In [4]:
```
from pandas import DataFrame
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
```

In [8]:
```
# Creating a random dataset of 30 elements with
# x and y variables using random function
# between 30 to 80 integers for x and 60 to 100 integers for y
import random
import numpy
x_col=random.sample(range(30, 80), 30)
print("X:",x_col)
y_col=random.sample(range(60, 100), 30)
print("Y:",y_col)
values={'X':x_col,'Y':y_col}
dt= DataFrame(values)
dt.head()
```

```
X: [40, 50, 35, 52, 57, 51, 44, 56, 72, 47, 34, 33, 78, 71, 62, 43, 69, 45, 54, 77, 46, 48, 39, 49, 32, 38, 60, 64, 53, 66]
Y: [83, 60, 96, 98, 65, 81, 67, 69, 66, 90, 64, 74, 78, 77, 93, 99, 70, 88, 79, 75, 62, 97, 91, 86, 95, 63, 68, 73, 82, 94]
```

Out[8]:

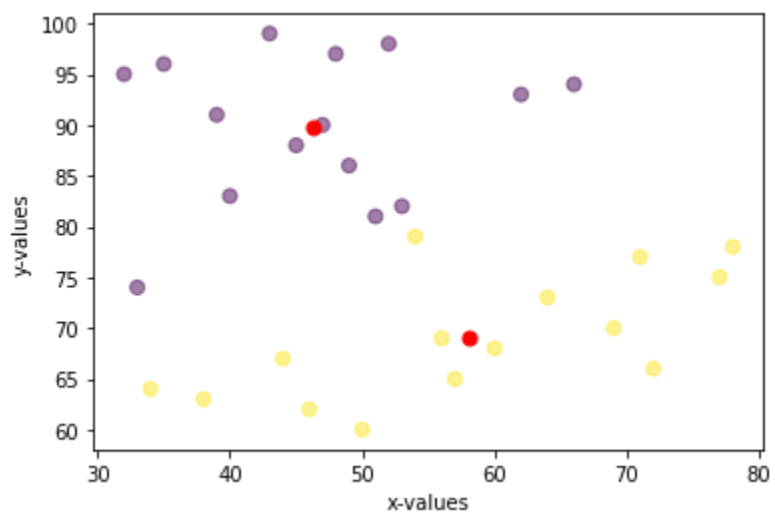
	X	Y
0	40	83
1	50	60
2	35	96
3	52	98
4	57	65

In [9]:
```
#Dividing the dataset into two clusters
kmeans = KMeans(n_clusters=2).fit(dt)
centroids = kmeans.cluster_centers_
print(centroids)

plt.scatter(dt['X'], dt['Y'], c= kmeans.labels_.astype(float), s=50, alpha=0.5)
plt.scatter(centroids[:, 0], centroids[:, 1], c='red', s=50)
plt.xlabel("x-values")
plt.ylabel("y-values")
plt.show()

# the center of each cluster (in red)
#represents the mean of all the observations that belong to that cluster.
```

```
[[46.33333333 89.8       ]
 [58.         69.06666667]]
```



In [11]:
```
#visualising the clusters using tkinter GUI
import tkinter as tk
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
root= tk.Tk()
canvas1 = tk.Canvas(root, width = 100, height = 100)
canvas1.pack()
label1 = tk.Label(root, text=centroids, justify = 'center')
canvas1.create_window(10, 90, window=label1)
figure1 = plt.Figure(figsize=(5,4), dpi=100)
ax1 = figure1.add_subplot(111)
ax1.scatter(dt['X'], dt['Y'], c= kmeans.labels_.astype(float), s=50, alpha=0.5)
ax1.scatter(centroids[:, 0], centroids[:, 1], c='red', s=50)
scatter1 = FigureCanvasTkAgg(figure1, root)
scatter1.get_tk_widget().pack(side=tk.LEFT, fill=tk.BOTH)

root.mainloop()
```

In [ ]:

In [ ]: