Central Limit Theorem - Part II

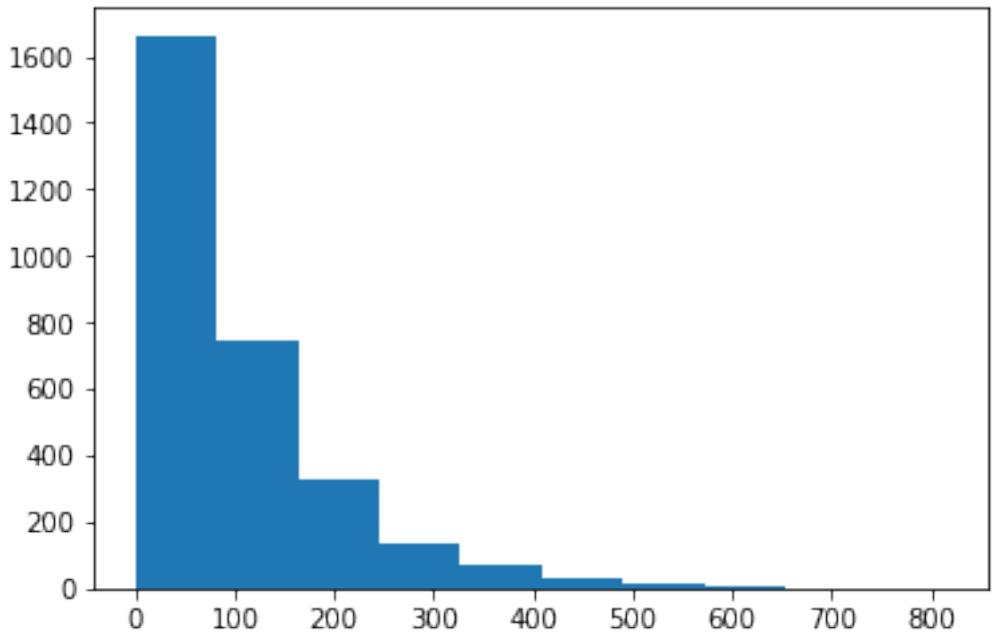
December 1, 2017

0.0.1 Central Limit Theorem - Part II

Work through the questions and use the created variables to answer the questions that follow below the notebook.

Run the below cell to get started.

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt

%matplotlib inline
np.random.seed(42)

pop_data = np.random.gamma(1,100,3000)
plt.hist(pop_data);
```



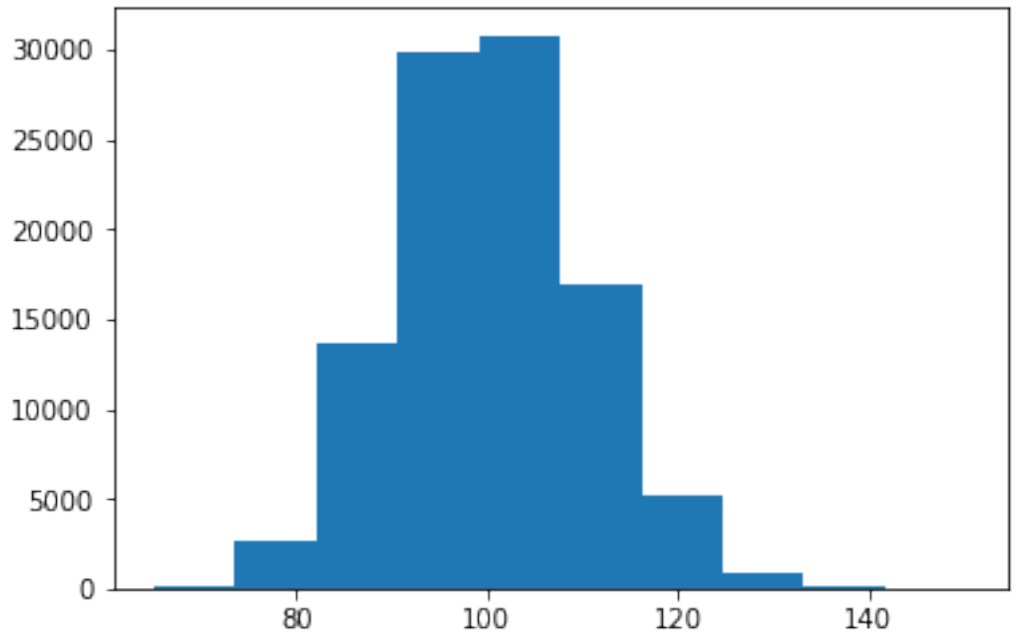
In [2]:
```
pop_data.mean()
```

```
Out[2]: 100.35978700795846

In [3]: np.std(pop_data)

Out[3]: 99.778601879689063

In [9]: np.var(pop_data)

Out[9]: 9955.7693930654896
```

1. In order to create the sampling distribution for the average of 100 draws of this distribution, follow these steps:

a. Use numpy's **random.choice** to simulate 100 draws from the pop_data array. b. Compute the mean of these 100 draws. c. Write a loop to simulate this process 10,000 times, and store each mean into an array called **means_size_100**. d. Plot a histogram of your sample means. e. Use **means_size_100** and **pop_data** to answer the quiz questions below.

```
In [4]: means_size_100 = []
        for _ in range(100000):
            means_size_100.append(np.random.choice(pop_data, size=100).mean())

In [5]: plt.hist(means_size_100);
```




```
In [7]: np.mean(means_size_100)

Out[7]: 100.40950210664138
```

```
In [8]: np.std(means_size_100)
```

```
Out[8]: 9.9710836414513651
```

```
In [10]: np.var(means_size_100)
```

```
Out[10]: 99.422508984819018
```