Building dynamic, operational data visualisations

Chris Horsley

CSIRT Foundry

Overview

- What types of visualisations?
- What type of tech?
- What toolkits?
- Building a visualisation
- Visualisation layouts

CSIRT foundry

Introduction

- Web dev background
- AusCERT
- JPCERT/CC
- CSIRT Foundry

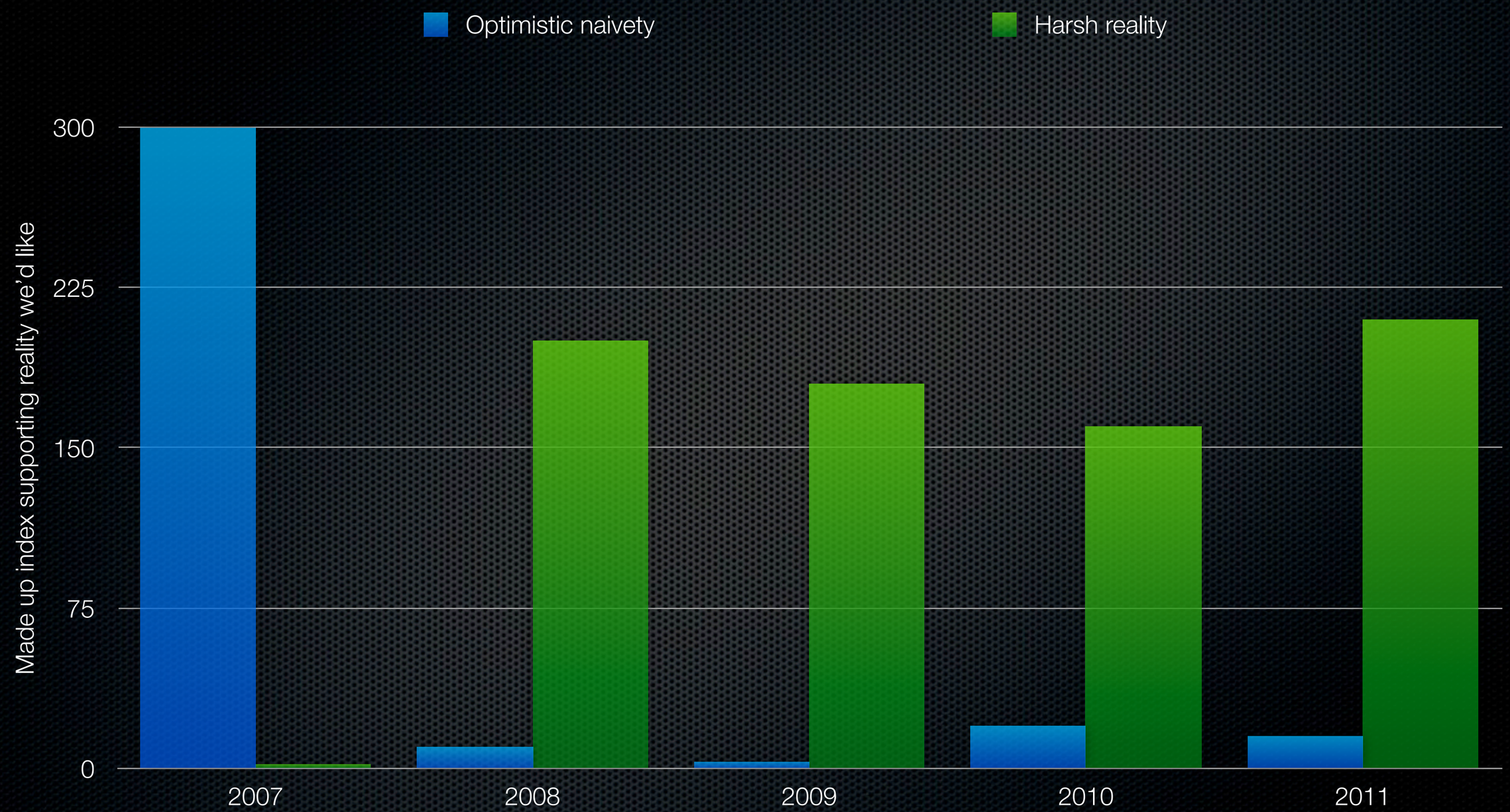
Who is this talk for?

- An analyst wanting to understand visualisation techniques
- Someone using visualisation, but looking for new options

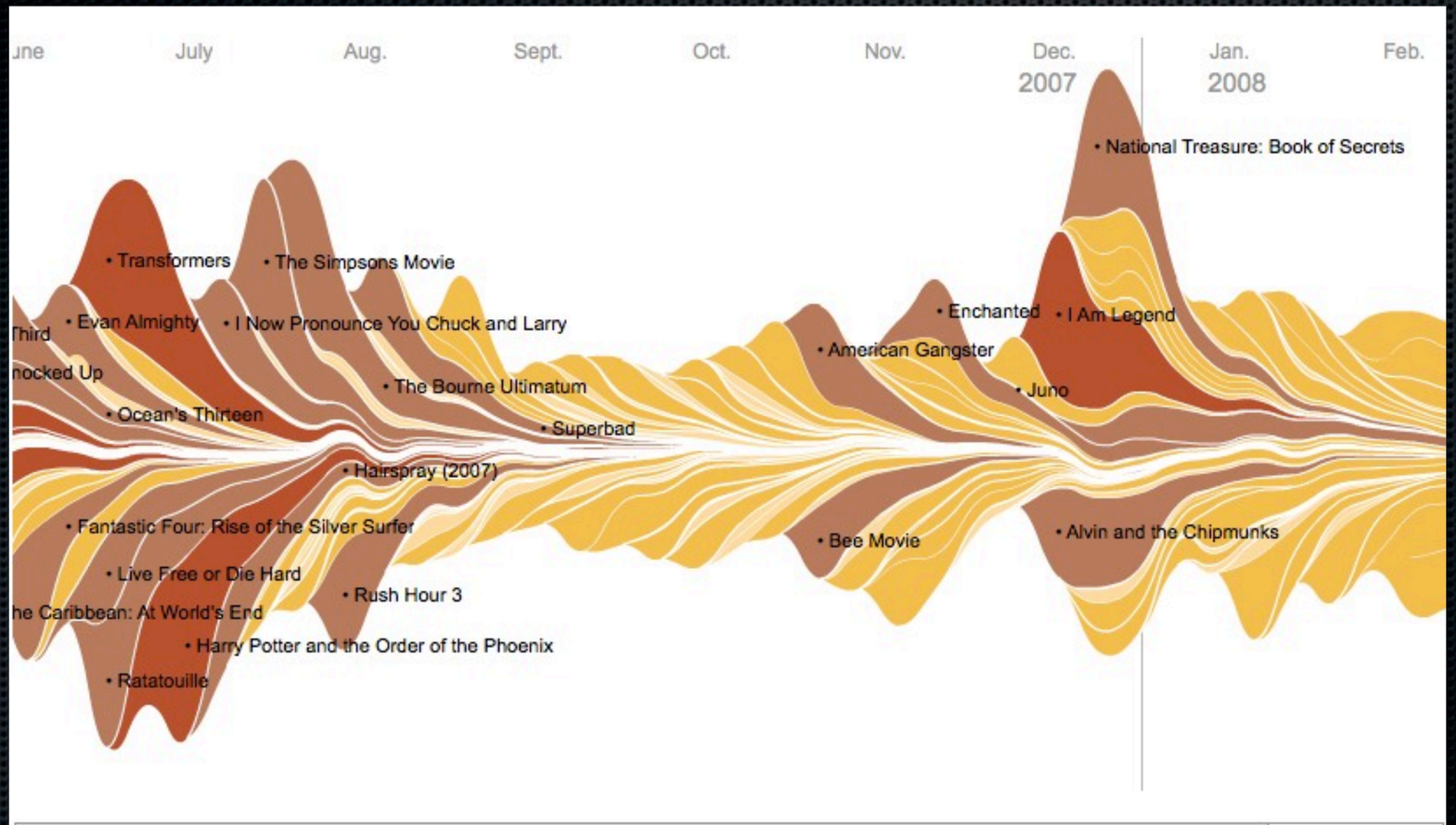
CSIRT foundry

When we talk about “visualisation”...

CSIRT foundry


Optimistic naivety
Harsh reality
Made up index supporting reality we'd like
300
225
150
75
0
2007
2008
2009
2010
2011

```
Beef   ##################
Cake   ######
Tom    ####
Yum    ##############################
Goong  #########
```

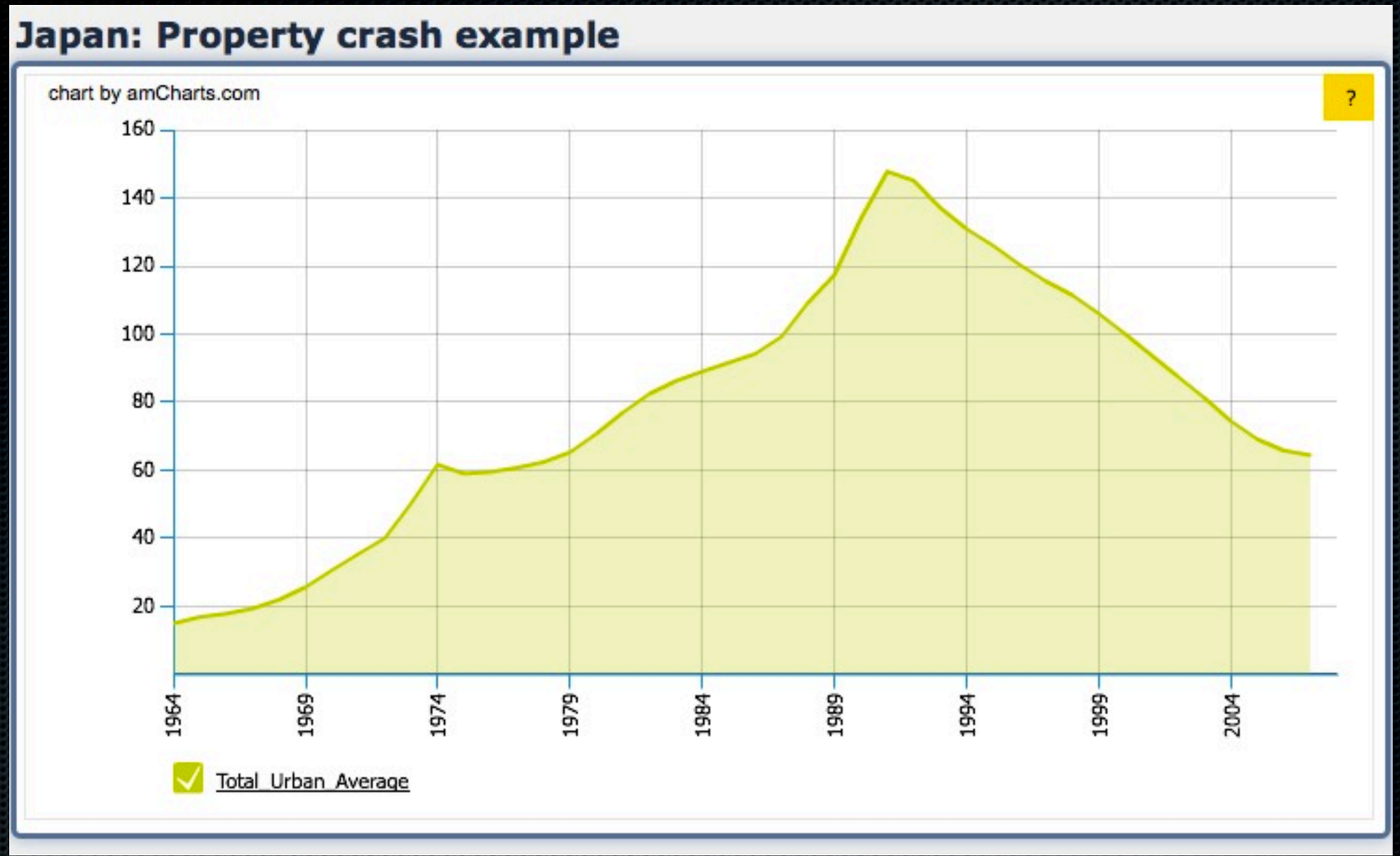

une
July
Aug.
Sept.
Oct.
Nov.
Dec.
2007
Jan.
2008
Feb.
• National Treasure: Book of Secrets
• Transformers
• The Simpsons Movie
• Evan Almighty
• I Now Pronounce You Chuck and Larry
Third
nocked Up
• Enchanted
• I Am Legend
• American Gangster
• The Bourne Ultimatum
• Juno
• Ocean's Thirteen
• Superbad
• Hairspray (2007)
• Fantastic Four: Rise of the Silver Surfer
• Bee Movie
• Alvin and the Chipmunks
• Live Free or Die Hard
• Rush Hour 3
he Caribbean: At World's End
• Harry Potter and the Order of the Phoenix
• Ratatouille

