Import some libraries. Right now, we're about to using PCA from decomposition from sklearn

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

import matplotlib.pyplot as plt
from pylab import rcParams

import sklearn
from sklearn import datasets
from sklearn import decomposition
from sklearn.decomposition import PCA
```

Configuring the canvas

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

Read Iris Datasets

In [3]:
```
iris = datasets.load_iris()
x = iris.data
variable_names = iris.feature_names
```

We're about to calculate the PCA. First thing first, we use decomposition.PCA() function, then fit_transform(x) and we use pca.explained_variance_ratio_ to explained first few component variance

In [4]:
```
pca = decomposition.PCA()
iris_pca = pca.fit_transform(x)
pca.explained_variance_ratio_
```

Out[4]:
```
array([0.92461872, 0.05306648, 0.01710261, 0.00521218])
```

And then, we calculate the sum(). Remember, make sure to retain at least 70% of dataset's

original information

In [5]:
```
pca.explained_variance_ratio_.sum()
```

Out[5]:
```
1.0
```

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

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 [8]:
```
sb.heatmap(comp)
```

Out[8]:
```
<matplotlib.axes._subplots.AxesSubplot at 0xc03c490>
```