In [15]:

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

In [7]:

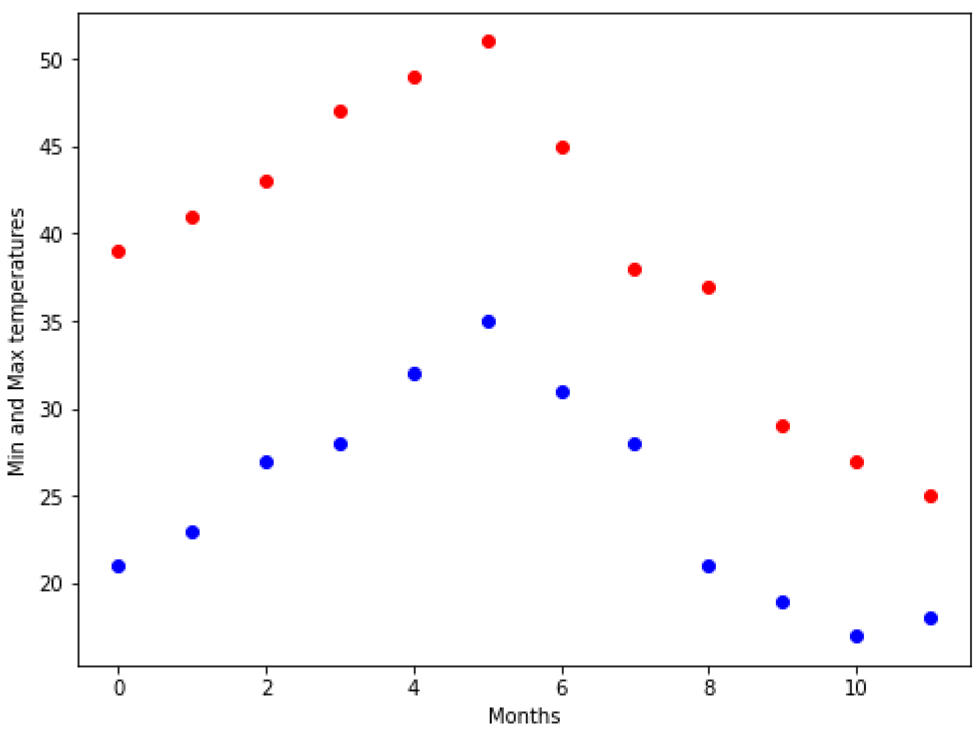
```
Max = np.array([39, 41, 43, 47, 49, 51, 45, 38, 37, 29, 27, 25])
Min = np.array([21, 23, 27, 28, 32, 35, 31, 28, 21, 19, 17, 18])
```

In [9]:

```
months = np.arange(12)
plt.figure(figsize = (8, 6))
plt.plot(months, Max, 'ro')
plt.plot(months, Min, 'bo')
plt.xlabel('Months')
plt.ylabel('Min and Max temperatures')
```

Out[9]:

```
Text(0, 0.5, 'Min and Max temperatures')
```



In [11]:

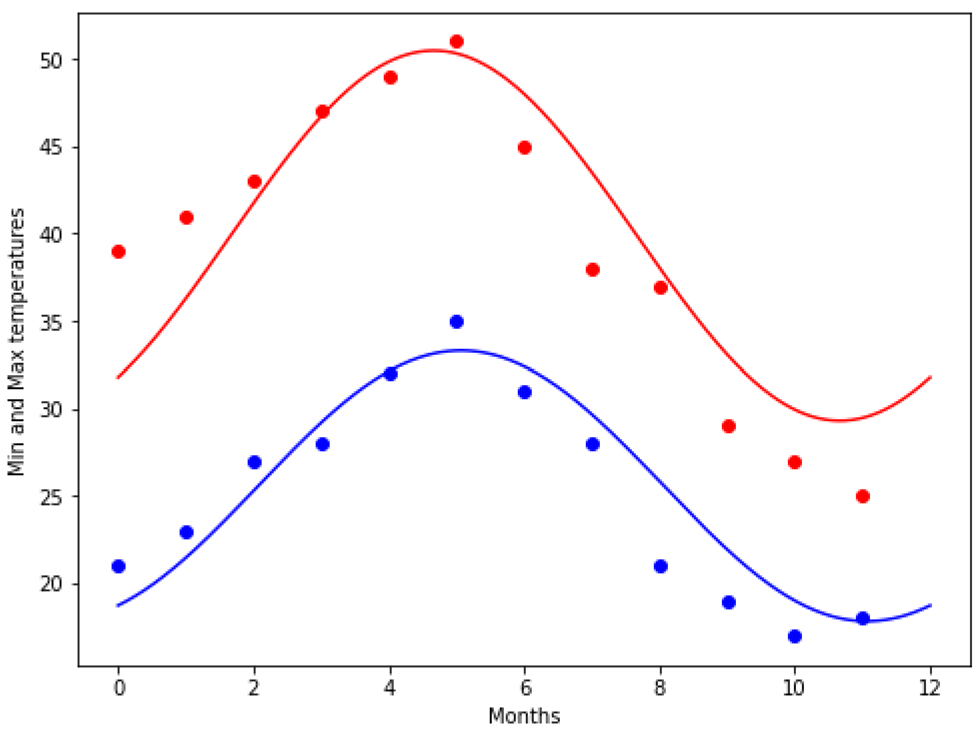
```
from scipy import optimize
def yearly_temps(times, avg, ampl, time_offset):
    return(avg + ampl * np.cos((times + time_offset) * 2 * np.pi/times.max()))
res_max, cov_max = optimize.curve_fit(yearly_temps, months, Max, [20, 10, 0])
res_min, cov_min = optimize.curve_fit(yearly_temps, months, Min, [-40, 20, 0])
```

In [12]:

```
days = np.linspace(0, 12, 365)
plt.figure(figsize = (8, 6))
plt.plot(months, Max, 'ro')
plt.plot(days, yearly_temps(days, *res_max), 'r-')
plt.plot(months, Min, 'bo')
plt.plot(days, yearly_temps(days, *res_min), 'b-')
plt.xlabel('Months')
plt.ylabel('Min and Max temperatures')
```

Out[12]:

```
Text(0, 0.5, 'Min and Max temperatures')
```



In [18]:

```
titanic = pd.read_csv('https://raw.githubusercontent.com/Geoyi/Cleaning-Titanic-Data/master/titanic_original.csv')
titanic.head()
```

Out[18]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.dest
0	1.0	1.0	Allen, Miss. Elisabeth Walton	female	29.0000	0.0	0.0	24160	211.3375	B5	S	2	NaN	St Louis, MO
1	1.0	1.0	Allison, Master. Hudson Trevor	male	0.9167	1.0	2.0	113781	151.5500	C22 C26	S	11	NaN	Montreal, PQ / Chesterville, ON
2	1.0	0.0	Allison, Miss. Helen Loraine	female	2.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON
3	1.0	0.0	Allison, Mr. Hudson Joshua Creighton	male	30.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	135.0	Montreal, PQ / Chesterville, ON
4	1.0	0.0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON

In [23]:

```
gender = titanic['sex'].astype('category')
gender = gender.cat.codes
gender.head()
```

Out[23]:

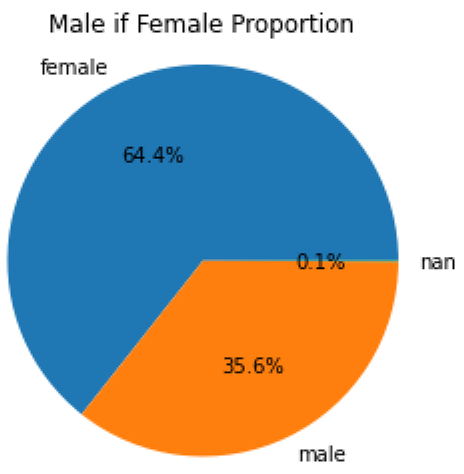
```
0    0
1    1
2    0
3    1
4    0
dtype: int8
```

In [24]:

```
plt.pie(gender.value_counts(), labels = titanic['sex'].unique(), autopct = '%1.1f%%')
plt.title('Male if Female Proportion')
plt.axis('Equal')
```

Out[24]:

```
(-1.1131210684116186,
 1.1006248129493268,
 -1.1156745899219302,
 1.1088629231702263)
```

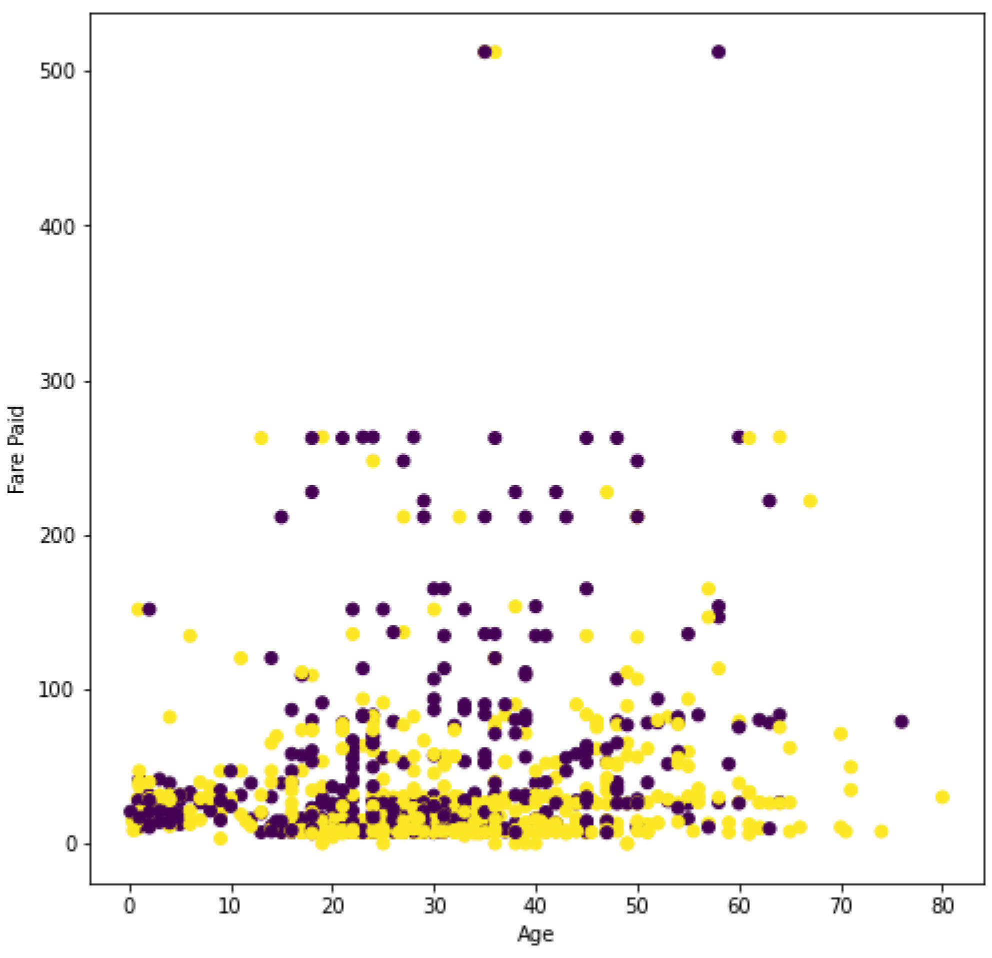


In [27]:

```
plt.figure(figsize=(8,8))
plt.scatter(titanic['age'], titanic['fare'], c = gender)
plt.xlabel('Age')
plt.ylabel('Fare Paid')
```

Out[27]:

```
Text(0, 0.5, 'Fare Paid')
```



In [ ]: