

VISUALISING DATA: INTRODUCTION AND COURSE OUTLINE

National Workshop
Accra, Ghana

What we are NOT

- 21st district in Paris
- (only) statistical nerds
- OECD - DAC



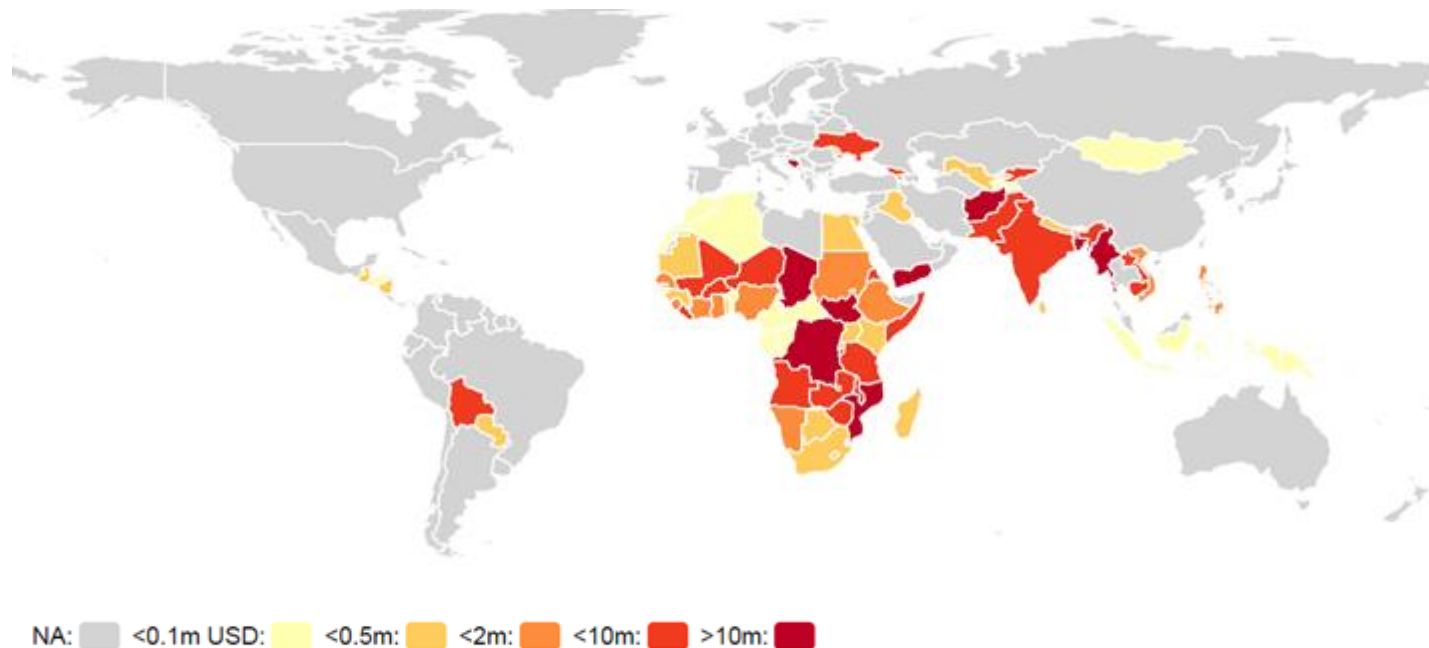
Partnership on stats capacity development

National & Regional

- Strategic planning (NSDS & RSDA)
- Advocacy
- Data (e.g. micro-data dissemination)

Global

- Co-ordination (BAPS, PRESS)
- Knowledge sharing





The data deluge

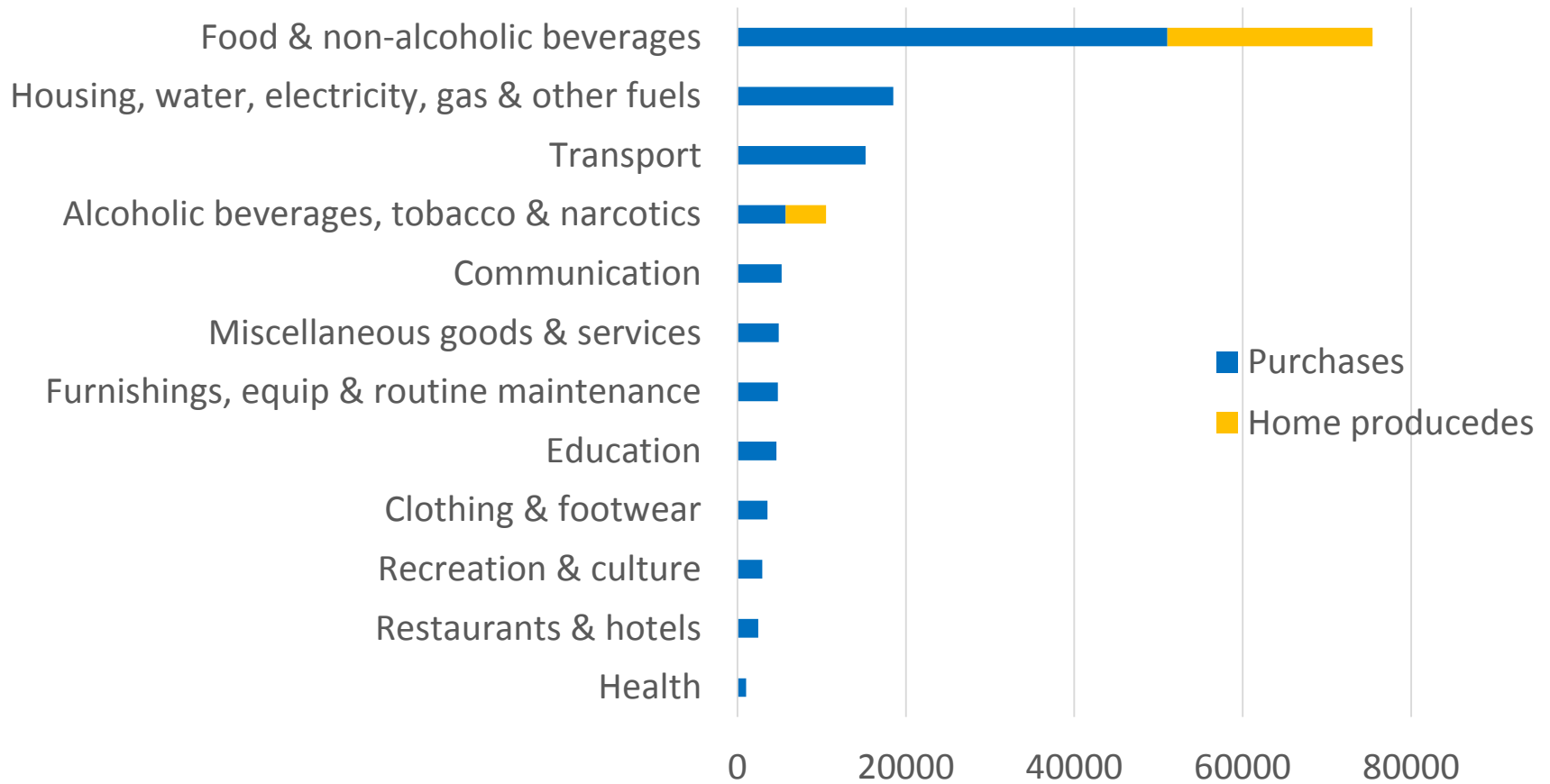
Brett Ryder - <http://www.econcnst.com/node/155797/17>

Household expenditure (1)

Table 5: Total annual household expenditure by expense category (US\$ 000)

Expense category	Purchase	Home produced
1 - Food & non-alcoholic beverages	51042	24374
2 - Alcoholic beverages, tobacco & narcotics	5739	4774
3 - Clothing & footwear	3580	0
4 - Housing, water, electricity, gas & other fuels	18515	0
5 - Furnishings, equip & routine maintenance	4806	0
6 - Health	1030	0
7 - Transport	15212	0
8 - Communication	5236	0
9 - Recreation & culture	2939	0
10 - Education	4619	0
11 - Restaurants & hotels	2457	0
TOTAL CONSUMPTION EXPENDITURE	120063	29148

Household expenditure (2)



Household expenditure (3)



How much do households spend per week?

Explore the data behind Family Spending 2013, UK

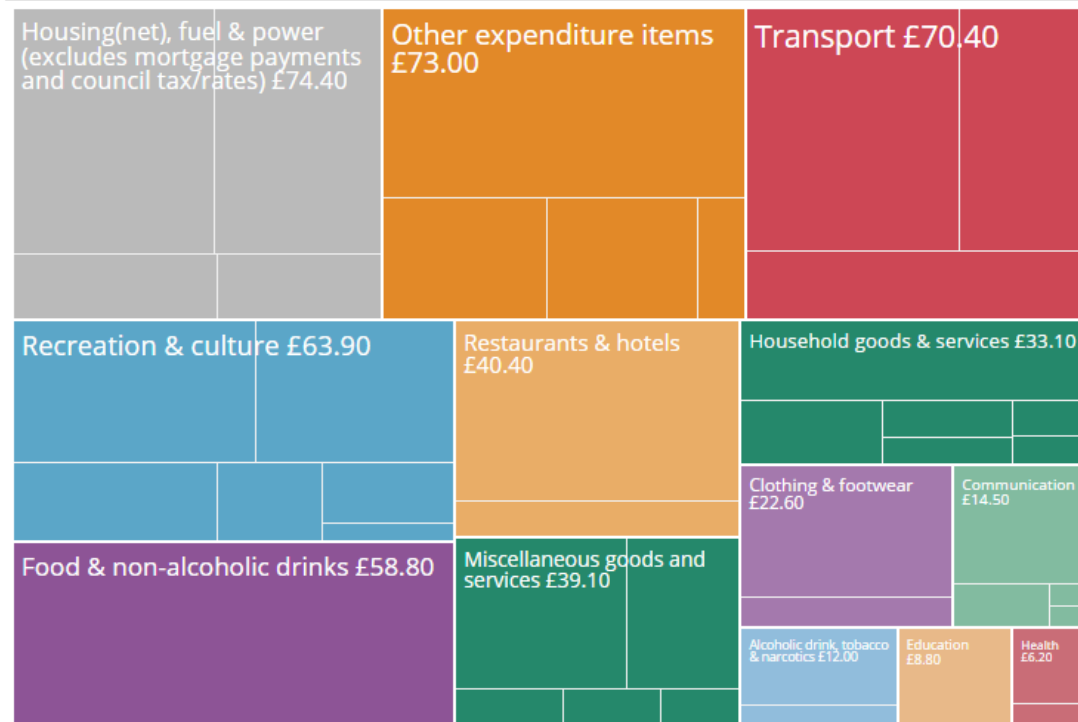
Show intro

Total expenditure

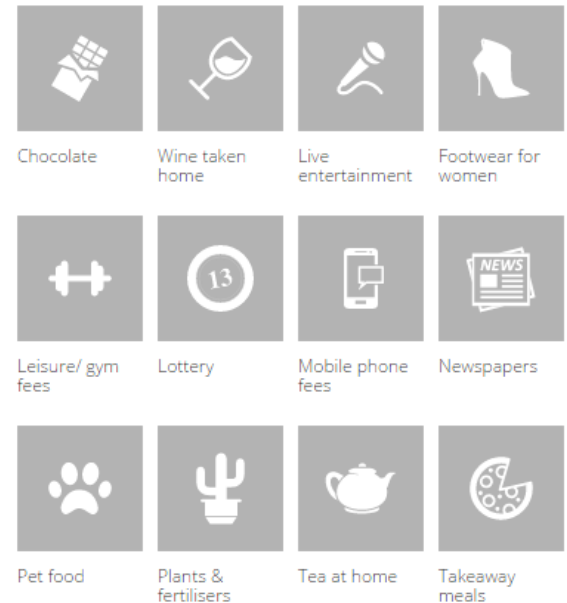
Up a level

Back to top

Housing(net), fuel & power (excludes mortgage payments and council tax/rates) > Average household spend of £74.40 per week.



How much did we spend on...



Source: ons.gov.uk/ons/interactive/

Outline

1. Examples
2. Visualisation tools
3. R as interface



1. Examples

How well do you know your area?

