# Benchmarks Online

**Computing and Information Technology Center**

CITC Home | Help Desk | Training | About Us | Publications | Our Mission

Volume 11 - Number 6 * June 2008

## Columns

- NetworkConnection
- Link of the Month
- Helpdesk FYI
- IRC News
- RSS Matters
- Short Courses
- Staff Activities

- About Benchmarks
- Search Archive
- Subscribe to Benchmarks

ACUTA Member Sites to See

## Campus Computing News

**Computing Information for Faculty**

**Summer Hours**

**Information Security Awareness**

**Get it while it lasts! UNT Support of SkillPort Computer Based Training Ends November 2008**

**TODAY'S CARTOON**

**Don't forget our monthly Columns!**

### By the Numbers

**Personal Computers currently supported by Microcomputer Maintenance Shop (MMS):**

UNT PC - 9,985 - 66%
APPLE - 1,062 - 7%
DELL - 4,000- 27 %
Total - 15,047 - 100%

Please Note: The University of North Texas will never ask for personal information by e-mail. If you receive an e-mail purporting to be from the University that asks for personal information or account passwords, do not respond. If there is any question regarding the authenticity of an email, please contact UNT Information Security at (940) 369-7800.

Return to top

# UNIVERSITY OF NORTH TEXAS

Network Connection | Link of the Month | IRC News | RSS Matters | Helpdesk FYI | Short Courses | Staff Activities

Computing and Information Technology Center Home | Help Desk | Training | About Us | Publications | Our Mission

**Available for RSS/XML syndication. See the list of all available xml/rss feeds.**

Questions, comments and corrections for this site: lynch@unt.edu
Site was last updated or revised : June 17, 2008

information on a specific topic, it may be best to search the UNT Website - http://www.unt.edu . You can also search *Benchmarks Online* - http://www.unt.edu/benchmarks/archives/back.htm as well as consult the UNT Helpdesk - http://www.unt.edu/helpdesk/ Questions and comments should be directed to benchmarks@unt.edu

# Benchmarks Online

# Campus Computing News

## Computing Information for Faculty

By Dr. Philip Baczewski, Director of  Academic Computing and User Services

When faced with an IT problem or question we all sometimes wonder, "where can I get help on this?"  Often faculty have questions that are not typical of the day-to-day issues faced by the administrative staff at the University. The CITC's Academic Computing and User Services division has been one area to support faculty computing needs over the years. A number of years ago we created a document to help new faculty understand the IT services and support available specifically for them. With the number of recent changes in the UNT IT environment, including the migration from GroupWise to Exchange that was recently completed, it may be useful to make all faculty aware of an online resource created to answer the question of what help is available to them.

Help is readily available at the following web page:

http://www.unt.edu/documentation/newfaculty.htm

Some highlights from this page include:

- http://www.unt.edu/helpdesk

  The Helpdesk staff provides support to all faculty, staff, and students at UNT who have any questions and/or problems of any sort concerning computing at UNT.

- http://www.unt.edu/rss/

  ACS Research and Statistical Support  provides a wide range of services for UNT researchers (faculty and graduate students).

- http://www.unt.edu/ACSUNIX/cluster/

  ACS maintains a number of facilities specifically to enable computation-intensive scientific computing in support of funded research projects.

- http://ecampus.unt.edu  [updated 7/25/08]

  The CITC Distributed Learning Support group along with the Center for Distributed Learning (CDL) provides support for the Blackboard Vista online course management system. Faculty who need assistance with Blackboard Vista should contact the Center for Distributed Learning:

Email: vistafac@unt.edu
Phone: 940-369-7394
In person: Chilton 112C

- Sometimes you just have a problem with the computer on your desk. Desktop computing support is provided by a College or Department-level IT support organization. If you don't already know who to call, consult the following web site:

  http://www.unt.edu/helpdesk/netmanDepartments.htm

- Finally, the CITC has implemented a new IT support request system that has items specifically designed for employees to request IT service. Some, such as the request for CITC RSS services, are specific to faculty needs. You can make a request by visiting the CITC IT Help Center:

  http://helpdesk.unt.edu/service

There are many IT support resources for faculty at UNT. Sometimes, however, it's hard to find out what those might be. Hopefully, this information will be a handy reference when those difficult questions come up.

Return to top

# Benchmarks Online

**Page One**

**Campus Computing News**

**Summer Hours**

**Information Security Awareness**

**Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008**

**Today's Cartoon**

**RSS Matters**

**The Network Connection**

**Link of the Month**

**Helpdesk FYI**

**WWW@UNT.EDU**

**Short Courses**

**IRC News**

**Staff Activities**

Subscribe to Benchmarks *Online*

# Summer Hours

**By Claudia Lynch**, *Benchmarks Online* Editor

The spring semester has ended and SUM , 3WK1, 8WK1 classes* began on Monday, May 12. Following are the hours for Computing and Information Technology Center-managed facilities during this time period and, in some cases, the summer. The Helpdesk plans, at this point, to be open their normal hours, including on May 26 and July 4. The University is officially closed for Independence Day, July 4.

- The **ACS General Access/Adaptive Lab** (ISB 110):

    **May 12 - August 8:**

    Sunday: 1 p.m. – 9:45 p.m.
    Monday – Thursday: 8 a.m. – 9:45 p.m.

    Friday: 8 a.m.– 8:45 p.m.

    Saturday: 10 a.m. – 8:45 p.m.

**Exceptions:**

    Friday, July 4: **Closed**
    Friday August 8: 8 a.m. – 5 p.m.
    Saturday, August 9: 10 a.m.– 5 p.m.

## Hours for Other Campus Facilities

### General Access Labs

- WILLIS:

    Open normal 24 hour schedule.

- SLIS:

    **May 12 - August 8:**
    Monday - Thursday: 10 a.m. 10 p.m.
    Friday: 10 a.m. - 6 p.m.
    Saturday: 10 a.m. - 6 p.m.
    Sunday : Noon - 8 p.m.

    **Exceptions:**
    Friday, July 4: **Closed**

MUSIC:

**May 12 - August 8:**
Monday - Thursday: 8 a.m. 9 p.m.
 Friday: 8 a.m. - 5 p.m.
 Saturday: 10 a.m. - 5 p.m.
 Sunday : 1 p.m. - 8 p.m.

- PACS Computing Center (Chilton Hall):

**May 12 - August 8:**

Sunday: Noon - 10 p.m.
 Monday - Thursday: 8 a.m. - 10 p.m.
Friday - Saturday: 8 a.m. - 6 p.m.
August 9-August 24: **Closed**
Monday, August 25: Resume normal hours

- SOVA:

Hours currently unavailable.

- COE:

Open normal hours except:

Friday & Saturday, July 4 & 5: **Closed**
August 9-August 24: **Closed**

- COBA:

Open normal hours.

**Exceptions:**
 Friday, July 4: **Closed**

- CAS:

**GAB 330**

> **June 2 - August 8:**
> Monday - Thursday: 8 a.m. - Midnight
> Friday: 8 a.m. - 5 p.m.
> Saturday: Noon - 8 p.m.
> Sunday: Noon - Midnight
>
> **Special Closings:**
> Independence Day: July 4
> Semester Break: August 9 - 24

**GAB 550 - Closed**

**Terrill 220**

> **June 2 - August 8:**

Monday - Thursday: 8 a.m. - 8 p.m.
Friday: 8 a.m. - 5 p.m.
Saturday: **Closed**
Sunday: **Closed**

**Special Closings:**
Independence Day: July 4
Semester Break: August 9 - 24

**Wooten 120**

**June 2 - August 8:**
Monday - Thursday: 8 a.m. - 10 p.m.
Friday: 8 a.m. - 5 p.m.
Saturday: **Closed**
Sunday: **Closed**

**Special Closings:**
Independence Day: July 4
Semester Break: August 9 - 24

- UNT [Dallas Campus](#) - 155A

  Open normal hours.

- **[Engineering](#) General Access Lab ([englab@unt.edu](mailto:englab@unt.edu), Research Park, B129, 891-6733)**

  **June 2 - August 8:**

  Monday - Thursday: 8 a.m. - 10 p.m.

---

* Terminology and schedules for classes offered in the summer has changed in recent years:

SUMmer=Entire Summer Session, 3WK1 = 3-week 1, 8WK1=8-week 1, 5WK1= 5-week 1, 10WK1= 10-week, 5WK2= 5-week 2. All summer sessions will be over by August 8 [this year](#).

- Summer Session 3W1**:** formerly May Minimester
- Summer Session 5W1: formerly Summer I
- Summer Session 5W2: formerly Summer II.

[Return to top](#)

# Benchmarks Online

# Information Security Awareness

By **Charlotte Russell, Director, Administration and Compliance (CITC), Information Security Officer (UNT)**

**F**aculty and staff who come into contact with University data (e.g., student records, personnel information, financial data, etc.) are strongly encouraged to either attend a classroom-based security awareness course or take the online security course. Both are offered by the CITC Information Security group. If you handle some type of protected information, these courses will help you to become more aware of how important it is to ensure the security of university information.

You can learn more about information security in a number of ways:

- Attend one of the monthly security classes offered by Information Security.

- Contact Gabe Marshall to register for the online "Defensive Computing" course.

- Visit the Information Security website: http://security.unt.edu/

- Contact Gabe Marshall for information about other training opportunities (Gabe.Marshall@unt.edu or dial ext: 4062. ).

Here are a few examples of methods that you can use to help ensure the protection of data and resources:

1. Save important files to the network rather than to your hard drive.

2. Use strong passwords and avoid using your UNT passwords on external systems.

3. Don't reply to e-mail requesting personal information, even if the source seems legitimate.

4. Keep your software up-to-date.

5. Social security numbers have been replaced by the empl id and should

*never* be collected or stored on desktops, workstations, or on web servers.

6. Credit cards numbers should never be transmitted via unencrypted means (e.g., e-mail, web forms, etc.). Departments who process credit card information must be authorized by UNT's Student Accounting and University Cashiering Services department.

7. If one of your job functions requires you to handle sensitive data, ensure that it is transmitted via secure channels only (ex: ssh, ipsec, ssl, etc.).

8. Research software and obtain permission from your supervisor and network manager before you install it. Look for known vulnerabilities by using websites such as www.secunia.com, www.securityfocus.com, or even popular search engines.

9. When using ssh, remember to disable root logins, use strong passwords (or even use key authentication instead of passwords), and filter connections to trusted sources.

10. If you are creating homegrown web applications, remember to keep best security practices in mind. Check the www.owasp.org community for current web application security standards or contact Central Web Services at cws@unt.edu for technical assistance.

11. Remember, if you administer a website, register it at http://web3.unt.edu/siteregistration. See this recent article in *Benchmarks Online* for background information on this policy.

12. Subscribe to an Information Security RSS Feed from http://security.unt.edu/news/general or http://security.unt.edu/news/vulnerabilities The feed icon is located at the bottom of each of the pages.

Return to top

# Benchmarks Online

Page One

Campus Computing News

Summer Hours

Information Security Awareness

**Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008**

Today's Cartoon

RSS Matters

The Network Connection

Link of the Month

Helpdesk FYI

WWW@UNT.EDU

Short Courses

IRC News

Staff Activities

Subscribe to Benchmarks *Online*

# Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008

By **Dr. Elizabeth Hinkle-Turner**, **Student Computing Services Manager**

As of November 2008, formal computer-based training for the UNT community via the SkillPort system will end. Though SkillPort and its predecessor, SmartForce, has served us well for almost ten years, there are now so many free online offerings and other training services provided at UNT that it is not cost-effective to keep this system in place as well.

In anticipation of this future move, I have presented a series of 'Free CBT' articles in *Benchmarks Online* including discussion of the extensive Safari technical books holdings of the UNT Library as well as the comprehensive no-charge and engaging training at both the Adobe and the Microsoft sites. The staff in Academic Computing and User Services will continue to track, update, and report on such offerings in the future.

In the meantime, use SkillPort while we have it! You can still request accounts from me - the CBT manager - at ehinkle@unt.edu. This is especially crucial at this time when various UNT offices will be moving towards the Microsoft SharePoint environment. SkillPort has a variety of course offerings in SharePoint.

However, to get everyone started in this new paradigm, I draw your attention to the past *Benchmarks Online* articles on free online training resources:

- Free and Legal: Copyright Advice and Training Online

- Adobe Provides Thorough Online Instruction at No Cost

- Using the Adobe Education Website - Revised November 2005

- No-Hassle CBT: Library Online Tech Book Resources

- Windows Vista Learning Resources

- Get Revved Up for Office and Outlook 2007!

Originally published, June 2008 --  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 search **Benchmarks Online** - http://www.unt.edu/benchmarks/archives/back.htm as well as consult the UNT Helpdesk - http://www.unt.edu/helpdesk/ Questions and comments should be directed to benchmarks@unt.edu

Return to top

# Benchmarks Online

**Page One**

**Campus Computing News**

**Summer Hours**

**Information Security Awareness**

**Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008**

**Today's Cartoon**

**RSS Matters**

**The Network Connection**

**Link of the Month**

**Helpdesk FYI**

**Short Courses**

**IRC News**

**Staff Activities**

**Subscribe to Benchmarks Online**
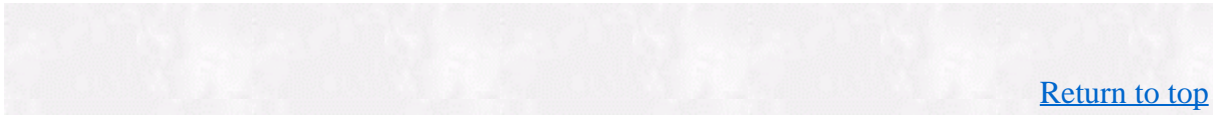
## Today's Cartoon

Copyright 2008 by Randy Glasbergen.
www.glasbergen.com

GLASBERGEN

"I prefer online dating. Deleting someone with one click is less exhausting than a long and painful breakup!"

From "Today's Cartoon by Randy Glasbergen", posted with special permission. For many more cartoons, please visit **www.glasbergen.com**.

Today's Cartoon

# Benchmarks Online

# Don't Forget Our Monthly Columns!

**By Claudia Lynch**, *Benchmarks Online* Editor

In addition to our feature articles, *Benchmarks Online* publishes monthly columns that are focused on specific aspects of computing here at UNT (and beyond, in some cases). Check out what is waiting for you this month:

- **By the Numbers** - Not really a column, but a new feature, giving you a glimpse behind the scenes of the volumes of data, spam, etc. processed, managed, and otherwise handled here at UNT.

- **RSS Matters** - "RSS Matters" is the monthly column written by the Research and Statistical Support Group in Academic Computing Services. Their articles focus on topics of a statistical and/or research methods nature. **This month, Dr. Mike Clark talks about "Getting Started with a Modern Approach to Regression." Check it out!**

- **The Network Connection** - "The Network Connection" may well be the longest running column in computer publishing history. Certainly in University of North Texas computer publishing history.

  **This month, Dr. Baczewski discusses document types in his article "Brand X." Click on the Network Connection link above for details.**

- **Link of the Month** - As it says on the top of the "Link of the Month" page, "Each month we highlight an online mailing list or website. Frequently the link is associated with UNT." **This month the Messaging Systems Group education site is featured. Click on the link above and check it out.**

- **Helpdesk FYI** - A new monthly feature from the CITC Helpdesk. Each month they will tackle a topic that has been of particular interest to callers/visitors to the Helpdesk. **This month  Jonathan "Mac" Edwards
  talks about "Outage Calendars." Check it out!**

- **Short Courses** - Every semester, Academic Computing Services (ACS) offers short courses on computer-related topics, many of them having to do with statistical research. This column keeps you up-to-date on what is being offered and when as well as other training opportunities.

  **Short Courses are close to being scheduled but aren't quite set. There are still lots of training opportunities, though. Also, if you have a group that needs a specific class, it may be possible to arrange a special class just for them. Click on the Short**

**Courses link above for more information.**

- **IRC News** - As their Webpage says, "the IRC is an advisory and oversight body created to foster communication and cooperation between and among UNT information resources providers and users." We publish the minutes of the IRC meetings each month, when they are available. **Minutes are published for April 15 and May 20 this month.**

- **Staff Activities** - This column focuses on new employees, people who are no longer employed at the Computing and Information Technology Center, awards and recognitions and other items of interest are featured here.

Return to top

# Benchmarks Online

# Network Connection

By **Dr. Philip Baczewski**, Director of  Academic Computing and User Services

## Brand X

**I**f you've used the newest version of Microsoft Office (2007 for Windows or 2008 for Mac OS), you may have noticed that it supports a new set of file formats for the various document types, all ending in "X". "X" used to stand for eXperimental (remember the X-1 and the X-15?) Now it stands for "XML", which is an eXtensible markup language for organizing or representing various kinds of online data.

Whether you realize it or not, there's been a bit of a storm brewing over document format. Back in 2005, the State of Massachusetts was considering adoption of open formats for State documents. This has wider implications for the applications used to produce documents. If an application doesn't support an open format, use of such an application would be discouraged.

For years, Microsoft Office document formats have been the de facto standard. Other applications have featured the ability to read and write documents in the Microsoft format. The new idea from the Massachusetts's initiative was that documents should be stored in a format that is agreed upon by a public standard-setting organizations. At the time, the Open Document Format (ODF), supported by the OpenOffice application suite, was one of the few standards that were acceptable under the new proposal.

Of course, if it's standards we want, it's standards that Microsoft will give us, at least their version. OOXML is the name of the format used for those "X" files written by Office 2007. Microsoft even got it adopted as a standard by Ecma, a European private standards organization. More recently, the OOXML was apparently accepted as a standard by the ISO, however, despite Microsoft's heralding approval of the standard, it has hit a snag in its approval. Because of appeals by four countries, ISO has put publication of OOXML on hold pending resolution of the appeals process.

Even if a standard is available for implementation by any software developer, it does not mean that it is particularly useful to the online community.  In a white paper comparing ODF and OOXML, Edward Macnaghten writes, "We are of the view that the format appears to be designed by Microsoft for Microsoft products, and to inter-operate with the Microsoft environment. Little thought appears to have been exercised regarding interoperability with non-Microsoft environments or compliance with established vendor-neutral standards." Others have alleged that Office 2007 is not even compliant with the OOXML standard as defined by Microsoft in their ISO submission.

All this talk of standards may seem esoteric, but it has a definite affect on the software we use on a daily basis.  Microsoft's pattern in the past has been to introduce new file formats

with the introduction of new software versions.  This has tended to cause confusion during a transition process between versions and has tended to force people or organizations to adopt the new version lest they be left behind and unable to read the new format. If all software used the same format for file storage, then competition would be based on features, and not on formats. This is true for HTML editors. Why can't it be true for Word Processors?

Return to top

# Benchmarks Online

-

# Link of the Month

*Each month we highlight an online mailing list or website. Frequently the link is associated with UNT.*

## Messaging Systems Group

The CITC Messaging Systems group has an education page containing a wealth of information for faculty and staff members. For example:

- **Running Outlook from Home**

- **Outlook/Entourage Tutorials** - A collection of tutorials and links online for users to become familiar with Microsoft Outlook, Microsoft Entourage for Mac, and Apple Mail.

- **Outlook Basic and Intermediate Classroom Instruction** - If you are interested in learning the Outlook 2007 client in a classroom setting, you will find information and materials relating to these classes listed here online

Check it out: **http://messaging.unt.edu/education**

The website **ncs.unt.edu,** which used to house information about GroupWise and the Outlook migration is now defunct. The Messaging Systems Group website should be consulted from now on for information relating to electronic communication systems at UNT.

---

Originally published, June 2008 -- 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 search **Benchmarks Online** - http://www.unt.edu/benchmarks/archives/back.htm as well as consult the UNT Helpdesk - http://www.unt.edu/helpdesk/ Questions and comments should be directed to benchmarks@unt.edu

Return to top

# Benchmarks Online

**Page One**

**Campus Computing**

**Summer Hours**

**Information Security Awareness**

**Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008**

**Today's Cartoon**

**RSS Matters**

**The Network Connection**

**Link of the Month**

**Helpdesk FYI**

**Short Courses**

**IRC News**

**Staff Activities**

**Subscribe to Benchmarks _Online_**

# Helpdesk FYI

**By Jonathan "Mac" Edwards, Assistant Manager of the CITC Helpdesk**

## Outage Calendars

**W**ith the upgrade to the new ITSM 7 reporting system, a new calendar feature has been added. In the past scheduled outages, and campus wide outages were listed in a table.

These same outages will now be found on an easily read calendar.

Three versions of this calendar are currently published.

1. Broadcasts and Critical Outages
2. Critical Resolved Outages
3. Combination of all Broadcasts and Outages (resolved and active)

## Event Types

One great feature of the new calendar is the Event Type selection.  Here we see the Event Type selection for the combined calendar.  This allows you to select which events you would like displayed on the calendar.

Unscheduled Outages: Outages which were not scheduled at an earlier date.

Scheduled Outages: Outages scheduled at an earlier date.

Broadcasts:  Used to inform users regarding system issues.  Generally not used for outages.

## Mini Calendar

The mini calendar sits above Event Types. This allows users to quickly select the month, and then day of the Outage or Broadcast in question. When clicking on a specific day the Calendar will display information for that day listed hourly.

## View Selection Tabs

These tabs, located in the upper right hand corner, allow you to toggle between different view modes.  Day will display an hourly breakdown of all outages and broadcasts.  Note that all day events will be displayed above the

time keeping area. The Week tab shows a less detailed breakdown of events, and the Month tab will give a broad over-view.

## Viewing Events

To view even details double click on the event. A window will pop-up listing important ticket information. A url is included allowing users to access the ticket if you have the correct permissions.

**Broadcasts and Critical Outages**

| | |
|---|---|
| Event Type | Unscheduled Outages |
| Name | Eaglemail unavailable. |
| Description | Accounts on eaglemail are being migrated to another mail server. Accounts on mailstore2 will be unavailable while this is happening and the system is expected to be online again by noon today. |
| URL **go→** | arsweb7.ars.unt.edu:8080/arsys/servlet/ViewFormServlet |
| Date | May 30, 2008 to May 30, 2008 |
| | ☑ All day event |

The new ITSM calendar is a very useful tool allowing users to quickly view issues reported within the ITSM system. You can find links to these calendars at the Helpdesk website ( http://www.unt.edu/helpdesk ), and the ITSM website ( http://itsm.unt.edu ).

Return to top

# Benchmarks Online

# IRC News

**Minutes provided by Sue Ellen Richey, Recording Secretary\***

## April 15, 2008

**VOTING MEMBERS PRESENT**:  PHILIP TURNER, CHAIR, TIM CHRISTIAN, ELIZABETH HINKLE-TURNER, JOHN HOOPER, GARY MATTHEWS, JUDITH ADKISON, RAMU MUTHIAH, SEAN MIKEL-FLOWERS (for JIM BYFORD), WIL CLARK (for JOHN PRICE), LOU ANN BRADLEY, BRUCE HUNTER, BILL JONES **NON-VOTING MEMBERS PRESENT**:   JOE ADAMO, PHILIP BACZEWSKI, SUE ELLEN RICHEY (Recording Secretary) **MEMBERS ABSENT**:  MAURICE LEATHERBURY, YUNFEI DU, FRANCES MAY, JON NELSON, DON GROSE, NOREEN GOGGIN, PATRICK PLUSCHT, WILL SENN, CENGIZ CAPAN, SCOTT WINDHAM, RAY BANKS, ABRAHAM JOHN, DONNA KEENER **GUESTS**:  CHARLOTTE RUSSELL

**J**udith Adkison moved for approval of the minutes of the March 11, 2008 meeting with a correction to the spelling of Wil Clark's name; Tim Christian seconded the motion, and the minutes were approved as corrected.

## Distributed Computing Support Management Team**

Philip Baczewski reported for the Distributed Computing Support Management Team that they met on April 4, 2008.  Representatives from Hewlett-Packard (HP) were on hand to provide an overview of some of their products which may be of interest to UNT.  They previewed a compact laptop computer which is smaller and lighter than others on the market.  They also provided an overview of the current HP printer line.  Krysta Kaye of CITC Data Communications presented information on the planned upgrade of the campus data network.  The five-year plan will result in the replacement of all data switches on the UNT campuses, with the older switches to be targeted first.  During the first year of this project, switches will be replaced in 20 buildings on campus, however, minimal network down time is expected. Also at this meeting, Philip Baczewski proposed that the new ITSM call tracking system be put into production operation as of May 12, 2008.  This date was accepted by the distributed managers and it was suggested that at least a week be provided to close out tickets in the old Remedy 5 call tracking system. The old system will be unavailable as of May 19, but all tickets from the old system will be available for reference in an archive view within the ITSM application. As an item of old business, Mike Hatch suggested that a group from DCSMT coordinate on  Exchange e-mail management now that a number of areas have made the migration from GroupWise. It was proposed that this idea be discussed at greater length at a future meeting of DCSMT.

Lou Ann Bradley asked for a status report of the Exchange migration.  Tim Christian said that Arts & Sciences is scheduled for May and June.  Judith Adkison reported that Education has completed their migration. Engineering is scheduled to migrate at the end of April.

## Learning Enhancement Planning Group

Dr. Turner reported for the Learning Enhancement Planning Group that the committee has labored diligently on the project to replace or renew WebCT Vista.  They are on track to bring their recommendation to the IRC at the May meeting.

## Communications Planning Group

Lou Ann Bradley reported that the Communications Planning Group will meet next Tuesday, April 22, 2008, and discuss VPN.

## EIS Planning Group

John Hooper reported for the EIS Planning Group that they are planning for the Financial upgrade which is scheduled for next spring.  They are starting preliminary planning to include the south Dallas campus in the EIS system.  They are about to place an order for more hardware for EIS; it has become apparent that additional memory is needed to improve response time.  The additional hardware is also needed to support a Blackboard/Vista database.  They are also working on a disaster recovery plan, which will involve getting a database machine for the Research Park for backup.  In addition, they are working on increasing the speed of the Cognos reports, and feel they are close to finding the cause of the slowdown.

## Standards & Policy Planning Group

Tim Christian reported for the Standards & Policy Planning Group that they are looking at the Project Management Policy and the handbook that goes with it.  They will run this by the DCSMT before bringing it to the IRC, but prior to that they need people to volunteer to use the Project Management program for a project over $10,000 (either an upcoming project, or a completed project).  He asked people to let him know if they want to volunteer.

## Student Computing Planning Group

Elizabeth Hinkle-Turner reported for the Student Computing Planning Group that they are finishing up their plans to implement a new student email system.  Microsoft Live has been tentatively chosen, and Elizabeth will be presenting it to the SGA next Wednesday.

There being no further business, the meeting was adjourned at 2:20 pm

# May 20, 2008

**VOTING MEMBERS PRESENT**:   PHILIP TURNER, CHAIR, TIM CHRISTIAN, ELIZABETH HINKLE-TURNER, JOHN HOOPER, JUDITH ADKISON, JIM BYFORD, LOU ANN BRADLEY, BRUCE HUNTER, BILL JONES, YUNFEI DU, FRANCES MAY, JON NELSON, DON GROSE, PATRICK PLUSCHT, CLEATUS ESTES (for WILL SENN)**NON-VOTING MEMBERS PRESENT**:   MAURICE LEATHERBURY, JOE ADAMO, PHILIP BACZEWSKI, SUE ELLEN RICHEY (Recording Secretary) **MEMBERS ABSENT**:   JOHN PRICE, RAMU MUTHIAH, GARY MATTHEWS, NOREEN GOGGIN, CENGIZ CAPAN, SCOTT WINDHAM, RAY BANKS, ABRAHAM JOHN, DONNA KEENER, SEAN-MIKEL FLOWERS **GUESTS**:   CHARLOTTE RUSSELL, JANE HIMMEL, RICH ANDERSON

**T**im Christian moved for approval of the minutes of the Tuesday, April 15, 2008 IRC meeting; Don Grose seconded the motion and the minutes were approved as distributed.

## Distributed Computing Support Management Team**

Philip Baczewski reported for the Distributed Computing Support Management Team that they met on May 2, 2008. This meeting featured a discussion regarding DCSMT membership and communication issues between distributed managers and the CITC. Some managers expressed that it was unclear who to call for certain issues related to central systems. Distributed managers were encouraged to call the CITC Helpdesk if they had a question regarding campus IT systems. Philip Baczewski agreed to do further investigation regarding membership issues and to investigate possibilities for using the DCSMT website as a more comprehensive communications channel for distributed managers. Also at that meeting, Philip Baczewski reminded managers that the new ITSM 7 call tracking system would be made live to customers on May 12, 2008. He demonstrated the "Broadcasts and Outages" and "Resolved Outages" links on the itsm.unt.edu web site which open calendar views of campus-wide outage and broadcast notifications. Philip Baczewski thanked Chris Strauss and Don McClure for their work on the ITSM application. Special commendation was directed to Richard Sanzone for his work on the Kinetic service request items to be used by IT customers and to Mac Edwards for his development of items in the knowledge base that is integrated into ITSM 7.

DCSMT again met on May 16, 2008. Vonn Miller and Ken Gray from Apple Inc. were on hand to provide information on recent updates to the MacBook line of laptop computers and to provide an overview of running other operating systems, such as Microsoft Windows, on Apple hardware. Also at that meeting, distributed managers discussed the need for an IT Help Center service item to allow a supervisor to request modifications to a computer account for their employees, including making a request to provision an account for new employees. Richard Sanzone agreed to create the service request item for the IT Help Center and then accept feedback from distributed managers regarding features needed for such a request. It was mentioned that there was often considerable delay in getting an EUID created for new employees and it was agreed that this issue should be reexamined if possible by a broader group, since there have been key personnel changes in a number of areas supporting this process. Also at that meeting, Charlotte Russell provided an update on the Computer Access Modification System that will notify IT support staff when an employee's computer access needs to be modified or revoked. The system will be pilot tested within the CITC this summer and should be available to the entire campus by August.

DCSMT is next scheduled to meet on June 6.

## LMS Evaluation Project

Patrick Pluscht presented the final report on the LMS Evaluation Project, which was electronically distributed about 3 weeks prior to the meeting so that everyone would have time to review it and be ready to vote on the issue today. This final report fulfills the charge to investigate a replacement for the WebCT Vista product, and recommends the adoption of the BlackBoard Vista product. In view of the need for a timely decision on this matter, Patrick moved for a suspension of the rules of order, which would allow for a vote on this recommendation on the same day it is presented. Tim Christian seconded the motion to suspend the rules and the motion passed unanimously.

Patrick Pluscht moved for approval of the report as presented; Jon Nelson seconded the motion; and it passed by a unanimous vote.

Patrick distributed another document which detailed additional recommendations regarding the renewal of UNT's contract with Blackboard, Inc. The document was a memorandum to Maurice Leatherbury and Tom McElwee from Jane Himmel, in which she notes important criteria to be included in the contract with Blackboard when it is re-newed. This was presented as an information item only.

## Communications Planning Group

Lou Ann Bradley reported for the Communications Planning Group that the committee met and received an update on VPN noting that it should be out this summer. The committee discussed the wireless network noting that it was initially installed as a convenience for students, but that it has become the network of choice for many, to the point that it has become very slow. They are preparing a white paper on the wireless network. Joe Adamo added that there are new developments in the wireless field, but that have not yet been certified, so they do not want to be too quick to adopt them. He explained that this is a big issue and they are looking into it. When questioned about the possibility of adding more nodes to help the speed of connection, Joe replied that if there is a need for more nodes by any department, that department just needs to make that request to CITC Communications/Datacom.

## EIS Planning Group

John Hooper reported for the EIS Planning Group that they are continuing to plan for the financial upgrade that will occur next Spring, and have hired consultants to help with that. John also noted that in their work on increasing the speed of running Cognos budget reports, they have found that Cognos runs a lot faster under Windows, so they are going to try to move in that direction.

## Standards & Policy Planning Group

Tim Christian reported for the Standards & Policy Planning Group that they are continuing to review the Project Management Policy and handbook.

## Microsoft Exchange Live

Elizabeth Hinkle-Turner presented Microsoft Exchange Live at the SGA meeting on April 23rd, and said that it was warmly received. Students had some questions about whether or not it works on Macs and about forwarding capability, but overall were enthusiastic about it. Elizabeth said that some of the system engineers will come here next week to help the Messaging Team set it up. When asked when it will be available, Elizabeth stated that alumni accounts for 2008 graduates will first be provisioned; on June 19th Freshman orientation will be in the system. URCM will have a campaign to have people switch over and the target for shutting down the Eagle mail system would be January 2009. New graduate student accounts would be provisioned in Fall of 2008.

## IT Advisory Committee

Maurice Leatherbury presented the restructuring of the IT Advisory Committee for discussion. He gave a quick summary of the policy which will establish two advisory committees for IT on campus. This Council will be called IT Council with not much change in the present membership. The policy also reconstitutes the IR Steering Committee which was organized by the Provost and made up of Vice Presidents and the Chair of the IRC. The policy expands the membership to include some Vice Presidents, Deans, and several others. He explained that this will be a much broader committee than the old IR Steering Committee used to be. The IT Steering Committee (ITSC) will be a high level advisory committee that will help to shape the whole direction of IT on campus. Under the Project Management policy it will have responsibility to approve projects over a certain threshold amount. The

IT Planning Groups will have a hand in approving IT projects.  Maurice's presentation served as a motion for approval.  Tim Christian made a friendly amendment to the motion which changed the composition of the ITSC to only the Vice President for Academic Affairs, the Chief Technology Officer, the Vice President for Research, the Vice President for Finance, and the ITC Chair.  Lou Ann Bradley seconded the motion.  Maurice explained that he had originally proposed the broader representation in order to obtain more participation on decisions to make changes that are adopted campuswide.  Tim Christian then offered a second friendly amendment which proposed that like all other divisions represented on the ITC, that Enrollment Management be granted only one representative.  Since there was no further discussion, and Maurice accepted both amendments offered by Tim Christian, the Chair called for the question and the motion passed unanimously.  This recommendation will now be taken up the ranks for approval as an official policy.

Maurice Leatherbury reported that his committee is making rather slow progress on the desktop replacement plan.  He has met with a partial committee but the decision was that since all of the data was not available they would put off meeting anymore until complete data was available and the financial implications of the plan could be investigated.  The CITC staff is working on that at the present time.

There being no further business the meeting was adjourned.

---

\* For a list of IRC Regular and Ex-officio Members click here (updated 3/14/08).

\*\*DCSMT Minutes can be found here.

# IRC Meeting Schedule

The IRC generally meets on the third Tuesday of each month, from 2-4 p.m., in the Administration Building Board Room. From time to time there are planned exceptions to this schedule. The schedule can be found here. All meetings of the IRC, its program groups, and other committees, are open to all faculty, staff, and students.

Return to top

IRC News

# Benchmarks online

## Research and Statistical Support
### University of North Texas

# RSS Matters

*Link to the last RSS article here: [Mapping And Data Representation In Stata: Part 2](#) - Ed.*

## Getting Started with a Modern Approach to Regression

**By [Dr. Mike Clark](#), Research and Statistical Support Services Consultant**

Multiple regression is one of the most widely employed statistical techniques in the social sciences, and unfortunately given its popularity, it is often poorly implemented as well. One can pretty much open any journal in which some study uses the technique and they will find all manner of issues from no mention of assumptions, poor attempts to ascertain whether assumptions have been met, no detail on outlier analysis and even if so, no usage of modern techniques to deal with it, poor approaches to determining variable importance, lack of validation etc. The purpose of this article is to provide a few examples of how one can deal with such issues, and the statistical package of choice will be R. For introductions on how to use it, start with the applications link here [www.unt.edu/rss](http://www.unt.edu/rss) and click on R.

### The issues

The reasons for such poor practice probably stem from two things primarily: the first is simple lack of knowledge regarding the underlying issues, and the second is using poor software for the analysis. Regarding the first, many believe typical analyses are 'robust' to even moderate violations of assumptions. Now if you ask them what *robust* actually means you may get a variety of answers, but some will answer "robust to type I error" and *in general* that is a correct assessment. However specifically the answer would be 'usually' robust to type I error, but not always, sometimes inflating and in other cases leading to a more stringent alpha level than one would want. Now what about type II error, retaining the null hypothesis when it should have been rejected? Typically disaster. Depending on the situation even slight deviations from normality, homoscedasticity etc. can destroy statistical power (a simple sim to play around with can be found [here](#)), which is often much more of a problem in our analyses even when data situations are not problematic (Cohen, 1992). Furthermore, violations of assumptions may lead to biased and inefficient estimates, making inference suspect at best and impossible at worst if corrective measures have not been taken. This suggests that not only should assumptions be tested and outcomes reported as a common practice[1], something should be done in the face of those problems as they may result in missed effects, biased and/or inefficient estimates, incorrect probabilities and in general 'bad things *man*'. Regarding the second issue I will use SPSS's menu system as an example as a. SPSS is very popular in social and other science research and b. I venture to wager that most applied social scientists that use SPSS do not use syntax unless absolutely necessary unless they are of the 'ol' school' as the kids say (and often the syntax does not even provide for any more options). Furthermore, as academic research reporting implies only a relative few are using any of the expensive add-ons, this will assume only the base package options.

If one looks at the menu options (Figure 1) it at first may appear there are many things to choose from, and while some of them are things we want, there are also redundant options (10+ measures of outlierness might be considered overkill by some but I personally like having as many as possible), poor options (e.g. pairwise deletion would lead to biased results), and several things that would only be feasible if doing a particular type of analysis (sequential or stepwise procedure). Its main dialog box is standard fare choosing of the model and a sequential or stepwise approach, the latter of which itself is fairly limited compared to other packages that provide subsets regression and keeping variables based on a variety of fit indices that correct for model complexity rather than just statistical significance. But as a reminder, we are doing 'simultaneous' regression here. In the statistics box one has access to some good things like interval estimates for coefficients, partial and semi-partial correlations (SPSS calls the latter 'part'), and collinearity diagnostics. The plots box might be fairly daunting to the uninitiated, and one could easily create uninformative graphs. The SPSS help file won't be of much use, since it tells of only a single graph one could make but not how to interpret it. But at least you can create the graphs, though other packages provide meaningful ones by default. Many things are in the 'save' box but you would not need but a few of them, e.g. saving the residuals and a couple outlier measures such as Cook's distance would suffice. The prediction intervals are nice for graphical display but that's best left to the graphics menu as you won't get such a graph here.

**Figure 1.**
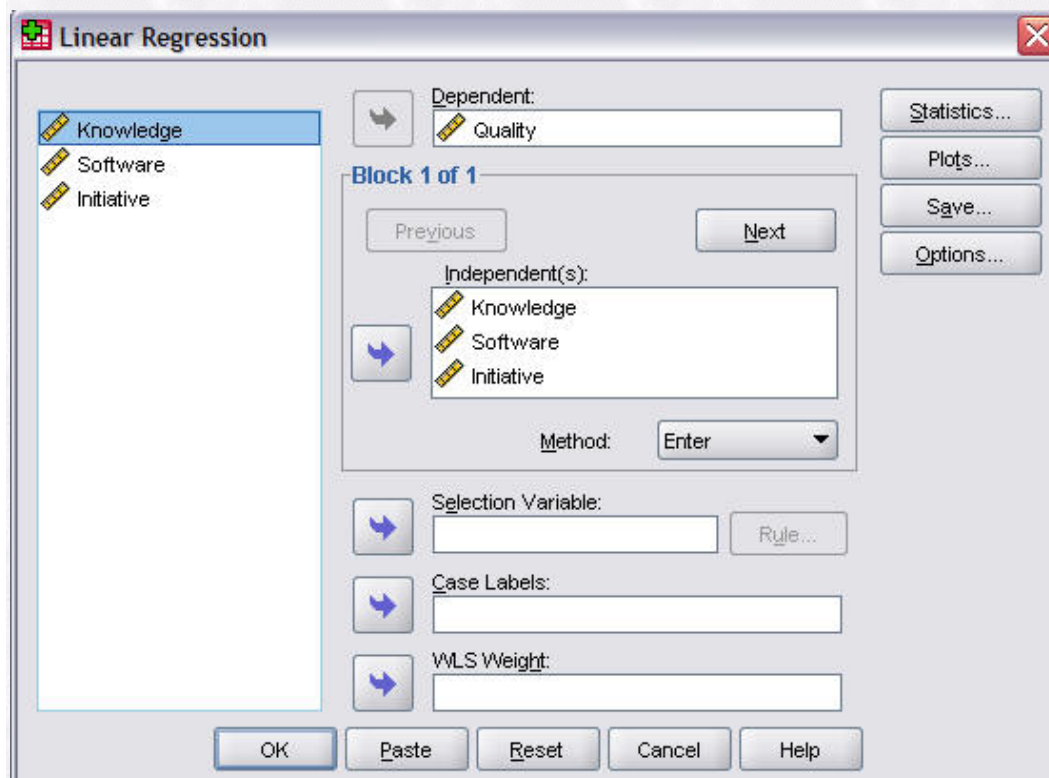


So, although there are several dialog boxes and seemingly many options, some useful, we have run into several problems.

1. One can get plots to assess normality and homoscedasticity assumptions, but no actual means to test the latter at all. For the former we'd have to save the residuals (not done by default) then assess those.

2. One can get outlier measures but no automatic visual display of influence.

3. There is no method of validating the model available.

4. There is no means to do anything if you actually have any problems (and that goes for the Regression syntax).

5. Options for stepwise methods and dealing with missing values are severely limited.

In the end, if an applied researcher even knew about the statistical issues with linear regression, if their statistical package know-how was limited to base SPSS menus/Regression syntax, there isn't any viable means to address nor deal with data problems. In short, a *good* regression analysis is not possible without serious programming skills, and additional, possibly inadequately tested macros and scripts one might obtain from the web. A good regression is possible in R even with minimal knowledge of the language and menus, and quite easily, so we turn to it now.

## Fitting the model

First off, the model. I have some made up data involving 4 variables: We'll call them *Knowledge* of statistical issues, Statistical *Software* expertise, and *Initiative*, with an outcome variable of *Quality* of analysis. The thing about regression is that we have to fit the model before we can test assumptions and search for outliers. For the following, relevant code will be in maroon, output in blue, comments in green. Links are provided to help documents for specific functions. We won't even need a specific package for fitting with the lm function, but we'll be some later.

```
#the name model i.e. "Model" is arbitrarily chosen by me

Model <- lm(QUALITY~INITIATI+KNOWLEDG+SOFTWARE, data=Dataset)

summary(Model)


Call:

lm(formula = QUALITY ~ INITIATI + KNOWLEDG + SOFTWARE, data = Dataset)


Residuals:

    Min       1Q   Median       3Q      Max
-2.46385 -0.69550 -0.03513 0.69538 2.00725


Coefficients:

             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.06079    0.58588  -3.517 0.000667 ***
INITIATI     0.05661    0.01229   4.605 1.26e-05 ***
KNOWLEDG     0.07079    0.01111   6.372 6.46e-09 ***
SOFTWARE     0.01381    0.01265   1.091 0.277881

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.9441 on 96 degrees of freedom

Multiple R-squared: 0.6158, Adjusted R-squared: 0.6038

F-statistic: 51.29 on 3 and 96 DF, p-value: < 2.2e-16
```

At this point, if you are planning on doing a good regression analysis, you wouldn't bother to look at the outcome, except maybe to note if there is anything seriously wrong. However, being human, most of us will, so go ahead. This model does seem to be fitting well (which makes sense since I made it up to do so), but the point is that everything that could be interpreted at this point could be wrong, from p-values to coefficients. So lets test our assumptions so that we can feel confident in our interpretation. I can do this with the R-commander menu (type library(Rcmdr) at the command line to bring it up) but the code will be listed.

## Diagnostics

The best way to start (and in my opinion, continue and end) is graphically, and graphs can be obtained through the *Models/Graphs/Basic diagnostic plots* menu, or simply typing plot(Model) at the command line.

**Figure 2.**
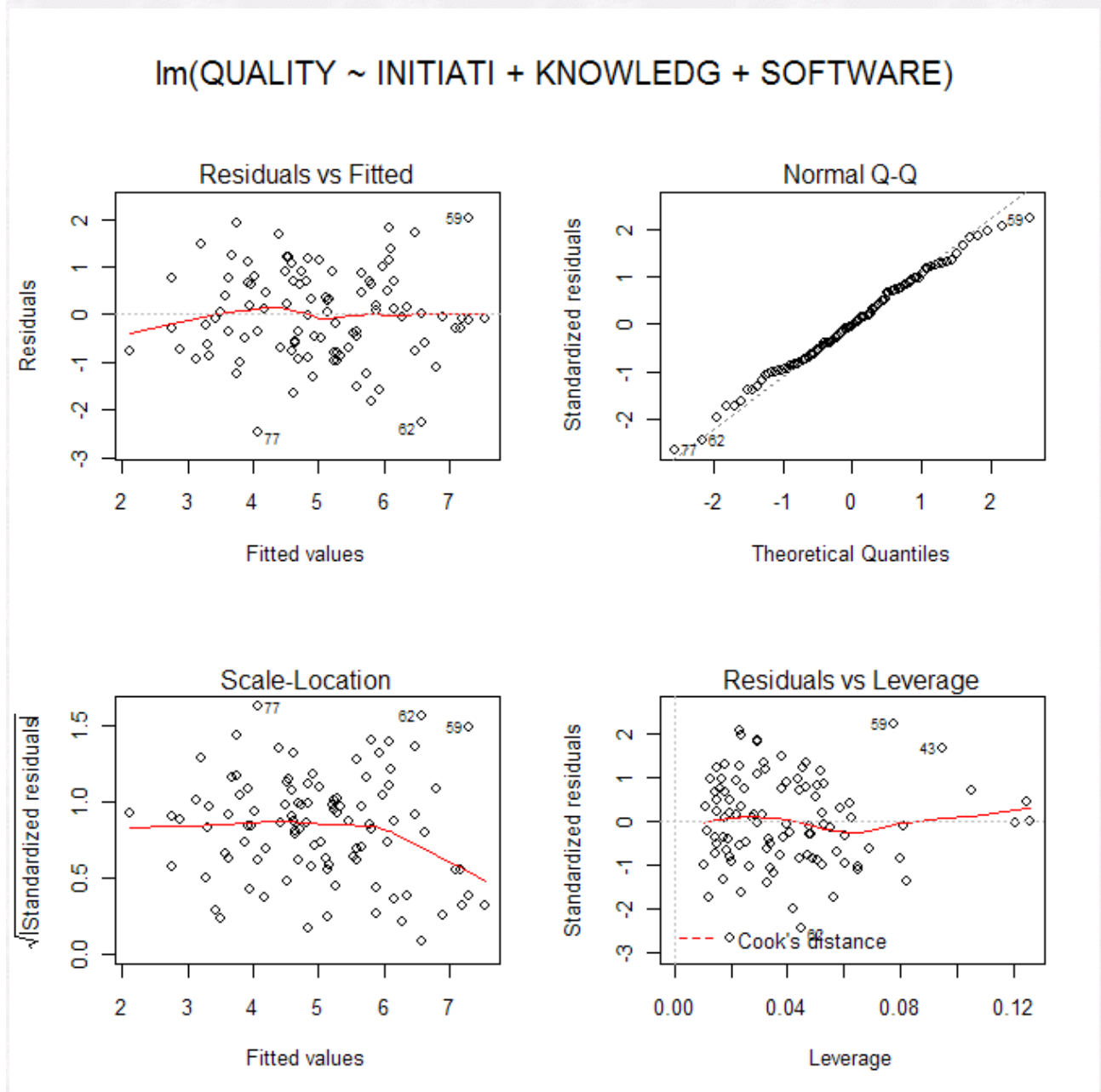


The Residuals vs. Fitted and Scale-Location are two versions of the same thing (the latter a more robust

of the former), and we are examining them with regard to two assumptions, linearity and homoscedasticity. The line should look like a child's freehand attempt at a straight line. The scatter of dots should look like the blob we see, no curvilinear pattern or fanning out on one or both ends. We're doing ok at this point for the most part, but we still will probably want to check the statistical assessment. The Normal Q-Q plot should also look pretty much as it is, with all the points lying on the indicated line (they typically start to get loose at the ends). Again, visual inspection would suggest we've met that assumption. The last graph regards outliers, but there is one I like better so I'll postpone discussion of it for now. First let's get actual statistical tests for the those graphs to back up what we see.

One can access the R-commander menus (Models/Numerical diagnostics/Breusch-Pagan...) but I will also provide the associated code. The following provides the Breusch-Pagan test for heteroscedasticity.

```
#As I've already got 'Model' specified, I don't need to put the formula.

library(lmtest)

bptest(Model, varformula = ~ fitted.values(Model), studentize=FALSE, data=Dataset)


Breusch-Pagan test

data: QUALITY ~ INITIATI + KNOWLEDG + SOFTWARE

BP = 0.1898, df = 1, p-value = 0.663
```

As with most tests of assumptions, statistical significance is not desired, but instead we will (illogically, actually) accept the null and go on our merry way. The following provides the Reset test for linearity and the Shapiro-Wilks test for normality (done on the residuals, which the normality assumption regards, not the predictors). The Reset test is in the same R-commander menu as the BP test, but while the Shapiro-Wilks test for normality is in the basic summaries menu, the residuals are not accessible as a variable there, and so testing that requires a visit to the command line. It is part of the base install of R and thus requires no additional package. A note for all you who have not been testing your multivariate normality assumption, there is mshapiro.test (and others) also that is just as easy to pull off.

```
#If you already had the lmtest library up, you would not need to call it again

     library(lmtest)      resettest(Model,      power=2:3,
type="regressor", data=Dataset)


      RESET test

 data: QUALITY  ~  INITIATI  +  KNOWLEDG  +  SOFTWARE
RESET = 1.0169, df1 = 6, df2 = 90, p-value = 0.4195


shapiro.test(Model$residuals)

      Shapiro-Wilk normality test

 data:  Model$residuals  W  =
0.9919, p-value = 0.8165
```

Again, we're doing fine as our graphics suggested earlier. The following is the Durbin-Watson test, which can assess whether there is serial correlation among the residuals, i.e. it is a means to test if our observations are independent. While usually linearity stays a theoretical assumption regarding the manner of data collection and the DW test is mostly seen with time-series data, I go ahead and provide it here. Again, it is available via the R-commander menus if one wishes.

```
dwtest(Model, alternative="two.sided", data=Dataset)

data: QUALITY ~ INITIATI + KNOWLEDG + SOFTWARE DW =
 1.8803,  p-value  =  0.5453  alternative  hypothesis:
true autocorelation is not 0
```

No problems there. We'll next assess multicollinearity, which is essentially redundant information

among the variables, such that one variable's variance is largely shared by the others. A way to test this is to simply regress the variable in question on the others and note the $R^2$ value. *Tolerance*, as given by SPSS for example, is 1- that $R^2$ just mentioned, and the *variance inflation factor* is the inverse of tolerance. Usually, if the VIF is in the neighborhood of 10, we might be concerned, as it would suggest that the vast majority of the variable's variance is accounted for by the other variables, and may lead to inefficient parameter estimates.

```
vif(Model)

INITIATI KNOWLDEG SOFTWARE
1.678325 1.370830 1.778163
```

Still doing well. Now back to that influence plot to help us assess potential outlying cases, which I obtained through that same Models/Graphs/ menu earlier or can get with influencePlot(Model) from the [car] package. The X axis is the leverage statistic, which is a measure of influence, the Y axis is the studentized residual value, and the size of the bubble reflects the case's Cook's distance. In this manner it provides three measures of 'outlierness' in one 2-d graph. Vertical reference lines are drawn at twice and three times the average hat value, horizontal reference lines at -2, 0, and 2 on the studentizedresidual scale. Case 59 for example, may be of concern. In any case we were going to run a robust version of regression for comparison anyway.

**Figure 3.**
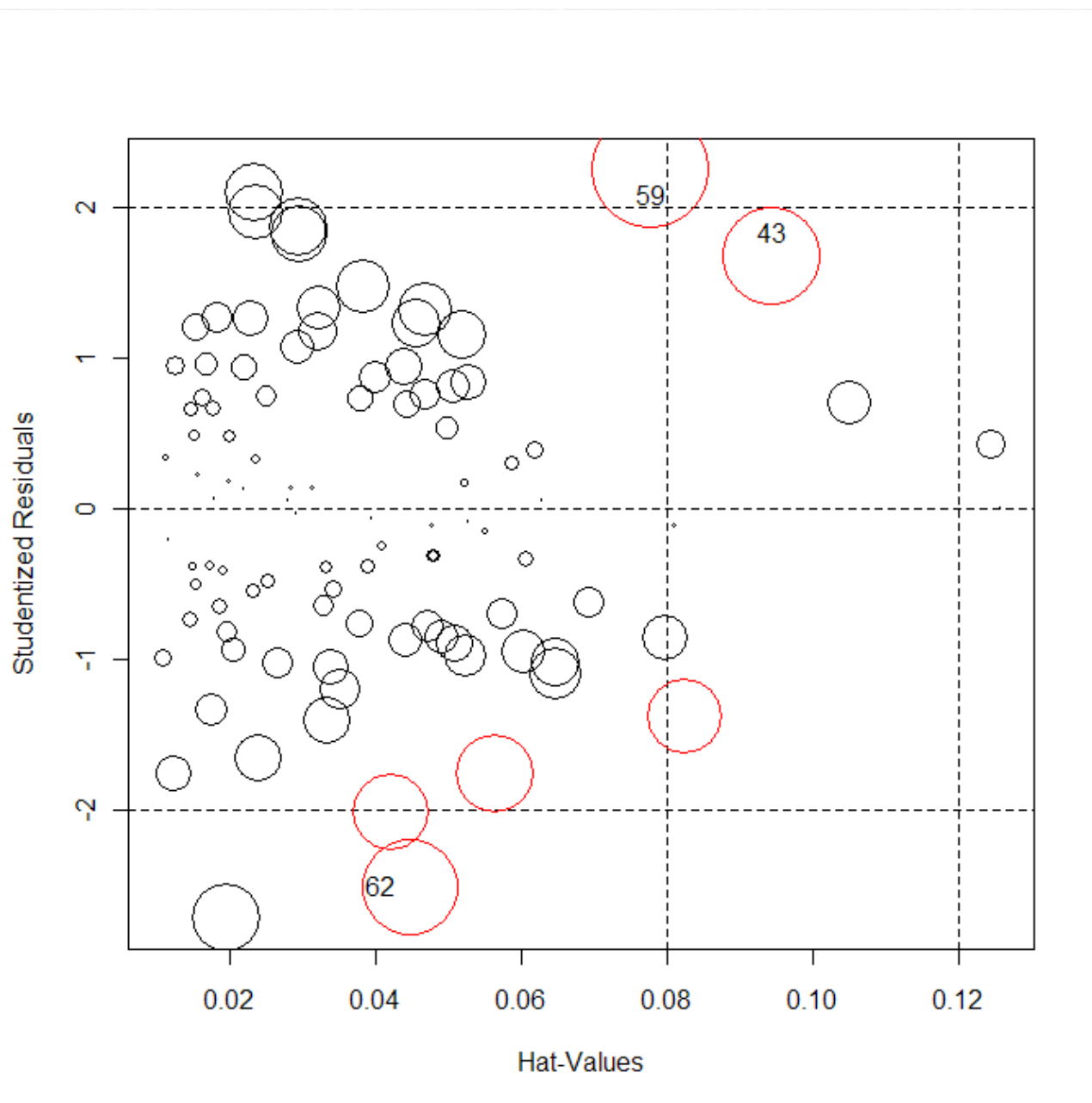
## Robust regression check

There are many ways in which to run a regression which is resistant to outliers and often performs better in heteroscedastic situations, and more are being developed (see Wilcox, 2001 for an introduction). Some are simple in concept, e.g. Least-trimmed squares regression, while others, for example those based on M-estimators, can get more technical. Here I follow Tukey's guideline:

> "… just which robust/resistant methods you use is not important – what is important is that you use some. It is perfectly proper to use both classical and robust/resistant methods routinely, and only worry when they differ enough to matter. But when they differ, you should think hard." (Tukey, 1979)

So for me, the thing to do is simply check, and R has whole packages devoted to robust techniques. I will use the 'robustbase' package for this example.

```
#modelrob is again an arbitrary name

library(robustbase)
```

```
modelrob=lmrob(Model)

summary(modelrob)


#Partial output

Coefficients:

           Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.04399    0.56972   -3.588 0.000527 ***

INITIATI     0.05526    0.01195    4.622 1.18e-05 ***

KNOWLEDG     0.07157    0.01045    6.851 6.99e-10 ***

SOFTWARE     0.01415    0.01253    1.130 0.261397

---

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



Robust residual standard error: 0.9919
```

Comparison to our previous output suggests hardly any change in coefficients, which, given the outcome of our previous diagnostic checks, was to be expected. Squaring the correlation of the robust fitted values to the Quality DV gave almost the same $R^2$ as the original. It is safe to assume at this point that our least squares estimates are okay and worth keeping.

### Validating the model and adjusting $R^2$

Some may see validation for stepwise exploratory endeavors though sadly often not even there, which renders them completely useless analyses in my opinion. However the point is that validation can and should be performed for *any* regression analysis that has an adequate sample size. Many are familiar with simple validation techniques, such as cross-validation using training and test sets. However sample sizes that are needed may be prohibitive with such approaches, and many others are not available in some standard statistical packages. The bootstrapping technique will allow us, again assuming you have a reasonable sample for the model in question to begin with, to use the cross validation technique multiple times by resampling (with replacement) from the original data set to create even several hundred training and test sets for validation. Details can be found in Harrell (2001), where further original references to Efron and others may be found also (Harrell created the [Design] package used here). In short, you are doing many, cross-validations, and getting an average estimate of bias in your $R^2$ metric.

```
library(Design)

valmodel=ols(formula=QUALITY~INITIATI+KNOWLEDG+SOFTWARE, data=Dataset, x=T, y=T)

validate(valmodel, method="boot", B=500)



Iteration 500
```

| | index.orig | training | test | optimism | index.corrected | n |
|---|---|---|---|---|---|---|
| R-square | 0.6158219 | 0.6188035 | 0.59984122 | 0.018962291 | 0.59685961 | 500 |
| MSE | 0.8557567 | 0.8286495 | 0.89135369 | -0.062704177 | 0.91846090 | 500 |
| Intercept | 0.0000000 | 0.0000000 | 0.01487108 | -0.014871076 | 0.01487108 | 500 |
| Slope | 1.0000000 | 1.0000000 | 0.99780735 | 0.002192654 | 0.99780735 | 500 |

The key statistic for our purposes regards the first row. Going across we see our original variance

accounted for, the average training set $R^2$, the average test set $R^2$, the optimism, which is the difference between training and test, and finally the corrected version of $R^2$ which is the optimism subtracted from the original. This serves as our bias-adjusted R square which is based on the specifics of the dataset rather than a heuristic as the typical "Adjusted $R^2$" reported is.

## Confidence Interval for the $R^2$

Like a mean or anything else we are estimating in a sample, point estimates do not allow us to get an estimate of the uncertainty in our guess, and effect sizes vary from sample to sample as do other statistics. To obtain that understanding of uncertainty we will construct a confidence interval for this $R^2$ using the MBESS package and basic information about the model such as the sample size and number of predictors as follows. Note that I have input the bias-corrected $R^2$.

```
library(MBESS)

ci.R2(R2=.597, N=100, K=3, conf.level=.95)

$Lower.Conf.Limit.R2
[1] 0.4475404

$Prob.Less.Lower
[1] 0.025

$Upper.Conf.Limit.R2
[1] 0.6996761

$Prob.Greater.Upper
[1] 0.025
```

So the 95% CI for the bias-adjusted $R^2$ is .448 to .700.

## Which variable is most important?

Satisfied with the model's integrity at this point, we might want to now take an explanatory approach to determine variable importance. With scales of different measures, raw coefficients can't provide this information by themselves, so we'll have to use something else. Many use standardized coefficients for this purpose, and if one is larger than another they claim it is more important. This is unsatisfactory, and equivalent to saying that just because one mean being higher than another entails it is statistically/meaningfully so. Given sampling variability, orderings that are not far apart have the potential to reorganize their ranks upon the next sample collected. The following uses a bootstrap approach to create confidence intervals around metrics of importance so that one can demonstrate statistical differences among them. The statistic of choice for me is the average semi-partial (Lindeman, Merenda, & Gold, 1980; Kruskal, 1987), which measures a variable's squared semi-partial at all possible points of entry into the model, and takes the average of those. It has the added bonus of decomposing $R^2$ into each variable's contribution to it. Furthermore, the relaimpo package will allow us to determine whether their contributions are statistically distinct.

```
library(relaimpo)

calc.relimp(Model)

Proportion of variance explained by model: 61.58%

Metrics are not normalized (rela=FALSE).

Relative importance metrics:

lmg

INITIATI 0.2138923

KNOWLEDG 0.2800459

SOFTWARE 0.1218837
```

These are the average semi-partials, called LMG after for the authors of the first known source of the

method. There are other options, e.g. Adding rela= TRUE would normalize them to sum to 100%, and each value would be the percentage of $R^2$ it accounts for. This may be preferable since emphasis would be put toward a bias corrected $R^2$ for the model. As it is above, they sum to the model's original $R^2$. Knowledge of statistical techniques *seems* to be most important in the model, but let's find out for sure.

The following will provide the confidence intervals for the above metrics as well as confidence intervals for the *difference between them*, and so provides a statistical test at a specified alpha level.

```
#lmgci is the arbitrary object name

lmgci=boot.relimp(Model, boot = 1000)

booteval.relimp(lmgci)


#Partial output

Confidence interval information ( 1000 bootstrap replicates, bty= perc ):

Relative Contributions with confidence intervals:


                              Lower   Upper

             percentage   0.95 0.95    0.95

INITIATI.lmg 0.2139      ABC   0.1279 0.3143

KNOWLEDG.lmg 0.2800      AB_   0.1980 0.3742

SOFTWARE.lmg 0.1219      _BC   0.0624 0.1972


Letters indicate the ranks covered by bootstrap CIs.

(Rank bootstrap confidence intervals always obtained by percentile method)

CAUTION: Bootstrap confidence intervals can be somewhat liberal.

Differences between Relative Contributions:


                              Lower   Upper

                 difference 0.95 0.95    0.95

INITIATI-KNOWLEDG.lmg -0.0662       -0.2335 0.0907

INITIATI-SOFTWARE.lmg  0.0920       -0.0258 0.2088

KNOWLEDG-SOFTWARE.lmg  0.1582      *   0.0440 0.2765

*indicates that CI for difference does not include 0.
```

The output first provides 95% CIs for the metrics themselves, while the second bit provides them for the difference between any two LMG statistics. Standard reporting based on standardized coefficients would have ranked them as knowledge first, initiative second, and software expertise as even being nonsignificant. However here we can see the low end of the software variable suggests a meaningful contribution (about 10% of our $R^2$), and the only difference among their orderings we might feel confident enough to make would be between statistical knowledge and software expertise, the former contributing more than the latter.

## Summary

While it may have seemed quite a bit to pull off, assuming we didn't use any menus the entire code is 15 lines including summary calls and takes only a few seconds to obtain *all* of the output. As a

comparison, for basic regression in SPSS it would have taken many more lines to produce as much as it could of the R output, but it simply would not be able to do most of it. So technically with less work one can perform real tests of assumptions, a robust check on the analysis, validation of the model, an interval estimate of the bias-corrected $R^2$, and obtain tests for a variable importance metric that decomposes $R^2$.

Here I did not do many of the options that were available for each function, and as was noted, this is 'pretty' data which is rare in the social sciences. But even with this minimal approach I was able to pull off an analysis that was just as or more interpretable and more accurate one than a standard one would have been, and one in which I can feel much more confidence regarding the results.

Given modern desktop computing capabilities, the time has long since passed to still be using methods without regard to the developments of the past 30 years. Standard texts have always pointed out the issues, but now even some introductory ones provide solutions, the implementation of which are found in many statistical packages, some of which, like R, are free. Just concluding 'caution must be taken in the interpretation of results' is a poor way of dealing with the problems of data. You could say that for any analysis despite the techniques employed including these. What would be a more accurate statement for those that do not use modern methods to get the sentiment across would be 'these results are completely suspect because care was not taken to perform the analysis well'.

