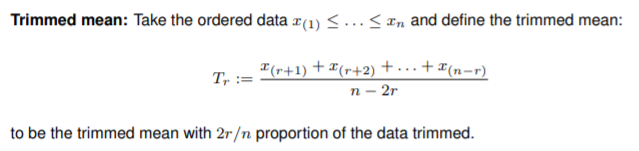
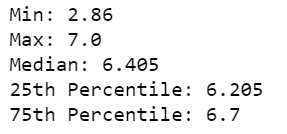
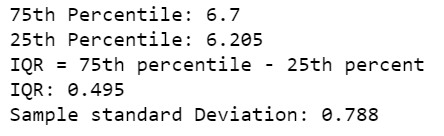
Name: Adam  
 ID: 1002010  
 Cohort: 1

**Statistics PSET 1:**

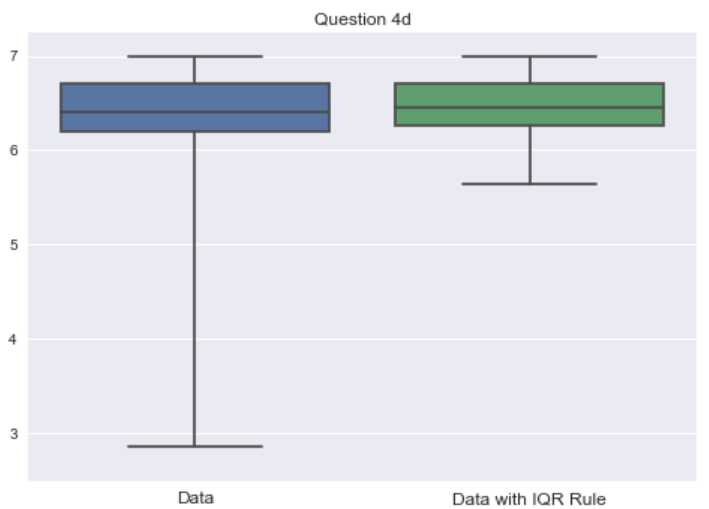
1. More likely to be biased because those who are driving will be less likely to answer the question and thus, drivers will be under represented in the poll
2. a.   
   Control: Initial Placebo Group  
   Treatment: Initial Aspartame Group  
     
   b.   
   Cross over design eliminates the possibility of biasness in the initial sampling.  
   In addition, fewer subjects might be required in order to attain the same level of statistical power as compared to a study with only one treatment  
     
   c.   
   No. It must be a quantitative argument to justify such i.e. *level of confidence*
3. a.   
   Systematic sampling  
     
   b.  
   Stratified sampling  
     
   c.  
   Simple random sampling
4. a.   
     
     
     
     
     
   b.

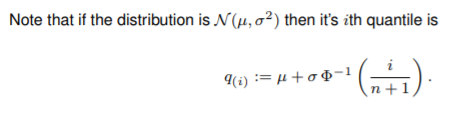


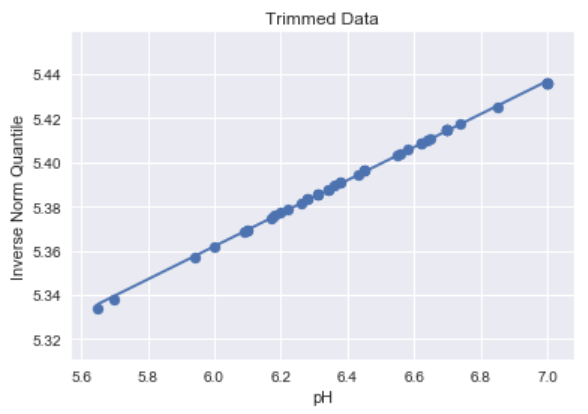
c.

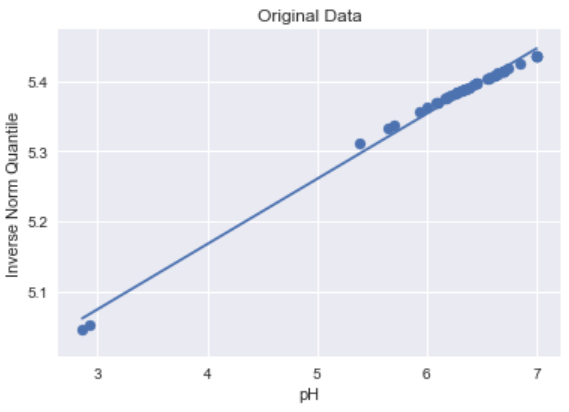


d.





e.



The original data is skewed to the right, the trimmed data is normal

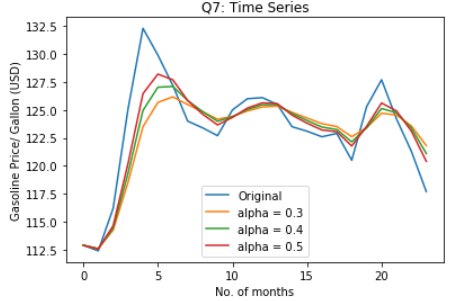
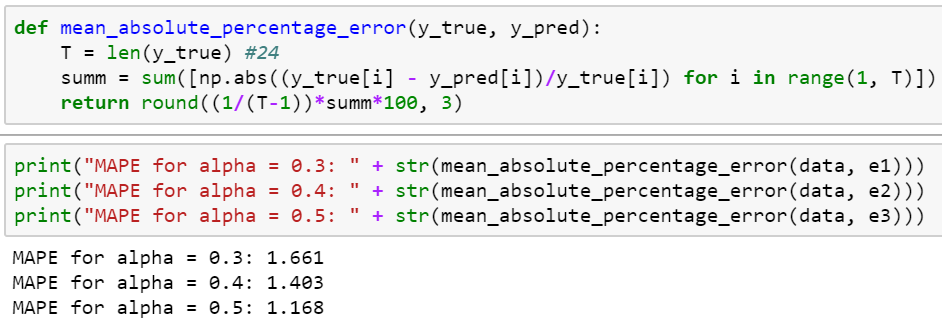
1. a.   
   Assume that participants are telling the truth   
     
   p = q\*pi + (1-q)/12  
     
   pi = (p – (1-q)/12) / q  
     
   b.   
   Yes it is possible.  
     
   c.   
   q=0, or q=1
2. a.  
   Hospital A has a far higher success rate  
   Success Rate = (Success/ Total)

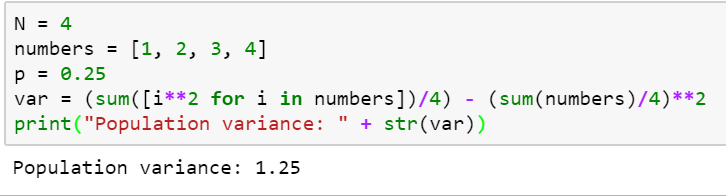
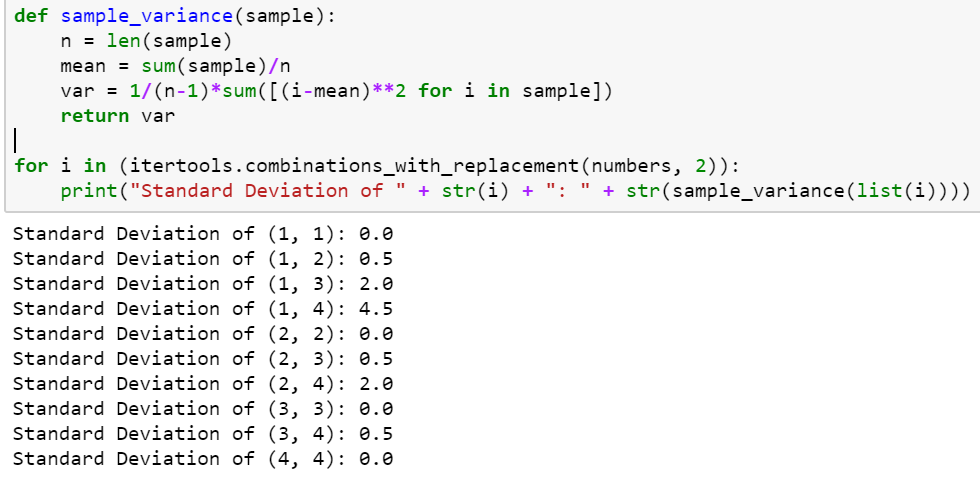
|  |  |  |
| --- | --- | --- |
|  | Low Risk | High Risk |
| Hospital A | 80% | 20% |
| Hospital B | 60% | 10% |

b.

|  |  |
| --- | --- |
| Hospital A | (400+160)/(500+800) = 0.431 |
| Hospital B | (300+20)/(500+200) = 0.457 |

c.  
Simpson’s paradox occurs when the group sizes are uneven.   
One explanation is that Hospital A engages in more high risk treatments, which tend to have lower success rate. Thus, on a whole, A seems worst.

1. a.   
     
   b.   
     
     
   MAPE for alpha = 0.5 is the lowest hence, we should use it to forecast the next price  
   Next forecast:   
   
2. a.

  
  
b.

c.   
Expected value: 1.25  
Same as population variance

1. a.  
   U follows a Normal Distribution with  
   mean: 40 and standard deviation: sqrt((15\*\*2)/50)  
     
   V follows a Normal Distribution with  
   mean: 40 and standard deviation: sqrt((15\*\*2)/5)  
     
   b.  
   V: V has a larger population size thus has less variance.   
   Thus, it has a larger probability of being around the mean  
     
   c.   
   