In [1]:
```
import pandas as pd
df_allez_spelling = pd.read_csv("allez_spelling.csv", sep=";")
df_allez_spelling
```

Out[1]:

	Id	Name	Gender	PreSpelling	PostSpelling	Gain
0	1	P1	M	10	80	70
1	2	P2	M	60	90	30
2	3	P3	M	0	50	50
3	4	P4	F	30	60	30
4	5	P5	M	10	100	90
5	6	P6	NaN	20	70	50
6	7	P7	M	10	40	30
7	8	P8	M	10	60	50
8	9	P9	F	40	80	40
9	10	P10	NaN	20	50	30

In [2]:
```
df_allez_spelling["PreSpelling"].mean()
```

Out[2]:
```
21.0
```

In [3]:
```
import numpy as np
list_preSpelling = []
list_postSpelling = []
list_gain = []
np_list_preSpelling = np.arange(10)
for index, row in df_allez_spelling.iterrows():
    print(row['PreSpelling'])
    list_preSpelling.append(row['PreSpelling'])
    list_postSpelling.append(row['PostSpelling'])
    list_gain.append(row['Gain'])
np_list_preSpelling = list_preSpelling
```

```
10
60
0
30
10
20
10
10
40
20
```

In [4]:
```
list_preSpelling
```

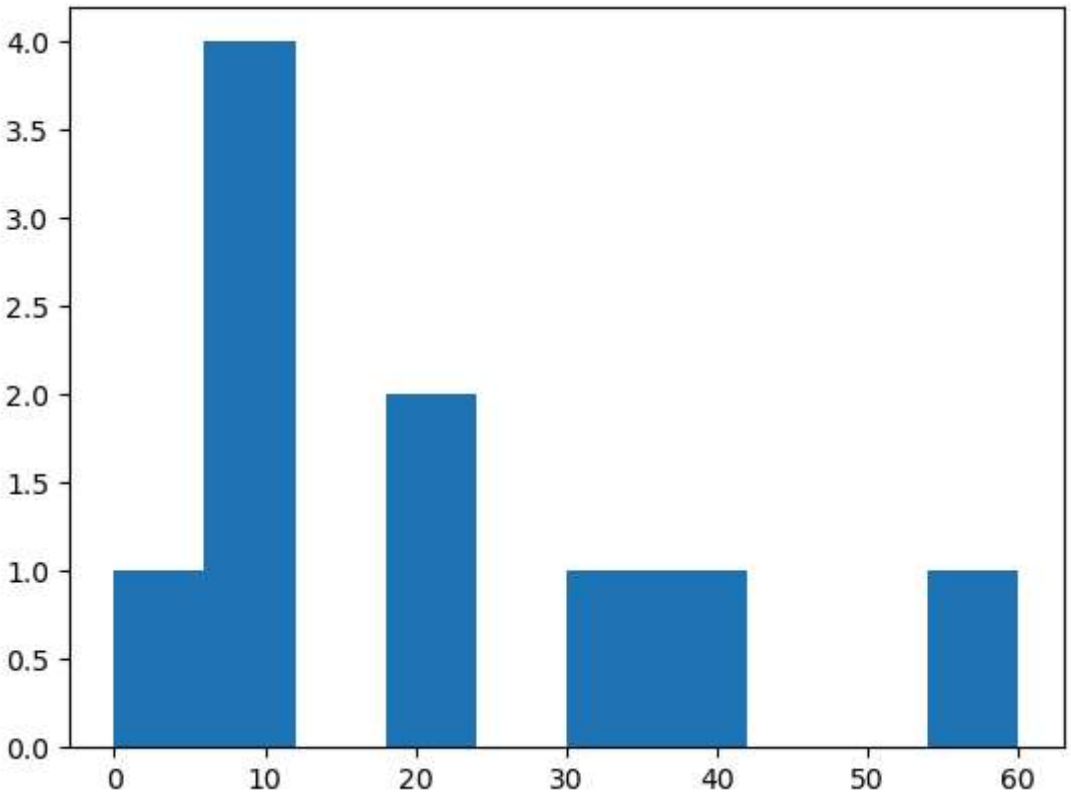
Out[4]:
```
[10, 60, 0, 30, 10, 20, 10, 10, 40, 20]
```

In [5]:
```
np_list_preSpelling
```

Out[5]:
```
[10, 60, 0, 30, 10, 20, 10, 10, 40, 20]
```

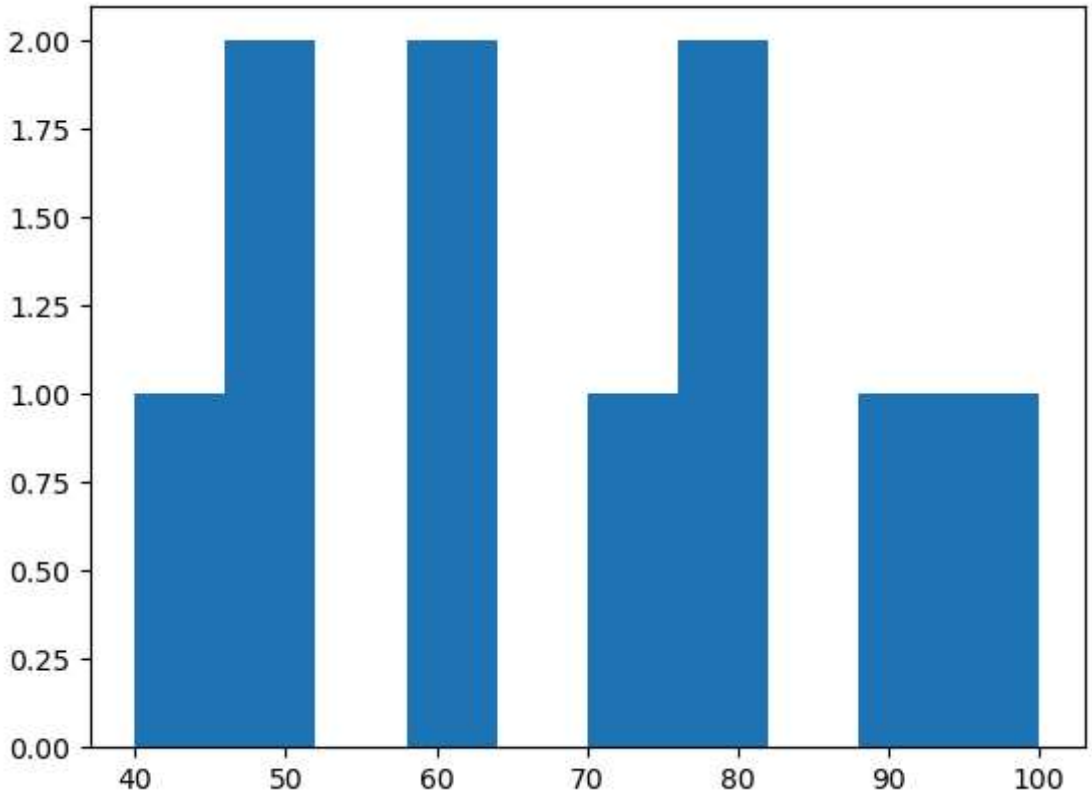
In [6]:
```
import matplotlib.pyplot as plt
plt.hist(list_preSpelling)
```

Out[6]:
```
(array([1., 4., 0., 2., 0., 1., 1., 0., 0., 1.]),
 array([ 0.,  6., 12., 18., 24., 30., 36., 42., 48., 54., 60.]),
 <BarContainer object of 10 artists>)
```



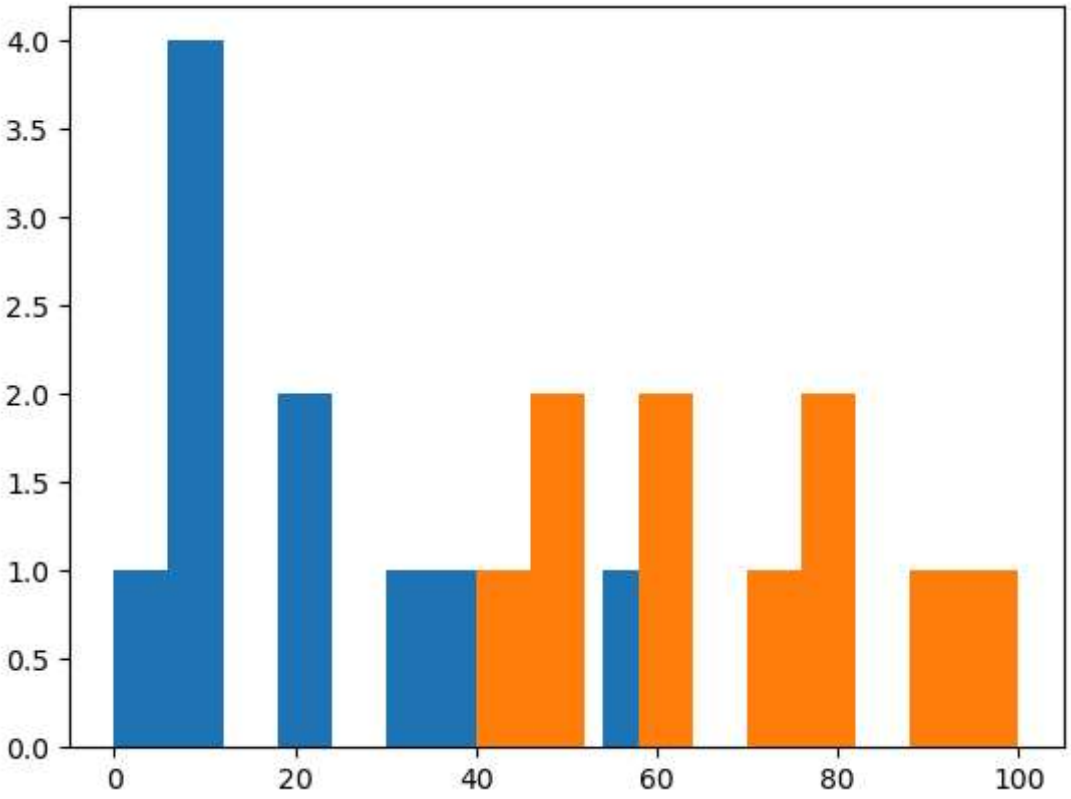
In [7]:
```
plt.hist(list_postSpelling)
```

Out[7]:
```
(array([1., 2., 0., 2., 0., 1., 2., 0., 1., 1.]),
 array([ 40.,  46.,  52.,  58.,  64.,  70.,  76.,  82.,  88.,  94., 100.]),
 <BarContainer object of 10 artists>)
```



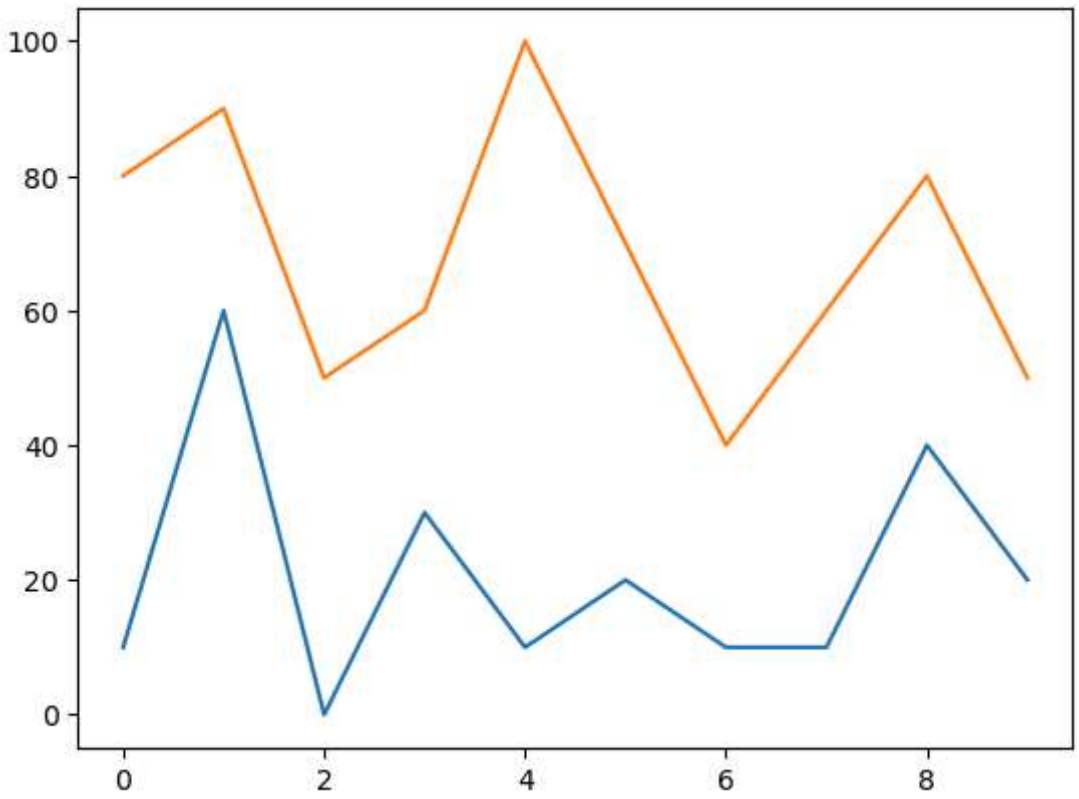
In [8]:
```
import matplotlib.pyplot as plt
plt.hist(list_preSpelling)
plt.hist(list_postSpelling)
```

