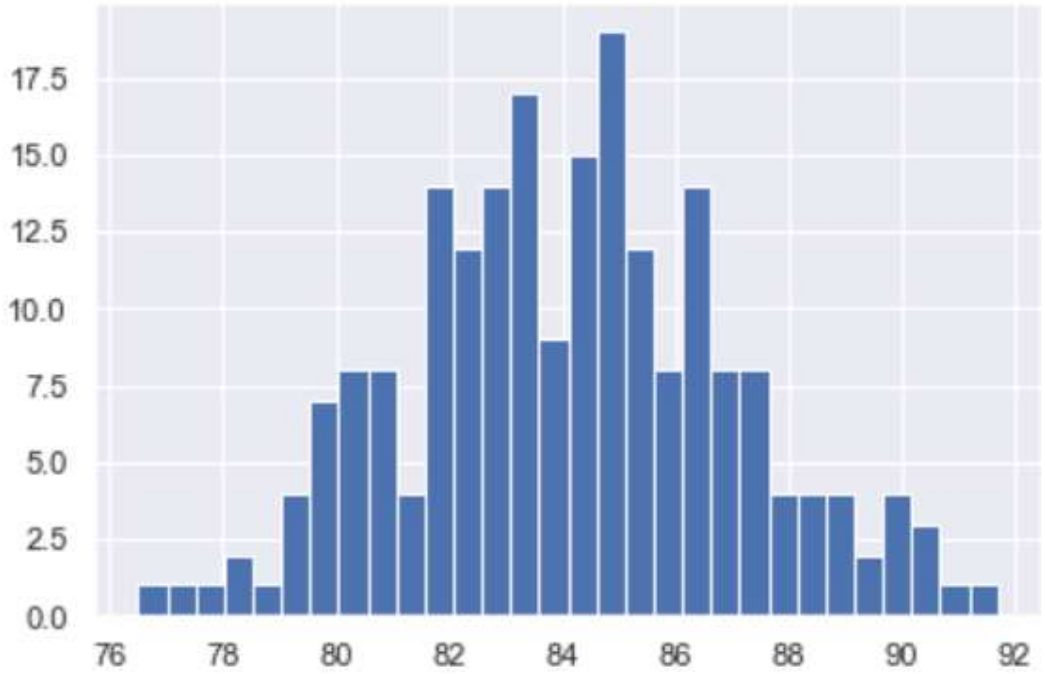
Exercise 6 Micro 110 Spring - Solutions

In this exercise, you will use Python to perform the analysis for the yield improvement exercise covered in slides 31 to 37 of Module 3.

2) Generate a histogram of obs. Record the mean and standard deviation



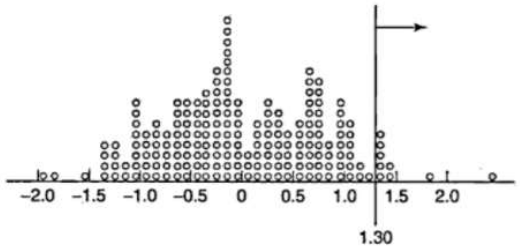

```
mean of obs: 84.12142857142857 standard deviation of obs: 2.8808836790793046
```

3) You will now create a new array (or column if you're using excel), with the goal of matching the values in the ave10 column. Create a new array called CalcAve10, and assign it a formula to calculate the moving average of the past 10 observations.

The values that you calculate should exactly match the values in ave10.

4) You will now create an array to generate a relative reference set of data by plotting the change in $y$bar for adjacent set of 10 observations. This process is demonstrated on slide 32 of module 2:

- **I generate a relative reference set of data by plotting the change in $\overline{y}$ for adjacent set of 10 observations**



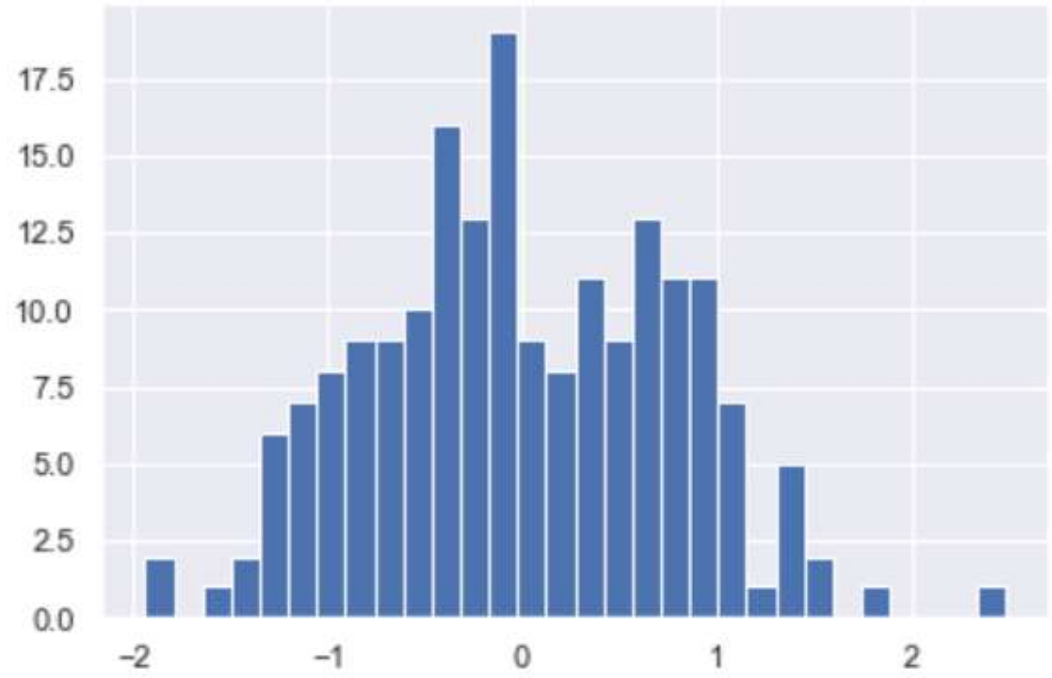
Only 9/147 observations support the null hypothesis (4.7% significance level), so the process probably works.

