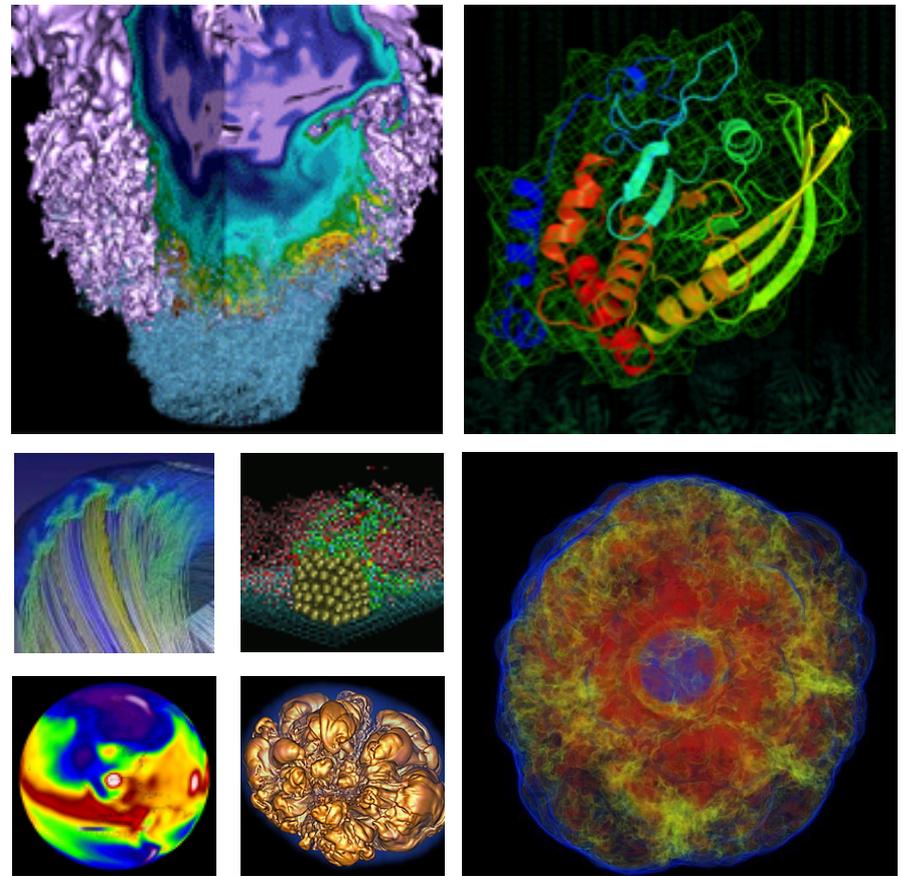


Energy technologies meets computing and data analysis

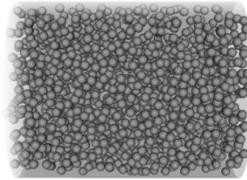
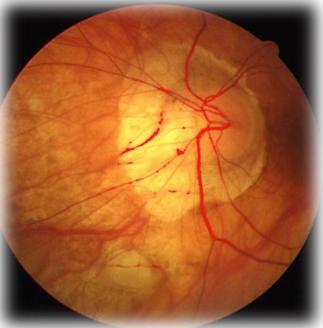
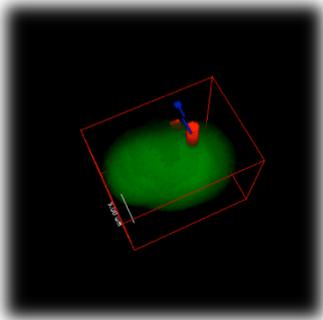


Dani Ushizima

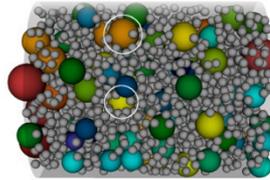
Deputy Team Leader
Visualization Group
CRD, LBL

dushizima@lbl.gov /
dani.lbnl@gmail.com

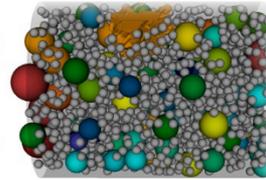
Dani's Projects



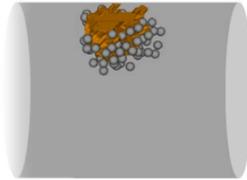
(a) Packed spheres



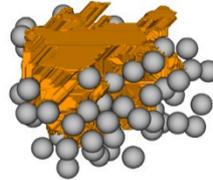
(b) Highlighting "pocket spheres"



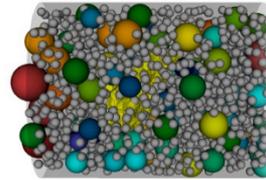
(c) Pocket boundary of the circled orange sphere



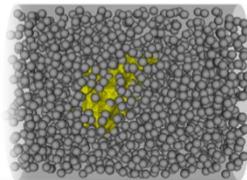
Restricting packed spheres to region around pocket



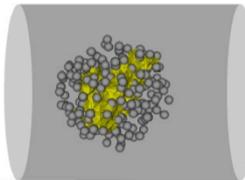
(e) Close-up view of pocket



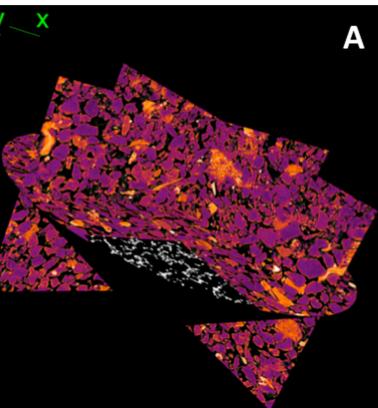
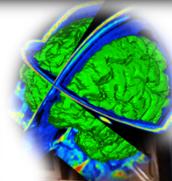
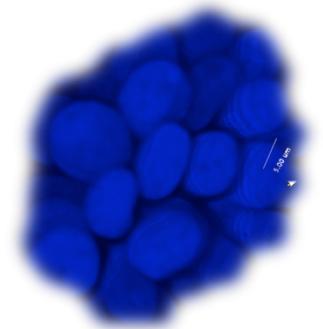
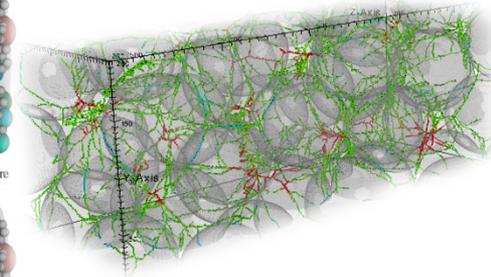
(f) Pocket boundary of the circled yellow sphere



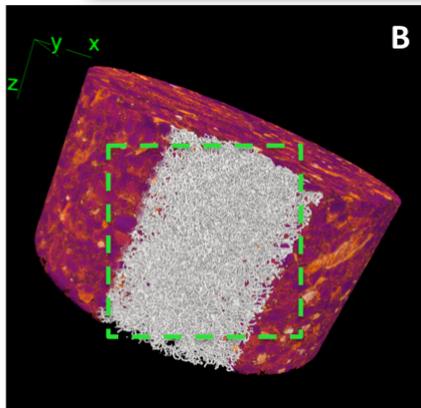
Restricting packed spheres to region around inner pocket



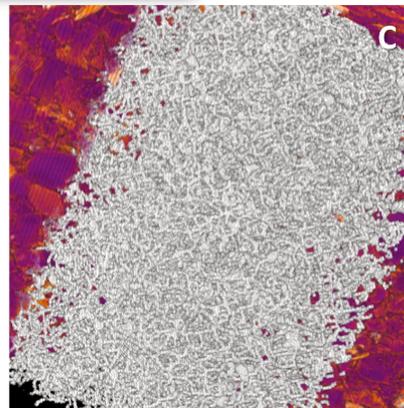
(f) Close-up view of inner pocket



A



B



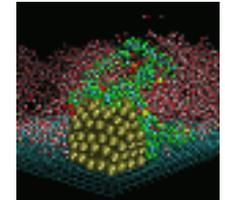
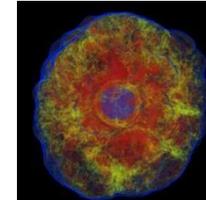
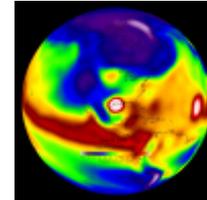
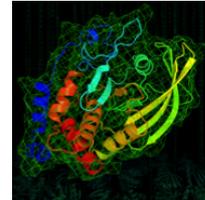
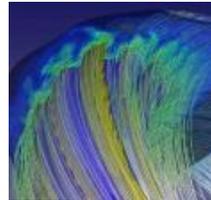
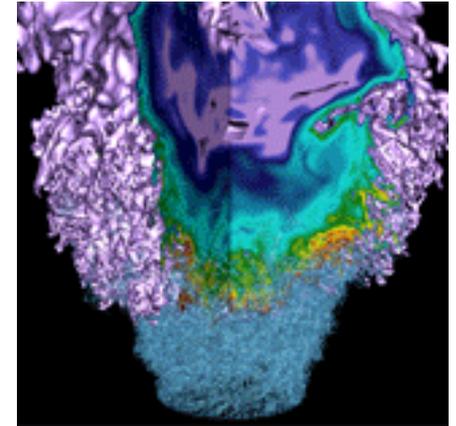
C



BERKELEY LAB

LAWRENCE BERKELEY NATIONAL LABORATORY

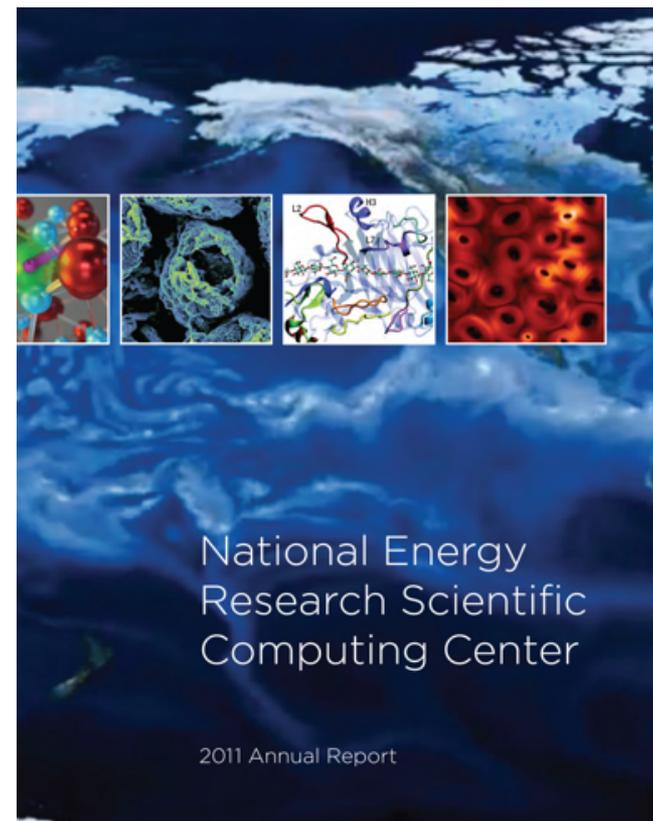
Q&A with EETD



1. What are the supercomputing resources?
2. What are the different approaches for accessing them, and how do they each work?
3. How do we get time on the computing resources?
4. How does one decide whether a task would be suitable for a supercomputer? Is the benefit more from computation time or the size of data storage?
5. What other computing resources are available at LBNL?
6. What does the visualization group do? How might this group be useful for our work?
7. Are there any statistics consulting resources available at LBNL?

1. What are the supercomputing resources?

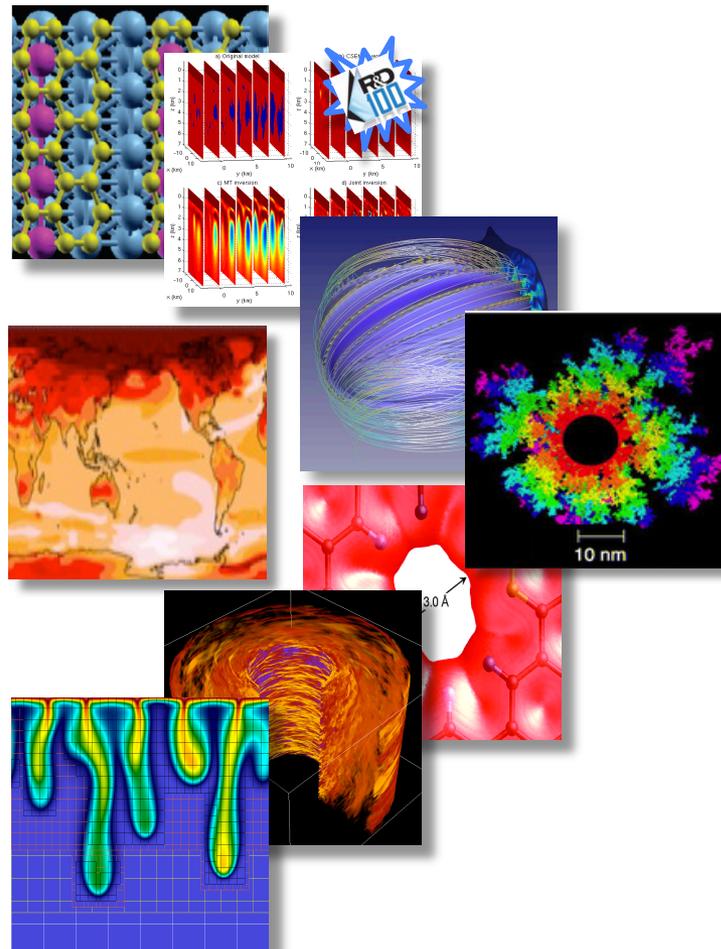
- National Energy Research Scientific Computing Center:
 - Established 1974, first unclassified supercomputer center;
 - Original mission: to enable computational science as a complement to magnetically controlled plasma experiment;
- **Today's mission:** Accelerate scientific discovery at the DOE Office of Science through high performance computing and extreme data analysis.



NERSC: Production Computing for the DOE Office of Science

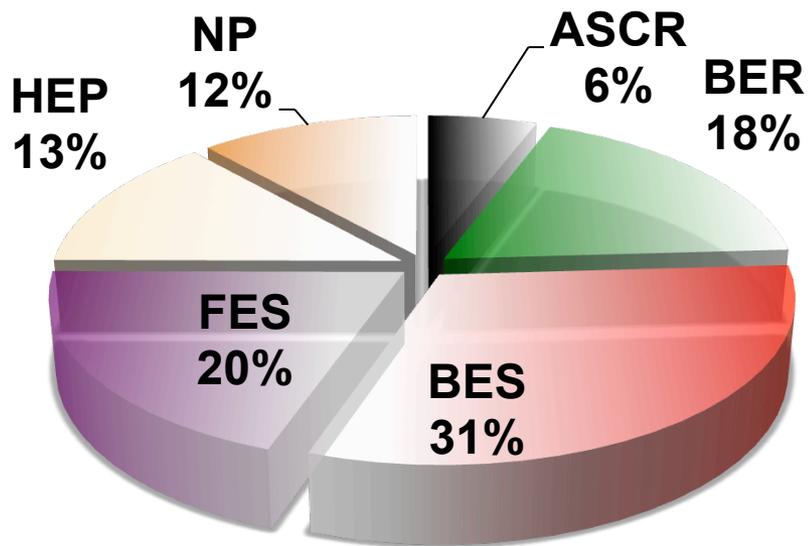


- Diverse workload:
 - 4,500 users, 600 projects;
 - 700 codes; 100s of users daily;
- Allocations controlled primarily by DOE;
 - 80% DOE Annual Production awards (ERCAP):
 - From 10K hour to ~10M hour;
 - Proposal-based; DOE chooses;
 - 10% DOE ASCR Leadership Computing Challenge;
 - 10% NERSC reserve (“NISE”).



2. What are the different approaches for accessing them, and how do they each work?

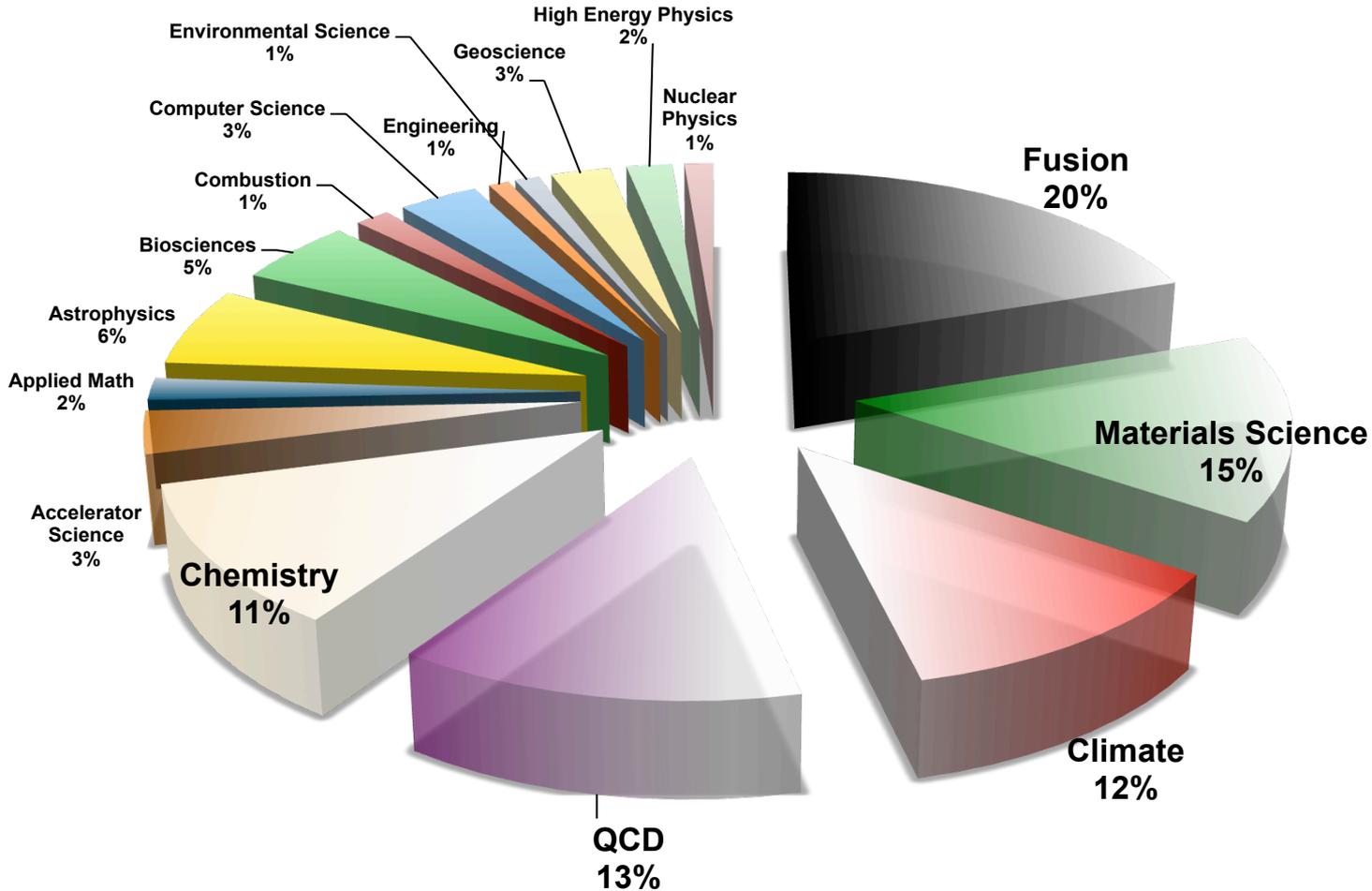
DOE View of Workload



ASCR	Advanced Scientific Computing Research
BER	Biological & Environmental Research
BES	Basic Energy Sciences
FES	Fusion Energy Sciences
HEP	High Energy Physics
NP	Nuclear Physics

NERSC 2013 Allocations By DOE Office

Science View of Workload



**NERSC 2013 Allocations
By Science Area**

3. How do we get time on the computing resources?

- <https://www.nersc.gov/users/accounts/allocations/request-form/>
- Requests to use NERSC resources are submitted annually via a web form known as the ERCAP (Energy Research Computing Allocations Process) Request Form.
- If you are **new** to NERSC and want to apply for an allocation as a **new Principal Investigator** please read Applying for your First NERSC Allocation.
- To access ERCAP, point your browser to nim.nersc.gov, and login to NIM. Click the My Stuff menu and select My ERCAP Requests.
- In general existing projects should be renewed for the next allocation year - you should not start a new request if you have an existing repository. Renewing carries forward most of the responses to the request questions and preserves the project's user list.

4. How does one decide whether a task would be suitable for a supercomputer?
Is the benefit more from computation time or the size of data storage?

NERSC Computing and Storage Capabilities



Large-Scale Computing Systems

Hopper (NERSC-6): Early Cray Gemini System

- 6,384 compute nodes, 153,216 cores
- 144 Tflop/s on applications; 1.3 Pflop/s peak



Edison (NERSC-7): Early Cray Aries System (2013)

- Over 200 Tflop/s on applications, 2 Pflop/s peak
- 333 TB of memory, 6.4 PB of disk



Midrange

275 Tflops peak



Carver

- IBM iDataplex cluster
- 10740 cores; 132TF

PDSF (HEP/NP)

- ~2300 core cluster; 30TF

GenePool (JGI)

- ~8200 core cluster; 113TF
- 2.1 PB Isilon File System

NERSC Global Filesystem (NGF)

Uses IBM's GPFS

- 8.5 PB capacity
- 15GB/s of bandwidth



HPSS Archival Storage

- 240 PB capacity
- 5 Tape libraries
- 200 TB disk cache



Analytics & Testbeds



IBM x3850

1TB, 2TB nodes

Dirac 50 Nvidia GPU nodes

Jesup IBM iDataPlex

Data Analytics; HTC

5. What other computing resources are available at LBNL?

Other resources at LBNL



IT Services

<https://commons.lbl.gov/display/itdivision/IT+Services>

[Archives and Records](#)

- [Records Storage](#)
- [Records Retrieval](#)
- [R&D Records Outreach Project](#)
- [Business Records Processing](#)
- [Records Management Training Workshops](#)

[Business Systems](#)

- [HR Systems](#)
- [Financial Systems](#)
- [Facilities Systems](#)
- [EHS Systems](#)
- [Other Business Systems](#)

[Computing Facilities](#)

- [Data Center Co-location facility](#)

[E-mail, Collaboration, Video, and Web Services](#)

- [Gmail](#)
- [Google Calendar](#)
- [Google Sites](#)
- [Google Docs](#)
- [Surveys and Google Forms](#)
- [Other Google Services](#)
- [Commons](#)
- [Mailing Lists](#)
- [eRoom](#)
- [LabAlert Emergency Msgs](#)
- [Lecture Capture](#)
- [Project Mgt Tools](#)
- [Video and Audio Conferencing](#)
- [Video Streaming Overview](#)
- [Web Hosting](#)

[Scientific Computing](#)

- [Advice, Consulting, Design](#)
- [Institution-Owned Cluster \(Lawrencium\)](#)
- [Linux Cluster Support \(SCS\)](#)
- [Scientific Workstation Support \(Linux\)](#)
- [UNIX Support](#)

[Networking](#)

- [Consulting and Advice](#)
- [DNS - Domain Name Service](#)
- [Firewalls](#)
- [Wired](#)
- [NTP - Network Time Protocol](#)
- [Remote Access](#)
- [VPN - Virtual Private Network](#)

[IT Help Desk](#)

[Passwords and User Accounts](#)

- [Account Requests](#)
- [Federated Authentication](#)
- [UCB Identities](#)
- [Web Authentication](#)

[Cyber Security](#)

- [Central System Logging](#)
- [One Time Passwords](#)
- [Computer Blocking \(Computer Isolation Removal\)](#)
- [Computer Vulnerability Scanning](#)
- [Certificates \(SSL\)](#)
- [Cyber Incident Handling](#)

[Mobile and Telephones](#)

- [Cellular Phone Service](#)
- [Smartphone & Tablet Info](#)
- [m.lbl.gov](#)
- [Cellular Data](#)
- [Desk Phones](#)
- [Calling Cards](#)
- [Conference Calling](#)
- [Pagers](#)
- [Voice Mail](#)

[Workstation Support](#)

- [Backups](#)
- [Workstation Acquisition and Advice](#)
- [Windows File Service \(CIFS\)](#)
- [Operations Desktop Support](#)

6. What does the
visualization group do?
How might this group be
useful?

Vis/Analytics Group



Wes Bethel, Prabhat, Gunther Weber, Daniela Ushizima, Hank Childs, Jeniiffer Horsman, Dmitriy Morozov, Joerg Meyer, Oliver Ruebel, Burlen Loring, David Camp, Hari Krishnam, Yushu Yao, Peter Nugent

Mission: scientific data understanding through visualization and analytics.

7. Are there any
statistics consulting
resources available
at LBNL?

Vis/Analytics Group



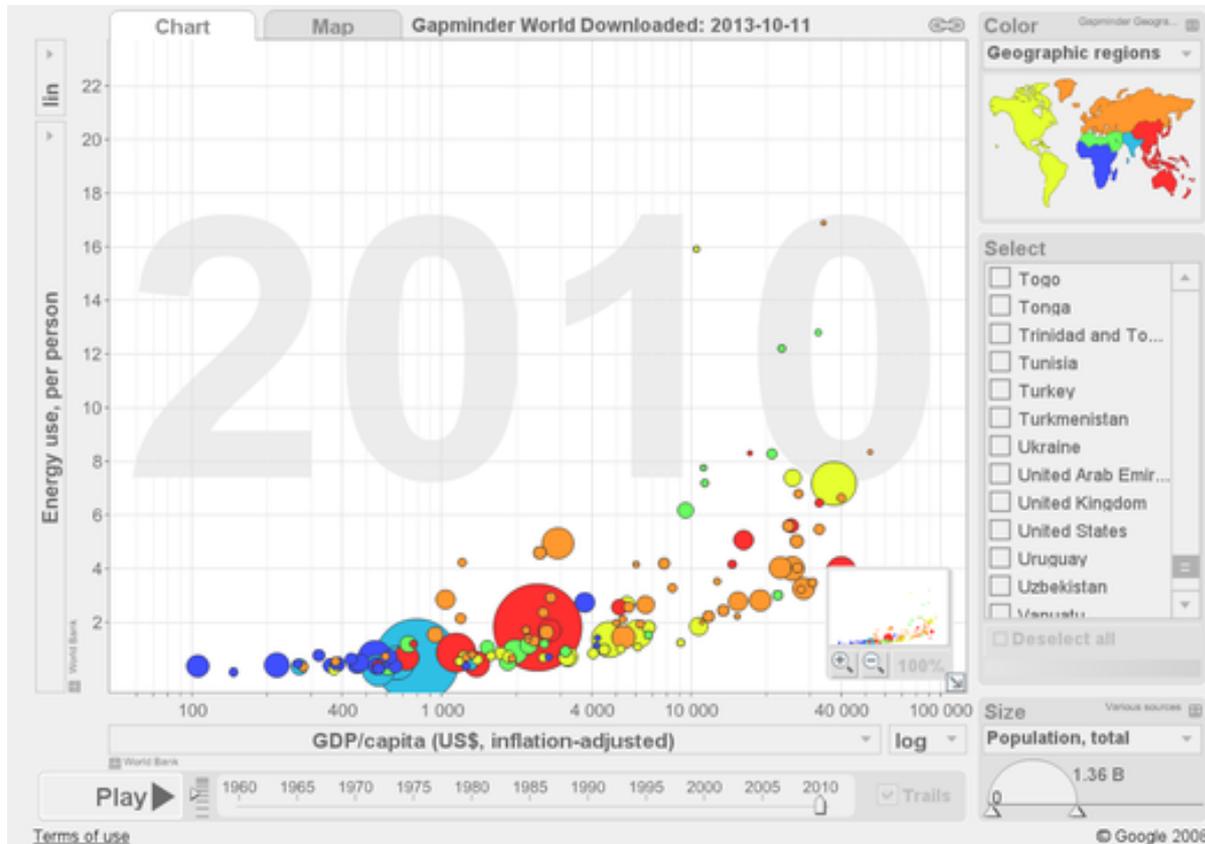
Wes Bethel, Prabhat, Gunther Weber, Daniela Ushizima, Hank Childs, Jeniiffer Horsman, Dmitriy Morozov, Joerg Meyer, Oliver Ruebel, Burlen Loring, David Camp, Hari Krishnam, Yushu Yao, Peter Nugent

Mission: scientific data understanding through visualization and analytics.

Statistical analysis and Visualization

Quick Tutorial

1. Googlevis – interactive statistics



- Trendalyzer in Sweeden: interactive charts;
- “Motion chart” and “geo map”
- Creating complex chart layouts with *gvisMerge*

1.1. Hands-on - Using it in R

Chart type
Change between bubble, bar and line chart.

Lin / Log scale
X- and y-axis scales can be linear or logarithmic. A log scale can make it easier to see trends.

Y-axis
Click here to select indicators for the y-axis.

Speed of animation
Drag to change the speed of the animation.

Play / Stop
Click Play/Stop to control the animation. (How the graph changes over time.)

Time
Click and drag to change year.

X-axis
Click here to select indicators for the x-axis. You can also choose to display time on this axis.

To zoom in:
1. Put your mouse in the chart area.
2. Hold down the left mouse button and draw a rectangle over the items that you want to zoom in.
3. Release the left mouse button.
4. In the menu that pops up, select 'Zoom in'.

To zoom out:
Click the 'Zoom out' link above the zoom thumbnail in the right panel.

Colour
Click to choose another indicator for colour.

Size Indicator
Select the indicator which represents the size of the bubble.

Select variables
Click boxes to select specific variables. (You can also click the bubbles.)

Trails
Click Trails to follow a selected country while the animation plays.

Settings
Change opacity of non selected items and further advanced settings.

Expenses

Sales

2010-07-26 22:03:08
R version 2.11.1 (2010-05-31)

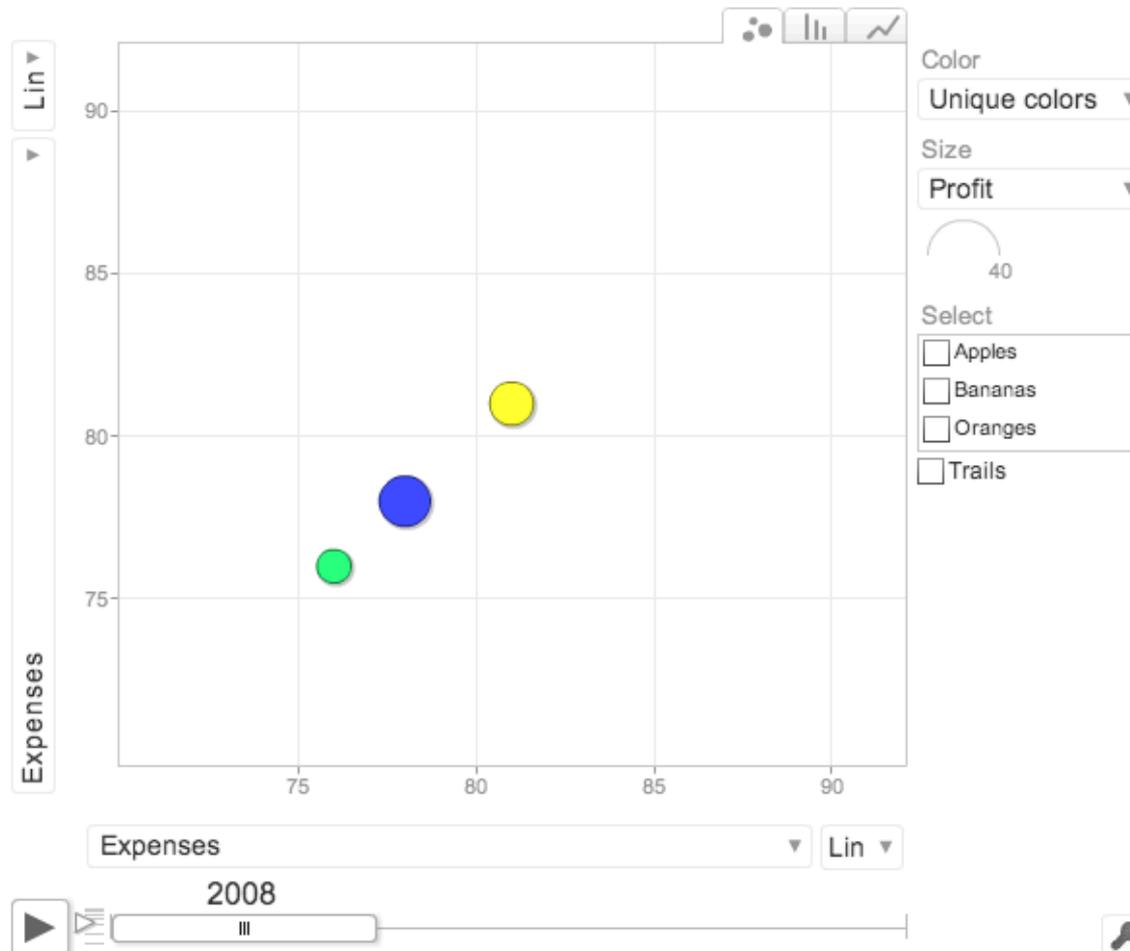
1.2. Hands-on - Using it in R

DATASET

```
> Fruits
```

	Fruit	Year	Location	Sales	Expenses	Profit	Date
1	Apples	2008	West	98	78	20	2008-12-31
2	Apples	2009	West	111	79	32	2009-12-31
3	Apples	2010	West	89	76	13	2010-12-31
4	Oranges	2008	East	96	81	15	2008-12-31
5	Bananas	2008	East	85	76	9	2008-12-31
6	Oranges	2009	East	93	80	13	2009-12-31
7	Bananas	2009	East	94	78	16	2009-12-31
8	Oranges	2010	East	98	91	7	2010-12-31
9	Bananas	2010	East	81	71	10	2010-12-31

1.3. Hands-on - Using it in R



Data: Fruits • Chart ID: [MotionChartID5f279ae813b](#)
R version 2.15.1 (2012-06-22) • [googleVis-0.4.2](#) • [Google Terms of Use](#) • [Data Policy](#)

1.4. Hands-on - Using it in R

```
#Specify the column names of the variable and time variable  
library(googleVis)  
data(Fruits)  
M <- gvisMotionChart(Fruits, idvar="Fruit", timevar="Year")  
#Creates html file in a temp folder using the input data  
plot(M)
```

2. Shiny – from R to the web



- Easy web applications in R;
- Shiny makes it super simple for R users to turn analyses into interactive web applications that anyone can use;
- Let your users choose input parameters using friendly controls like sliders, drop-downs, and text fields;
- Easily incorporate any number of outputs like plots, tables, and summaries;
- Great to visualize survey results, annual reports.

2.1. Stocks examples

- <http://glimmer.rstudio.com/winston/stocks/>

Stocks

Stocks

Apple (AAPL)
 Microsoft (MSFT)
 IBM (IBM)
 Google (GOOG)
 Yahoo (YHOO)

Chart type
Candlestick

Date range (back from present)

Time number
1 6 24

Time unit
Months

log y axis



References



- [The Visualization Group - http://vis.lbl.gov/](http://vis.lbl.gov/)
- [Dani's webpage - - http://vis.lbl.gov/~daniela/](http://vis.lbl.gov/~daniela/)
- [Nersc: http://www.nersc.gov](http://www.nersc.gov)
- [Kjiersten Fagnan - https://www.nersc.gov/users/computational-systems/genepool/genepool-training-and-tutorials/](https://www.nersc.gov/users/computational-systems/genepool/genepool-training-and-tutorials/)
- [GoogleVis: https://code.google.com/p/google-motion-charts-with-r/#Introduction](https://code.google.com/p/google-motion-charts-with-r/#Introduction)



Thank You