**Paper 145-27**

# Analyze the Stock Market Using the SAS® System
Luis Soriano, Qualex Consulting Services, Inc.
Martinsville, VA

## ABSTRACT
Financial Managers, Analysts, and Investors need to find the better opportunities of investment on a huge global market where every day it's more difficult to understand the behavior and tendencies of the economy.

The access and organization of information is crucial for the financial sector in order to have fundamental elements that will help Financial Manager, Analysts, and Investors to understand market behavior and how well decisions are taken.

This paper explains in a very general manner how the SAS System could be used to manage stock market and financial data in an organization giving a few examples of this.

The paper contains examples of the data warehouse process from the access to the data to the presentation of valuable information to analysts using stock market data and explaining how to accomplish some of these goals with SAS programs.

In some of the examples, it's explained the SAS modules or products that are used including some characteristics of them.

New SAS users can benefit from this paper to understand different SAS products and how they can be used. Experienced SAS users and business managers can benefit reading this document to obtain more experience on how the SAS System could be used in the financial sector.

Several SAS products will be highlighted and experience is not necessary to derive benefit from this paper.

## INTRODUCTION
SAS Software offers a variety of tools to companies who are interested to access, organize, and deliver useful information for the benefit of their business and to take decisions that will help them to obtain better financial results in their organizations and against the competence.

We are living the age of information technology where the best-informed and prepared participants win the battles. It's crucial for financial organizations to have information elements that will help them to take fast and better decisions in the battle field (The Market) otherwise they are condemned to a tragic fail sooner or later.

The investment in information technology is always considerable but the return of that investment will be more considerable if those resources are used effectively.

That's why this paper will focus to show the readers in a very general manner the benefits that financial companies could obtain from using SAS Software to address the different issues that organizations deal when they try to deliver useful information to their organization.

This paper is a mix of business and technical concepts, ideas, tips, and strategies.

I will guide the readers with examples to understand the "elements" that are needed to be considered in the construction of a Data Warehouse in the financial sector. When I'm referring to "elements", I'm talking about all those things that are related to the business itself and that will help us to build our information technology structure.

This paper is based on my personal experience and I hope that it will be of great help to all in their effort to provide better solutions to their organization.

### THE ANALYSTS OBJECTIVE
Wall Street is one of the most important stock market centers in the U.S.A and in the world where daily market operations take place every day between brokers that produce a huge volume of data.

We don't have enough time to talk about the characteristics and diverse financial instruments that are sold or bought here as Bonds, Futures, Warrants, Commodities and shares just to mention a few.

Let's focus on the operations that correspond to stocks or shares.

Fundamental Analysts usually have the responsibility to understand and follow the behavior of a particular sector or sectors; for example, Transportation, Technology, Services, Healthcare, Financial, and Energy to identify and recommend the strategy that the investors need to follow respect to a specific stock or share.

In TABLE-1 we can see the structure of a daily stock operation:

| Variable | Value | Type | Description |
|---|---|---|---|
| Date | 10SEPT2001 | N | Date of operation |
| Hour | 11:33 AM | N | Time of operation |
| Stock | IBM | C | Name of the stock |
| Price | $96.47 | N | Price of the stock |
| Volume | 1,300 | N | Number of stocks |
| Total | $125,411 | N | Total of the operation |
| Buyer | JP Morgan | C | Buyer of the stock |
| Seller | Merrill Lynch | C | Seller of the stock |

**TABLE- 1**

The interpretation of this daily stock operation explains that in September 10, 2001 at 11:33 AM the broker house Merrill Lynch sold to JP Morgan 1,300 IBM stocks at a price of $96.47 USD; the total amount of the operation was for $125,411.00 USD.

Thousands of operations of this type occur every day on Wall Street. With this volume of information the IT department has the challenge to deliver this data to analysts.

With this data the IT department and the business analysts can perform the following tasks:

- Create daily operations databases
- Create daily prices databases
- Evaluate portfolios
- Benchmark a group of stocks in a period of time
- Evaluate the recommendations of analysts
- Find strategies of the competence
- Fundamental analysis
- Graphical Technical Analysis
- Deliver this information using an intranet

These are just a few tasks that can be performed using the SAS® System. In the next sections of the paper we will see how to perform some of these tasks using the SAS® System.

**ACCESS AND VALIDATION OF DAILY STOCK OPERATIONS**
Broker houses receive the stock market operations during the day while the market is in operation. The data that they received is not limited only to stock market operations and usually they receive other types of data; for example, the financial statement of companies every quarter of the year.

This is possible due an agreement between the broker houses and the institution which records, handles, and controls the market operations. A monthly fee needs to be paid to receive this data and usually someone who has the hardware infrastructure and money to pay the fee can get access to the data.

Over the world, the format of how this information is sent to the broker houses varies and is changes as technology does. Usually, the preferred format to send these data is flat files, ASCII files or text files. Other scenario could be that broker houses have access to a relational data base where they can extract the information that they need; however, this scenario is not recommended. It's preferable that every broker house creates and maintains its own data bases.

That's why is crucial for financial institutions or business to use the correct tools to handle the data. SAS Software is used and could be used over the world to handle this type of data without distinction of the file format or platform where the data needs to be extracted or collected.

FIGURE-1 represents text lines (file p_09102001.txt) as a possible format of how the data is sent to the broker houses for the daily stock operations.

```
091020011133SUN0000037250000038300JPML
091020011134IBM0000096470000001300BACG
091020011134HP 0000028660000450000CGJP
091020011134SUN0000037000000038300BAML
091020011134GTW0000009250002000000JPML
091020011135SUN0000036750000038300BACG
99999999
```
**FIGURE-1**

To read this text file we will use a DATA step; an element of SAS/BASE® software and where you can see how easy is to program and create a SAS data set:

```
data operations(keep=date time stock_symbol stock price volume
                    buyer seller amount_of_operation);
infile 'c:\stock market\p_09102001.txt';
length stock buyer seller $ 20;
array broker_house{2} $ buyer seller;
array broker_code(2)  $ buyer_code seller_code;
input control 1-8 @;
if control ne 99999999 then
do;
    input month 1-2 day 3-4 year 5-8 hour 9-10 minute 11-12
          stock_symbol $ 13-15 price 16-25 volume 26-35
          buyer_code $ 36-37 seller_code $ 38-39;
          date=mdy(month,day,year);
          time=hms(hour,minute,0);
          select(stock_symbol);
           when('GTW')
                    stock='Gateway';
           when('HP')
                    stock='Hewlet Packard';
           when('IBM')
                    stock='IBM Corporation';
           when('SUN')
                    stock='Sun Microsystems';
           otherwise
          end;
          price=price/100;
          amount_of_operation=price*volume;
          do i=1 to 2;
```

```
          select(broker_code{i});
           when('JP')
                broker_house{i}='JP Morgan';
           when('ML')
                broker_house{i}='Merrill Lynch';
           when('BA')
                broker_house{i}='Bank of America';
           when('CG')
                broker_house{i}='City Group';
           otherwise
          end;
    end;
    label stock_symbol = 'stock symbol'
          amount_of_operation='amount of operation';
    format date date9. amount_of_operation price dollar20.2
          volume comma20. time time8.;
    output;
  end;
run;
```

In our first data step example we have performed two phases of the datawarehouse process, access and transformation of data. This is a valuable characteristic of the SAS System. The SAS System is a tool that allows organization to build data warehouse projects from the backend to the front end. In our example, the transformation rules are very simple. These rules are usually called "Business Rules". The definition of the business rules represents the business model essential to the datawarehouse process. The transformation phase includes cleaning, validation, and quality insurance of data. The integration and summarization of data is also considered part of the transformation phase.

After we have read our input file we can display our SAS data set in a tabular format using SAS/FSP® software. (FIGURE-2)


**FIGURE-2**

SAS/FSP software goes beyond data manipulation; it lets you interact with your data using integrated tools for data entry, computation, query, editing, validation, display, and retrieval.

In case the input data were in another format; for example, ORACLE, we could use the SAS/ACCESS® software to ORACLE to access and extract this data. We are not going to talk about SAS/ACCESS software to ORACLE but it's important to tell the readers that SAS can read multiple files formats as relational databases, non-relational databases, ERP systems (Enterprise Resource Planning), and other data sources.

The SAS/ACCESS software interfaces are the foundation of the SAS data access solution. These interfaces enable your SAS solutions to read, write, and update data regardless of its native database or platform.

SAS/ACCESS software offers interfaces to the following relational databases:

| | |
|---|---|
| DB2 under OS/390 | OLE DB |
| DB2 under VM | ORACLE |
| DB2 under UNIX or PC | Oracle Rdb |
| CA-OpenIngres | SYBASE |
| Informix | MS SQL Server |
| ODBC | Teradata |

2

## THE DAILY OPERATION AND PRICES DATABASES

The daily stock operations produce the daily operation database; the daily operation database will be the input for the prices database. The prices database consists of the summarization of all the stock operations; this database contains information per day of each stock. Each stock will be summarized to high, low, and close price, total volume and total amount operated.

The following SAS program is used to create the daily operation database:

```
proc sort data=operations;
        by date stock_symbol time;
data close(keep=stock_symbol price);
        set operations;
        by date stock_symbol;
        if last.stock_symbol;
proc summary data=operations nway order=data;
        class date stock_symbol stock;
        var volume amount_of_operation price;
        output out=daily_resume(drop=_type_ _freq_)
        sum(volume amount_of_operation)=total_volume
        total_amount_of_operation
        max(price)=high
        min(price)=low;
        label volume='total volume'
        amount_of_operation='total amount of operation';
proc sort data=daily_resume;
        by stock_symbol;
proc sort data=close;
        by stock_symbol;
data daily_resume;
        merge daily_resume close(rename=(price=close));
        by stock_symbol;
run;
```

FIGURE-3 shows the prices database of our example.



FIGURE-3

## DIVIDENDS AND SPLITS

Financial databases as the daily operation database and the prices database suffer changes in their numeric values when the companies that own the stock enact a split or a dividend, so the price of the stocks is altered. In order to maintain these databases, you can write SAS MACROS that will alter the historic values of the stock. The following program is the example of the MACRO dividend:

```
%macro dividend(dividend,last_price,stock_symbol,date);
proc sql;
update market.daily_prices
  set high=high -(high *(&dividend/&last_price)),
     low=low  -(low  *(&dividend/&last_price)),
     close=close-(close*(&dividend/&last_price)),
     amount_of_operation=amount_of_operation –
     (amount_of_operation*(&dividend/&last_price))
  where (stock_symbol eq "&stock_symbol" and date lt "date"d;
update market.daily_operations
  set price=price-(price*(&dividend/&last_price))
  where (stock_symbol eq "&stock_symbol" and date lt "date"d;
update market.daily_operations
  set amount_of_operation=price*volume;
   where (stock_symbol eq "&stock_symbol" and date lt "date"d;
%mend;
```

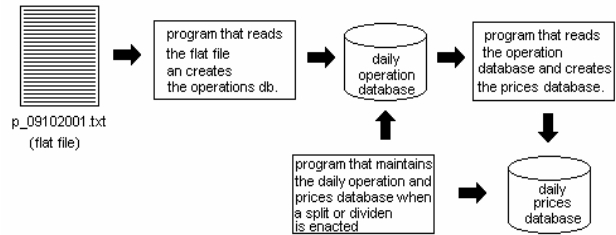At this point, we can represent the flow of information with FIGURE-4.



FIGURE-4

It's important to mention that these databases are updated appending the information of each day to a master or historic database; this is possible because the stock operations are independent per day. The operations and prices database are used in a diversity of tasks to perform market analysis.

## COMPARING STOCKS FROM THE SAME SECTOR

Financial Analysts need to find the best instruments to invest. The common method that helps to understand the behavior of stocks is to graphic them together.

TABLE-2 and FIGURE-5 shows some of the most important stocks in the Computer-Hardware Technology sector:

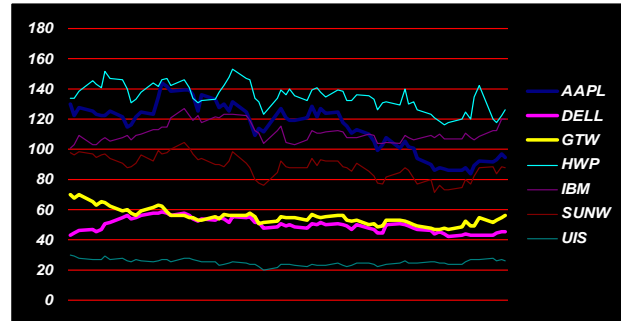| STOCK SYMBOL | NAME |
|---|---|
| AAPL | Apple Computer, Inc |
| DELL | Dell Computer Corporation |
| GTW | Gateway, Inc. |
| HWP | Hewlett-Packard Company |
| IBM | IBM |
| SUNW | Sun Microsystems, Inc. |
| UIS | Unisys Corporation |

