

## Part 2 : Other Interesting Python Packages

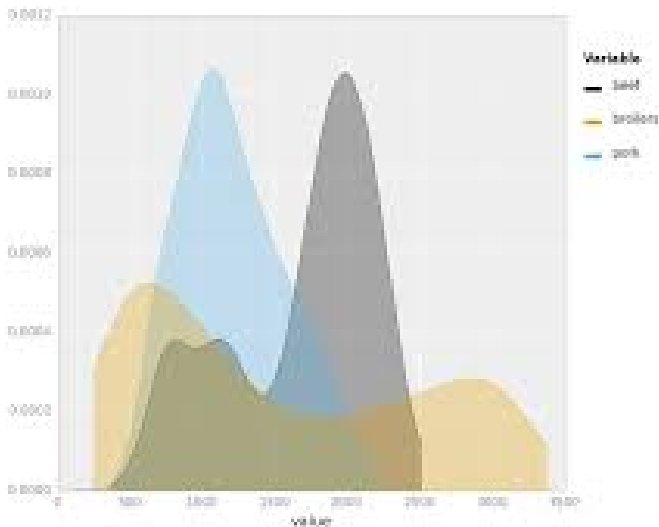
# statsmodels

- ▶ statsmodels provides a large range of cross-sectional models aswell as time-series models.
- ▶ statsmodels uses a model descriptive language (provided via the Python package patsy) to formulate the model when working with pandas DataFrames.
- ▶ Models supported include linear regression, generalized linear models, limited dependent variable models, ARMA and VAR models.

# Bokeh Data Visualization



# Bokeh Data Visualization



# Bokeh Data Visualization

- ▶ interactive graphics for the web
- ▶ designed for large data sets
- ▶ Designed for streaming data
- ▶ Native interface in python
- ▶ Fast javascript components
- ▶ DARPA funded
- ▶ v.01 relase imminent



Figure:



- ▶ scikit-learn is an open source machine learning library for the Python programming language.
- ▶ scikit-learn features various classification, regression and clustering algorithms including support vector machines, logistic regression, naive Bayes, random forests, gradient boosting, k-means and DBSCAN.
- ▶ scikit-learn is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

## pytz and babel

pytz and babel provide extended support for time zones and formatting information.



rpy2 provides an interface for calling R 3.0.x in Python, as well as facilities for easily moving data between the two platforms.

# PyTables and h5py

PyTables and h5py both provide access to HDF5 files, a flexible data storage format optimized for numeric data.