

GroupBy

November 17, 2016

```
In [1]: import pandas as pd
```

```
In [2]: drinks = pd.read_csv('http://bit.ly/drinksbycountry')
```

```
In [3]: drinks.head()
```

```
Out[3]:
```

	country	beer_servings	spirit_servings	wine_servings	\
0	Afghanistan	0	0	0	
1	Albania	89	132	54	
2	Algeria	25	0	14	
3	Andorra	245	138	312	
4	Angola	217	57	45	

	total_litres_of_pure_alcohol	continent
0	0.0	Asia
1	4.9	Europe
2	0.7	Africa
3	12.4	Europe
4	5.9	Africa

```
In [4]: drinks.beer_servings.mean()
```

```
Out[4]: 106.16062176165804
```

```
In [5]: drinks.groupby('continent').beer_servings.mean() # shows mean by continent
```

```
Out[5]:
```

continent	
Africa	61.471698
Asia	37.045455
Europe	193.777778
North America	145.434783
Oceania	89.687500
South America	175.083333

Name: beer_servings, dtype: float64

```
In [6]: drinks[drinks.continent == 'Africa']
```

```

Out[6]:
      country  beer_servings  spirit_servings  wine_servings
2      Algeria             25              0              1
4      Angola            217             57              4
18     Benin              34              4              1
22    Botswana           173             35              3
26  Burkina Faso           25              7              1
27    Burundi            88              0              1
28  Cote d'Ivoire          37              1              1
29   Cabo Verde          144             56              1
31   Cameroon           147              1              1
33  Central African Republic  17              2              1
34         Chad           15              1              1
38   Comoros              1              3              1
39    Congo              76              1              1
47   DR Congo            32              3              1
49   Djibouti            15             44              1
53    Egypt              6              4              1
55  Equatorial Guinea       92              0             23
56    Eritrea             18              0              1
58    Ethiopia            20              3              1
62    Gabon             347             98              5
63    Gambia              8              0              1
66    Ghana             31              3              1
70    Guinea              9              0              1
71  Guinea-Bissau         28             31              2
88    Kenya            58             22              1
95   Lesotho            82             29              1
96   Liberia            19            152              1
97    Libya              0              0              1
100  Madagascar          26             15              1
101    Malawi              8             11              1
104    Mali               5              1              1
107  Mauritania           0              0              1
108   Mauritius          98             31              1
114    Morocco           12              6              1
115  Mozambique          47             18              1
117   Namibia          376              3              1
123    Niger              3              2              1
124   Nigeria           42              5              1
142    Rwanda           43              2              1
148  Sao Tome & Principe     56             38             14
150    Senegal              9              1              1
152  Seychelles          157             25              5
153  Sierra Leone         25              3              1
158    Somalia              0              0              1
159  South Africa         225             76              8
162    Sudan              8             13              1
164  Swaziland           90              2              1

```

172	Togo	36	2
175	Tunisia	51	3
179	Uganda	45	9
183	Tanzania	36	6
191	Zambia	32	19
192	Zimbabwe	64	18

	total_litres_of_pure_alcohol	continent
2	0.7	Africa
4	5.9	Africa
18	1.1	Africa
22	5.4	Africa
26	4.3	Africa
27	6.3	Africa
28	4.0	Africa
29	4.0	Africa
31	5.8	Africa
33	1.8	Africa
34	0.4	Africa
38	0.1	Africa
39	1.7	Africa
47	2.3	Africa
49	1.1	Africa
53	0.2	Africa
55	5.8	Africa
56	0.5	Africa
58	0.7	Africa
62	8.9	Africa
63	2.4	Africa
66	1.8	Africa
70	0.2	Africa
71	2.5	Africa
88	1.8	Africa
95	2.8	Africa
96	3.1	Africa
97	0.0	Africa
100	0.8	Africa
101	1.5	Africa
104	0.6	Africa
107	0.0	Africa
108	2.6	Africa
114	0.5	Africa
115	1.3	Africa
117	6.8	Africa
123	0.1	Africa
124	9.1	Africa
142	6.8	Africa
148	4.2	Africa

150	0.3	Africa
152	4.1	Africa
153	6.7	Africa
158	0.0	Africa
159	8.2	Africa
162	1.7	Africa
164	4.7	Africa
172	1.3	Africa
175	1.3	Africa
179	8.3	Africa
183	5.7	Africa
191	2.5	Africa
192	4.7	Africa

```
In [7]: # for each country in Africa..use group by
drinks[drinks.continent == 'Africa'].beer_servings.mean() # mean of drinks
```

```
Out[7]: 61.471698113207545
```

```
In [10]: drinks[drinks.continent == 'Asia'].beer_servings.mean() # mean of drinks for
```

```
Out[10]: 37.045454545454547
```

```
In [11]: drinks[drinks.continent == 'Europe'].beer_servings.mean() # mean of drinks
```

```
Out[11]: 193.77777777777777
```

```
In [12]: drinks.groupby('continent').beer_servings.max() # shows max of all serving
```

```
Out[12]: continent
Africa          376
Asia            247
Europe          361
North America   285
Oceania         306
South America   333
Name: beer_servings, dtype: int64
```

```
In [13]: drinks.groupby('continent').beer_servings.agg(['count', 'min', 'max', 'mean'])
```

```
Out[13]:
```

	count	min	max	mean
continent				
Africa	53	0	376	61.471698
Asia	44	0	247	37.045455
Europe	45	0	361	193.777778
North America	23	1	285	145.434783
Oceania	16	0	306	89.687500
South America	12	93	333	175.083333

```
In [14]: drinks.groupby('continent').mean() # this will find mean across all columns
```

```
Out[14]:
```

	beer_servings	spirit_servings	wine_servings	\
continent				
Africa	61.471698	16.339623	16.264151	
Asia	37.045455	60.840909	9.068182	
Europe	193.777778	132.555556	142.222222	
North America	145.434783	165.739130	24.521739	
Oceania	89.687500	58.437500	35.625000	
South America	175.083333	114.750000	62.416667	

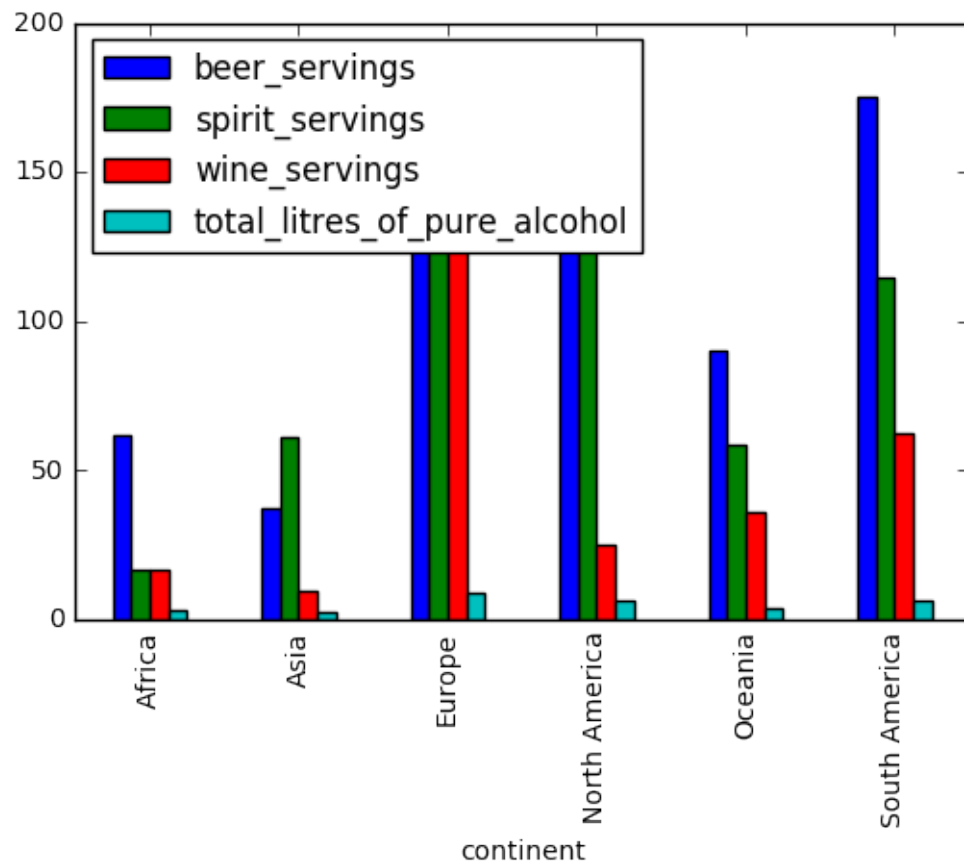
	total_litres_of_pure_alcohol
continent	
Africa	3.007547
Asia	2.170455
Europe	8.617778
North America	5.995652
Oceania	3.381250
South America	6.308333

```
In [15]: %matplotlib inline
```

```
/opt/conda/lib/python3.5/site-packages/matplotlib/font_manager.py:273: UserWarning:
  warnings.warn('Matplotlib is building the font cache using fc-list. This may take
/opt/conda/lib/python3.5/site-packages/matplotlib/font_manager.py:273: UserWarning:
  warnings.warn('Matplotlib is building the font cache using fc-list. This may take
```

```
In [16]: drinks.groupby('continent').mean().plot(kind='bar')
```

```
Out[16]: <matplotlib.axes._subplots.AxesSubplot at 0x7f55e92f2898>
```



In []: