Assignment 1

Problem 1

Solve the following linear system using Numpy and round the answer to 2 decimals:

3a + 7b + 5c = 1

2 + 2b + 3c = 8

5a + 6c = 4

In [5]:

```
import numpy as np
a = np.array([[3, 7, 5], [0, 2, 3], [5, 0, 6]])
b = np.array([1, 6, 4])
x = np.linalg.solve(a, b)
np.around(x, decimals=2)
```

Out[5]:

```
array([-2.15, -0.69,  2.46])
```

Problem 2

Find the 4th largest value in the following array:

In [10]:

```
import numpy as np

np.random.seed(42)
array = np.random.rand(20)
print(array)
```

```
[0.37454012 0.95071431 0.73199394 0.59865848 0.15601864 0.15599452
 0.05808361 0.86617615 0.60111501 0.70807258 0.02058449 0.96990985
 0.83244264 0.21233911 0.18182497 0.18340451 0.30424224 0.52475643
 0.43194502 0.29122914]
```

Hint: look for functions to sort arrays in Numpy documentation!

In [13]:

```
array.sort()
#print(array)
print ('The 4th largest value in the array is ' + str(array[-4]))
```

```
The 4th largest value in the array is 0.8324426408004217
```

Problem 3

Manipulate the Pandas dataframe following the instructions.

In [1]:

```
import pandas as pd
```

Read a `csv` file which is called `assignment.csv`.

In [2]:

In [2]:

```
assignment = pd.read_csv(r"assignment.csv", sep=',')
```

View the first 5 rows in this dataframe.

In [3]:

```
assignment.head(5)
```

Out[3]:

	Order ID	Order Type	Source	Destination	Begin Time	End Time
0	A5007D	retrieval	M0086	M3111	2017/10/22 00:20	2017/10/22 01:44
1	A5101D	reshuffle	M0026	M0045	2017/10/28 23:35	2017/10/29 00:12
2	A5101D	retrieval	M0045	M3111	2017/10/29 00:17	2017/10/29 02:47
3	A5101D	storage	M7312	M0026	2017/10/28 05:46	2017/10/28 06:01
4	A5152D	retrieval	M0396	M3111	2017/10/22 00:21	2017/10/22 01:46

Sort the full dataframe by `Order Type`.

In [4]:

```
assignment.sort_values(by=['Order Type'])
```

Out[4]:

	Order ID	Order Type	Source	Destination	Begin Time	End Time
1	A5101D	reshuffle	M0026	M0045	2017/10/28 23:35	2017/10/29 00:12
14	A5559D	reshuffle	P0286	P0336	2017/10/31 06:55	2017/10/31 06:57
0	A5007D	retrieval	M0086	M3111	2017/10/22 00:20	2017/10/22 01:44
15	A5559D	retrieval	P0336	P7213	2017/10/31 07:34	2017/10/31 08:23
11	A5499D	retrieval	P0976	R7212	2017/10/24 07:33	2017/10/24 08:00
10	A5499D	retrieval	N0076	P0055	2017/10/24 04:32	2017/10/24 05:15
18	A5656D	retrieval	P0036	Q1111	2017/10/24 02:32	2017/10/24 04:24
19	A5656D	retrieval	Q0317	R7212	2017/10/24 07:34	2017/10/24 08:13
5	A5384D	retrieval	G0126	G7213	2017/10/29 00:19	2017/10/29 02:27
4	A5152D	retrieval	M0396	M3111	2017/10/22 00:21	2017/10/22 01:46
2	A5101D	retrieval	M0045	M3111	2017/10/29 00:17	2017/10/29 02:47
7	A5440D	retrieval	N0947	N7712	2017/10/31 15:15	2017/10/31 15:44
8	A5440D	storage	P1312	N0946	2017/10/31 04:56	2017/10/31 05:20

	Order ID	Order Type	Source	Destination	Begin Time	End Time
					04:56	05:30
6	A5384D	storage	G7412	G0126	2017/10/28 08:02	2017/10/28 08:11
12	A5499D	storage	P0075	P0976	2017/10/24 06:26	2017/10/24 06:46
13	A5499D	storage	P3212	N0076	2017/10/24 01:49	2017/10/24 03:26
3	A5101D	storage	M7312	M0026	2017/10/28 05:46	2017/10/28 06:01
16	A5559D	storage	P0055	P0286	2017/10/31 05:09	2017/10/31 05:29
17	A5559D	transfer	P3422	Q1112	2017/10/31 03:36	2017/10/31 04:12
9	A5440D	transfer	P7312	P0065	2017/10/31 02:23	2017/10/31 02:41

Select all rows whose `Order Type` is `storage`.

In [5]:

```
assignment.loc[assignment["Order Type"] == 'storage']
```

Out[5]:

	Order ID	Order Type	Source	Destination	Begin Time	End Time
3	A5101D	storage	M7312	M0026	2017/10/28 05:46	2017/10/28 06:01
6	A5384D	storage	G7412	G0126	2017/10/28 08:02	2017/10/28 08:11
8	A5440D	storage	P1312	N0946	2017/10/31 04:56	2017/10/31 05:30
12	A5499D	storage	P0075	P0976	2017/10/24 06:26	2017/10/24 06:46
13	A5499D	storage	P3212	N0076	2017/10/24 01:49	2017/10/24 03:26
16	A5559D	storage	P0055	P0286	2017/10/31 05:09	2017/10/31 05:29

Transform `Begin Time` and `End Time` to date format.

In [6]:

```
assignment['Begin Time'] = pd.to_datetime(assignment['Begin Time'] )
```

In [7]:

```
assignment['End Time'] = pd.to_datetime(assignment['End Time'] )
```

Group the dataframe by `Order ID` and `Order Type`, and count the size.

In [8]:

```
assignment.groupby(['Order ID'] and ['Order Type']).count()
```

Out[8]:

	Order ID	Source	Destination	Begin Time	End Time
Order Type					

Order Type	Order ID	Source	Destination	Begin Time	End Time
reshuffle	2	2	2	2	2
retrieval	10	10	10	10	10
storage	6	6	6	6	6
transfer	2	2	2	2	2

Use `Processing Time` = `End Time` - `Begin Time`, add a new column called `Process Time`.

In [9]:

```
assignment['Processing Time'] = assignment['End Time'] - assignment['Begin Time']
```

Compute the mean and median of `Process Time`.

In [10]:

```
print ('The mean of Process Time is ' + str(assignment['Processing Time'].mean()))
print ('The median of Process Time is ' + str(assignment['Processing Time'].median()))
```

```
The mean of Process Time is 0 days 00:51:42
The median of Process Time is 0 days 00:36:30
```

Problem 4

Use Matplotlib to draw a basic line plot.

- **use `np.arange(0.0, 1.0, 0.01)` as x-values**
- **use `x ** n` for `n=1, 2, 3, 4` as 4 group y-values**
- **plot four lines together in one graph with legend**

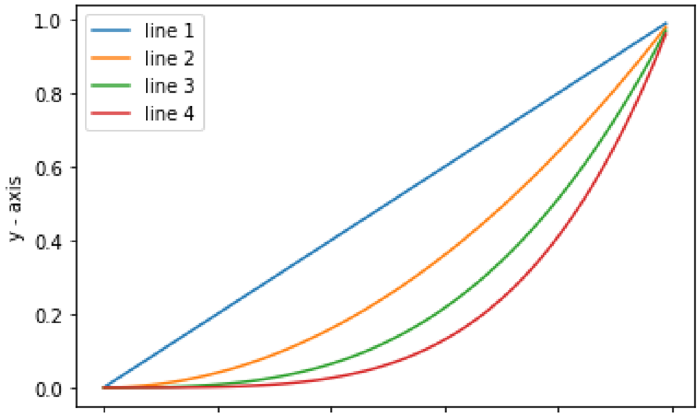
In [11]:

```
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np

# Draw the line plot
x = np.arange(0.0, 1.0, 0.01)
y1 = x**1
y2 = x**2
y3 = x**3
y4 = x**4

plt.plot(x, y1, label = "line 1")
plt.plot(x, y2, label = "line 2")
plt.plot(x, y3, label = "line 3")
plt.plot(x, y4, label = "line 4")
plt.xlabel('x - axis')
plt.ylabel('y - axis')
plt.legend()

plt.show()
```



Problem 5

Suppose your data is `np.random.randn(1000, 3)` where the numpy seed is 19680801 (`np.random.seed(19680801)`)

You need to provide the histogram with 10 bins based on your data.

Let's color them via `'red', 'tan', 'lime'`.

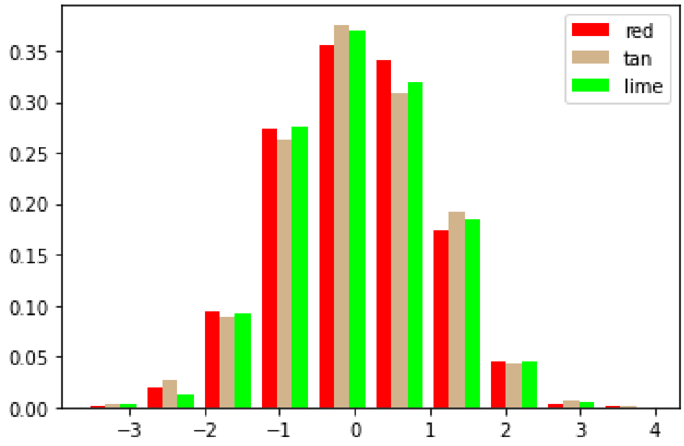
Also set your graph title as `bars with legend`.

In [12]:

```
np.random.seed(19680801)

n_bins = 10
x = np.random.randn(1000, 3)
colors = ['red', 'tan', 'lime']

# Draw the bar chart
plt.hist(x, n_bins, density=True, histtype='bar', color=colors, label=colors)
plt.legend(prop={'size': 10})
plt.show()
```


red
tan
lime