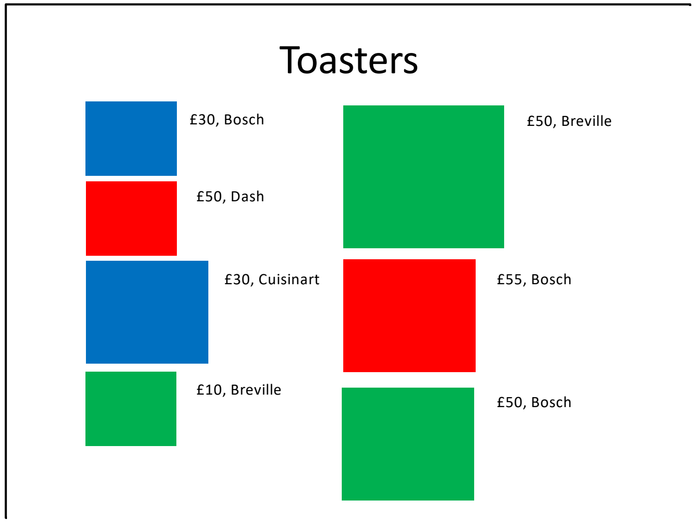
Design Space

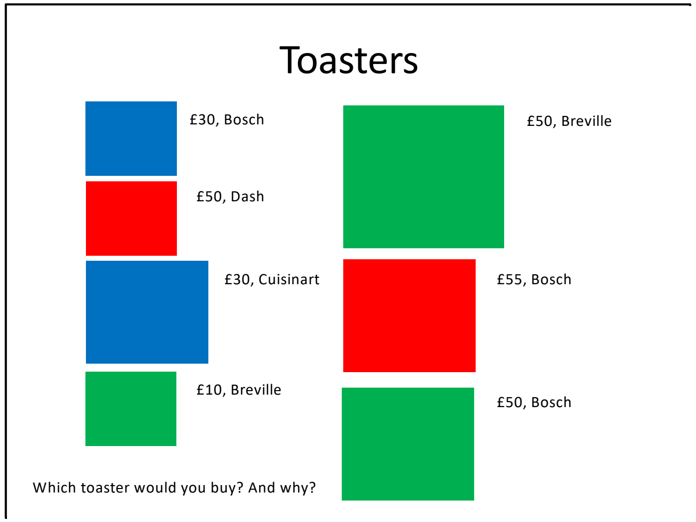
Design space

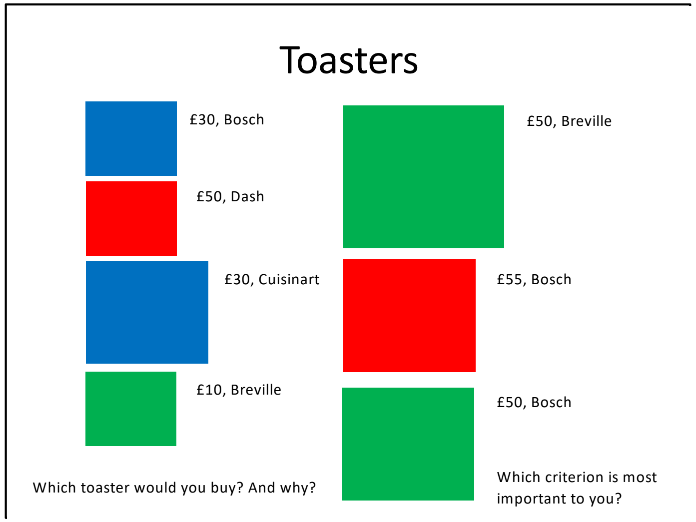
Design is all about **making decisions**

"What is common among design spaces is that they make design decisions explicit, summarize what is possible, and what is under-explored."

Dimensional Reasoning and Research Design Spaces,
S. MacNeil, J. Okerlund, C. Latulipe,
C&C 2017, June 27–30, 2017, Singapore

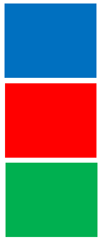

Toasters
£30, Bosch
£50, Dash
£30, Cuisinart
£10, Breville
£50, Breville
£55, Bosch
£50, Bosch

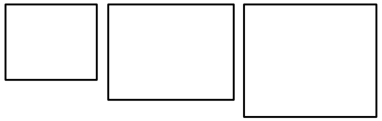

Toasters
£30, Bosch
£50, Dash
£30, Cuisinart
£10, Breville
£50, Breville
£55, Bosch
£50, Bosch
Which toaster would you buy? And why?


Toasters
£30, Bosch
£50, Dash
£30, Cuisinart
£10, Breville
£50, Breville
£55, Bosch
£50, Bosch
Which toaster would you buy? And why?
Which criterion is most important to you?

Choosing a toaster: the nature of choice

Price: £10, £30, £50, £55

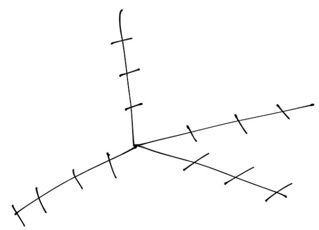


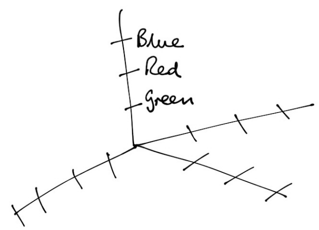


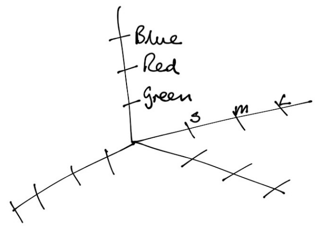
Size: small, medium large

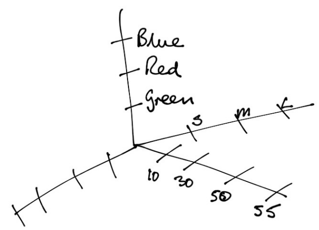
Colour: blue, red, green

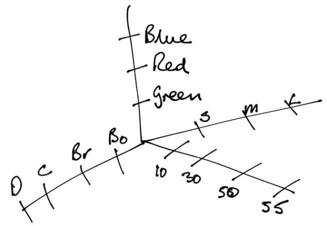
Brand: Bosch, Cuisinart, Dash, Breville

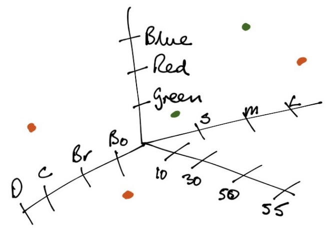



Blue
Red
Green


Blue
Red
Green
s
m
r


Blue
Red
Green
S
M
L
10
30
50
55


Blue
Red
Green
S
M
L
D
C
Br
Bo
10
30
50
55

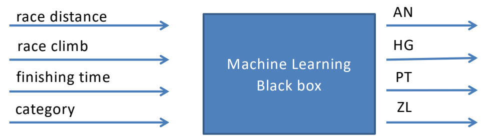

Blue
Red
Green
S
M
D
C
Br
Bo
10
30
50
55

Design space: definition

- Each *decision* is a dimension
- Each *dimension* has a range of *values*
- Each *design* is a point in n-dimensional space
- Dimensions may *interact* with each other
- *Constraints* may indicate that some of the space is not available
- Some areas of the space might be preferable to to others

Design justification explains why one particular point has been chosen instead of another

Identify the runner



	distance	climb	time	category
AN	30	1500	3.54	M40
AN	30	1500	3.21	M30
PT	25	2400	5.32	MOPEN
HG	25	2400	4.27	MOPEN
ZL	30	1500	4.32	F40
HG	10	470	1.06	F50
HG	5	330	0.48	F50
……more rows about races these runners have taken part in				

Visualising the results,
for 500 ML trials

	output categorisation					total
input data	AN	PT	HG	ZL		
AN	385	0	100	15		500
PT	109	323	45	23		500
HG	10	0	432	58		500
ZL	0	0	140	360		500

Design is all about good *decisions*

Choices… with justifications

- – which data to present
- – which visualisation method to use on that data
- – which encoding of the data attributes in the chosen visualisation: what colours, fonts, sizes, marks, symbols...

The process we are going to get into now is for when you have a discrete set of choices to make, in terms of the data to present and the visualization method to use on that data.

We will look at the data issues first, and then the visualization methods after that.

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

(4)	output
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(5)	output			
input	AN	PT	HG	ZL
AN	385	0	100	15
PT	109	323	45	23
HG	10	0	432	58
ZL	0	0	140	360

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

(4)	output
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(5)	output			
input	AN	PT	HG	ZL
AN	385	0	100	15
PT	109	323	45	23
HG	10	0	432	58
ZL	0	0	140	360

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

(4)	output
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

(6)	output (as %)			
input	AN	PT	HG	ZL
AN	77	0	20	3
PT	22	65	9	4
HG	2	0	86	12
ZL	0	0	28	72

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(5)	output			
input	AN	PT	HG	ZL
AN	385	0	100	15
PT	109	323	45	23
HG	10	0	432	58
ZL	0	0	140	360

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

(4)	output
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

(6)	output (as %)			
input	AN	PT	HG	ZL
AN	77	0	20	3
PT	22	65	9	4
HG	2	0	86	12
ZL	0	0	28	72

(7)	output	
input	HG	not-HG
HG	432	68
not HG	285	1215

Data choice

(1)	output
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	output
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(5)	output			
input	AN	PT	HG	ZL
AN	385	0	100	15
PT	109	323	45	23
HG	10	0	432	58
ZL	0	0	140	360

(3)	output
input	% correct
AN	77
PT	65
HG	86
ZL	72

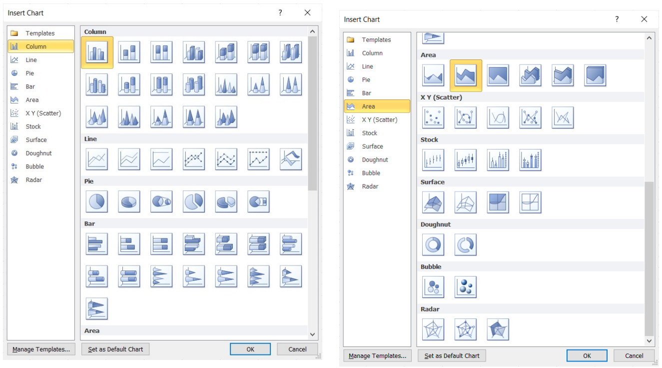
(4)	output
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

(6)	output (as %)			
input	AN	PT	HG	ZL
AN	77	0	20	3
PT	22	65	9	4
HG	2	0	86	12
ZL	0	0	28	72

(7)	output	
input	HG	not-HG
HG	432	68
not HG	285	1215

(8)	output
input	%correct
all	75

Visualisation choices: from Excel



Now we move on to the visualization choices.

Note that here we are assuming that you are going to use a library of visualization tools, each of which can be parameterized to some small extent, e.g. the colours, fonts, marks and symbols in it. You can't really change the core encoding of attributes and dimensions in these tools. Instead, you have to choose ones that already fit with your data choices.

Like here in Excel, you have a large variety of methods to use… but you're not going to go in deep and recode them. You're just going to feed data into them.

Then the first step is to choose a small set of potential tools or methods. There might be many reasons for this, but you may initially just be looking to make a variety of prototypes, to evaluate as you start to narrow down your choices… or you may have a number of methods that you think are good solid methods to use, for other reasons

Set of visualisation choices

(A)		clustered column
(B)		100% stacked column
(C)		stacked column
(D)		stacked line
(E)		line
(F)		pie
(G)		radar

You need to narrow down the set of choices here. You have 8 x 7 = 56 options, which is too many… and so you need to make good design choices that narrow down the space of options.

