

Tools for Data Visualisation

Python and R

Dr. Rebecca Lange
Curtin Institute for Computation



Why use R
or Python?



Python and R are supported by a large open source community

The screenshot shows the Stack Overflow interface for the 'python' tag. The top navigation bar includes links for Questions, Developer Jobs, Documentation (marked BETA), Tags, Users, and a search bar containing '[python]'. On the right, there are links for Log In and Sign Up. The 'python' tag is highlighted in the top navigation. Below the navigation bar, the 'Tag Info' section is active, showing 'info', 'newest', '30 featured', 'frequent', 'votes', 'active', and 'unanswered' tabs. The 'About python' section contains a text box with the following text: 'Python is a dynamic and strongly typed programming language designed to emphasize usability. Two similar but mostly incompatible versions of Python are in widespread use (2 and 3). If you have a version-specific Python question consider using the [python-2.7] or [python-3.x] tags in addition to the [python] tag. If you are using a python variant such as jython, pypy, iron-python, etc., please tag appropriately.' Below this, a paragraph states: 'Python is a [dynamic and strongly typed](#) programming language that is used for [a wide range of applications](#). It is a general-purpose, high-level programming language that is designed to emphasize usability.' Another paragraph follows: 'Python allows programmers to express concepts in fewer lines of code than would be possible in many other languages such as C, and the language has constructs intended to be used to create clear programs in a variety of domains.' The final sentence reads: 'Python was created by [Guido van Rossum](#).' To the right of the 'Tag Info' section, the statistics for the 'python' tag are displayed: '797,595 questions tagged', 'python' (tag name), 'Synonyms' (pythonic, python-interpreter, python-shell), 'Stats' (created 9 years ago, viewed 77080 times, active 10 days ago, editors 127), and 'Top Answerers' (listing Martin Dorey).

The screenshot shows the Stack Overflow interface for the 'r' tag. The top navigation bar includes links for Questions, Developer Jobs, Tags, Users, and a search bar containing '[r]'. On the right, there are links for Log In and Sign Up. The 'r' tag is highlighted in the top navigation. Below the navigation bar, the 'Tag Info' section is active, showing 'info', 'newest', '7 featured', 'frequent', 'votes', 'active', and 'unanswered' tabs. The 'About r' section contains a text box with the following text: 'R is a free, open-source programming language and software environment for statistical computing, bioinformatics, visualization and general computing. Provide minimal, reproducible, representative example(s) with your questions. Use dput() for data and specify all non-base packages with library calls. Do not embed pictures for data or code, use indented code blocks. For statistics questions, use <http://stats.stackexchange.com>.' Below this, the section is titled 'R Programming Language'. A paragraph follows: 'R is a free, open-source programming language and software environment for [statistical computing](#), [bioinformatics](#), [information graphics](#), and general computing. It is a multi-paradigm language and dynamically typed. R is an implementation of the [S programming language](#) combined with lexical scoping semantics inspired by [Scheme](#). R was created by [Ross Ihaka](#) and [Robert Gentleman](#) and is now developed by the [R Development Core Team](#). The R environment is easily extended through a packaging system on [CRAN](#), the Comprehensive R Archive Network.' To the right of the 'Tag Info' section, the statistics for the 'r' tag are displayed: '238,005 questions tagged', 'r' (tag name), 'Synonyms' (rstats, r-language), 'Stats' (created 9 years, 8 months ago by David Locke, viewed 65369 times, active 20 days ago, editors 80), and 'Top Answerers'.

Python and R are free

- No paywalls
 - No subscription costs
 - Open source software is free (as in freedom)
- ➡ everyone can afford their vegetables this week!

The code is readable

JAVA

```
public class Main {  
    public static void main(String[] args) {  
        System.out.println("hello world");  
    }  
}
```

PYTHON

```
print('hello world')
```

Analysis is scalable

```
fig = plt.figure(figsize=(20, 30))
markersize = 1
markertype = ',' # pixel
markercolor = '#325CA9' # blue
markeralpha = .8 # a bit of transparency

m = Basemap(
    projection='mill', lon_0=bbox['lon'], lat_0=bbox['lat'],
    llcrnrlon=bbox['ll_lon'], llcrnrlat=bbox['ll_lat'],
    urcrnrlon=bbox['ur_lon'], urcrnrlat=bbox['ur_lat'])

# Avoid border around map.
m.drawmapboundary(fill_color='ffffff', linewidth=0)

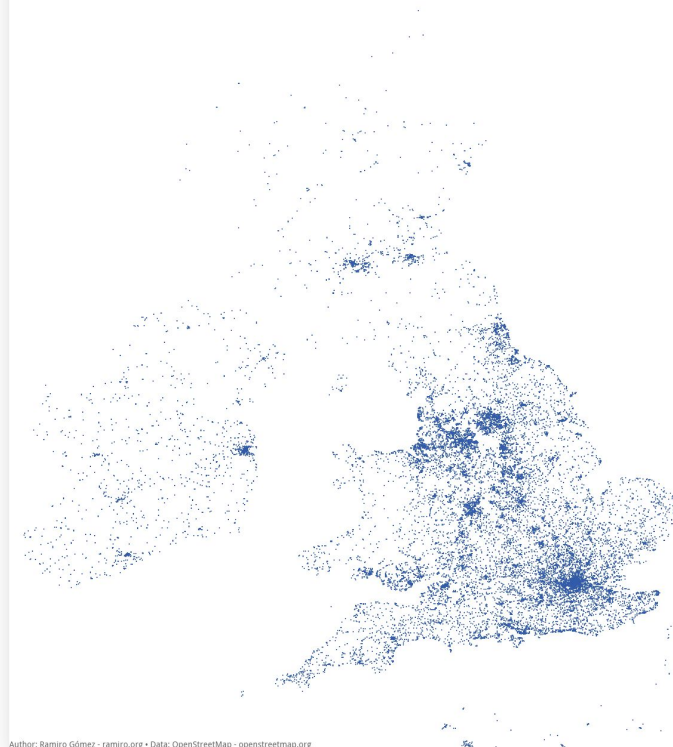
# Convert locations to x/y coordinates and plot them as dots.
lons, lats = zip(*locations)
x, y = m(lons, lats)
m.scatter(x, y, markersize=marker, color=markercolor, alpha=markeralpha)

# Set the map title.
plt.annotate('Britain & Ireland\ndrawn from pubs',
            xy=(0, .87),
            size=120,
            xycoords='axes fraction',
            color='#888888',
            family='Gloria')

# Set the map footer.
plt.annotate('Author: Ramiro Gómez - ramiro.org • Data: OpenStreetMap - openstreetmap.org',
            xy=(0, 0),
            size=14,
            xycoords='axes fraction',
            color='#666666',
            family='Droid Sans')

plt.savefig('img/britain-ireland-drawn-from-pubs.png', bbox_inches='tight')
```

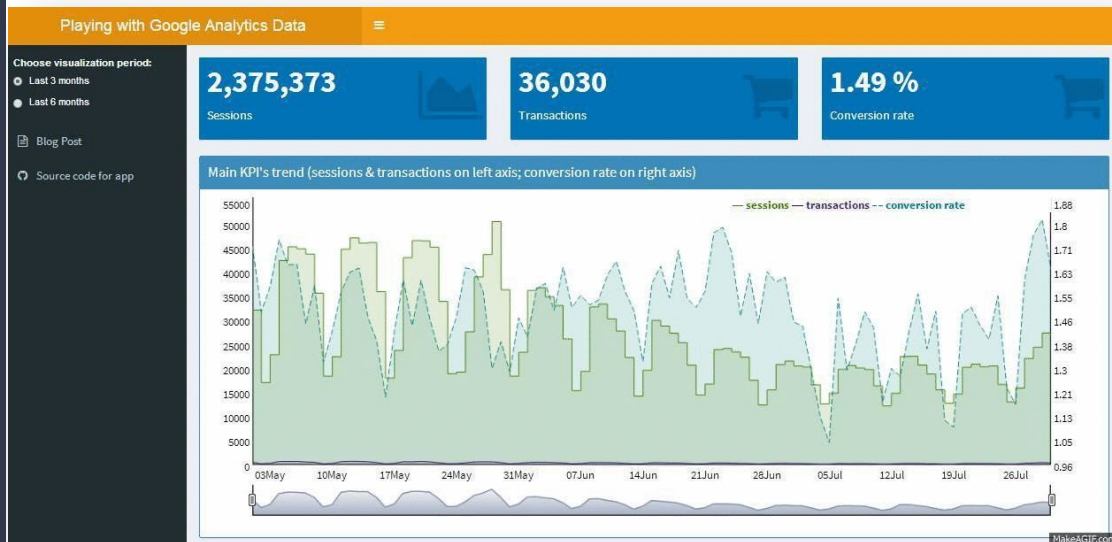
Britain & Ireland
drawn from pubs



Author: Ramiro Gómez - ramiro.org • Data: OpenStreetMap - openstreetmap.org

<http://ramiro.org/notebook/mapping-pubs/>

Results are shareable



Visualisation examples



Visualisation Examples

The **type of variable** determines the **kind of visual**, e.g.:

- Diagrams
 - barcharts, histograms, boxplots
- Plots
 - simple plots, trend plots, scatter plots
- Figures
 - photos, drawing, schematic

Visualisation Examples

Live coding visualisation examples

Python (seaborn) - https://github.com/CurtinIC/GRASP_workshop/blob/master/Python%20Vis%20Examples.ipynb

R (ggplot2) - https://github.com/CurtinIC/GRASP_workshop/blob/master/R%20Vis%20Examples.ipynb

Interactive plots examples

Python (Bokeh) - <https://demo.bokeh.org/gapminder>

R (ggplot2 + googlevis) - http://remi-daigle.github.io/2016-04-15-UCSB/viz/#14_interactive:_motion_plot

3D plotting examples

Python (matplotlib) - <https://jakevdp.github.io/PythonDataScienceHandbook/04.12-three-dimensional-plotting.html>

R (plotly) - <https://plot.ly/r/3d-scatter-plots/>

Support from the CIC!



- **CIC drop-in session**

- Every Tuesday 3pm-4pm, CIC office, B216:202

- **ResBaz**

- ResBaz is a worldwide festival promoting the digital literacy emerging at the center of modern research
- 2 - 4 July @ Curtin Medical school
- <https://resbaz.github.io/resbaz2019/perth/>

- **CIC Workshops**

- We teach Carpentries workshops and introduction to ML/DL
- Sign up to our newsletter



@CurtinIC



www.computation.curtin.edu.au

Join our newsletter!



curtinic@curtin.edu.au