

# Week 4: Introduction to Probability

Download the “lab4\_prob.ipynb” python notebook from Surrey Learn. Follow the sections of this notebook of the code as explained in the tutorial slides.

## Exercise:

Experiment with the following code (as python code using any development environment of your choice (PyCharm or Spider) or Python notebooks using Jupyter, or command line). Use the weather data for New York Park you can download from Surrey Learn, in Lab 1 material section. Write a report to explain the findings and share it in the discussion forum of Lab 4 in Surrey Learn.

```
%matplotlib inline

import numpy as np #matrices and data structures
import scipy.stats as ss #standard statistical operations
import pandas as pd #keeps data organized, works well with data
import matplotlib.pyplot as plt #plot visualization

#read a csv
nyw = pd.read_csv('NYC-CParkWeather.csv')
nyw = nyw.set_index('year')

#represents observations

nyw.head()

nyw.describe()

#plot some data

nyw_lastten = nyw[:10]

plt.bar(nyw_lastten.index,nyw_lastten['mean'])
plt.ylabel('mean temperature')
plt.xlabel('year')
plt.show()

#directly plotting with pandas wrapper

nyw_lastten['mean'].plot(kind='bar')

nyw_lastten[['mean', 'low', 'high']].plot(kind='bar')

plt.show()

#plot a histogram
```

```
nyw_avgtemp = nyw[ 'mean ' ]

nyw_avgtemp.hist(bins=15)
#hint: try other bin sizes!

nyw_avgtemp.min(), nyw_avgtemp.max()

plt.show()

#plot a normal distribution with parameters fit to the data:
fit = ss.norm.pdf(sorted(nyw_avgtemp), nyw_avgtemp.mean(), nyw_avgtemp.std())
plt.plot(sorted(nyw_avgtemp), fit)

plt.show()

#create a custom distribution (based on x-squared): class definition

class sq_dist(ss.rv_continuous):

    def _pdf(self,x):
        return 3*x**2.0

#create an instance of the random variable:
sq_crv = sq_dist(a=0.0, b=1.0, name="sq_pdf") #a to b is the range

#set a range to plot
#x = np.arange(0.0, 1.0, .05)
x = np.arange(-0.1, 1.1, .001)

#to see the edges of the distribution, use this

plt.plot(x, sq_crv.pdf(x))
```