In [1]:

```
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:

```
data = pd.read_csv("../data/CoffeeAndCodeLT2018.csv")
data
```

Out[2]:

	CodingHours	CoffeeCupsPerDay	CoffeeTime	CodingWithoutCoffee	CoffeeType	CoffeeSc
0	8	2	Before coding	Yes	Caffè latte	Sc
1	3	2	Before coding	Yes	Americano	
2	5	3	While coding	No	Nescafe	
3	8	2	Before coding	No	Nescafe	
4	10	3	While coding	Sometimes	Turkish	
...	...	...	...	...	...	
95	6	2	Before coding	Yes	Nescafe	
96	4	1	Before coding	Sometimes	Nescafe	Sc
97	10	3	Before coding	Yes	Cappuccino	
98	2	2	While coding	Sometimes	Espresso (Short Black)	Sc
99	10	4	Before coding	Sometimes	Double Espresso (Doppio)	Sc

100 rows × 9 columns

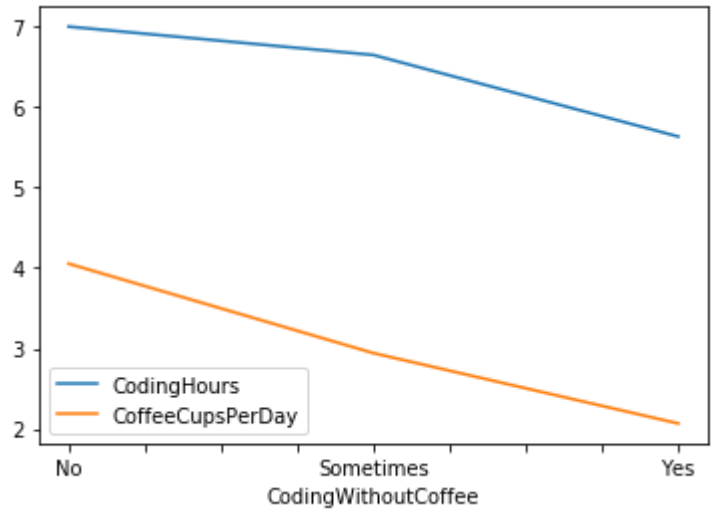
Время кодинга напрямую зависит от количества выпитого кофе

In [3]:

```
data.groupby(['CodingWithoutCoffee']).mean().plot()
```

Out[3]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1cc67e98>
```



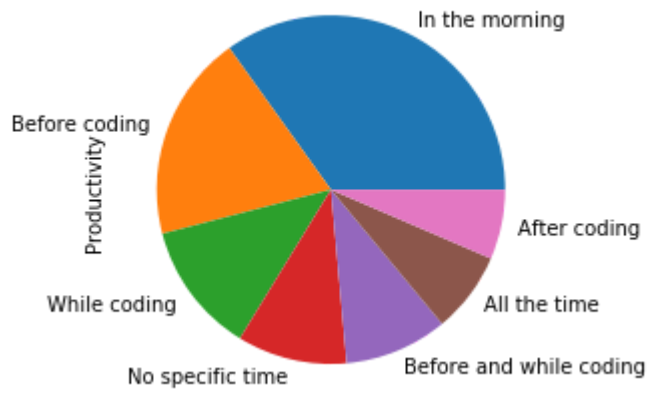
Эффективность кофе в зависимости от времени его употребления

In [4]:

```
data_with_kpd = data.groupby(['CoffeeTime']).mean()
data_with_kpd['Productivity'] = data_with_kpd['CodingHours']/data_with_kpd['CoffeeCupsPerDay']
data_with_kpd['Efficiency'] = data_with_kpd['Productivity']/data_with_kpd['Productivity'].max()
data_with_kpd = data_with_kpd.sort_values('Efficiency', ascending=False)
data_with_kpd['Productivity'].plot.pie()
```

Out[4]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1dd21a48>
```



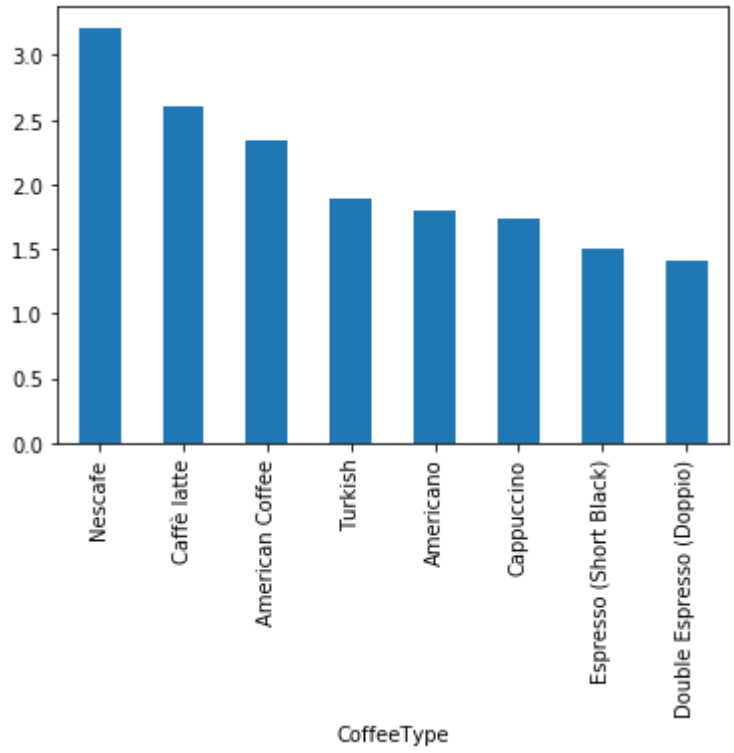
Эффективность кофе в зависимости от его типа

In [5]:

```
coffee_type = data.groupby(['CoffeeType']).mean()
coffee_type['Productivity'] = coffee_type['CodingHours']/coffee_type['CoffeeCupsPerDay']
]
coffee_type['Efficiency'] = coffee_type['Productivity']/coffee_type['Productivity'].max
()
coffee_type = coffee_type.sort_values('Efficiency', ascending=False)
coffee_type['Productivity'].plot.bar()
```

Out[5]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1dd794d8>
```



In [6]:

```
gender = data.groupby(['Gender']).mean()
gender['Productivity'] = gender['CodingHours']/gender['CoffeeCupsPerDay']
gender
```

Out[6]:

Gender	CodingHours	CoffeeCupsPerDay	Productivity
Female	5.384615	2.269231	2.372881
Male	6.770270	3.108108	2.178261