Ξεκινάμε το Jupyter γράφοντας στο terminal

/opt/pyenv-3.7.5/bin/jupyter notebook

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

Part 1 - Τυχαίο δείγμα 300 επιβατών του Τιτανικού

In [2]:
```
dataset = pd.read_csv("./titanic_300.csv")
```

Με την head() λαμβάνουμε τα 5 πρώτα στοιχεία (preview)

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

Out[3]:

	Survived	Pclass	Name	Sex	Age	Fare
0	0	3	Braund, Mr. Owen Harris	male	22.0	7.2500
1	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	71.2833
2	1	3	Heikkinen, Miss. Laina	female	26.0	7.9250
3	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	53.1000
4	0	3	Allen, Mr. William Henry	male	35.0	8.0500

Με την tail() λαμβάνουμε τα 5 τελευταία στοιχεία (preview)

In [4]:
```
dataset.tail()
```

Out[4]:

	Survived	Pclass	Name	Sex	Age	Fare
295	1	1	Harder, Mr. George Achilles	male	25.0	55.4417
296	0	3	Wiklund, Mr. Jakob Alfred	male	18.0	6.4958
297	0	3	Beavan, Mr. William Thomas	male	19.0	8.0500
298	0	1	Ringhini, Mr. Sante	male	22.0	135.6333
299	0	3	Palsson, Miss. Stina Viola	female	3.0	21.0750

Εμφάνιση τυχαίου στοιχείου του δείγματος

In [5]:
```
dataset.iloc[np.random.randint(300)]
```

Out[5]:
```
Survived                          0
Pclass                            3
Name        Green, Mr. George Henry
Sex                            male
Age                              51
Fare                           8.05
Name: 179, dtype: object
```

Εφαρμογή του κανόνα του Sturges

$$K = 1 + 3.322 * \log(N)$$

In [6]:
```
K = int(1+3.322*np.log10(300)) + 1
```

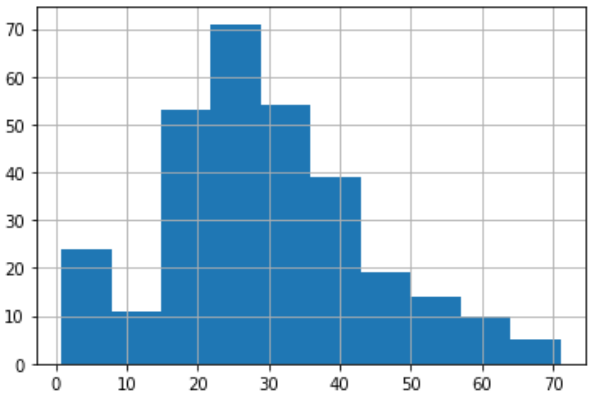
In [7]:
```
K
```

Out[7]:
```
10
```

Ιστόγραμμα των ηλικίων θεωρώντας K ηλικιακές κλάσεις (από τον κανόνα του Sturges)

In [8]:
```
dataset["Age"].hist(bins=10)
```

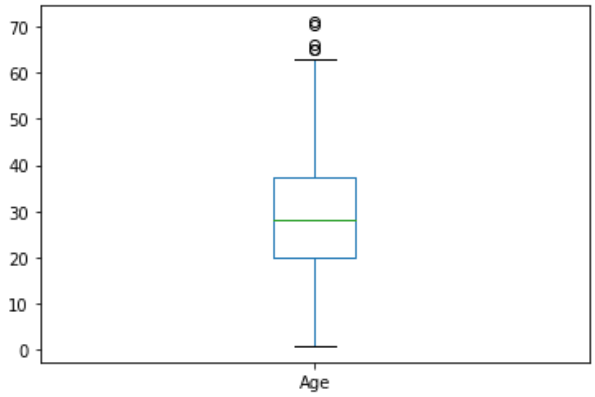
Out[8]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f89a13f32b0>
```



Box-and-Whisker plot των ηλικιών

In [9]:
```
dataset["Age"].plot.box()
```

Out[9]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f28ab70>
```



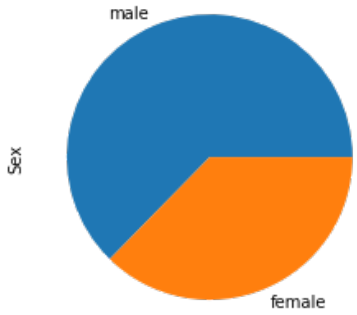
Απαρίθμηση των εμφανίσεων κάθε φύλου στο δείγμα

In [10]:
```
dataset["Sex"].value_counts()
```

Out[10]:
```
male      188
female    112
Name: Sex, dtype: int64
```

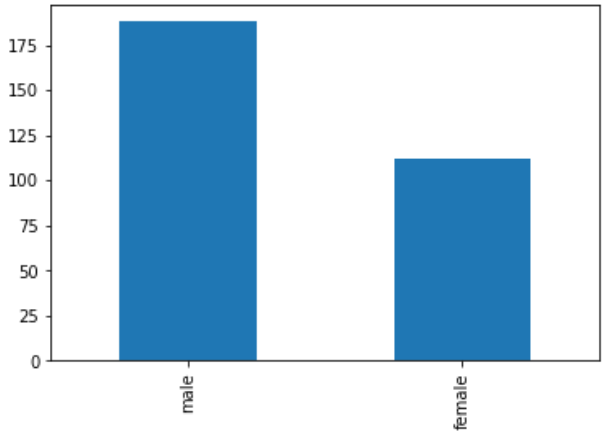
In [11]:
```
dataset["Sex"].value_counts().plot.pie()
```

Out[11]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f212c18>
```



In [12]:
```
dataset["Sex"].value_counts().plot.bar()
```

Out[12]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f1d7e10>
```



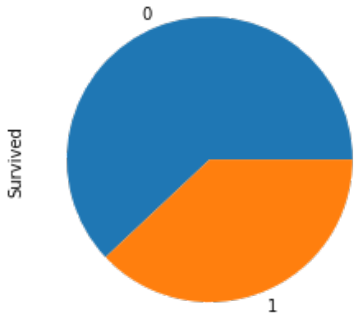
Απαρίθμηση των επιζόντων και των θυμάτων (0-θύμα, 1-επιζών) στο δείγμα

In [13]:
```
dataset["Survived"].value_counts()
```

Out[13]:
```
0    186
1    114
Name: Survived, dtype: int64
```

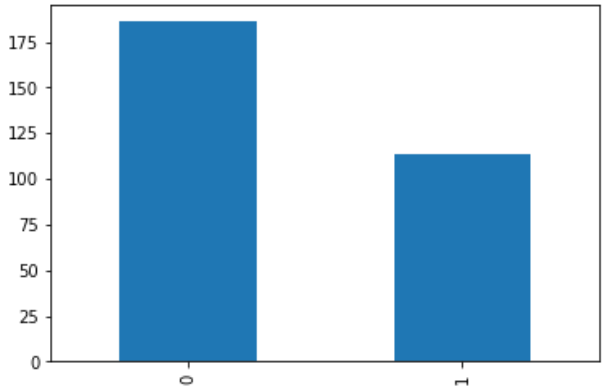
In [14]:
```
dataset["Survived"].value_counts().plot.pie()
```

Out[14]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f148c50>
```



In [15]:
```
dataset["Survived"].value_counts().plot.bar()
```

Out[15]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f100ef0>
```



Part 2 - Πείραμα μέτρησης ταχύτητας του φωτός (3η Διάλεξη)

In [16]:
```
L = [28, 26, 33, 24, 34,-44, 27, 16, 40, -2, 29,
     22, 24, 21, 25, 30, 23, 29, 31, 19, 24, 20,
     36, 32, 36, 28, 25, 21, 28, 29, 37, 25, 28,
     26, 30, 32, 36, 26, 30, 22, 36, 23, 27, 27,
     28, 27, 31, 27, 26, 33, 26, 32, 32, 24, 39,
     28, 24, 25, 32, 25, 29, 27, 28, 29, 16, 23]
```

In [17]:
```
x = pd.DataFrame(L)
```

Υπολογισμός του πρώτου τεταρτημόριου Q1

In [18]:
```
Q1 = x.quantile(0.25)
```

In [19]:
```
Q1
```

Out[19]:
```
0    24.0
Name: 0.25, dtype: float64
```

Άσκηση 1 : Υπολογίστε Διάμεσο M, το Q3 και το IQR

In [20]:
```
M = x.median()
Q3 = x.quantile(0.75)
```

In [21]:
```
M
```

Out[21]:
```
0    27.0
dtype: float64
```

In [22]:
```
Q3
```

Out[22]:
```
0    30.75
Name: 0.75, dtype: float64
```

In [23]:
```
IQR = Q3-Q1
```

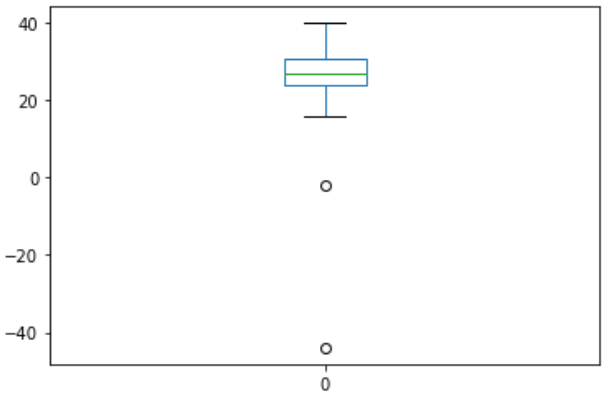
In [24]:
```
IQR
```

Out[24]:
```
0    6.75
dtype: float64
```

Άσκηση 2: Δημιουργήστε το Box-and-Whisker plot

In [25]:
```
x.plot.box()
```

Out[25]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f899f08a2b0>
```



In [ ]:
```

```