Notes:

Stratification

Key Learning Points

1. Describe the importance of Stratification.
2. Explain how to Stratify Data.
3. Utilize Stratification in improvement projects.

What is Stratification?

Stratification is the separation of data. It's most frequent use is during analysis of theories to identify which theories contribute to the problem being solved.

Stratification is the basis for other tools like Pareto Analysis, and it is used in conjunction with other tools such as Scatter Diagrams to make them more powerful.

Stratification Variable

The characteristic used to separate data is called a stratification variable. Each stratification variable has two or more values.

Categories

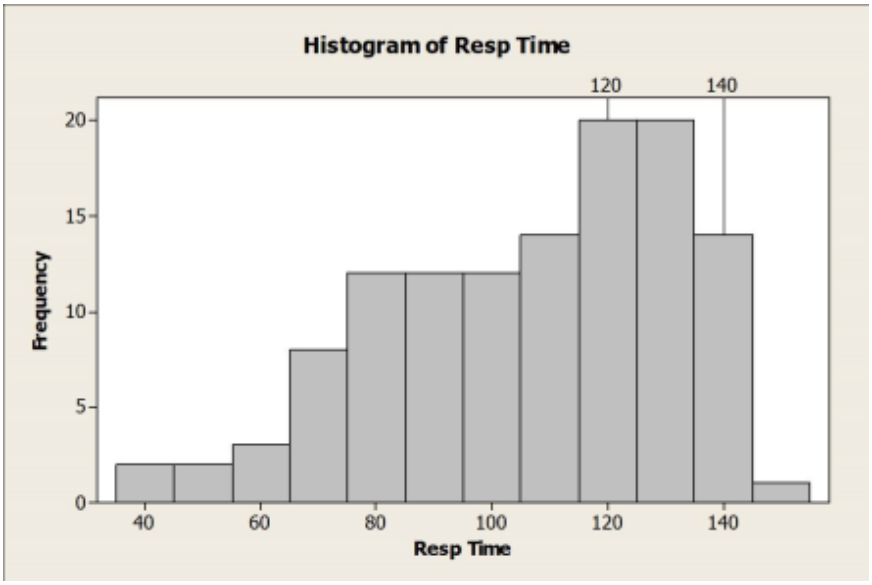
A category is defined by specific combinations of ranges of variables.

Points Out

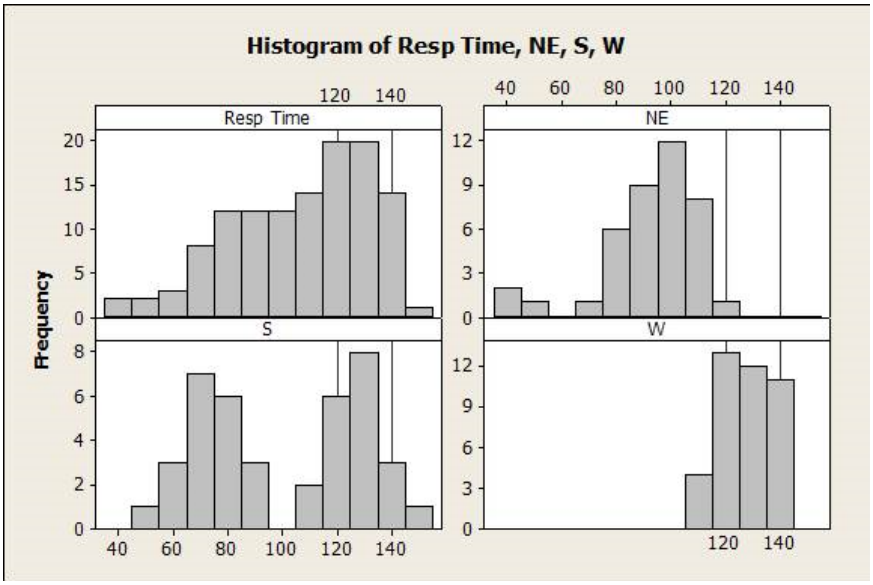
Stratification alone will not normally point out exactly what the root causes of a problem are. It will, however, point out which areas need further investigation in

order to have a full understanding of how the problem originated.

Example Stratification



Stratified Data



Steps in Stratifying

1. Select the stratification variables. If new data are to be collected, be certain that all potential stratification variables are collected as identifiers.
2. Establish categories that are to be used for each stratification variable. The categories may be either discrete values or ranges of values.
3. Sort observations into the categories of one of the stratification variables. Each category will have a list of the observations that belong to it.
4. Calculate the phenomenon being measured for each category. These

Notes:

calculations are usually either a count of the number of observations in the category or a calculation of an average for those observations. Sometimes stratification is applied to an entire tool, such as a histogram or scatter diagram.

5. Display the results. Bar graphs are usually the most effective, unless stratified histograms, or scatter diagrams are involved.
6. Prepare and display the results for other stratification variables. Repeat Steps 2 through 5. Do second-stage stratification as appropriate.
7. Plan for additional confirmation. Usually additional data, other techniques, and/or controlled experiments will be used to confirm the initial stratification results.

Interpretation

Look across the categories of a variable to see whether one or more of the categories stand out. If results appear to indicate the likely source of the phenomenon being studied, validate results by gaining further details on the underlying cause.
If initial stratification does not yield useful results, there are two possible courses of action:

- Conduct a two-stage stratification (stratify by a 2nd variable within each category of a 1st variable).
- Stratify by other variables.

When Should Stratification Be Used?

- In Measure through the creation of Pareto Diagrams
- In Analyze when testing theories

Pitfalls to Avoid

The most likely pitfall is to conclude too much from the data. Small differences among classes should not be given undue weight. If the team is looking at a major quality problem, they should not expect that the causes will be found in small differences.

Another mistake is to jump to the conclusion that the anomalous category is the cause of the problem. The anomalous category is probably where you should go looking for the cause, but the category itself is not necessarily the cause.

If the team is going to collect new data, they should make an effort to collect as much identifying information as they think could possibly prove useful for stratification. (Use the data-collection plan discussed earlier.)

Notes: