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

In [2]:
```
df = pd.read_csv('world-happiness-report.csv')
```

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

Out[3]:

	Country	year	Life Ladder	Log GDP	Social support	life expectancy	life choices	Generosity	Perceptions of corruption	Positive affect	Negative affect
0	Afghanistan	2008	3.724	7.370	0.451	50.8	0.718	0.168	0.882	0.518	0.258
1	Afghanistan	2009	4.402	7.540	0.552	51.2	0.679	0.190	0.850	0.584	0.237
2	Afghanistan	2010	4.758	7.647	0.539	51.6	0.600	0.121	0.707	0.618	0.275

In [4]:
```
df.iloc[2]
```

Out[4]:
```
Country                          Afghanistan
     year                               2010
Life Ladder                            4.758
Log GDP                                7.647
Social support                         0.539
life expectancy                         51.6
life choices                             0.6
Generosity                             0.121
Perceptions of corruption              0.707
Positive affect                        0.618
Negative affect                        0.275
Name: 2, dtype: object
```

In [5]:
```
df.loc[2]
```

Out[5]:
```
Country                          Afghanistan
     year                               2010
Life Ladder                            4.758
Log GDP                                7.647
Social support                         0.539
life expectancy                         51.6
life choices                             0.6
Generosity                             0.121
Perceptions of corruption              0.707
Positive affect                        0.618
Negative affect                        0.275
Name: 2, dtype: object
```

In [6]:
```
df[:2]
```

Out[6]:

	Country	year	Life Ladder	Log GDP	Social support	life expectancy	life choices	Generosity	Perceptions of corruption	Positive affect	Negative affect
0	Afghanistan	2008	3.724	7.37	0.451	50.8	0.718	0.168	0.882	0.518	0.258
1	Afghanistan	2009	4.402	7.54	0.552	51.2	0.679	0.190	0.850	0.584	0.237

In [7]:
```
df[3:6]
```

Out[7]:

	Country	year	Life Ladder	Log GDP	Social support	life expectancy	life choices	Generosity	Perceptions of corruption	Positive affect	Negative affect
3	Afghanistan	2011	3.832	7.620	0.521	51.92	0.496	0.162	0.731	0.611	0.267
4	Afghanistan	2012	3.783	7.705	0.521	52.24	0.531	0.236	0.776	0.710	0.268
5	Afghanistan	2013	3.572	7.725	0.484	52.56	0.578	0.061	0.823	0.621	0.273

In [10]:
```
to_drop=['Generosity','Negative affect']
df.drop(to_drop, inplace=True, axis=1)
```

In [11]:
```
df
```

Out[11]:

	Country	year	Life Ladder	Log GDP	Social support	life expectancy	life choices	Perceptions of corruption	Positive affect
0	Afghanistan	2008	3.724	7.370	0.451	50.80	0.718	0.882	0.518
1	Afghanistan	2009	4.402	7.540	0.552	51.20	0.679	0.850	0.584
2	Afghanistan	2010	4.758	7.647	0.539	51.60	0.600	0.707	0.618
3	Afghanistan	2011	3.832	7.620	0.521	51.92	0.496	0.731	0.611
4	Afghanistan	2012	3.783	7.705	0.521	52.24	0.531	0.776	0.710
...	...	...	...	...	...	...	...	...	...
94	Azerbaijan	2007	4.568	9.386	0.753	62.26	0.522	0.871	0.521
95	Azerbaijan	2008	4.817	9.465	0.684	62.64	0.601	0.715	0.578
96	Azerbaijan	2009	4.574	9.534	0.736	63.02	0.498	0.754	0.544
97	Azerbaijan	2010	4.219	9.569	0.687	63.40	0.501	0.858	0.527
98	Azerbaijan	2011	4.680	9.540	0.725	63.64	0.537	0.795	0.536

99 rows × 9 columns

In [12]:
```
df.aggregate(['sum', 'min'])
```

Out[12]:

	Country	year	Life Ladder	Log GDP	Social support	life expectancy	life choices	Perceptions of corruption	Positive affect
sum	AfghanistanAfghanistanAfghanistanAfghanistanAf...	199295	550.714	953.461	77.402	6512.6	67.306	71.514	66.674
min	Afghanistan	2005	2.375	7.370	0.420	50.8	0.374	0.357	0.351

In [13]:
```
df.aggregate({"Social support":['sum', 'min'],
              "life expectancy":['max', 'min'],
              "life choices":['min', 'sum'],
              "Positive affect":['sum']})
```

Out[13]:

	Social support	life expectancy	life choices	Positive affect
sum	77.402	NaN	67.306	66.674
min	0.420	50.8	0.374	NaN
max	NaN	74.2	NaN	NaN

In [15]:
```
left = pd.DataFrame({
   'id':[1,2,3,4,5],
   'Name': ['Alex', 'Amy', 'Allen', 'Alice', 'Ayoung'],
   'subject_id':['sub1','sub2','sub4','sub6','sub5']})
right = pd.DataFrame({
   'id':[1,2,3,4,5],
   'Name': ['Billy', 'Brian', 'Bran', 'Bryce', 'Betty'],
   'subject_id':['sub2','sub4','sub3','sub6','sub5']})
print(pd.merge(left, right, on='subject_id', how='left'))
```

```
   id_x  Name_x subject_id  id_y Name_y
0     1    Alex       sub1   NaN    NaN
1     2     Amy       sub2   1.0  Billy
2     3   Allen       sub4   2.0  Brian
3     4   Alice       sub6   4.0  Bryce
4     5  Ayoung       sub5   5.0  Betty
```

In [16]:
```
print(pd.merge(left, right, on='subject_id', how='inner'))
```

```
   id_x  Name_x subject_id  id_y Name_y
0     2     Amy       sub2     1  Billy
1     3   Allen       sub4     2  Brian
2     4   Alice       sub6     4  Bryce
3     5  Ayoung       sub5     5  Betty
```

In [ ]:
```

```