



UNIVERSITY OF
CANBERRA

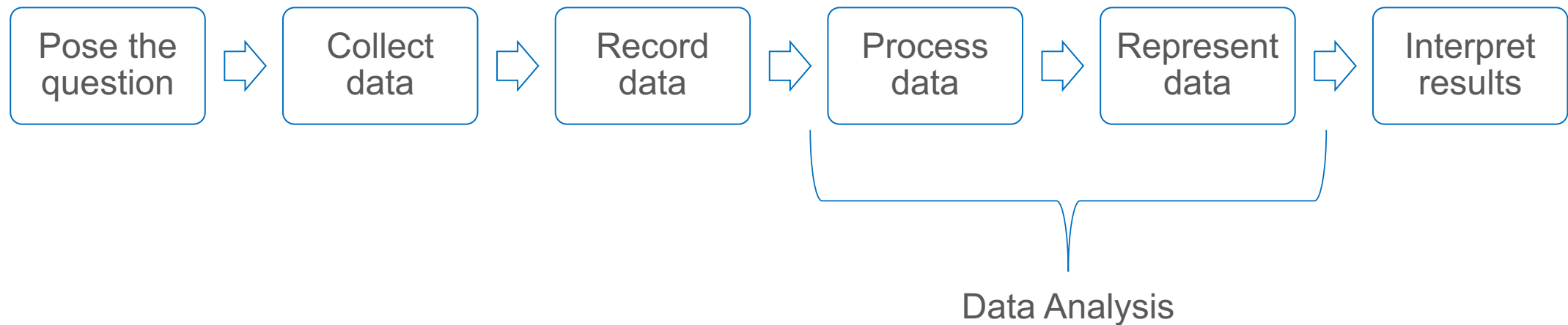
Getting started with data analysis

Jocelyn Mara

Discipline of Sport and Exercise Science

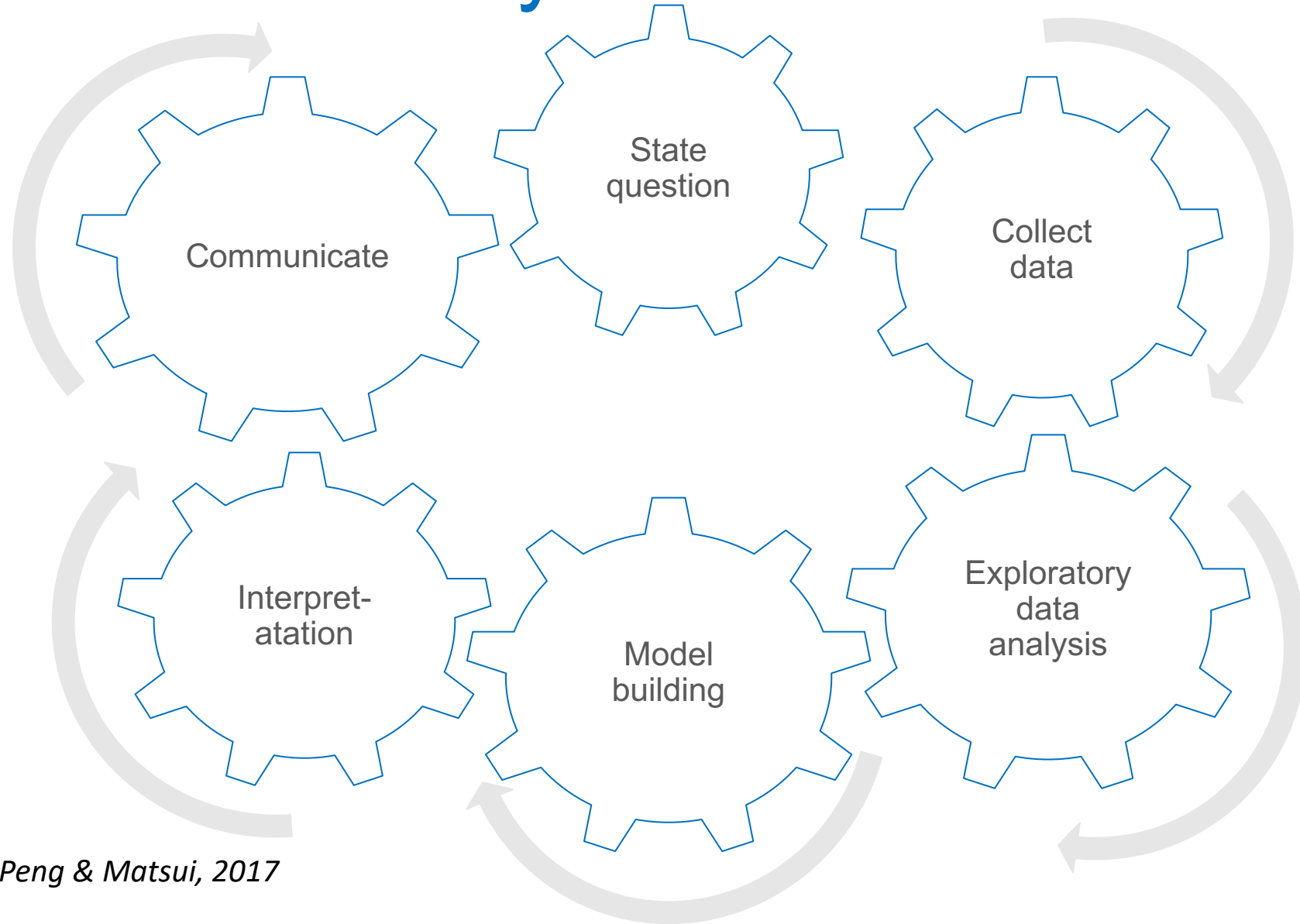


Stages of Analysis





“Epicycles” of Analysis



Types of Questions¹

1. Descriptive
2. Exploratory
3. Inferential
4. Predictive
5. Causal
6. Mechanistic

Descriptive

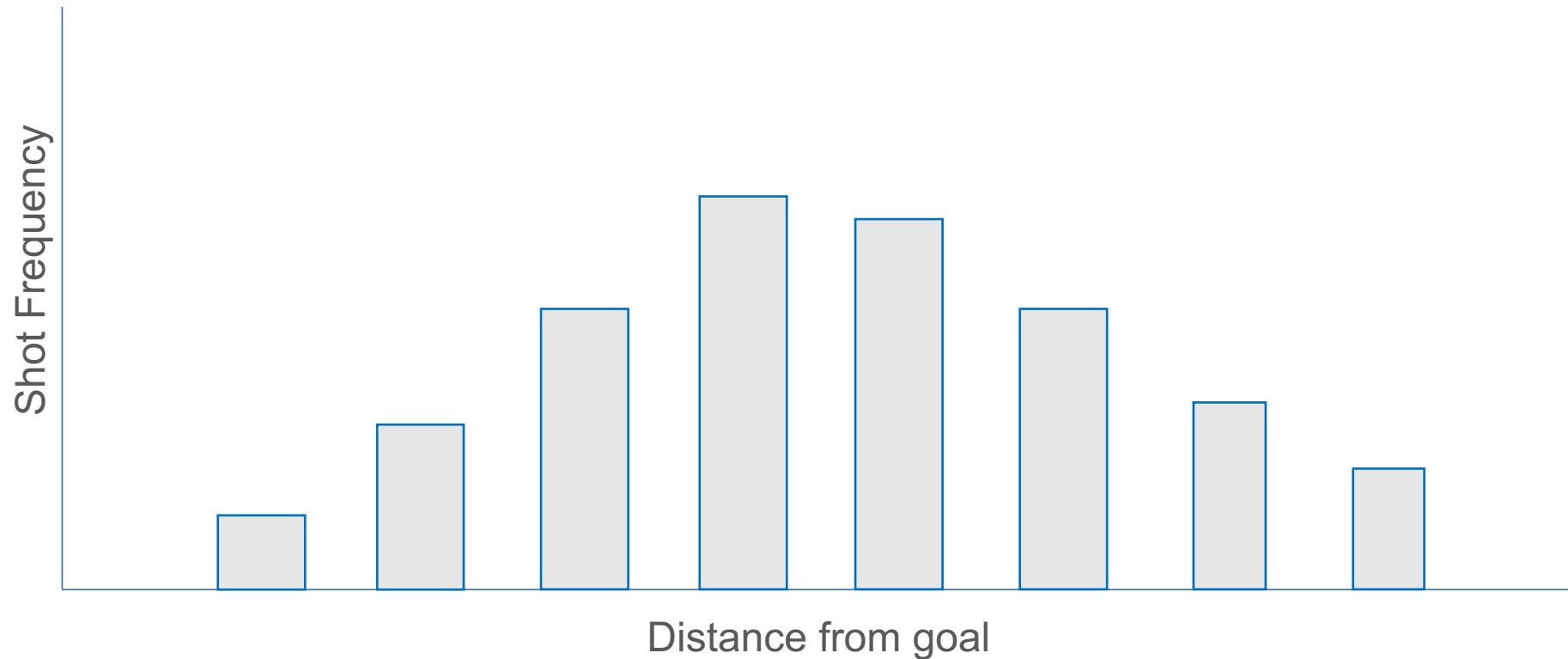
- Aims to summarise a characteristic of a set of data¹

Goals	Shots at Goal	% Goals
34	324	10.4

Mean shots per game	SD	Min - Max
9.5	2.1	5 - 20

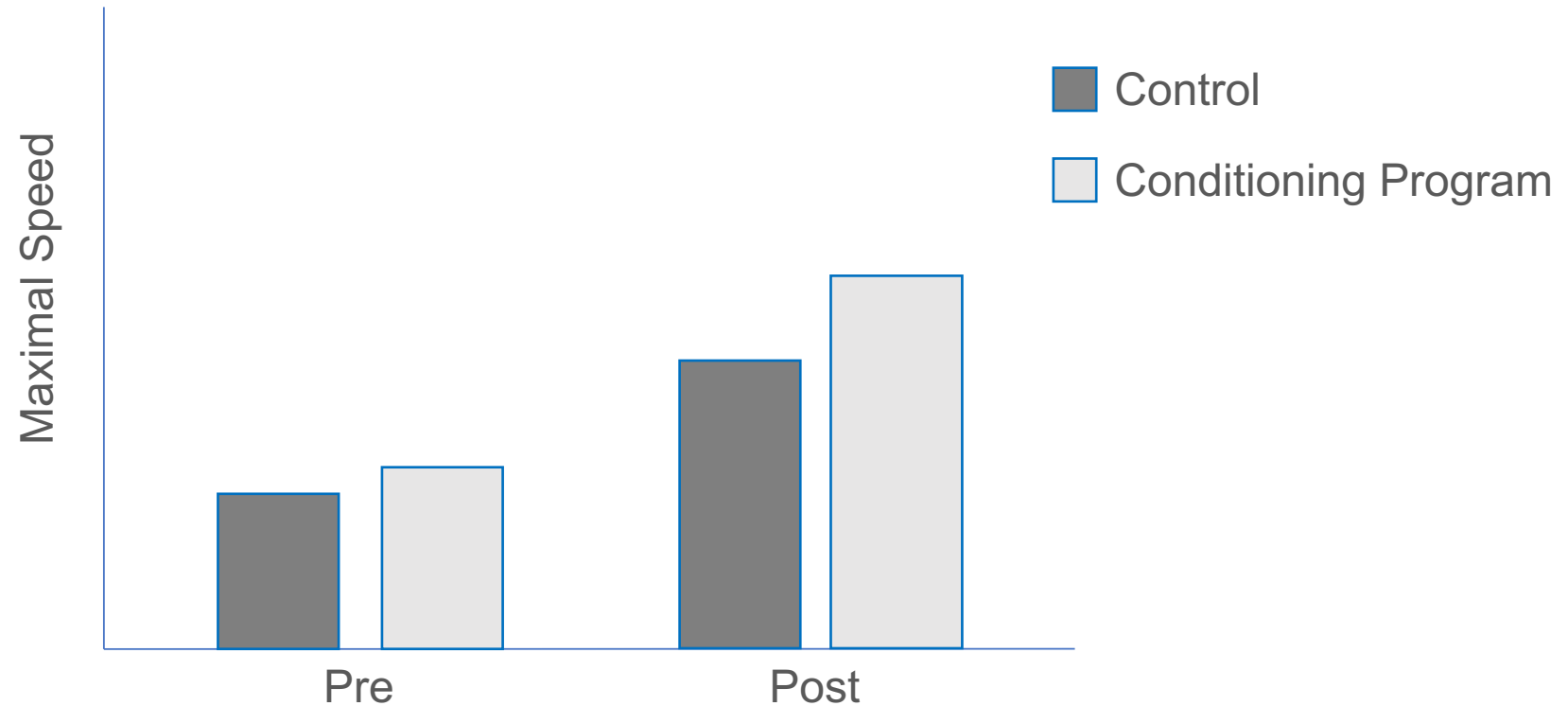
Exploratory

- Analyse data to determine patterns, trends or relationships¹



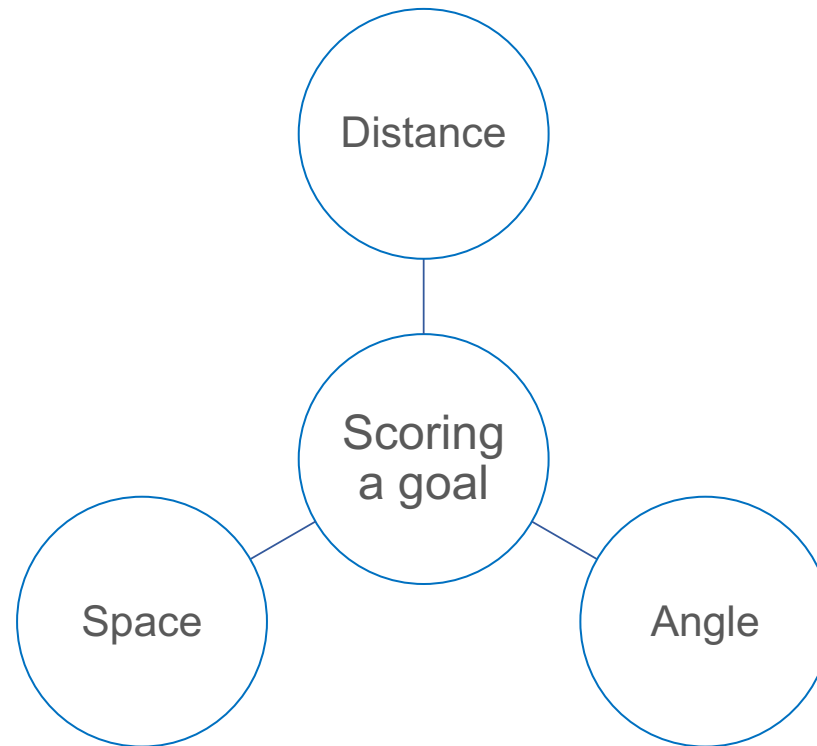
Inferential

- Seeks to determine whether results you find in a sample are representative of the population



Predictive

- Aims to predict an outcome based on particular variables

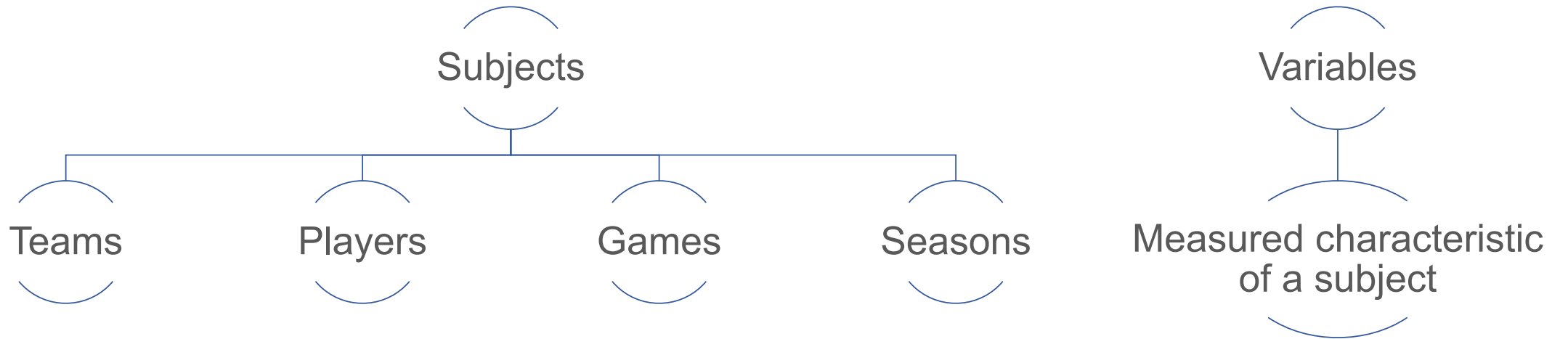


Why is this important?

- Because the type of analysis you conduct is determined by what question you are trying to answer.....

.....And what *types of data* you have

Sports data



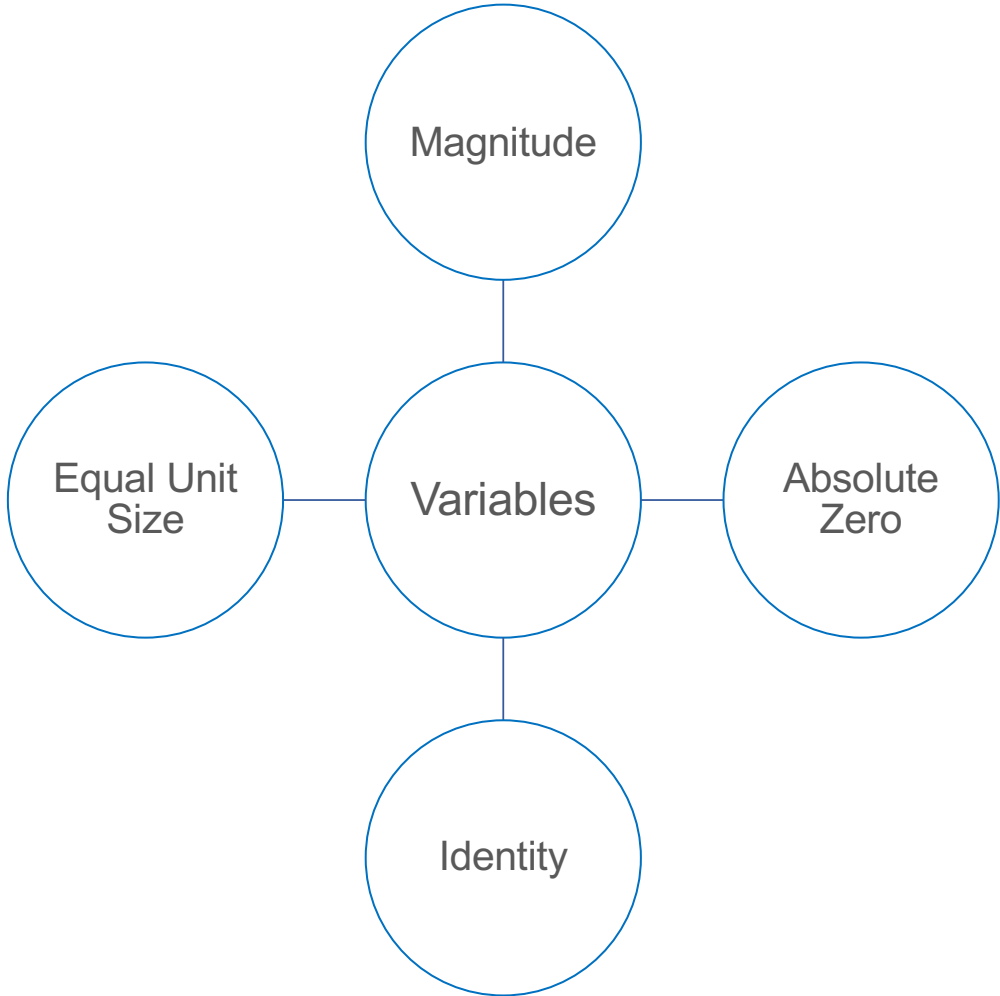
Sports data

Name	Team	Position	Games Played	Minutes Played	Field Goals Attempted	Threes Made
Aaron Afflalo	DEN	SG	82	2228	585	108
Alexis Ajinca	CHA	C	6	29	10	0
Lamarcu Aldridge	POR	PF	78	2919	1169	5
Joe Alexander	DEN	SG	8	30	6	0

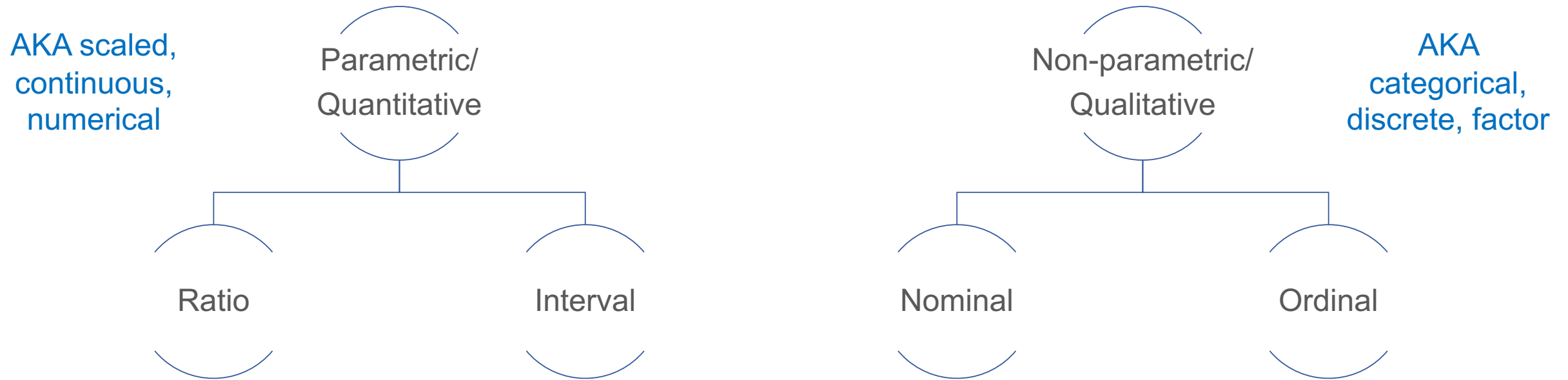
Sports data

Team	Games Played	Field Goals Made	Field Goals Attempted	Threes Made	Threes Attempted
ATL	62	227	485	37	103
BOS	56	206	426	34	97
CHA	57	195	429	29	86
CHI	56	215	479	25	77

Variable properties



Scales of measurement



Scales and properties of measurement

Scale of Measurement	Example	Properties	Mathematical Operations
Ratio	Time Distance Speed Frequencies	Identity Magnitude Equal Unit Size Absolute Zero	Add Subtract Multiply Divide
Interval	Temperature	Identity Magnitude Equal Unit Size	Add Subtract Multiply Divide
Nominal	Pitch Type Playing Position Shot Outcome	Identity	None
Ordinal	RPE Wellbeing	Identity Magnitude	Rank Order

Operational Definitions

Variable	Scale of Measurement	Operational Definition
Speed	Ratio	Metres per second
Distance	Ratio	Metres travelled
Distance	Ordinal	Short = 0-10m Medium = 10-20m Long = >20m
Shot Type	Nominal	F = Forehand B = Backhand S = Smash V = Volley

Something to consider...

Runs (Ordinal)
No Hit (NH)
Single (S)
Double (D)
Triple (T)
Home Run (HR)

Runs (Ratio)
0
1
2
3
4

Raw vs summarised data

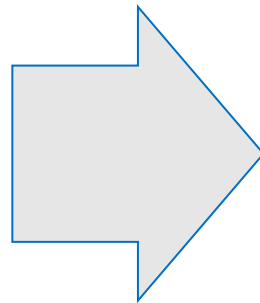
Pitch #	Runs
1	NH
2	NH
3	S
4	HR

Team	NH	S	D	T	HR
KAN	139	1016	300	42	139
MIN	156	872	277	44	156
CLE	141	922	303	29	141
CHI	136	958	260	27	136

Independent and dependent variables

Influences

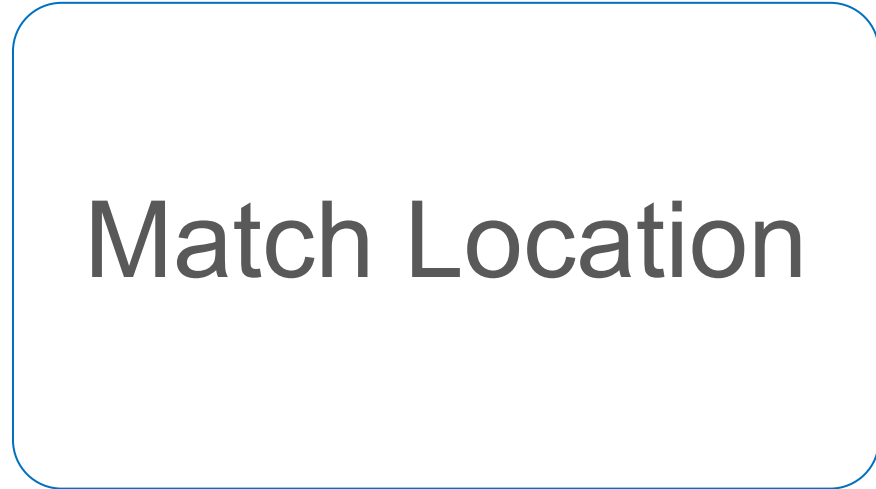
Independent
Variable



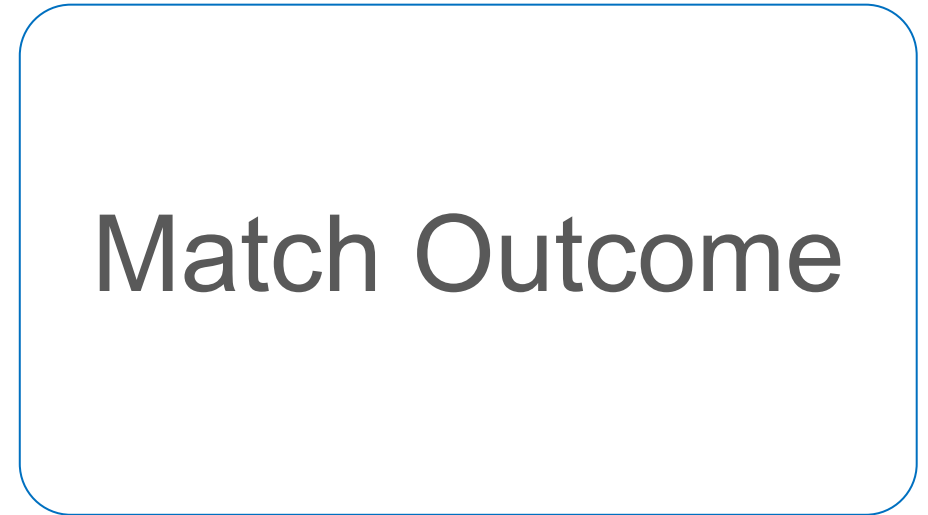
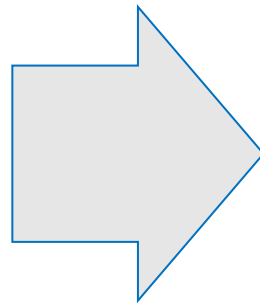
Dependent
Variable

Independent and dependent variables

Influences



Independent
Variable

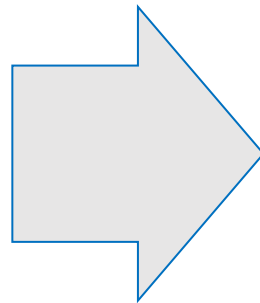


Dependent
Variable

Independent and dependent variables

Influences

Distance from
goal



Shot Outcome

Independent
Variable

Dependent
Variable

Check your understanding...

- Descriptive, exploratory, inferential and predictive questions
- Properties of measurement
- Scales of measurement
- Independent and dependent variables