PANDAS CHEATSHEET

df - A pandas Dataframe object
pd – alias of pandas
s - A pandas Series object

Importing Data	
pd.read_csv(filename)	From a CSV file
pd.read_table(filename)	From a delimited text file (like TSV)
pd.read_excel(filename)	From an Excel file
pd.read_sql(query, connection_object)	Reads from a SQL table/database
pd.read_json(json_string)	Reads from a JSON formatted string, URL or file.
pd.read_html(url)	Parses an html URL, string or file and extracts tables to a list of data frames.
pd.read_clipboard()	Takes the contents of your clipboard and passes it to read_table()
pd.DataFrame(dict)	From a dict, keys for columns names, values for data as lists.

Exporting Data	
df.to_csv(filename)	Write to a CSV file
df.to_excel(filename)	Writes to an Excel file
df.to_sql(table_name, connection_object)	Write to a SQL Table
df.to_json(filename)	Writes to a file in JSON Format.
df.to_html(filename)	Saves as an HTML table.

Selecting Data	
df[col]	Returns column with label col as Series
df[[col1, col2]]	Returns columns as a new dataframe

s.iloc[0]	Selection by position
s.loc[0]	Selection by index
df.iloc[0, :]	First row
df.iloc[0, 0]	First element of first column

Statistics on Data	
df.describe()	Summary statistics for numerical columns
df.mean()	Returns the mean of all columns
df.corr()	Returns the correlation between columns in a Data Frame
df.count()	Returns the number of non-null values in each Data Frame column
df.max()	Returns the highest value in each column
df.min()	Returns the lowest value in each column
df.median()	Returns the median of each column
df.std()	Returns the standard deviation of each column

JOINS on Data	
df1.append(df2)	Adds the rows in df1 to the end of df2
pd.concat([df1, df2],axis=1)	Adds the columns in df1 to the end of df2
df1.join(df2,on=col1,how='inner')	SQL-style joins the columns in df1 with the columns on df2 where the rows for col have identical values. “how” can be one of 'left', 'right', 'outer', 'inner'