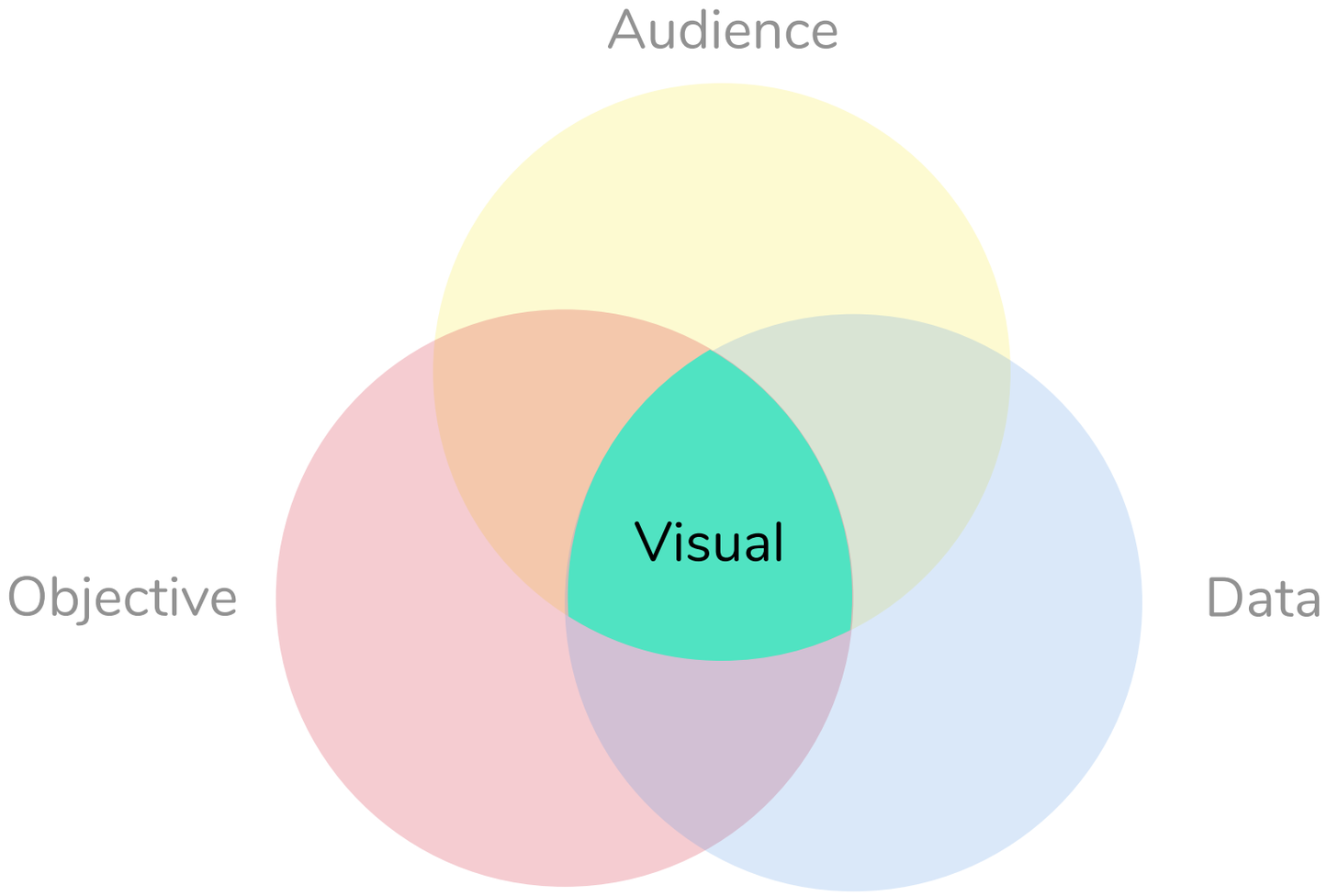
Design Context

What should we know before we visualise?



1. Who is your audience?
2. What is the objective?
3. What is the data?

1. Who is your audience?

2. What is the objective?

3. What is the data?

Sub-questions?

Who is your audience?

- What is their role?

  General public, Trainee, CEO, Scientist

- What is their motivation?

  Entertainment, Money, Couldn’t care less

- What is their prior knowledge?

  None, Fleeting, Subject expert

- How often will they use it?

  Once, Once a decade, Sometimes, Often, Bread/Butter

- Where will they view it?

  Print, Web, Presentation

What is the objective?

- What are you trying to do?

  Explore, Explain, Persuade

- How do you want to achieve this?

  Factual, Emotional, Opinionated

- What are the constraints?

  Time, Money, Resources

- How can you measure the impact?

  How do we know we achieved the objective?

What is the data?

- How big is the data?

  I can read all values in 30 seconds, Distributed; Count dimensions

- What class of data is it?

  Diagram, Network, Spatial

- Can I trust the data?

  Who owns it, how did it get collected, how clean is the data

- Does it change?

  Static vs dynamic data, how often does it change, do we need to reflect the change?

Discussion

What data do you most regularly come in contact with?

Who do you analyse this data for?

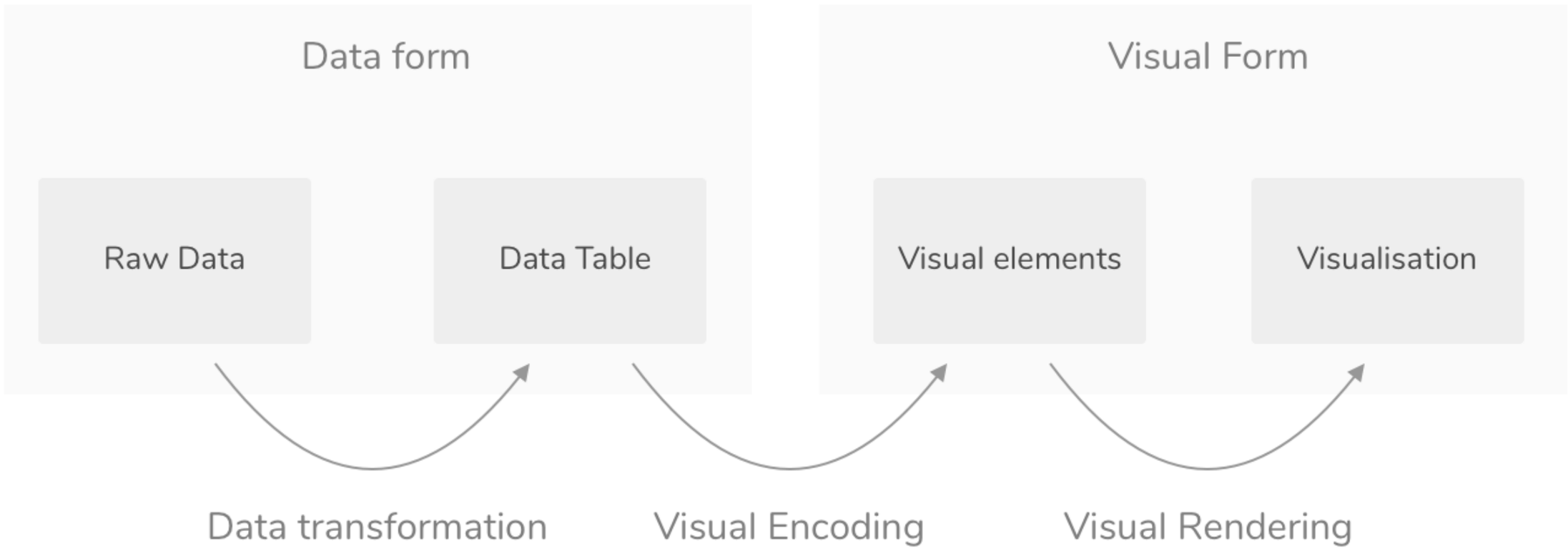
What do you have to or want to convey with the results?

Discover, Describe, Persuade?

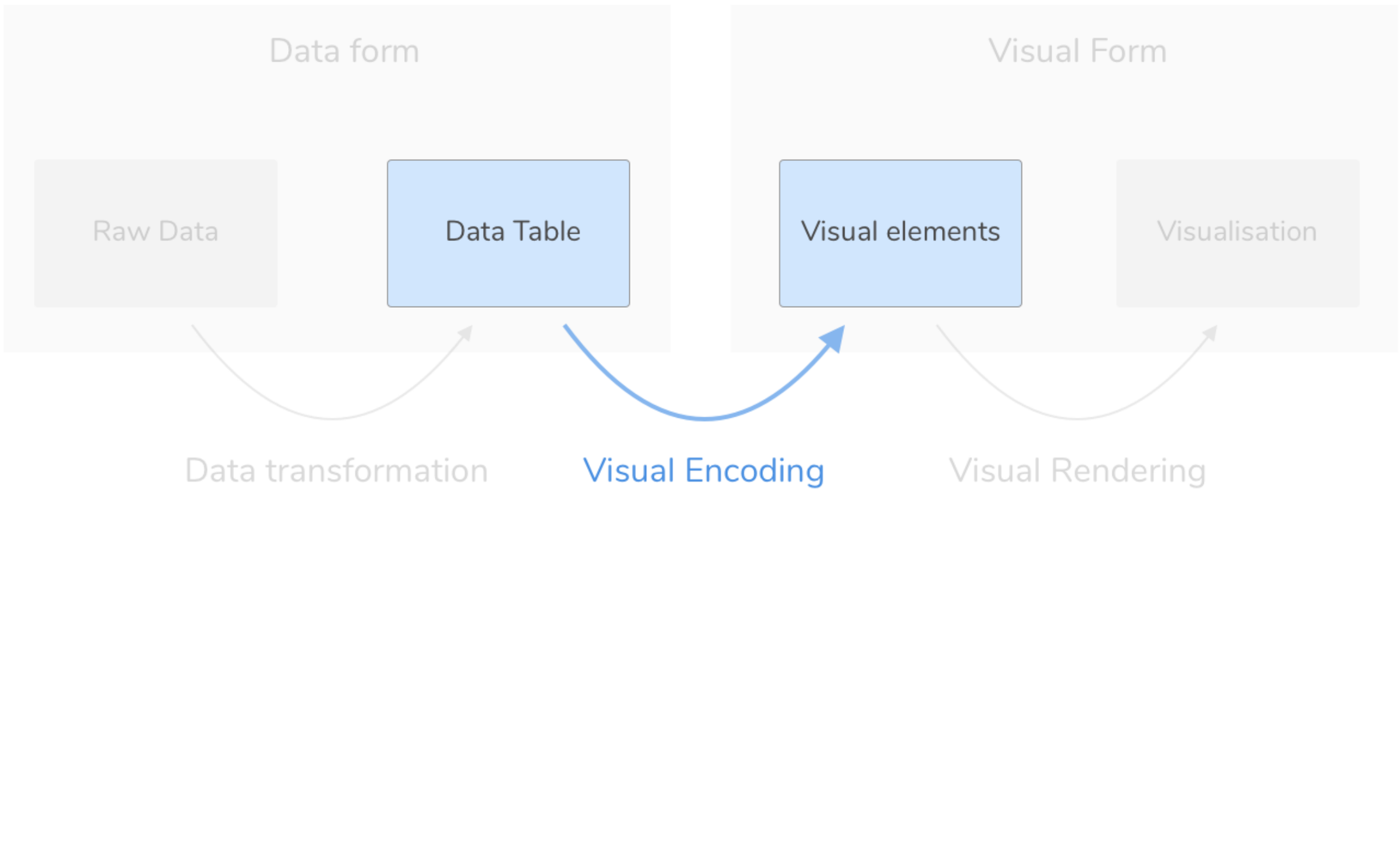
Visual Encoding

Mapping **data** elements to **visual** elements

The visualisation pipeline



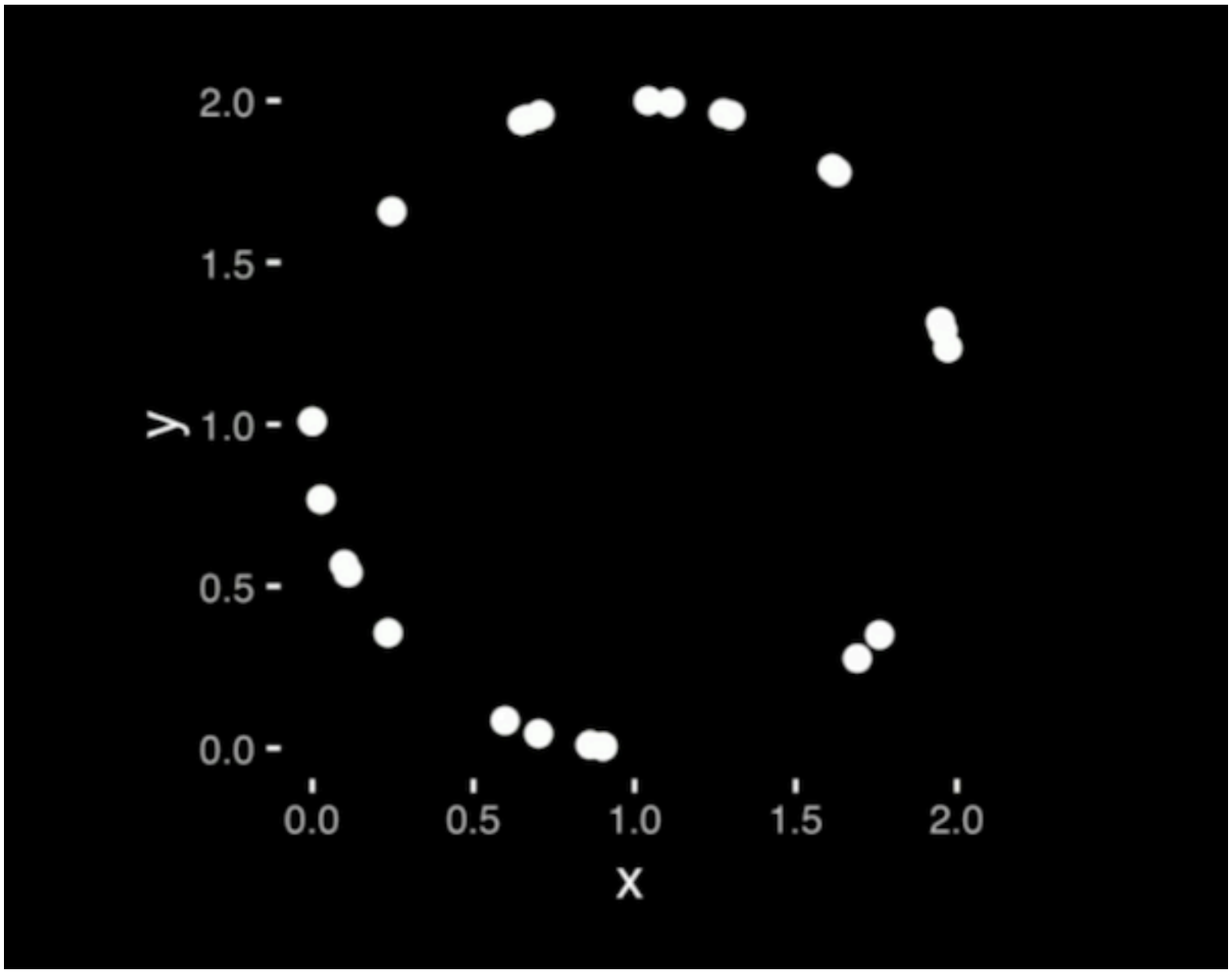
The visualisation pipeline

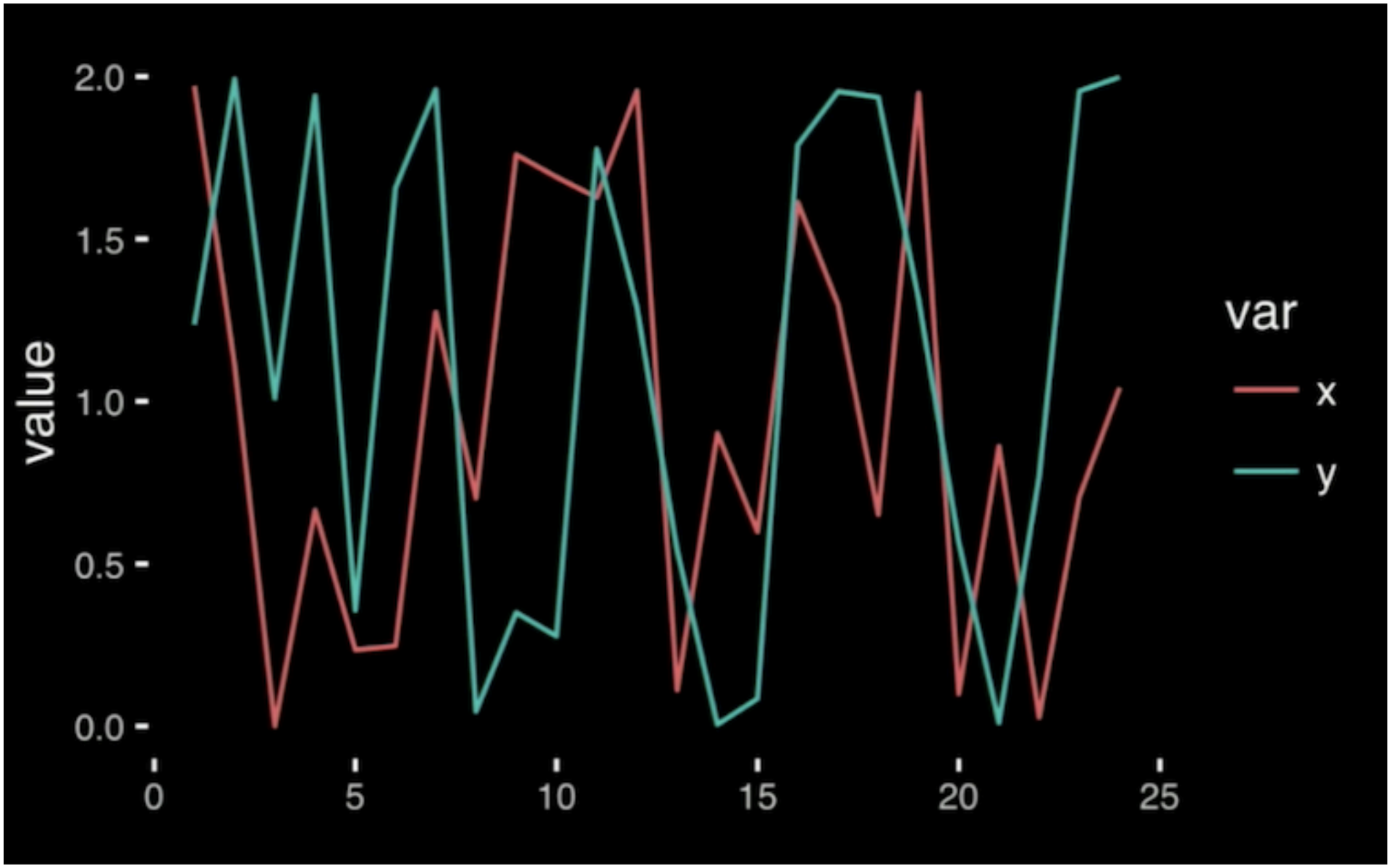


This is the design part.

Why?

x	y	x	y
1.972	1.236	0.111	0.542
1.112	1.994	0.902	0.005
0.000	1.009	0.598	0.085
0.665	1.942	1.613	1.790
0.235	0.356	1.298	1.955
0.247	1.658	0.651	1.937
1.275	1.961	1.949	1.316
0.702	0.045	0.099	0.567
1.760	0.350	0.862	0.010
1.691	0.277	0.027	0.768
1.628	1.778	0.706	1.956
1.957	1.290	1.042	1.999


2.0
1.5
1.0
0.5
0.0
Y
0.0
0.5
1.0
1.5
2.0
X


2.0
1.5
1.0
0.5
0.0
value
0
5
10
15
20
25
var
x
y

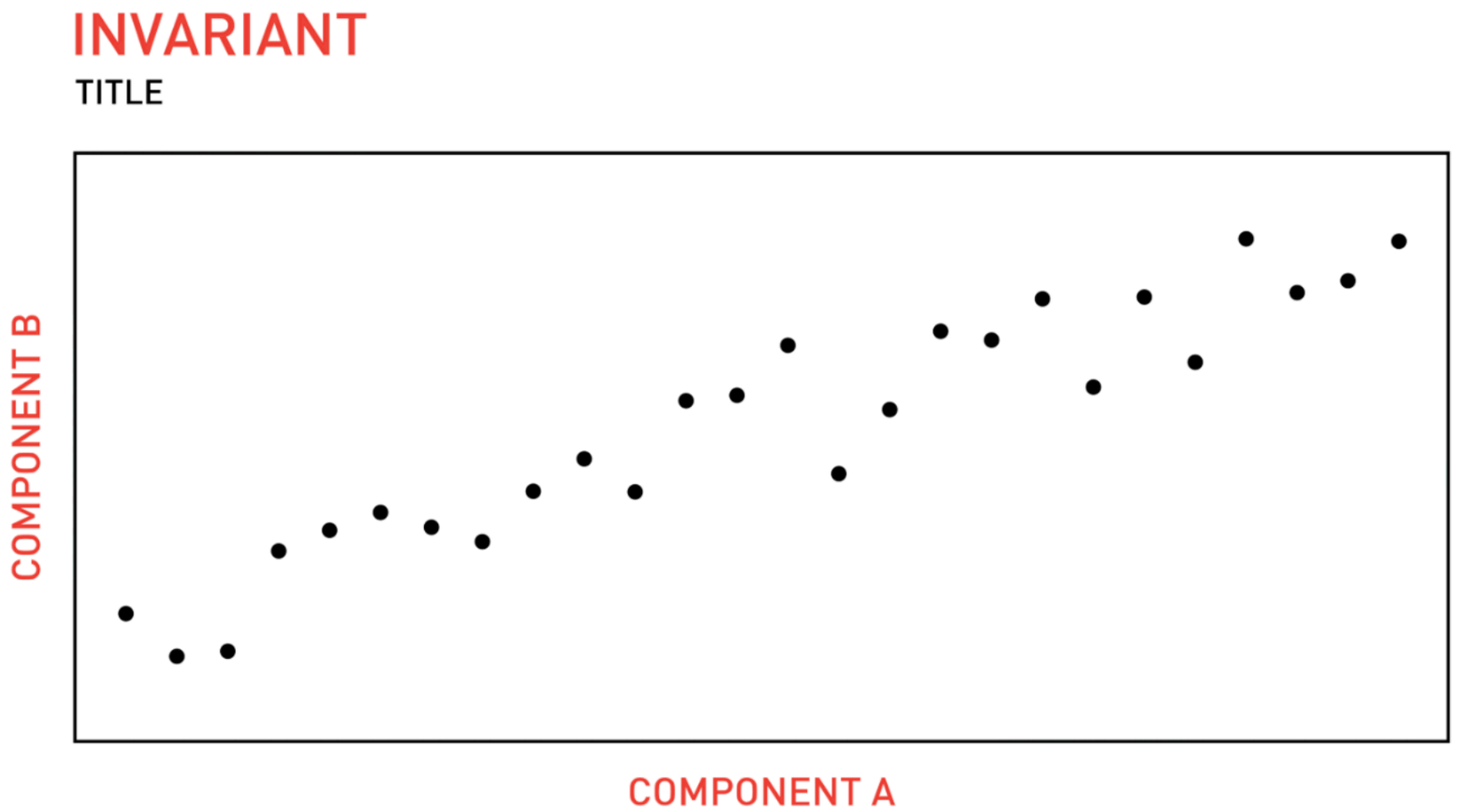
What to encode?

What is data?

Data = Data Object + Data Attributes

Data = Data Object + Data Attributes

Invariant Component



Data = Data Object + Data Attributes

id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9
30	Netta	Female	22	25	BASIC	West Midlanc	2
31	Wallie	Female	34	38	BASIC	West Midlanc	2
32	Jules	Male	21	21	BASIC	Scotland	1
33	Hermia	Female	27	20	BASIC	London	2

Data = Data Object + Data Attributes

id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9
30	Netta	Female	22	25	BASIC	West Midlanc	2
31	Wallie	Female	34	38	BASIC	West Midlanc	2
32	Jules	Male	21	21	BASIC	Scotland	1
33	Hermia	Female	27	20	BASIC	London	2

Objective I want to see the relationship between Age and Revenue

Data Object ?

Data Attributes ?

Data = Data Object + Data Attributes

id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9
30	Netta	Female	22	25	BASIC	West Midlanc	2
31	Wallie	Female	34	38	BASIC	West Midlanc	2
32	Jules	Male	21	21	BASIC	Scotland	1
33	Hermia	Female	27	20	BASIC	London	2

Objective I want to see the relationship between Age and Revenue

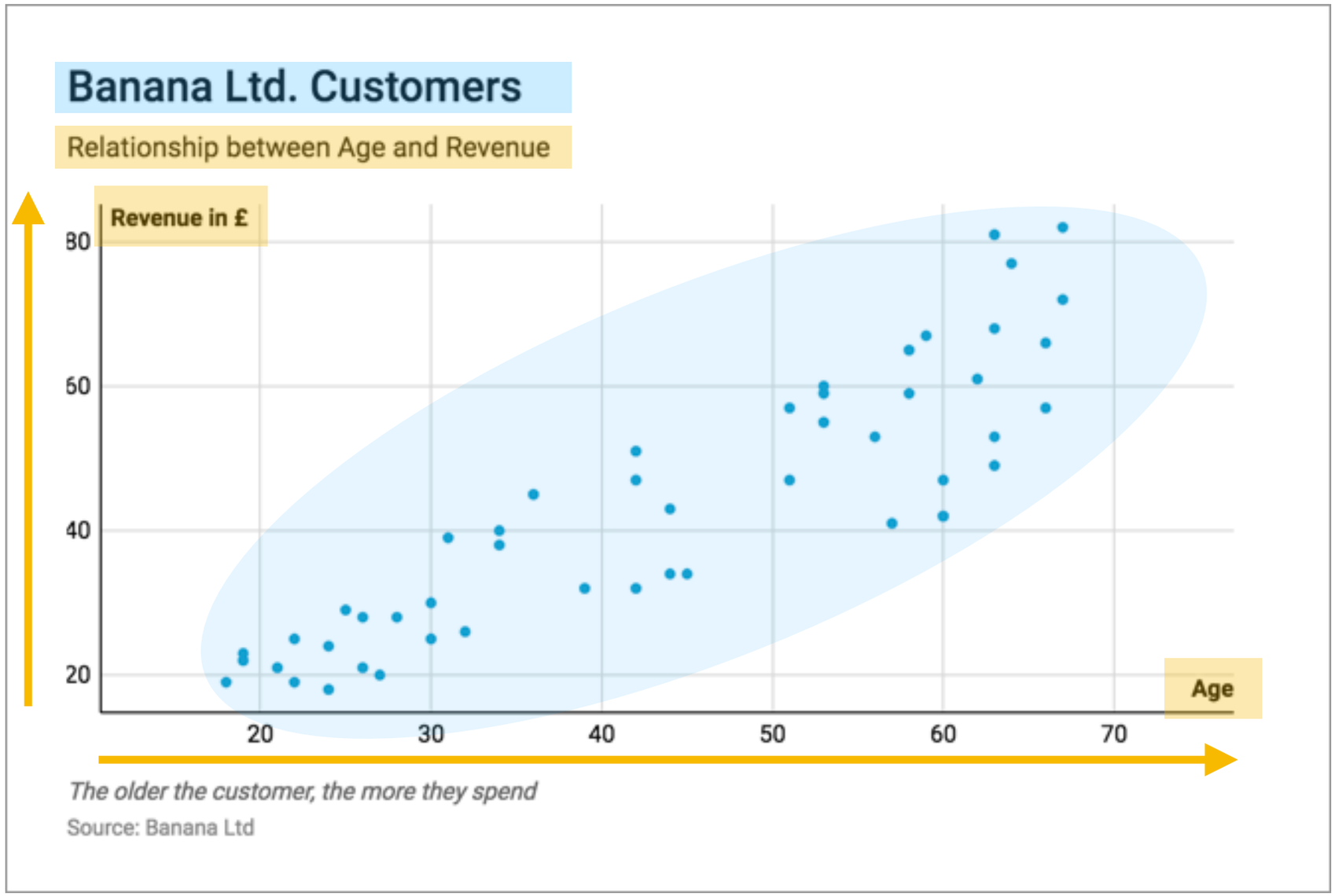
Data Object customers

Data Attributes age, revenue

Data = Data Object + Data Attributes

Data Object customers

Data Attributes age, revenue



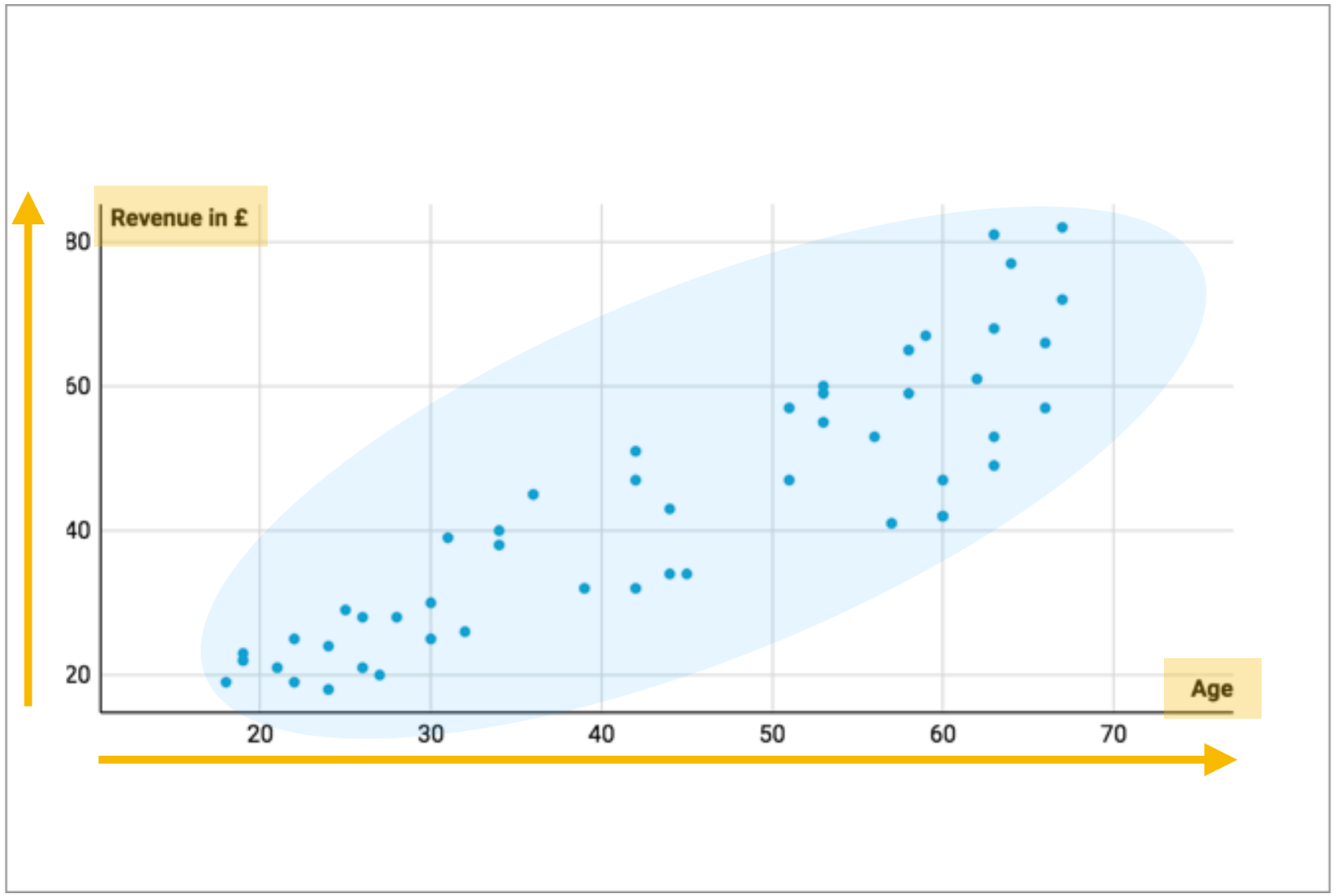
Marks and channels

Data Object customers

Data Attributes age, revenue

inform *Visual Elements* == **Marks**

inform *Visual States* == **Channels**



Marks: Points, Lines and Areas

Channels: Size, Position, Colour, ...

Marks and channels

Attributes → Channels Size, Position, Colour...

id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9

Rows → Marks

Points, Lines, Areas

How do we pick the right channel?

Visual channels relate to **data types**.

Rule #1 Follow the **expressiveness principle**

Data types

	Description	Examples
Categorical data	Different categories, but no order	Gender, Ethnicity, Type of Fruit
Ordered data	Different categories in an order without a meaningful difference	Medals, Star ratings, Age groups
Quantitative data	Ordered values with a meaningful difference	Age, Income, Size, Weight

Data types ?

id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9
30	Netta	Female	22	25	BASIC	West Midlanc	2

Data types ?

Ordered	Categ.	Categ.	Quant	Quant	Ordered	Categ.	Ordered*
id	first_name	gender	age	revenue	subscription	region	rating
1	Dorotea	Female	34	40	SILVER	London	4
2	Arthur	Male	63	68	GOLD	Yorkshire	8
3	Tim	Male	51	47	SILVER	London	4
4	Billie	Female	19	22	BASIC	Wales	1
5	Jarvis	Male	66	57	SILVER	NI	5
6	Rudie	Male	63	81	GOLD	East Midlands	8
7	Allistair	Male	24	18	BASIC	London	3
8	Sarah	Female	42	51	SILVER	Wales	4
9	Pia	Female	26	21	BASIC	Scotland	3
10	Sanders	Male	19	23	BASIC	Scotland	3
11	Lorilyn	Female	60	42	SILVER	Yorkshire	4
12	Bill	Male	56	53	SILVER	Scotland	6
13	Gillan	Female	63	49	SILVER	West Midlanc	5
14	Julita	Female	51	57	SILVER	Wales	4
15	Adelaide	Female	28	28	BASIC	East Midlands	1
16	Joni	Female	18	19	BASIC	East Midlands	3
17	Aime	Female	45	34	BASIC	Scotland	2
18	Elwood	Male	25	29	BASIC	Scotland	3
19	Angie	Female	63	53	SILVER	London	4
20	Dalton	Male	62	61	GOLD	London	10
21	Fabiano	Male	58	65	GOLD	Scotland	9
22	Marice	Female	67	72	GOLD	West Midlanc	10
23	Gabriella	Female	64	77	GOLD	London	10
24	Marge	Female	60	47	SILVER	London	4
25	Kriste	Female	59	67	GOLD	London	8
26	Shina	Female	22	19	BASIC	East Midlands	2
27	Hildegarde	Female	24	24	BASIC	East Midlands	2
28	Ruby	Female	26	28	BASIC	Scotland	2
29	Shurlock	Male	66	66	GOLD	Scotland	9
30	Netta	Female	22	25	BASIC	West Midlanc	2

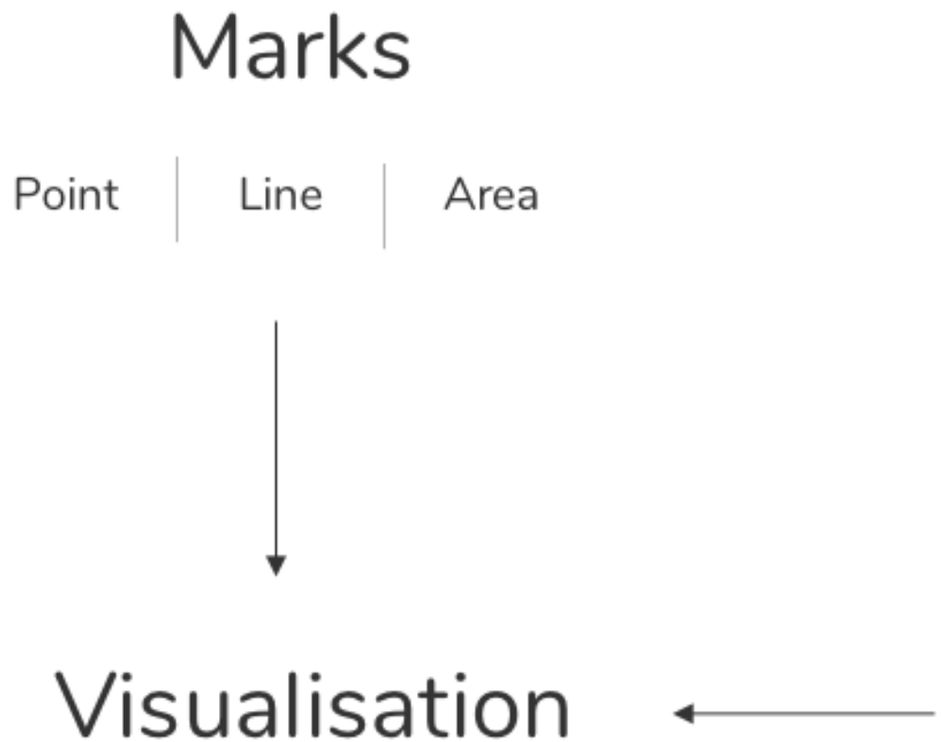
Data type ⟶ Visual channel

		Categorical	Ordered	Quantitative
Position	X	●	●	●
	Y	●	●	●
Size	Area		●	●
Colour	Luminance		●	●
	Hue	●		

Data type ⟶ Visual channel

		Categorical	Ordered	Quantitative
Position	X	●	●	●
	Y	●	●	●
Size	Area		●	●
Colour	Luminance		●	●
	Hue	●		

Data type ⟶ Visual channel



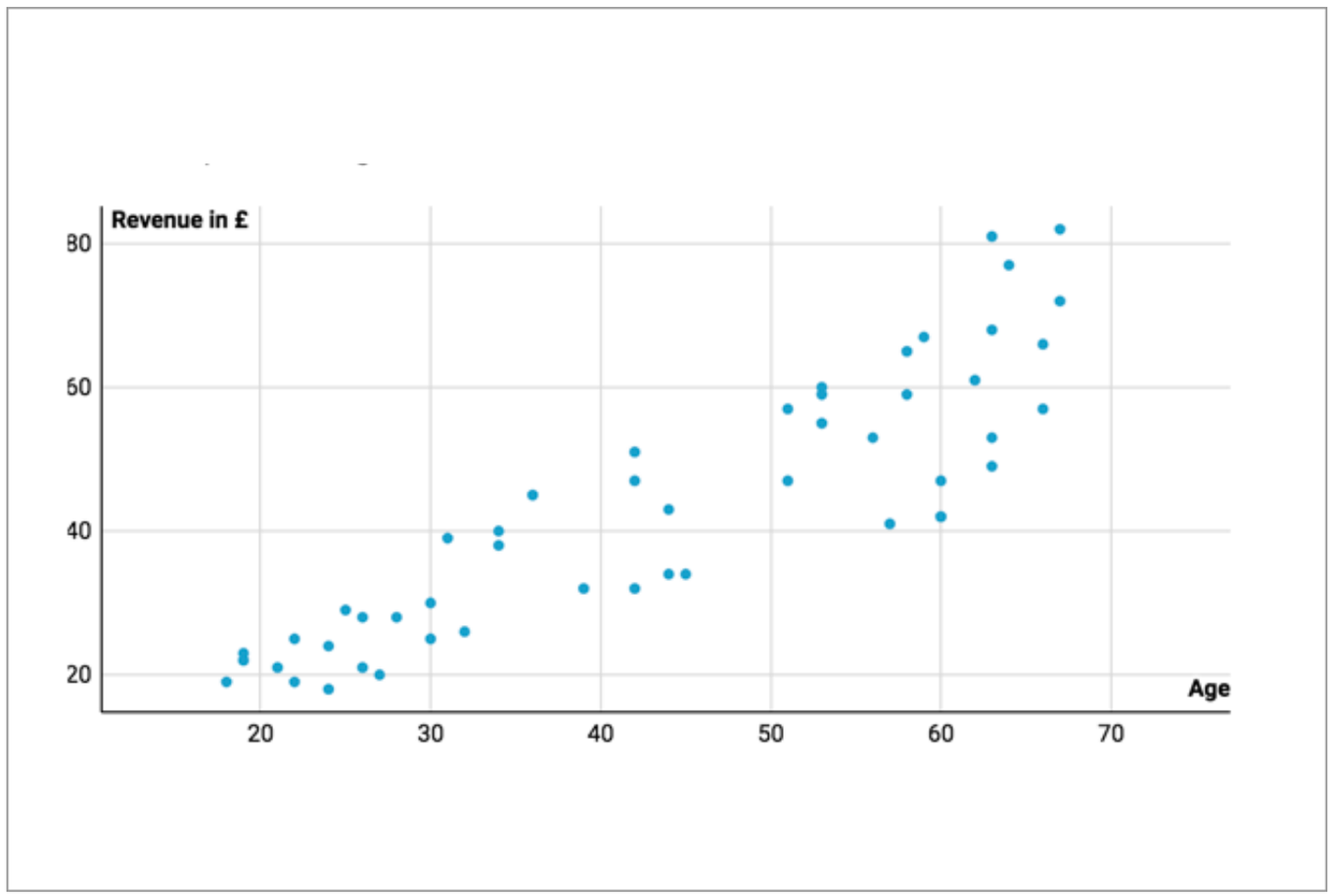
		Categorical	Ordered	Quantitative
Position	X	●	●	●
	Y	●	●	●
Size	Area		●	●
Colour	Luminance		●	●
	Hue	●		

Data Object: Banana Ltd customers 2018

Data Attributes: Revenue, Age

Data Elements: Customers

Marks: Point | Line | Area



Channels

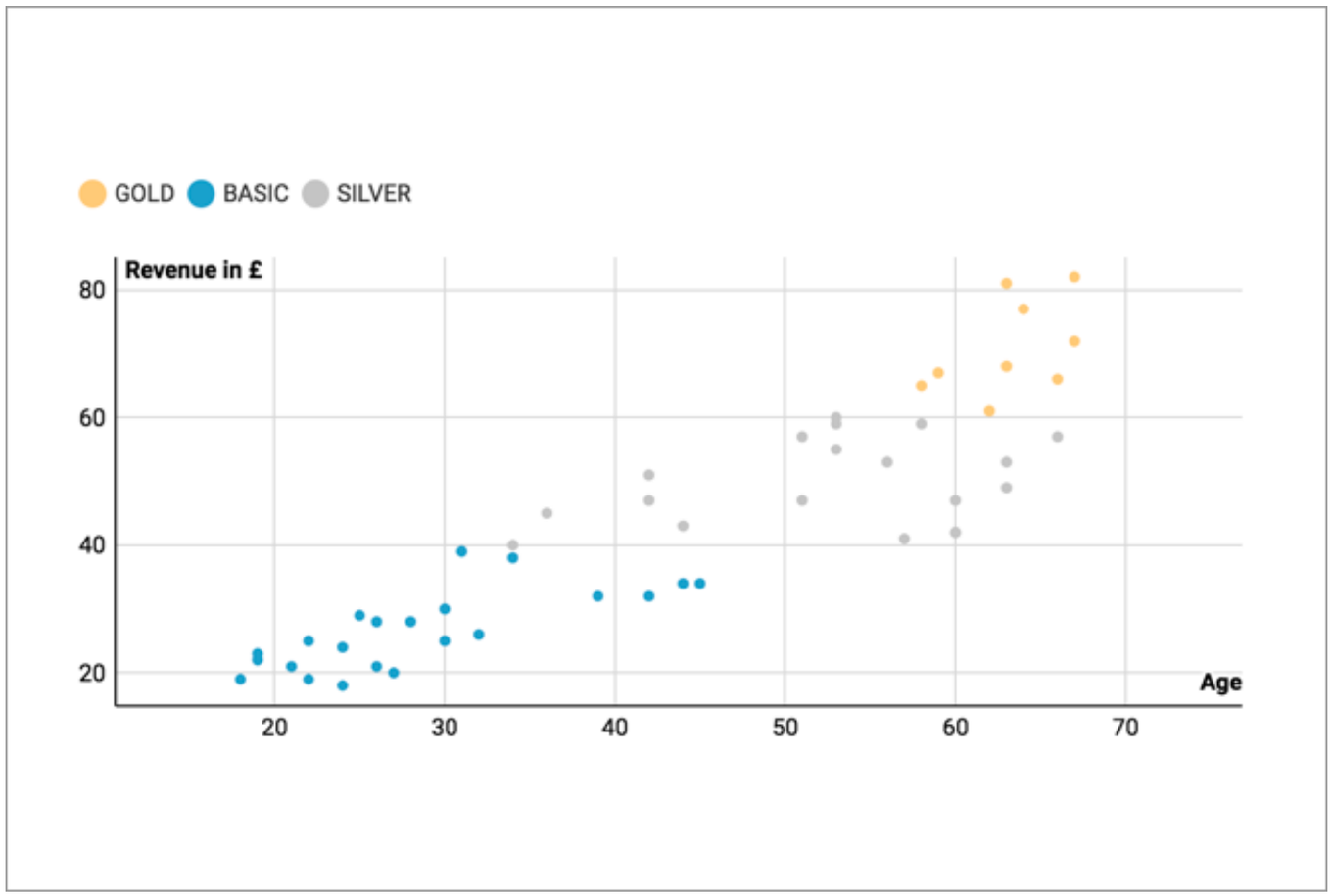
	Categorical	Ordered	Quantitative
Position X			●
Position Y			●
Area			
Colour Lum			
Colour Hue			

Data Object: Banana Ltd customers 2018

Data Attributes: Revenue, Age, Subscription

Data Elements: Customers

Marks: Point | Line | Area



Channels

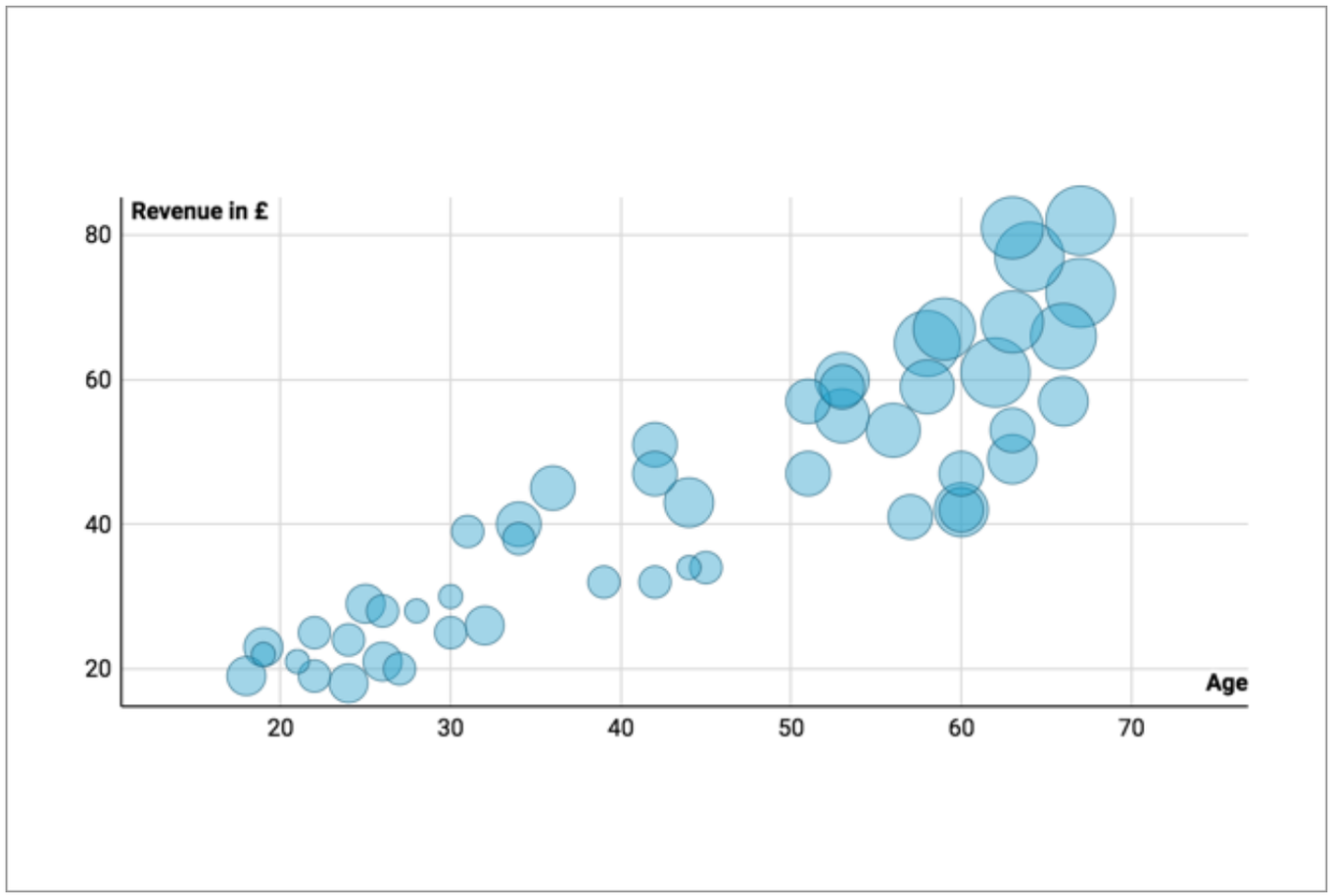
	Categorical	Ordered	Quantitative
Position X			●
Position Y			●
Area			
Colour Lum			
Colour Hue			●

Data Object: Banana Ltd customers 2018

Data Attributes: Revenue, Age, Rating

Data Elements: Customers

Marks: Point | Line | Area



Channels

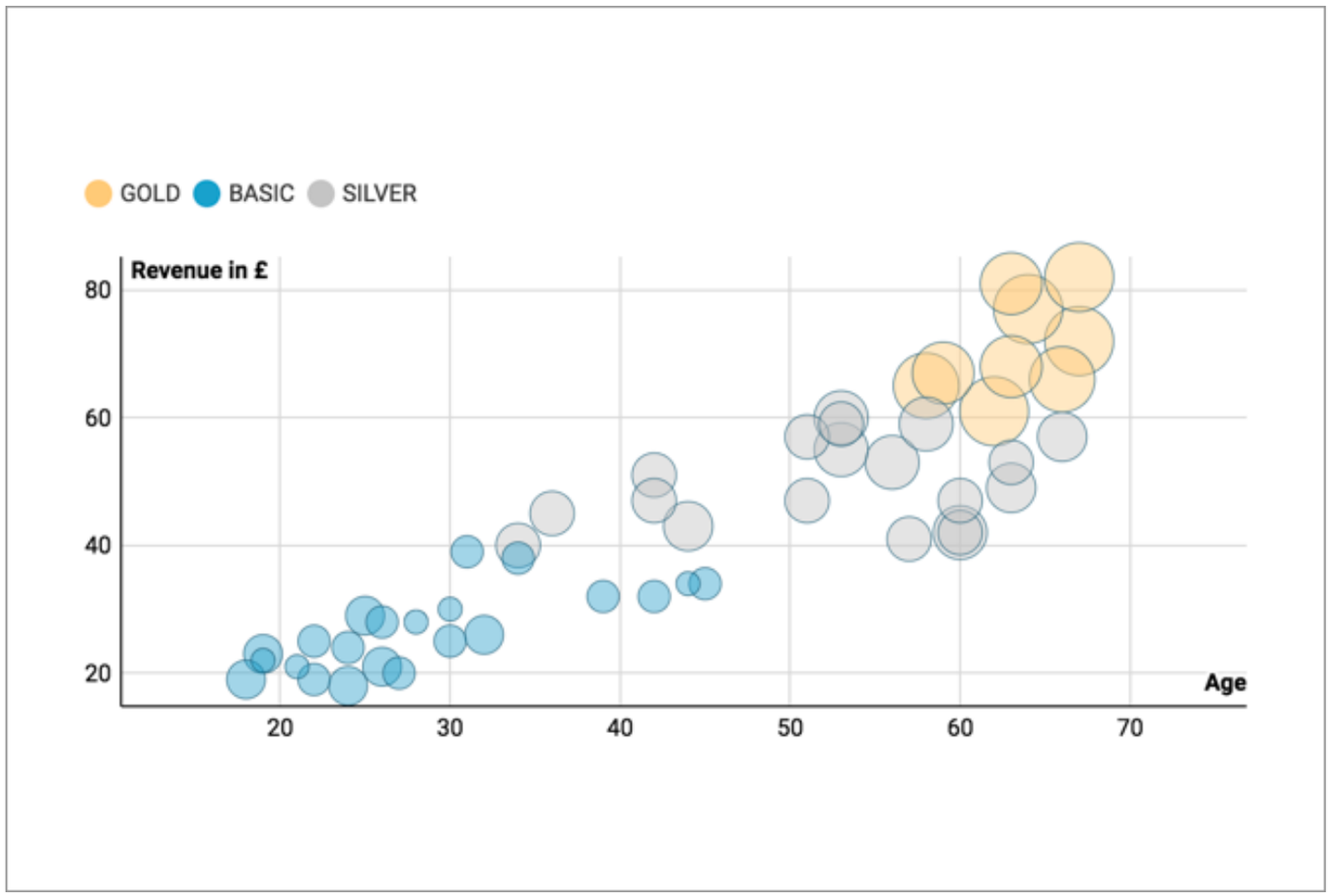
	Categorical	Ordered	Quantitative
Position X			●
Position Y			●
Area			●
Colour Lum			
Colour Hue			

Data Object: Banana Ltd customers 2018

Data Attributes: Revenue, Age, Subscription, Rating

Data Elements: Customers

Marks: Point | Line | Area



Channels

	Categorical	Ordered	Quantitative
Position X			●
Position Y			●
Area			●
Colour Lum			
Colour Hue			●

Banana Ltd Stock Price 2018

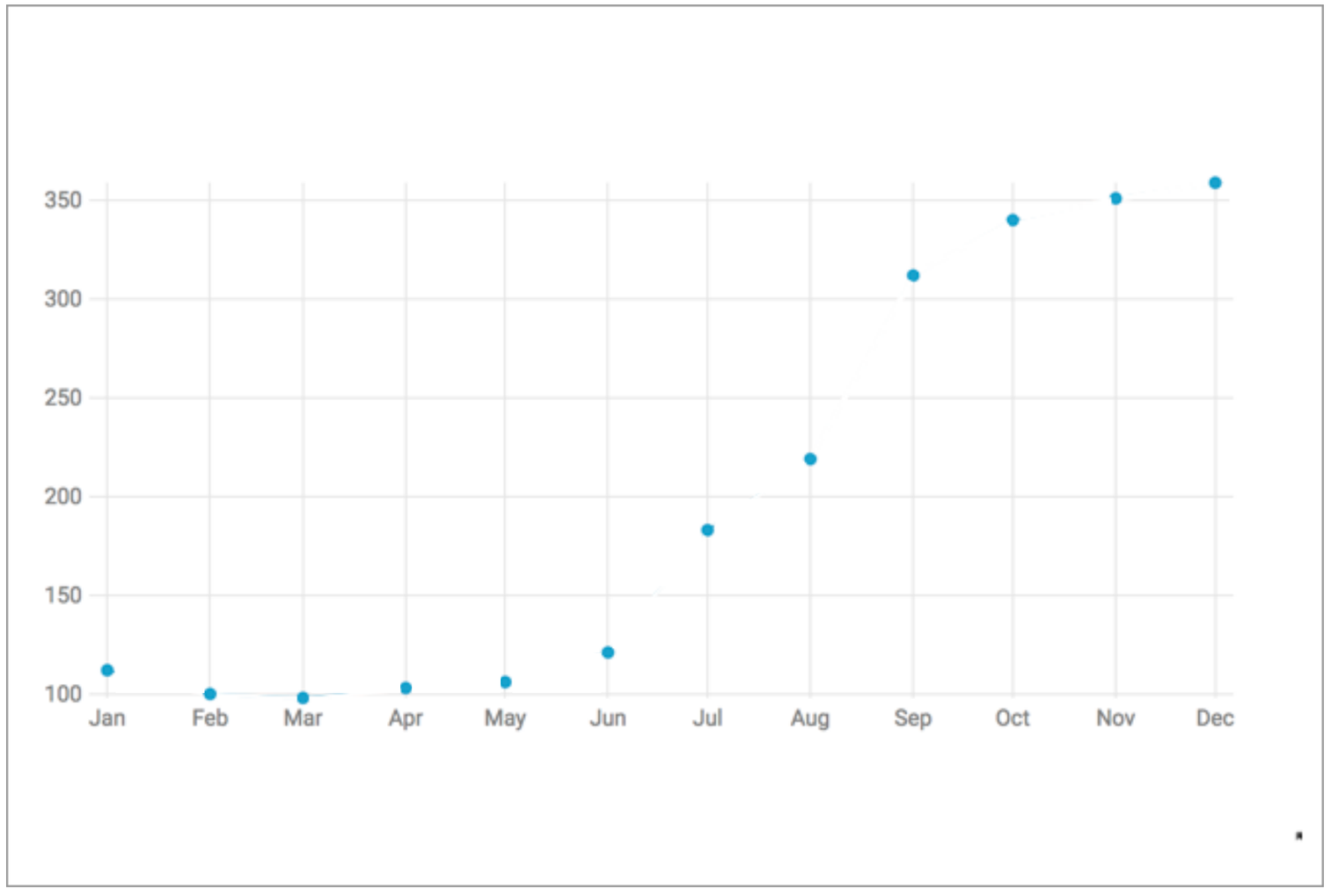
Date	Value
Jan	112
Feb	100
Mar	98
Apr	103
May	106
Jun	121
Jul	183
Aug	219
Sep	312
Oct	340
Nov	351
Dec	359

Data Object: Average monthly Banana Ltd stock quotation 2018

Data Attributes: Price, Time

Data Elements: Revenue values per month

Marks: Point | Line | Area



Channels

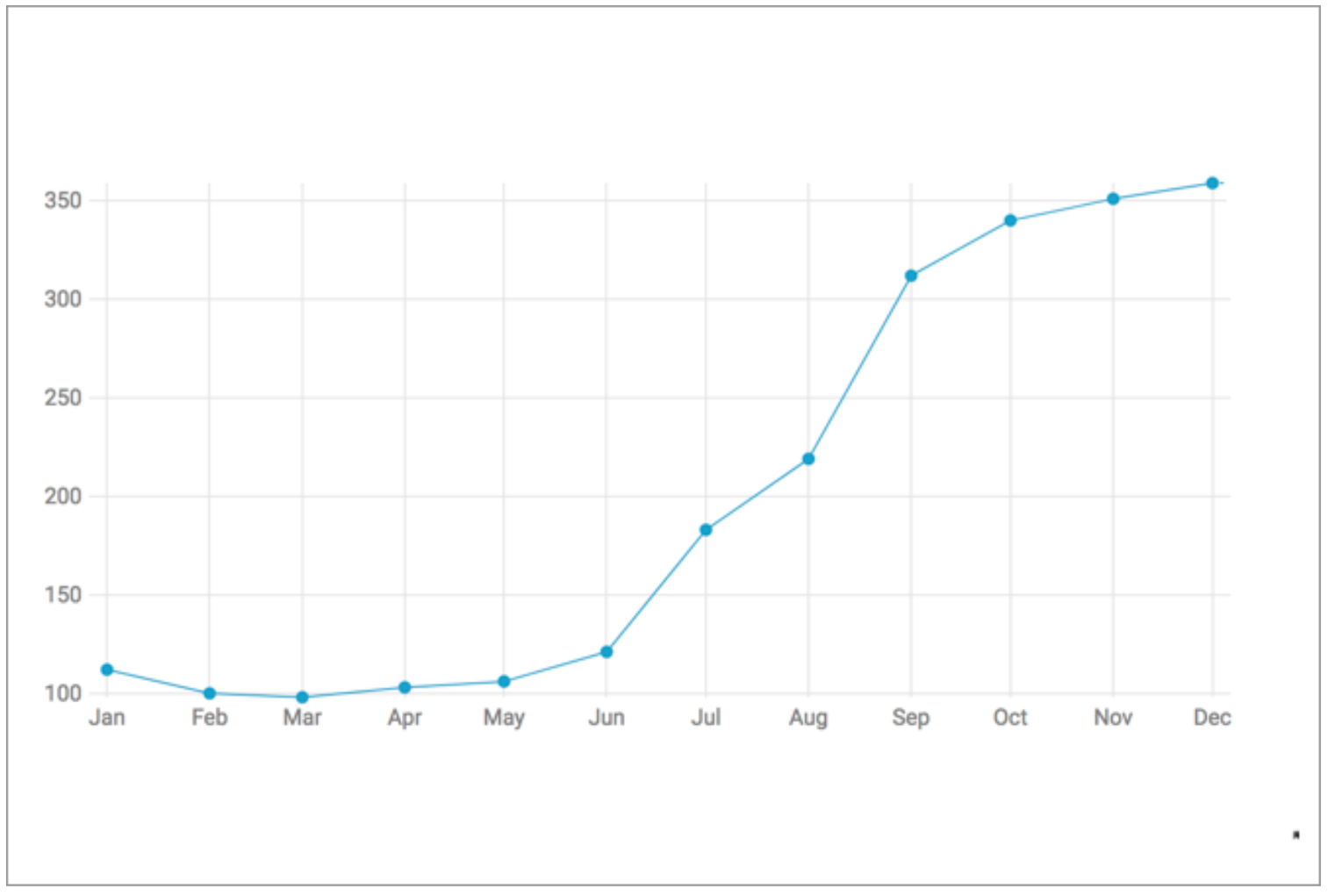
	Categorical	Ordered	Quantitative
Position X		●	
Position Y			●
Area			
Colour Lum			
Colour Hue			

Data Object: Average monthly Banana Ltd stock quotation 2018

Data Attributes: Price, Time

Data Elements: Revenue values per month

Marks: Point | Line | Area



Channels

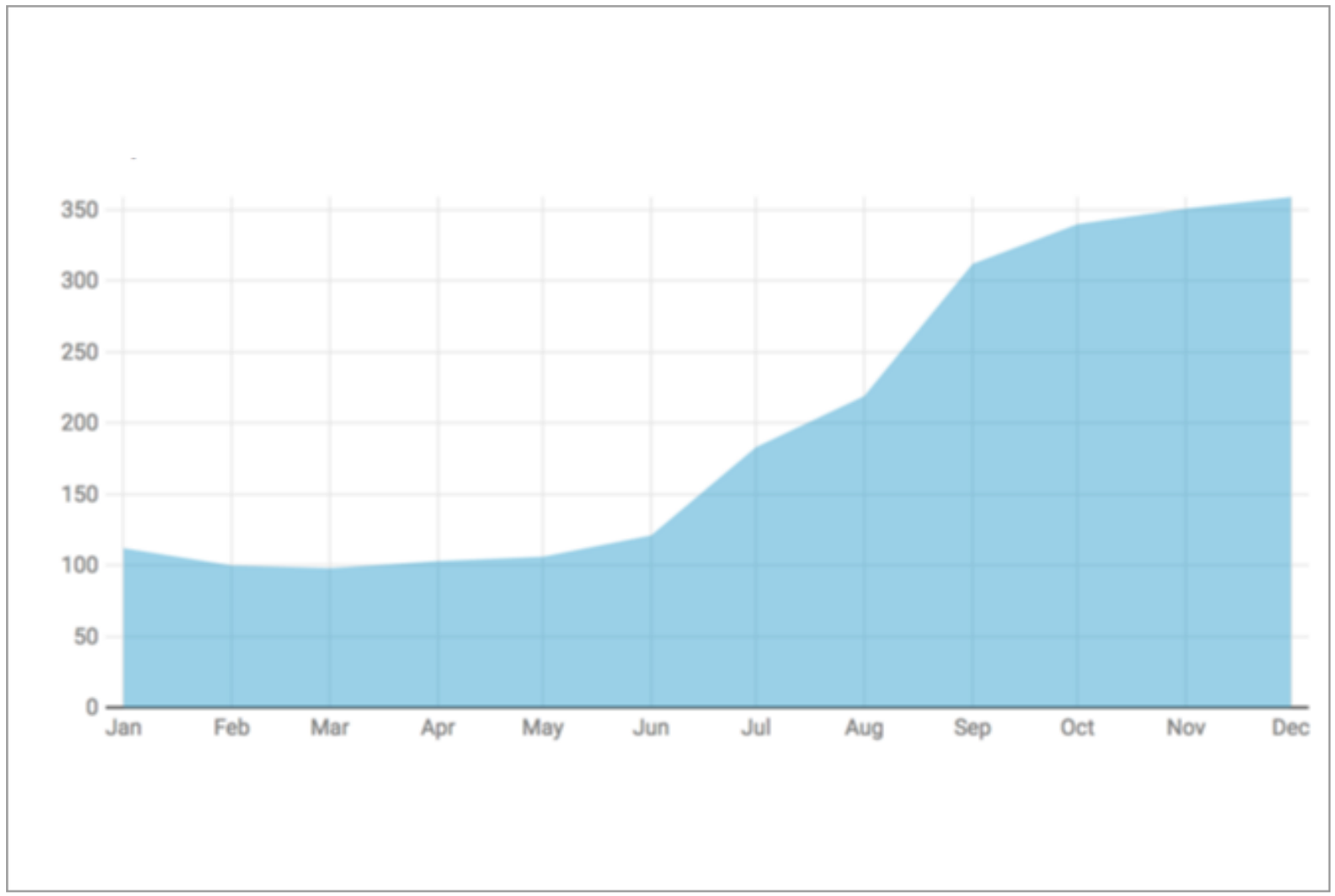
	Categorical	Ordered	Quantitative
Position X		●	
Position Y			●
Area			
Colour Lum			
Colour Hue			

Data Object: Average monthly Banana Ltd stock quotation 2018

Data Attributes: Price, Time

Data Elements: Revenue values per month

Marks: Point | Line | Area



Channels

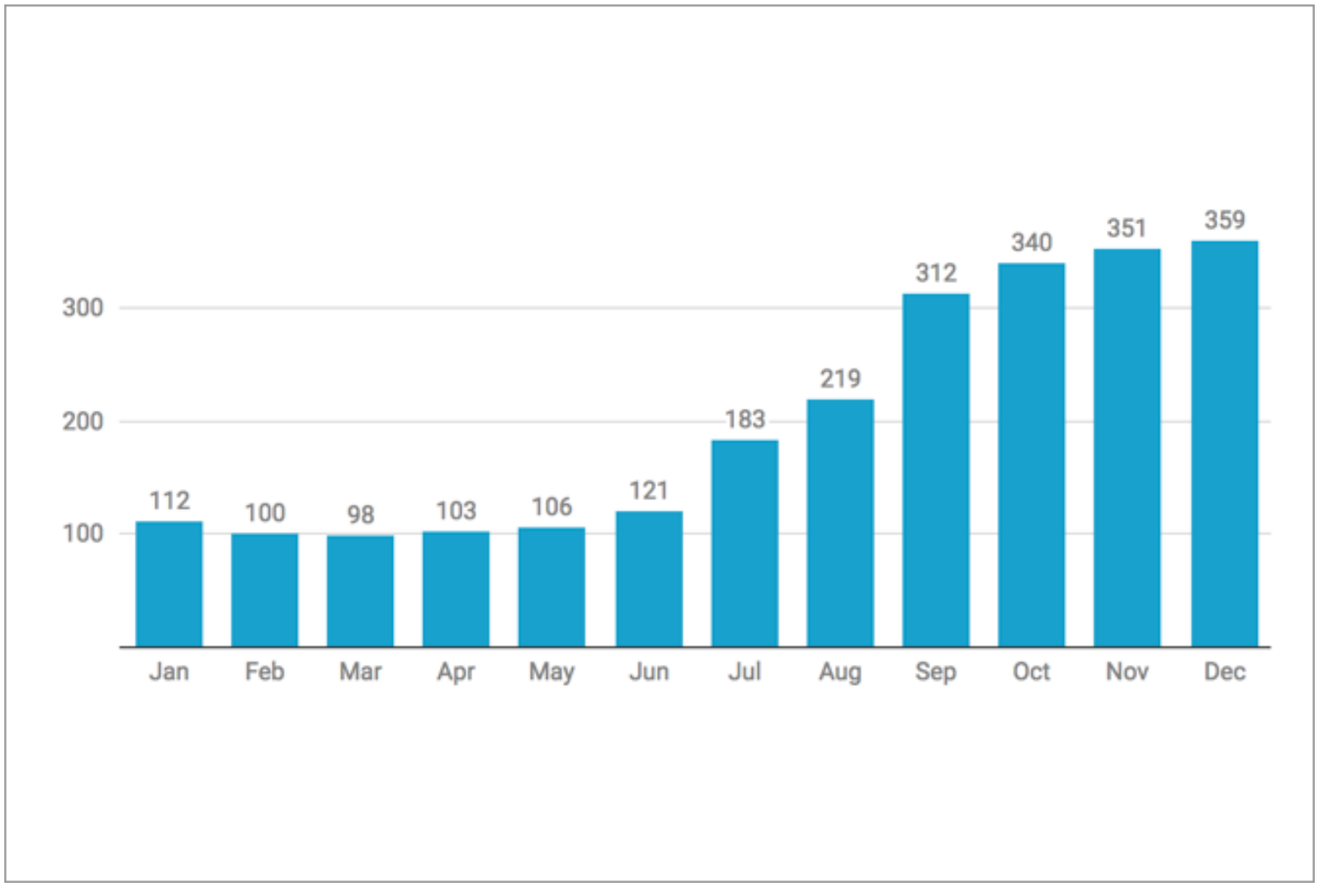
	Categorical	Ordered	Quantitative
Position X		●	
Position Y			●
Area			
Colour Lum			
Colour Hue			

Data Object: Average monthly Banana Ltd stock quotation 2018

Data Attributes: Price, Time

Data Elements: Revenue values per month

Marks: Point | Line | Area



Channels

	Categorical	Ordered	Quantitative
Position X		●	
Position Y			
Area			
Colour Lum			
Colour Hue			
Length			●

Banana Ltd Avg. Revenue by Region

Avg. revenue per 1 mil. people

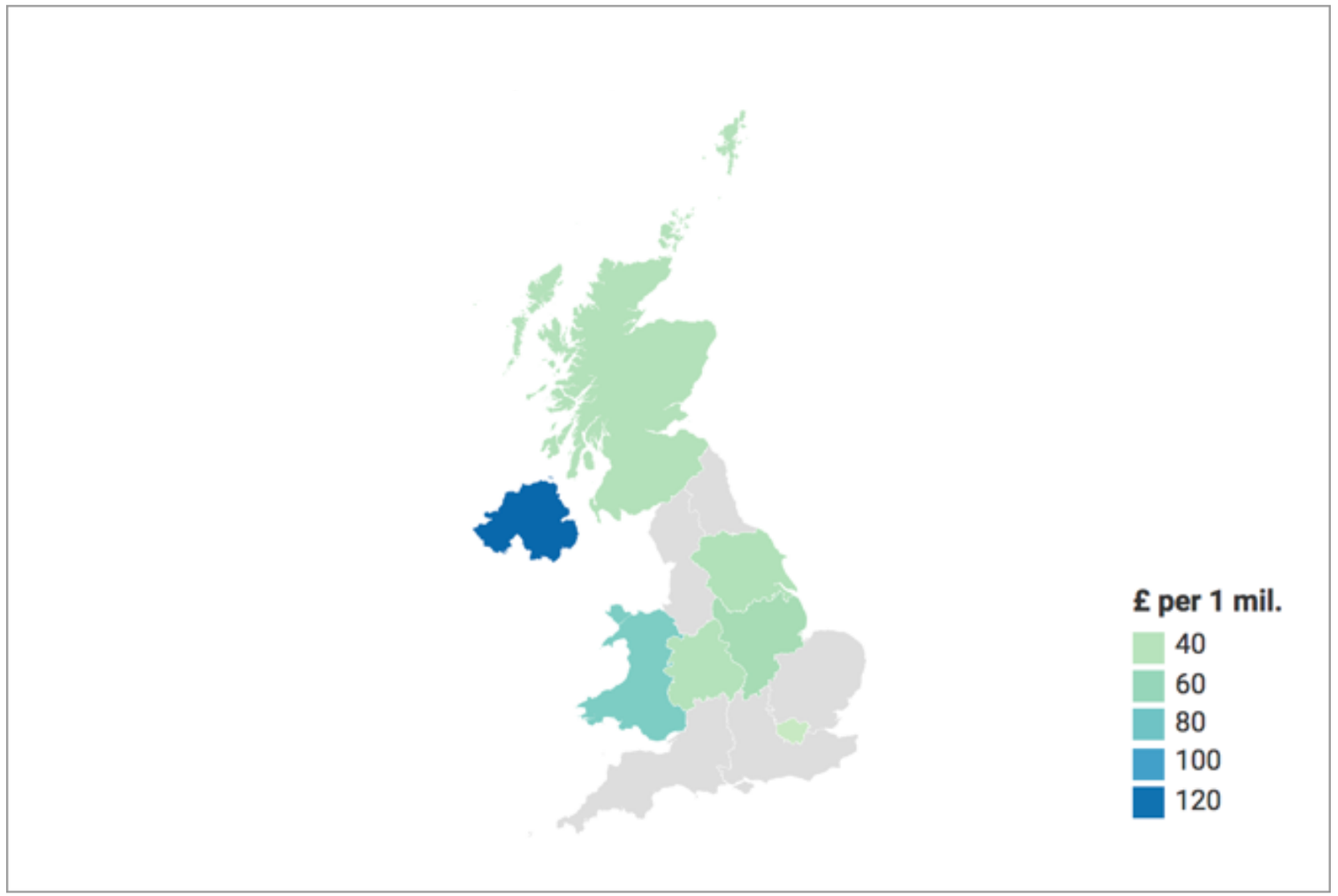
Region	Avg. Revenue
East Midlands	51
London	28
NI	124
Scotland	43
Wales	74
West Midlands	41
Yorkshire	44

Data Object: Average Banana Ltd customer spend per 1 mil people, 2018

Data Attributes: Regions, Revenue

Data Elements: Regions

Marks: Point | Line | Area



Channels

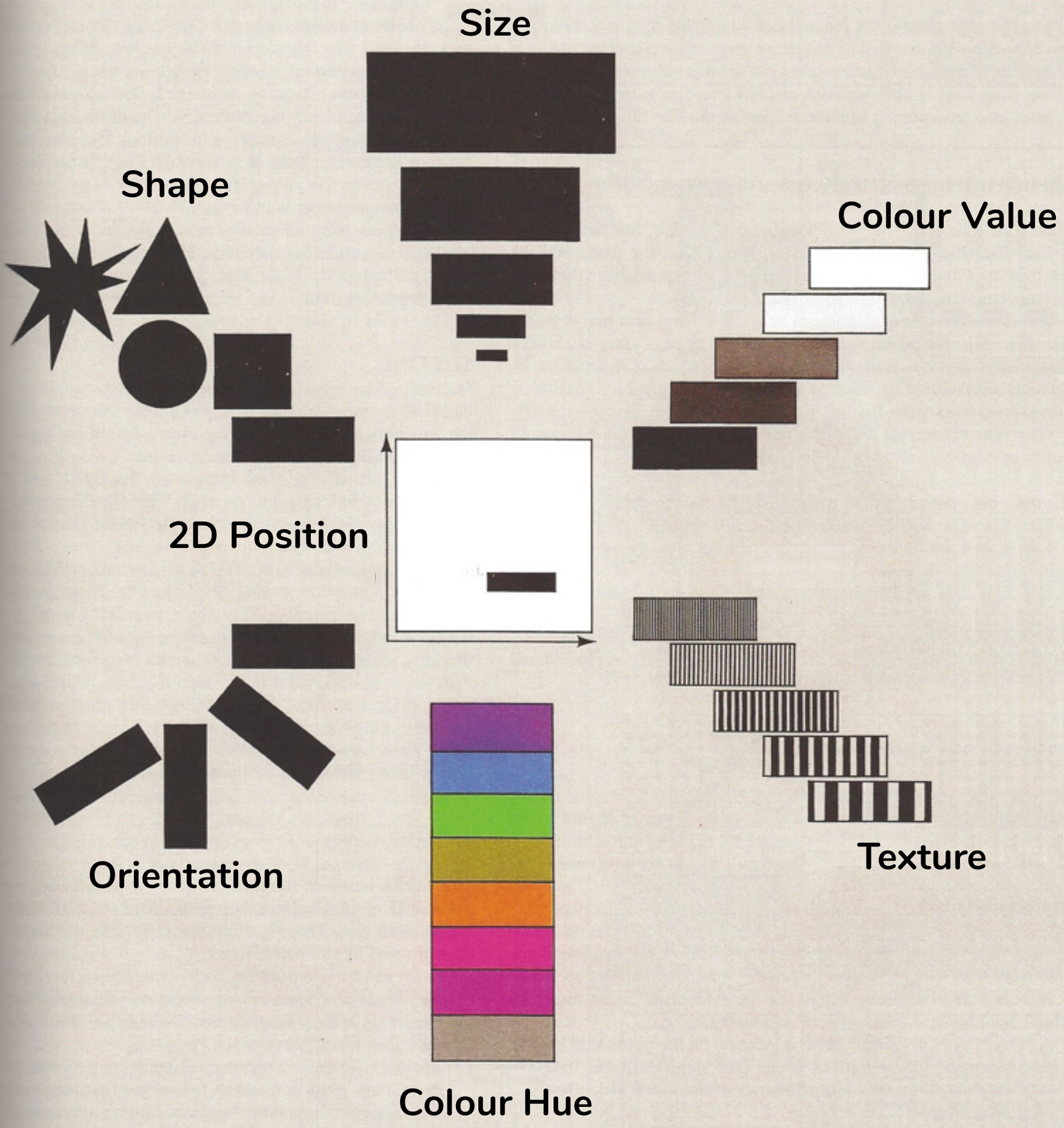
	Categorical	Ordered	Quantitative
Position X			●
Position Y			●
Area			
Colour Lum			●
Colour Hue			●

How do we know which channel to take?

Visual channels are ranked by **effectiveness**.

Rule #1 Follow the **expressiveness principle**

Rule #2 Follow the **effectiveness principle**



	Categorical			
	Association	Selection	Ordered	Quantitative
Planar Dimension	●	●	●	●
Size	○	●	●	●
Value	○	●	●	
Texture	●	●	●	
Colour	●	●		
Orientation	●	○		
Shape	●			

Magnitude Channels for Ordered Attributes

Position on common scale

Position on unaligned scale

Length (1D size)

Tilt/angle

Area (2D size)

Depth (3D position)

Color luminance

Color saturation

Curvature

Volume (3D size)

Same

Same

Same

Effectiveness

More

Less

Identity Channels for Categorical Attributes

Spatial region

Color hue

Motion

Shape

1. Be expressive

Match channel to data type

2. Be effective

Encode the most important attributes with the highest ranked channels

Depends on the objective

Tamara Munzner, Visualization Analysis and Design, 2015

Summary

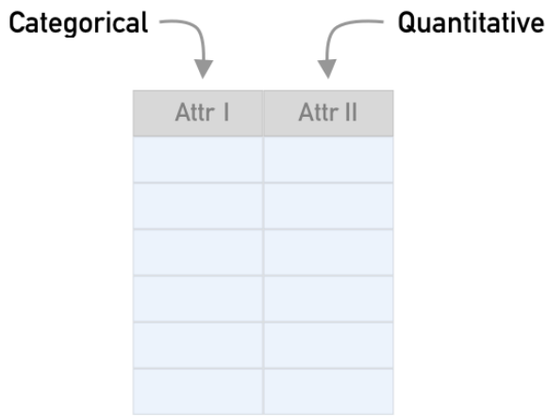
6 stepped Recipe

1. Identify Data Object and Data Attributes to visualise.

Object: the invariant thing this visual is about….

Attr I	Attr II	Attr III	Attr IV

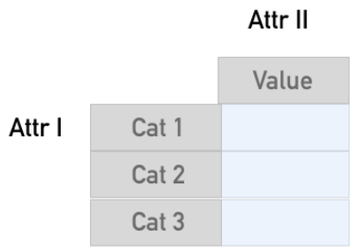
2. Identify the data types of your attributes.



3. Rank attributes by importance. 2-3 are great, careful with more.

4. Map the attributes to channels according to their importance.

5. Get data into the right shape.



6. Visualise.