Out[8]:
```
(array([1., 2., 0., 2., 0., 1., 2., 0., 1., 1.]),
 array([ 40.,  46.,  52.,  58.,  64.,  70.,  76.,  82.,  88.,  94., 100.]),
 <BarContainer object of 10 artists>)
```

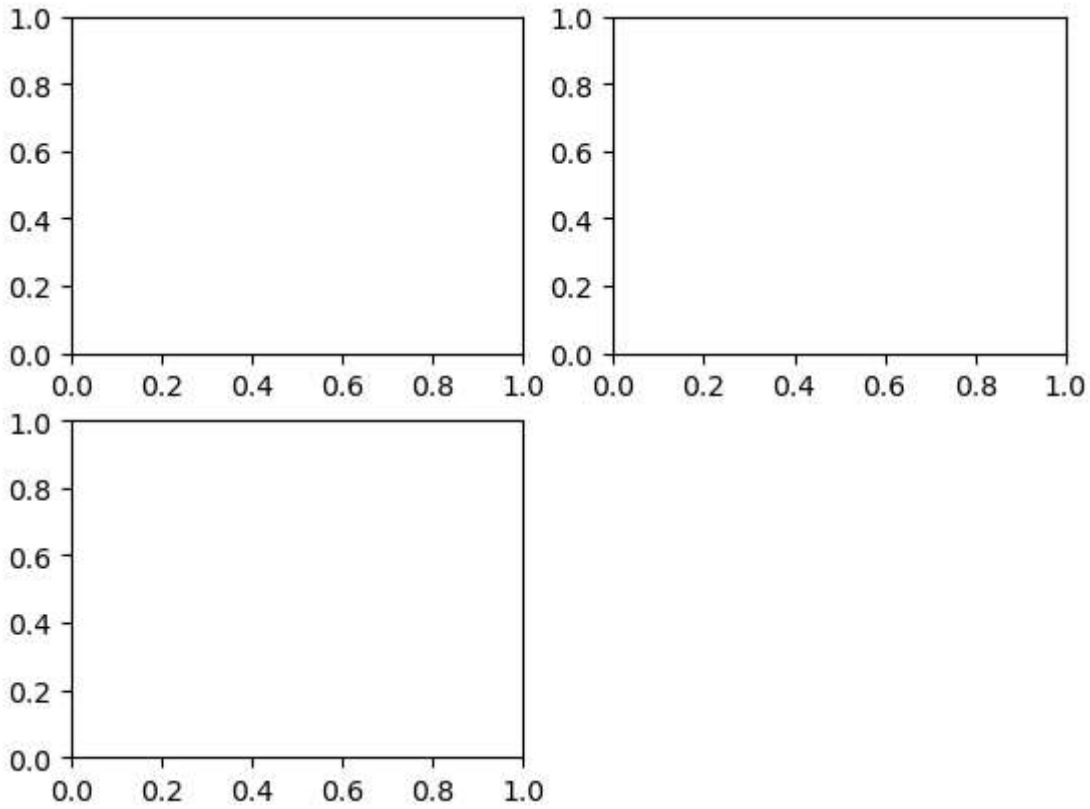


In [9]:
```
plt.plot(np_list_preSpelling)
plt.plot(list_postSpelling)
```

Out[9]: [<matplotlib.lines.Line2D at 0x2b52c867e10>]

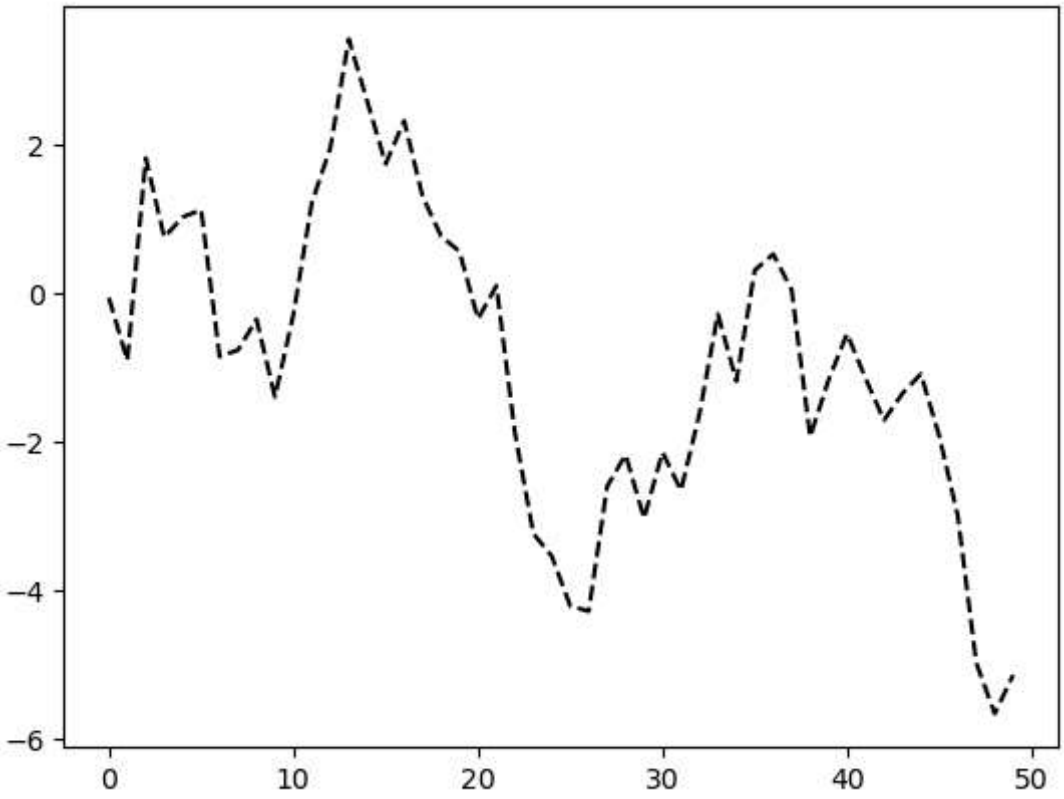


In [10]:
```
fig = plt.figure()
ax1 = fig.add_subplot(2,2,1)
ax2 = fig.add_subplot(2,2,2)
ax3 = fig.add_subplot(2,2,3)
```



In [11]:
```
plt.plot(np.random.randn(50).cumsum(),'k--')
```

Out[11]:
```
[<matplotlib.lines.Line2D at 0x2b52c5e5190>]
```



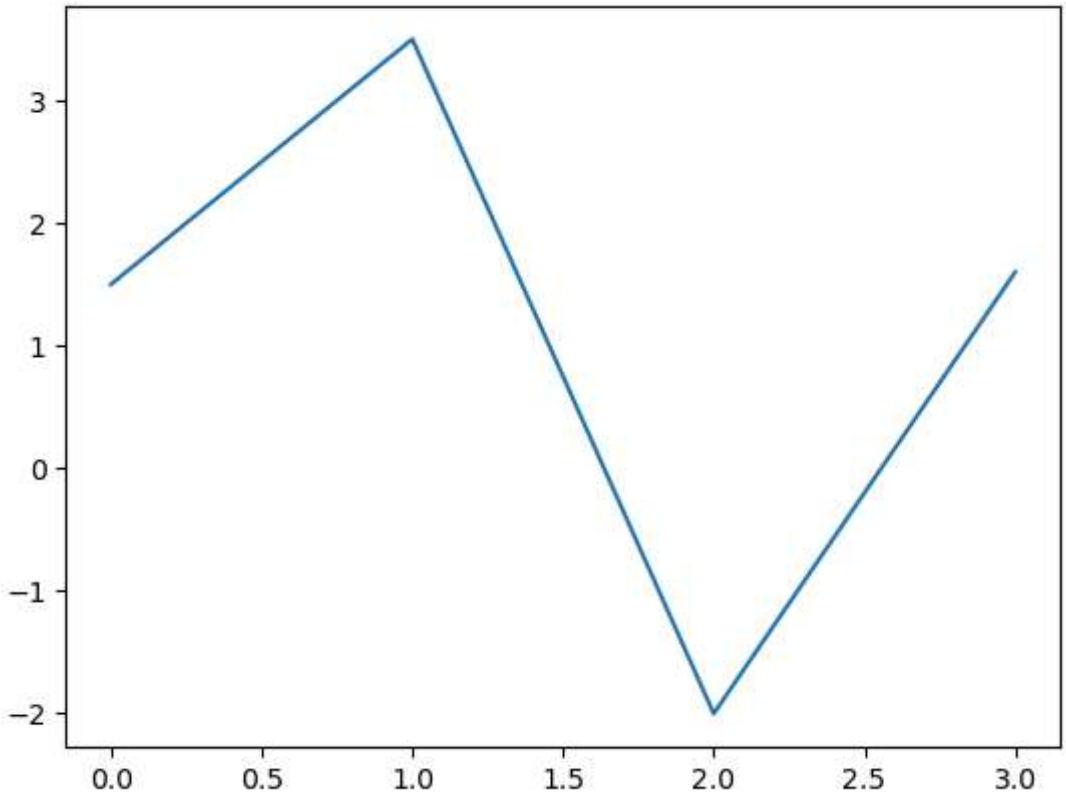
In [12]:
```
_ = ax1.hist(np.random.randn(100), bins = 20, color = 'k', alpha=0.3)
```

In [13]:
```
ax2.scatter(np.arange(30),np.arange(30) + 3 * np.random.randn(30))
```

Out[13]:
```
<matplotlib.collections.PathCollection at 0x2b52c5c6fd0>
```

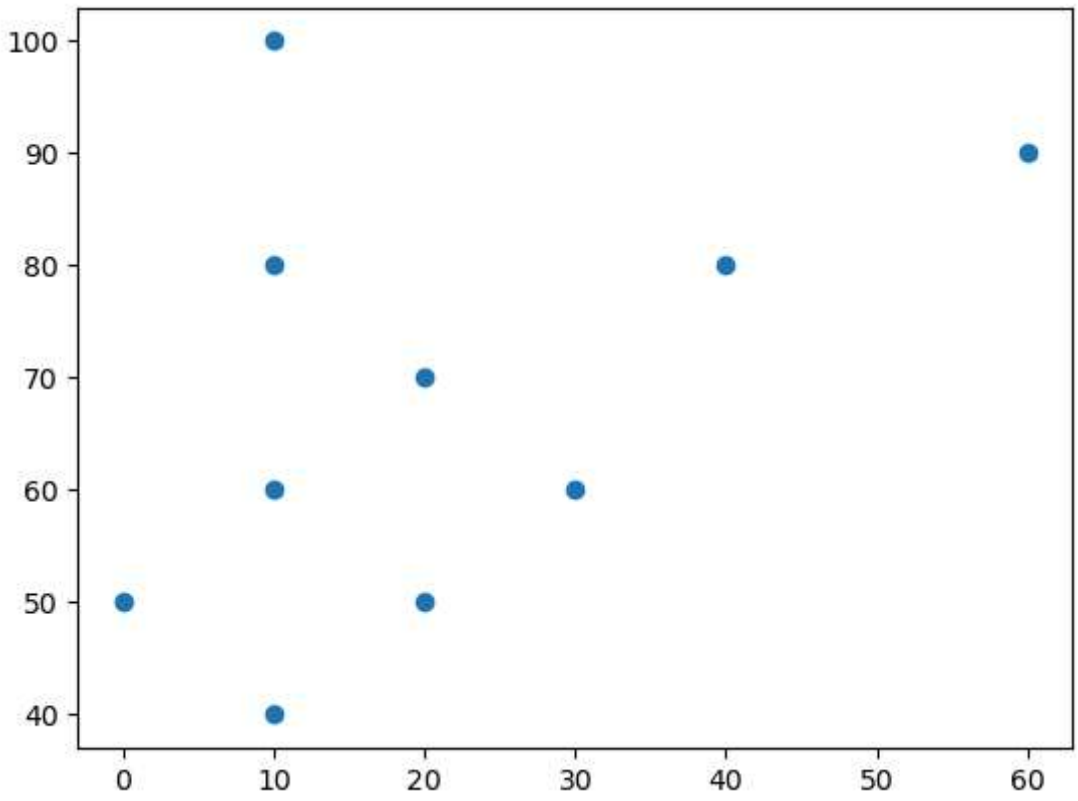
In [14]:
```
plt.plot([1.5, 3.5, -2, 1.6])
```

Out[14]:
```
[<matplotlib.lines.Line2D at 0x2b52da6ed50>]
```



In [15]:
```
plt.scatter(list_preSpelling, list_postSpelling)
```

Out[15]:
```
<matplotlib.collections.PathCollection at 0x2b52bd06d90>
```



In [16]:
```
import pandas as pd
df_cvt_spelling = pd.read_csv("cvt_spelling.csv", sep=";")
```

```
df_cvt_spelling
```

Out[16]:

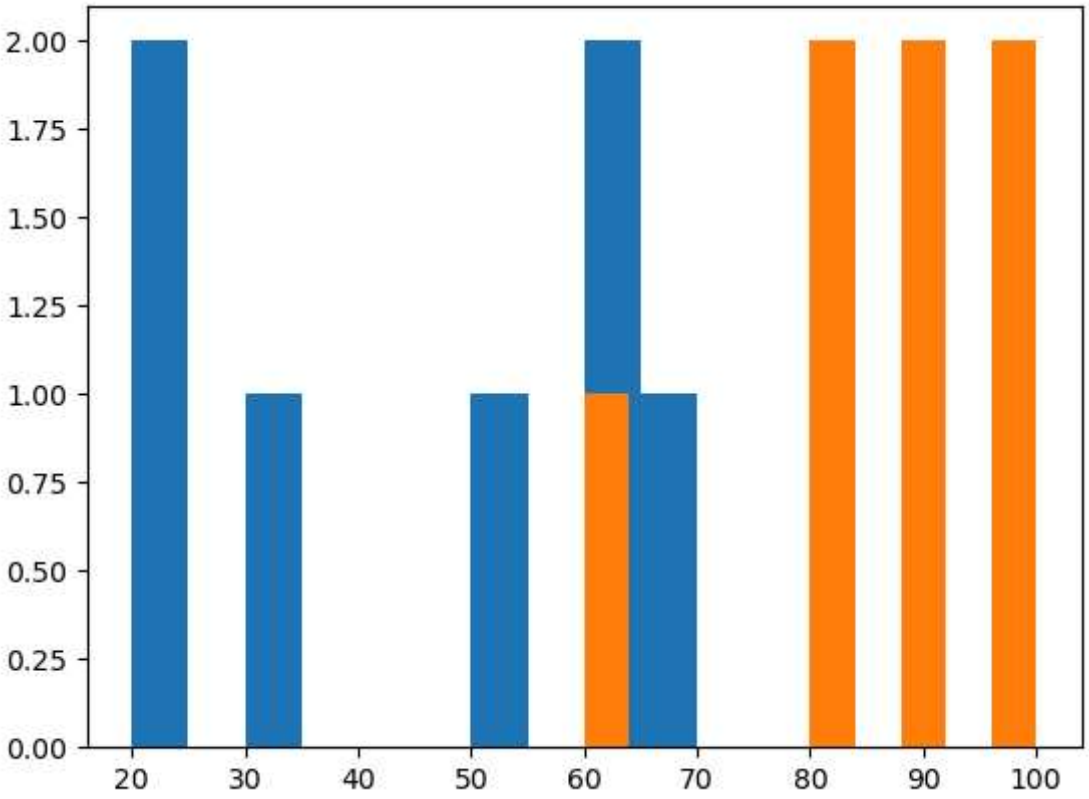
	Unnamed: 0	Nama	Unnamed: 2	PreSpelling	PostSpelling	Gain
0	1	P1	F	60	90	30
1	2	P2	F	70	80	10
2	3	P3	M	60	80	20
3	4	P4	M	30	60	30
4	5	P5	F	20	100	80
5	6	P6	F	20	100	80
6	7	P7	F	50	90	40

In [17]:
```
import numpy as np
cvt_list_preSpelling = []
cvt_list_postSpelling = []
cvt_list_gain = []
np_cvt_list_preSpelling = np.arange(10)
for index, row in df_cvt_spelling.iterrows():
    print(row['PreSpelling'])
    cvt_list_preSpelling.append(row['PreSpelling'])
    cvt_list_postSpelling.append(row['PostSpelling'])
    cvt_list_gain.append(row['Gain'])
np_cvt_list_preSpelling = cvt_list_preSpelling
```

```
60
70
60
30
20
20
50
```

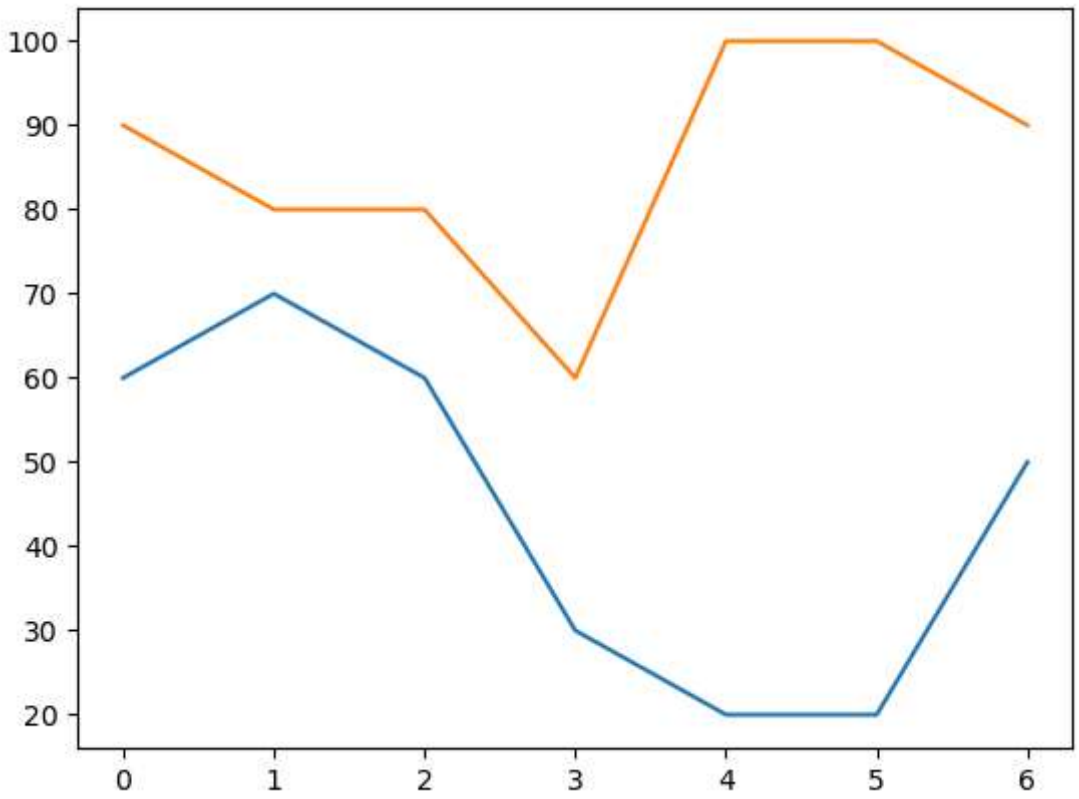
In [18]:
```
import matplotlib.pyplot as plt
plt.hist(cvt_list_preSpelling)
plt.hist(cvt_list_postSpelling)
```

Out[18]:
```
(array([1., 0., 0., 0., 0., 2., 0., 2., 0., 2.]),
 array([ 60.,  64.,  68.,  72.,  76.,  80.,  84.,  88.,  92.,  96., 100.]),
 <BarContainer object of 10 artists>)
```



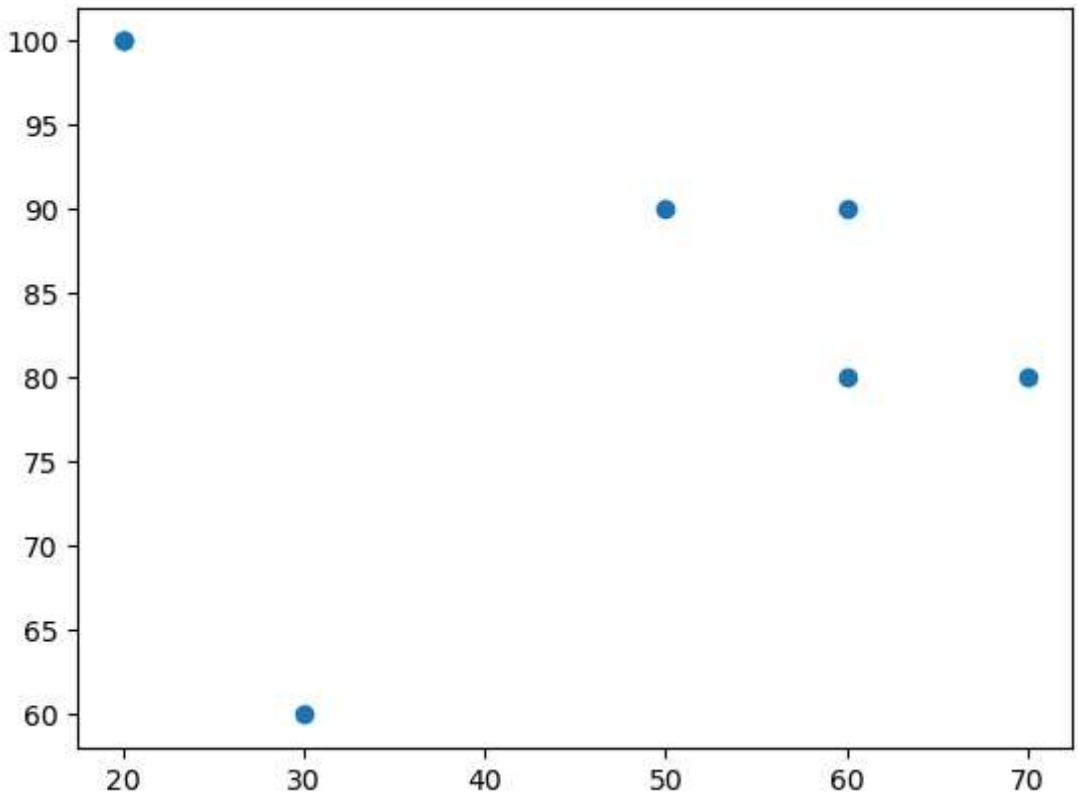
In [19]:
```
plt.plot(cvt_list_preSpelling)
plt.plot(cvt_list_postSpelling)
```

Out[19]: [<matplotlib.lines.Line2D at 0x2b52dbf5890>]



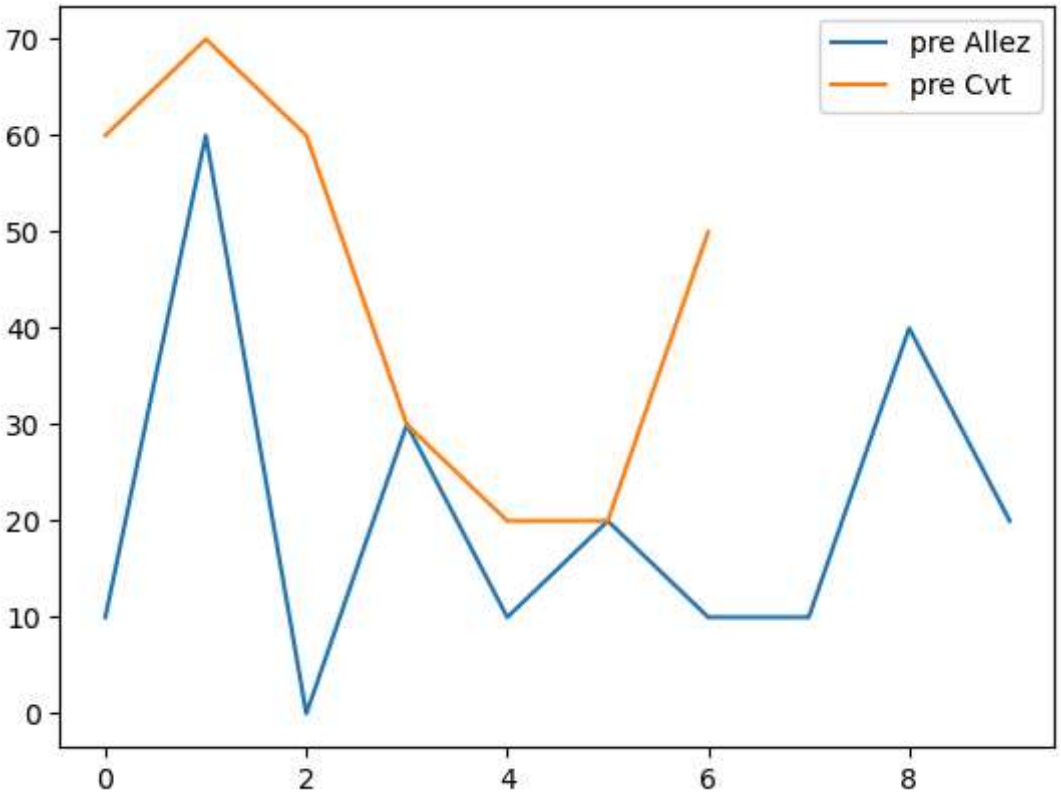
In [20]:
```
plt.scatter(cvt_list_preSpelling, cvt_list_postSpelling)
```

Out[20]:
```
<matplotlib.collections.PathCollection at 0x2b52d928690>
```



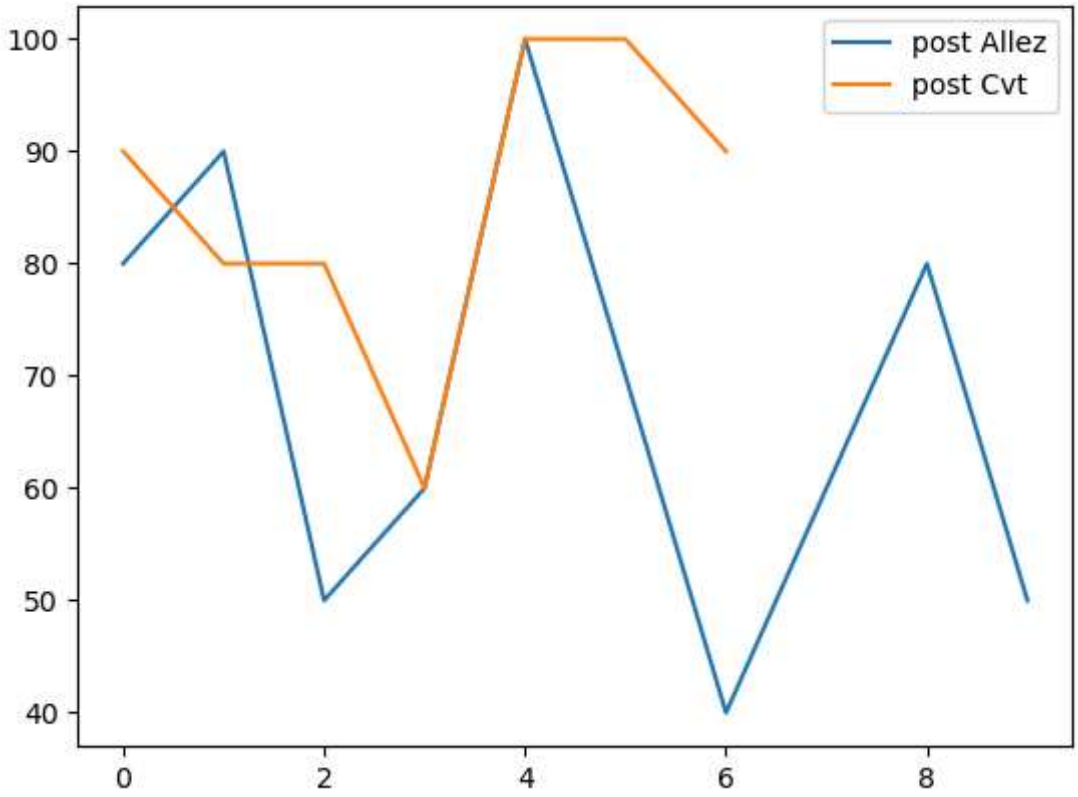
In [21]:
```
plt.plot(list_preSpelling, label="pre Allez")
plt.plot(cvt_list_preSpelling, label = "pre Cvt")
plt.legend(loc = 'best')
```

Out[21]:
```
<matplotlib.legend.Legend at 0x2b52d9a3e50>
```



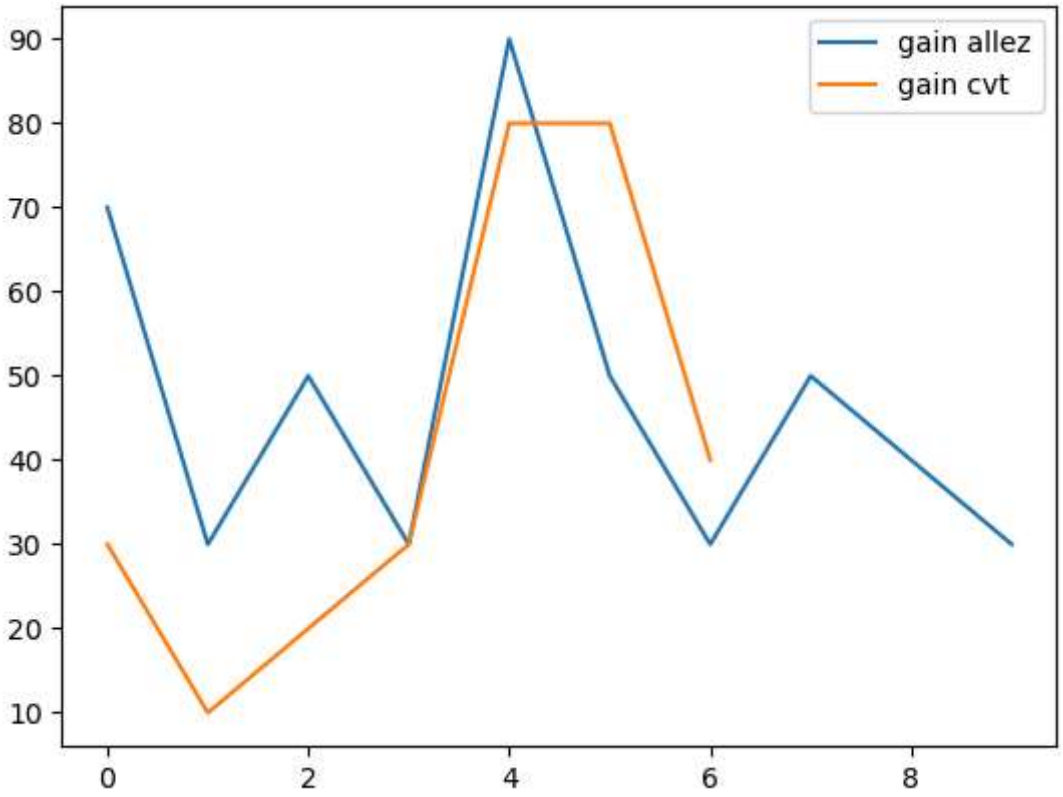
In [22]:
```
plt.plot(list_postSpelling, label="post Allez")
plt.plot(cvt_list_postSpelling, label = "post Cvt")
plt.legend(loc = 'best')
```

