Paper 202

Three Methods to Produce Descriptive Statistics and Histogram Using SAS

Bing Deng, Technology Department of Baoshan Iron &. Steel Corporation, Shanghai China

ABSTRACT

SAS software is a great powerful tool of statistics and analysis. It is fun, easy to learn and use. This paper will present three ways to produce d**escriptive statistics** and **histogram** using **SAS/QC, SAS/INSIGHT, SAS/LAB** under win95. It will help the beginner to master the way quickly and finish their job easily.

INTRODUCTION

SAS software is used by people all over the world—in over 100 countries including China. The SAS System is a modular, integrated, hardware-independent system of software for enterprise-wide information delivery. It integrates many elements such as data access, data management, data analysis, data presentation into one powerful, flexible, and easy-to-use software system.

I am working at quality control for several years. SAS software becomes my best friend. It helps me complete many complex quality control jobs easily. I am grateful to SAS Institute for providing the great tool to us. In my routine work, I usually need to compute descriptive statistics (such as mean and standard deviation) and create a histogram to examine the distribution of some variables. So I have some experience on it .I hope that I have the honor to share it with the SAS user all over the world.

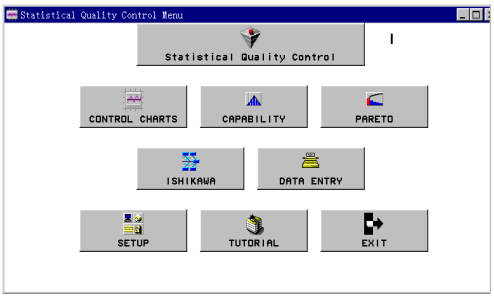
One Method: Using SAS/QC

SAS/QC software is a comprehensive system of statistical tools for quality improvement. You can use these tools to

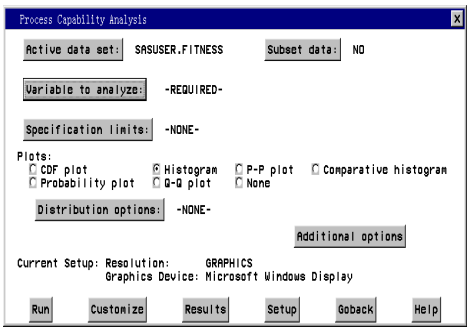
- plan and organize quality improvement efforts
- design industrial experiments for product and process improvement
- apply Taguchi methods for quality engineering
- establish and maintain statistical control of a process
- develop and evaluate acceptance sampling plans.

To get started, you must invoke the SQC menu system using either the SAS Display Manager System, SAS/ASSIST, or an operating system command issued by menu bar or by entering a command on the command line.

1. to invoke the SQC menu using the command line, type **sqc** on the command line and press ENTER. Then the Statistical Quality Control Menu window appears.



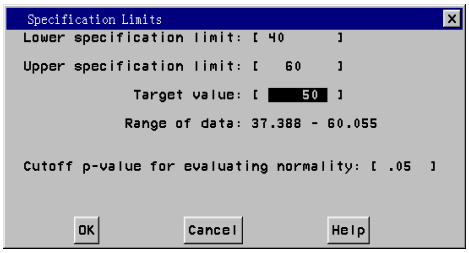
2. click on CAPABILITY in the Statistical Quality Control Menu window. The process Capability Analysis window appears.



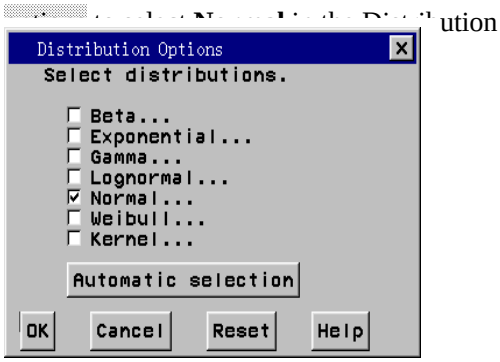
Then click on Active Data set to select the data set SASUSER.FITNESS which is distributed with SAS/ASSIST.

3. click on Variable to analyze then select the variable OXYGEN as the analysis variable.

4. Click on Specification limits in the Process Capability Analysis window. In the Specification Limits window type 40 as the lower specification limit and 60 as the upper specification limit. Then click on OK to confirm your selection and close the window.

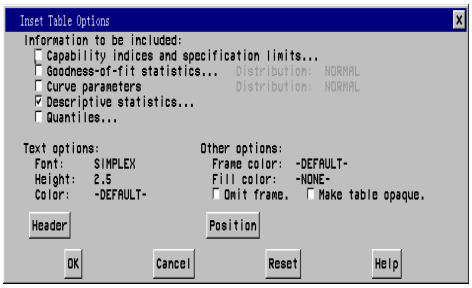


If you want to determine whether the data are normally distributed, then click on Distribution

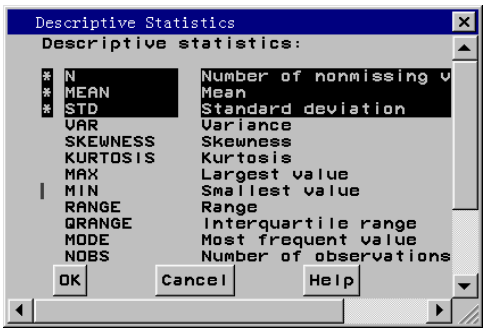


Options window and OK in the Normal Parameters window.

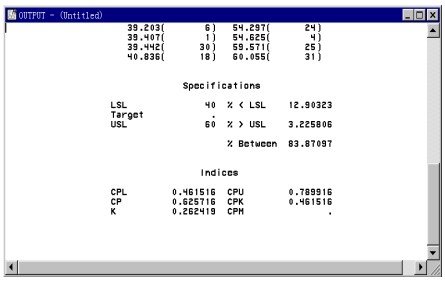
5. The menu system can inset one or two tables of statistics on the histogram. Click on Additional options in the Process Capability Analysis window and select Inset table options in the Additional options window, then click on First inset table. The Inset Table Options windows appeals.



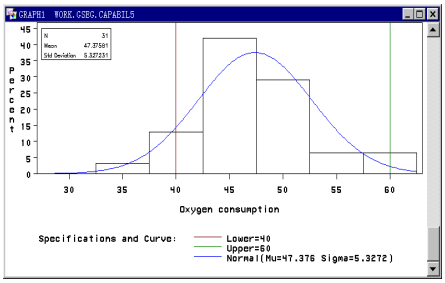
6. To add an inset table with summary statistics. Select **Descriptive statistics**.
   In the Descriptive statistics window you can select the item you want to display. Then select **N, MEAN STD** and OK in the Descriptive Statistics window.



7. Select Run in the Process Capability Analysis window.
   The statistical output for the capability analysis of OXYGEN. Moments and descriptive statistics for the fitness data appear first. The quantiles follow, along with a listing of the five lowest and five highest observations in the data. Finally, the menu system lists the percentage of data that fall above, below, and between the specification limits, together with capability indices.



The histogram allows you to visualize the distribution of the data, to examine if the data conform to specification limits and display the descriptive statistics you selected.

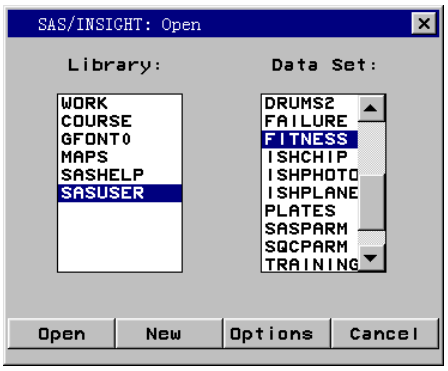


The Second Method: Using SAS/INSIGHT

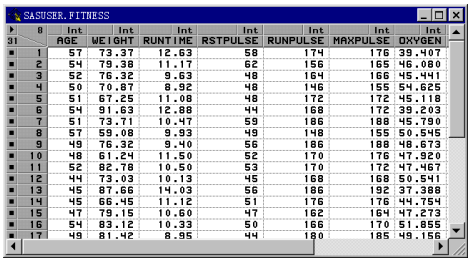
SAS/INSIGHT software is a tool for data exploration and analysis. With it you can explore data through graphs and analyses linked across multiple windows. You can analyze univariate distributions, investigate multivariate distributions, and fit explanatory models using analysis of variance, regression, and the generalized linear model.

To get started, you must invoke SAS/INSIGHT software in any of three ways: type **insight** on the command line; choose Globals: Analyze: Interactive data analysis; or use the SAS procedure.

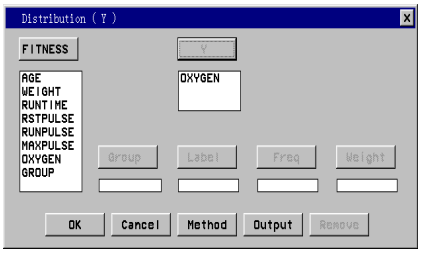
Upon invoking SAS/INSIGHT software, you are prompted with a data set dialog. Select the data set sasuser.fitness and click on Open.



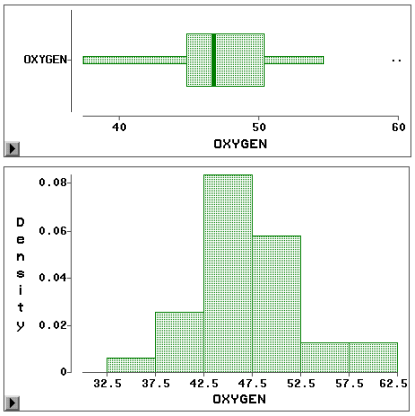
You will get the contents of data set SASUSER.FITNESS as follows;



In the pop menu select Analysis:Distribution
The Distribution (Y) window appears;
Then select the variable you want to analysis.
Select OXYGEN and Y;



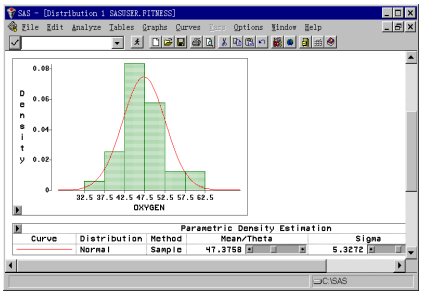
Then click OK and you will get a box plot and a histogram and the moments and quantile tables.



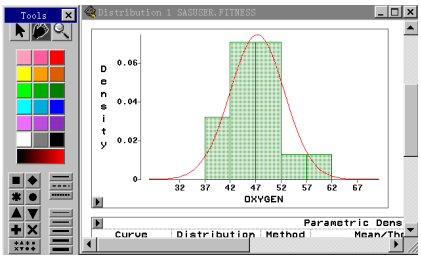
Moments			
N	31.0000	Sum Wgts	31.0000
Mean	47.3758	Sum	1468.6500
Std Dev	5.3272	Variance	28.3794
Skewness	0.4363	Kurtosis	0.6303
USS	70429.8597	CSS	851.3815
CV	11.2446	Std Mean	0.9568

Quantiles			
100% Max	60.0550	99.0%	60.0550
75% Q3	50.3880	97.5%	60.0550
50% Med	46.7740	95.0%	59.5710
25% Q1	44.8110	90.0%	54.2970
0% Min	37.3880	10.0%	39.4420
Range	22.6670	5.0%	39.2030
Q3-Q1	5.5770	2.5%	37.3880
Mode	37.3880	1.0%	37.3880

If you want to add a normal density curve, select Curve: Parametric Density. The parametric density estimation dialog will disappear. Click OK in the dialog. The density curve is superimposed on the histogram.



INSIGHT provide a easy tool to change the shape of the distribution. Choose Edit: Windows: Tools. Then click on the hand in the tools windows. The cursor changes shape from an arrow to a hand. Move the cursor back to the distribution window and click on the histogram. When you move horizontally, the bins start at different locations. When you move vertically, the bin width changes. Release the mouse button when you find a histogram that satisfied you.



Click on the arrow in the tools window before proceeding.

The Third Methods: Using SAS/LAB

SAS/LAB, a fully integrated component of the SAS System, is menu-driven, task-oriented software designed to guide the analysis of data and the interpretation of results. You can use SAS/LAB software quickly and easily without having to write SAS programs. Performing an

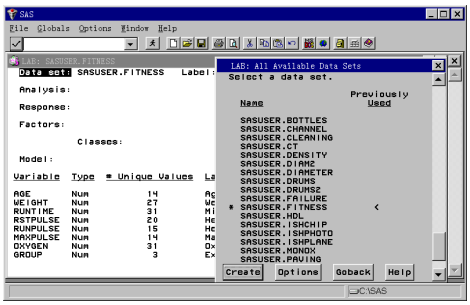
entire analysis can be as easy as making a few selections from menus and buttons.

With SAS/LAB software, you can conduct routine analyses with minimal statistical training. The summarization and analysis capabilities are linked with graphical tools that produce scatter plots, histograms, box-and-whisker plots, and contour plots. SAS/LAB software presents appropriate choices at each step of the analysis, provides a journal facility that enables you to store text and graphs with software-generated interpretations and personal comments, and offers extensive help.

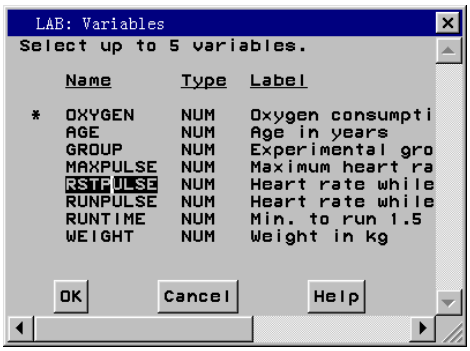
To invoke SAS/LAB, you can use any of three ways.

- type **Lab** in the command line and press **ENTER**;
- From the pop menu select Globe→Assist→DATA ANALYSIS→INTERACTIVE→Guide Data Analysis;
- From the pop menu select Globe→Invoke application→Data analysis→SAS/LAB.

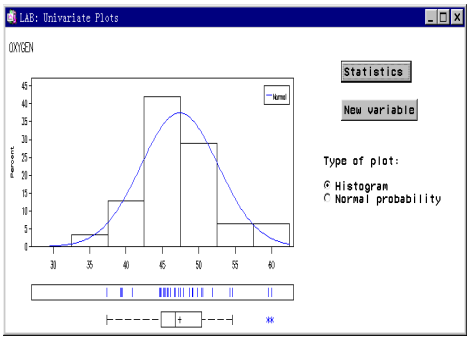
When you invoke the LAB main window, double click on the Data set:, then the ALL Available Data Sets window displays. You can select the data set you want to analysis. Here we select the sasuser.fitness as we selected before.



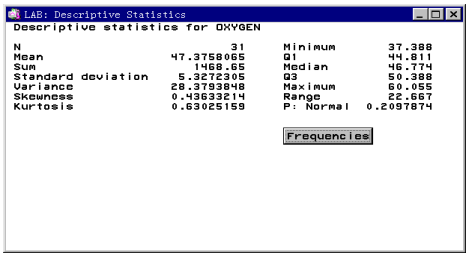
After you select the data set, then go back to the LAB main window. Click on Summarize to display the Variables window to select the variable for analysis. Then select OXYGEN;



Click on OK then Univariate Plots window appears as follows:



If you want to see the descriptive statistics, click on Statistics in the Univariate Plots window and you get the Descriptive Statistics windows.



CONCLUSION

This introduction had provided three ways to product a histogram and descriptive statistics. SAS software is the great powerful tool. The more you study, the more you will find it excellent. I love it. I hope the people around the world will share my happiness. Remember: Practice is the best teacher. Just do it.

REFERENCES

SAS/LAB Software: User's Guide,Version 6 First Edition, Cary, NC 27513: SAS Institute Inc.

SAS/INSIGHT: User's Guide, Version 6 Third Edition, Cary, NC 55582: SAS Institute Inc.

SAS/QC Software: SQC Menu System for Quality Improvement, Version 6, Second Edition, Cary, NC 55242: SAS Institute Inc.

AUTHOR CONTACT

Dengbing
Technology department of Baoshan iron & steel corporation
Guoyuan, Baoshan, Shanghai China

Tele: (8621) 56648648-2922
FAX: (8621)56198801