



because good research needs good data.

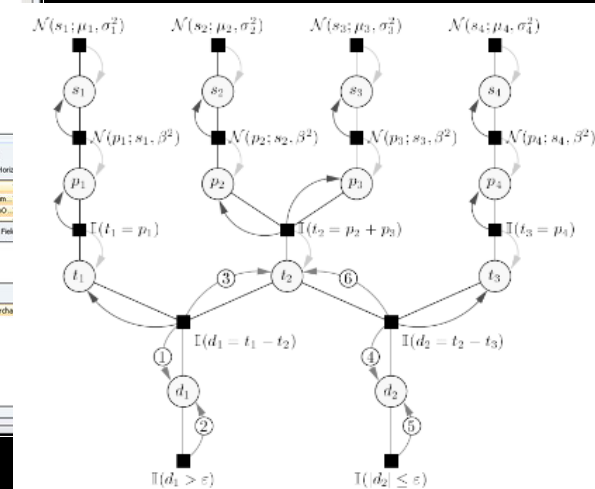
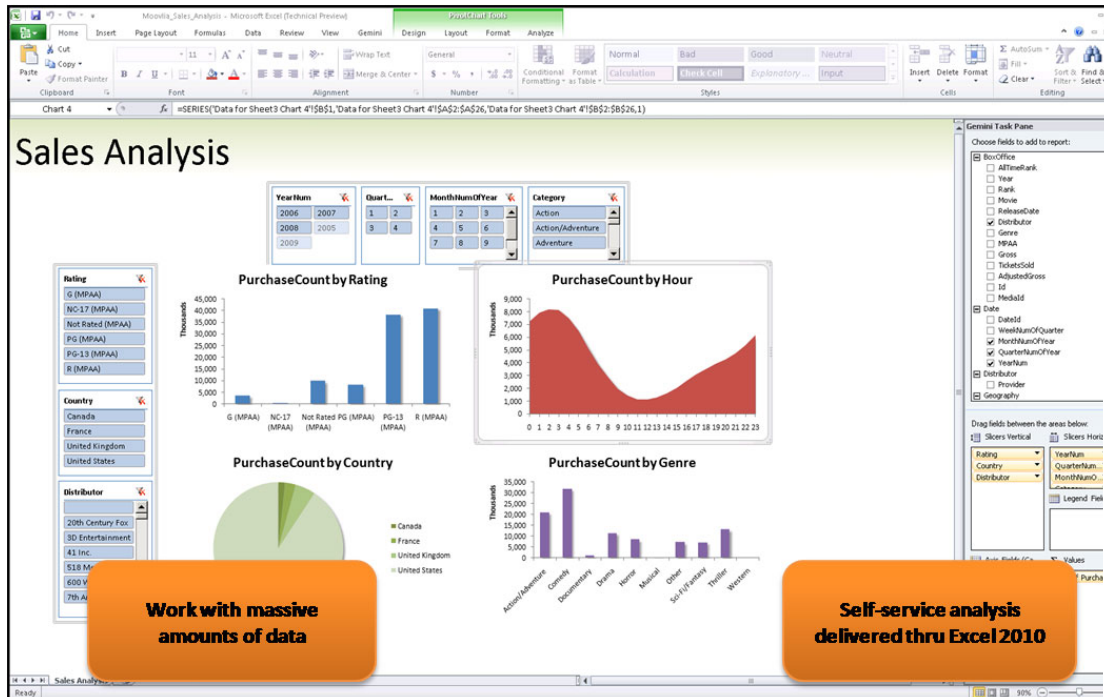
What is a Data Scientist?

IDCC Symposium, Amsterdam, January 2013

Chair: Dr Liz Lyon,
Associate Director, Digital Curation Centre,
Director, UKOLN, University of Bath, UK



Data Science



- Data as a commodity
- Data as infrastructure
- Data workflows, data tools, data publications
- Science which is data intensive, data driven
 - Science at web-scale
- Data as research substrate
 - Data as a science utility

Data Science

A close-up photograph of three ice cream cones. The foreground features a chocolate cone on the left and a vanilla cone with raspberries on the right. A third chocolate cone is visible in the background. The image is used as a metaphor for the 'flavours' of a data scientist's roles.

Flavours of “data scientist”

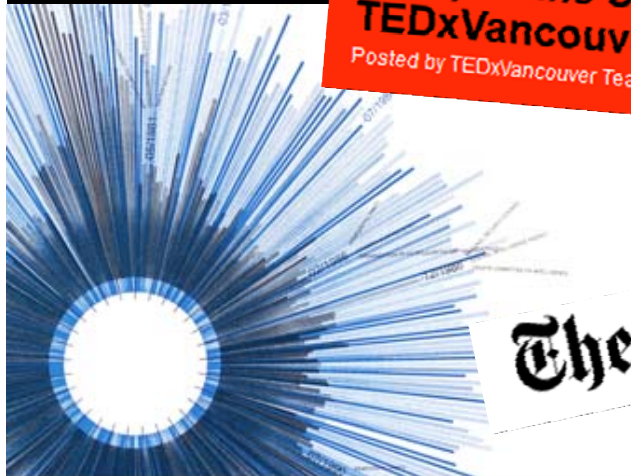
Family of roles?

- ***data engineer*** - focus on software development, coding
- ***data analyst*** – focus on business/scientific analytics and statistics
- ***data librarian*** – focus on advocacy, research data management / informatics
- ***data steward*** – focus on long term digital preservation, repositories, archives, data centres
- ***data journalist*** – focus on telling stories and news
- ***data publisher*** – focus on openly sharing data

New York Times Data Artist in Residence, Jer Thorp Joins Stellar Cast of Speakers at TEDxVancouver 2011

Posted by TEDxVancouver Team on October 17th, 2011 · No Comments »

The New York Times



Jer Thorp: Hope / Crisis, NYT Word Frequency



Stephanie Wright
University of Washington



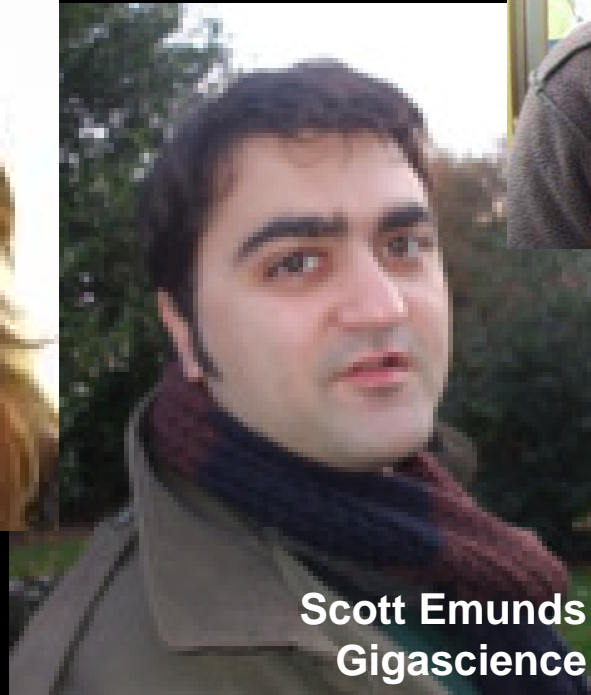
4 Data Scientists... in the Wild



Francine Bennett
Mastodon C



Louise Corti
UK Data Archive



Scott Emunds
Gigascience