to do it?



What is Data Normalization?

and why do we need to do it?

Min Max Norm.

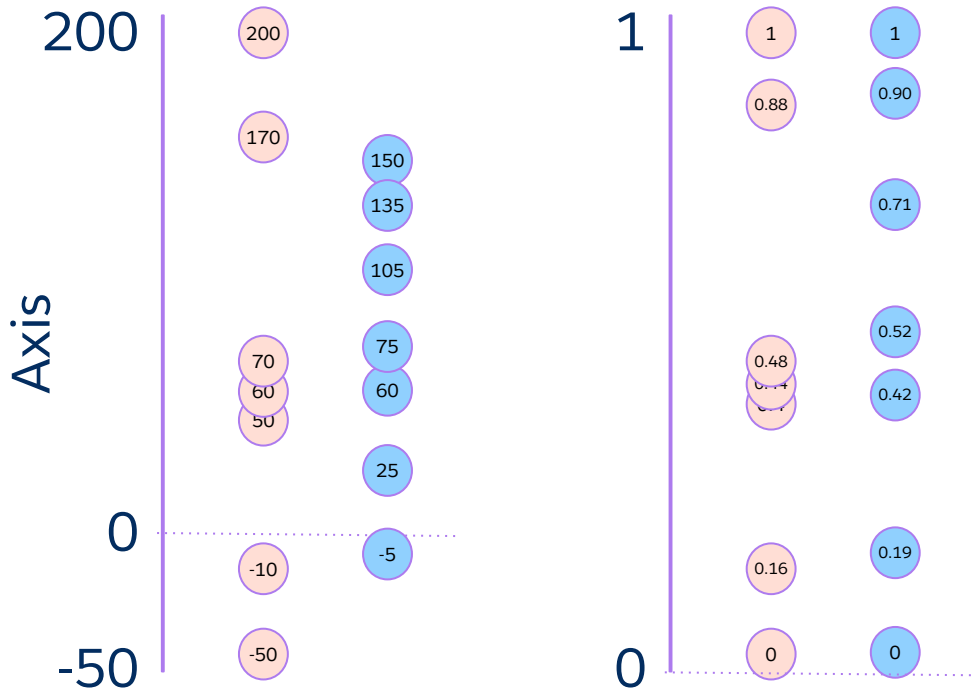


What is Data Normalization?

and why do we need to do it?

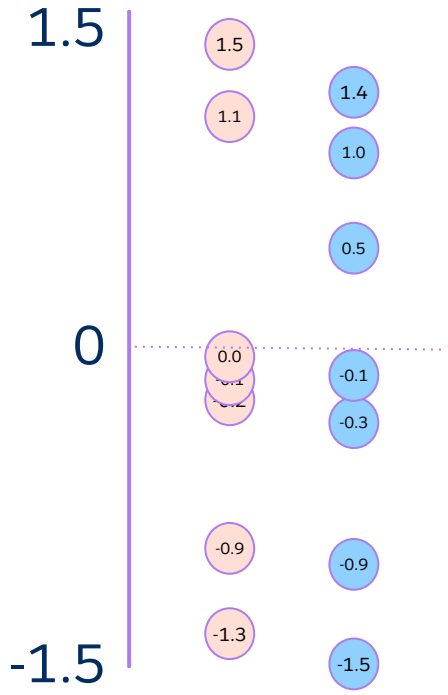
TABLEAU CONFERENCE

Min Max Norm.



