GERMAN TANK PROBLEM

Names: _______________________________________________________

During World War II, Allied intelligence reports on Germany’s production of tanks varied widely and were somewhat contradictory. Statisticians set to work on improving the estimates. In 1943, they developed a method that used the information contained in the serial numbers stamped on the tanks. They discovered that the serial numbers on the tanks were consecutive. That is, the tanks were numbered in a manner equivalent to 1, 2, 3, ……, N. Capturing a tank was like randomly drawing an integer from this sequence.

One sample collected in 1943 found tanks with the following serial numbers: 21 123 201 297 342

What is your best estimate for the number of tanks, N? ______. Is this value a parameter or a statistic?

In this assignment, your job is to develop a method to estimate the total number of German tanks based on a random sample and the assumption that the tanks are numbered sequentially starting from 1.

For example, one possible method would be to take the median value and double it. In the previous sample, median • 2 = 201 • 2 = 402. Thus, our estimate for the total number of tanks is \( \hat{N} = 402 \).

For this activity, you and your group will create a total of three statistics to estimate the total number of tanks. You may use any combination of the summary statistics we already know (mean, median, min, max, quartiles, IQR, standard deviation, etc.) or invent your own. The goal is to find a relatively simple statistic that reliably predicts the total number of tanks.

List your three statistics to estimate the number of tanks \( N \).

1. __________ 2. __________ 3. __________

To determine which of your three statistics gives the best predictions, you will perform a simulation to generate sampling distributions for each of your three statistics. For the purposes of the simulation, assume that there are 36 tanks total (\( N = 36 \)) and that you will be taking samples of size 5 (\( n = 5 \)). For each trial of your simulation, generate 5 random numbers from 1–36 and compute the values of each of your three statistics. Graph each of these values on a separate dotplot and repeat many times.

What you need to turn in:
- An introduction describing the choice of your 3 statistics
- Three dotplots showing the sampling distributions for each of your statistics. Make sure they are on the same scale so they can be easily compared.
- A written description/comparison of the three distributions.
- A description of which statistic you think is best and why.

NOTE: In class we will have a competition to see which group has the best statistic. The nominating order will be randomly selected and no statistic will be allowed more than once. During each round of the competition, I will secretly choose \( N \) and then give you a sample of \( n \) numbers from 1–\( N \). Each group will compute the value of their statistic based on the sample of \( n \) numbers and the closest estimate gets 10 points, the second closest gets 5 points, and the third closest gets 3 points. After several rounds, the team with the most points will get 10 extra credit points per member.
GERMAN TANK PROBLEM

Names: 1) _________________________________ 2) _________________________________

3) _________________________________ 4) _________________________________

Introduction:

Description/comparison of the three distributions:

Which statistic does your group think is best? Why?
Test your first statistic. Round your estimate to the nearest whole tank.

Sample #1 ______, ______, ______, ______, ______. Estimate: ______
Sample #2 ______, ______, ______, ______, ______. Estimate: ______
Sample #3 ______, ______, ______, ______, ______. Estimate: ______
Sample #4 ______, ______, ______, ______, ______. Estimate: ______
Sample #5 ______, ______, ______, ______, ______. Estimate: ______
Sample #6 ______, ______, ______, ______, ______. Estimate: ______
Sample #7 ______, ______, ______, ______, ______. Estimate: ______
Sample #8 ______, ______, ______, ______, ______. Estimate: ______
Sample #9 ______, ______, ______, ______, ______. Estimate: ______
Sample #10 ______, ______, ______, ______, ______. Estimate: ______
Sample #11 ______, ______, ______, ______, ______. Estimate: ______
Sample #12 ______, ______, ______, ______, ______. Estimate: ______
Sample #13 ______, ______, ______, ______, ______. Estimate: ______
Sample #14 ______, ______, ______, ______, ______. Estimate: ______
Sample #15 ______, ______, ______, ______, ______. Estimate: ______
Sample #16 ______, ______, ______, ______, ______. Estimate: ______
Sample #17 ______, ______, ______, ______, ______. Estimate: ______
Sample #18 ______, ______, ______, ______, ______. Estimate: ______

Make a dotplot of your estimates below. Interval widths are already set up.

What is the mean of your estimates? ________ What is the standard deviation? ________
Test your **second** statistic. Round your estimate to the nearest whole tank.

Sample #1 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #2 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #3 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #4 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #5 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #6 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #7 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #8 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #9 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #10 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #11 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #12 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #13 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #14 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #15 ______ , ______, ______ , ______, ______ . Estimate: ______

Sample #16 ______ , ______, ______, ______, ______ . Estimate: ______

Sample #17 ______ , ______, ______, ______, ______ . Estimate: ______

Sample #18 ______ , ______, ______, ______, ______ . Estimate: ______

Make a dotplot of your estimates below. Interval widths are already set up.

What is the mean of your estimates? ________ What is the standard deviation? ________
Test your third statistic. Round your estimate to the nearest whole tank.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Data Points</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Make a dotplot of your estimates below. Interval widths are already set up.

What is the mean of your estimates? ________  What is the standard deviation? ________