









Using Free and Open Source Software for Data Analysis in the Classroom

Chadd Williams
Pacific University

chadd@pacificu.edu



Overview

Goal: What can be done, not how



- Why do I think this is important?
 - You have future CS Majors in your classes!
 - Everyone needs to know how to analyze data

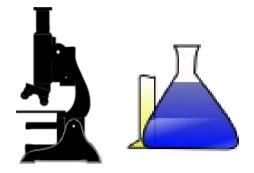
From: What can the computer do?

To: What can I make the computer do?





Process

















Explore Data



Report Data



Open Source Software

- Community developed software
- Application and Source Code given away freely
 - Firefox web browser
 - Pigdin instant messenger
 - OpenOffice office software
 - Linux operating system kernel









(Valid) Concerns

You get what you pay for!



The software cannot do X like Excel can.

Often these are expandable

- Open Office
 - -~8 years old
 - Backed by Sun Microsystems
- SciLab
 - ~ 14 years old
 - Backed by Consortium of companies/research labs

Plugins or downloadable toolbox

Software

- OpenOffice.org
 - Free Open Source Software
- Google Docs
 - Free but not open source
- SciLab
 - Free Open Source Software
- Many others
 - IBM Lotus Symphony
 Peepel
 - NeoOffice (Mac)ThinkFree
 - Maxima



OpenOffice.org

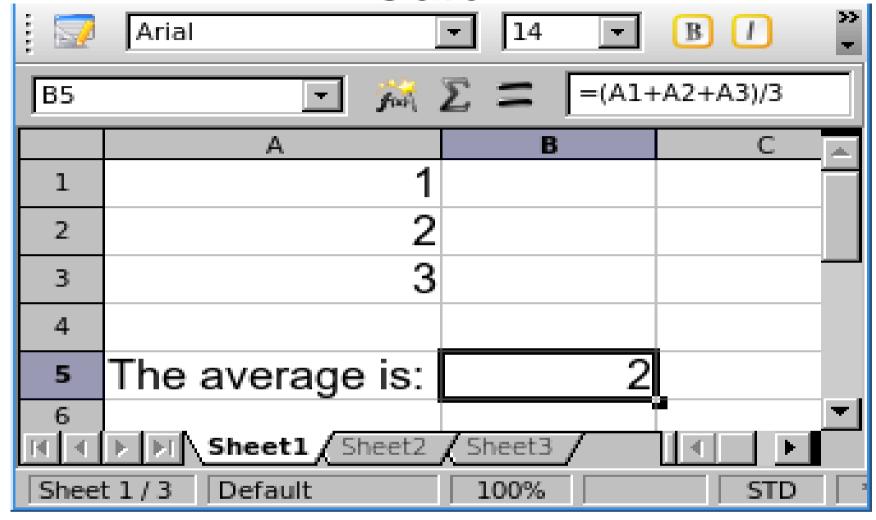
- Microsoft Office replacement
- Applications
 - Writer word processor



- Calc spreadsheet
 - Impress presentation maker
 - Draw diagram builder
 - Base database
 - Math equation editor

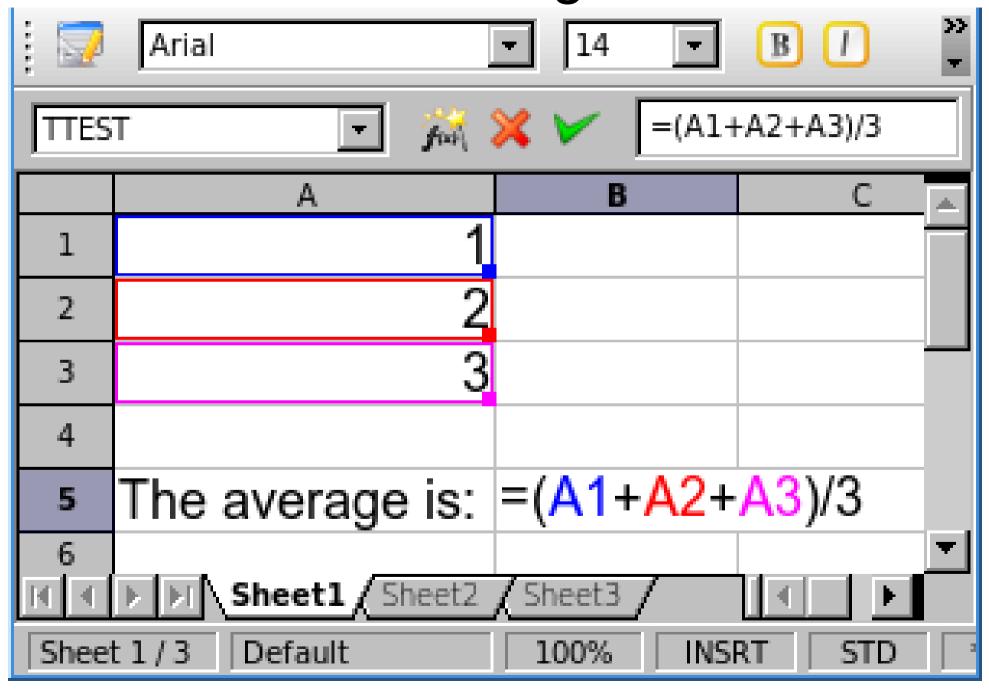


Calc



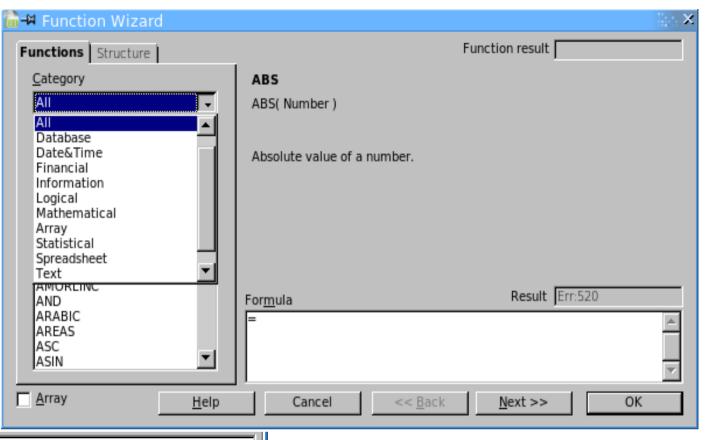
- Each cell can contain
 - Data
 - Formula to manipulate the data

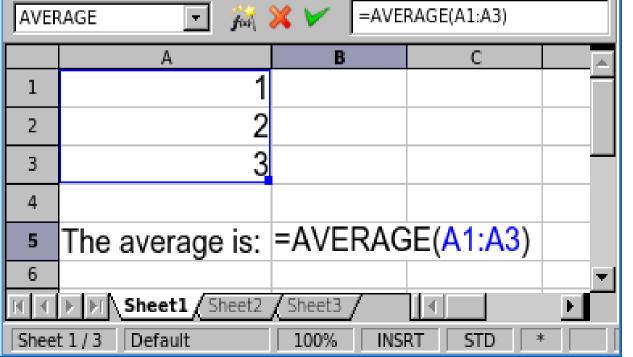
Referencing Data



Functions

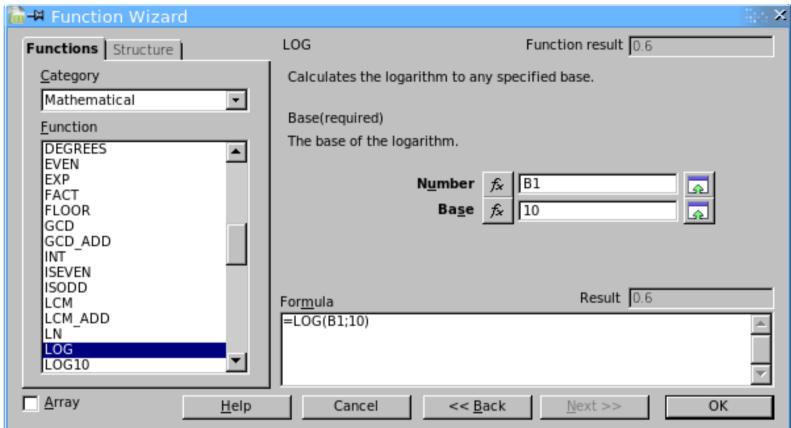
 Large number of pre-defined functions are available



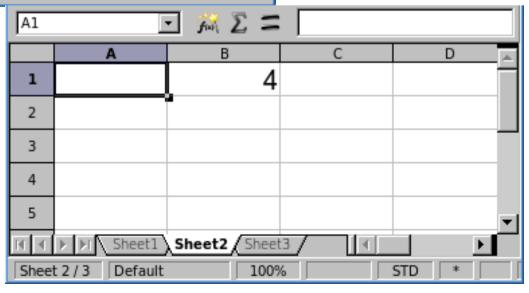


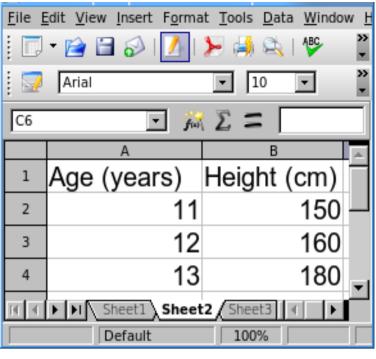
- A1:A3
 - All cells from A1 to A3
- Formulas always begin with =

Inserting a Function

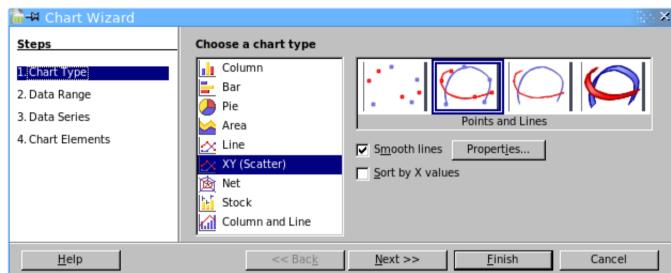


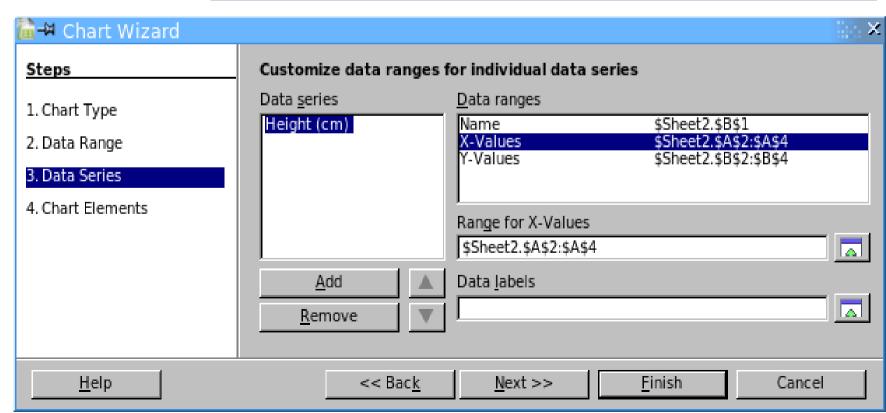
- Valid entries
 - Cell (B1)
 - Number (10)
 - Function

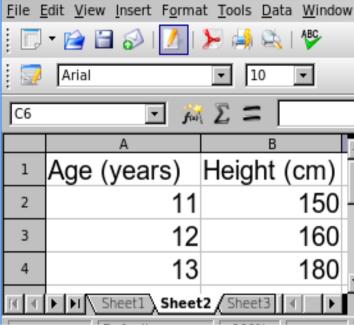




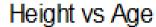
Charts

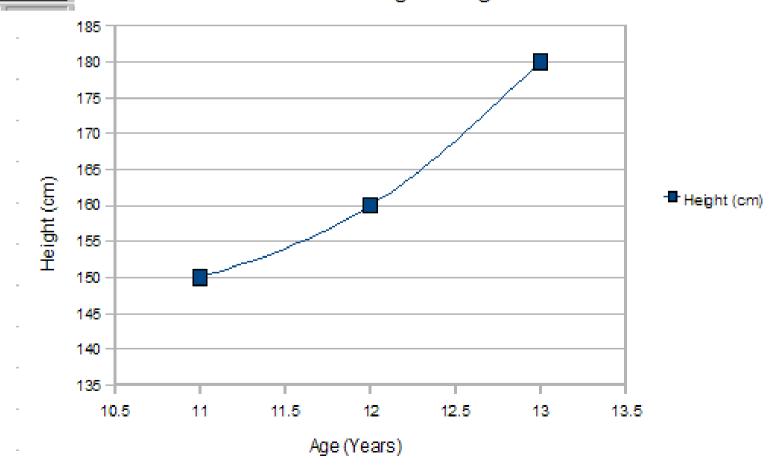




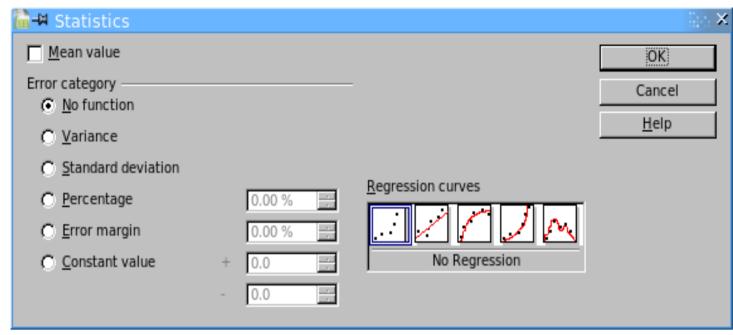


Result

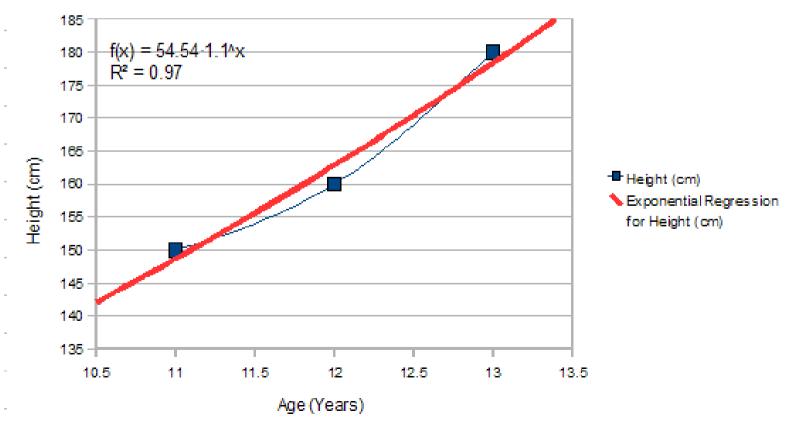




Regression

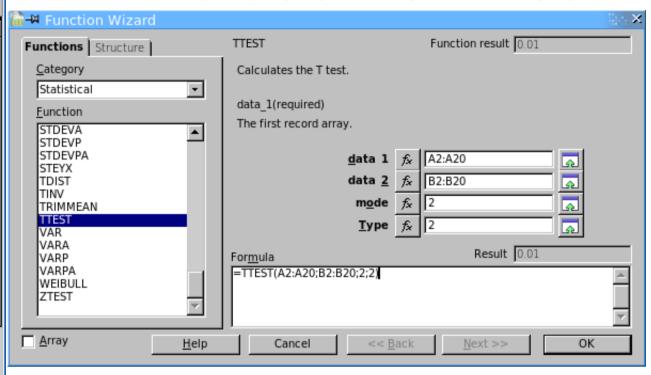


Height vs Age



MODE =TTEST(A2:A20;B2:B20;2;2) Exam1 Exam2 58 Exam1 58 Median 58 Mode 62 Exam2 100 Median 71 Mode Average Average 83.16 72.74 =TTEST(A2:A20;B2:B20;2;2)_ T-Test

Student's T-Test

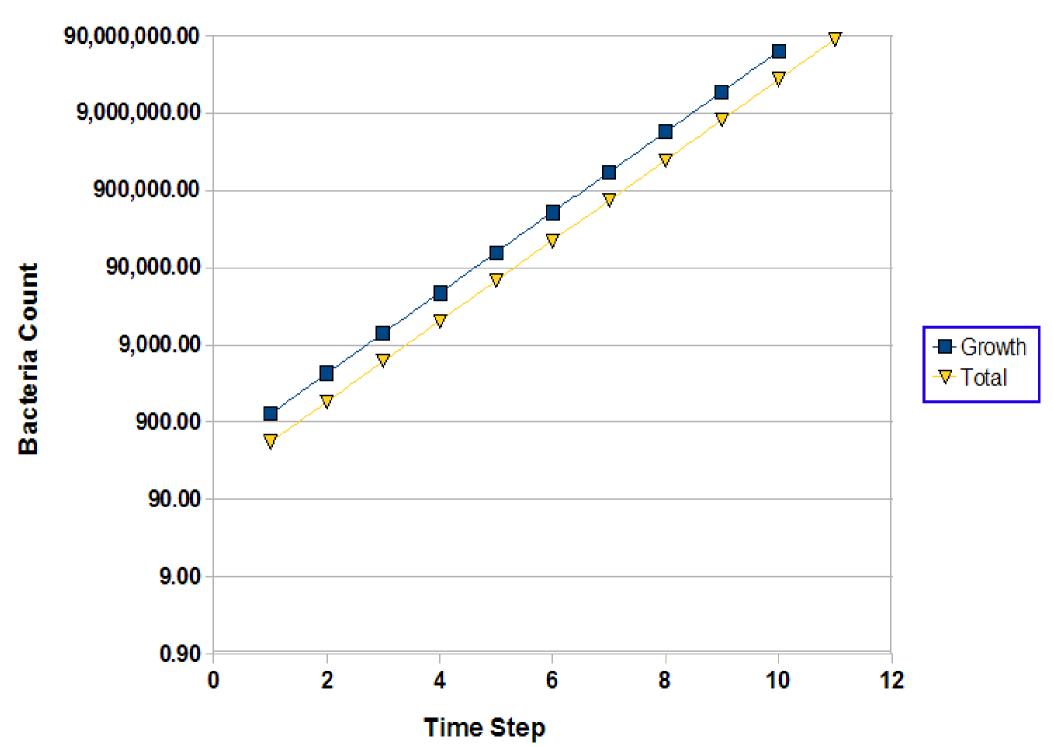


- Other Stat tests
 - Chi²
 - F Test
 - Pearson

| | | Δ | | B | | С | D | |
|----|-----|---------------------|-------|------------------|---------|------------|------------|------|
| 1 | Ва | cteria | Popul | ation(t) = A * e | e^(k*t) | - | | |
| 2 | | | • | | | | | |
| 3 | nit | tial population (A) | Growt | th Factor (k) | | | | |
| 4 | | 500 | | 1.01 | | | | |
| 5 | | | | | | | | |
| 6 | \ | / 1 \ | سنس م | - Danulation | Orough | | ρ. | |
| 7 | | Initial populatio | n (A) | Growth Fac | tor (k) | | | |
| 8 | | | 500 | | 1.01 | | | |
| 9 | | | | | | | | |
| 10 | | Time (t) | | Staring Pop | ulation | Growth | | Po |
| 11 | | | 0 | | 500.00 | | 0.00 | |
| 12 | | | 1 | | | | P(B\$4*A8) | |
| 13 | | <u> </u> | 400 | 000 00 | | | | |
| 14 | | 4 | ,433, | ,093.03 | 11 | 38411.68 |)]=B17+C | 17 |
| 15 | | ď | | 500,U/4.UZ | Ί, | UZ0,54Z.0U | 1,014,010 | 0.02 |
| 16 | | Q | | 1,614,616.62 | 2, | 818,476.41 | 4,433,093 | 3.03 |
| 17 | | 1 | | 4,433,093.03 | 7, | 738,411.69 | 12,171,504 | 1.71 |

| | A | В | С | D |
|----|------------------------|--------------------------------|--------------|---------------|
| 1 | Bacteria | Population(t) = $A * \epsilon$ | e^(k*t) | |
| 2 | | | | |
| 3 | Initial population (A) | Growth Factor (k) | | |
| 4 | 500 | 1.01 | | |
| 5 | | | | |
| 6 | Time (t) | Staring Population | Growth | Population(t) |
| 7 | 0 | 500.00 | 0.00 | 500.00 |
| 8 | 1 | 500.00 | 872.80 | 1,372.80 |
| 9 | 2 | 1,372.80 | 2,396.36 | 3,769.16 |
| 10 | 3 | 3,769.16 | 6,579.45 | 10,348.62 |
| 11 | 4 | 10,348.62 | 18,064.56 | 28,413.17 |
| 12 | 5 | 28,413.17 | 49,598.06 | 78,011.23 |
| 13 | 6 | 78,011.23 | 136,176.49 | 214,187.72 |
| 14 | 7 | 214,187.72 | 373,886.30 | 588,074.02 |
| 15 | 8 | 588,074.02 | 1,026,542.60 | 1,614,616.62 |
| 16 | 9 | 1,614,616.62 | 2,818,476.41 | 4,433,093.03 |
| 17 | 10 | 4,433,093.03 | 7,738,411.69 | 12,171,504.71 |

Bacteria Growth



Google Docs

- http://docs.google.com
 - Requires free Google account
- On-line Office Suite
 - Word Processor



Spreadsheet



Presentation Maker





Runs in your web browser



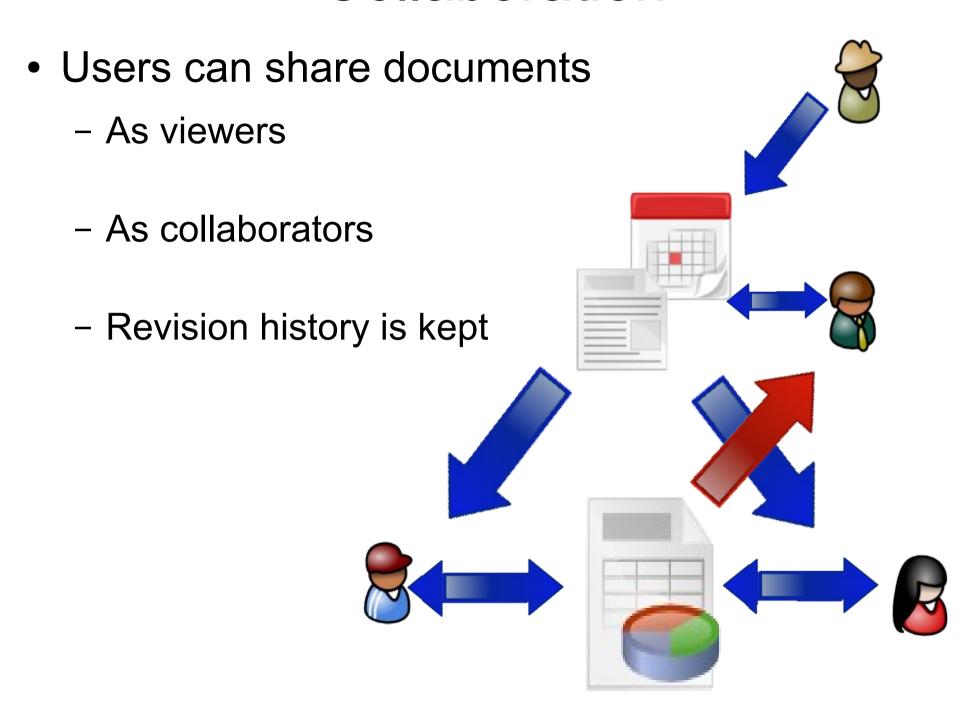
The difference...

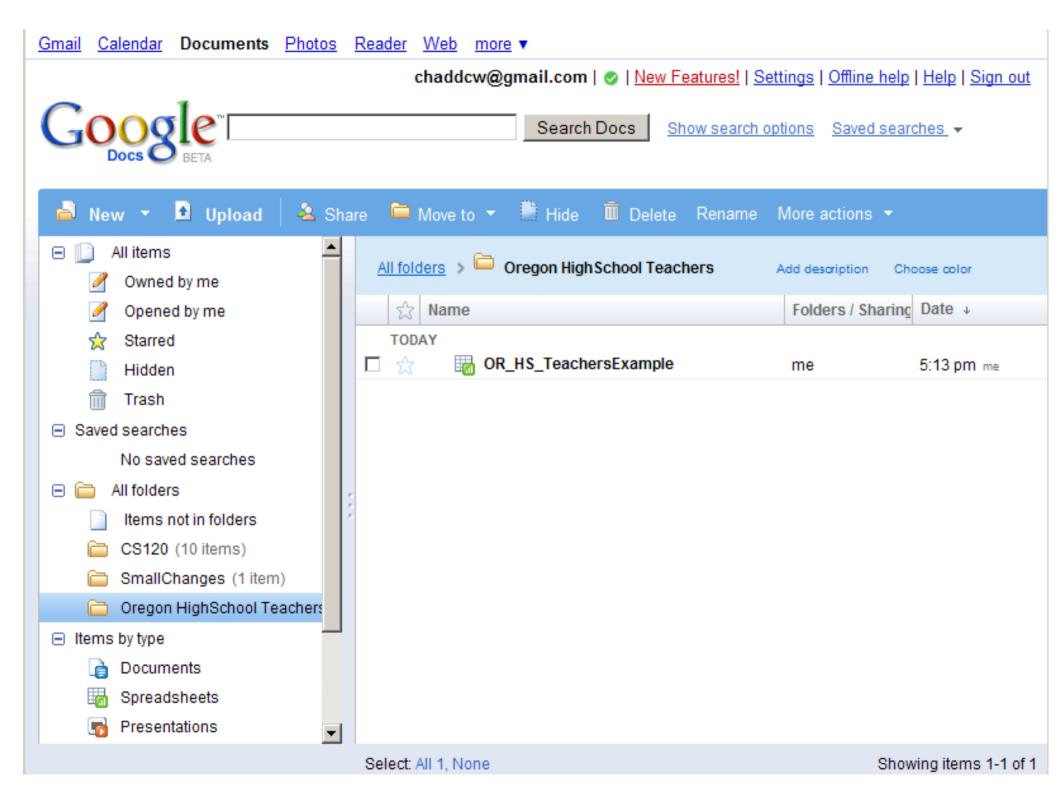
Data stored at Google

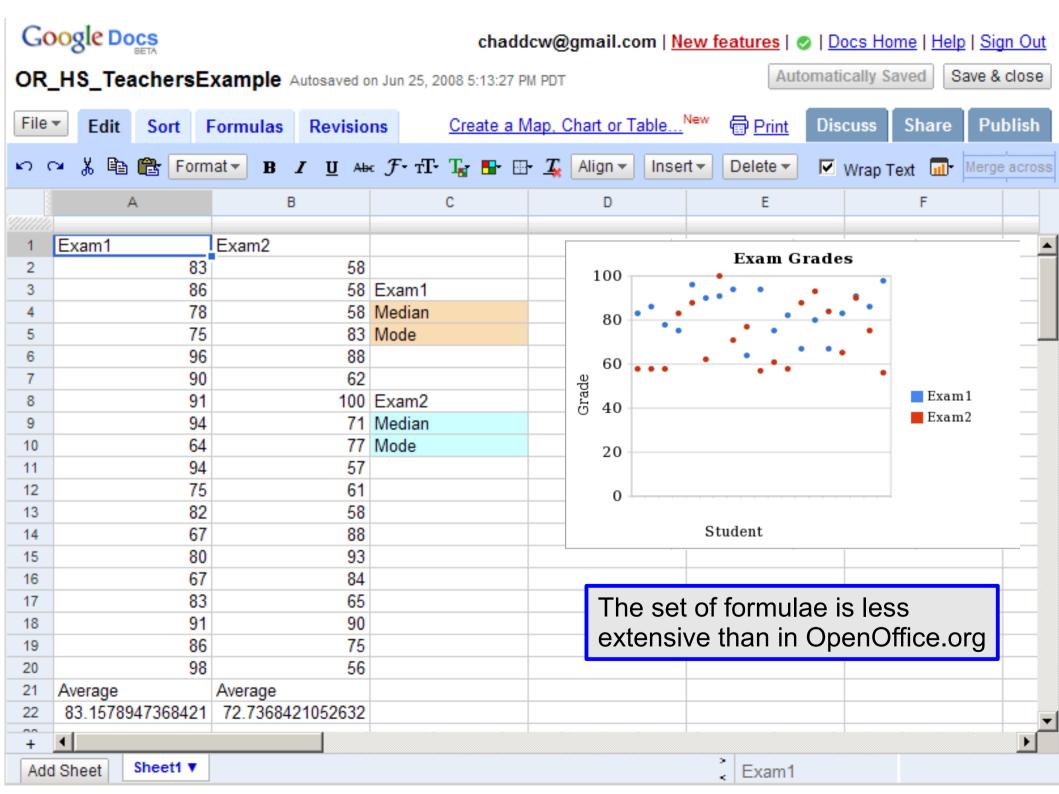


- Documents can be shared and edited
- Privacy concerns
- Documents can be downloaded to your home or school computer
- http://www.google.com/educators/p_docs.html

Collaboration









chaddcw@gmail.com | New features | ♥ | Docs Home | Help | Sign Out

OR_HS_TeachersExample Autosaved on Jun 25, 2008 5:13:27 PM PDT

Automatically Saved

Save & close



Edit

Sort

Notify me at chaddcw@gmail.com

Every day, with a summary of changes 🔻

Formulas Revisions

Create a Map, Chart or Table... New

′ (

Print

Discuss

Share

Publish

×

Share this document

| Invite people: | | | | | | | | |
|---|------------|--------------------|--|--|--|--|--|--|
| as collaborators | as viewers | to fill out a form | | | | | | |
| | | | | | | | | |
| Separate email addresses with commas | | | | | | | | |
| Choose from contacts | | | | | | | | |
| Invite collaborators | | | | | | | | |
| Advanced Options | | | | | | | | |
| ✓ Collaborators can invite others | | | | | | | | |
| Invitations may be used by anyone Allows mailing lists Learn more | | | | | | | | |
| Anyone can view this document WITHOUT LOGGING IN at: | | | | | | | | |

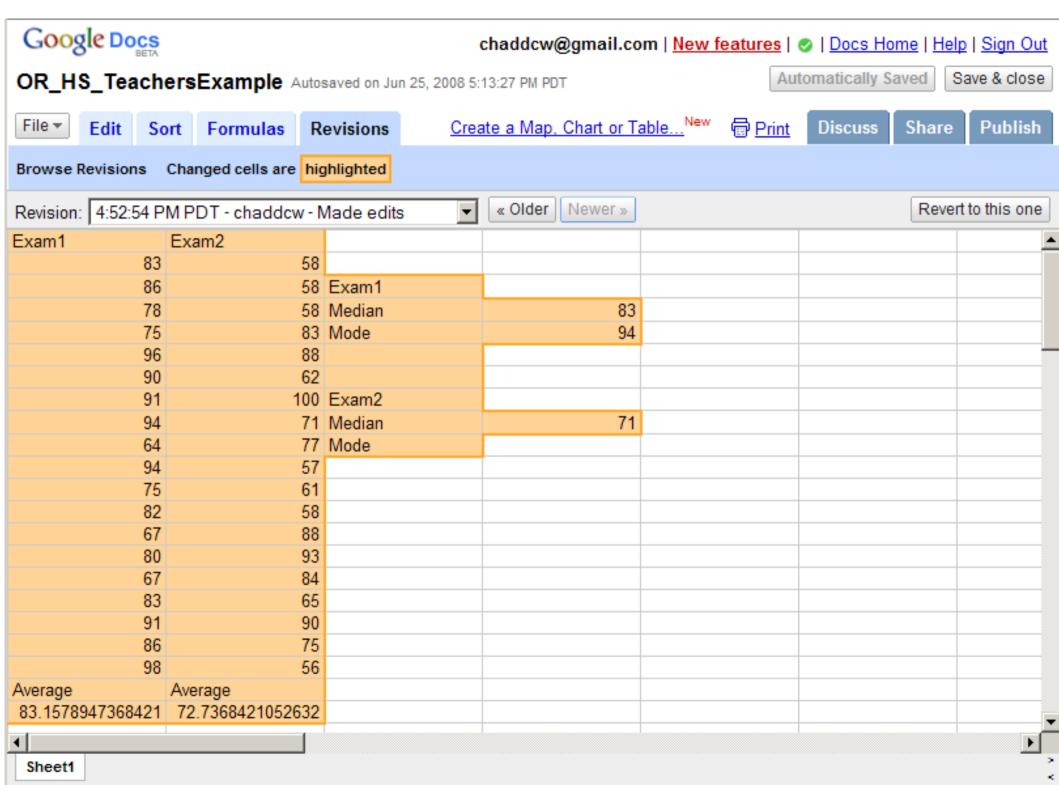
This document is not shared.

Collaborators (0)

Collaborators may edit the document and invite more people.

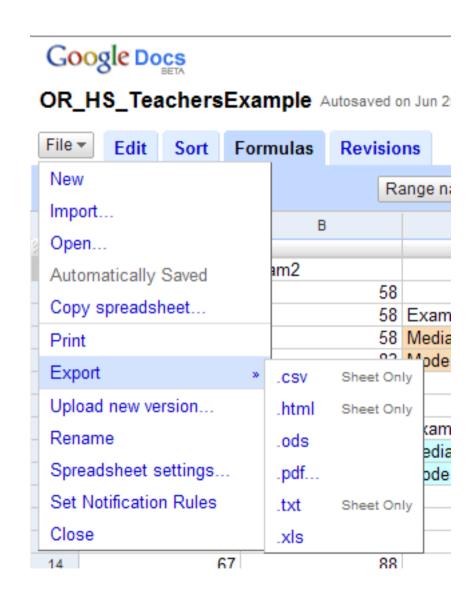
Viewers (0)

Viewers may see the document but not edit it.



Backup your Data!

- You data is stored on Google's computers
- Save them to your computer using Export
 - .ods OpenOffice.org
 - .xls Microsoft Excel
 - .pdf Acrobat Reader



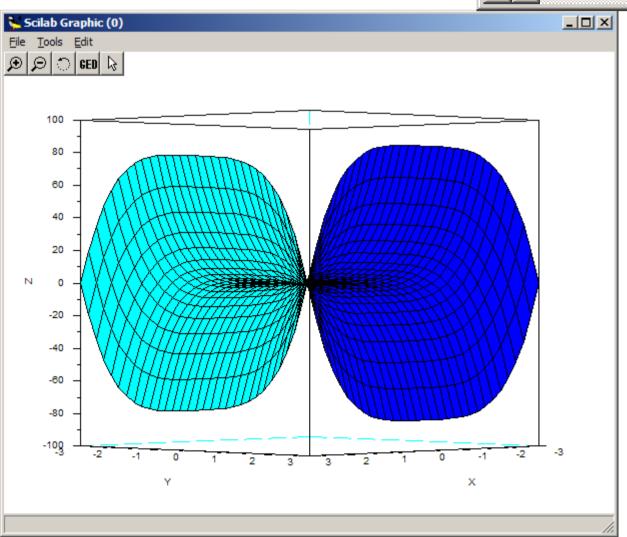
SciLab



- Numerical computation package
 - Similar to MATLAB
 - Write small programs to analyze data
 - Fortran & C interface
 - Many built in functions
 - Statistics
 - Linear algebra
 - FFT
 - Sound analysis
 - Image processing

```
M scilab-4.1.2 (0)
   Edit Preferences Control Editor Applications ?
                           scilab-4.1.2
                   Copyright (c) 1989-2007
               Consortium Scilab (INRIA, ENPC)
Startup execution:
  loading initial environment
-->x=1:
-->y=2;
-->X+V
 ans =
    3.
-->X = [0 1 2 3 4 5];
-->Y = [0 2 4 6 8 10]:
-->X+Y
 ans
       3. 6. 9.
                              12.
                                      15.
-->X * Y'
 ans
    110.
```

```
-->deff('z=f(x,y)', 'z=x^4-y^4')
-->x=-3:0.2:3 ;y=x ;
-->clf() ;fplot3d(x,y,f)
-->_
```



Resources

These slides will be on the wiki!

chadd@pacificu.edu

http://www.openoffice.org

http://docs.google.com



http://www.scilab.org