TABLE-2



FIGURE-5

In FIGURE-5 is difficult to determine which stocks are the best to invest, because all of them are in a different scale. To resolve this issue we can follow a very simple technique to produce a new graphic which will compare each of the stocks in the same scale as FIGURE-6.
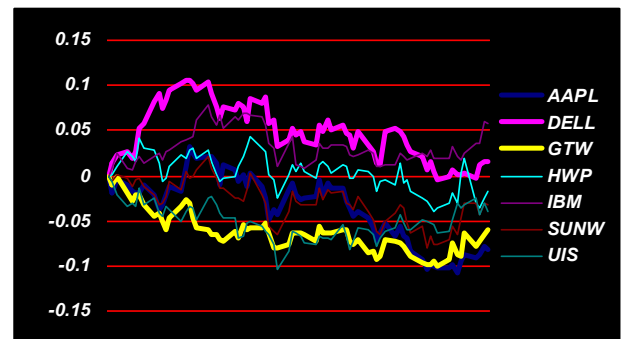


FIGURE-6

3

FIGURE-6 represents the performance of the stocks in a selected period of time and it starts comparing them from cero. The stocks which have the highest value at the end of the graphic are the ones that would give an investor the higher yield in that period of time.

The steps to generate this graphic are the following:

1. Select the group of stocks that you want to analyze from the daily operations database indicating a period of time; for example, our graphic shows 7 stocks from March 1, 2000 to June 8, 2000.

2. Determine which of the stocks has the lowest price in the close price variable in the first date of the series; for example, Unisys Corporation shows the lowest close price of $30.188 for March 1, 2000. (TABLE-3)

| STOCK | DATE | CLOSE |
|---|---|---|
| Apple Computer, Inc | March 1, 2000 | $ 130.312 |
| Dell Computer Corporation | March 1, 2000 | $ 43.016 |
| Gateway, Inc. | March 1, 2000 | $ 69.812 |
| Hewlett-Packard Company | March 1, 2000 | $ 133.625 |
| IBM | March 1, 2000 | $ 100.438 |
| Sun Microsystems, Inc. | March 1, 2000 | $ 97.938 |
| Unisys Corporation | March 1, 2000 | $ 30.188 |

**TABLE-3**

3. Determine a factor value dividing the first close price of each stock series and the lowest close price of step 2. (TABLE-4)

| STOCK | FORMULA | FACTOR |
|---|---|---|
| Apple Computer, Inc | 130.312/30.188 | 4.3166821 |
| Dell Computer Corporation | 43.016/30.188 | 1.4249371 |
| Gateway, Inc. | 69.812/30.188 | 2.3125745 |
| Hewlett-Packard Company | 133.625/30.188 | 4.4264277 |
| IBM | 100.438/30.188 | 3.3270836 |
| Sun Microsystems, Inc. | 97.938/30.188 | 3.2442692 |
| Unisys Corporation | 30.188/30.188 | 1 |

**TABLE-4**

4. Finally, divide each series between the corresponding factor minus the lowest close price value determine in step 2 and divide the result by a hundred.

Close price= ((close price / factor)–lowest close price) / 100

After these steps, you can compare all the series in a graphic as FIGURE-6. The following SAS MACRO program is used to generate FIGURE-6:

```
%macro graph_compare(from,to,sector);

/* selection of the stocks */

proc sql noprint;
create table temp as
select date, close, stock_symbol
from   work.daily_resume
where  date   ge "&from"d and date   le "&to"d   and  sector eq
"&sector";

/* determine the lowest close price */

select min(close) into : minimum
from   temp
where  date   eq "&from"d;

/* determine the factor */
create table factor as
select stock_symbol, close/&minimum as factor from   temp
where  date   eq "&from"d
order by stock_symbol;
```

```
quit;run;

/* sort by stock_symbol and date */

proc sort data=temp;
by stock_symbol date;
run;

/* divide the series between the factor */

data temp(drop=factor);
merge temp factor;
by stock_symbol;
close=((close/factor)-&minimum)/100;
run;

/* sort the series before the transpose procedure */

proc sort data=temp;
by date stock_symbol;
run;

/* transpose the dataset for graphical purpose */

proc transpose data=temp out=temp2(drop=_name_ _label_);
by date;
id stock_symbol;
var close;
quit;run;

/* graphic options */

goptions reset=(axis, legend, pattern, symbol, title, footnote)
norotate hpos=0 vpos=0 htext= ftext= ctext= target= gaccess=
gsfmode= device=WIN  ctext=blue graphrc nterpol=join;

symbol1 c=DEFAULT i=SPLINE ci=GREEN  w=5 v=NONE;
symbol2 c=DEFAULT i=SPLINE ci=BLUE   w=5 v=NONE;
symbol3 c=DEFAULT i=SPLINE ci=YELLOW w=5 v=NONE;
symbol4 c=DEFAULT i=SPLINE ci=ORANGE w=5 v=NONE;
symbol5 c=DEFAULT i=SPLINE ci=BROWN  w=5 v=NONE;
symbol6 c=DEFAULT i=SPLINE ci=BLACK  w=5 v=NONE;
symbol7 c=DEFAULT i=SPLINE ci=PINK w=5 v=NONE;

axis1 color=blue width=2.0;
axis2 color=blue width=2.0;
axis3 color=blue width=2.0;

/* procedure to produce the graphic */
proc gplot data=WORK.TEMP2;
plot (AAPL DELL GTW HWP UIS SUNW IBM) * DATE / overlay
haxis=axis1 vaxis=axis2 frame;
run; quit;
%mend;

%graph_compare(01MAR00,08JUN00,Computer-Hardware
Technology);
```

As you see with a few SAS code lines we have generated a program that will help us to understand graphically the behavior of a group of stocks and a view of the possible best options to invest. PROC GPLOT is part of SAS/GRAPH® software.

**EVALUATION OF PORTFOLIOS**
One of the principal activities of financial institutions is to determine the present value of the investments of their customers; for example, an investor decides to invest $10,000 dollars in IBM stocks in March 1, 2000. What's the value of this portfolio in January 7, 2002?

The price of IBM in March 1, 2000 was $100.438, so the initial amount of shares bought was 99.56; this is the result of dividing $10,000 between $100.438.

4

TABLE-5 shows the value of the portfolio in January 7, 2002.

| | January 7, 2002 |
|---|---|
| Initial Investment | $10,000.00 |
| Instrument | IBM Shares |
| NAV | $124.90 |
| Shares | $99.56 |
| % Yield | 24.35% |
| Value | $12.435.04 |

**TABLE-5**

The value of the portfolio is determined by multiplying the number of shares and the Net Asset Value (NAV) of the instrument. The example shows a very simple portfolio; however, it describes how it works. The problem starts when a financial institution has thousand of customers who buy and sell instruments on-line, when they use other methods, and when the composition of the portfolio is more complex.

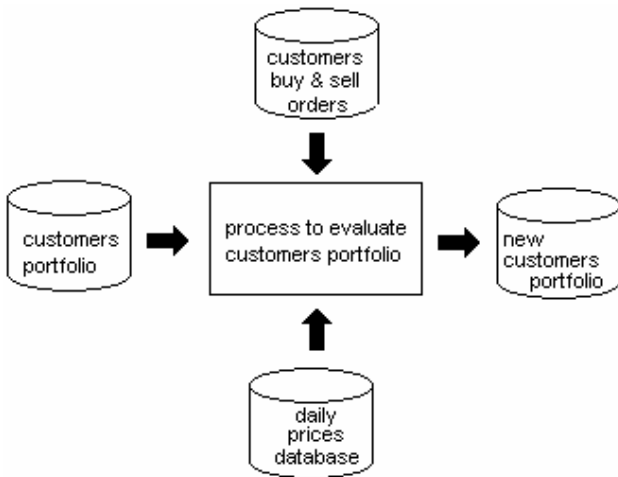FIGURE-7 describes the databases used in this process.



**FIGURE-7**

Financial Institutions require tools and methodologies that will help them to handle the information of their customers effectively. This is a key for business success but also it's needed to offer to customers additional guide and recommendations that will help them to obtain profit from their investments.

SAS Software posses a diversity of tools that will help Financial Institutions to handle and manage huge volumes of information and on the other hand SAS has tools for analysts that will help them to perform analysis.

A variety of strategies could be implemented in the management of customers' portfolios; for example, to set personalized alarms when a stock has reached certain value indicating that and adjustment should be made in the portfolio. At certain level, this means that and intelligent algorithm will prevent customers of certain situations giving them the opportunity to adjust their investments in order to obtain the best profits.

### FINANCIAL DATAWAREHOUSE
I have shown examples of how to use the SAS System in order to analyze or handle financial data; however, I haven't talked about how to integrate all this individual pieces in a common structure. This common structure that I'm talking about is called data warehouse.

Financial Institutions have access to diverse data sources; the challenge here is to organize those sources in a way that anyone in the institution could be able to utilize that information. The organization of information and the automation of processes will give analysts more time to do what they are best to do which is "Analysis". It's relatively common to find worldwide institutions that have manual processes to collect financial data or automated processes that take to much time to deliver data to analysts or institutions which deliver wrong data to their analysts without knowing it. Maybe, these institutions are operable; however, they are not being efficient.

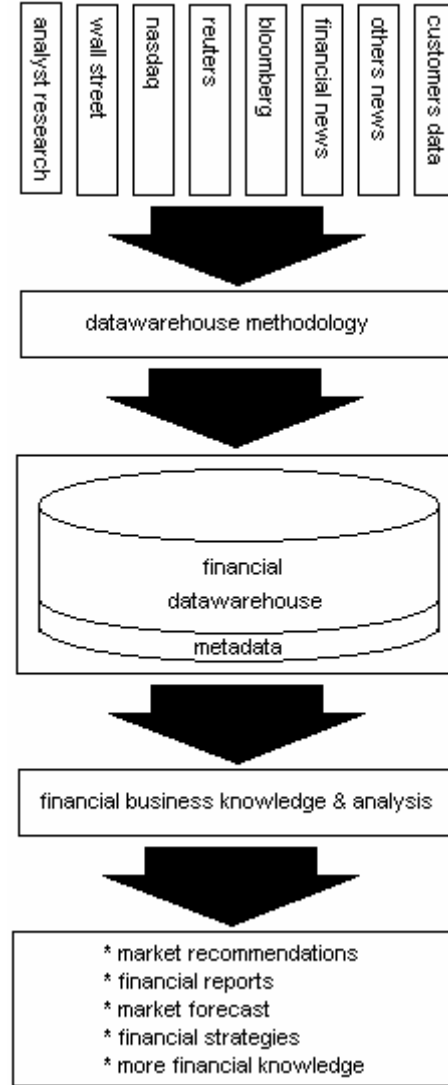FIGURE-8 shows in a high level view the process of data warehouse.



**FIGURE-8**

The idea here is first to identify all those data sources that could be used by the institution to create the datawarehouse like stock market data, electronic news, real time financial data, and customers data (portfolios). Then with a previous defined methodology, those sources will be accessed, transformed, validated, and organized based on the rules of the business in order to create the datawarehouse. This warehouse will have a data layer (metadata), which is the dictionary of the warehouse, used to register the information contained in the warehouse and the processes that update that information.

Finally, this information will be delivered to the different departments in the organization by different ways; for example, a web application or other mechanism that will permit to the owners of the information to take better decisions in benefit of their institution and customers.

## CONCLUSION

With a few examples, I have shown how the SAS system could be used in financial institutions in order to manipulate, present, and provide valuable information to Financial Analysts or Financial Managers to take financial decisions or to follow market behavior. SAS has extensive tools that will help Financial Institutions organize and analyze data. The examples showed here refer to an approach of how to organize data and some analysis through data management or numeric techniques.

## AUTOR BIOGRAPHY

Luis Soriano was a Senior Analyst at Scotiabank Inverlat in Mexico City. Scotiabank is one of North America's leading financial institutions, and Canada's most international bank. Now, he is working as a SAS contractor in The U.S. Census Bureau. Luis is employee of Qualex Consulting Services, Inc. (SAS Gold Quality Partner) and he is assigned to work in conjunction with the SAS Public Sector Group.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged.
Contact the author at:

|  |  |
|---|---|
| Author Name | Luis Soriano |
| Company | Qualex Consulting Services, Inc. |
| Address | 2015 Key Blvd # 601 |
| City state ZIP | Arlington, VA 22201 |
| Work Phone: | (703) 456 - 1438 |
| Fax: | (703) 832 - 8447 |
| Email: | luis.soriano@qlx.com |
| Web: | http://www.qlx.com/ |

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brands and product names are registered trademarks or trademarks of their respective companies.