Chapter 5.a - Box Plots and Comparing Groups

1) Quick Assignment Discussion

2) Quiz 4

3) Notes 5.a - Box Plots and Fence

4) Notes 5.a - Comparing Box Plots

---

**Describing the Distribution/Association**

**G - Gaps(Clusters)**: are there any gaps in the data, where do they occur; is there a clear reason why?

**S - Shape**: multiple modes, skewed, symmetric?

**O - Outliers**: where/what are they, are they affecting the stats

**C - Center**: Mean or Median, which is appropriate, why

**S - Spread**: SD or IQR, which is appropriate, what does it tell you
15. 
   a) Set 2 has the greater standard deviation. Both sets have the same mean, 6, but set two has values that are generally farther away from the mean.
      \[ \text{SD(Set 1)} = 2.24 \quad \text{SD(Set 2)} = 3.16 \]
   
   b) Set 2 has the greater standard deviation. Both sets have the same mean (15), maximum(20), and minimum (10), but 11 and 19 are farther from the mean than 14 and 16.
      \[ \text{SD(Set 1)} = 3.61 \quad \text{SD(Set 2)} = 4.53 \]
   
   c) The standard deviations are the same. Set 2 is simply Set 1 + 80. Although the measures of center and position change, the spread is exactly the same.
      \[ \text{SD(Set 1)} = 4.24 \quad \text{SD(Set 2)} = 4.24 \]

1. A survey conducted in a college intro stats class asked students about the number of credit hours they were taking that quarter. The number of credit hours for a random sample of 16 students is given in the table.

<table>
<thead>
<tr>
<th>10</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>15</th>
<th>15</th>
<th>15</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>17</td>
<td>19</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>22</td>
</tr>
</tbody>
</table>

a. Sketch a histogram of these data

b. Find the mean and standard deviation for the number of credit hours.

c. Find the median and IQR for the number of credit hours.

d. Is it more appropriate to use the mean and standard deviation or the median and IQR to summarize these data? Explain.
Creating Box Plots
-sometimes shown vertically for comparative statistics

1) Create a five number summary (Min, Q1, Med, Q3, Max)

2) Create the "box", lines at Q1, Q3 and inside at Med

3) Find fences and draw dotted lines to help
   Upper Fence = Q3 + 1.5(IQR)   Lower Fence = Q1 - 1.5(IQR)

4) Draw "whisker" at lowest and highest point within fence

5) Mark outliers that are data outside of the fences
   [and far outliers]

Example
Use the following data and create a box plot with outliers

```
2  2
17 17
12 14 15 15 15 15 15
17 19 20 20 20 20 22
```

UL = 25
LI = Min = 2
LO = 14.5
Med = 16
UQ = 20
Max = 22

IQR = 5.5
1.5 IQR = 8.25

--- End of Document ---
Example
Use the following data and create side by side box plot

| 10 | 10 | 12 | 14 | 15 | 15 | 15 | 15 |
| 17 | 17 | 19 | 20 | 20 | 20 | 20 | 22 |

Assignment (Due Friday 9/20)

1) Read Chapter 5, Pg. 80-85

2) Pg. 95, #5, 7, 8, 9, 12, 16, 17, 20a-d
   (but think about e)

3) Read TI Tips (pg. 86) and be able to do boxplots