DATA ENTRY

The Procedure

The Project analyst

-Set up the Performance Analyzer with the indicators, and "compliant answers" and "aggregate groups" for each indicator
-Reviewed checklists
-Identified the checklists for "applicable facilities"
-Reviewed checklists for completeness and consistency, with particular attention on the use of NA or blanks
-Contacted states with any questions
-Scanned checklists
-Entered data into the Performance Analyzer

The Quality Assurance Check of Data

The Project Analyst:
-Checked tables in Performance Analyzer containing checklist answers against the scanned checklists and made necessary corrections

The Quality Assurance Officer
-Checked the Performance Analyzer set up

Quality Assurance Check of Performance Analyzer

The Project Analyst and Project Quality Assurance Officer:
-Ran charts and tables and statistical outputs to see if they
- Worked
- Were stable (provided the same answers repeatedly)
- Produced internally consistent results (When different tables and charts had the some of the same data elements, the values of those data elements were the same in the two different tables or charts
- Produced accurate results
- tables and charts t same data elements similar data elements containing the same data the data tables
 -Identified problems
-Worked with the contractor to fix identified problems

PERFORMING THE STATISTICAL ANALYSIS OF THE DATA

The Procedure:
The Performance Analyzer is generates export files for statistical analysis. Two modules were used for this project:

Module B2 Export - Between "Rounds" - Inspectors which arrays the data as needed by the JMP program to calculate any two states' SQG mean facility scores and achievement rates and the associated confidence intervals. A report was run for each possible state to state comparison as follows:

R1 v R2	R1 v R2	R1 v R2
MA v CT	ME v NH	VT v CO
MA v ME	ME v NY	VT v RI
MA v NH	ME v VT	CO v RI
MA v NY	ME v CO	
MA v VT	ME v RI	
MA v CO	NH v NY	
MA v RI	NH v VT	
CT v ME	NH v CO	
CT v NH	NH v RI	
CT v NY	NY v VT	
CT v VT	NY v CO	
CT v CO	NY v RI	
CT v RI		

Module A Export – Inspectors which arrays the data as needed by the JMP to calculate a state's SQG achievement rates and associated confidence intervals for each indicator. This report was run for each state.

Separate Folders were created for each state's Module A and B statistical analyses, into which the "raw" Module A and B exports from the Performance Analyzer were placed. Module B Reports which compare two states' results were put into a subfolder indicating which two states were being compared in the folder corresponding to the state assigned as R1.

The Performance Analyzer assigns the exports a name that indicates which of the two statistical modules was being run as follows:

Data ERP_ModuleB2_Y_Inspector.xls
Data ERP_ModuleA_Inspector.xls

Note that the name does NOT indicate which two states were being compared. In order to ensure that spreadsheets for different state pairings were not confused with each other   "_ModuleB2_Y_Inspector" on each Module B report was replaced with the names of the R1 and R2 states, using the convention in the following example:

Data ERPyr1-MA vs yr2-RI.xls.

A statistical script that runs on JMP software was developed to analyze the excel spreadsheets exported by the Performance Analyzer.   he appropriate DataERP_____.xls report was copied to the desktop of the computer on which the JMP software was installed. Because the JMP analytical software requires the input file to be named Data ERP.xls, the extra characters were deleted.

The JMP script was then run at the 85%, 90%, 95% confidence levels.
The Wilson Score was used for the Module A analysis .

The Module A JMP output for each state and the Module B JMP output for each state to state comparison at each confidence level was copied and pasted into a "Master Data Source" workbook This workbook included formulae that automatically translated the JMP expression of a statistically significant difference (YR2 != YR1) into columns that displayed which of the two states in state to state comparison had a higher score, and which comparisons were statistically significant at the 85%, ,90%, and 95% confidence levels. This Excel workbook became the "Master Data Source" for all of the Statistical Project Results Charts and Tables.

The Data ERP.xls and JMP outputs were then moved back into the appropriate folder. For Module B reports the appendage indicating which states were being compared was added back to the Data ERP.xls chart and to each of the JMP outputs.

The Quality Assurance Check:
1) Each state's results were included on seven different Module B "Data ERP.xls" reports generated by the Performance Analyzer. The "Master Data Source" workbook was checked to make sure that the JMP analytical results for the achievement rates and number of facilities for each indicator, and the mean facility scores and number of facilities for each aggregate group for any given state were the same regardless of which state it was paired with, and on both the "Module B" and "Module A" reports. These comparisons ensured that the same raw data for a state was used each time the state was compared to another state. It also indicated that the raw data used for the Module A reports was the same as the raw data used for the Module B reports

   Note however, because Indicator 6 did not apply in NH and RI, it was excluded from the mean score calculations for all interstate pairings involving NH and RI. As a result the mean facility scores on EBPI's, Regulatory, and All indicators were different when states were compared with NH and RI than when they were compared with other states. Beyond Compliance scores were not affected because Indicator 6 is a regulatory indicator, and as such is not included in the Beyond Compliance Indicator mean facility score calculations

2) One Module B export that included each state was generated from the performance analyzer, and the four resulting DataERP.xls worksheets were rerun with the JMP analytical software. The newly calculated achievement rates, mean facility scores, and numbers of facilities for each state were then compared to the results for that state on the "Master Data Source". Matching results indicated that the "right" raw data was used for each Module B state to state comparison and for the Module A results for each state on the "Master Data Source". It is possible to draw this conclusion, because Step 1 showed that for each state, the same raw data was used each time that state was compared to any other state, and for the Module A report for that state.

CREATION OF THE PROJECT RESULTS STATISTICAL CHARTS AND TABLES

The Procedure:

Data from the "Master Data Source" workbook was copied and pasted into the tables in the "Project Statistical Results Workbooks". Excel' "insert charts" feature was used to create charts.

Quality Assurance Check:
The Project Analyst, Quality Assurance Officer, and an EPA project member checked

1) The data in the "Project Statistical Results Workbooks" were checked against the "Master Data Source" to make sure the data had been copied and pasted from the "Master Data Source properly.

2) The data shown in the charts in "Project Statistical Results Workbooks" were compared to the "Master Data Source" to make sure that the charts were drawn from the correct data in the "Project Statistical Results Workbook" tables.

CREATION OF THE CHARTS SHOWING THE DISTRIBUTION OF FACILITY SCORES IN A STATE

The Procedure:
The Performance Analyzer calculates the distribution of facility scores for any pair of states. Each state was included in a comparison, and the resulting distribution for each aggregate group of indicators for each state was copied and pasted into a "Facility Score Distribution" workbook, and the data was used to create the charts.

The Quality Assurance Check:
The distributions were rerun for each state, using a different pairing of states and resulting tables were compared to those in the "Facility Score Distribution" workbook.

DOCUMENT VERSION CONTROL
The creation date was created was included in the name of each chart.. When the data in the chart was modified (for whatever reason) or when the format was substantially changed. it was saved under a new name with the latest date. The prior version was then archived, and the most recently modified version was left in the "active" folder.