Benchmarks Online
Archives
February 2016
April 2016
Next Edition: Aug. 15

Publication Schedule
Subscribe / Opt Out

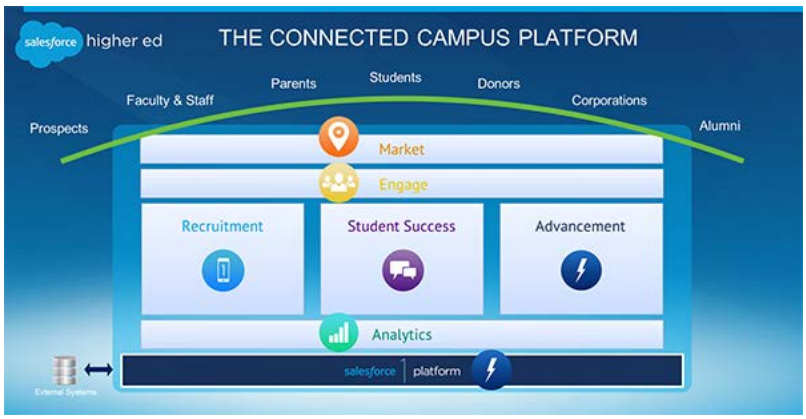
Social Media

Tech Insider: **How to Block Companies from Tracking You on Facebook**

Benchmarks Online, June 2016

Campus Computing News

Beyond Academic Advising: UNT Expands Student Success Technology to Admissions, UNT-International and URCM



By Jennifer Lee, IT manager and director, Student Success Technology Services, UIT

On May 10, 2016, UNT launched a new chapter in its vision to create a 360-degree view of the student experience. The offices of undergraduate and graduate admissions, UNT-International, and University Relations, Communications and Marketing took campus recruitment and marketing initiatives to new heights by becoming the latest members of the campus Salesforce family.

The strategic move will help the University of North Texas create a connected college experience for the student lifecycle. **Read more about it.**

CLEAR: UNT Purchases Site License for New Student-Response System

By Jane Himmel, associate director, Center for Learning Enhancement, Assessment and Redesign

REEF Polling by i>clicker replaces Turning Point as UNT's centrally supported student-response system, effective with the fall 2016 semester. Read more about it.

UNT's Sky Theater Touts Cool Technology and Temperatures this Summer

People in the Zone

UNT UIT HOTSPOT Benchmarks Online

Jeff Anderson

Jennifer Lee

Jennifer Spillman



It may be a starry, starry night, but you can see only what the city lights, hot nights and the naked eye permit. A refreshing 70-degree cosmic safari at the UNT Sky Theater is the hot ticket this summer! **Randall Peters**, planetarium director in the UNT Physics Department, uses advanced computer graphics technologies to teach, engage and entertain UNT students, K-12 visitors and the general public year round. Read more about the technology, the site and the sights!

EDUCAUSE: Registering Now for The Best Thinking in Higher Education IT Conference

The EDUCAUSE Annual Conference, Oct. 25-28, 2016, unites the best thinking in higher education IT and is derived from member-driven content that is organized across overarching IT domains. **EDUCAUSE registration is now open** for events at the Anaheim Convention Center in Anaheim, California. And, yes, Anaheim is the home of the original Disneyland, only 2.8 miles from the convention center!

A $2.5 million upgrade to the convention center's Wi-Fi network system makes it one of the most technically advanced convention centers in the U.S. The newly installed high-density network consists of nearly 700 access points throughout the facility. This first-of-its-kind system increases the amount of coverage and bandwidth available to ACC meeting and event attendees. The Anaheim Convention Center is now the first convention center on the west coast with 10-gigabit internet connectivity.

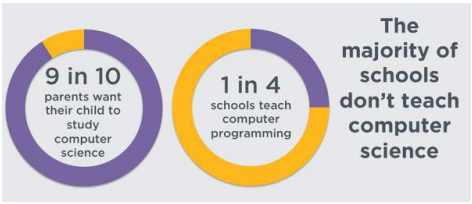
Teaching Assistant Surprises Class with a Remarkable Revelation

Jill Watson, one of nine teaching assistants assigned to help answer questions about coursework and projects from the 300 students enrolled in the advanced artificial intelligence course at Georgia Tech, really struggled during the first few weeks of the semester. As the semester continued, Jill was getting better at her assigned tasks; then she surprised everyone. Read through the article for the surprise in the story: a hint is in the name of the course. **NOTE:** If you are interested in artificial intelligence, check out UNT course CSCE 4310, Introduction to AI, taught by Professor Kathy Swigger, College of Engineering.

Video: Hackers Show Vulnerabilities in the U.S. Power Grid

Professional computer hackers were hired to break into the U.S. electrical power grid in the Midwest as part of the company's offensive security efforts. By using social engineering, finding weak spots in physical security and breaking in, the technicians gain access, install malware, hack computers and more – all captured on a Go Pro. Check it out.

Bill Gates' Online Petition Promotes Teaching of Computer Science in K-12 #CSForAll



In a recent tweet about teaching computer science in the U.S., **Bill Gates**, technology advisor at Microsoft, stated that every student deserves a chance to learn this essential 21st-century skill. Seeking 150,000 signatures, see the petition here: http://www.SeSeechange.org/computerscience. In addition to being a business magnate, inventor and philanthropist, you may have noticed that now Gates' official title is technology advisor. Moving away from his former duties as CEO at Microsoft in 2014, Gates is now free to pursue his pet projects, such as making his first virtual reality video and reading science fiction among others. Check it out his blog, Gates Notes, with the #VR

video and his summer reading list.

DYK?

VIDEO: Your iPhone has a hidden childproof mode: watch how to enable it.

Five Big Takeaways from Meeker at the Internet Trends 2016 Code Conference

Longtime Morgan Stanley analyst and investor Mary Meeker has released her annual Internet Trends report, a widely read data-filled compendium of where the technology industry is headed. Internet growth is flat globally: The number of internet users globally is flat, according to Meeker's report. In 2014 and 2015, usage was up 9 percent both years. Also, data as a platform is exploding: The growth of data is increasing while the cost of storing it is declining. Analyzing this data is increasingly crucial for businesses to manage their operations including security, communication, and customer insights. Meeker is an American venture capitalist and former Wall Street securities analyst. Her primary work is on internet and new technologies. She is a partner at the Silicon Valley venture capital firm Kleiner Perkins Caufield & Byers. Click here for Meeker's Internet Trends slide show.

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

REEF Polling

The mobile-optimized engagement system

With REEF Polling, instructors can present with anything, utilize a floating toolbar, view a participant list, and send images and polling questions to student devices. Students respond to questions with any device, receive correct and incorrect answers, and can review session history as a study guide outside of class. REEF Polling is also a hybrid product, allowing students in the same classroom to use their smartphone, tablet, laptop or an i>clicker remote.

INSTRUCTORS: GET STARTED

STUDENTS: GET STARTED

EXISTING I>CLICKER USERS

Campus Computing News, June 2016

Beyond Academic Advising: UNT Expands Student Success Technology to Admissions, UNT-International and URCM

By Jennifer Lee, IT manager and director for Student Success Technology Services, UIT | April 2015 Salesforce article

On May 10, 2016, UNT launched a new chapter in its vision to create a 360-degree view of the student experience. The offices of undergraduate and graduate admissions, UNT–International, and University Relations, Communications and Marketing took campus recruitment and marketing initiatives to new heights by becoming the latest members of the campus Salesforce family. The strategic move will help the University of North Texas create a connected college experience for the student lifecycle.

The project launch was a culmination of months of work for the Salesforce cross-functional project team that included the leadership and key members from Admissions, URCM, Academic Advising, UNT–International, Student Success Technology Services, University Information Technology and UNT System IT Shared Services.

Two implementation partners also were involved in the project. The group worked with TargetX based in Conshohocken, Pa., one of the leading higher education customer relationship management companies, to build its recruitment platform in Salesforce. Also, the group collaborated with DEG based in Overland Park, Kansas, a digital agency purpose-built to meet clients with data-driven marketing, commerce and collaboration solutions to implement a comprehensive, multi-channel marketing strategy plan driven by data analytics.

Regarding recruitment, Salesforce will help UNT manage prospective students, schools, campaigns, events and all engagement with future students in one system. It will improve outreach efforts by providing better data intelligence. The recruitment solution also offers predictive intelligence on yield efforts.

When it comes to marketing initiatives, the university will use its marketing tools to move potential students along the recruiting journey with persuasive messages about the benefits of attending UNT. The CRM software also allows personalization of each message based on the interests of the prospective student while tracking the effectiveness of its marketing strategies.

As more higher education institutions have turned to CRM solutions to manage the student lifecycle, from prospective students to alumni, the move has created an opportunity for campus users to share and use student data in ways that were not possible before. With UNT Academic Advising at the helm of the first wave of deployment a year earlier, Salesforce enabled advising offices on campus to share notes as well as develop strategies to promote student success and student engagement. With the recent additions of recruitment and marketing applications on the same platform, UNT can use Salesforce as a blueprint for academic excellence whether it is tracking recruitment, advising, retention or student data.

Project Team Role	Name
Sponsor	Neal Smatresk
Steering Committee	Shannon Goodman, Rebecca Lothringer, Deborah Leliaert, Kenn Moffitt, Phillip Baczewski, Kem Marcum, Mike McKay
Project Manager	Schenita (Shay) Floyd
Business / Functional Leads	Shannon Goodman, Rebecca Lothringer, Deborah Leliaert, Kenn Moffitt, Phillip Baczewski, Mike McKay, Jeri Takimoto, Pieter Vermeulen, Dana Mordecai
Technical Architect	Martin (Marty) Miller
Technical Leads	Kem Marcum and Monika Botha
Programmer/Developers	Kok Chuan Koh, John Dysart, Rachel Richey
Business/Functional Team	Maureen Saringer, Craig Howard, Jennifer Lee, Jennifer Spillman, Austin Milner, Dickie Hargrave, Madison Davis
Salesforce Systems Administrators	Jennifer Lee, Jennifer Spillman
Vendors	DEG, Salesforce, TargetX
Vendor Project Manager	Liz Sullivan, DEG, and Creighton Dent, TargetX
Vendor Implementation Consultant	Kristen Hewlett, DEG, and James Seymour, TargetX

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Contact Us:

University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:

Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

UNT System:

- UNT Home
- UNT System
- UNT Dallas
- UNT Health Science Center

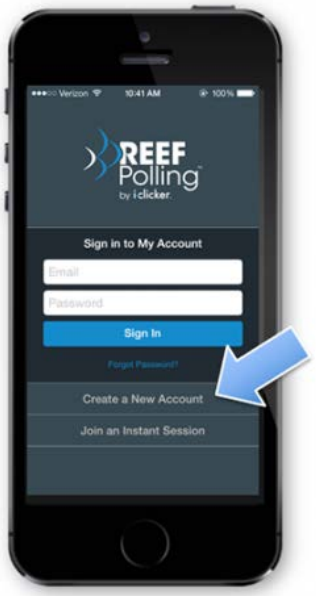
Site last updated on June 15, 2016

UNT Adopts New Student Response System

UNT Purchases Site License to REEF Polling

By Jane Himmel, associate director, Center for Learning Enhancement, Assessment and Redesign

REEF Polling by i>clicker replaces Turning Point as UNT's centrally supported student response system, effective with the fall 2016 semester.



UNT has purchased a site license to **REEF Polling** by **i<Clicker**, a mobile engagement platform that enables faculty to increase class participation and gather instant feedback by sending polling questions to students' devices. Students respond with their smartphone, tablet, or laptop and can review session history on their REEF account outside of class.

With this purchase of a **REEF Polling** site license, UNT is providing this software to **every student free of charge**.

Effective with the fall 2016 semester, **REEF Polling** will replace Turning Point as UNT's centrally supported student-response system.

During the spring 2016 semester, a faculty-led committee thoroughly evaluated and tested leading student-response systems in an effort to recommend a replacement for Turning Point 5, which the company had announced it would soon stop supporting. The committee recommended adoption of **REEF Polling** citing ease-of-use, integration with Blackboard and top-notch support, among other factors.

Training on REEF Polling for faculty began on June 1, 2016 and continues throughout the summer in both face-to-face and webinar formats.

For webinar times and dates, please visit our GoToTraining session list to register. http://bit.ly/25q4aFt

For in-person sessions, please RSVP via this Qualtrics survey: http://bit.ly/1U0to27

Additionally, REEF offers an Instructor Quick Start Guide, a Syllabus Template, and a First Day of Class Template, all designed to help make the transition easy for faculty.

For more information, go to http://clear.unt.edu/reef, or call **Kathy Roberts**, instructional technology training and support, 940-369-5201.

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Archive

- 1980 (2)
- 2009 (4)
- 2010 (12)
- 2011 (12)
- 2012 (12)
- 2013 (12)
- 2014 (12)
- 2015 (6)

Issues from before September 2009 or in years not shown on this page may be found here.

Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

Subscribe to the Benchmarks Online Mailing List

If you would like to know when the latest issue is available and get a preview of it's contents, you can subscribe to the **Benchmarks Online** mailing list, unt-benchmarks.

SUBSCRIBE

To subscribe to this list, send an email as indicated below.

- Address your email to listserv@unt.edu.
- In the body of the message, type: **SUBscribe unt-benchmarks + your name**
- Send your email.

UNSUBSCRIBE

To insubscribe/opt out, send an email as indicated below.

- Address your email to listserv@unt.edu
- In the body of the message type: **Sign Off unt-benchmarks + your name**
- Send your email.

Questions? Contact the Benchmarks Online editor, by email or telephone, 940-565-4068.

Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

Network Connection, June 2016

The Once and Proper Internet

By Philip Baczewski, senior director, University Information Technology

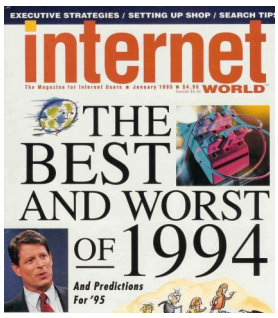
The Internet has passed yet another milestone. Starting on June 1, 2016 the *AP Stylebook* will stipulate that the Internet will no longer be considered a proper noun, but will be referred to simply as the internet. As reported by ABC News, both the Internet and the Web no longer warrant capitalization, placing internet in a similar category to electricity and telephone. Of course, Internet was never trademarked as kleenex and thermos were, but we've capitalized it all these years anyway. It was not just any internet -- it was the Internet.

As pointed out in The Atlantic, we don't call the Internet the Information Superhighway anymore either. However, the Internet still seems to carry a more distinctive role in our lives than does the electric grid. Without the electric grid we wouldn't likely have any internet, but we don't place as much cultural stock in it as we do in the Internet. If you google (non-proper verb) them both on Google (proper noun), you'll get 668 million results for "internet news", but only 17 million results for "electrical grid news." (You may also find out that a monkey fell into a transformer, knocking out Kenya's entire power grid.)

Looking Back...

Perhaps it's just a symptom of getting old, but I continue to marvel at the Internet and the impact, good and bad, that it has on our society. Or perhaps it's the fact that I've been an ongoing observer of the Internet's growth and development almost since its inception as a widely available resource that makes me reluctant to relegate it to the small-i Internet. I've been writing this column since 1988, and it's been appearing on the Internet since 1994 -- OK, I guess I am getting old. But, looking back at this column, it's possible to construct a narrative that may illustrate that the ubiquitous internet that we have today wasn't always a future certainty. What follows is just a sample in four-year increments.

- *A Brave New Internet* (January/February, 1994)



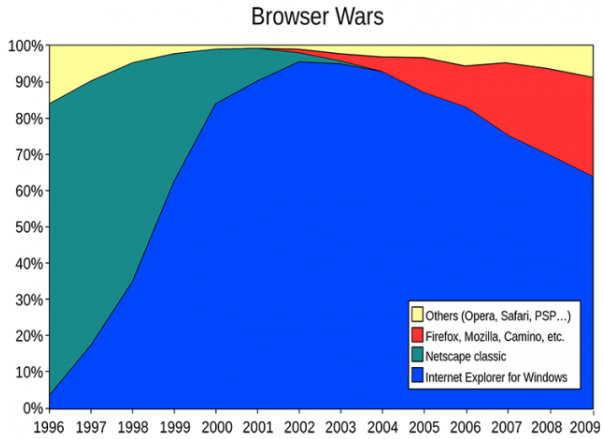
  You may have caught news reports of Vice President Gore's proposal to restructure the communications industry in the U.S. The administration's proposal seeks to reduce regulation on the telephone and cable TV industries and encourage the development of a new level of electronic information access for U.S. households. If implemented and successful, the new policies will have a definite effect on who and how many will have access to the Internet.

  One of the main benefits of the change in the Internet should be the extensive development of private Internet services. It is already possible to contract with a private service provider to get Internet access from your home. The trend toward more reliance on commercial network providers should greatly increase the number and extent of services available to private households. The year 2000 may see us truly in the midst of the Information Age.

- *What's all this Fuss* (June, 1998)

  Why should we be concerned about Internet browsers. After all, as long as you can still use your favorite one, why should there be a fuss?

  ... your browser, through its built-in features, can influence you towards certain Web sites at the exclusion of competing and



possibly better alternatives. Without competition, one company determines how the Internet will look to you.

So the fuss is really about what the Internet will look like in 20 years. Will it be a thriving and egalitarian international communications and commerce media, as accessible in libraries as it is in executive offices? Or will it be "television," A.K.A. a "wasteland" of commercial and meaningless information driving a consumption-based economy and culture, filling the coffers of a few very rich companies and people? All this fuss about browsers might just be the deciding factor.

- *The New "World Book"* (April, 2002)

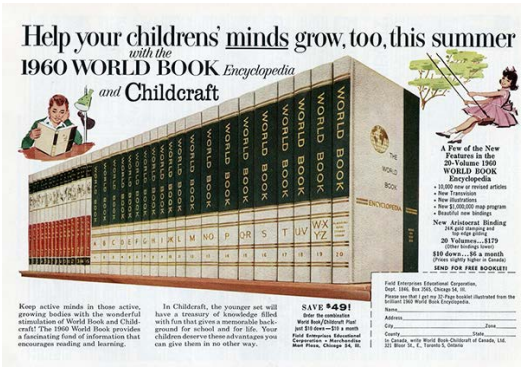
  The "information age" in which we find ourselves has caused a cultural shift which has obliterated the encyclopedia as we knew it from our collective conscious. Why buy a set of paper books which will be obsolete before the ink dries on your check when an electronic source can be updated continually with minimal expense? ... Even the concept of the encyclopedia as a unit of measure is no longer useful. In the olden days of computing (1988 or so), we used to measure the capacity of magnetic tapes in terms of the number of encyclopedia volumes it could hold. Today, that set of encyclopedias is just a spec on your 80 Gigabyte hard drive.

- *The Once and Future Internet* (June, 2006)

  If we future beings have learned one thing about the Internet, it is that we are not creating a global village. One thing the Internet does support is many communities that can extend beyond the geographical limitations of the past. Whether it's a community of learners or a community of software developers (as in open source), the impact is not from the technology itself, but rather from how people can use it to their advantage. That's not a utopian vision of life-apprentice learning, but it is a significant and sometimes positive change. We'll just need some more future to tell for sure.

- *The Problem with the Internet* (February, 2010)

I have to admit that when I was a kid, I liked to read randomly the encyclopedia and wished that such random and extensive knowledge didn't require such heavy books. In my college years navigating through numerous volumes of abstracts and indexes, I thought that there had to be a better way to access the tremendous volumes of scholarship produced by the academic world. I revel in the fact that I now have my wish in both cases and it's via an elegant little electronic package that rides in my pocket. My rotten brain never worked better.



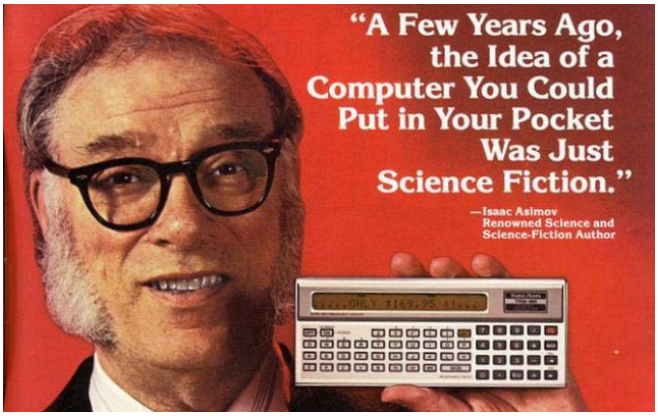
- *Big Data -- Big Problem* (January, 2014)

The convergence of universal Internet connectivity, online transaction processing, and inexpensive data storage has the potential to turn our Internet information dream into a dystopian privacy nightmare. ... It seems that we've created the technology of data aggregation without much understanding or even consideration of the ethics and impact of its use.

I suppose it's too late to bring back pay phones, printed encyclopedias, dumb power grids, and cash transactions. It would seem that this would be the time for some reasoned debate to determine some reasonable legislation to protect individual privacy in this age of aggregate data. However, today's U.S. legislative climate would seem to exhaust any hope of reason.

Looking Forward...

An advantage of longevity is that it turns you into a bit of a time traveler. Much of the promise of the early Internet has indeed been fulfilled. But as with all technologies, there are unforeseen issues that develop and unexpected turns along the way. The development of mobile networks and devices has moved much of our Internet consumption to devices we carry in our pockets. Facebook and Twitter have created enormous networks of people in ways not anticipated in the days of Listserv and IRC. And data aggregation and analytics now seem to question the concept of free will, a concept that has broad implications for human behavior.



Another futurist dealt with this concept many years ago. Isaac Asimov portrayed the concept of Psychohistory in his Foundation series of short stories and novels first published in 1942. Psychohistory was the embodiment of an idea that laws of statistics could predict the progression of future events. It's axioms included "that the population whose behaviour was modeled should be sufficiently large" and "that the population should remain in ignorance of the results of the application of psychohistorical analyses." Asimov, in a short story titled, *Franchise* (1955), posited an election where only one voter is needed to allow the ultimate computer, Multivac, to determine the outcome.

It seems that Asimov was a genius in predicting today's technology, yet in reality, he was simply extrapolating human behavior. The technology has not changed our human tendencies -- it has just allowed things to happen at a faster pace. So if the Internet is now subsumed into some general concept of internet, that doesn't change the nature of the impact it has had on our society and daily lives. Information technology has changed the way we live. But hopefully, we will still attempt to exercise our free will, even if the internet questions the fact that we have it.

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Contact Us:

University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:

Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

UNT System:

- UNT Home
- UNT System
- UNT Dallas
- UNT Health Science Center

Site last updated on June 15, 2016

Help Desk FYI

Use the Clutter Feature to Help Sort Low-priority Emails in Ou

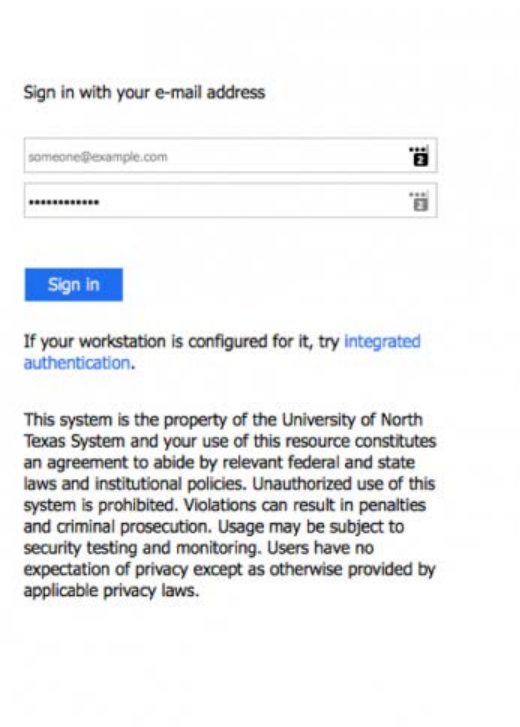
By Sharukh Mithani, manager, UIT Help Desk

June 15, 2016 – Clutter is a mailbox management feature in Microsoft® Outlook that is designed to help you see your most important emails first. Based on how you manage your email, your Clutter folder will begin to fill automatically with emails that you frequently ignore. You also can help determine **clutter** by dragging emails from your inbox to the Clutter folder. This feature is different from junk filtering that is designed to prevent spam and unsolicited email from reaching you.

Although the Clutter feature may be useful for many people, it can be disabled easily should it become a distraction. Please follow these instructions to disable or re-enable Clutter.

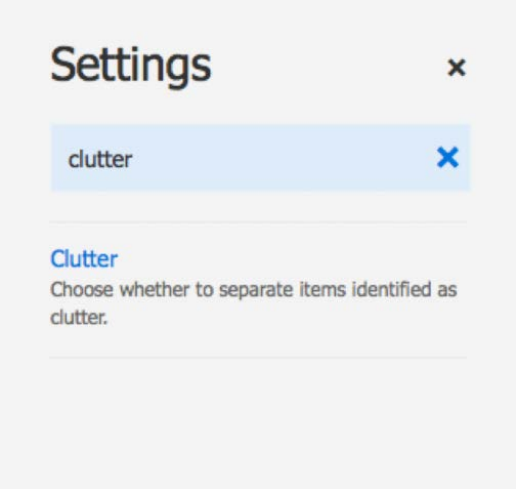
To disable / enable Clutter:

1. Sign into Outlook through the **web interface**.

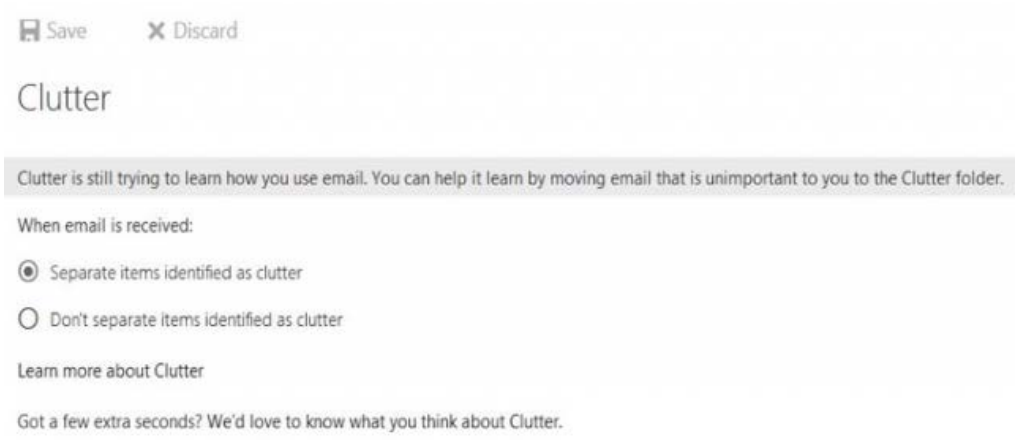


2. Click on the "**Settings**" icon near the top right corner.

3. Type "**Clutter**" in the search box, and then click on the "**Clutter**" option that appears in the search.



4. Uncheck "**Separate items identified as clutter**" to disable Clutter. Or, check "**Separate items identified as clutter**" to enable Clutter.



5) Select "Save"

NOTE: Clutter is a recently introduced feature of Microsoft® Outlook and has been enabled on all accounts by default. If you have noticed missing emails recently, it may have been moved to Clutter.

Please contact the UIT Help Desk should you have any questions.

Telephone: 940-565-2324 | Email: helpdesk@unt.edu | Sage Hall, Room 130

Telephone Availability:
Sunday: noon-midnight
Monday-Thursday: 8 a.m.-midnight
Friday: 8 a.m.-8 p.m.
Saturday: 9 a.m.-5 p.m.

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

RSS Matters

R_stats

Explicit Bayes: Working Concrete Examples to Introduce the Bayesian Perspective.

Re-published June 2016; originally published January 2015.

Link to the last RSS article here: Identifying or Verifying the Number of Factors to Extract using Very Simple Structure. -- Ed.

By Dr. Jon Starkweather, Research and Statistical Support Consultant Team

We use the term *explicit* because we are going to calculate these examples *by hand* with programing rather than simply loading a package and using functions to estimate parameters. The purpose of using these explicit methods is to hopefully convey a better understanding of what it means to *do* Bayesian statistics.

First, we must present a little bit about Bayesian statistics. Very, very briefly, Bayesian statistics requires three elements: a prior, likelihood, and a posterior. The prior is a distribution specified by the researcher which represents all *prior* information regarding the parameter the researcher is attempting to estimate. The prior represents an educated, best guess at the parameter (e.g. the mean of the prior) and the degree of certainty or confidence in that educated, best guess (e.g., the variance and shape of the prior distribution). The prior is specified before (i.e. *prior*) to data collection. The prior is then combined with the likelihood (a representation of the data at hand) to create a more informed, empirical distribution of the parameter being estimated. We call this last distribution the *posterior* distribution. The mean of the posterior is our estimate of the parameter. Interval estimates can then be calculated from the posterior which truly will represent the interval which contains the actual population parameter; we call those intervals, *credible intervals* (rather than confidence intervals – which *do not* tell you the probability of the population parameter being contained in this interval).

Let's say we want to estimate **the mean** IQ scores on the Weschler Adult Intelligence Scale (WAIS) of a small town, X.Town, which has a population of 10000 individuals. Let's start by importing the X.Town data.

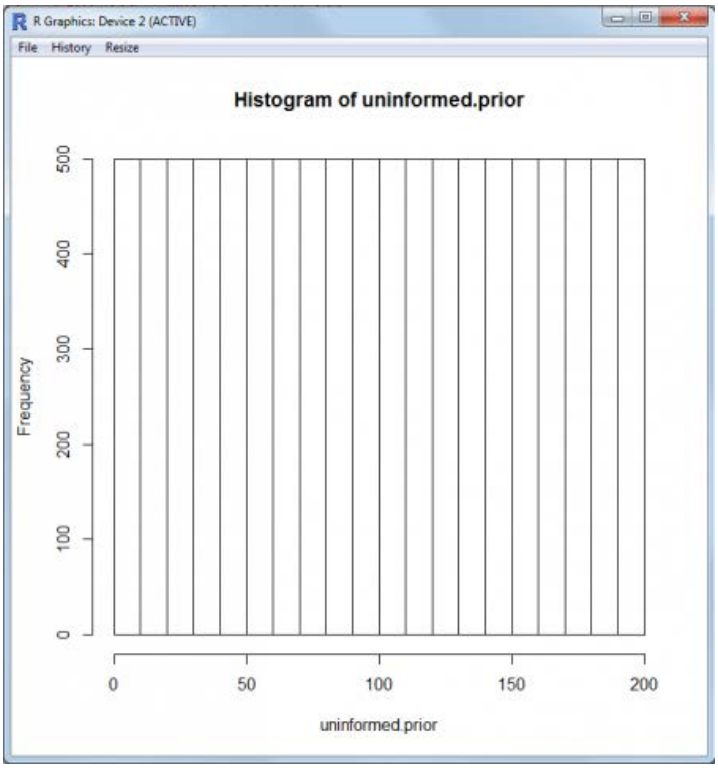
```
x.town.df <- read.table("http://www.unt.edu/rss/class/Jon/ExampleData/X.Town.sample.txt",
        header = TRUE, sep = ",", dec = ".", na.strings = "NA")
nrow(x.town.df)
[1] 10000
```

We know from a mountain of normative data and prior research that the U.S. population distribution of WAIS scores has a mean (μ) of 100 and a standard deviation (σ) of 15. This information represents a best case scenario; where we *know* the population distribution and that distribution is normally distributed with an identified mean and standard deviation. Generally, we would not have such great prior information; so consider an alternative where we have virtually no prior information accept to know the WAIS questions / procedures which allow a possible score to range from 1 to 200. In such a case, our specification of a prior distribution would mean each score in that range is equally likely -- which prompts us to specify a *uniform* distribution (i.e. a distribution in which each value has an equal probability of being represented). A uniform prior is also known as an un-informative or un-informed prior. In both examples below we are using a population of 10000 individuals.

```
uninformed.prior <- rep(seq(1:200), 50)
length(uninformed.prior)
[1] 10000
summary(uninformed.prior)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1.00   50.75  100.50  100.50  150.20  200.00
```

```
hist(uninformed.prior)
```



However, with the WAIS and the knowledge of the U.S. population, we can specify a Gaussian (i.e. normal) distribution as our prior.

```
informed.prior <- rnorm(10000, mean = 100, sd = 15)
```

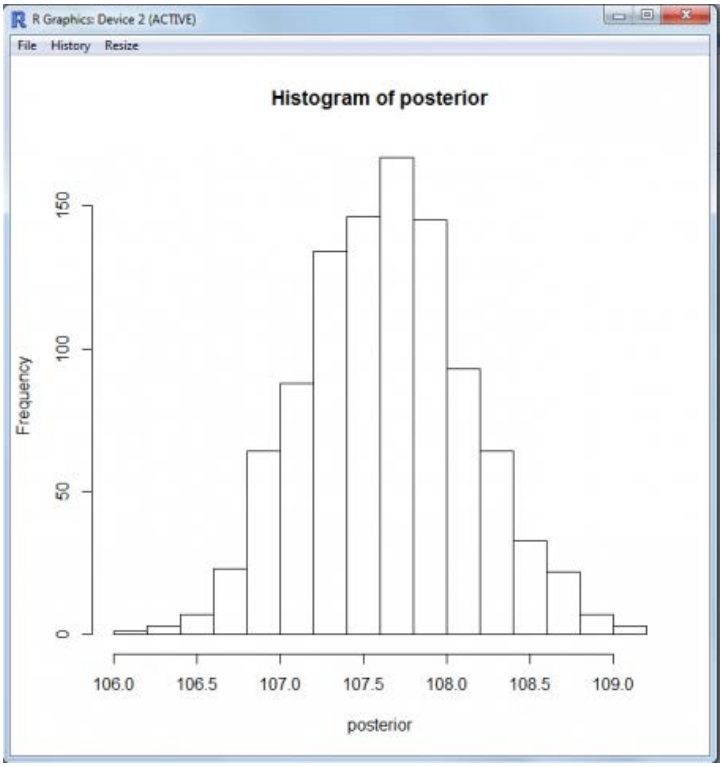
```
length(informed.prior)
```

```
[1] 10000
```

```
summary(informed.prior)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  37.51   89.93  100.10  100.10  110.40  157.30
```

```
hist(informed.prior)
```



Clearly; the two example priors above are extremes (i.e. worst case and best case); there are a variety of other distributions which can be specified as priors (e.g. Cauchy, Poisson, beta, etc.) and the prior **is not** required to be symmetrical. For more information on the variety of distributions, see: http://en.wikipedia.org/wiki/List_of_probability_distributions

Our research questions are as follows: What is the mean WAIS score of the population (*n* = 10000) of X.Town; and, does that mean differ from the larger (U.S.) population? In more precise terms, what is the population mean of X.Town WAIS scores and is that mean *larger* than the known U.S. population mean. To be clear, there are two populations we are referring to here; the population of X.Town (*N* = 10000) and the larger population of the U.S.

It is unrealistic to think we would have all 10000 adult citizens' data from X.Town; we would generally have a sample of that town's data. Note; the 7th column of our X Town data file contains the WAIS scores. Here we randomly sample (*n* = 1000) cases from the entire X.Town data (*N* = 10000):

```
wais.sample <- sample(x.town.df[,7], 1000, replace = FALSE)

length(wais.sample)

[1] 1000
```

Traditional Frequentist Perspective: Null Hypothesis Significance Testing (NHST).

In a traditional *frequentist* setting, we would begin by simply calculating the sample mean as our best estimate of the entire X.Town population mean WAIS score:

```
M <- mean(wais.sample)

M

[1] 107.6305
```

and the standard error of that mean if we wanted confidence intervals for that estimate (of the entire X.Town's mean):

```
std.err <- sqrt(15^2 / length(wais.sample))

std.err

[1] 0.4743416
```

Then using an alpha value (e.g. 0.05) look up the associated critical value (i.e. +/-1.96) in a table; then calculate the lower and upper bounds of the confidence interval for our estimate (i.e. the confidence interval for the estimated mean of X.Town).

```
lower.bound <- (-1.96*std.err) + M
```

```
lower.bound
```

```
[1] 106.7008
```

```
upper.bound <- (+1.96*std.err) + M
```

```
upper.bound
```

```
[1] 108.5602
```

Then, we would run a one sample t-test using our random sample of X.Town adults' WAIS scores, comparing **the mean** of the sample scores (*M*; as our best estimate of the entire X.Town's mean) to the mean of the U.S. population (mu*:* μ); using the standard error of the mean (*std.err*) and some pre-designated probability cutoff (e.g. 0.05) to determine statistical significance.

```
t.test(wais.sample, alternative = 'greater', mu = 100, conf.level = .95)
```

```
        One Sample t-test

data:  wais.sample

t = 17.0653, df = 999, p-value < 2.2e-16

alternative hypothesis: true mean is greater than 100

95 percent confidence interval:

 106.8944      Inf

sample estimates:

mean of x

 107.6305
```

It is important to recall (or review) what the above test is doing. We have drawn a random sample of data from X.Town and we are testing **the mean** of that sample against a known (U.S.) population mean to determine if the sample indeed comes from that population (i.e. the null hypothesis). Notice we are using the sample mean (*n* = 1000) as a representation of the entire X.Town's WAIS scores (*N* = 10000).

Bayesian Perspective: Bayesian Statistics; Bayesian Inference; Bayesian Parameter Estimation.

All three of the above terms are often used to refer to Bayesian data analysis. The examples below were all adapted from Kaplan (2014). Our example explores the normal prior for the normal sampling model in which the variance σ^2 (sigma squared) is assumed to be known. Thus, the problem is one of estimating **the mean** μ (mu). Let *y* denote a data vector of size *n* (*y* = the sample of 1000 WAIS scores). We assume that *y* follows a normal distribution shown with the equation below:

$$p(y|\mu, \sigma^2) = (1/\text{sqrt}(2*p*\sigma)) * \exp(-((y - \mu)^2) / (2*\sigma^2))$$

To clarify and show an example in R, we use the following:

```
mu <- 100
```

```
o <- 15
```

```
y <- wais.sample
```

We use the word 'output' to refer to $p(y|\mu, \sigma^2)$ from above; which is read as the probability of *y*, given a mean of mu (μ), and variance of sigma squared (σ^2).

```
output <- (1/sqrt(2*pi*o)) * exp(-((y - mu)^2) / (2*o^2))
```

```
summary(output)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.

0.000289 0.047630 0.078600 0.069690 0.096360 0.103000
```

Next, we specify the prior. We have plenty of confidence that our prior distribution of the mean is normal with its own mean and variance hyper-parameters, *k* and t^2 (using *t* in R code to refer to tau: τ), respectively, which for this example are known. The prior distribution can be written as:

$$p(\mu|k,t^2) = (1/\text{sqrt}(2* p*t^2)) * \exp(-((\mu - k)^2) / (2*t^2))$$

The term, $p(\mu|k,t^2)$, can be read as the probability of μ given *k* and t^2.

```
k <- mean(y); k
```

```
[1] 107.6305

t <- sd(y); t

[1] 14.13976

n <- length(y); n

[1] 1000

prior.mean <- (1/sqrt(2*pi*t^2)) * exp(-((mu - k)^2) / (2*t^2))

prior.mean

[1] 0.02439102
```

Combine the prior information with the likelihood of the data (given the population variance; sigma squared [σ^2] and the sample size [n]) to create the posterior distribution. Using some algebra, the posterior distribution can be obtained as:

$$p(\mu|y) \sim N[\ ((k/t^2)+(n*\text{mean}(y)/\sigma^2))\ /\ ((1/t^2)+(n/\sigma^2)),\ (t^2*\sigma^2)/(\sigma^2+(n*t^2))\ ]$$

Thus, the posterior distribution of mu (μ) is normal with a mean:

