EMPIRICAL BOOTSTRAP BASED CONFIDENCE INTERVAL

* This is a simple code snippet to demonstrate bootstrap sampling method to obtain C.I for median of a population given sample of size 10.

* Dataset : 1 sample of size 10 of heights (in cm)

* Objective: Obtain Confidence Interval of population median from given sample using bootstrap method

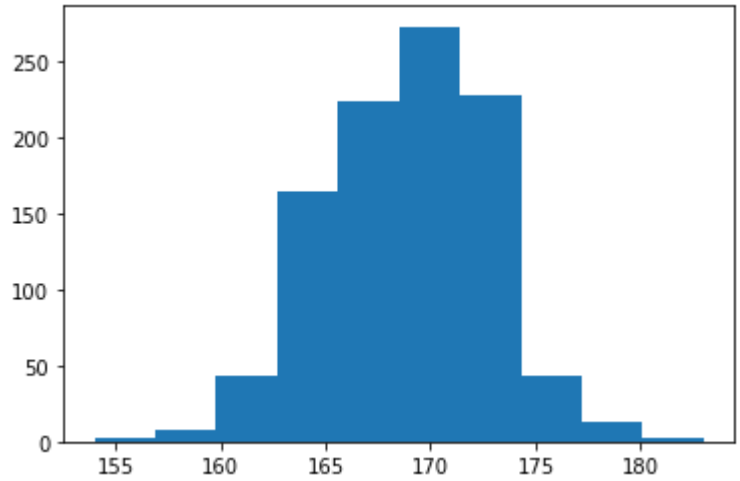
note: this is part of a code walkthrough of Applied AI Course.

In [3]:

```
#importing required libraries
import pandas as pd
import numpy as np
from sklearn.utils import resample
from sklearn.metrics import accuracy_score
from matplotlib import pyplot
```

In [9]:

```
#Dataset
x = np.array([180,162,158,172,168,150,171,183,165,176])

#Configure bootstrap
n_iterations = 1000
n_size = int(len(x))

#run bootstrap
medians = list()
for i in range(n_iterations):
    #prepare train and test sets
    s = resample(x,n_samples = n_size) #bootstrap sample
    m = np.median(s)
    #print m
    medians.append(m)

#plot scores
pyplot.hist(medians)
pyplot.show()

#confidence intervals
alpha = 0.95
p = ((1-alpha)/2)*100
lower = np.percentile(medians,p)

q = (alpha + ((1-alpha)/2))*100
upper = np.percentile(medians,q)

print("{}% Confidence Interval(C.I) of medians is between {} and {} ".format(alpha*100,lowe
```




```
95.0% Confidence Interval(C.I) of medians is between 162.0 and 176.0
```

In conclusion:

- Distribution of heights was unknown(population)
- Random sample of size 10 was given (sample)
- Objective was to calculate C.I for **median of population** given only a sample
- Using Bootstrap Method , objective is satisfied

references:

- Bootstrapping (https://en.wikipedia.org/wiki/Bootstrapping_(statistics))
- Confidence Interval(C.I) (https://www.mathsisfun.com/data/confidence-interval.html)