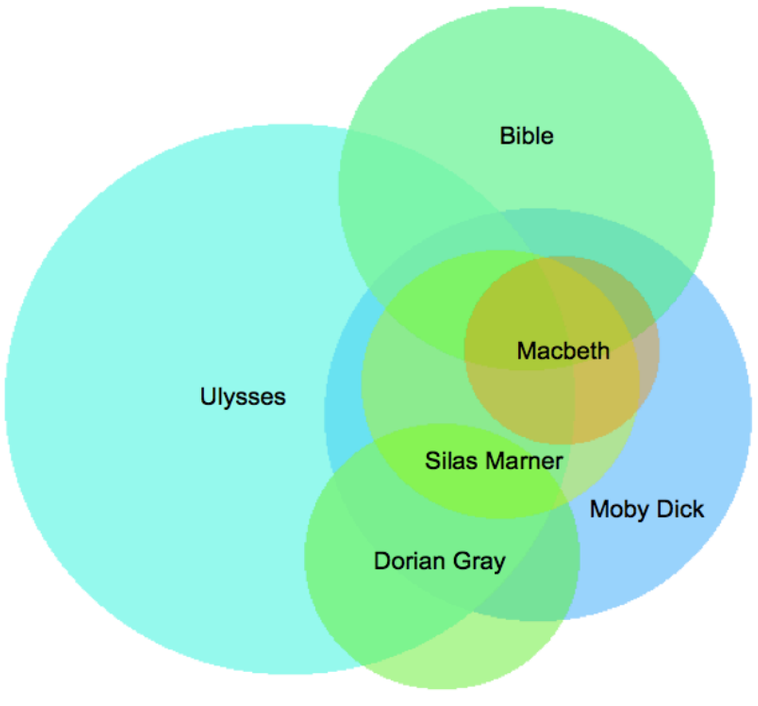
D3 extension:

Area-proportional Venn Diagrams

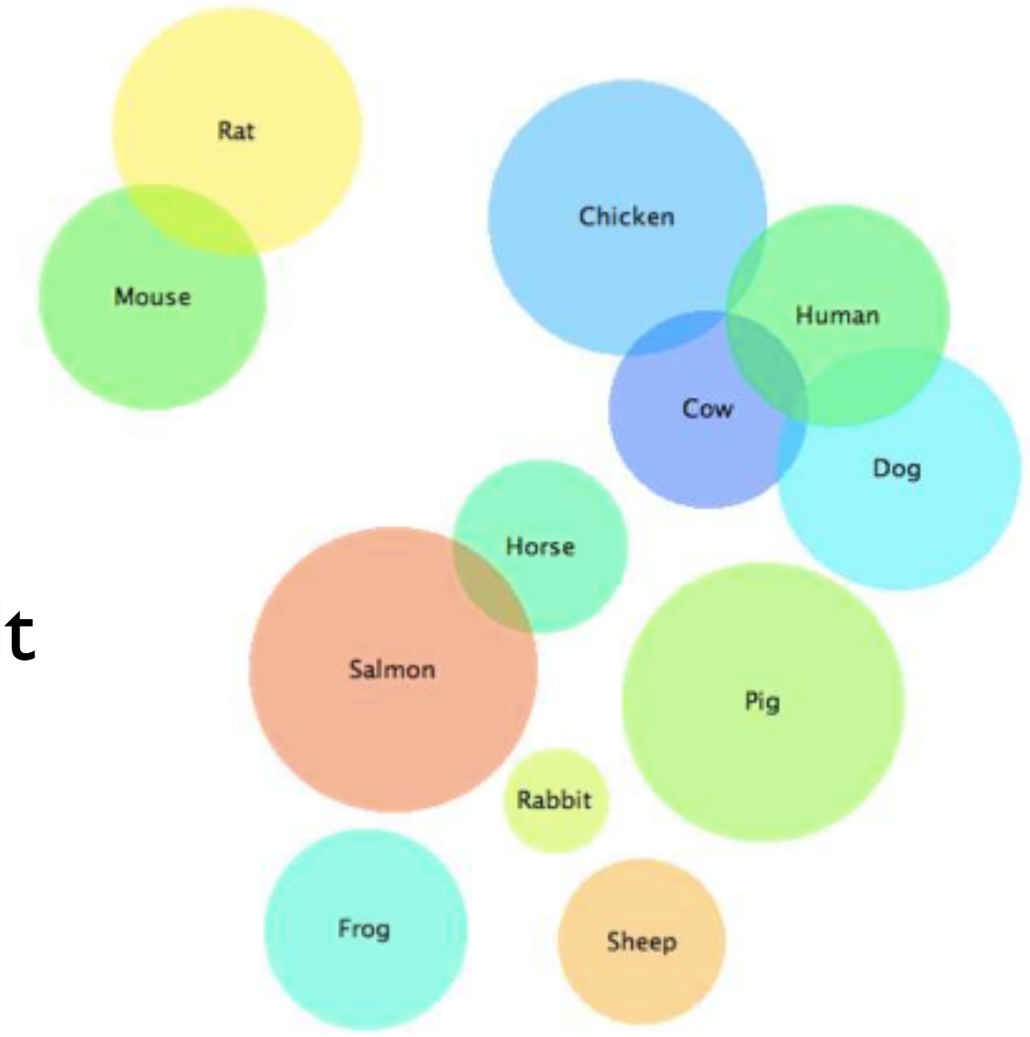


Team:
Helen Lu

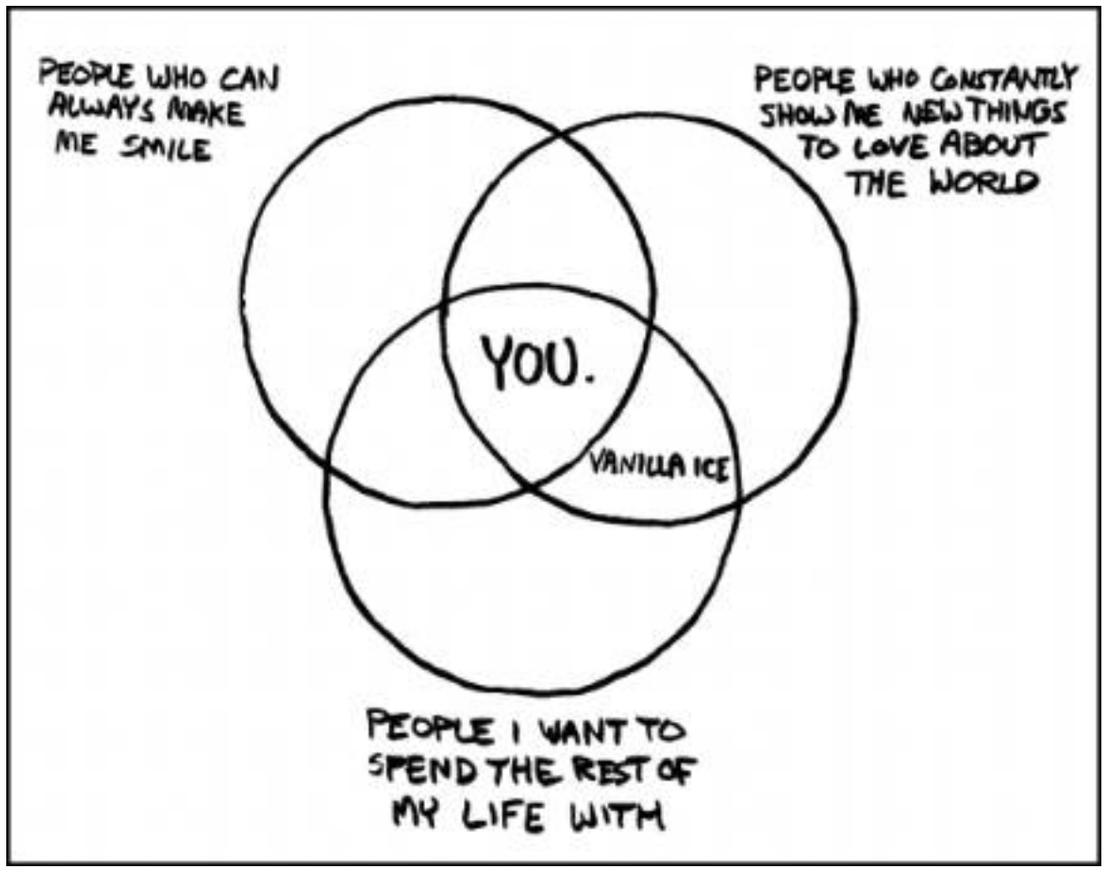
Problem

Area-proportional Venn diagrams show relative sizes of sets and their unions.

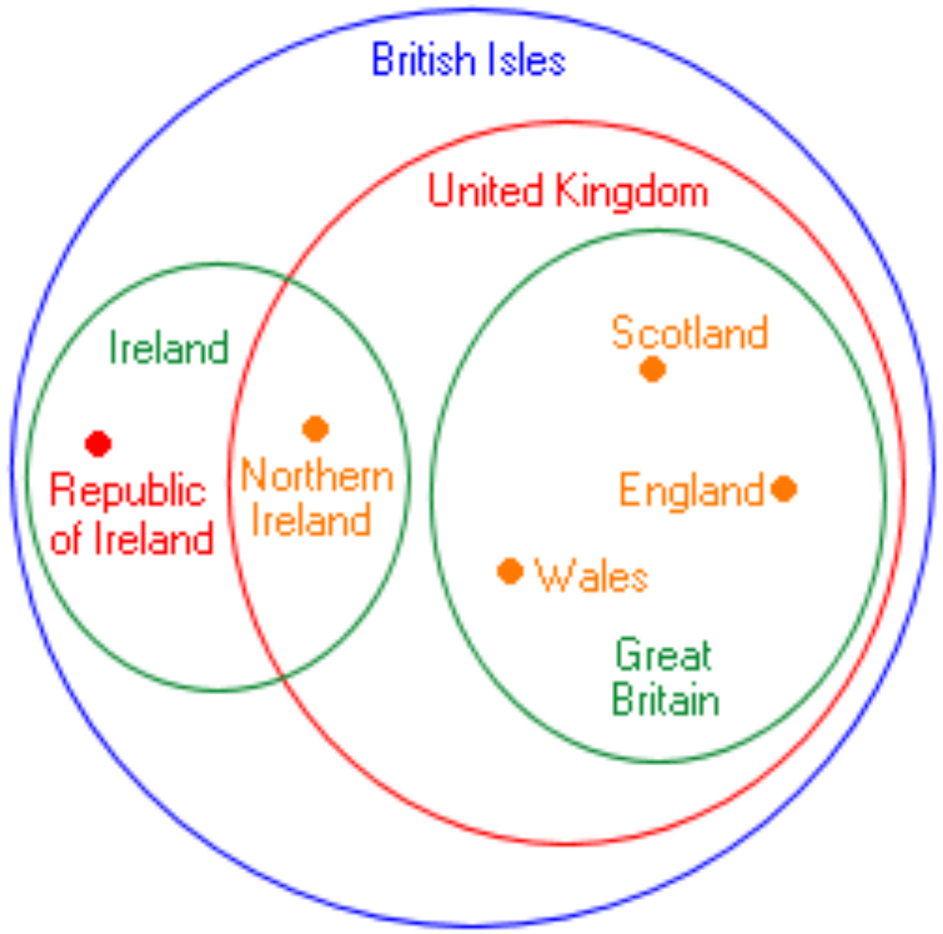
D3 does not include this feature. It is also nontrivial for clients to implement themselves.



Venn & Euler diagrams

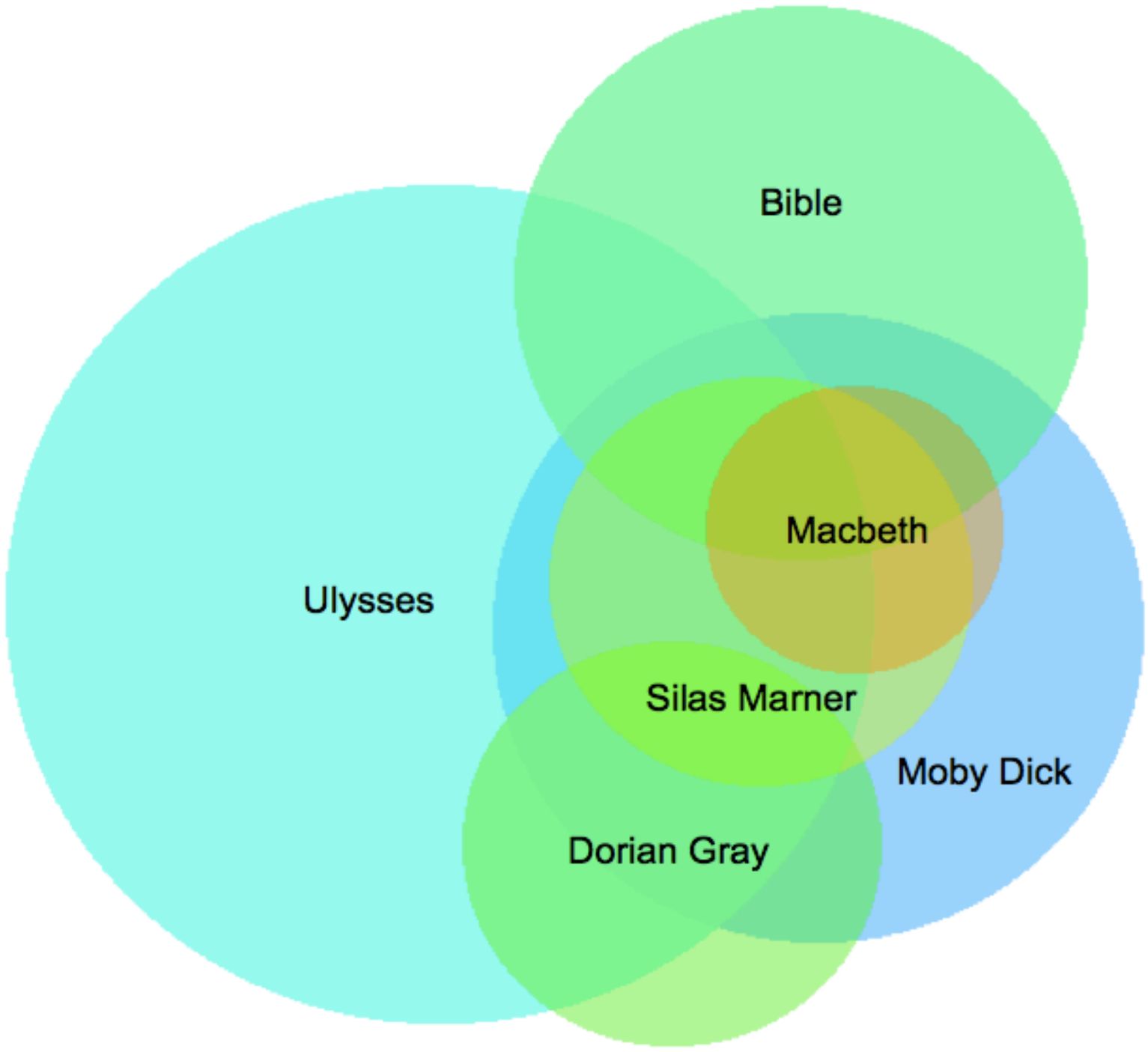


Venn diagram



Euler diagram

Area-proportional Euler diagrams



Prior work

- `venneuler()` - R library
- eulerAPE - Java application, Venn-3 only
- BioVenn - webapp, Venn-3 only
- I will use the same algorithm as `venneuler()`, described in this paper: http://www.cs.uic.edu/~wilkinson/Publications/venneuler.pdf
- Other algorithms exist

Set Image Parameters

title	BioVenn	Courier New bold	24	Black
subtitle	Example diagram	Courier New bold	18	Black
x title	ID Set X	Courier New bold	12	Black
y title	ID Set Y	Courier New bold	12	Black
z title	ID Set Z	Courier New bold	12	Black
print numbers (absolute nrs / percentages)		Courier New bold	12	Black

ID Set X	1007_s_at 1053_at 117_at	<- Copy and paste your IDs .. Or input a file with IDs: Choose File No file chosen	Red
ID Set Y	1255_g_at 1294_at 1316_at	<- Copy and paste your IDs .. Or input a file with IDs: Choose File No file chosen	Lime
ID Set Z	1007_s_at 1405_i_at 1255_g_at	<- Copy and paste your IDs .. Or input a file with IDs: Choose File No file chosen	Blue

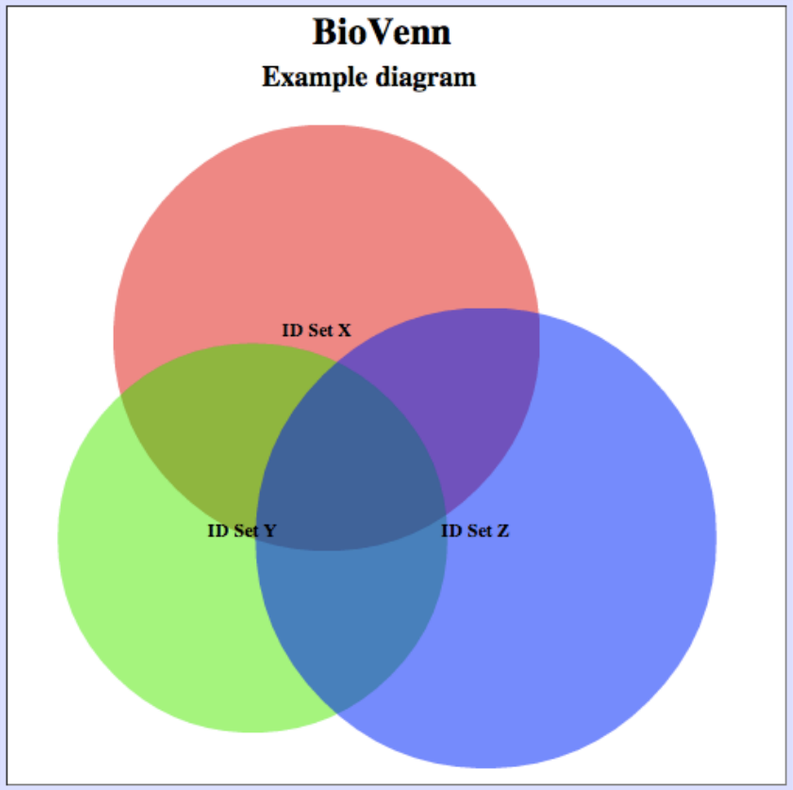
ID Type: Automatic — map Affymetrix/EntrezGene IDs to Ensembl Gene IDs

background transparency — background color: White

image width 500 image height 500

Embedded SVG | Embedded PNG | SVG Only | PNG Only | Reset | Full Reset

(Click here to see an example)

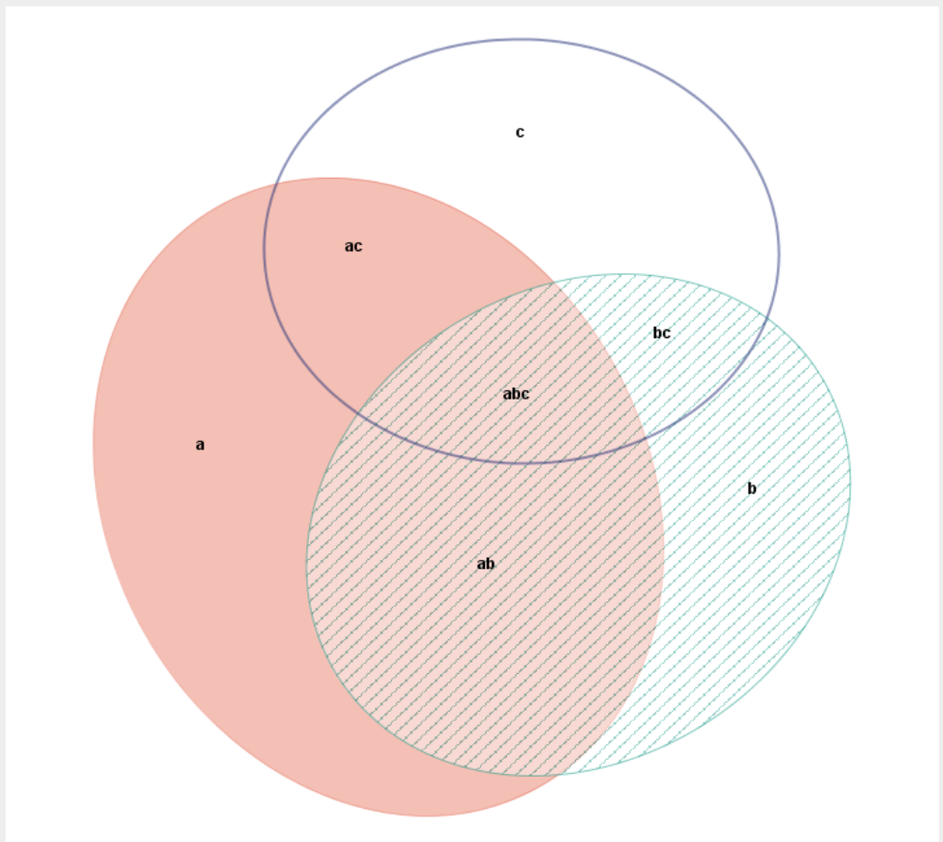


Current Image Statistics

(click on a number to get the corresponding ID list)

x total	6	x only	3	x-y total overlap	2	x-y only overlap	1	x-y-z overlap	1
y total	5	y only	2	x-z total overlap	2	x-z only overlap	1	image width	500
z total	7	z only	4	y-z total overlap	2	y-z only overlap	1	image height	500

eulerAPE: Drawing Area-Proportional Euler and Venn Diagrams using Ellipses



region	required area	actual area	required-actual area
a	35754.013	35754.0395	-0.0265
b	19659.5199	19659.5769	-0.057
c	25875.2088	25875.1624	0.0464
ab	31804.3018	31804.3508	-0.049
ac	12767.035	12767.0325	0.0025
bc	6146.8151	6146.8013	0.0138
abc	10660.841	10660.8571	-0.0161

random | Fitness: 1.0E-7

load | clear all

SAVE TO FILE

Directory C:\eulerAPE2-0 browse

File name example

LABELS yes no COLOUR yes no

CURVES FOR SETS ellipses circles

VIEW SEARCH yes no

Starting Diagram | RUN

Exact 00:00:01:469

New functions for D3

`d3.layout.venn` - Produce an area-proportional Venn/Euler diagram.

Data format:

2D array: [[A1, A2, ... An], [B1, B2, ...Bn], ..., [K1, K2, ...Kn]]

1D array: [|A|, |B|, |A ∩ B|, |C|, |A ∩ C|, |B ∩ C|, |A ∩ B ∩ C|, ...]

`venn.size` - Get or set the width/height of the visualization.

`venn.stress` - Retrieve the stress (goodness-of-fit measure) of the Venn diagram.

Example client use

```
var groups = ["A", "B", "C"];
var data = [33/*A*/, 46/*B*/, 12/*AB*/, 20/*C*/, 18/*AC*/, 7/*BC*/, 1/*ABC*/];
var color = d3.scale.category10();
var venn = d3.layout.venn();
var circle = d3.svg.arc().innerRadius(0).startAngle(0).endAngle(2*Math.PI);

var vis = d3.select("body")
  .append("svg")
    .data([data])
    .attr("width", 800).attr("height", 600);

var circles = vis.selectAll("g.arc")
    .data(venn)
  .enter().append("g")
    .attr("class", "arc")
    .attr("transform", function(d, i){ return "translate(" + d.x + "," + d.y + ")"; });
circles.append("path")
    .attr("fill", function(d, i) { return color(i); })
    .attr("opacity", 0.6)
    .attr("d", circle);
circles.append("text")
    .attr("text-anchor", "middle")
    .text(function(d, i) { return groups[i]; });
```

Questions

- If a client wants to create an area-proportional Venn diagram, what format might the data be in?
- Do the function calls make sense?
- What are some desirable extra features?