Patient Experience Data Self-Assessment
Answer Key

1. What is a **run chart?**

Answer: True

Explanation:
"A run chart is a graphical display of data plotted in some type of order. The horizontal axis is most often a time scale (e.g., days, weeks, months, quarters) but could also include sequential patients, visits or procedures. The vertical axis represents the quality indicator being studied (e.g., infection rate, number of patient falls, readmission rate). Usually, the median is calculated and used as the chart's centerline." —Perla et al (2011), "The run chart: a simple analytical tool for learning from variation in healthcare processes", BMJ Qual Saf; 20; 46-51

Why do we emphasize the use of run charts in our Patient Experience work?
You are almost always better able to understand your numerical patient experience data when you plot summary numbers in time order. To go one step beyond a time-ordered plot, you can add a median as a reference line and annotations to create a run chart. Then you are ready to apply a handful of rules to detect unusual patterns worthy of investigation. This second step is what we mean when we use the term *run chart analysis*. You can use more technically advanced tools called **control charts** to analyze data in time order, too.

2. The median of the HCAHPS Nursing Communications top box scores in the accompanying chart…

Answer: The median is 77.

Explanation:
You can find the median on the graph using the definition of median (below). Look at the vertical axis, this scale gives in size order from lowest value to highest value in the 13 month series.

Read off the numbers in size order (from bottom of the graph to top of the graph) as: 75,75,75,76,76,77,77,78,79,79,79, 80,80.

The median is the 7th value, 77, counting from either end of the ranked series.

Review: How do I find the median?

1. Order the values in your data set from small to large.
2. If there is an **odd number** of values (e.g. 13) count in from the lowest value to the middle value, (e.g. the 7th value). That's the median.

   Example:
   56, 56, 58, 60, 62, 63, 63, 65, 67, 70, 70, 73, 74

   Median = 63

   If there is an **even number** of values (e.g. 12) count in from the lowest value to the two middle values (e.g. 6th and 7th values). Average those two values. That's the median

   Example:
   56, 56, 58, 60, 62, 63, 65, 67, 70, 70, 73, 74

   Median = 64
3. In Excel® there is a median function!

For those of you who want a formula to find the position of the median in a set of *n* numbers, here it is:

Position of median = one half of $n+1$

For example, if $n=13$, then the position of the median is one half of 14, or 7. If $n= 20$, then the position of the median is one half of 21, or 10.5 – midway between the 10th and 11th value.

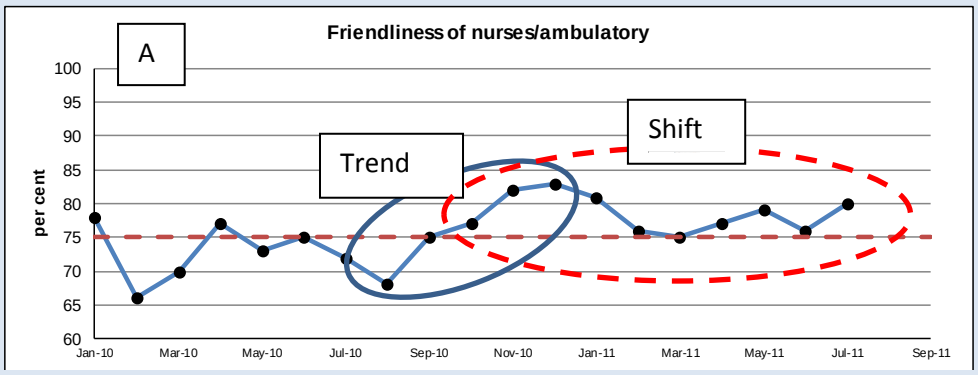
The use of a graph of points to find the median corresponds to a quick graphical method best deployed with a transparent straight-edge, you avoid copying over the values into an ordered list.

1. Count (or estimate) the number of points graphed.
2. Calculate the median position.
3. With the edge of the straight-edge parallel to the horizontal axis, slide the straight-edge up until your edge passes through the median position.

3. **Using the standard run chart rules for shift, trend, and "astronomical point," examine each chart for signs of improvement or degradation. The dashed line is the median value to be used for reference. The first chart uses a median based on the first 12 months, also known as a "baseline reference" median.**

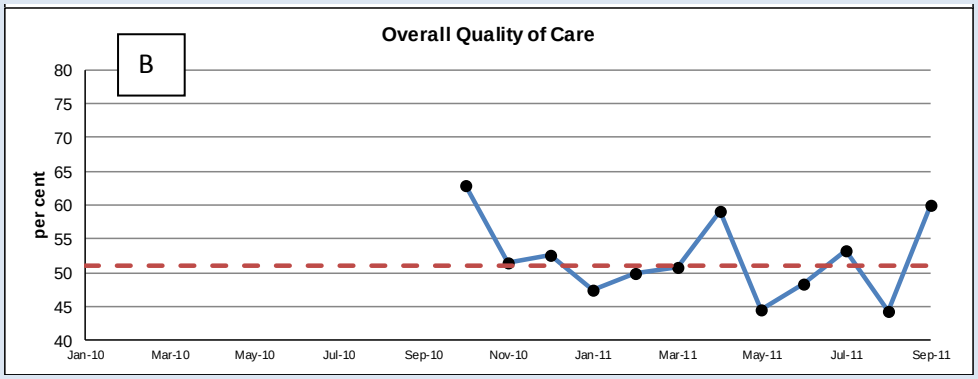
Answers:

Chart A:



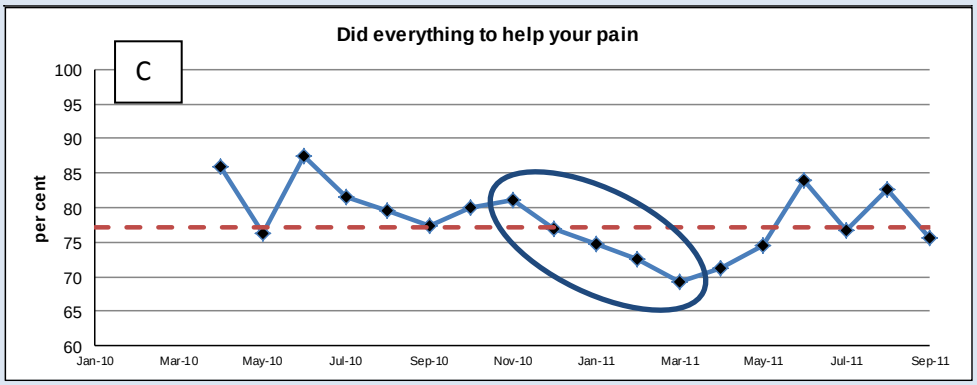
Evidence of improvement: **trend** and **shift**. Note the shift rule applies when we exclude the values that fall on the median.

Chart B:



No signs of change. This does not mean the performance is "acceptable," just that the data series at this point gives no evidence of improvement or degradation.

Chart C:



Signal of degradation, negative trend signal. The visual pattern supports this belief, with two more points below the median after the trend ends.

Explanation:

The BMJ article (reference) outlines four basic rules to detect unusual patterns in a run chart. The presence of unusual patterns may give you clues to system changes or to provide evidence to help you judge whether your interventions have made a difference. A common set of rules helps reduce needless discussions and waste of resources about whether or not performance has actually changed. The run chart rules, while a bridge to more sophisticated tools like control charts, can be used and understood by virtually all healthcare professionals. The rules discussed here are incorporated by IHI in its Extranet web application and a variety of statistical software tools.

The two simplest rules detect presence of a **shift** or a **trend.** The BMJ article defines the terms this way:

> **Shift:** Six or more consecutive points either all above or all below the median. Values that fall on the median do not add to nor break a shift.
>
> **Trend:** Five or more consecutive points all going up or all going down. If the value of two or more consecutive points is the same, only count the first point and ignore the repeating values; like values do not make or break a trend.

4. Interpreting a New Month's Value

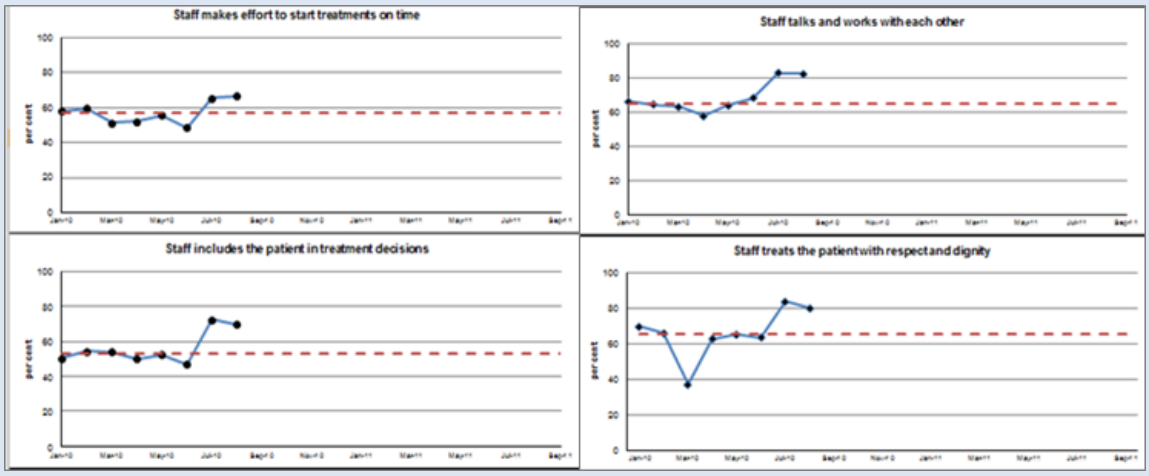
Answers:

a. Yes
b. Yes
c. No
d. No
e. Yes

Explanation:

a. See discussion of run charts earlier in this pre-test and the BMJ reference.

b. Tracking patterns across a family of related measures often provides more insight than considering a measure in isolation. This is the same logic applied to a dashboard of performance measures for a department, hospital or hospital system. Here is an example from an internal survey that shows all four measures moving up, even though none of the run chart rules signal. Taken together, these four charts support a belief that the system is improving performance.

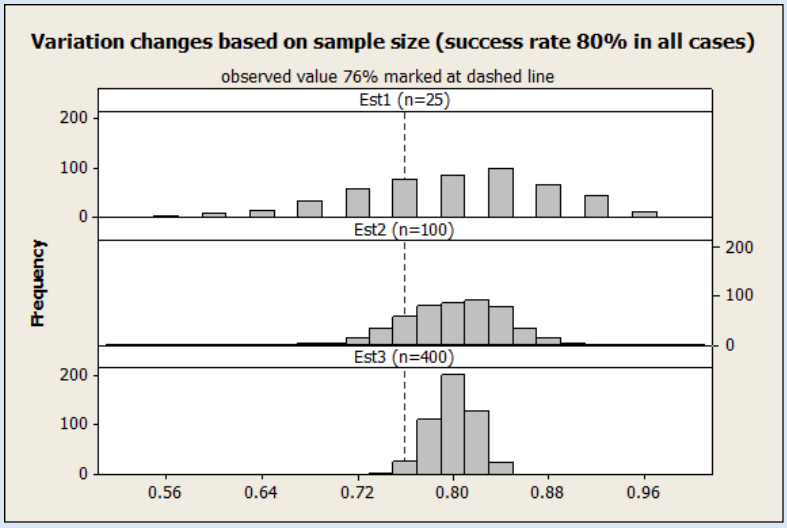


c. While it is true that 76% is lower than the recent average of 80%, the value 76% needs to interpreted in context, for example in a run chart, and also in terms of the number of surveys that drive the calculation (see item e). In the absence of specific tools and consistent management practices and structure to drive continuous improvement, an exhortation is unlikely to have a positive impact on performance.

d. Similar to item c, this action is likely to be a waste of organizational energy if the focus is exclusively on the most recent value of 76%.

Explanation (continued):

e. Your interpretation should recall basic statistical ideas. Data like HCAHPS survey results often can be approximated by random sampling of data from a bowl of beads with two colors (e.g. green for "top box" and red for all other responses). The variation in the observed percent of green beads in repeated random sampling from the bowl depends on both the proportion of green beads in the bowl and the sample size. For example, if the true proportion of green beads in a bowl is 80% here's a picture of the variation of the sampled proportion of green beads. We generated 500 samples of size 25, 100, and 400:



For samples of size 25 (say, the survey results for one department or unit), the variation is much greater than for the results of an entire hospital (n = 400) even if the underlying success rate is the same. For sample size n=25, observing 76% is not surprising; when sample size is 400, you might start to wonder if the system has changed if the new month's value is 76%.

The perspective illustrated by the figure above underlies the mathematics of control charts, which show calculated limits on top of a plot of values in sequence order.

Even if you have not yet learned about control charts or you don't regularly use them, it pays to keep in mind the sample size effect when interpreting month-to-month changes in proportions.

For those people who like formulas, the variability in sampling for a proportion like HCAHPS data, expressed in terms of "standard deviation," is given by

$$\sqrt{p(1-p)/n}$$

where *p* is the proportion of success and *n* is the sample size. For example, if *p*=0.8 and *n*=100, the standard deviation is

$$\sqrt{p(1-p)/n} = .4/10 = .04$$

Explanation (continued):

There is a table on the HHS site that makes the same point.

http://www.hospitalcompare.hhs.gov/staticpages/for-professionals/poc/confidence-interval.aspx (accessed 6 February 2012).

Confidence intervals can be used to estimate the precision of the calculated rates for an individual hospital. A confidence interval is the range of values, within which an estimated value or rate is likely to fall. A confidence interval is a statistical determination of the degree of certainty associated with an estimated value. As can be seen in the table of estimated values (below), large differences between individual hospitals' rates may be significant, and small differences between hospitals are usually not significant.

The smaller the sample size, the greater the difference in rates must be order for that difference to be statistically meaningful. Also, as sample size varies between hospitals, it is difficult to precisely compare their rates, without considering the confidence intervals.

Over time, as the quality data base is expanded, a full four quarters of data will be posted for each measure. The number of cases used to determine hospitals' rates will likely increase, thereby increasing the reliability and stability of the rates.

Estimating Confidence Intervals for the Process of Care Measures: Estimated Values for Proportion Data

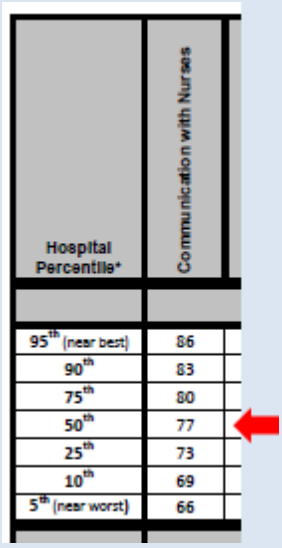
SAMPLE SIZE	OBSERVED RATE								
	10%	20%	30%	40%	50%	60%	70%	80%	90%
<25	--	--	24.9%	26.6%	27.2%	26.6%	24.9%	--	--
25-75	8.3%	11.1%	12.7%	13.6%	13.9%	13.6%	12.7%	11.1%	8.3%
76-125	5.9%	7.8%	9.0%	9.6%	9.8%	9.6%	9.0%	7.8%	5.9%
126-175	4.8%	6.4%	7.3%	7.8%	8.0%	7.8%	7.3%	6.4%	4.8%
176-225	4.2%	5.5%	6.4%	6.8%	6.9%	6.8%	6.4%	5.5%	4.2%
226-275	3.7%	5.0%	5.7%	6.1%	6.2%	6.1%	5.7%	5.0%	3.7%
276+	2.9%	3.9%	4.5%	4.8%	4.9%	4.8%	4.5%	3.9%	2.9%

Source: CMS/OCSQ/QIG: The values in the table are the approximate amount to add and subtract from the observed rate to estimate a 95 percent confidence interval for the given sample size. (Interpolation between the values in the table is appropriate.) Estimates of an interval in these cells exceed the natural limits for proportions.

5. Interpreting HCAHPS Percentile Tables

Answers:

a. T
b. T
c. F or ?
d. ?
e. T
f. T



Hospital Percentile*	Communication with Nurses
95th (near best)	86
90th	83
75th	80
50th	77
25th	73
10th	69
5th (near worst)	66

Explanation:

a. **True.** This follows from the definition of percentile: 10% (about 383 hospitals) of the 3837 surveyed hospitals scored at or above 83 and 5% (about 162 hospitals) scored at or above 86. So at least 150 hospitals had a score at or above your hospital's score of 83.

b. **True.** The top box score is the percent of patients who respond to the "top box" options on the elements of the survey contributing to the composite. A score of 83 means 83% of respondents checked "top box" for the composite, so 100%-83% = 17% were not at top box.

c. **False or ?** The national percentile table is based on patients discharged between April 2010 and March 2011. Your most recent monthly data will be based on patients discharged after the survey period so your monthly result is not directly comparable to the national table. Hence our answer **False.** However, you may think this point is minor, that the national numbers don't change that much over time. Here's a table of three years of the percentile values for top box *Recommend the Hospital*:

	percentiles						
Year	5th	10th	25th	50th	75th	90th	95th
Apr 08-Mar 09	50	-	61	68	75	-	85
Apr 09-Mar 10	52	56	63	69	76	82	85
Apr 10-Mar 11	53	57	63	70	76	82	85

Unless the current year is very different than the past three years, we agree that the 75th percentile is likely to be lower than 78 based on 12 months of data. So, can you conclude that your hospital is better than 75% of the hospitals? Not quite. Remember, you have to take into account variation in the monthly numbers (refer back to the explanation attached to question **4e**. Unless you know the sample size, and better yet, the history of your top box numbers plotted on a run chart, you can't quite make the leap stated in the question. Hence our answer: **?**

Explanation (continued):

d. ? You need additional information. At either the bottom of the percentile scale or the top of the percentile scale, a difference of 6 points does NOT translate to a percentile difference of 15%. In the middle of the scale, a difference of 6 points translates to a percentile difference of 25% or more.

Hospital Percentile*	Overall Hospital Rating
95th (near best)	83
90th	79
75th	73
50th	68
25th	62
10th	57
5th (near worst)	52

Implication: You need to keep top box scores distinct in your mind from percentiles of top box scores. The percentile scale is "non-linear" and so can mess up simple intuition and arithmetic.

The point about sampling variation raised in the explanations for **4e** and **5c** also applies, although a year's worth of survey data will typically lead to variability in your score estimates that is less than one third of the variability in a single month's score (as measured by standard deviation).

e. True. Survey professionals take care to account for the effect of all the people who do not respond but still it is sobering to realize that only about 1 in 3 patients make the effort to respond to the HCAHPS survey. This level of non-response is one more reason to think about multiple ways to gather timely and meaningful patient experience data.

f. True. You have greater sensitivity when you stratify HCAHPS scores and percentiles by department type, e.g. comparing the OB department at your hospital to a reference set of OB departments. The overall scores mask relatively strong performance and obscure opportunities for improvement.

A Note About HCAHPS "Boxes"

HCAHPS results are publicly reported on Hospital Compare as "top-box," "bottom-box" and "middle-box" scores. The **"top-box"** is the most positive response to HCAHPS survey questions. The "top-box" response is *"Always"* for five HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, Pain Management, and Communication about Medicines) and two individual items (Cleanliness of Hospital Environment and Quietness of Hospital Environment), *"Yes"* for the sixth composite, Discharge Information, *"'9' or '10' (high)"* for the Overall Hospital Rating item, and *"Would definitely recommend"* for the Recommend the Hospital item.

The **"bottom-box"** is the least positive response category for HCAHPS measures. The "bottom-box" response is *"Sometimes or never"* for five HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, Pain Management, and Communication about Medicines) and two individual items (Cleanliness of Hospital Environment and Quietness of Hospital Environment), *"No"* for the sixth composite, Discharge Information, *"'6' or lower (low)"* for the Overall Hospital Rating item, and *"Would not recommend"* for the Recommend the Hospital item.

The **"middle-box"** captures intermediate responses to HCAHPS survey items. The "middle-box" response is "Usually" for five HCAHPS composites (Communication with Nurses, Communication with Doctors, Responsiveness of Hospital Staff, Pain Management, and Communication about Medicines) and two individual items (Cleanliness of Hospital Environment and Quietness of Hospital Environment), *"'7' or '8' (medium)"* for the Overall Hospital Rating item, and *"Would probably recommend"* for the Recommend the Hospital item. There is no "middle-box" response in the Discharge Information composite.

http://www.hcahpsonline.org/SummaryAnalyses.aspx

6. Correlations: Friend or Foe?

Answers:

a. T

b. ?

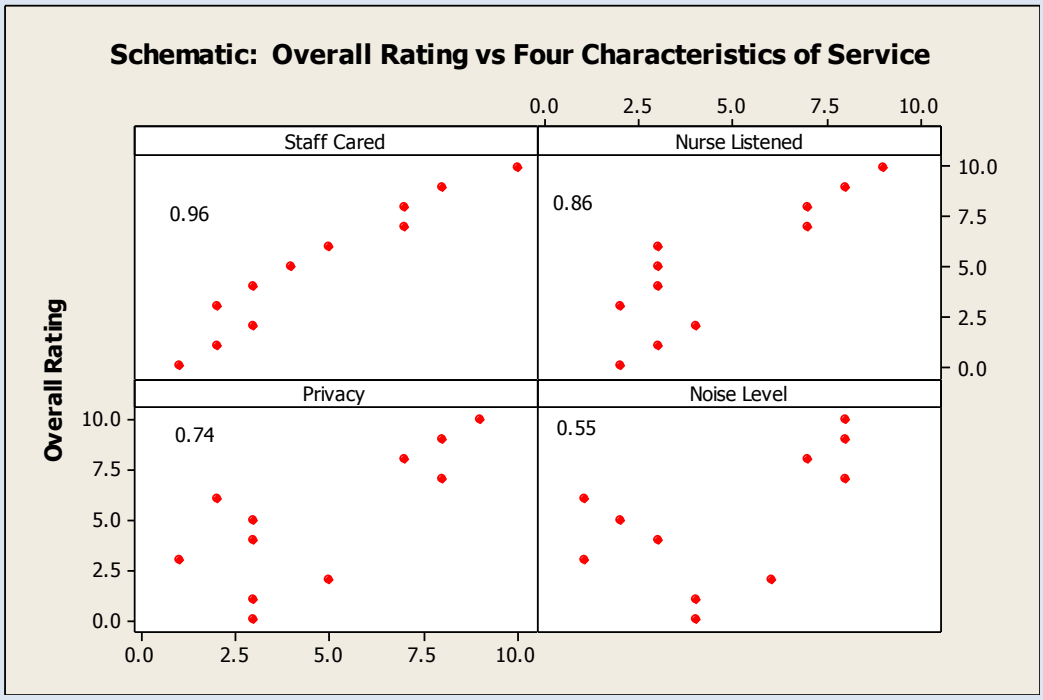
c. T

d. F/?

Explanation:

a. **True.** Remember that a correlation, based on either scores or ranks, measures strength of association and ranges from 1 (perfect positive linear or rank order relationship) to 0 (no linear or rank relationship) to -1 (perfect negative linear or reverse rank order relationship.)

Here's a picture that shows some invented data, with the correlation coefficient ranging from 0.96 to 0.55

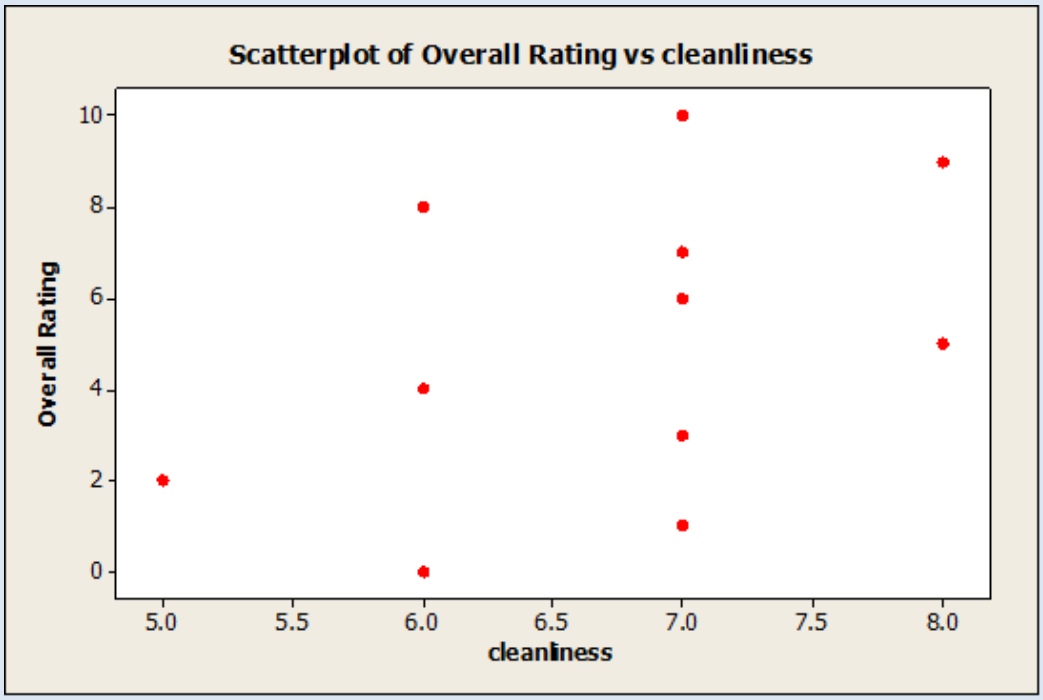


b. **?** There's an extra step and a logical leap you need to take before you conclude that the factors with the highest correlation have the most impact. Extra step: What do the data from your hospital indicate? (factor relationships in your hospital may differ from the national score.) Logical Leap: correlations show association, not proof of causation. You need to apply your knowledge of behavior and management when interpreting these correlations. To have a data-driven foundation for cause-and-effect, we need to look at appropriately designed experiments.

Explanation (continued):

c. **True or ?.** The correlation ranking suggests that the staff including patients in treatment (clustered with the other factors that look at staff interactions with patients and families regarding inconvenience and emotional needs) is a better use of limited resources than a focus on noise reduction. You might answer "**?**" if you are not ready to make the logical leap from association to causation. Still, the ranking of factors has face validity with our faculty experts and provides guidance for testing changes.

d. **False or ?** While cleanliness has the lowest correlation, you can't conclude necessarily that the factor has no association. A low correlation might mean that cleanliness ratings are clustered tightly together (every hospital is ranked somewhere between a 5 and a 7 on the 11 point scale, for example). Here's a plot with a correlation of 0.43:



It is possible that if cleanliness slipped (as perceived by patients and their families, who may have increasing expectations), it would negatively affect the overall rating.

WARNING: Correlations obscure time order! Don't forget to anchor your thinking with plots that show values plotted in time order.