Out[22]: <matplotlib.legend.Legend at 0x2b52dc6be50>



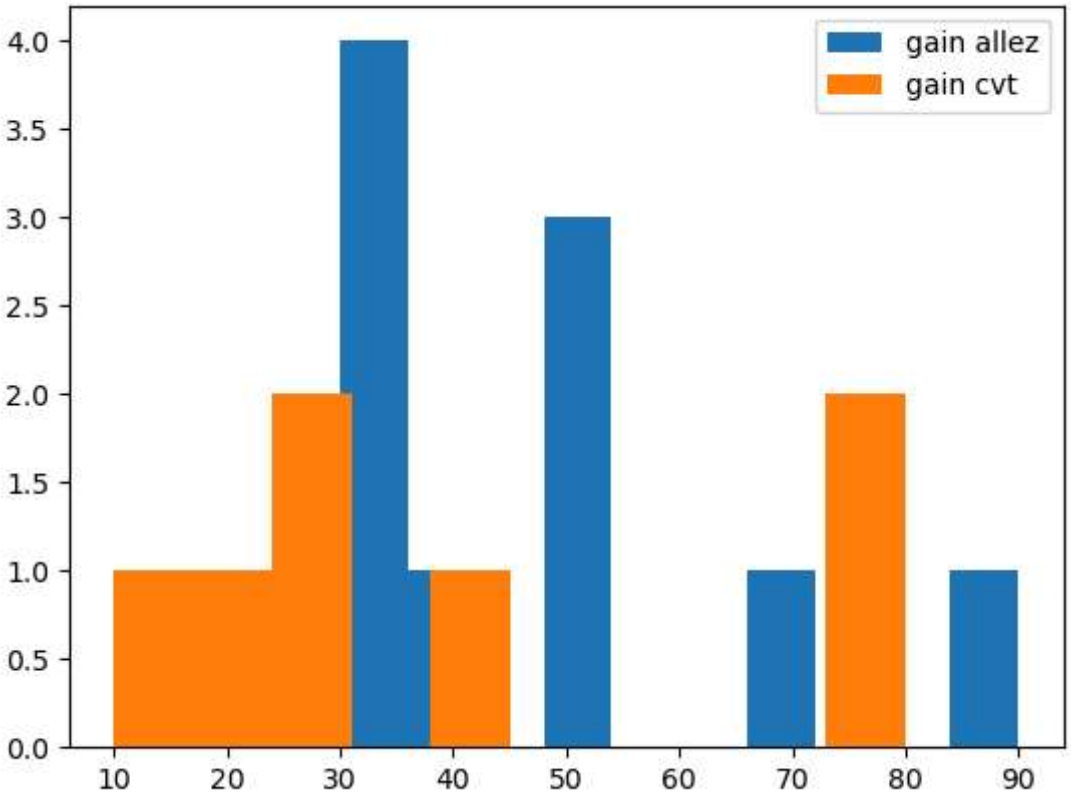
In [23]:
```
plt.plot(list_gain, label="gain allez")
plt.plot(cvt_list_gain, label = "gain cvt")
plt.legend(loc = 'best')
```

Out[23]:
```
<matplotlib.legend.Legend at 0x2b52dcc3e50>
```



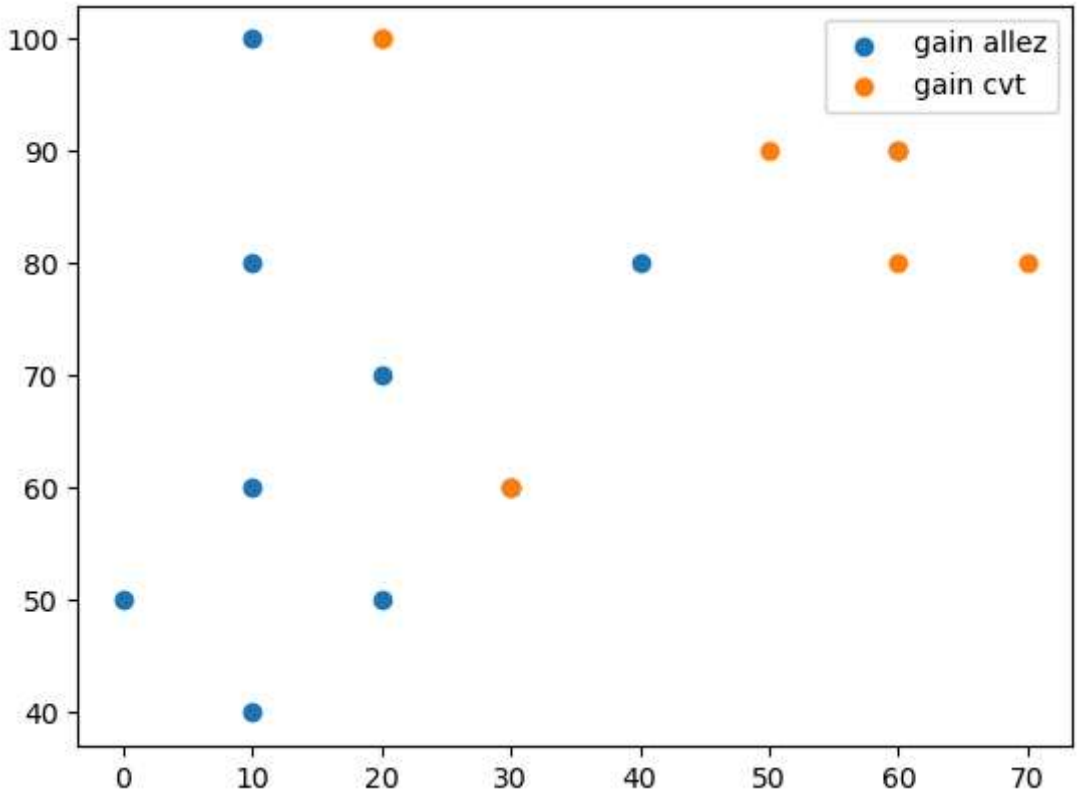
In [24]:
```
plt.hist(list_gain, label = "gain allez")
plt.hist(cvt_list_gain, label = "gain cvt")
plt.legend(loc = 'best')
```

Out[24]:
```
<matplotlib.legend.Legend at 0x2b52de8be50>
```



In [25]:
```
plt.scatter(list_preSpelling, list_postSpelling, label = "gain allez")
plt.scatter(cvt_list_preSpelling, cvt_list_postSpelling, label = "gain cvt")
plt.legend(loc = 'best')
```

Out[25]: <matplotlib.legend.Legend at 0x2b52de65b50>



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: