USING SAS/QC®TO AUTOMATE THE "MAGNIFICENT SEVEN" TOOLS OF QUALITY

Rod Lashley
STATKING Consulting, Inc.
Fairfield, Ohio

The goal of statistical process control (SPC) is to improve process performance through the reduction of variability. There are seven graphical tools to help the quality practitioner analyze his/her process in order to see how the process may be improved. They are check sheets, Pareto charts, cause and effect (Ishikawa) diagrams, defect concentration diagrams, scatter diagrams, control charts and histograms.

By implementing these seven tools, process improvement teams can get immediate or on-line feedback of information about the production process. All seven tools of quality can be automated using SAS®, SAS/GRAPH®, and SAS/QC® software. Thorough descriptions of these tools can be found in Montgomery(1991) and Burr(1993). This paper will show examples of how SAS software can be used to produce the seven tools of process improvement.

Check Sheets

The first of the seven quality improvement tools we will look at is the *check sheet*. The check sheet is a tabular summary of defect information collected on a process over a specified period of time.

The check sheet takes the form of a two way table with the row entries giving the defect types and the column entries giving the time periods during which these defects were observed. The time period may be days, weeks, shifts, hours, etc. Other pertinent information about the process is included either at the top or bottom of the chart. These items may include an identification of the process or product under study, the operator (if machines are involved) and the QC inspector.

We can easily use the TABULATE procedure in base SAS software to summarize a set of defect data. By looking at defect data over time, the check sheet may identify trends in the type and time of occurrence of certain defects.

We will use data from Benjamin and Shaw (1993) as an example of the use of check sheets from the medical care industry. A medical group submits patient insurance claim forms to a clearinghouse that checks the forms for validity and then sends the forms on to the proper insurance carrier. The medical group is then reimbursed from the insurance carrier for the covered services. Forms that are filled out incorrectly are returned to the medical group for correction. This slows the reimbursement process.

A SAS® data set named CLAIMS is created to house the data. The data and the data step code used to create the CLAIMS data set is shown in the Appendix. Each observation in the data set represents a returned form. The variable REASON gives the field on the form that was incorrectly entered. The date the defect was recorded is given in the DATE variable. The LOCATION variable gives the branch location.

The TABULATE procedure constructs a check sheet with the reasons the forms are returned as the rows and the day the form was returned as the columns. The call to PROC TABULATE and the check sheet produced is shown in the Appendix.

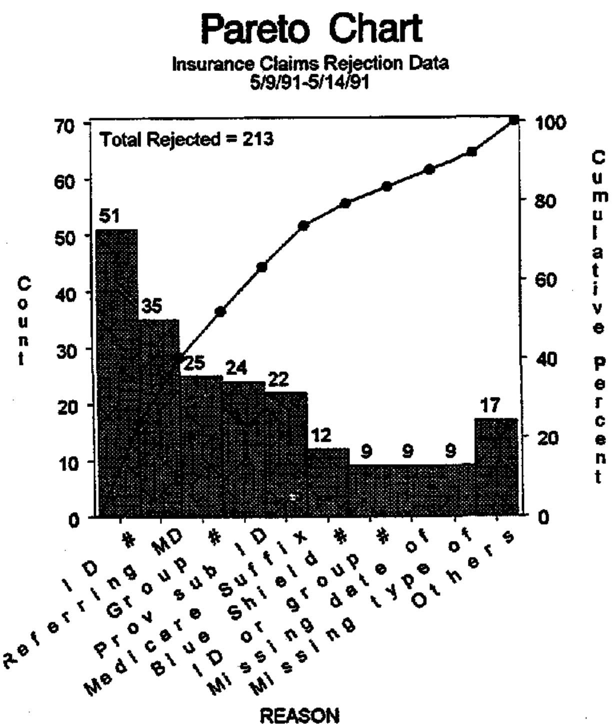
From the chart, it is readily apparent that an incorrect patient identification number is the most commonly occurring problem. We can also see that the number of returned forms peaks on the thirteenth of the month.

Pareto Chart

A graphical method that is commonly used in conjunction with check sheets is the *Pareto chart*. The Pareto chart is a bar chart of the number of defects observed on a process over a specified time period. Each bar of the chart represents a defect type. The height of the bar gives the number of times the defect occurred during the study period. The bars are arranged from left to right in descending order of frequency giving a quick visual indication of the most commonly occurring defects. In many cases, a second scale is added to the right of the chart to display the cumulative percentage of defects across the defect classes.

PROC PARETO, available in the SAS/QC® software, constructs Pareto diagrams of defect data. The procedure contains many options to enhance the graphical appearance of the chart. For a summary of all the options available with PROC PARETO, see *SAS® QC Software: Usage and Reference, Version 6, First Edition.*

Using PROC PARETO, a Pareto chart for returned insurance forms can be constructed. A complete listing of the call to PROC PARETO is given in the Appendix. We can see from the chart that incorrect entry of the ID number and referring doctor account for over 40% of the returned forms.



Pareto Chart for Insurance Claims Data

Ishikawa Diagrams

Cause and effect diagrams, sometimes called *Ishikawa diagrams,* are a graphical method for displaying the root causes of defects in finished products. Once a defect has been identified as most harmful to product performance, an Ishikawa diagram of the evolution of the defect can be constructed.

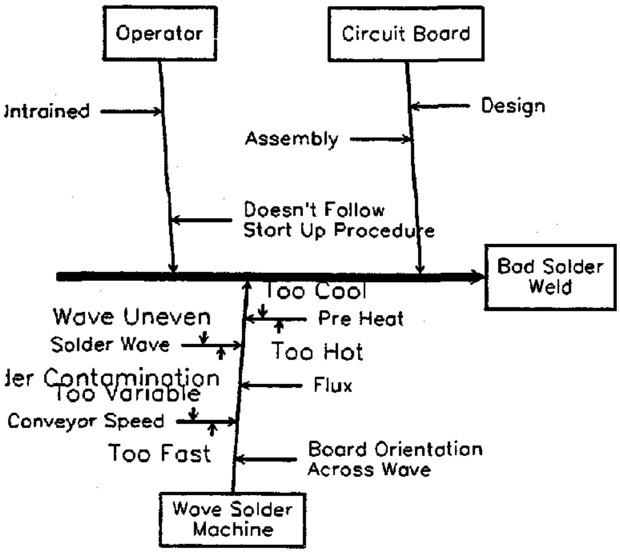
Ishikawa diagrams are most often used to summarize information and ideas expressed at a meeting of those most knowledgeable of the process. With the ISHIKAWA software, you can construct the Ishikawa diagram through a series of point and click instructions.

Ishikawa diagrams can be constructed using the SQC menu system. From the main menu of the SQC Menu System, click on the ISHIKAWA icon. By utilizing the point and click functionality of the ISHIKAWA environment, the team can interactively build a cause and effect diagram for the bad solder weld problem. This allows their thoughts concerning the causes of the problem to be recorded and immediately diagrammed. They are also able to save the diagram so that it can be updated and refined as more input is obtained about the problem. More information about the finer points of using the ISHIKAWA environment are given in *SAS/QC® Software: Language and Reference, Version 6, First Edition.* The following example will illustrate the use of the Ishikawa diagram.

The manufacturer of electronic circuit boards for airplane guidance systems involves the insertion of a number of components called dual inline processors (DIPS) into the circuit board. A series of prongs, called leads, extend from either side of the DIPS . The DIPS are placed on the board by a robot arm and the leads are crimped from the underside of the board to hold the DIPS in place. The board then passes over a device called a wave solder in order that the components be soldered in place.

Because the wave solder involves a combination of sensitive settings, many solder welds on each board are not soldered properly and are considered defective. The quality engineering personnel wish to improve the wave solder process. Using the ISHIKAWA software, the bad solder weld defect problem for the circuit board assembly process can be diagrammed by the process improvement team.

The Ishikawa diagram resulting from the process improvement team's work is shown below.



Ishikawa Diagram for Wave Solder Process

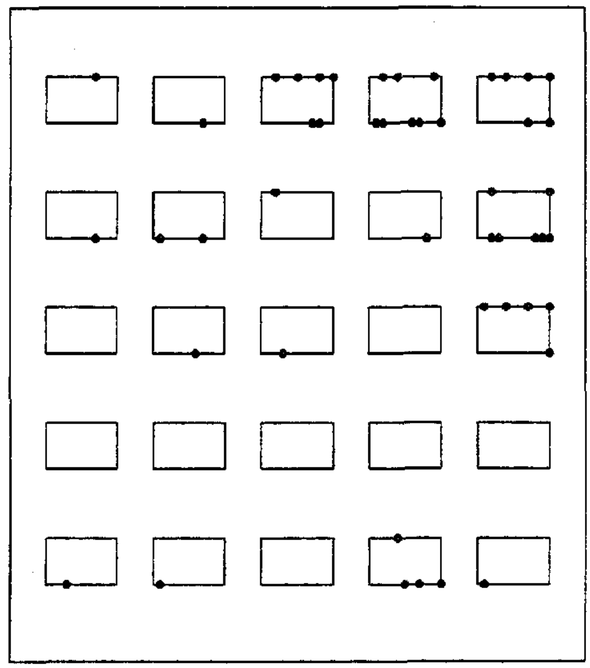
Defect Concentration Diagrams

Defect concentration diagrams give a picture of the item being manufactured and show the position of relevant defects for a series of product units. By plotting the defect positions from enough of the units, certain patterns or trends may emerge that lead to the discovery of why these defects occur.

Referring to our electronic circuit board example, another common problem associated with the boards is bent leads. The dips are inserted into the circuit board by robot arms. If the alignment of the board and the dips are such that the leads do not enter the holes in the board correctly the leads are bent and must be fixed manually.

The SAS/GRAPH® component of the SAS® system can be used to draw a picture of a circuit board and mark the location of bent leads in order to form a defect concentration diagram for this process. In order to draw figures using SAS/GRAPH, a data step consisting of a specific set of variables describing the drawing coordinates must be created. This data set is then used by the ANNOTATE facility in the SAS/GRAPH® software to draw the figure. A detailed description of these variables and the ANNOTATE facility is given in *SAS/GRAPH® Software: Reference, Version 6, First Edition, Volume 1.*

Defect Concentration Diagram for Circuit Board
Board Style VX025B7 --- July 7–July 15



Defect Concentration Diagram

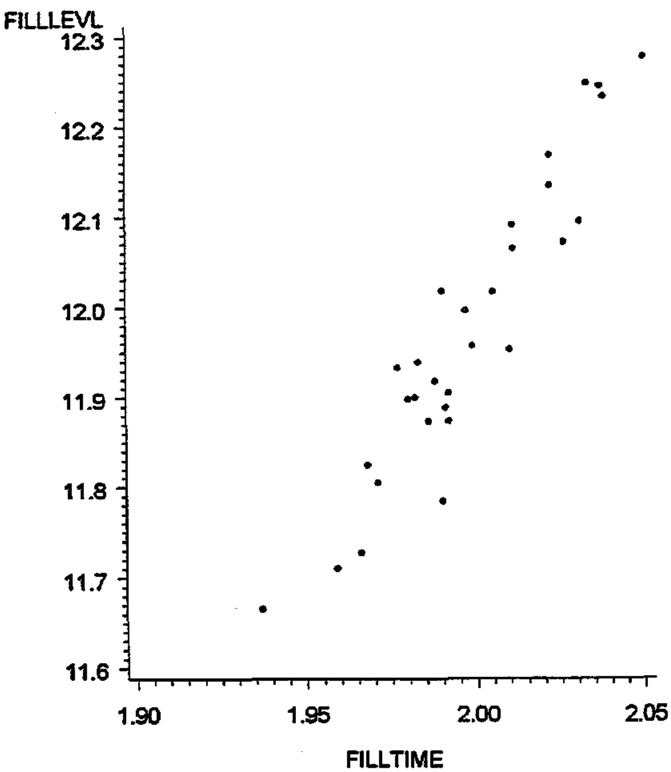
Scatter Diagrams

The *scatter diagram* displays relationships between pairs of process variables. A quality characteristic measured on finished goods may be a function of one or more variables affecting the process. A first step towards identifying these relationships is to record the values of the variables of interest for a number of units of finished product and plot each pair of variables on an x-y coordinate system. The scatter or shape of the points gives an indication of the relation of the two variables. The following example shows an example application of scatter diagrams.

The bottlers of a particular soft drink product wish to study the relationship of the fill level in their 12 ounce soft drink cans to the length of time the fill nozzle is open above the can and the pressure maintained in the filling bowl above a set of nozzles. They record the fill level in ounces, the time in seconds that the nozzle is open above the can and the fill pressure in pounds per square inch for 30 cans of product.

PROC GPLOT in the SAS/GRAPH software allows you to construct and display scatter diagrams. For the bottle fill data, stored in the SAS data set FILLS, FILLLEVL and FILLTIME can be displayed in a scatter diagram using the code in the Appendix.

Soft Drink Filling Data



Scatter Diagram for Soft Drink Filling Process

The scatter diagram shows that the fill level is positively correlated with the time the valve is open over the can.

Control Charts

The control chart is a plot of a statistic used to summarize a quality characteristic versus a sample number or time. Overlaid on the plot are the centerline and upper and lower control limits (UCL, LCL). The centerline of the control charts marks the center of the distribution of the quality characteristic. It may be a targeted (known) value for the process or it may be the estimated process distribution center. The UCL and LCL are chosen so that if a process is in control, nearly all sample points will lie between them.

A process is said to be operating in an in-control state if it is operating only under chance causes of variability. When a shift occurs in the process due to the presence of assignable causes then the process is said to be "out-of-control". When the assignable cause is identified and eliminated, the process resumes operating in an in-control state.

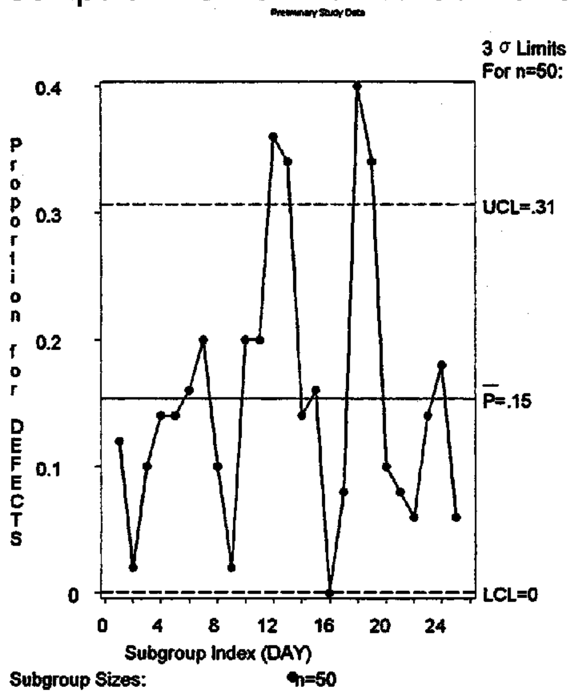
Control charts fall into two general categories, *variables control charts* and *attribute control charts*. Variables control charts are constructed for quality characteristics measured on a continuous scale. Attribute control charts are constructed for quality characteristics which are of a conforming/non conforming nature or counts of nonconformities. Control charts to monitor both attribute and continuous variables can be created using the SAS/QC® procedure SHEWHART. Consider the following scenario for the construction of a control chart.

The XYZ Corporation manufactures cabinets for computer monitors. The frame that holds the computer screen in place is made of molded plastic. As each batch of frames is produced, a portion of the frames are inspected for proper size and uniformity of surface. Some of the frames must be rejected. To monitor the proportion of non conforming frames over a period of time, we can construct and maintain a P chart for the molding process. The P chart is an attribute type chart. The plotting statistic for the P chart is $\hat{p}$, an estimate of the proportion of non conforming items in the population of interest.

Since the current nonconformity levels are unknown, the first step is to conduct a preliminary study of the process in order to establish current nonconformity levels and control limits. We will sample 50 frames from each day's production and record the number of nonconforming frames for 25 days of production. The results of the preliminary study are contained in the SAS® data set FRAMEPRE, shown in the Appendix.

The control chart produced from the preliminary study data is shown below. Note that on four days the process was out of control.

Computer Monitor Frames Control Chart

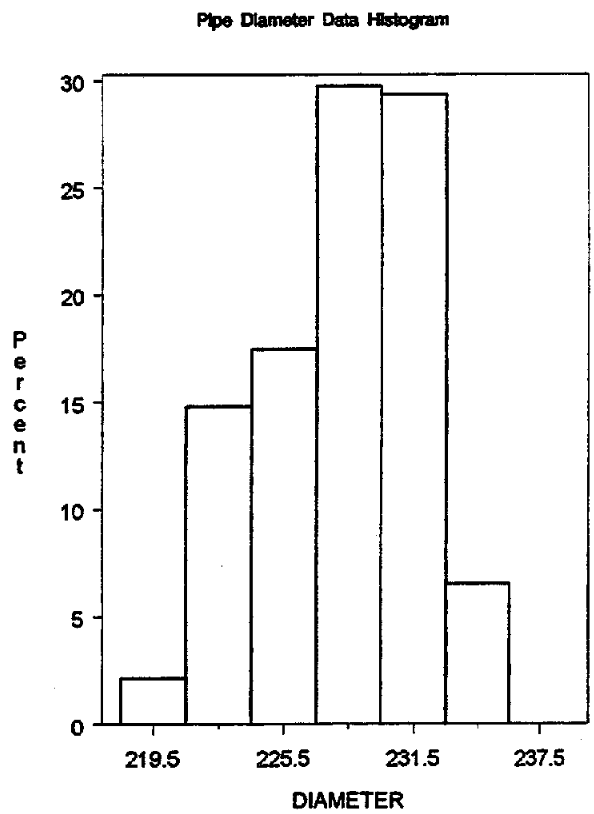


Histogram

The histogram is the seventh and final member of the "magnificent seven" tools for quality improvement. The *histogram* is a bar chart of the observed proportion of samples falling in each of several ranges of values of the process characteristic. These ranges are sometimes called class intervals or bins. Using a sample of process data, the histogram gives a picture of the underlying distribution of the quality characteristic under study.

To illustrate the construction of the histogram, a data set containing the inside diameters of 750 drainage pipes used in road construction is shown in the Appendix. The pipes were sampled after being arranged in lots for transport. The SAS® data set containing this data is named DIAMS and is shown in the Appendix. The diameters are measured to the nearest tenth of a millimeter. The target diameter is 230 millimeters.

In order to construct a histogram for the pipe diameter data, the midpoints and widths of the class intervals of the histogram must be selected. One way to choose the midpoints and the interval widths, w (=2h), is to set the number of intervals (and therefore, the number of midpoints) at some arbitrary number between 5 and 20 and then compute the width of these intervals by finding the range of the data and dividing the range by the number of intervals. This is the procedure recommended in introductory statistics texts such as Bhattacharyya and Johnson (1977) and Ott (1977).



For the pipe diameter data, choose six as the number of intervals. A scan of the DIAMS data set reveals that the maximum diameter value shown in the data set is 236 millimeters while the minimum diameter is 218 millimeters so the width of the intervals, w=(236-218)/6=3 and h=1.5. These intervals can be specified on the HISTOGRAM statement with the MIDPOINTS option. With midpoints a distance of three units apart beginning at 219.5 and ending at 234.5, six intervals of width 3 are constructed. The code for this histogram is shown in the first HISTOGRAM statement in the code in the Appendix. The distribution of the diameters appears to be skewed left with a mean somewhere around 230 millimeters.

Summary

We have used examples from a variety of industries to show how SAS® and SAS/QC® can be used to produce the "magnificent seven" tools of quality. These methods are graphical in nature and provide immediate feedback of information about the production process.

Contact Information

Rod Lashley
Statistical Programmer
STATKING Consulting, Inc.
780 Nilles Rd. Suite E-2
Fairfield, Ohio 45014
Phone/Fax: (513) 858-2989
e-mail: 74242.3526@compuserve.com

References

Benjamin, M. and Shaw, J.G. (1993). "Harnessing the Power of the Pareto Principle", *Quality Progress*, September 1993, 103-107.

Bhattacharyya, G.K. and Johnson, R.A. (1977). *Statistical Methods and Concepts*, New Your, NY: John Wiley & Sons.

Burr, J.T. (1993). *SPC Tools for Everyone*, Milwaukee, WI: ASQC Quality Press.

Montgomery, D.C. (1991). *Introduction to Statistical Quality Control*, New York, NY: John Wiley & Sons, second edition.

Ott, L. (1977). *An Introduction to Statistical Methods and Data Analysis*, North Scituate, MA: Duxbury Press.

APPENDIX

A1. Partial listing of Insurance Claims Data Set.

```
OBS  REASON              DATE     LOCATION
 1   ID #               05/09/91 Downtown
 2   ID #               05/09/91 Downtown
 3   Group #            05/09/91 Downtown
 4   Medicare Suffix    05/09/91 Downtown
 5   ID #               05/09/91 Oakridge
              .
              .
              .
 210 Prov sub ID        05/14/91 Downtown
 211 Medicare Suffix    05/14/91 Downtown
 212 Group #            05/14/91 Downtown
 213 ID or group #      05/14/91 Berryessa
```

A2. SAS code for creation of Pareto Chart.

```
goptions hsize=6 vsize=7;

proc pareto data=claims graphics;
  vbar reason /
         scale=count   maxcmpct=95
         other='Others'    last='Others'
         angle=35   barlabel=value
         pbars=m5x45  anchor=bl
         nlegend='Total Rejected'
         axisfactor=1.3;
  symbol w=3 v=dot;
  title1 'Pareto Chart';
  title2 'Insurance Claims Rejection
             Data';
  title3 '5/9/91-5/14/91';
run;
```

A3. SAS code for drawing of Defect Concentration Diagrams

```
data board;
    length function style $ 8;
    xsys='5'; ysys='5';
    style='empty'; line=0;
/* Create Outline */
    function='move'; x=10; y=10; output;
/* of Circuit     */
    function='bar '; x=90; y=95; output;
/* Board          */

    do row=15,30,45,60,75;
/* Draw Five Rows*/
      do col=20,35,50,65,80;
/* of Five Dips  */
        function='move'; x=row; y=col;
output;
        function='bar '; x=row+10; y=col+6;
output;
      end;
    end;
run;

data leadpos;
    length function style color $ 8;
    xsys='5'; ysys='5';

    input x y lead @@;
    x=15*x+lead; y=15*y+5;
    if lead>10 then do; y=y+6; x=x-10;
    end;
    style='solid'; color='red';
    rotate=360;
    function='pie '; size=.2;
    output;
  cards;
1 1 3    3 5 7    4 5 14    5 5 12    3 4 12
2 1 1    3 5 15   3 5 20
4 5 7    4 4 8    4 5 12    3 5 8     2 1 1
5 4 9    4 1 7    4 5 6
2 5 7    5 4 8    3 5 12    3 5 18    2 4 1
5 4 12   4 1 5    4 5 2
5 5 7    5 4 3    4 5 10    4 5 2     2 3 6
5 4 2    4 1 14   4 5 19
5 5 17   5 5 14   5 5 10    5 5 20    2 4 7
5 5 17   4 1 10   4 5 1
5 3 17   5 3 14   5 3 10    5 3 20    1 4 7
1 5 17   5 1 1    5 3 11
5 4 10   5 4 20   3 3 3
run;

data circuit; set board leadpos; run;

proc gslide annotate=circuit;
  title1 f=swissx h=1 'Defect
Concentration Diagram for Circuit Board';
  title2 f=swissx h=1 'Board Style VX025B7
--- July 7-July 15';
run;
```

A4. SAS code for drawing Scatter Diagram.

```
proc gplot data=fills;
  plot filllevl*filltime;
  symbol v=dot h=.5;
run;
```

A5. SAS code for control chart.

```
proc shewhart data=framepre graphics;
  pchart defects*day / subgroupn=samples;
  symbol v=dot h=.8 i=join ;
  title 'Computer Monitor Frames Control
Chart';
  title2 h=.5 'Preliminary Study Data';
run;
```

A6. SAS code for histogram

```
proc capability data=diams noprint
graphics;
```

```
  histogram diameter / midpoints=219.5 to
234.5 by 3;
  title h=.8 'Pipe Diameter Data
Histogram';
run;
```

A7. SAS code to create check sheet for insurance claim data.

```
proc tabulate data=claims format=5.0 ;
  class reason day;
  label reason='Reason for Rejection';
  var form;
  table reason all='Total',day*(form='   '*sum='   ') all='Total' / rts=30 printmiss
misstext='  ';
  title1 'Check Sheet';
  title2 'Insurance Claim Rejection Data';
  title3 '5/9/91-5/14/91';
run;
```

Check Sheet
Insurance Claim Rejection Data
5/9/91-5/14/91

	DAY						Total
	9	10	11	12	13	14	N
Reason for Rejection							
Blue Shield #		2	3	2	4	1	12
CPT4/HCPCS		1	1	1	1		4
Group #	2	1	5	6	6	5	25
ID #	7	8	5	9	19	3	51
ID or group #			1	3	3	2	9
Invalid carrier zip	1	2	2		1	1	7
Invalid patient zip		1		1	1		3
Medicare Suffix	2	6	5	2	6	1	22
Missing date of accident		1	1	5	2		9
Missing type of service	1	1	4	2	1		9
Prov sub ID	1	6	6	2	8	1	24
Referring MD	1	6	9	11	6	2	35
Service not payable	2			1			3
Total	17	35	42	45	58	16	213