**Pandas - The R of Python**

#Pandas is an open source library developed by some smart and generous people who felt the need to introduce R-like functionality in python as numpy wasn't able to handle more complex data that well.

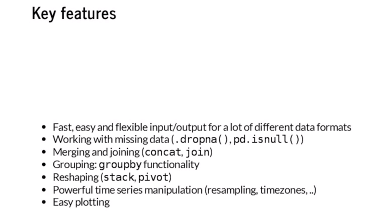
#The main contributor of Pandas is [Wes Mckinney](http://wesmckinney.com/)and there are various tutorials available by him.

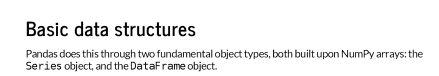
#Numpy is more of a numerical analysis package. The scope of numpy is much wider. It can handle lot of scientific calculations, structural analysis and differential equations.

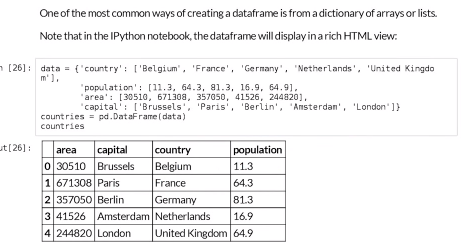
Pandas can be thought of as Numpy arrays with labels for rows and columns and better support for heterogenous data types but its also much more than that

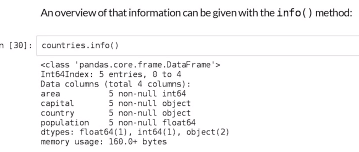
Pandas can be thought of R’s data.frame in python

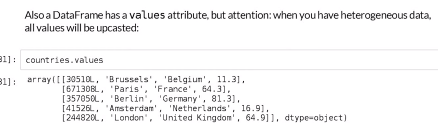
Powerful for working with missing data, working with time series data for reading and writing your data, for reshaping, grouping and merging the data

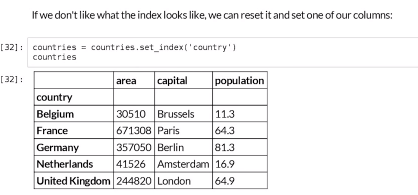


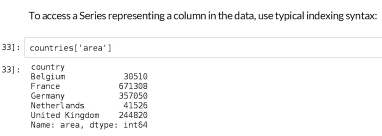


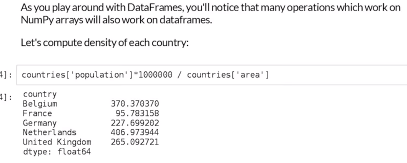


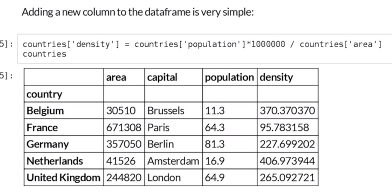


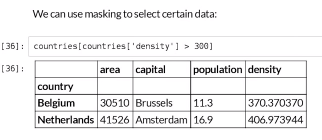


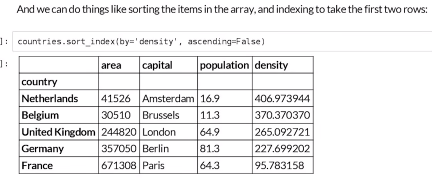


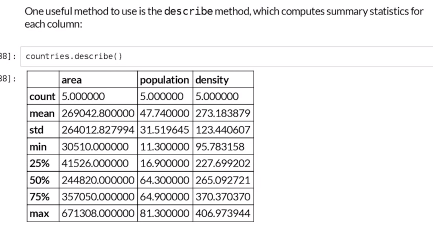


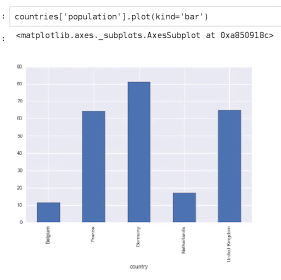


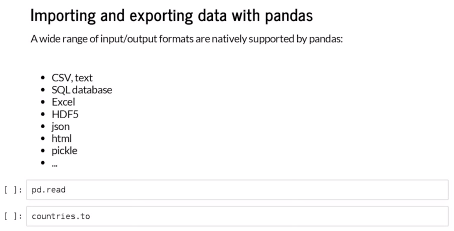












#Basic structure for pandas is series pd.series(), the indexing starts from 0

# describe() gives the statistical summary of the data

#.columns gives the names of the columns in the data

#Given the subset command - market\_data[72:108], what will be the number of rows in this subset? (Hint: Use len command on the code)

The number of rows = end - start (includes start index but excludes end index)

# loc works on labels in the index, iloc works on the position in the index

Pandas offers various merging functions like append, concat and merge. The difference is simple.

* Append directly connects dataframes, similar to picking up an entire dataframe and attaching it at the end of the other. This should be used when the indexing doesn't matter.
* Concat is the function to use when the indexing matters and you have a variety of joins to use like inner, outer, left and right.
* Merge provides a database style of merging where the merging happens on a "key" column.