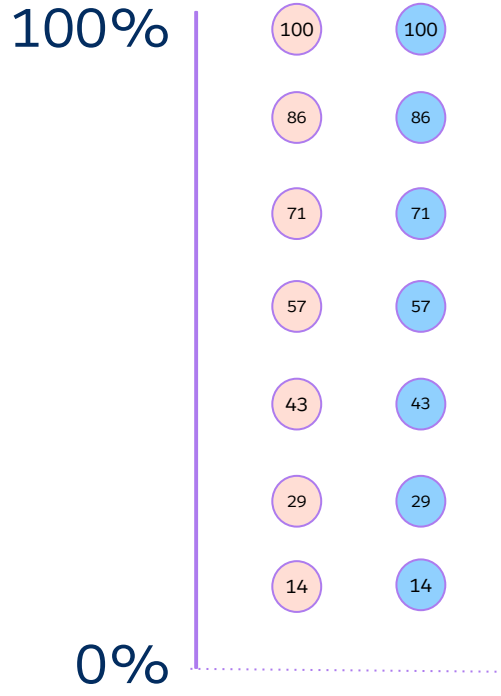
$$= \frac{value - min}{max - min}$$

Z-Score



$$= \frac{value - avg}{stdev}$$

Percentile Rank



$$= \frac{position}{n} +$$

Examples in the Wild

NRL Player Stats

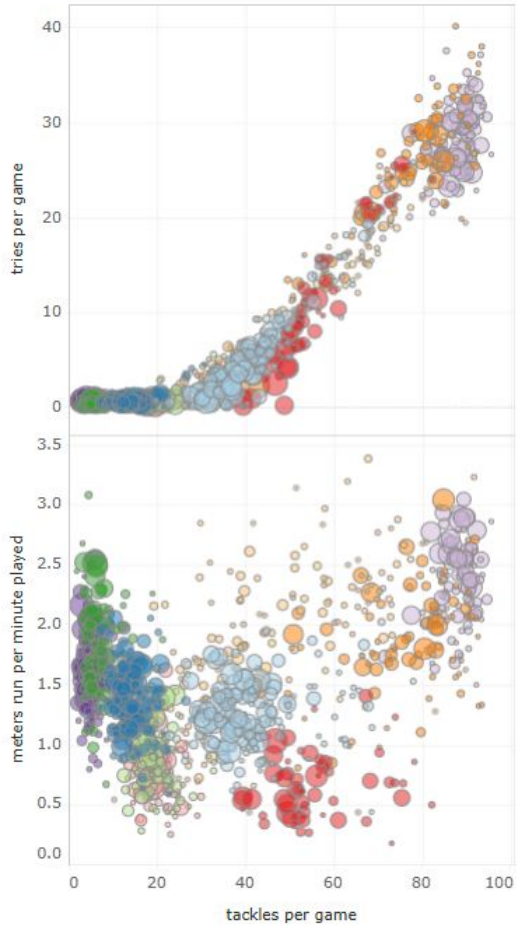
| Data Normalization and Percentile Calculations

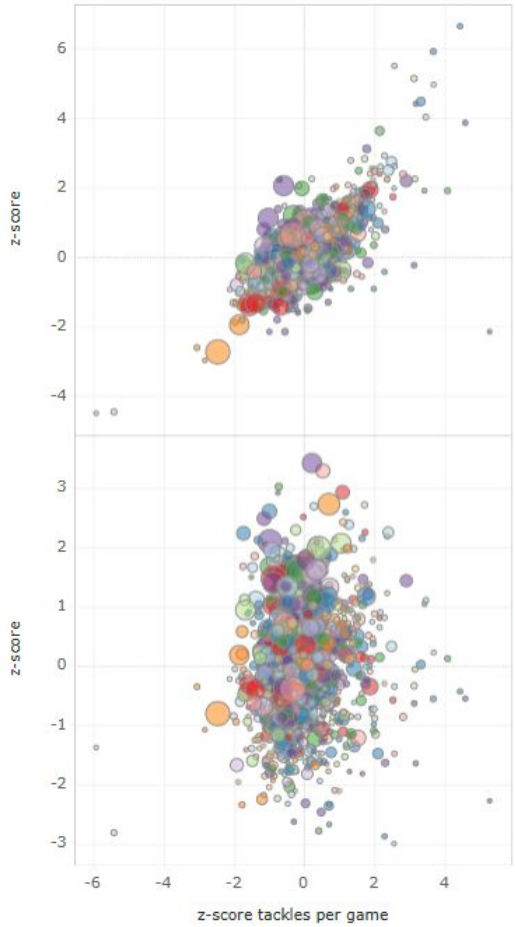
METRIC DETAILS FILTERS

TABLEAU CONFERENCE

2nd Row | Centre | FiveEighth | Fullback | Halfback | Hooker | Interchange | Lock | Prop | Winger

circles sized by games played

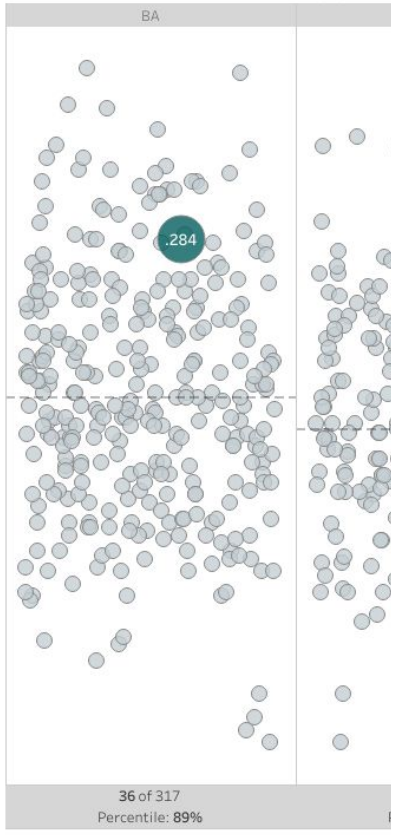




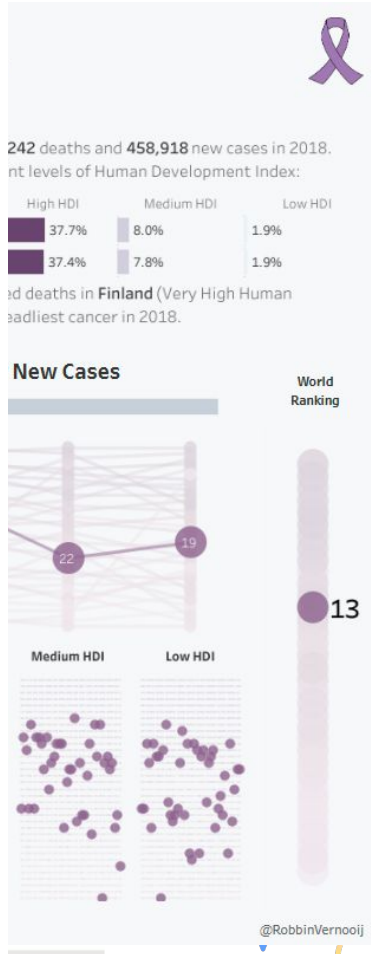
Data: https://www.nrl.com/

Part of the **Explaining with** blog series: https://github.com/Imba456/explaining_with

@RobbinVernooij



and thrown for over 2000 yds in total.

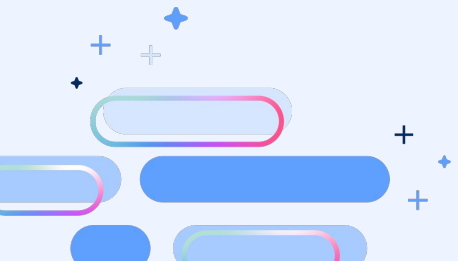


na_Brown

Calculating Z Scores

What is a Z-Score?

A z-score is a statistical measure that describes how many standard deviations a data point is from the mean of a dataset.



- It helps to understand the relative position of a value in a distribution.
- Finding common patterns within different groups of data points.
- They are used to derive t- and p-values (Math..shivers).

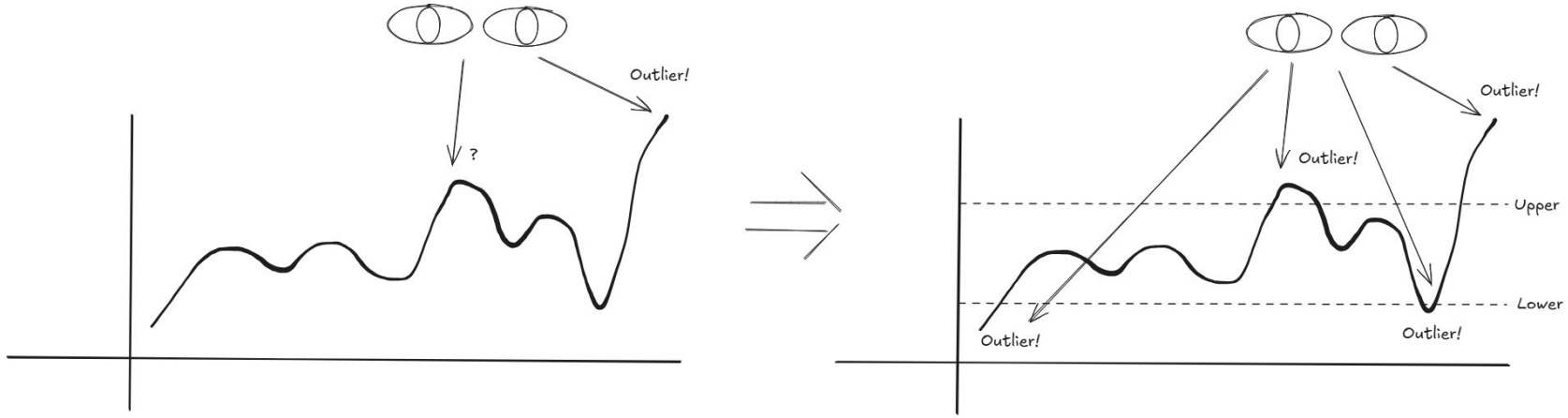
Interpretation:

- A z-score of 0 means the data point is exactly at the mean.
- A positive z-score indicates the data point is above the mean.
- A negative z-score indicates the data point is below the mean.

Why would you calculate Z-Scores?

(Other than for calculating t- or p-values….)

- Do we have outliers? Control Charts
- Peer to Peer comparison, Scatter Plots
- Enhance Box Plots
- Comparing Multiple Data Points



How to calculate a Z-Score

TABLEAU CONFERENCE

$$\frac{(\text{Actual Value} - \text{Average Value})}{\text{Standard Deviation of Values}}$$

Let’s do this in Tableau ....

DEMO

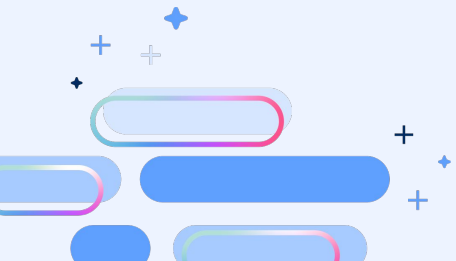
TABLEAU CONFERENCE

Parallel Coordinates & Radar Charts

Using Standardization

What is a Parallel Coordinate Chart?

A parallel coordinate chart is a data visualization technique used to display multi-dimensional data

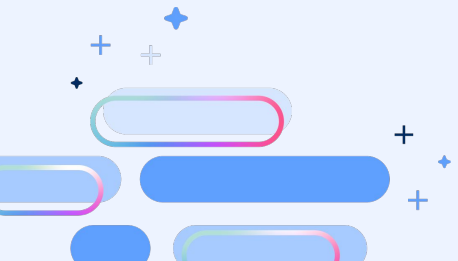


- Multiple Dimensions
- Data Points as Lines
- Parallel Axes
- Visualizing Trends
- Overlapping Lines

Why would you use one?

TABLEAU CONFERENCE

- Handling High-Dimensional Data
- Detecting Patterns and Relationships
- Visualizing Trends Across Multiple Variables
- Comparing Multiple Data Points
- Exploratory Data Analysis (EDA)
- Handling Complex Interactions



Requires Standardization

$$\frac{(\text{Actual Value} - \text{Minimum Value})}{(\text{Maximum Value} - \text{Minimum Value})}$$

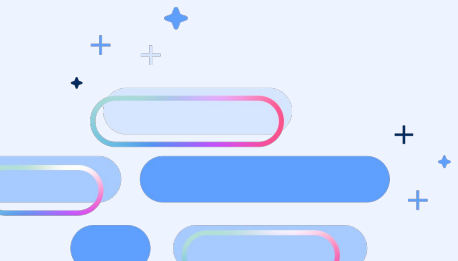
Creates a score between 0 - 1

Let's do this in Tableau

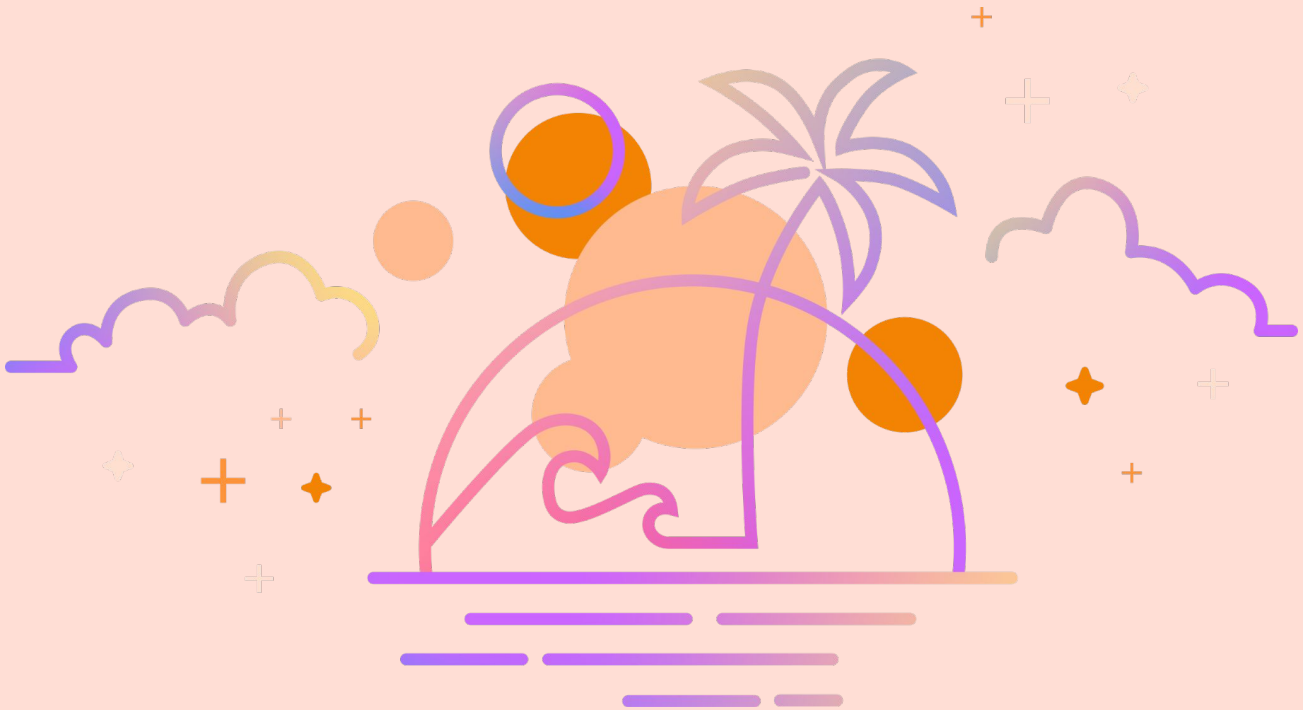
DEMO

What is a Radar Chart?

- Represents the same as a Parallel Coordinate Chart but in a circular format
- Common in performance metrics, especially sports



DEMO

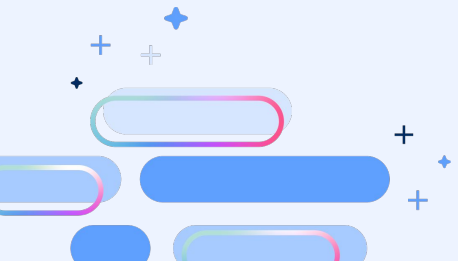


Preparing the Data with Tableau Prep

Why use Prep?

TABLEAU CONFERENCE

- If no filtering is needed on the Viz itself or if filtering is not to affect the normalization.
- Some calcs can be slow depending on the number of data points being visualised (Table Calcs or LoDs)
- Can still be refreshed with new data by just running the Prep Flow first



Let’s show you in Prep!

Key Takeaways

TABLEAU CONFERENCE

Z-Scores

It helps to understand the relative position of a value in a distribution.

Good at group to group comparison and spotting outliers.

Normalization

Comparing and ranking of values on the same axis.

Good at visualizing highly dimensional data and bring forward trends & patterns.

Using Prep

Normalize data pre-vizzing, to reduce complexity and potential load times.

Keeps the baseline the same regardless of filters.

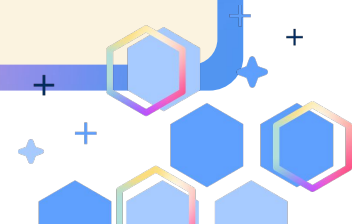


TABLEAU
CONFERENCE
Thank you

TABLEAU CONFERENCE

Coffee's on us.

The first 1,000 DataFam to provide feedback on this event will receive a $5 Starbucks gift card.*

1. Download the Salesforce Events app.
2. Scan the QR code or navigate to My Event, then My Surveys.
3. Take four session surveys and the event survey.
4. Redeem your gift card at Badge Pickup on Thursday.

*Restrictions apply. See terms and conditions at sforce.co/survey-terms.

TABLEAU CONFERENCE