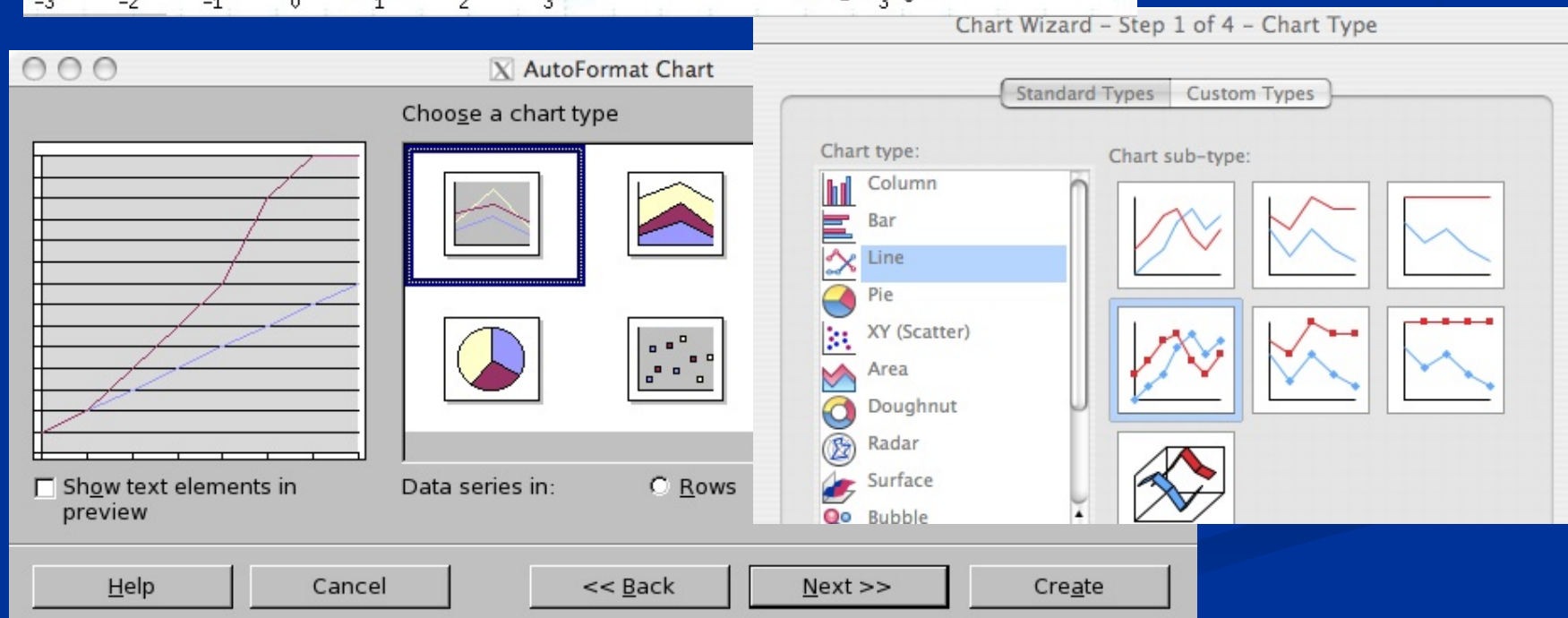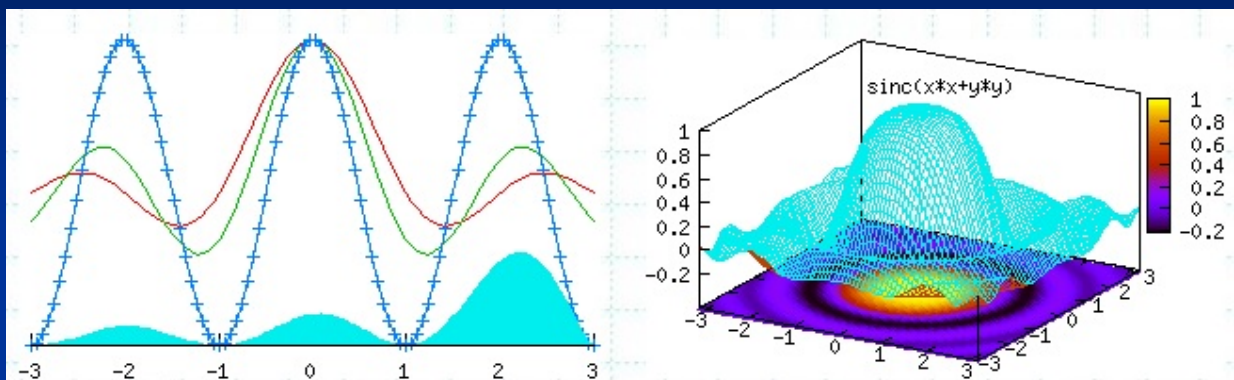# Trivial And Non-Trivial

Data Analysis for Geant4

Paul Guèye, HU

Joseph Perl, SLAC

Analysis Choices

Simplest using text (ASCII) files

- Geant4 does not attempt to provide its own data analysis tools, focusing instead on its central mission as a simulation toolkit.

- The Geant4 user is expected to provide her/his own code to output results to an appropriate analysis format.

- Simplest possible approach = comma-separated values (.csv files)

- The resulting files can be analyzed by tools such as: Gnuplot, Excel, OpenOffice …

- This talk focuses on this approach

Advanced using ntuple files

- Allows to control what plot you want with modular choice of conditions
  Ex: energy of electrons knowing that (= cuts): (1) position/location, (2) angular window, (3) primary/secondary …

- Tools: Root or AIDA-compliant (PI, JAS3 and OpenScientist)

- See: Thursday @ 14:00 (Analysis 2)

- Quick overview at end of talk

Gnuplot

n **"Gnuplot** is a **portable** command-line driven interactive data and function plotting utility for UNIX, IBM OS/2, MS Windows, DOS, Macintosh, VMS, Atari and many other platforms."

n "The software is copyrighted but freely distributed (i.e., you don't have to pay for it)."

n http://www.gnuplot.info/
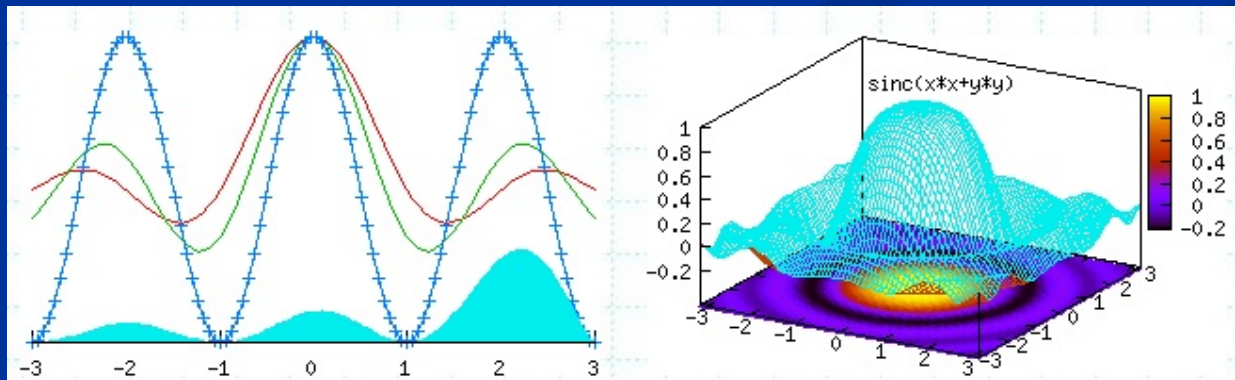
Plotting Data from a .csv File in Gnuplot

- n **plot mydata.csv**

- n **(that's all you need)**

Excel

- Comes from a small company called MicroSoft.
- Part of Microsoft Office.

Importing Data from a .csv File Into Excel

- From the "Data" menu, select "Get External Data"…"Import Text File…"
  - Select the file
  - The Text Import Wizard will come up
  - Check the "Delimited" box
  - Hit "Next"
  - Check the "Comma" box
  - The bottom part of the Wizard should now show your data nicely separated into appropriate columns.
  - Hit "Finish"
  - You will see a dialog titled "Import Data" which will ask where to put the data. Accept the defaults:
    - Existing worksheet:
    - =A1
  - Hit "OK"
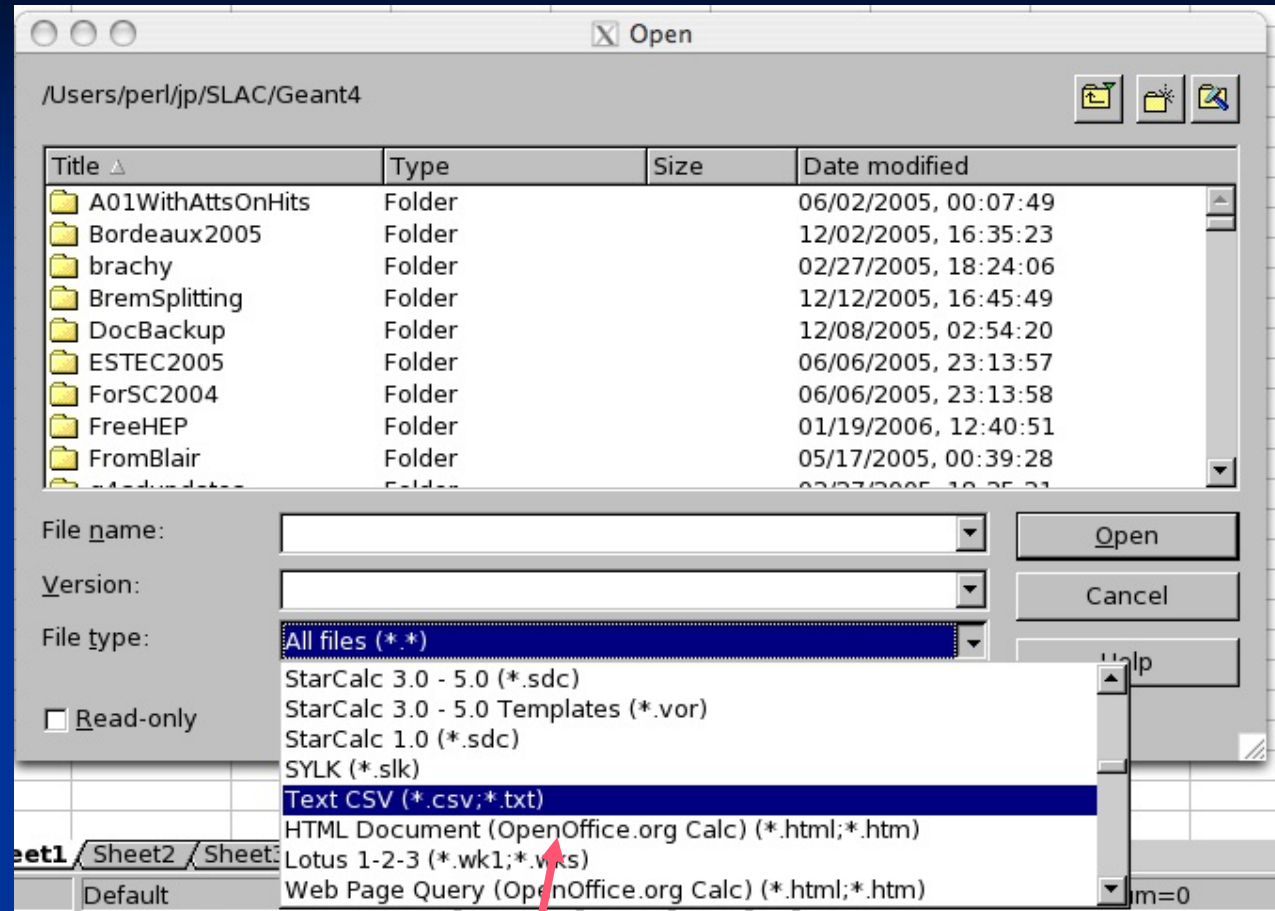- The data should now show up in the spreadsheet.

Plotting with Excel

n Now that the data is in the spreadsheet, you can plot it as follows:

  n Select at least one item in the filled data area.

  n From the "Insert" menu, select "Chart…"

  n Select Chart type: "Line"

  n Hit "Finish"

  n You're done
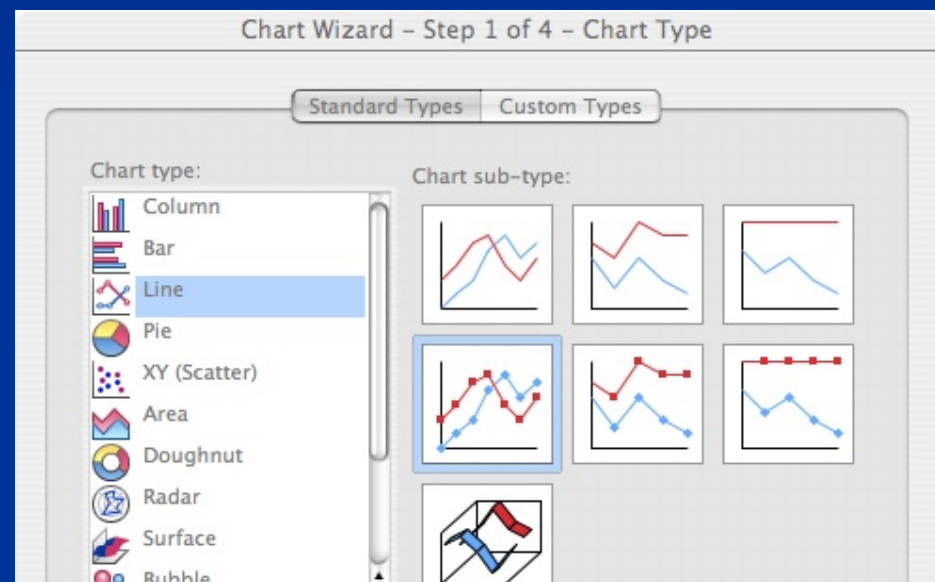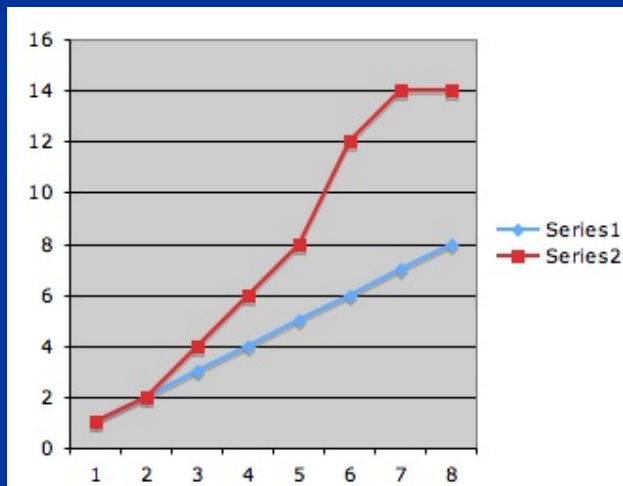
OpenOffice

- OpenOffice.org is a <u>multiplatform</u> and <u>multilingual</u> <u>office suite</u> and an <u>open-source</u> project.

- Compatible with all other major office suites, the product is free to download, use, and distribute.

- http://www.openoffice.org/

Importing Data from a .csv File Into OpenOffice

n From the "File" menu, select "Open…"

    n In the "File type" box, select "Text CSV"
    Note that this is NOT the same as just "Text"

    n Then select your file

Importing Data from a .csv File Into OpenOffice

- n Once you select your file:
    - n The Text Import dialog will come up
    - n Check the "Comma" box
    - n The bottom part of the dialog should now show your data nicely separated into appropriate columns.
    - n Hit "OK"
- n The data should now show up in the spreadsheet.

June 4-8, 2007
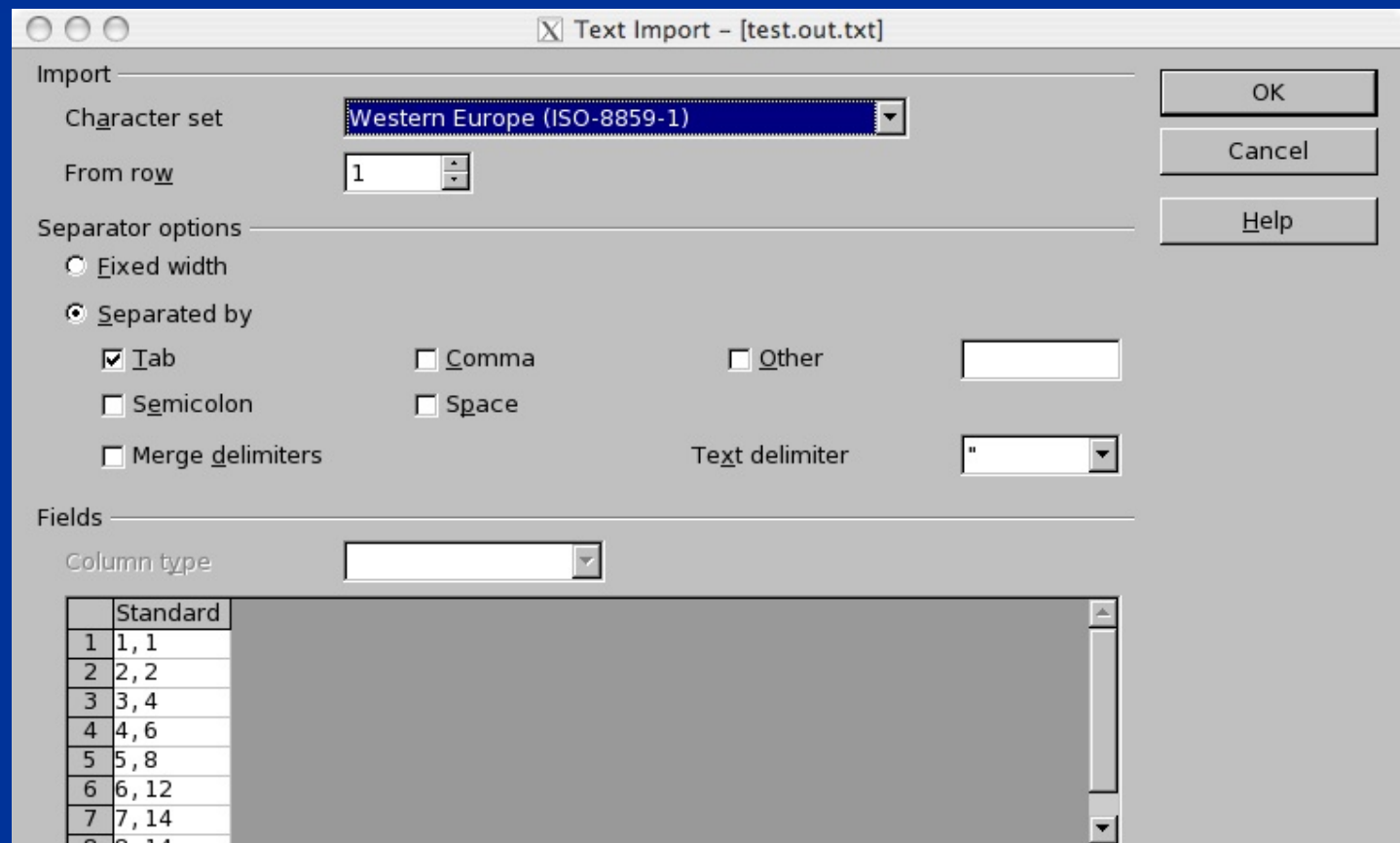
Plotting with OpenOffice

n Now that the data is in the spreadsheet, you can plot it as follows:

- n Select at least one item in the filled data area.
- n From the "Insert" menu, select "Chart…"
- n The defaults should be fine
- n Hit "Next"

- n Select the icon for "Line" chart
- n Hit "Create"
- n You're done
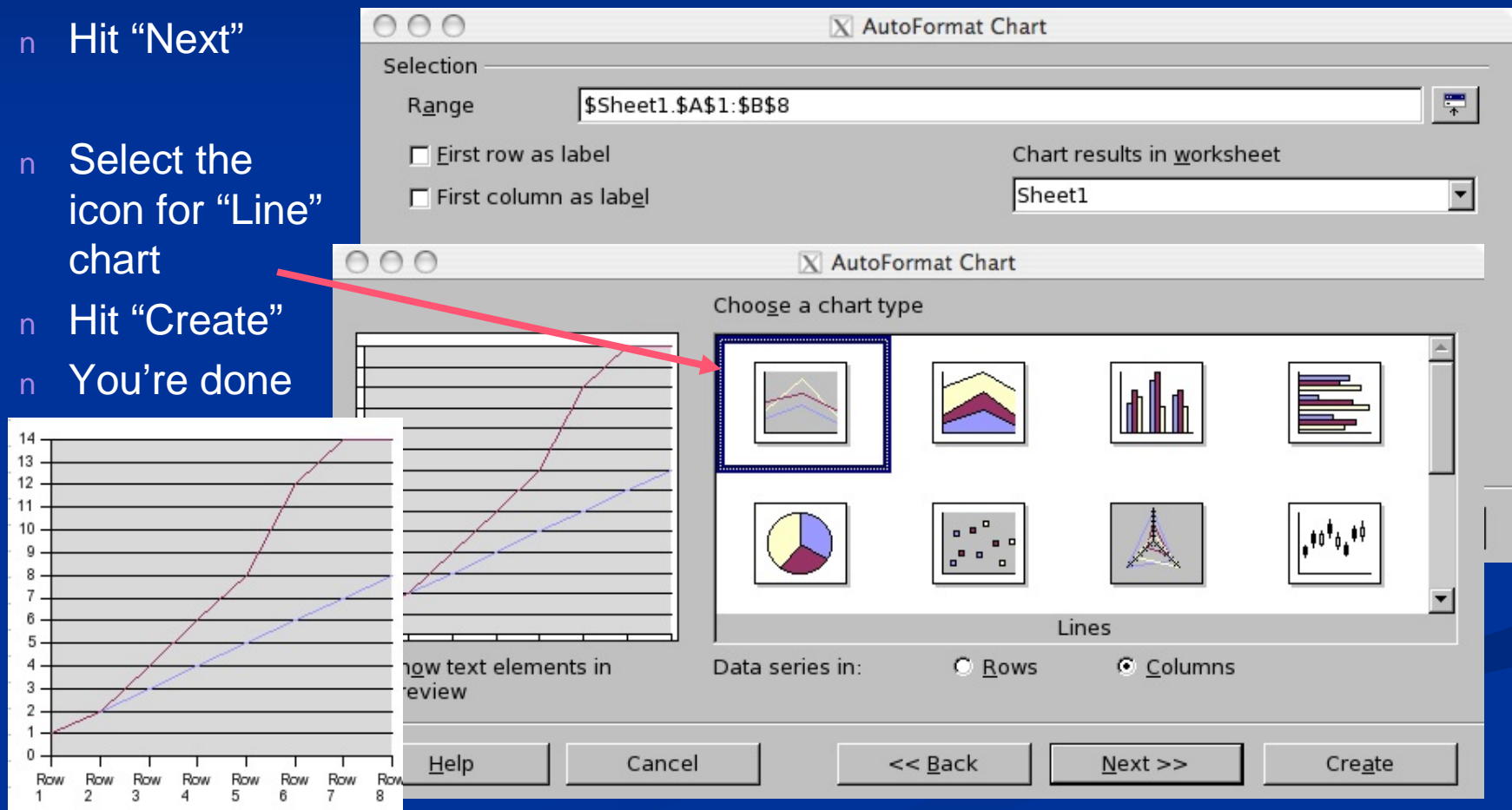
Pointers to More Advanced Data Analysis Options

n   While the Geant4 collaboration does not maintain any data analysis system, many data analysis tool developers have interfaced their systems to Geant4 and each of the following has an active Geant4 user community.

n   A good way to get started thinking about any of these is to look in the Geant4 Analysis Forum:

>   n   http://geant4-hn.slac.stanford.edu:5090/HyperNews/public/get/analysis.html

ROOT

n ROOT is an Object Oriented Data Analysis Framework.

n It is heavily used in High Energy Physics.

n Freely available.

n http://root.cern.ch/

Analysis Interface for Data Analysis (AIDA)

n AIDA is an interface standard used by several advanced data analysis tools. It provides a single standard way to create your data

  n Commands to instantiate an nTuple

  n Commands to fill an nTuple

  n Commands to instantiate a Histogram

  n Commands to fill a Histogram

n Depending on which AIDA-complaint tool you link in to Geant4, these commands will one or the other analysis tool

n You can switch from one AIDA-compliant tool to another without having to change how you book and fill your histograms

  n http://aida.freehep.org/

The Three AIDA-Compliant Analysis Systems

- PI (Physicist Interface)
    - Part of CERN LHC Computing Grid (LCG) Applications
    - http://lcg-pi.web.cern.ch/lcg-pi/
    - C++ and Python

- JAS3
    - http://jas.freehep.org/jas3/
    - Java

- OpenScientist
    - http://openscientist.lal.in2p3.fr/
    - C++

Use Whatever Analysis Tool You Like

- Geant4 does not attempt to provide its own data analysis tools, focusing instead on its central mission as a simulation toolkit.
- The Geant4 user is expected to provide her/his own code to output results to an appropriate analysis format.