ggplot2

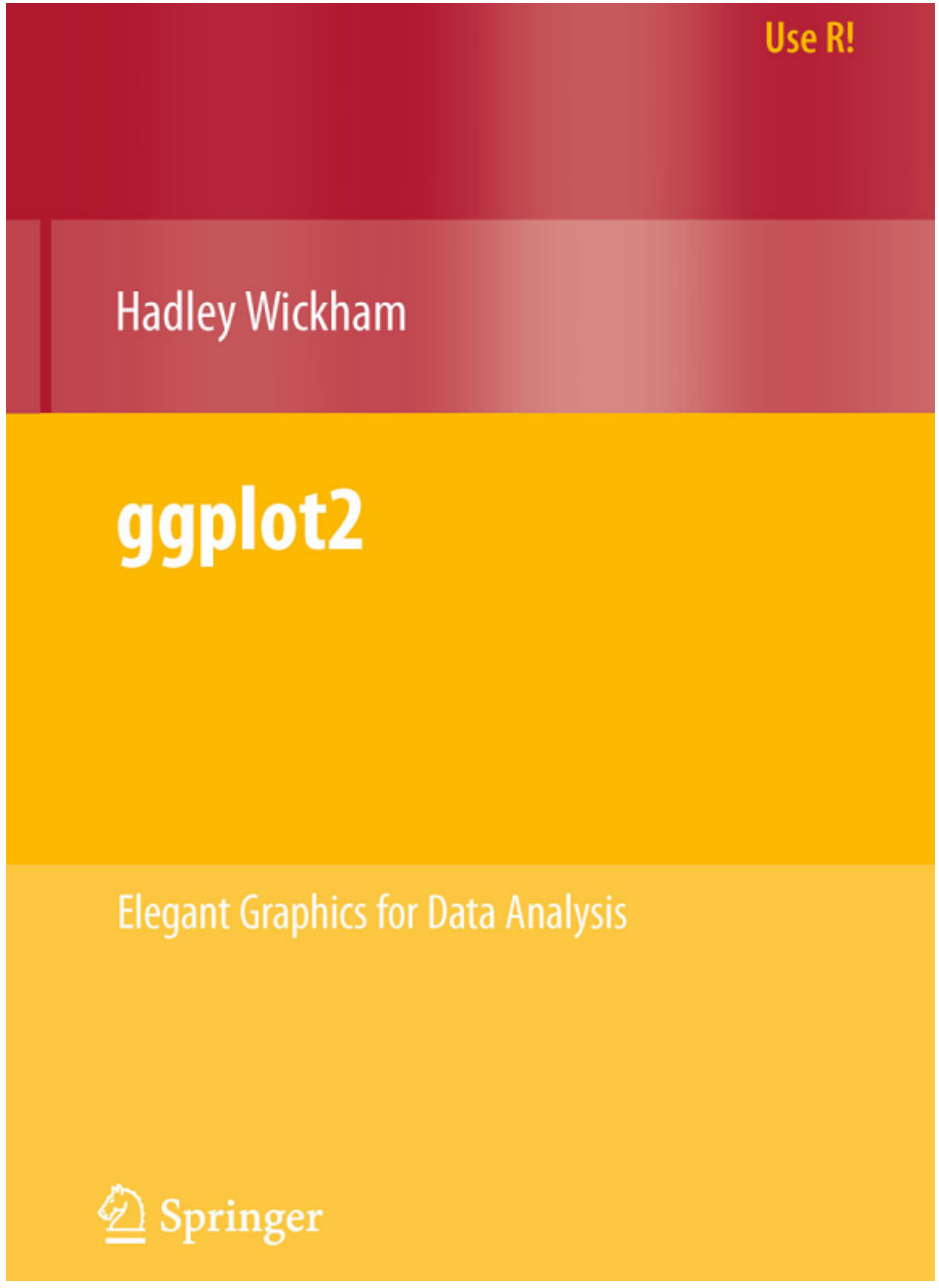
ggplot2 is a plotting system for R, based on the grammar of graphics, which tries to take the good parts of base and lattice graphics and none of the bad parts. It takes care of many of the fiddly details that make plotting a hassle (like drawing legends) as well as providing a powerful model of graphics that makes it easy to produce complex multi-layered graphics.

Documentation

ggplot2 documentation is now available at docs.ggplot2.org.

website: www.had.co.nz

Hadley Wickham (Chief Data Scientist, RStudio)



ggplot2: Elegant Graphics for Data Analysis

ggplot from ŷhat

About | Installation | How It Works | Docs | Gallery

`ggplot` is a plotting system for Python based on R's `ggplot2` and the *Grammar of Graphics*. It is built for making profressional looking, plots quickly with minimal code.

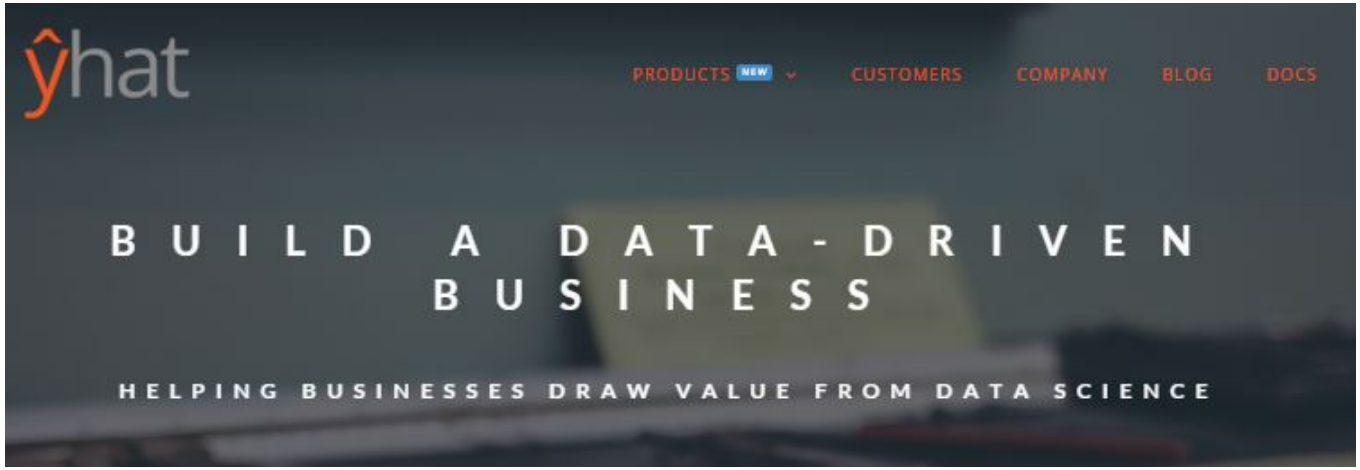
`ggplot` is easy to learn

```
from ggplot import *

ggplot(aes(x='date', y='beef'), data=meat) +\
    geom_line() +\
    stat_smooth(colour='blue', span=0.2)
```

Important :

for Python, the name is simply "**ggplot**".



Yhat (*pronounced y-hat*) is a data science technology company that provides tools and systems that allow enterprises to turn data insights into data-driven products.

About ggplot

ggplot is a graphics package for Python that aims to approximate R's ggplot2 package in both usage and aesthetics.

Authors: Greg Lamp and Austin Ogilvie

What are Yhat saying?

1. ggplot is easy to learn [*1*]
2. ggplot is fun
3. ggplot is powerful [*2*]

[1] Lots of learning resources, mainly intended for the R environment, that can applied to Python also.

[2] Less code required to compute high-level publication quality plot

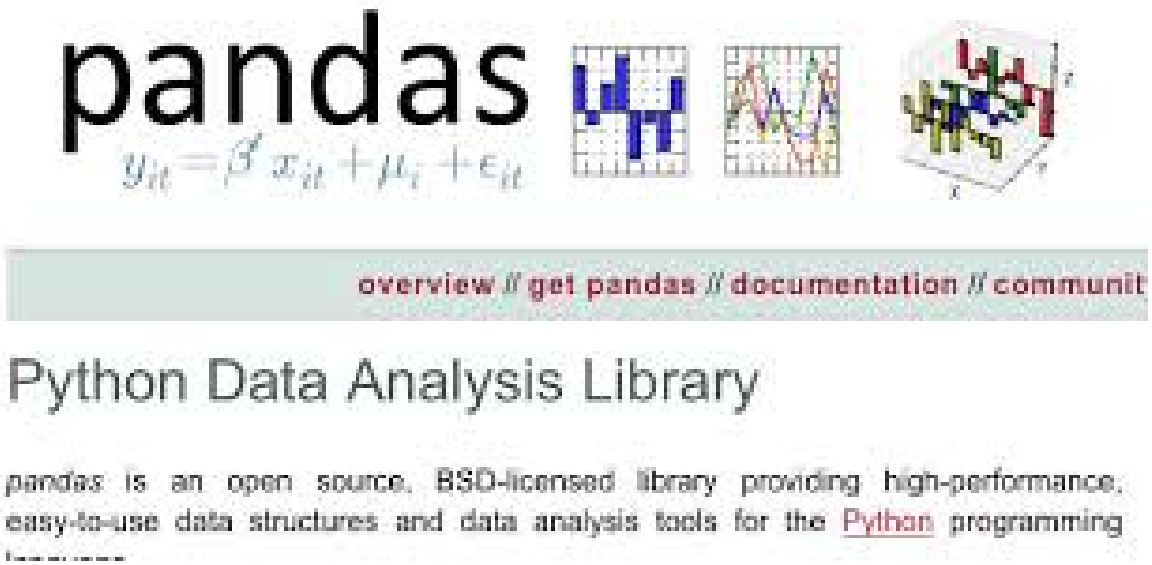
1. Installing ggplot

```
pip install ggplot
```

2. Getting Set Up on Jupyter Notebook

```
In [5]: from ggplot import *

In [6]: %matplotlib inline
```



Important: ggplot accepts data in the form of a `pandas Dataframe`, so you need to configure all data accordingly first.

Watch out for this operator

```
.... +\ .....
```

```
ggplot(aes(x='date', y='beef'), data=meat) +\
    geom_line() +\
    stat_smooth(colour='blue', span=0.2)
```

- The main command is `ggplot()`.
- The name comes from "**grammar of graphics**", a book by Leland Wilkinson
- A very "high-level" approach to data visualization.

> *A grammar of graphics is a tool that enables us to concisely describe the components of a graphic. Such a grammar allows us to move beyond named graphics (e.g., the scatterplot) and gain insight into the deep structure that underlies statistical graphics.*

A Layered Grammar of Graphics

Hadley WICKHAM

A grammar of graphics is a tool that enables us to concisely describe the components of a graphic. Such a grammar allows us to move beyond named graphics (e.g., the "scatterplot") and gain insight into the deep structure that underlies statistical graphics. This article builds on Wilkinson, Anand, and Grossman (2005), describing extensions and refinements developed while building an open source implementation of the grammar of graphics for R, `ggplot2`.

Hadley Wickham.
A layered grammar of graphics.
Journal of Computational and Graphical Statistics,
vol. 19, no. 1, pp. 328, 2010.

A **plot** is made up of multiple layers.

A **layer** consists of **data**, a set of **mappings** between variables and aesthetics, a **geom**etric object and a **stat**istical transformation

Scales control the details of the mapping.

All components are independent and reusable.