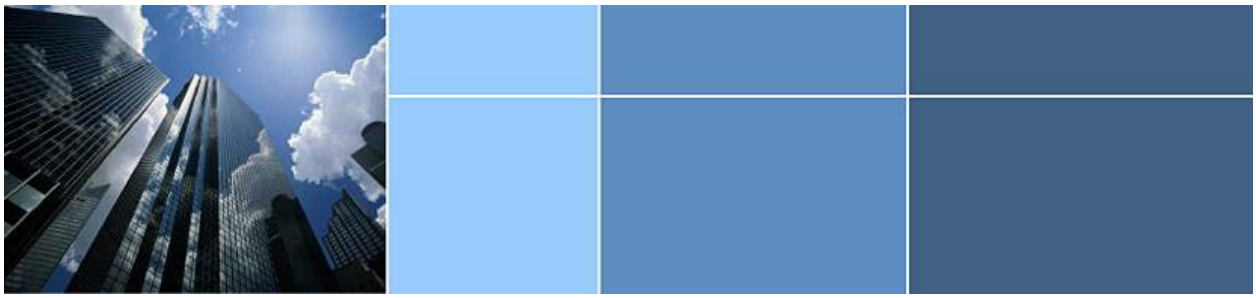
June 18, 2009

Newmont Mining Corporation Statistical Analysis Final Report

Newmont Mining Corporation



Newmont Contact:	**Mines Contacts:**
Shawn Horton Manager TSS System Support 303-708-4323 Office .Horton@Newmont.com	Elliott Hoefflin @mines.edu Andrew Lass @mines.edu Monica Noring @mines.edu

ABSTRACT

The Newmont Statistical Analysis Program (NSAP) is a program that was requested by Newmont Mining Corporation geostatisticians and executives to assist in determining the economic feasibility of a potential mine sight. Previous functionality of the software is scattered in multiple programs which are outdated and not user-friendly. For these reasons, the old software functionality has been combined and updated in order to better assist the geostatisticians in their work. The final product allows expanded functionality for the users, including new plot options and an additional type of plot, while also giving the software an updated Office 2007® / Windows Vista® look and feel. By using C# in coordination with Charting and Graphing tools from Infragistics, the new version of the software is professional and modern, while still guaranteeing speed and accuracy.

INTRODUCTION

Newmont Mining Corporation was founded in 1921 and has been publicly traded since 1925. Newmont is one of the world's largest gold producers and is the only gold company included in the S&P 500 Index and Fortune 500. Headquartered near Denver, Colorado, the company has approximately 34,000 employees and contractors worldwide[i]. They have a number of operational gold, silver and copper mines along with a number of potential mines in various phases of startup around the globe.

One of the issues surrounding the determination of the economic feasibility of these potential mine sites is the use of complex statistics based on data gathered at each site. The software used by geostatisticians at Newmont Mining already existed, but was outdated, not user friendly and required the use of several command-line driven programs for complete analysis. The new software application, which is known as the Newmont Statistical Analysis Program (NSAP), is a single user friendly application that performs the required data analyses, while also allowing the user to output the results of their analysis in a file format that can easily be saved or printed.

The newly created NSAP application includes an improved graphical user interface (GUI), a feature requested by the key stakeholders of this project. This new GUI allows users to choose the input file they would like to analyze, as well as several other parameters which will allow the users to view the data exactly as they need. A new feature implemented will allow users to graphically analyze multiple sets of data on a single plot. Along with the graphical representation of the data, the final product also outputs the resulting statistics in textual form. The process of GUI implementation has been guided by the needs of geostatisticians—the primary users the software.

A second component of the new NSAP focuses on the output, or the statistics calculated by the new application. In order to verify both the correctness and completeness of this statistical output, the NSAP performed diagnostics between both the existing software and those of the final product.

In summary, the newly created software product has added value by:

- performing more statistical operations than the previous version
- being more user friendly which decreases new user ramp-up time
- allowing fewer steps to produce an output file which reduces the possible number of potential errors
- generating fewer errors which increases user confidence in the statistical results

PROJECT REQUIREMENTS

USER INTERFACE

GUI

The graphical user interface for Newmont's Statistical Analysis Program (NSAP) uses Windows Presentation Foundation (WPF) as the graphical system in which the components of the GUI will be implemented. WPF uses the Extensible Application Markup Language (XAML), which will designate the appearance of each component of the GUI. After the user enters the required fields and runs the program, the results appear in a new WPF output window.

User Input

In the primary use case, the user will choose a file from either of the file types supported, enter several other required parameters, and click "Run" to start the program. If the user wishes to plot multiple data sets on a single plot, or if they simply want to control the number of the bars in the histogram, the program will allow entry of these values. In the secondary use case of plotting multiple data sets on a single graph, this will normalize the data so it can be shown efficiently on the single plot output.

FILE FORMATS

Geostatisticians use three common input data file types, two of which will be supported by the new Newmont Statistical Analysis Program (NSAP). Both are described below.

rf

The first file type supported is the .rf random file. This file type contains point data from any number of spatial points which are defined by x, y, and z coordinates, along with the gold grade and other parameter values at that location.

GRD

The other file type supported within the application is the .grd file. This file contains a cubic grid representing the potential mine site. Each cube contains data pertaining to various geological statistics for the represented area. The .GRD file saves space in that it does not store the x, y, and z coordinates directly—they can be calculated from the cube number and block size.

STATISTICS

Previous versions of the software allowed for a number of statistics to be calculated for each of the data file types. This functionality continues in the new version of the software and supports the following calculations.

rf Functionality

1. Minimums and maximums for each x, y, and z coordinate of the points, as well as any user defined values (such as gold grade)
2. Total number of points and total number of undefined points for each of the user defined fields
3. Average value and variance for each of the fields
4. Standard deviation, percent standard deviation and standard deviation error for each of the fields

GRD Functionality

1. Total number of samples
2. Minimum and maximum values
3. Average value and variance
4. Standard deviation

PLOTTING

Plot Types

Previous functionality also allowed the user to generate a histogram and cumulative frequency plot of a single data set. The final product continues to allow this functionality, while also adding options for plotting multiple data sets on a single set of axes. In addition, functionality for another type of plot, the Q-Q plot, which is useful in measuring the correlation of a set of data, has been added. Definitions of these plot types exist in the Glossary.

Infragistics Tools

Infragistics[ii] is a company that specializes in the look and feel of today's Microsoft applications, such as Office® 2007 and Windows® Vista®. This development toolkit is used for Charting and Graphing capabilities, which incorporates the Windows Presentation Foundation (WPF). By using these tools, the final product has a professional, modern look and feel, while also being fast and accurate.

PLATFORM

For the final product, the latest in Microsoft .NET languages has been used—C# (see Glossary for definitions). This will allow total .NET compatibility with the C++/CLI wrappers for the file format libraries, while also allowing flexible and powerful GUI development and keeping the speed necessary to run the program efficiently. In order to develop in C#, Microsoft Visual Studio 2008 has been used.

OUTPUT

Program Output

The last step of the final program will be to show the output, specifically the statistics and plots, to the user. This will include opening a window which will graphically display the statistics and plots on the screen.

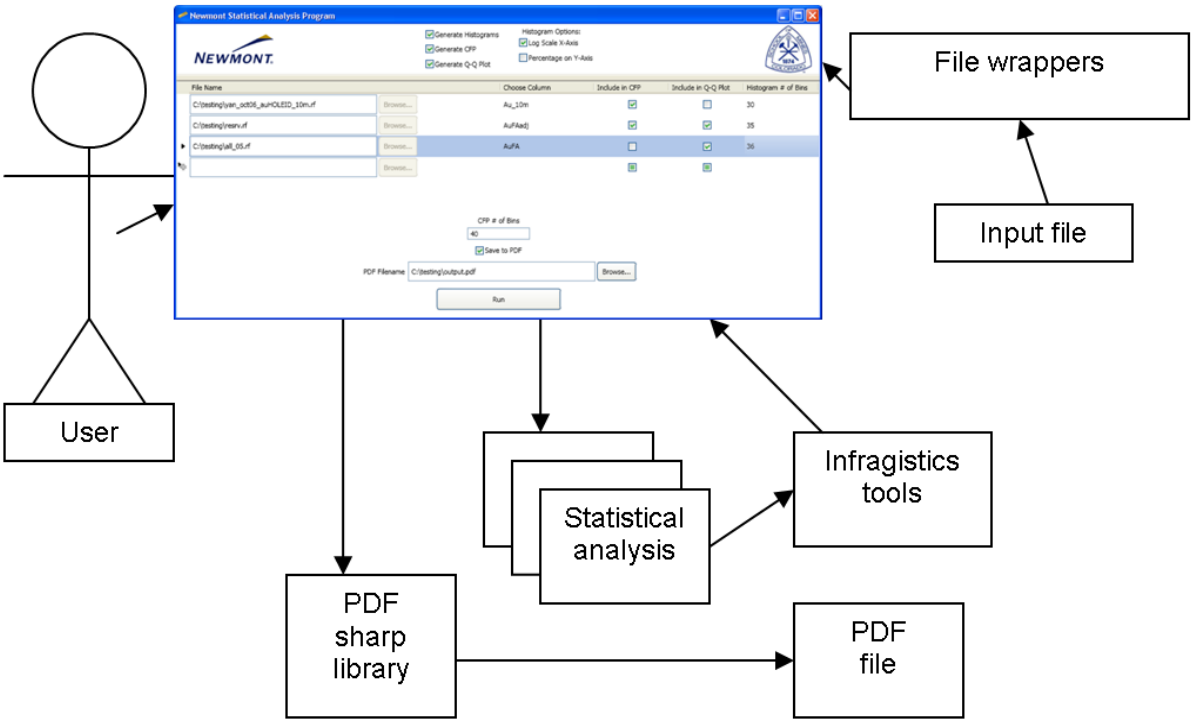
Save to PDF

Upon user approval of the program input, a check box can be activated so the results can be saved to a PDF file by simply specifying a file name and location. The PDF file is the opened by the NSAP, so it can be easily printed, sent or published.

PROJECT DESIGN

HIGH-LEVEL DESIGN

The NSAP program is a stand-alone application that links to various other components from Newmont software packages, as well as several external packages for plotting and PDF generation. When the program is initially started, the GUI has fields in which the user will enter the input files. Once the user has entered the data files and several other parameters regarding the form of the output, the user presses the "Run!" button. When the button is pressed, the program calculates the required statistics and generates the plots. The statistics and plots are then displayed to the screen, and a PDF file is generated so it can be published or printed. The following diagram shows a high-level overview of the system.

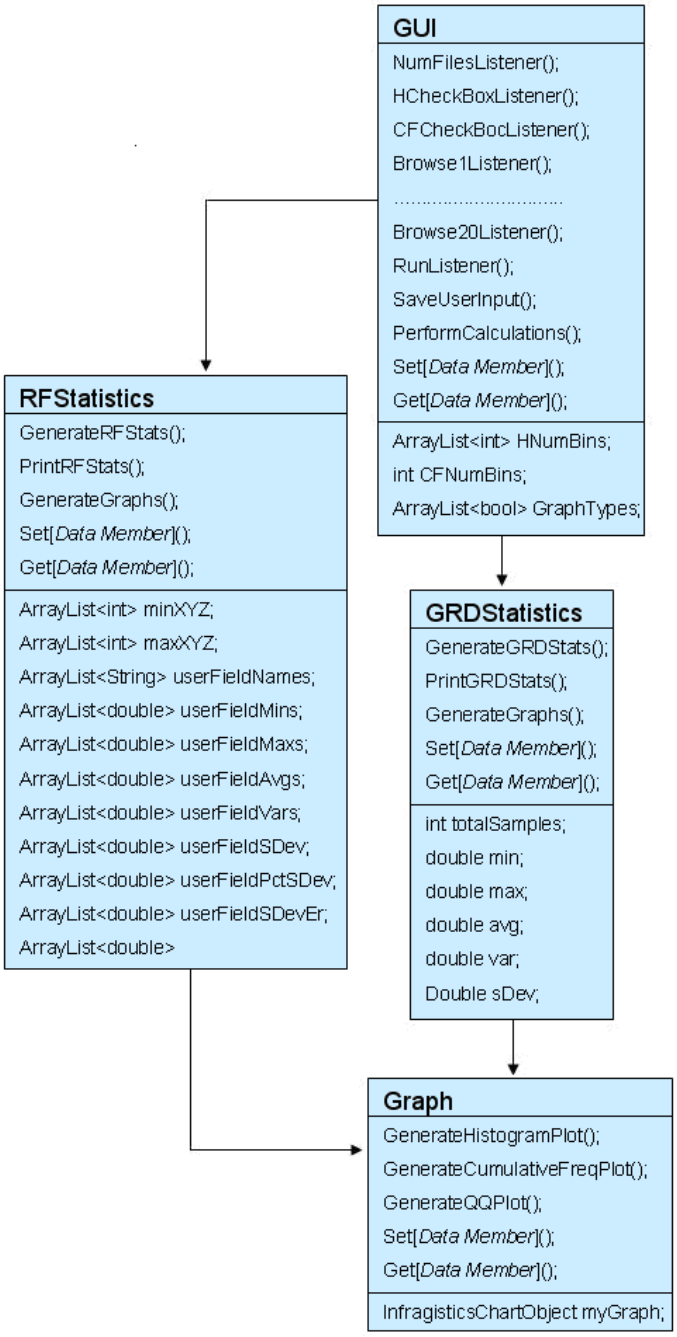


Note: A larger screenshot of our final product is shown in Appendix A.

DETAIL DESIGN

UML Diagram

The UML diagram on the following page shows the interaction of classes within the NSAP program.

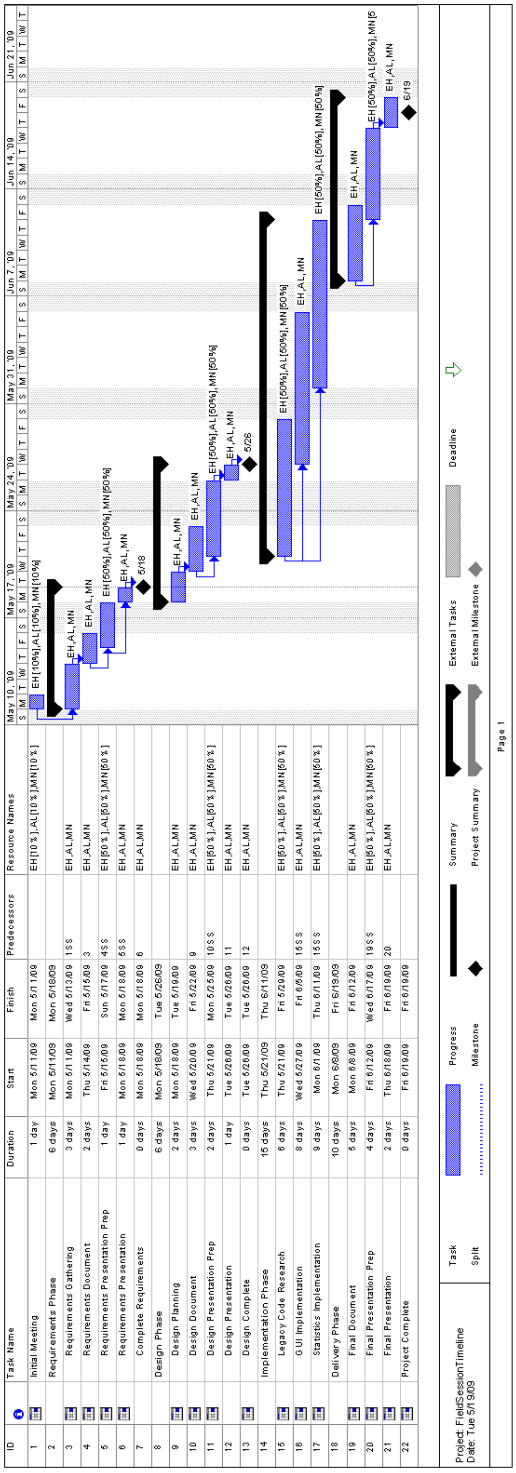


Legacy Code

The statistics functionality that has been put into the final NSAP program has been rewritten into C# based on legacy C code that already existed within Newmont. The exact functionality is detailed in the Statistics section, which is located on page 5.

SCHEDULE AND TIME ESTIMATES

The following diagram shows a timeline of work for the entirety of the field session project.



Note: This Microsoft Project document appears after Appendix A, enlarged to show details.

IMPLEMENTATION AND RESULTS

IMPLEMENTATION

Details of Implementation

As previously mentioned, we chose to use the .NET language of C#, which incorporates the latest technology from Microsoft. This will allow total .NET compatibility with the C++/CLI wrappers for the file format libraries, while also allowing flexible and powerful GUI development and keeping the speed necessary to run the program efficiently.

In order to develop in C#, Microsoft Visual Studio 2008 has been used. This development tool allows full functionality with the Windows platform, while also allowing visual GUI development and a source control server to backup our progress.

For development of graphs, including histograms, CFP, and Q-Q plots, WPF Controls from Infragistics are used. These provide powerful charting tools, while maintaining quickness and giving a professional look and feel.

For implementation of PDF functionality, an open source library called PDFSharp, which was been specifically designed for saving PDF files using C#, has been used. This allows professional, proven PDF functionality, while also allowing us the flexibility to choose the format and amount of information placed into a give PDF file.

Steps of Implementation

In order to implement the solution outlined above in the requirements and design portions of this document, we followed these steps, in general:

1. Analyzed legacy C code in order to update it
2. Rewrote these outdated programs in C#
3. Built the initial GUI window allowing data entry
4. Researched the Infragistics tools in order to determine the types of charts needed and the formatting required to achieve our desired output
5. Designed and implemented several file format classes that allow easy and efficient execution of our statistical analysis program, including calculation of the data required to generate the plots
6. Researched and configured the PDFsharp library that allows for generation and saving of PDF files
7. Built a graph class in order to simplify the code required in each file class
8. Wrote code pertaining to PDF file generation
9. Completed programming, including testing and bug correction
10. Walked through with client and users in order to verify that program requirements were met

Throughout this entire process, we worked closely along-side the users and client in order to guarantee that the work we did was inline with their thoughts and visions of the project. This allowed very rapid feedback about the quality of the project and quick response to any questions or issues we had. This also allowed the users to be very closely tied into the design process, adhering to the Agile methodology that Newmont strives for.

Issues During Implementation

Several issues arose during the implementation process that required some relatively major changes to our initial design.

The first issue we had was the assumption that the WPF Controls were actually just pieces of a GUI that could be added to a WinForm, which is an instance of a Windows::Form class that allows a programmer the power to visually design a GUI using the Visual Studio environment. That assumption turned out to be incorrect, which required us to re-design and re-implement the entire GUI window, using WPF from the start and incorporating the Infragistics Controls.

The second issue that we ran in to, which also pertained to the GUI, was the distinction between the initial input window and the output window. This issue arose from the fact that we could not easily add charts and data results to the initial window, while also being dynamic enough to allow for any number of files that the user chooses to enter. This issue was fixed by separating the input and output GUI's, which allowed us to dynamically add data and charts to a window, then build it after everything had been added.

RESULTS

Analysis

In the conversion of the Newmont legacy C code to C#, we performed an analysis of the old code and spoke to the geostatisticians in order to understand the statistics being calculated. The specific calculations are described in the Statistics functionality sections in the Requirements part of this document.

The other analysis we performed was on the current program that allows the plotting of data from random files. In that analysis, we determined the form of data required in order to generate the histograms, cumulative frequency plot, and Q-Q plot. The specific plots are described in the Plot Types section in the Requirements part of this document, above.

Code Requirements

The result of our implementation has some specific requirements and dependencies in order to run properly. The first of which comes from the fact that we used C#. This is the .NET Runtime Environment 3.5 or higher, which allows our executable to be run properly on any Windows computer with this Runtime installed.

The second specific requirement of our program is the need for the Infragistics tools to be used. In our case, the dll's (see Glossary) are used as references in our project files, so each individual user does not need to purchase a license from Infragistics. They are required, however, to install a redistributable package from Infragistics. On the contrary, any programmer who will be compiling the programs will need to purchase a license of the Infragistics Controls to do this.

Another requirement of our program is the .NET file wrappers from Newmont. These come as external dll's that are installed on the programmer's computer and are used as references in the program. Again, these require an installation of the wrappers on each user's computer.

The final requirement of our Newmont Statistical Analysis Program is the PDFsharp open-source C# library. This is a download that is available online, and is used as another reference for our program. As with the other requirements, any user will need to have the dll's available to run the program, and any programmer will need them as references in order to compile.

SCOPE CHANGES

Incomplete Requirements

There were several initial requirements of the project that we were unable to complete before the deadline for a number of reasons. The first of these is the option to save the Workspace. Initially, we intended to give the user the option to save all of the input data to a data file so it can be easily reloaded and recalculated. Unfortunately, we were unable to complete this request because it required the serialization of data to an XML format which gets very complicated with the use of the Infragistics Grid Control. As this was not an important part of the program, our client advised us to leave that as an enhancement.

The second request we didn't meet was to allow the user to generate data for DHL files. Early on in the implementation process, our client requested that we remove DHL as a supported file format, because the benefit it brought to the final product was outweighed by the amount of work. This too will be left for a future enhancement.

The final implementation detail that we did not complete was the support of a probability scale on the cumulative frequency plot. A probability scale is a logarithmic-type scale in which the outer bounds of the x-axis are expanded so details at the very beginning and very end of the data set are detailed. Unfortunately, this type of axis is not currently supported by the Infragistics tools, so it would require the generation of a complicated user control, which was far outside the scope of this Field Session.

FUTURE DIRECTIONS AND CONCLUSIONS

FUTURE WORK

Enhancements

In completion of this initial version of our program, we have opened the door for a number of enhancements to be completed in the future. These include the several requests we were unable to complete from our initial design:

1. An option to save the state of an instance of our program so it can be reloaded and rerun without requiring manual input of all program options
2. DHL file format compatibility so that the users can also calculated statistics and plots based on drill hole data
3. A user control to allow the users an option to make the x-axis a logarithmic-based probability scale

Finally, there have been requests by the user to make the final output more actively dynamic. This includes the ability to resize the graphs after generation, change the colors of bars or lines, or even change the axes scales dynamically. As with the probability scale, this lies far outside the scope of the current project, but could be revisited as a future enhancement.

CONCLUSIONS

User Feedback

After brief feedback meetings and demonstrations with the primary users of the software, the geostatisticians at Newmont Mining Corporation, the initial feedback received regarding the NSAP has been extremely positive. Users feel that it will significantly increase their productivity as it will shorten the amount of time required to generate statistics and plots of any amount of given data.

Preliminary testing feedback has shown that the NSAP could cut users amount of time by about two-thirds, as it does not require the use of complicated command-line driven applications or multiple programs to achieve a single goal. Another large factor in this speed-up is in the quality of the output, which is immensely improved over the simple, command-line charts that are complicated and difficult to understand.

[i] http://www.newmont.com/about
[ii] http://www.infragistics.com/