Hopefully those who read this can see how easy it can be to pull off, and begin their journey to doing better statistical analyses.

## Summary of code

```r
#red lines could have been done via R-commander menus, assumes data has already been imported
 and called 'Dataset'

library(Rcmdr) #brings up the menu system

library(lmtest) #diagnostics

library(robustbase) #robust regression

library(Design) #validation, bias assessment

library(MBESS) #CI for the R^2

library(relaimpo) #variable importance


Model <- lm(QUALITY~INITIATI+KNOWLEDG+SOFTWARE, data=Dataset)

summary(Model)

bptest(Model, varformula = ~ fitted.values(Model), studentize=FALSE, data=Dataset)

resettest(Model, power=2:3, type="regressor", data=Dataset)

shapiro.test(Model$residuals)

dwtest(Model,alternative="two.sided",data=Dataset)

vif(Model)

modelrob=lmrob(Model)

summary(modelrob)

valmodel=ols(formula=QUALITY~INITIATI+KNOWLEDG+SOFTWARE, data=Dataset, x=T, y=T)

validate(valmodel, method="boot", B=500)

ci.R2(R2=.597, N=100, K=3, conf.level=.95, Random.Predictors=TRUE)

calc.relimp(Model)
```

```
lmgci=boot.relimp(Model, boot = 1000)

booteval.relimp(lmgci)
```

## Footnotes

[1]   I have at times actually stopped reading research reports that don't even bother to mention anything about the testing of assumptions, since I can never know if the results are what they say they are, much less any theoretical conclusions based on them.

### References

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Harrel, F.E. (2001). *Regression Modeling Strategies.*

Lindeman, R.H., Merenda, P.F. and Gold, R.Z. (1980). *Introduction to Bivariate and Multivariate Analysis.*

Kruskal, W. (1987). Relative importance by averaging over orderings. *The American Statistician,41,* 6-10.

Tukey, J.W. (1979). "Robust techniques for the user". In R.L. Launer & G.N. Wilkinson. (Eds.), *Robustness in Statistics*.

Wilcox, R. (2003). *Applying Contemporary Statistical Techniques.*

Wilcox, R. (2002). *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy.*

Wilcox, R. (1997). *Introduction to Robust Estimation and Hypothesis Testing.*

[Return to top](#)

# Benchmarks Online

Page One

Campus Computing News

Summer Hours

Information Security Awareness

Get it while it lasts! UNT Support of SkillPort Computer-Based Training Ends November 2008

Today's Cartoon

RSS Matters

The Network Connection

Link of the Month

Helpdesk FYI

Short Courses

IRC News

Staff Activities

Subscribe to Benchmarks Online

# Short Courses

**By Claudia Lynch**, *Benchmarks Online* Editor

Short Courses are currently being scheduled for the summer. We hope to have them up "any day now." Meanwhile, surf over to the Short Courses page to see what sort of classes will likely be offered starting, probably, late June/early July. If you have a group that needs a specific class, it may be possible to arrange a special class just for them . See "Customized Short Courses" below for further information.

## Customized Short Courses

Faculty members can request customized short courses  geared to their class needs from ACS. Other groups can request special courses also. Contact ACS for more information (ISB 119, 565-4068, lynch@unt.edu).

## Especially for Faculty and Staff Members

In addition to the ACS Short Courses, which are available to students, faculty and staff, staff and faculty members can take courses offered through the Human Resources Department, and the Center for Distributed Learning. Additionally, the Center for Achievement and Lifelong Learning offers a variety of courses, usually for a small fee.

EIS training is available. Questions or comments relating to EIS training should be sent to EIStrn@unt.edu.

### Moving from GroupWise to Microsoft Outlook Training

The GroupWise to Microsoft Exchange Migration is almost complete. If you still need it, a useful place for finding resources to prepare for this transition can be found here.

### Central Web Support

Consult Central Web Support for assistance in acquiring "Internet services and support." As described on their website:

- Services include allocating and assisting departments, campus organizations and faculty with web space and associated applications. Additionally, CWS assists web developers with databases and associated web applications, troubleshooting problems, support and service.

- . . .

  our staff will continually add additional courses to fit the needs of our faculty, staff and students.

## Center for Distributed Learning

The Center for Distributed Learning offers courses especially for Faculty Members. A list of topics and further information can be found [here](#).

The center also offers a "Brown Bag" series which meets for lunch the first Thursday of each month at Noon in Chilton 245. The purpose of this group is to bring faculty members together to share their experiences with distributed learning. One demonstration will be made at each meeting by a faculty member with experience in distributed learning. More information on these activities can be found at the [Center for Distributed Learning](#) Website.

## UNT Mini-Courses

There are a variety of courses offered, for a fee, to UNT faculty, staff and students as well as the general public. For additional information surf over to [http://www.unt.edu/minicourses/](http://www.unt.edu/minicourses/)

# Information Security Awareness

- Wednesday, July 23  2-3:30 p.m. ESSC Room 152

The UNT Information Security team is currently offering Information Security Awareness courses to all UNT faculty and staff. Topics to be covered will include workstation security, sensitive data handling, copyright infringement issues, identity theft, email security, and more. For more information, or if you would like to request a customized course to be taught for your department, contact Gabe Marshall at x4062, or at [security@unt.edu](mailto:security@unt.edu).

Also, Information Security Training  is [now available](#) through WebCT Vista.

# Alternate Forms of Training

Many of the General Access Labs around campus have tutorials installed on their computers. See [http://www.gal.unt.edu/](http://www.gal.unt.edu/) for a list of labs and their locations. The Willis Library, for example, has a [list of Tutorials and Software Support](#).

The [Training Website](#) has all sorts of information about alternate forms of training. Computer Based Training (CBT) and Web-based training are some of the alternatives offered.

For further information on CBT at UNT, see the CBT [website](#). Note, also, some *Benchmarks Online* articles that have been published in the recent past:

- [Computer-Based Training for Microsoft Project 2007 is Now Available](#)

- [More Windows Vista and Microsoft 2007 Server Courseware Added](#)

- [SkillPort Training Site Update](#)

- [Free and Legal: Copyright Advice and Training Online](#)

- [The Gift that Keeps on Giving: Even More Outlook and Office 2007 Training Posted on the CBT Website](#)

- [New Titles Added in SharePoint Server 2007, Publisher, Project, and Office 2007](#)

## State of Texas Department of Information Resources

Another possible source of training for staff and, perhaps, faculty members is the Texas Department of Information Resources. A look at their Education and Training [website](#) reveals some interesting possibilities. For example, under [Conferences, Briefings, and Events](#) is a "Microsoft Training Series" described as "free training classes ... delivered by Microsoft and hosted by DIR as part of the Technology Today Series (TTS)."

[Return to top](#)

# Benchmarks Online

# Staff Activities

## Transitions

### New Employees:

- **Andrew Babb**, Distributed Computing and Imaging Services Student Assistant (part-time).

- **Nathaniel McGraw**, Computer Systems Manager, EIS Technical Services (AIS).

- **Matthew D. Bishop,** Computer Equipment Operator, Computer Operations (part-time).

- **Anna West**, Student Assistant, Administration and Planning (part-time).

### No longer working in the Computing and Information Technology Center:

- **Pravi Ponnaiah,** Student Assistant**,** Data Communications (part-time).

- **Emily Schlebach**, Student Assistant, Planning and Administration (part-time).

- **Jennifer Moore**, Student Assistant, Administration and Planning (part-time).

- **J.P. Williams**, Computer Systems Manager, Central Web Support, Enterprise Systems Technical Services.

- **Graham Pocta**, Distributed Computing and Imaging Services Student Computer Technician (part-time).

## Changes, Awards, Recognition, Publications, etc.

- **Jesse White**, Programmer/Analyst, Student Records Data Systems went to the SPEEDE users group conference in Austin in May, along with Bobby Lothringer (Assistant Registrar), to learn about the latest changes and updates to the SPEEDE software, which is the electronic transcript software used by all Texas public institutions to send transcripts to other Texas public institutions.

- **Kevin Cox**, Data Warehouse & Reporting Infrastructure Team Manager, participated in Cognos Forum in Las Vegas, May 12 – 15; one was a panel discussion, and the other was a focused lunch group led discussion for higher education and public sector.

Abstracts for the sessions are as follows:

> **Session 1503 – IBM Cognos 8 Planning Customer Panel (University of North Texas, Constar International, Quiznos)** – Join Cognos customers as they discuss their deployment of IBM Cognos Planning within their organizations. Share in the best practices these leading companies have utilized to help them increase efficiency and effectiveness to transform their planning processes and increase the speed of decision making.

> **Networking Lunch : Public Sector** – Share best practices for implementing an enterprise-wide performance management system, and hear how the University of North Texas successfully deployed performance management in a higher education enterprise environment.

## Service to UNT

The following people were recognized for their service to UNT in *InHouse* recently:

**10 years of service**

- **Roy Allen Akers, Jr.**, Programmer Analyst, EIS Technical Services (AIS).

## New Baby

Last month they had a baby shower, and right after that they had a baby! Meet Micah Eugene Franklin, David and Danja Franklin's baby boy, born on May 22. Danja is a Production Control Specialist in PeopleSoft Application Infrastructure Management (AIS) and David is a Computer Equipment Operator in Computer Operations (AIS).

**Stay informed!**

**Faculty/Staff Announcements**

# announce.unt.edu

Staff Activities