Note: to generate this reference set, I needed a lot of prior data, which isn't always available

a. Calculate the difference in the moving averages.

```
[-0.43 -1.24 -0.15 -0.02 -0.08 -0.15 -0.79 -0.38 -0.26 -0.1   0.82  0.9
 -0.68 -0.66 -1.25 -0.27  0.13  0.21  0.24  0.29 -0.18  0.43  1.47  1.33
  2.48  1.01  1.33  0.29  0.57  0.95 -0.42 -0.36 -0.52 -1.33 -1.81 -0.36
 -1.02  0.21 -0.29 -0.91  0.64 -0.17 -0.17  0.96  0.78 -0.13  0.3  -0.34
  0.71  0.68  0.53  1.01  1.46  0.76  1.04  1.35  1.37  0.88 -0.12  0.2
 -0.12 -0.37 -1.38 -0.9  -0.8  -1.04 -1.94 -0.9  -0.76 -0.63 -0.94 -0.32
 -0.21 -0.36 -0.93 -0.75  0.13  0.39  0.38 -0.22  0.2  -0.37 -0.16  0.12
  0.8   0.54  0.08 -1.01 -0.55 -0.05 -0.3   0.33  0.79 -0.11 -0.42  0.3
  1.13  1.25  0.97  0.68  0.68 -0.45 -0.62 -0.03  0.54 -0.43 -1.24 -0.64
 -0.86 -1.1  -0.16  1.09  0.87  1.11 -0.12  0.67  1.01  0.74  0.98  1.87
  0.66 -0.04 -0.6  -0.93  0.02 -0.5  -0.51 -0.67 -0.78 -1.15 -1.07 -0.3
  0.78  0.95 -0.17  0.61  0.74  0.67  0.79  0.66  1.   -0.11 -0.4  -0.45
  0.1  -0.3  -0.97 -0.82 -1.53 -1.2  -1.1  -0.43 -1.32 -1.3  -0.64 -0.58
  0.37  0.03  0.75  0.44  0.17 -0.23  0.97  0.72  0.98 -0.21 -0.81  0.29
  0.49 -0.58 -0.3  -0.01 -0.61  0.4  -1.06 -0.13 -0.52 -1.07 -1.4   0.11
  0.46 -0.01  0.33 -0.87 -0.18  0.51  1.39  0.61  0.5   0.64 -0.53]
```

b. Generate a histogram of this data. What is the mean and standard deviation for this histogram?




```
mean of moving average difference: -0.010732984293193778
standard deviation of moving average difference: 0.7763290721899867
```

5) Now, suppose I perform a supposed process improvement, and then compare pre- and post-improvement sample means for sets of 10 adjacent samples. I find that the change in sample means is 1.30, as shown below:

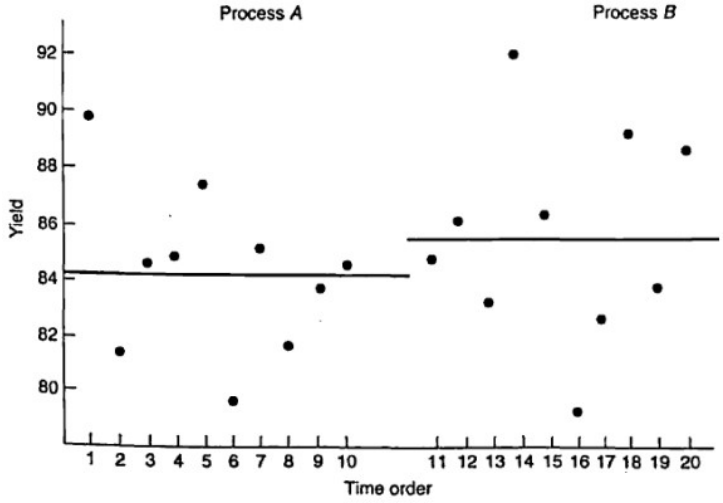


Table 3.1. Yield Data from an Industrial Experiment

Time Order	Method	Yield
1	*A*	89.7
2	*A*	81.4
3	*A*	84.5
4	*A*	84.8
5	*A*	87.3
6	*A*	79.7
7	*A*	85.1
8	*A*	81.7
9	*A*	83.7
10	*A*	84.5
11	*B*	84.7
12	*B*	86.1
13	*B*	83.2
14	*B*	91.9
15	*B*	86.3
16	*B*	79.3
17	*B*	82.6
18	*B*	89.1
19	*B*	83.7
20	*B*	88.5

$\bar{y}_A = 84.24 \quad \bar{y}_B = 85.54$
$\bar{y}_A - \bar{y}_B = 1.30$

a) What is the probability that the process B actually improved things? HINT: You should determine the probability that the reference data from part 4 above could actually result in a sample mean

change for sets of 10 adjacent samples of 1.30. Perform this test by comparing against the empirical data from part 4 and also comparing against a normal distribution with the mean and standard deviation of part 4.

```
empirical probability that sample mean improves by 1.3: 0.047120418848167539

normal distribution probability that sample mean improves by 1.3: 0.04644
```