Let's say that you have then chosen these seven as potential visualization methods to use. 56 is far too many to really work with, so have to narrow down from the 56 to a more sensible and practical number… possibly even just 1!

Data choice

(1)	1x4
input	correct
AN	385
PT	323
HG	432
ZL	360

(2)	1x4
input	incorrect
AN	115
PT	177
HG	68
ZL	140

(5)	4x4			
input	AN	PT	HG	ZL
AN	385	0	100	15
PT	109	323	45	23
HG	10	0	432	58
ZL	0	0	140	360

(3)	1x4
input	% correct
AN	77
PT	65
HG	86
ZL	72

(4)	1x4
input	%incorrect
AN	23
PT	35
HG	14
ZL	28

(6)	4x4			
input	AN	PT	HG	ZL
AN	77	0	20	3
PT	22	65	9	4
HG	2	0	86	12
ZL	0	0	28	72

(7)	2x2	
input	HG	not-HG
HG	432	68
not HG	285	1215

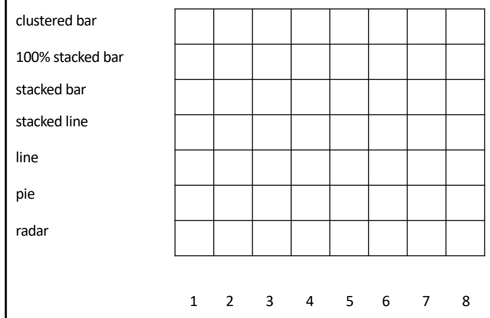
(8)	1x1
input	%correct
all	75

Two dimensions of choice

(A)		clustered bar
(B)		100% stacked bar
(C)		stacked bar
(D)		stacked line
(E)		line
(F)		pie
(G)		radar

(1)	1x4 (correct)
(2)	1x4 (incorrect)
(3)	1x4 (%correct)
(4)	1x4 (%incorrect)
(5)	4x4 (all data)
(6)	4x4 (all, as %)
(7)	2x2 (HG)
(8)	1x1 (%correct)

A 2D space of 56 design options



1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. 1x1 (%correct)

Each location in the space is a particular design – a combination of a data choice and a visualization choice

A 2D space of 56 design options

	1	2	3	4	5	6	7	8
clustered bar								X
100% stacked bar								X
stacked bar								X
stacked line								X
line								X
pie								X
radar								X

1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. **1x1 (%correct)**

We can how use this to help us think about what design choices are available, and what choices are better (and why)… and what choices are not possible or not desirable.

For example, option number 8 is just one number! So we don't need a visualization for that… and so we can mark that entire column as not possible or desirable. It would be best also to take some notes as to why that is so, in our design documents.

A 2D space of 56 design options

	1	2	3	4	5	6	7	8
clustered bar	X	X	X	X				X
100% stacked bar	X	X	X	X				X
stacked bar	X	X	X	X				X
stacked line	X	X	X	X				X
line								X
pie								X
radar								X

1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. 1x1 (%correct)

Similarly, if we consider data choices 1, 2, 3 and 4… they all have 1x4 arrays of data… and so the visualization methods that combine multiple attributes are not appropriate.

For example, a clustered bar chart makes no sense when we only have one attribute for a data item. The same applies for the two types of stacked bar chart and the 100% stacked line chart. So, we can mark those cells as out.

A 2D space of 56 design options

	1	2	3	4	5	6	7	8
clustered bar	X	X	X	X				X
100% stacked bar	X	X	X	X				X
stacked bar	X	X	X	X				X
stacked line	X	X	X	X				X
line								X
pie					X	X	X	X
radar								X

1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. 1x1 (%correct)

A pie chart only shows univariate data, so it will not be suitable for the data sets that are bivariate – 5, 6 and 7.

A 2D space of 56 design options

	1	2	3	4	5	6	7	8
clustered bar	X	X	X	X				X
100% stacked bar	X	X	X	X				X
stacked bar	X	X	X	X				X
stacked line	X	X	X	X	X	X	X	X
line	X	X	X	X	X	X	X	X
pie					X	X	X	X
radar	X	X	X	X			X	X

1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. 1x1 (%correct)

Line charts show trend data, but what we have are discrete values… and so line charts are not appropriate.

Also, radar charts need more than two attributes or dimensions, so they are not appropriate for most of the data choices

Again, we can cut out some design options. I think we have just 15 options left.

A 2D space of 56 design options

	1	2	3	4	5	6	7	8
clustered bar	X	X	X	X			●	X
100% stacked bar	X	X	X	X			●	X
stacked bar	X	X	X	X			●	X
stacked line	X	X	X	X	X	X	X	X
line	X	X	X	X	X	X	X	X
pie					X	X	X	X
radar	X	X	X	X			X	X

1. 1x4 (correct)
2. 1x4 (incorrect)
3. 1x4 (%correct)
4. 1x4 (%incorrect)
5. 4x4 (all data)
6. 4x4 (all, as %)
7. 2x2 (HG)
8. 1x1 (%correct)

We have already narrowed down our potential choices a lot. We may now bring in other design constraints, when we think about what data we want to show. For example, if we want to be able to see data on just one individual runner, and no other, then we should check which data choices have that.

Only option 7 really stays now -- the 2x2. We exclude the two 'all data' options… 5 and 6… as we don't want to see data on multiple runners. The same issue arises with 1, 2, 3 and 4.

We can then see that we have three options left here, shown with green circles … and then we might focus in on other design constraints, such as more detailed tasks or questions we expect users to have, which might then make us give priority to one or more of these remaining options.

Overall, I've tried to show in this example how there may be many potential design options, but there are ways to narrow down, or focus in on, a smaller and smaller set of possible options… with clear reasons at each step for each decision.

Design is all about good *choices*

Choices… with justifications

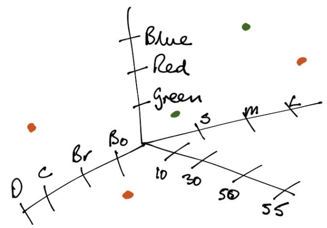
- which data to present
- which visualisation method to use on that data
- How to present the data attributes in the chosen visualisation: what colours, fonts, sizes, marks, symbols...

The kind of process we have been through here is for very early-stage design work, with discrete options for the data to present and the methods to use.

It is about making clear your high-level system design choices… justifying *why* you made those choices. When you start to work with a real data set, and real visualisation methods, you will see that there are many options with regard to how to you initially process and present the data, before visualization, and also what visualisation methods you use… and so it is good to have a clear process for making a good decision at the start of your work.

Once you have made the first two choices here, other issues become important, too. For example, when you've decided on the data and the method, then there will be different ways to show the attributes in that data, for example colours, fonts, mark sizes, and so on. We won't get into that in this lecture, though.

Every design is a single point in the multidimensional design space



Remember the toaster example? That had even more dimensions… and so it is more difficult to draw and work with.

In the next lecture, we will talk about how to handle the situation where we have many design dimensions. We will use one of the techniques we discussed before, for visualising multivariate data… next time.

Design space

"What is common among design spaces is that they make design decisions explicit, summarize what is possible, and what is under-explored"…

…and what is not possible or not desirable

Dimensional Reasoning and Research Design Spaces,
S. MacNeil, J. Okerlund, C. Latulipe,
C&C 2017, June 27–30, 2017, Singapore

Here is a quote from the research that was used to guide this lecture. It describes what these three authors see as the strengths of their method.

I must also add that if you have the option to create or adapt the code for the visualization methods, then this whole process may not apply! You don't have discrete and clearly different choices for visualization methods, like the Excel library, but a more open and flexible set of possibilities for what you make. In a later lecture, we will cover a different design process that deals with the core of that, based on Bertin's work on semiology of graphics… but that's another lecture…

Design Space