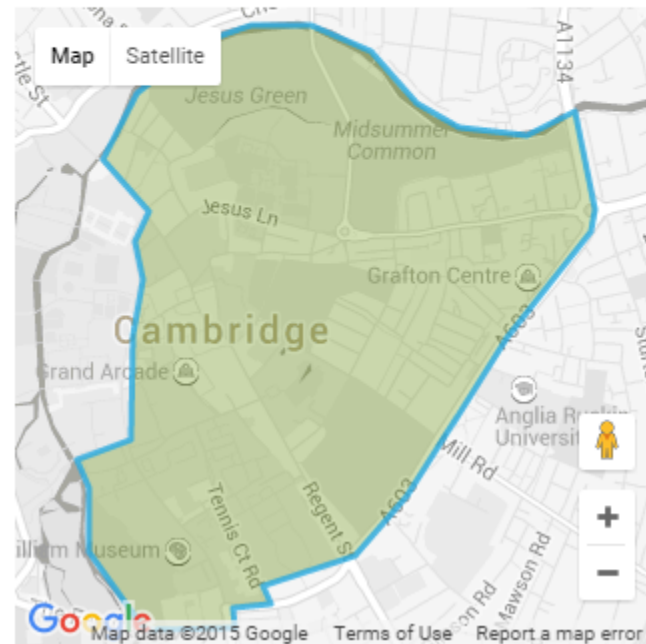
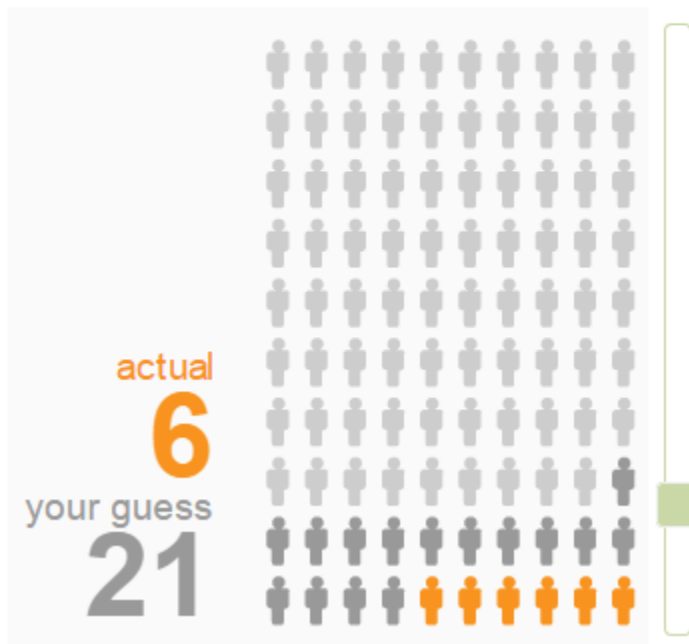
Quiz prepared for the ward of **Market** in **Cambridge**

Question 1 of 7 about Market:

For every 100 people, how many are aged under 16?

You were **15** over the actual value - press 'next' to continue...

next



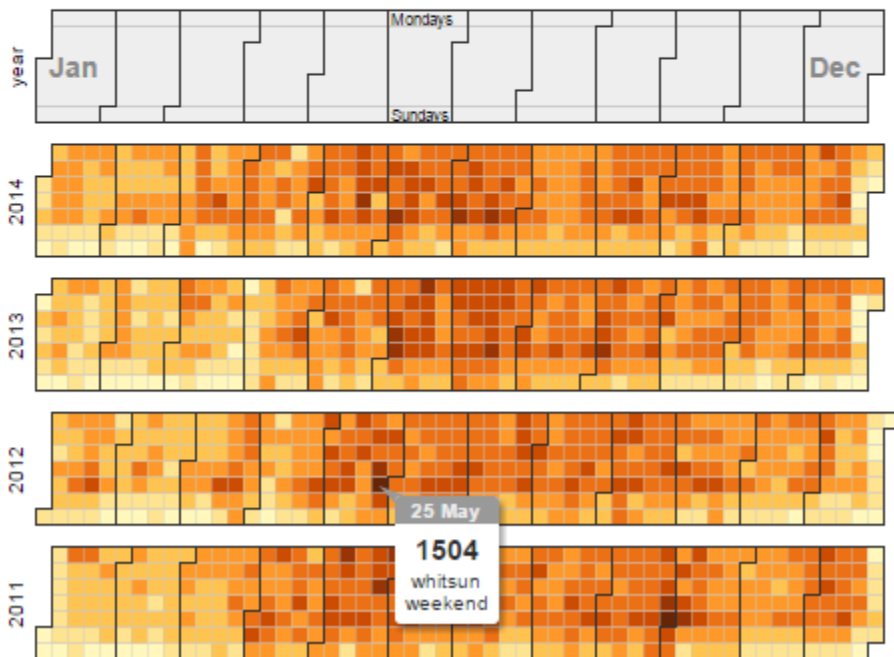
Road traffic accidents

Road traffic accidents: the calendars for 2014 and previous years

Compare accidents involving personal injury to alcohol-related accidents
 motorcycle accidents

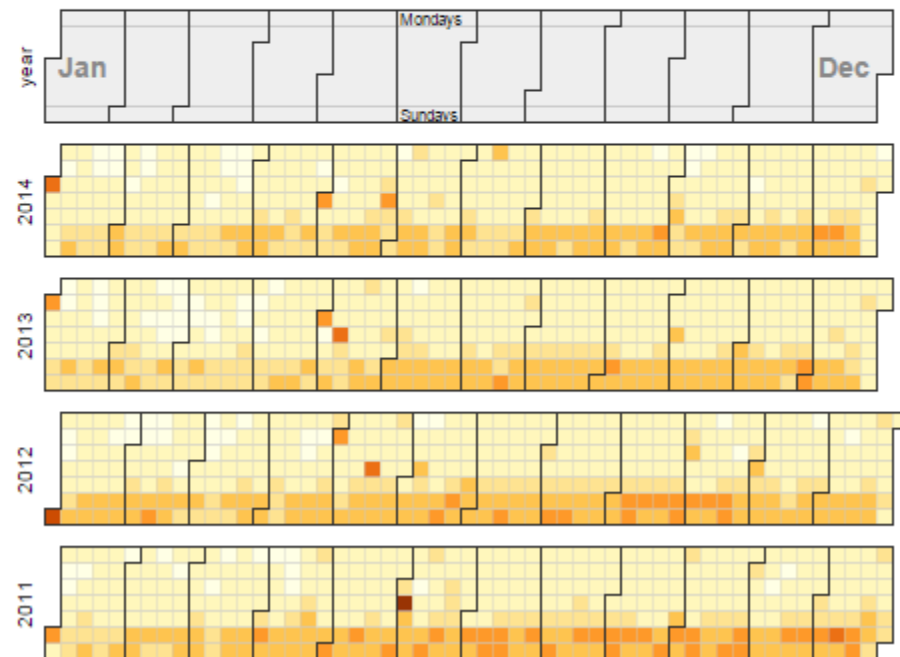
Personal injury: downward trend, but very frequent on Fridays

Friday, 30 September 2011, saw 1640 of such accidents, which was a sad record in the last ten years.



Alcohol-related accidents: be careful on New Year's Day and at weekends

The most alcohol-related accidents (458) happened on 20 May 2004 (Father's Day).



OECD Better Life Index

English

Español

Français

Русский

Deutsch

Português

Italiano

Contact Us

oecd.org

Index

Responses

Countries ▾

Topics ▾

FAQ

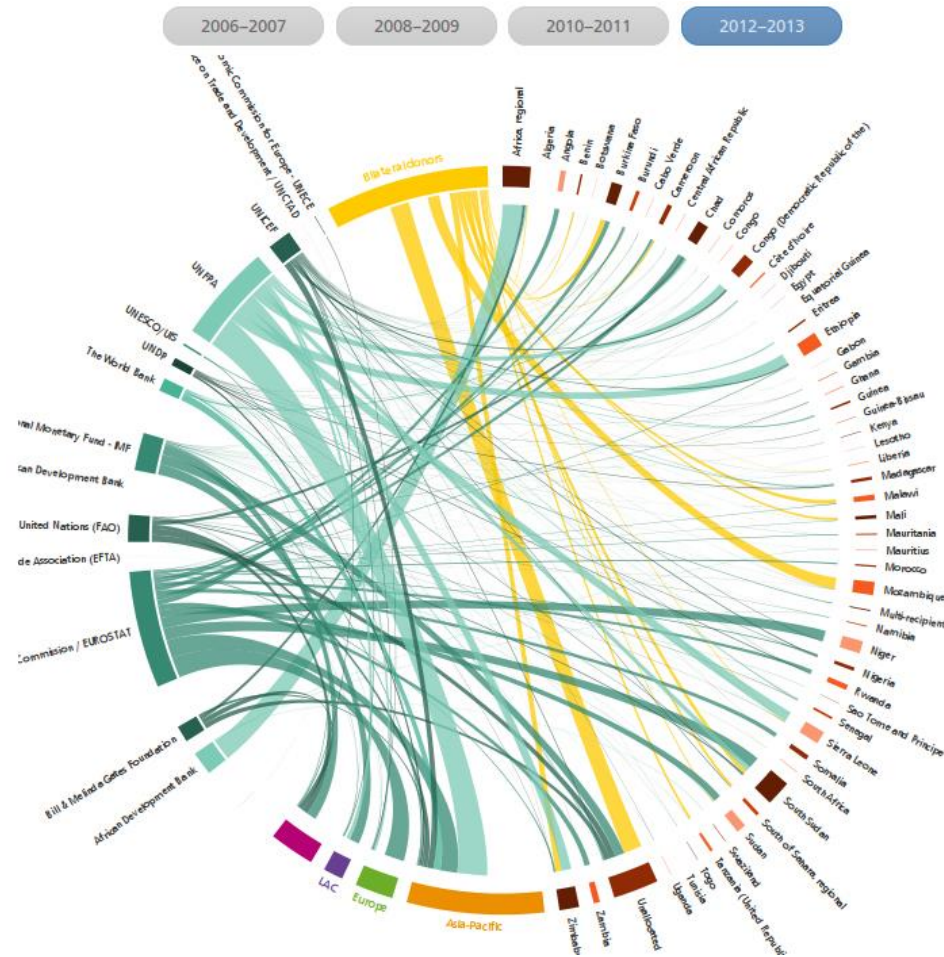


Create Your Better Life Index

Rate the topics according to their importance to you:

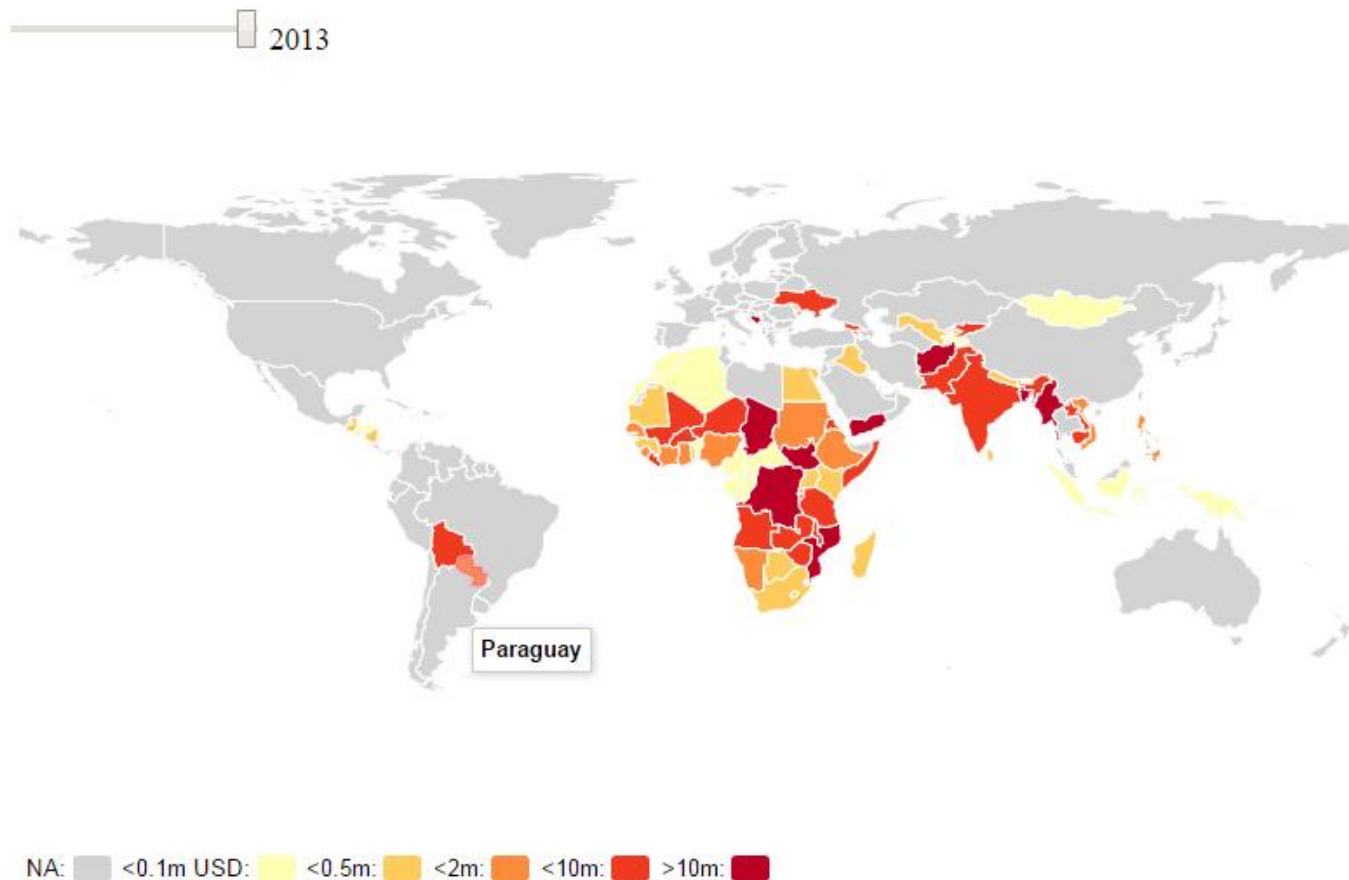
	-	+
 Housing		<input type="checkbox"/>
 Income		<input type="checkbox"/>
 Jobs		<input type="checkbox"/>
 Community		<input type="checkbox"/>
 Education		<input type="checkbox"/>

Global aid flows



Global aid recipients

Who were the main beneficiaries from 2006 to 2013?



Aims

- Note: good graphs are self-explanatory
 - The key to understand a graph should not be hidden somewhere in the text!



2. Visualisation tools

Visualisation tools

- What are we looking for in a viz software?
 - Quality output
 - Vector graphics (SVG) vs PNG/JPEG
 - Interactive output
 - JavaScript
 - “Free” (as in free beer)
 - Well documented
 - Convenient/ quick to learn

A comparison of tools

	Excel	Raw/ Datawrapper	Tableau	D3.js/ Google Charts	R
Quality	✗	✓	✓	✓	✓
Interactive	✗	✗	✓	✓	✓
Free	✗	✓	✗	✓	✓
Documentation	✓	✓	✓	✓	(✓)
Convenience	✓	✓	✓	✗	(✓)

Sources:

raw.densitydesign.org / datawrapper.de

tableau.com

d3js.org

r-project.org

Google Charts: pie chart

Charts

HOME

GUIDES

REFERENCE

SUPPORT

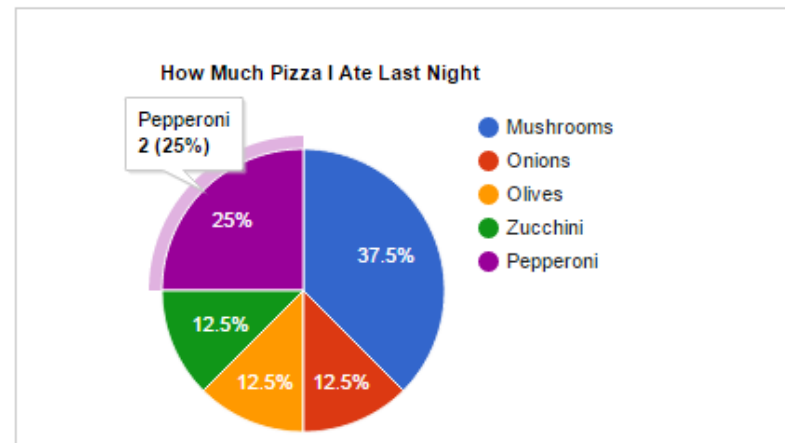
Load the Libraries
Prepare the Data
Customize the Chart
Draw the Chart

Chart Types

Chart Gallery
Annotation Charts
Area Charts
Bar Charts
Bubble Charts
Calendar Charts
Candlestick Charts
Column Charts
Combo Charts

Quick Start

Here's a simple example of a page that displays a [pie chart](#):



Google Charts: code comparison

- Ajax code

```
<html>
<head>
  <!--Load the AJAX API-->
  <script type="text/javascript" src="https://www.google.com/jsapi"></script>
  <script type="text/javascript">

    // Load the Visualization API and the piechart package.
    google.load('visualization', '1.0', {'packages':['corechart']});

    // Set a callback to run when the Google Visualization API is loaded.
    google.setOnLoadCallback(drawChart);

    // Callback that creates and populates a data table,
    // instantiates the pie chart, passes in the data and
    // draws it.
    function drawChart() {

      // Create the data table.
      var data = new google.visualization.DataTable();
      data.addColumn('string', 'Topping');
      data.addColumn('number', 'Slices');
      data.addRows([
        ['Mushrooms', 3],
        ['Onions', 1],
        ['Olives', 1],
        ['Zucchini', 1],
        ['Pepperoni', 2]
      ]);

      // Set chart options
      var options = {'title':'How Much Pizza I Ate Last Night',
                    'width':400,
                    'height':300};

      // Instantiate and draw our chart, passing in some options.
      var chart = new google.visualization.PieChart(document.getElementById('chart_div'));
      chart.draw(data, options);
    }
  </script>
</head>

<body>
  <!--Div that will hold the pie chart-->
  <div id="chart_div"></div>
</body>
</html>
```

- R code

```
## read data
```

```
pizza = read.csv("http://klein.uk/R/pizza.csv")
```

```
## install googleVis package
```

```
install.packages("googleVis")
```

```
## create pie chart
```

```
library(googleVis)
```

```
pizza = gvisPieChart(pizza)
```

```
plot(pizza)
```

```
## find help
```

```
?gvisPieChart
```



3. R as Interface

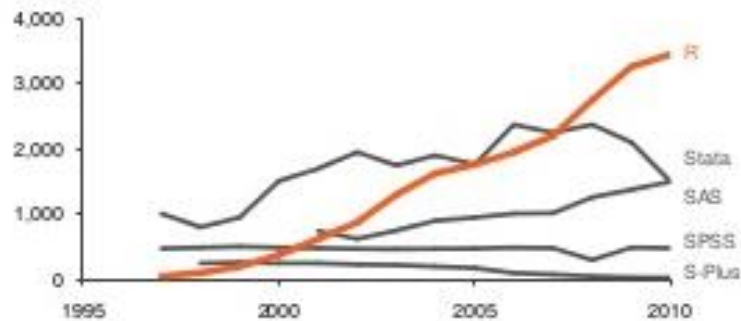
R as interface

- Many open source libraries
 - [googleVis](#)
 - [rCharts](#)
 - [rMaps](#)
 - [leaflet](#)
 - [htmlwidgets](#)
 - [plotly](#)
- R combines the functionality of these packages in one interface

R is exploding in popularity & functionality

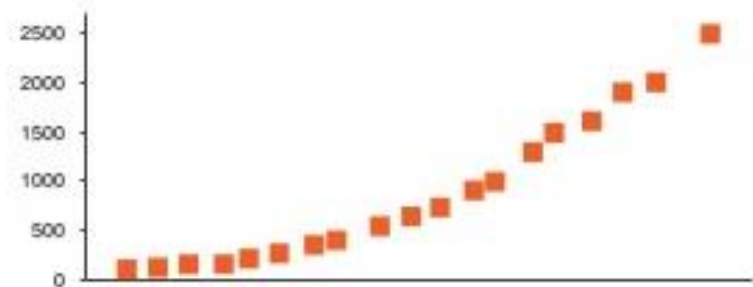
Internet Discussion

Mean monthly traffic on email discussion list



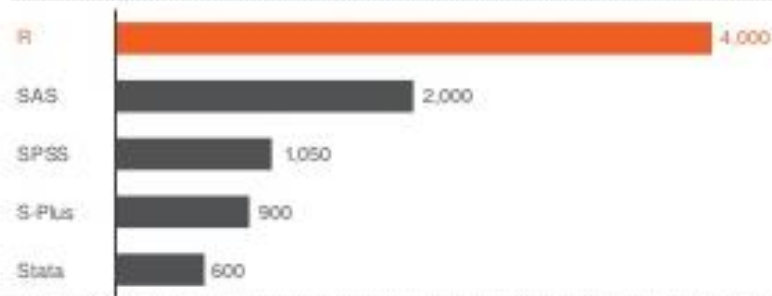
Package Growth

Number of R packages listed on CRAN



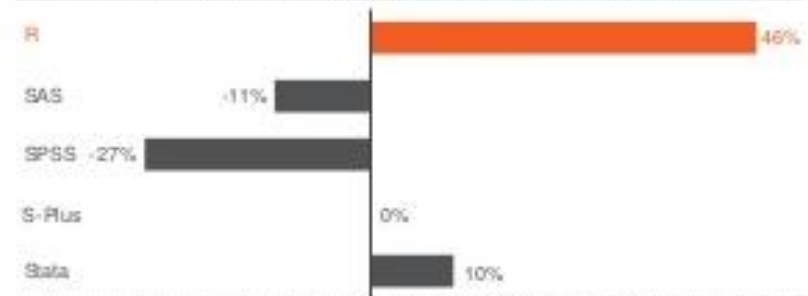
Web Site Popularity

Number of links to main web site



Scholarly Activity

Google Scholar hits (05-'09 CAGR)



Your workflow

1. use your tools of choice for data manipulation and analysis
2. import into R from Excel, Stata, SAS, SPSS, ... using R libraries 'xlsx' or 'foreign'
3. produce graphics using code snippets available online
4. export in static format (SVG, PNG, JPG) or dynamic graphs (HTML, JavaScript)
5. embed in your reports and websites

Training outline

- Good practices and workflow of statistical data analysis
- Hands-on group work
- Producing reports and blogs using knitr
- Next steps: finding help and resources