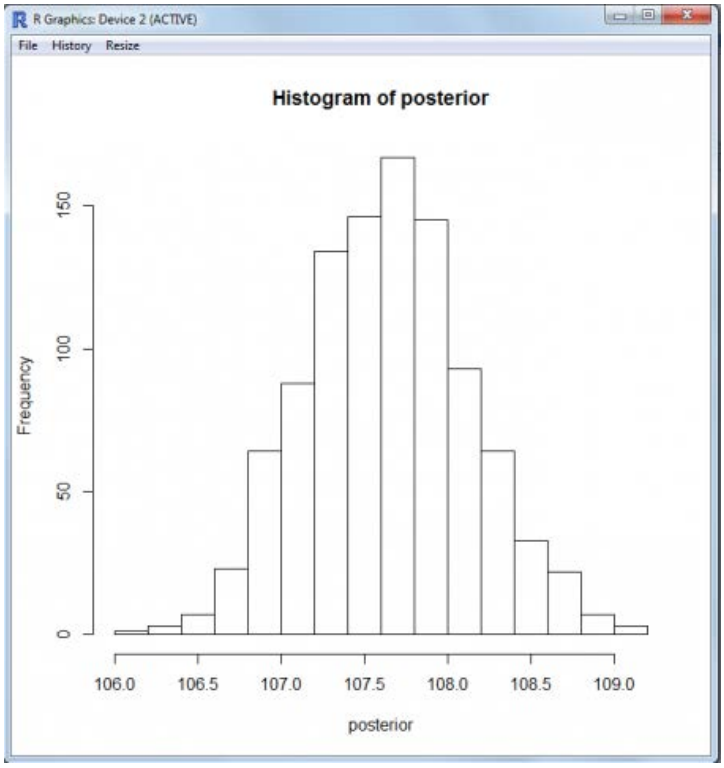
```
posterior.mu <- ((k/t^2)+(n*mean(y)/o^2)) / ((1/t^2)+(n/o^2))

posterior.mu

[1] 107.6305
```

and variance:

```
posterior.o2 = (t^2*o^2)/(o^2+(n*t^2))

posterior.o2

[1] 0.2247471
```

So, the posterior distribution can be simulated using these two parameters (and n = 1000); which in R, should be:

```
posterior <- rnorm(n = length(y), mean = posterior.mu,
                   sd = sqrt(posterior.o2))

hist(posterior)
```



In a traditional frequentist analysis, one would be required to report both the estimated mean (i.e. mean of the sample) and a confidence interval with lower and upper bounds of that mean. However, a frequentist confidence interval only tells us; if this same study was repeated 100 times, we would expect the sample mean to be between the upper and lower bounds 95 times (if using a 95% confidence interval). It **does not** tell us the probability of the population parameter being included in the interval. Here in the Bayesian setting, we use the posterior distribution

and simply take the quantiles (i.e. probabilities) to compute the lower and upper bounds of a *credible interval* – which does give us the probability that the actual population parameter is included in this interval.

```
quantile(posterior, c(.05,.95))

    5%    95%

106.8662 108.4625
```

It is critically important to recognize, the above example is **only** interested in estimating the mean of X.Town's WAIS scores. The example is NOT attempting to estimate the entire X.Town's distribution of WAIS scores. So let's compare the actual mean of X.Town's WAIS scores to the sample mean, and the mean of the posterior distribution (of course, in a real research situation you would not have the 'actual' parameter -- i.e. mean of the entire population of X.Town).

```
mean(x.town.df$wais)

[1] 107.8662

mean(wais.sample)

[1] 107.6305

mean(posterior)

[1] 107.6389
```

Undoubtable readers will notice the virtually identical estimates provided by the mean of the posterior (i.e. Bayesian estimate) and simply the mean of the sample (i.e. frequentist estimate); and both of those are very, very close to the X.Town population mean. There are two very important reasons for this. First, the Bayesian and Frequentist methods will result in virtually the same parameter estimate(s) with large samples. The prior is weighted very lightly and the likelihood (a representation of the data at hand) contributes the bulk of the weight to the estimation when large samples are used in a Bayesian analysis. Second, the data used in the examples above is simulated data and a truly random sample (n = 1000) was taken from the entire population (N = 10000). Therefore, our results here have very low bias as a result of the truly random sample and the fact that 10% of the population was contained in the sample. Most research is not conducted on a truly random sample and very few research endeavors include 10% of the population as the sample.

Lastly, hypothesis testing and statistical significance are not foreign to the Bayesian perspective. For example, if one were interested in conducting a Bayesian *t*-test, you would use something called Bayes Factors which has been covered on the RSS Do-it-yourself Introduction to R web site and specifically here in Module 11. Bayes Factors were also discussed in a previous RSS Matters article (Adobe.pdf version).

Until next time, "*knowledge is freedom and ignorance is slavery.*"

-- The above quote is attributed to Miles Dewey Davis III (1926 – 1991):
http://www.goodreads.com/author/quotes/54761.Miles_Davis

Highly Recommended Reference

Kaplan, D. (2014). *Bayesian Statistics for the Social Sciences*. New York: The Guilford Press.

Other Important Resources

Albert, J. (2007). *Bayesian Computation with R.* New York: Springer Science + Business Media, LLC.

Berry, D. A. (1996). *Statistics: A Bayesian Perspective.* Belmont, CA: Wadsworth Publishing Company.

Berry, S. M., Carlin, B. P., Lee, J. J., & Muller, P. (2011). *Bayesian Adaptive Methods for Clinical Trials.* Boca Raton, FL: Taylor & Francis Group, LLC.

Bolker, B. M. (2008). *Ecological Models and Data in R.* Princeton, NJ: Princeton University Press.

Bolstad, W. M. (2004). *Introduction to Bayesian Statistics.* Hoboken NJ: John Wiley & Sons, Inc.

Broemeling, L. D. (2007). *Bayesian Biostatistics and Diagnostic Medicine.* Boca Raton, FL: Taylor & Francis Group, LLC.

Congdon, P. (2005). *Bayesian Models for Categorical Data.* West Sussex, UK: John Wiley & Sons, Ltd.

Congdon, P. (2006). *Bayesian Statistical Modeling.* West Sussex, UK: John Wiley & Sons, Ltd.

Dey, D. K., Ghosh, S., & Mallick, B. K. (2011). *Bayesian Modeling in Bioinformatics.* Boca Raton, FL: Taylor & Francis Group, LLC.

Efron, B. (1986). Why isn't everyone a Bayesian? *The American Statistician, 40*, 1 – 5.

Gelman, A., & Hall, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York: Cambridge University Press.

Gelman, A., & Meng, X. (2004). *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives.* West Sussex, UK: John Wiley & Sons, Ltd.

Gelman, A., & Shalizi, C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology, 66*, 8 – 38.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (2nd ed.).Boca Raton, FL: Chapman & Hall/CRC.

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics.* Hoboken, NJ: John Wiley & Sons, Inc.

Ghosh, J. K., Delampady, M., & Samanta, T. (2006). *An Introduction to Bayesian Analysis: Theory and Methods.* New York: Springer Science + Business Media, LLC.

Hoff, P. D. (2009). *A First Course in Bayesian Statistical Methods.* New York: Springer Science + Media, LLC.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences.* West Sussex, UK: John Wiley & Sons, Ltd.

Jeffreys, H. (1939). *Theory of Probability* (1st ed.). London: Oxford University Press.

Jeffreys, H. (1948). *Theory of Probability* (2nd ed.). London: Oxford University Press.

Koop, G. (2003). *Bayesian Econometrics.* Hoboken, NJ: John Wiley & Sons, Inc.

Koop, G., Poirier, D., & Tobias, J. (2007). *Bayesian Econometric Methods.* New York: Cambridge University Press.

Kruschke, J. K. (2011). *Doing Bayesian Data Analysis.* Burlington, MA: Academic Press.

Lancaster, T. (2004). *An Introduction to Modern Bayesian Econometrics.* Malden, MA: Blackwell Publishing.

Lee, P. M. (2004). *Bayesian Statistics: An Introduction* (3rd ed.). New York: Oxford University Press Inc.

Link, W. A., & Barker, R. J. (2010). *Bayesian Inference with Ecological Applications.* London: Academic Press (Elsevier Ltd.).

Lynch, S. M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists.* New York: Springer Science + Business Media, LLC.

Mallick, B., Gold, D. L., & Baladandayuthapani, V. (2009). *Bayesian Analysis of Gene Expression.* West Sussex, UK: John Wiley & Sons, Ltd.

Marin, J., & Robert, C. P. (2007). *Bayesian Core: A Practical Approach to Computational Bayesian Statistics.* New York: Springer Science + Business Media, LLC.

McGrayne, S. B. (2011). *The Theory that Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy.* New Haven, CT: Yale University Press.

Rossi, P. E., Allenby, G. M., & McCulloch, R. (2005). *Bayesian Statistics and Marketing.* West Sussex, UK: John Wiley & Sons, Ltd.

Rouder, J. N., Speckman, P. L., Sun, D., & Morey, R. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225 – 237.

Sorensen, D., & Gianola, D. (2002). *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics.* New York: Springer Science + Business Media, LLC.

Stone, L., D. (1975). *Theory of Optimal Search.* Mathematics in Science and Engineering, Vol. 118. New York: Academic Press, Inc.

Tanner, M. A. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions.* New York: Springer-Verlag, Inc.

Taroni, F., Bozza, S., Biedermann, A., Garbolino, P., & Aitken, C. (2010). *Data Analysis in Forensic Science: A Bayesian Decision Perspective.* West Sussex, UK: John Wiley & Sons, Ltd.

Williamson, J. (2010). *In Defense of Objective Bayesianism.* Oxford, UK: Oxford University Press.

Woodworth, G. G. (2004). *Biostatistics: A Bayesian Introduction.* Hoboken, NJ: John Wiley & Sons, Inc.

Originally published January 2015 -- Please note that information published in *Benchmarks Online* is likely to degrade over time, especially links to various Websites. To make sure you have the most current information on a specific topic, it may be best to search the UNT Website - http://www.unt.edu . You can also consult the UNT Helpdesk - http://www.unt.edu/helpdesk/. Questions and comments should be directed to benchmarks@unt.edu.

Staff Activities

April 15-June 15, 2016

Arrivals

- **Yuguang Ma**, IT manager, High-Performance Computing, Research IT Services, General Academic Building, late May
- **Robin Melendez-Martin**, pictured right, administrative coordinator, University IT, Sage Hall, early June

Departures

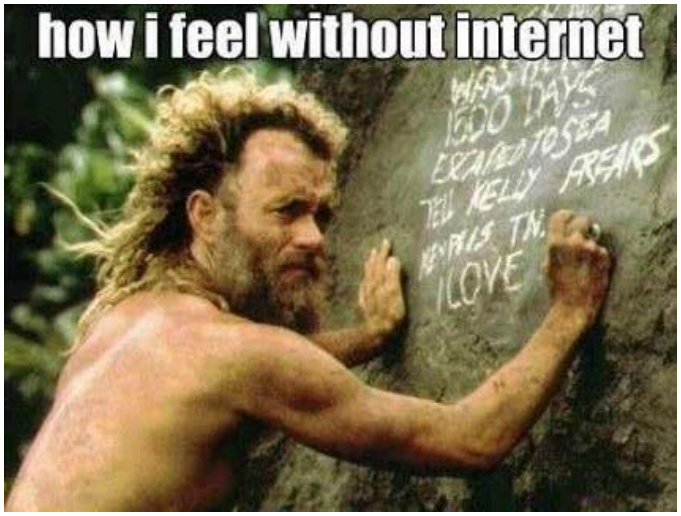
- **Cody Hardi**, administrative coordinator, late April

Congratulations to the May 2016, UNT graduates!

- **Katrina Carpenter**, IT Specialist, UIT Help Desk and Adaptive Computer Lab – Bachelor of Arts in Professional and Technical Communication
- **Rene Hernandez**, UIT Help Desk: Bachelor of Science in Electrical Engineering
- **Andrew Hoyt**, UIT Help Desk: Bachelor of Science in Geography
- **Amanda Johnson**, Sage Hall Adaptive Computer Lab, Bachelor of Science in Biology
- **Katie Welch**, UIT Help Desk: Bachelor of Science in Mechanical and Energy Engineering

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Today's Cartoon



Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu

UIT Reading List

Send in your reading list to share; suggestions may be education- or tech-related – or not.

June 2016

Name of Contributor | Name of Reading Material

Anonymous: Magic and Loss: The Internet as Art, by Virginia Heffernan

Jeff Anderson, systems administrator, CWS: Warhammer 40K Story Expansions

Jacob Flores, manager, host computing user services, UIT: The Stormlight Archive, series by Brandon Sanderson

Bill Gates, technical advisor, Microsoft: Hyperbole and a Half, by Allie Brosh; The Magic of Reality, by Richard Dawkins; What If?, by Randall Munroe; On Immunity, by Eula Biss; How to Lie With Statistics, by Darrell Huff; and Should We Eat Meat?, by Vaclav Smil.

Jennifer Lee, IT manager and director, student success technology, UIT: Top 100 literary classics

Jennifer Spillman, IT programmer/analyst, UIT: "Health Informatics for Medical Librarians," by Ana D. Cleveland and Donald B. Cleveland, and "No Excuses: The True Story of a Congenital Amputee Who Became a Champion in Wrestling and in Life," by Kyle Maynard

Jonathan Starkweather, research and statistical support consultant, UIT: The Lady Tasting Tea: How Statistics Revolutionized Science in the 20th Century, by David Salsburg

Editor's Note: Please note that information in each edition of Benchmarks Online is likely to change or degrade over time, especially links to various websites. For current information on a specific topic, search the UNT website, UNT's UIT Help Desk or the World Wide Web. Email your questions and comments to the UNT University Information Technology Department or call 940-565-4068.

Contact Us:
University Information Technology
1155 Union Circle #310709
Denton, TX 76203 USA
Voice: 940-565-4068
Fax: 940-565-4060

Visit Us:

Sage Hall, Room 338
http://it.unt.edu/benchmarks/

Email us:
Have questions on content or technical issues? Please contact us.
unt.uit@unt.edu