Art + science

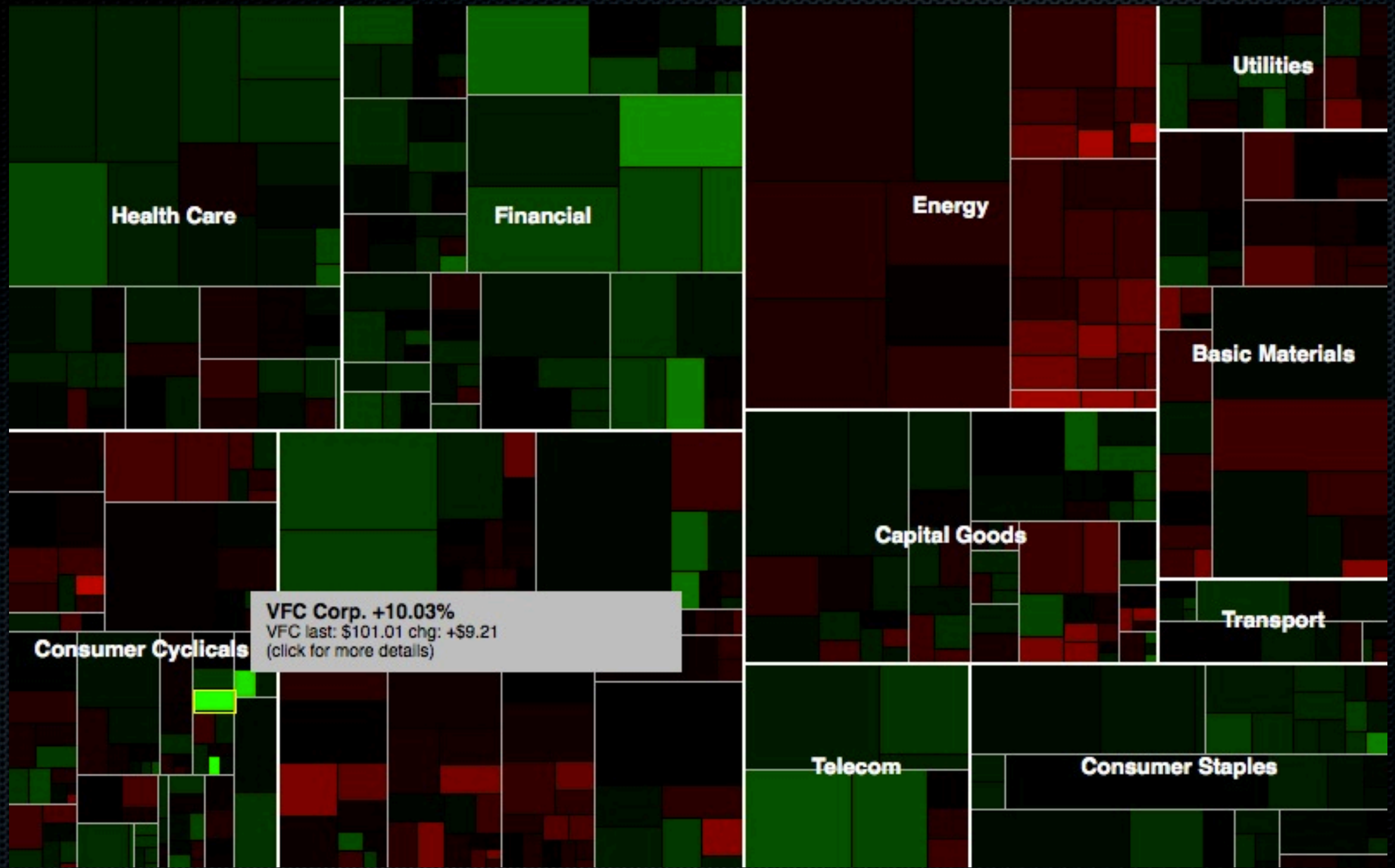
CSIRT foundry

What should a good visualisation do?

The same as a good photo

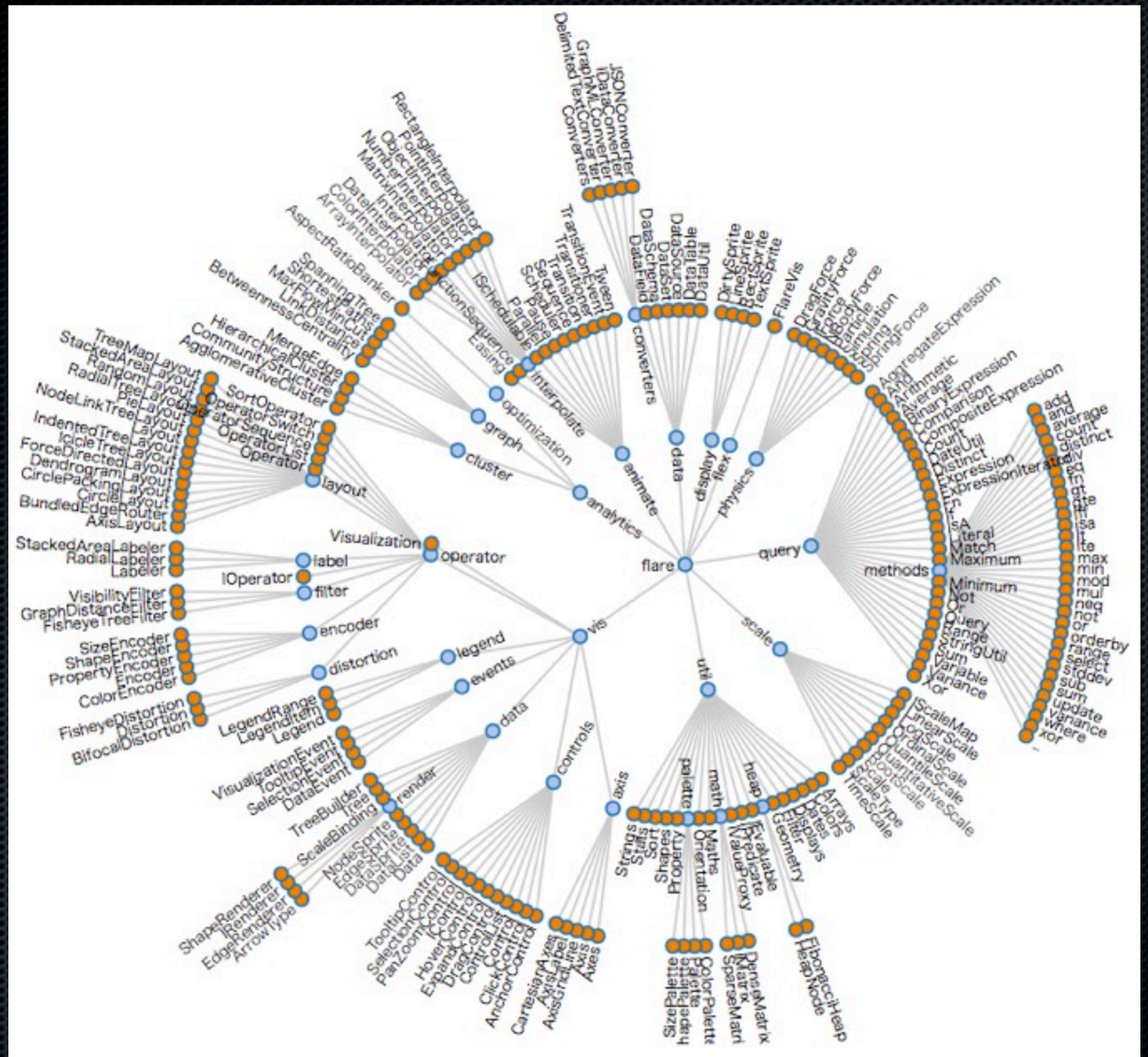

Japan: Property crash example
chart by amCharts.com
?
160
140
120
100
80
60
40
20
1964
1969
1974
1979
1984
1989
1994
1999
2004
Total_Urban_Average

Tell a story

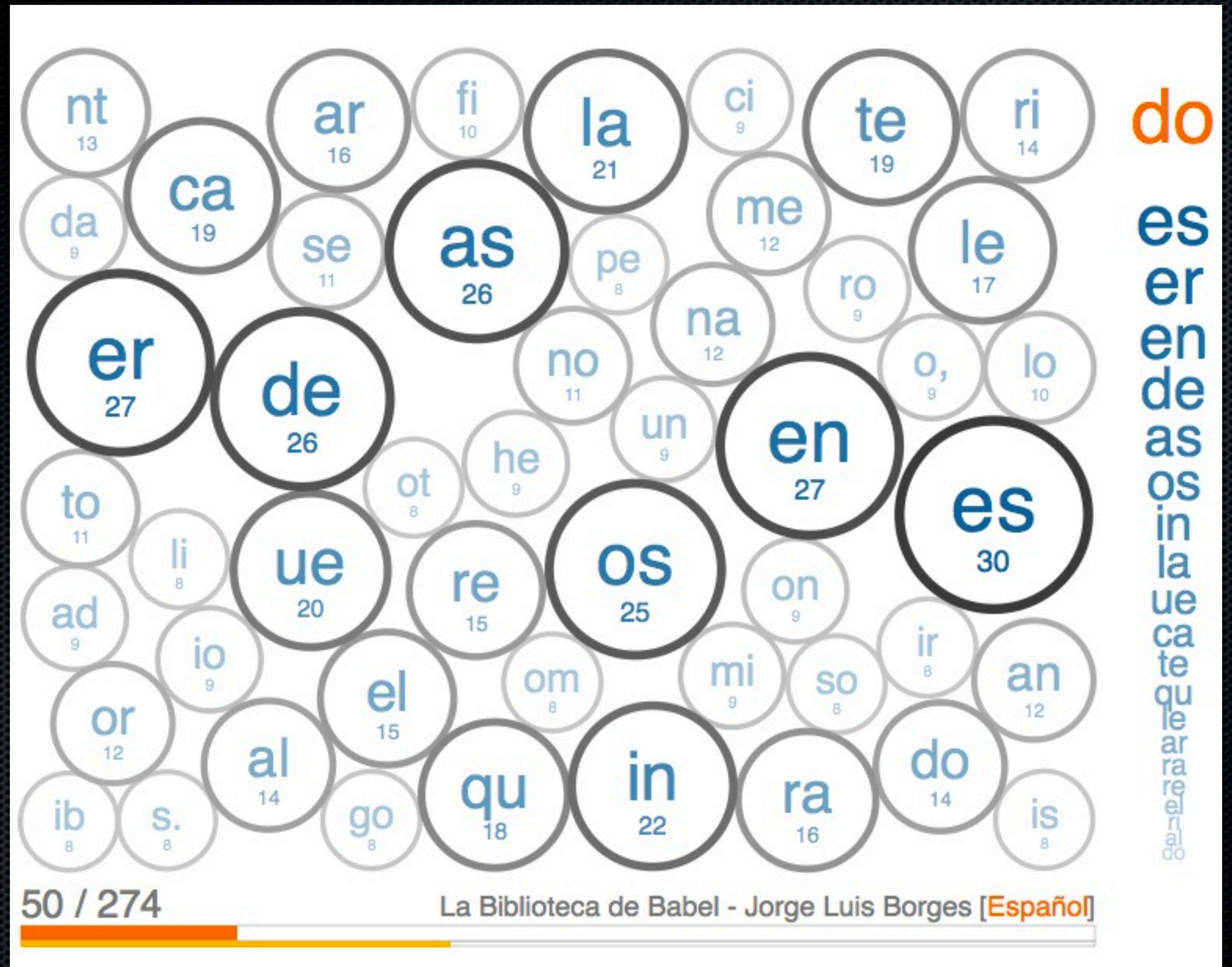


What do we look at next?

CSIRT foundry



Discover new relationships



Be attractive

Think:

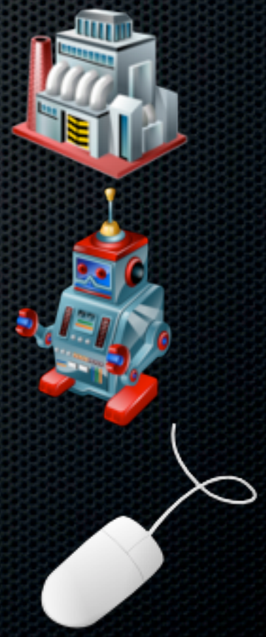
Same as writing: what does the reader take away?

CSIRT foundry

What types of visualisations are we talking about today?

CSIRT foundry

- Reusable
  - Load new data to the same tool
- Automated
  - Run via cron or realtime
- Interactive
  - Layers of information available
- Animated (if you like)
  - Not for every vis, but nice to have



CSIRT foundry

There's visualisations and there's *visualisations*

What kind of visualisation are we **not** talking about today?

CSIRT foundry

- ★ “Infographics”
- ★ One-off visualisations made manually
- ★ Static images
- ★ Flash
- ★ Desktop tools

So, the best toolkit?

CSIRT foundry

So, the best toolkit?

HTML + Javascript

CSIRT foundry

Why web tech?

- Reusable - programmable, code driven
- Automated - pre-generated code or load data realtime
- Interactive - SVG, events
- Animation - SVG, DOM manipulation
- Lots of helper libraries available
- Data format in wide use - JSON
- Cross-platform

CSIRT foundry

Getting started with web visualisations

CSIRT foundry

Tools of the trade

- Toolkits
- JSON
- ...as well as HTML + JS, of course

CSIRT foundry

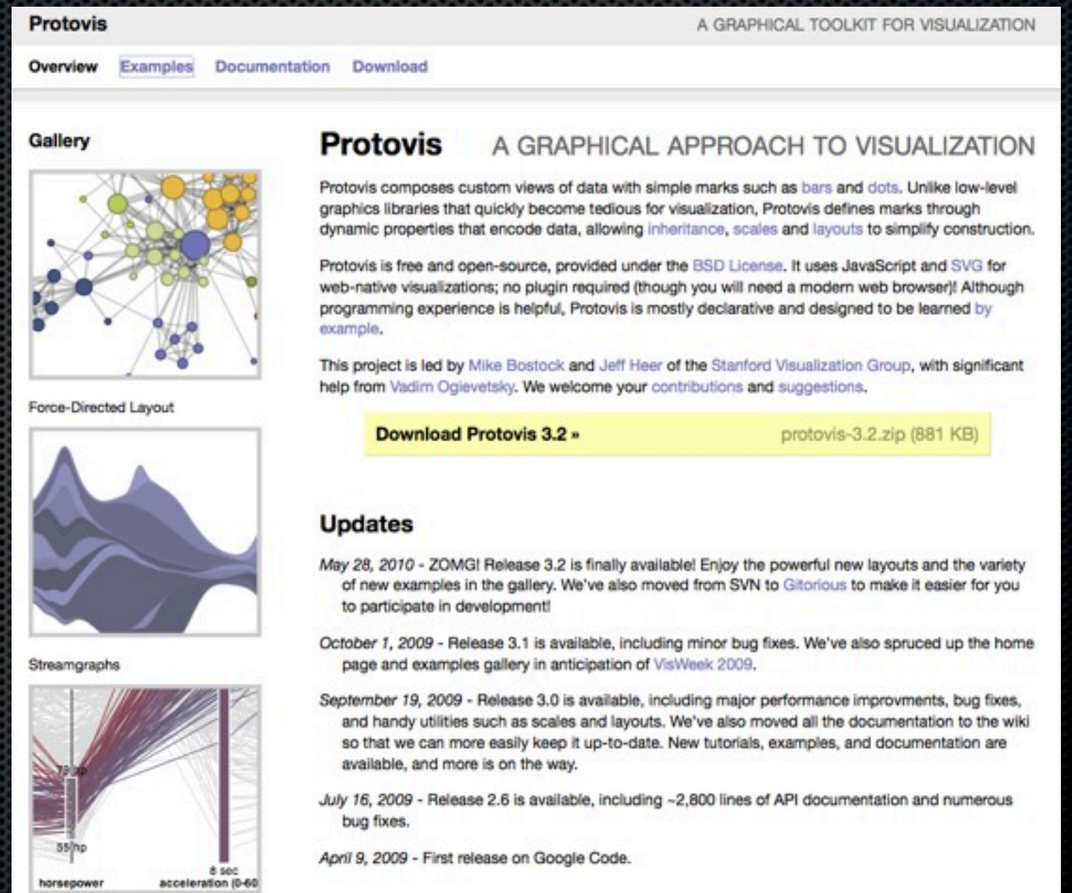
Toolkits

- Protovis
- JIT
- ProcessingJS
- D3

CSIRT foundry

Toolkits - Protovis

- High level API
- Lots of stock types
  - Treemaps, streamgraphs, vector graphs
- SVG
- Animation needs high CPU

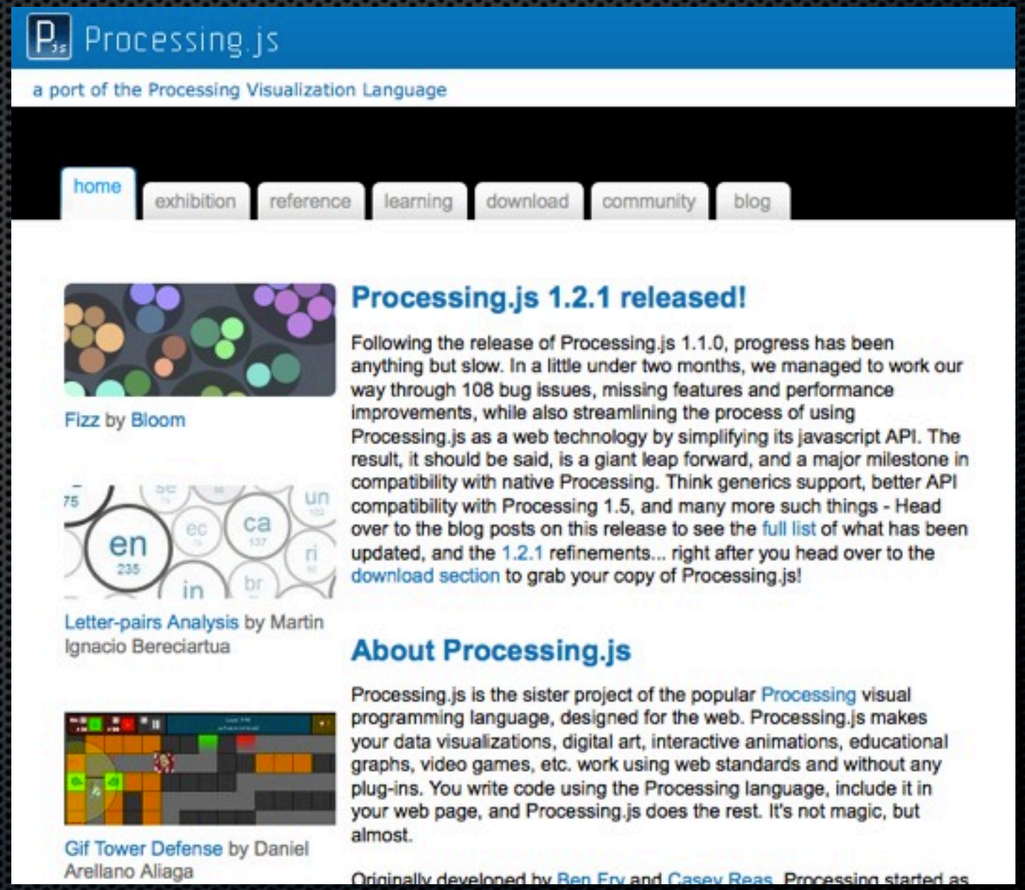


Toolkits - JIT

- Similar in aims to Protovis
- Different API style
- Nicer transitions
- Different layouts

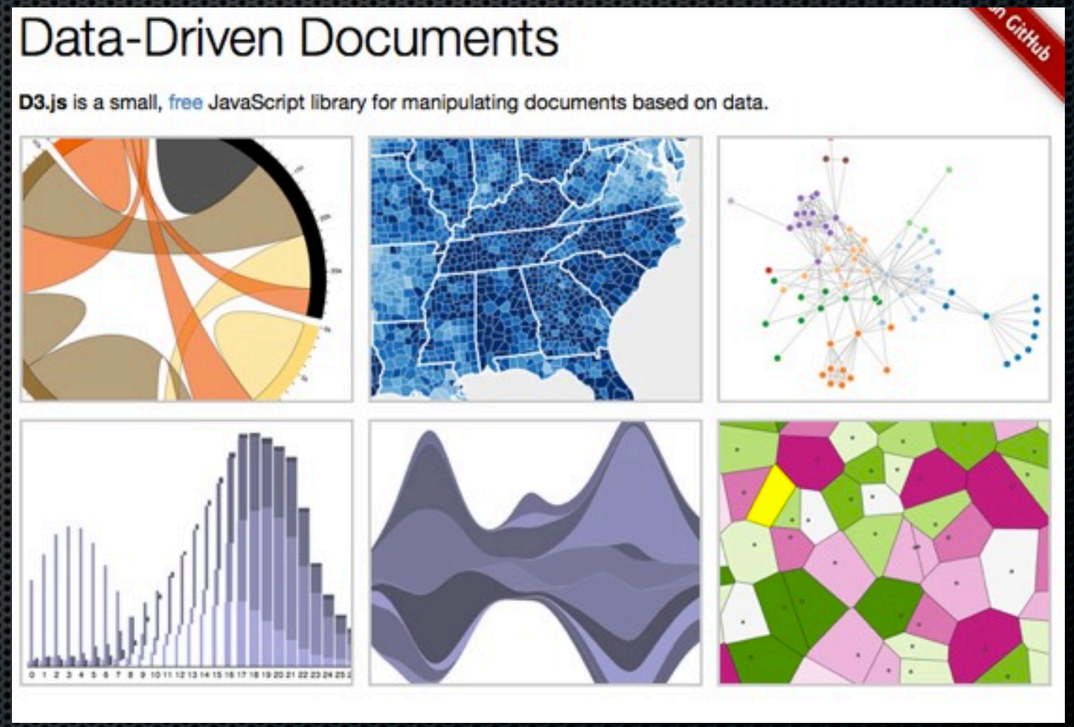
Toolkits - ProcessingJS

- Lower level
- Port from Java to JS
- Wide use by non-techs and artists
- Very flexible



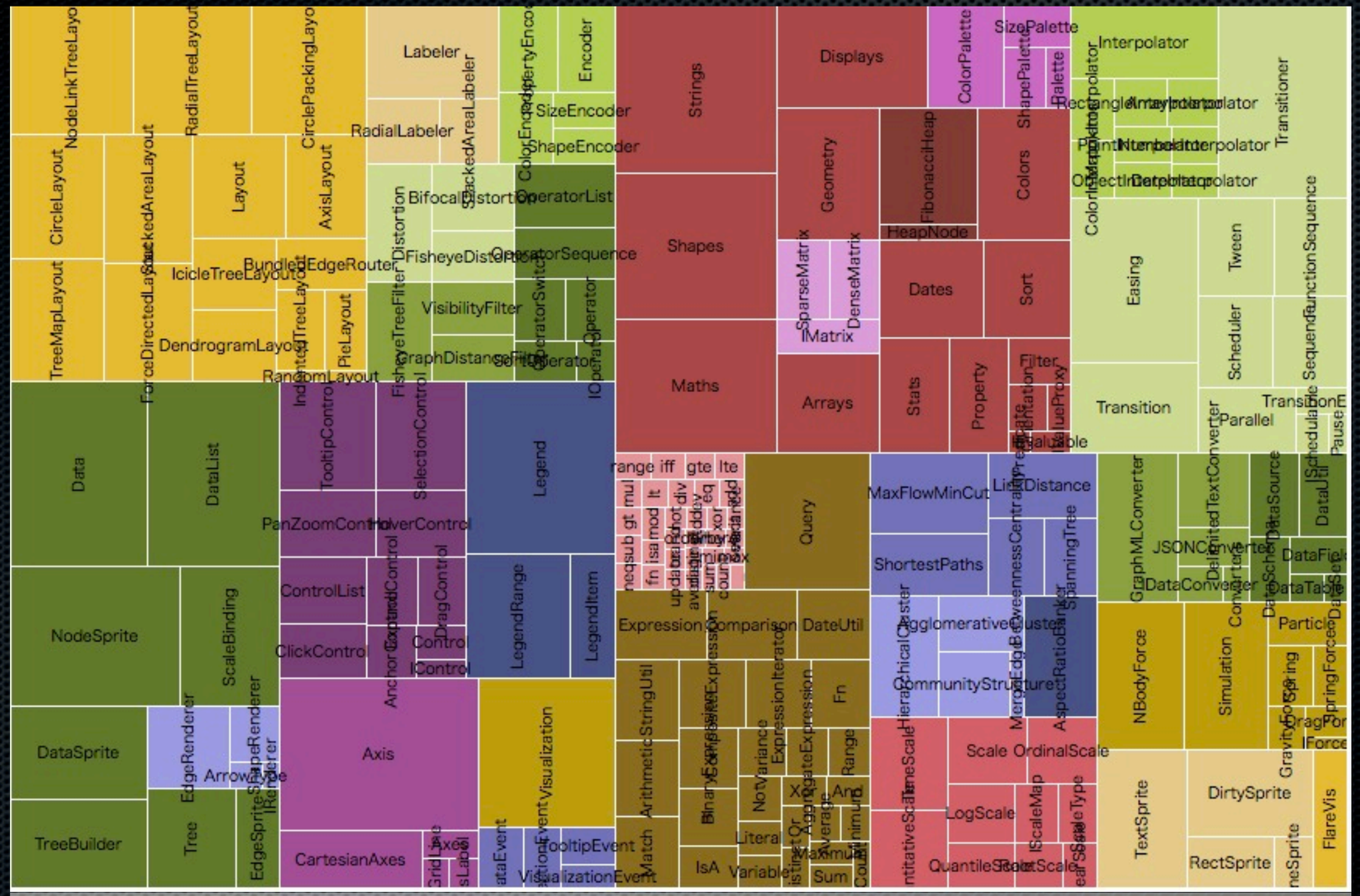
Toolkits - D3

- Same developer as Protovis
- Lower level, more flexible
- SVG + DOM manipulation
- Designed for smooth transitions



Getting your hands dirty

Protovis in action


NodeLinkTreeLayout
RadialTreeLayout
CirclePackingLayout
CircleLayout
StackedAreaLayout
Layout
AxisLayout
TreeMapLayout
ForceDirectedLayout
IcicleTreeLayout
BundledEdgeRouter
DendrogramLayout
IndentedTreeLayout
PieLayout
RandomLayout
Labeler
RadialLabeler
StackedAreaLabeler
ColorEncoder
PropertyEncoder
Encoder
SizeEncoder
ShapeEncoder
Distortion
BifocalDistortion
FisheyeDistortion
OperatorList
OperatorSequence
OperatorSwitch
Operator
IOperator
FisheyeTreeFilter
VisibilityFilter
GraphDistanceFilter
SortOperator
Strings
Shapes
Maths
Displays
ColorPalette
SizePalette
ShapePalette
Palette
Interpolator
RectangleInterpolator
ArrayInterpolator
PointInterpolator
NumberInterpolator
ObjectInterpolator
ColorInterpolator
DateInterpolator
Transitioner
Geometry
FibonacciHeap
HeapNode
Colors
SparseMatrix
DenseMatrix
IMatrix
Dates
Sort
Easing
Tween
Sequence
FunctionSequence
Scheduler
SequenceScheduler
Arrays
Stats
Property
Filter
ValueProxy
Transition
TransitionEvent
Parallel
Pause
Data
DataList
TooltipControl
SelectionControl
Legend
PanZoomControl
HoverControl
ClickControl
AnchorControl
ControlList
DragControl
ControlEvent
IControl
LegendRange
LegendItem
range iff gte lte
mul gt sub mod lt div eq neq xor and or add
fn isa
update
avg min sum max count
Query
Expression
Comparison
DateUtil
StringUtil
ArithmeticExpression
BinaryExpression
ExpressionIterator
Fn
Range
NotVariance
Variance
And
Literal
Match
IsA
Variable
Sum
MaxFlowMinCut
ShortestPaths
ListDistance
MaxCentrality
BetweennessCentrality
SpanningTree
GraphMLConverter
JSONConverter
DelimitedTextConverter
DataConverter
DataSource
DataUtil
DataField
DataTable
DataSchema
Converter
Simulation
NBodyForce
Particle
Spring
SpringForce
Gravity
DragForce
IForce
HierarchicalCluster
AgglomerativeCluster
CommunityStructure
MergeEdge
AspectRatioBanker
Scale
OrdinalScale
TimeScale
LogScale
ScaleMap
ScaleType
QuantitativeScale
QuantileScale
ThresholdScale
LinearScale
Data
NodeSprite
DataSprite
TreeBuilder
Tree
EdgeRenderer
ArrowType
ScaleBinding
EdgeSprite
IRenderer
Axis
CartesianAxes
Axes
GridLine
AxisLabel
DataEvent
Visualization
VisualizationEvent
TooltipEvent
TextSprite
DirtySprite
RectSprite
FlareVis

The data

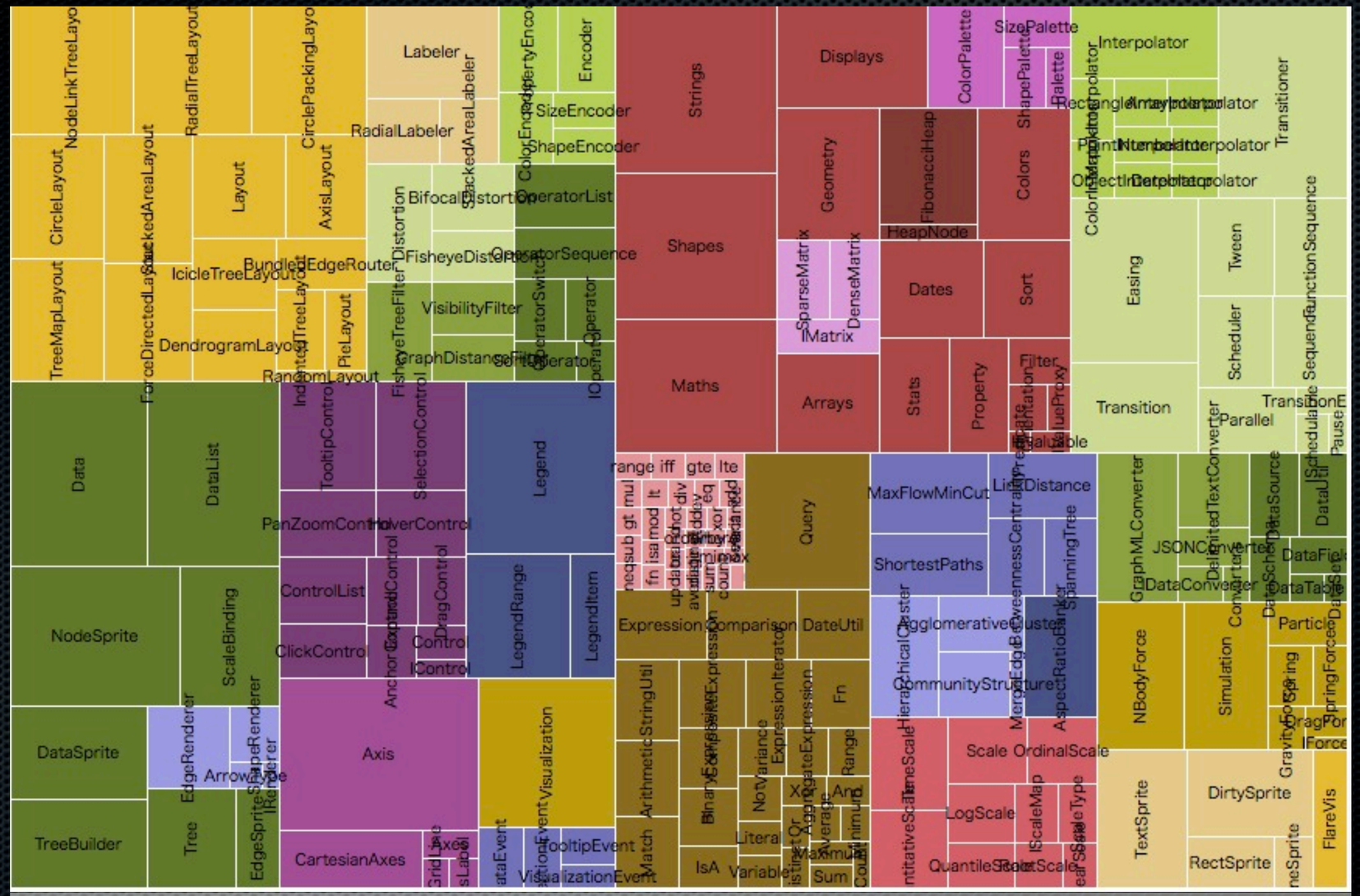
```
var flare = {
analytics: {
  cluster: {
    AgglomerativeCluster: 3938,
    CommunityStructure: 3812,
    HierarchicalCluster: 6714,
    MergeEdge: 743
  },
  graph: {
    BetweennessCentrality: 3534,
    LinkDistance: 5731,
    MaxFlowMinCut: 7840,
    ShortestPaths: 5914,
    SpanningTree: 3416
  },
  optimization: {
    AspectRatioBanker: 7074
  }
},
...
```

The HTML

```
<html>

  <head>
    <title>Treemap</title>
    <link rel="stylesheet" type="text/css" href="ex.css?3.2"/>
    <script type="text/javascript" src="../protovis-r3.2.js"></script>
    <script type="text/javascript" src="flare.js"></script>
    <style type="text/css">
    </style>
  </head>
  <body><div id="center"><div id="fig">
    <script type="text/javascript+protovis">
      <!-- Protovis code goes here -->
    </script>
  </div></div></body>
</html>
```

The code

```
    // Protovis code

var re = "",
    color = pv.Colors.category19().by(function(d) d.parentNode.nodeName)
    nodes = pv.dom(flare).root("flare").nodes();

var vis = new pv.Panel()
    .width(860)
    .height(568);

var treemap = vis.add(pv.Layout.Treemap)
    .nodes(nodes)
    .round(true);

treemap.leaf.add(pv.Panel)
    .fillStyle(function(d) color(d).alpha(title(d).match(re) ? 1 : .2))
    .strokeStyle("#fff")
    .lineWidth(1)
    .antialias(false);

treemap.label.add(pv.Label)
    .textStyle(function(d) pv.rgb(0, 0, 0, title(d).match(re) ? 1 : .2));

vis.render();
```


NodeLinkTreeLayout
RadialTreeLayout
CirclePackingLayout
CircleLayout
StackedAreaLayout
Layout
AxisLayout
TreeMapLayout
ForceDirectedLayout
IcicleTreeLayout
BundledEdgeRouter
DendrogramLayout
IndentedTreeLayout
PieLayout
RandomLayout
Labeler
RadialLabeler
StackedAreaLabeler
ColorEncoder
PropertyEncoder
Encoder
SizeEncoder
ShapeEncoder
Distortion
BifocalDistortion
FisheyeDistortion
OperatorList
OperatorSequence
OperatorSwitch
Operator
IOperator
FisheyeTreeFilter
VisibilityFilter
GraphDistanceFilter
SortOperator
Strings
Shapes
Maths
Displays
ColorPalette
SizePalette
ShapePalette
Palette
Interpolator
RectangleInterpolator
ArrayInterpolator
PointInterpolator
NumberInterpolator
ObjectInterpolator
ColorInterpolator
DateInterpolator
Transitioner
Geometry
FibonacciHeap
HeapNode
Colors
SparseMatrix
DenseMatrix
IMatrix
Dates
Sort
Easing
Tween
Sequence
FunctionSequence
Scheduler
SequenceScheduler
Arrays
Stats
Property
Filter
ValueProxy
Transition
TransitionEvent
Parallel
Pause
Data
DataList
TooltipControl
SelectionControl
Legend
PanZoomControl
HoverControl
ClickControl
AnchorControl
ControlList
DragControl
ControlEvent
IControl
LegendRange
LegendItem
range iff gte lte
mul gt sub mod lt div eq neq xor and or add
fn isa
update
avg min sum max count
Query
Expression
Comparison
DateUtil
StringUtil
ArithmeticExpression
BinaryExpression
ExpressionIterator
Fn
Range
NotVariance
Variance
And
Literal
Match
IsA
Variable
Sum
MaxFlowMinCut
ShortestPaths
ListDistance
MaxCentrality
BetweennessCentrality
SpanningTree
GraphMLConverter
JSONConverter
DelimitedTextConverter
DataConverter
DataSource
DataUtil
DataField
DataTable
DataSchema
Converter
Simulation
NBodyForce
Particle
Spring
SpringForce
Gravity
DragForce
IForce
HierarchicalCluster
AgglomerativeCluster
CommunityStructure
MergeEdge
AspectRatioBanker
Scale
OrdinalScale
TimeScale
LogScale
ScaleMap
ScaleType
QuantitativeScale
QuantileScale
ThresholdScale
LinearScale
Data
NodeSprite
DataSprite
TreeBuilder
Tree
EdgeRenderer
ArrowType
ScaleBinding
EdgeSprite
IRenderer
Axis
CartesianAxes
Axes
GridLine
AxisLabel
DataEvent
Visualization
VisualizationEvent
TooltipEvent
TextSprite
DirtySprite
RectSprite
FlareVis

Visualisation layouts

Well, just one...

A word on maps

Is your data REALLY geographical?

CSIRT foundry

I have some IP addresses

I'll map them to lat long!

Questions

- Which country has the most infections?
- Which continent has the most infections?
- How many infections are there in Sydney, Australia?

CSIRT foundry

- What new insights did we get?
  - Computers got infected where people live?
- We lose 70% of space
- Don’t work well for comparisons
- Networks are largely topological, not geographical

CSIRT foundry

Better options for country listings

- Treemaps
- Bar charts
- Interactive listings
- Always ask: what point do we want to illustrate?

CSIRT foundry

Summary

- HTML + JS is a very flexible option for visualisation
- Lots of mature and emerging toolkits
- Choose your subject and technique carefully

CSIRT foundry

Thank you!

- chris.horsley@csirtfoundry.com
- http://csirtfoundry.com
- Twitter: @Parsify and @CSIRTFoundry