```
Chapter 5 - Dimensionality Reduction Methods

Part 2 - Principal component analysis (PCA)
```

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

import matplotlib.pyplot as plt
import pylab as plt
import seaborn as sb
from IPython.display import Image
from IPython.core.display import HTML
from pylab import rcParams

import sklearn
from sklearn import datasets
```

In [3]:
```
from sklearn import decomposition
from sklearn.decomposition import PCA
```

In [2]:
```
%matplotlib inline
rcParams['figure.figsize'] = 5, 4
sb.set_style('whitegrid')
```

```
PCA ont the iris dataset
```

In [8]:
```
iris = datasets.load_iris()
x = iris.data
variable_names = iris.feature_names

x[0:10,]
```

Out[8]:
```
array([[5.1, 3.5, 1.4, 0.2],
       [4.9, 3. , 1.4, 0.2],
       [4.7, 3.2, 1.3, 0.2],
       [4.6, 3.1, 1.5, 0.2],
       [5. , 3.6, 1.4, 0.2],
       [5.4, 3.9, 1.7, 0.4],
       [4.6, 3.4, 1.4, 0.3],
       [5. , 3.4, 1.5, 0.2],
       [4.4, 2.9, 1.4, 0.2],
       [4.9, 3.1, 1.5, 0.1]])
```

In [9]:
```
pca = decomposition.PCA()
iris_pca = pca.fit_transform(x)

pca.explained_variance_ratio_
```

Out[9]:
```
array([0.92461872, 0.05306648, 0.01710261, 0.00521218])
```

In [10]:
```
pca.explained_variance_ratio_.sum()
```

Out[10]:
```
1.0
```

In [11]:
```
comps = pd.DataFrame(pca.components_, columns=variable_names)
comps
```

Out[11]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	0.361387	-0.084523	0.856671	0.358289
1	0.656589	0.730161	-0.173373	-0.075481
2	-0.582030	0.597911	0.076236	0.545831
3	-0.315487	0.319723	0.479839	-0.753657

In [12]:
```
sb.heatmap(comps, cmap='Blues', annot=True)
```

Out[12]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1cebed4c5c8>
```