Statistics 8-9: Associations

Gwyneth Hughes

©Initiative for Developing Mathematical Thinking 2015
Suggestions for usage
This webinar is designed to assist teachers, administrators, and instructional coaches in understanding how foundational ideas about statistics might help in the instruction of the Common Core State Standards around data at the middle (and high school) level.

Materials Needed
We recommend using the Voicethread version which will run off the web: http://voicethread.com/share/5904122/ If you won’t have internet access during the webinar, you can download a quicktime movie of the webinar ahead of time from this Google Drive folder https://drive.google.com/folderview?id=0B_4LR-7_JQNgcVkwdUxEazBiTVE&usp=sharing. Right click on the movie and select “Download.”

Before beginning the webinar, teachers should complete the introductory activities that are located in the above Google Drive. They can/should work together to create and discuss their conclusions and representations. Teachers should also have paper and pencil/pen.

Facilitation Ideas
This webinar might be viewed in one workshop, or it could be broken up over multiple meetings to go more in depth into specific topics. To make this easier, the table below lays out the major topics and the associated slides.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Slides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to statistics and the framework for statistical thinking (This is repeated from the 6-9 Distributions webinar)</td>
<td>1-12</td>
</tr>
<tr>
<td>Statistical questions about associations</td>
<td>13-17</td>
</tr>
<tr>
<td>Categorical v. quantitative variables in associations</td>
<td>18-30</td>
</tr>
<tr>
<td>Representing data: Association of two quantitative variables</td>
<td>31-53</td>
</tr>
<tr>
<td>Representing data: Association of two categorical variables</td>
<td>54-59</td>
</tr>
<tr>
<td>Interpreting results, common pitfalls</td>
<td>60-61</td>
</tr>
<tr>
<td>Summary and conclusion</td>
<td>62-66</td>
</tr>
</tbody>
</table>

Full Webinar Facilitation
1. Begin the Webinar.
2. The first 12 slides are a repeat from the distributions webinar so if teachers have just worked through that webinar they could be skipped or run through quickly.
4. Possible pause on slide 26. The difference between quantitative and qualitative variables is really key in terms of thinking about associations but can get a bit tricky. It may be good to pause briefly on that slide to have a discussion about the two types of variables, possibly having teachers brainstorm examples of both types.
5. Slide 33 – Give enough time for teachers to really think about their process in describing the plots and in finding their line of best fit for a few. If possible, give participants a piece of spaghetti to model the line of best fit and ask them to think carefully about how they decide where to place it. Before moving on, have teachers discuss their methods.

6. Slide 42 – Have teachers compare their methods to those discussed. It’s important to note that none of these are how a computer program actually fits a line, as is described on the next few slides.

7. Slides 50-52 focus on the correlation coefficient. Depending on grade level of teachers present, these slides might be examined quickly or discussed in more detail.

8. Slide 55 – Before moving on, have teachers compare their representations and conclusions they made on their own.

9. Slide 58 – This might be a good place to pause and compare the different representations, particularly in thinking about some displays can be misleading.

10. Slide 61 – Have participants think of examples of when a correlation was interpreted as causation in their own lives. Such instances are ubiquitous in education.

**Extensions**

1. Ask teachers to create their own statistical questions that would be appropriate for their students to investigate. Brainstorm ways in which that question could be addressed with data and think about how students might collect and analyze it.

2. Find or collect data and think about what statistical questions could be answered by that data set, and what representations would be appropriate given the type of data.