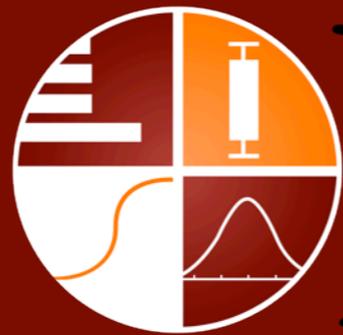




Food and Agricultural Education
Information System

+



LISA

FAEIS Statistical Initiatives and Discussion

FAEIS Peer Panel Meeting
Dallas, Texas
November 16, 2010

Eric Vance
ervance@vt.edu
Virginia Tech

From 2002 to 2009 the number of entries in FAEIS doubled.

Four times as many institutions now submit data to FAEIS.

With this increase in the quantity of data and the number of institutions reporting, new challenges have arisen:

- More data mean more outliers
- More institutions and programs mean more changes and transitions within the programs
- New institutions added mean more “gaps” in the data for previous years

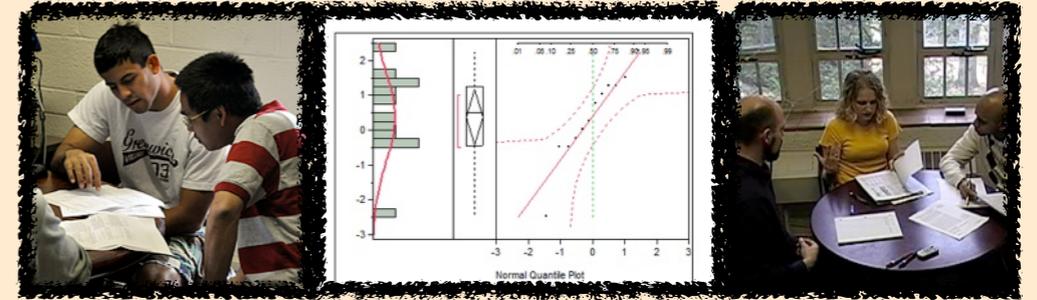
What other issues affect the use of FAEIS data?

In an effort to improve the quality of the FAEIS data and prepare the data for rigorous statistical analysis, the FAEIS team asked LISA to collaborate and provide leadership on statistical issues from an independent, outside perspective.



What/Who is LISA?

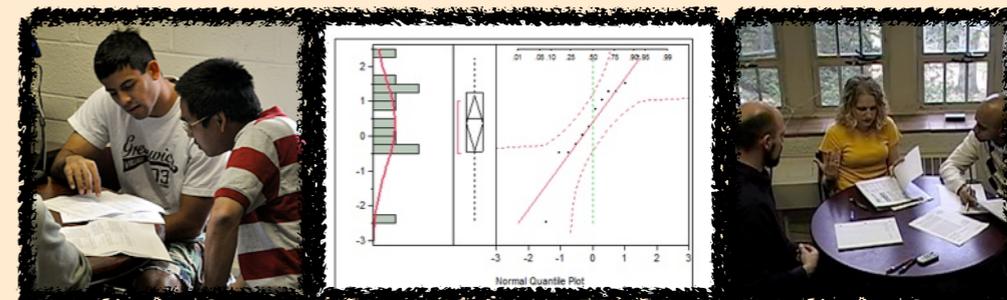
LISA helps VT researchers benefit
from the use of **Statistics**



Experimental Design • Data Analysis • Interpreting Results
Grant Proposals • Software (R, SAS, JMP, SPSS...)

Our goal is to make statistics a strength of
research at Virginia Tech, not a roadblock.

LISA helps VT researchers benefit
from the use of **Statistics**

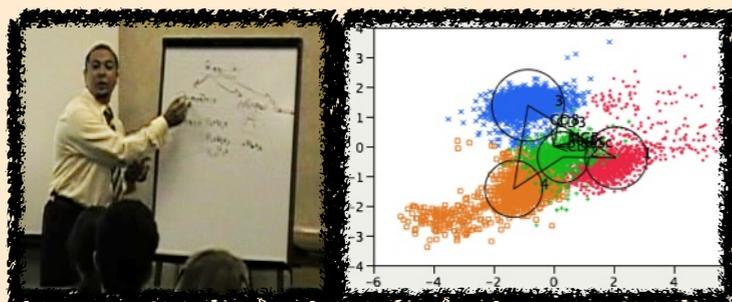


Experimental Design • Data Analysis • Interpreting Results
Grant Proposals • Software (R, SAS, JMP, SPSS...)

Collaboration

From our website request a meeting
for personalized statistical advice

Great advice right now:
*Meet with LISA **before** collecting your data*



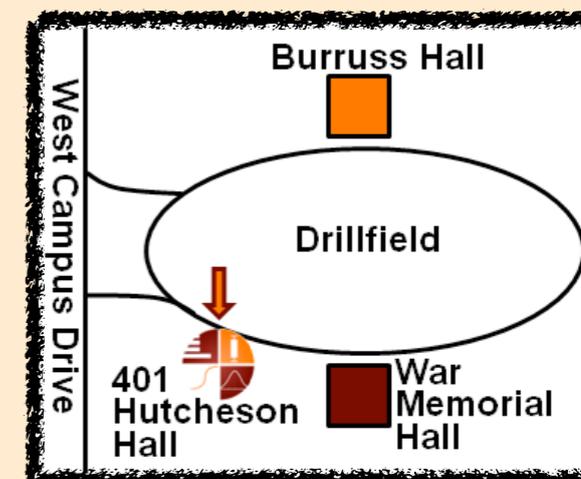
Short Courses

Designed to help graduate students
apply statistics in their research

Walk-In Consulting

Monday—Friday* 12-2PM for
questions requiring <30 mins

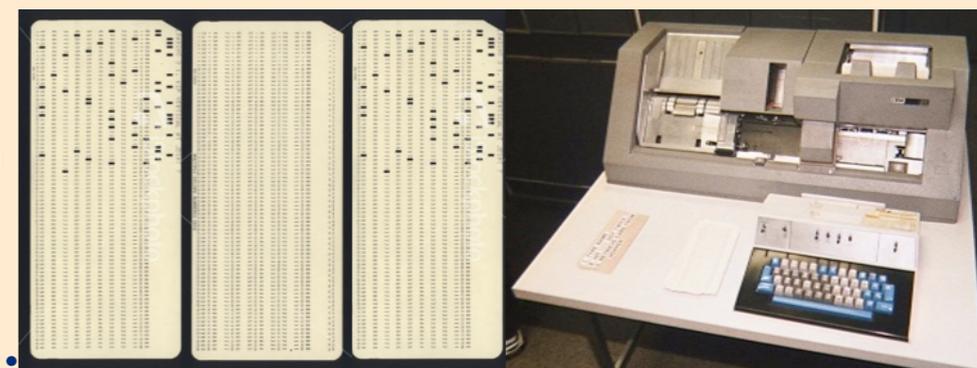
**Mon—Thurs during the summer*



1948: The Statistical Laboratory was founded as a division of the Virginia Agricultural Experiment Station to serve the needs of agricultural and biological research in Virginia and VPI.



1973: The Statistical Laboratory was re-formed as the Statistical Consulting Center to assist with statistical analyses in every college of Virginia Polytechnic Institute & State University (VPI&SU).

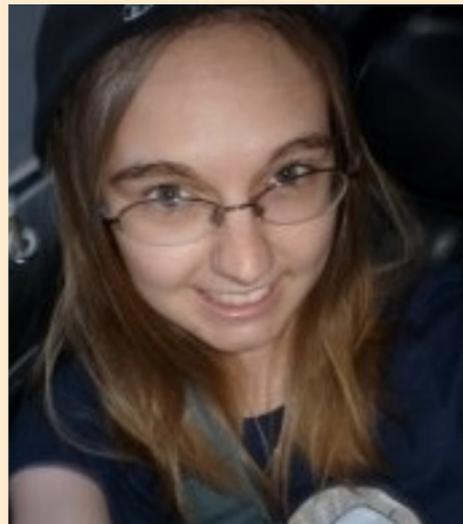


2008: The Statistical Consulting Center was re-organized as the Laboratory for Interdisciplinary Statistical Analysis (LISA) to collaborate with researchers across the Virginia Tech.





Eric Vance



Tonya Pruitt



Chris Franck



3 Lead Collaborators (20 hours/week)

~15 MS and PhD Associate Collaborators (5-10 hours/week)

LISA
collaborators
meet weekly to
discuss projects
such as FAEIS and
to learn from
each other.



LISA supports the FAEIS team. We deal with their statistical issues.

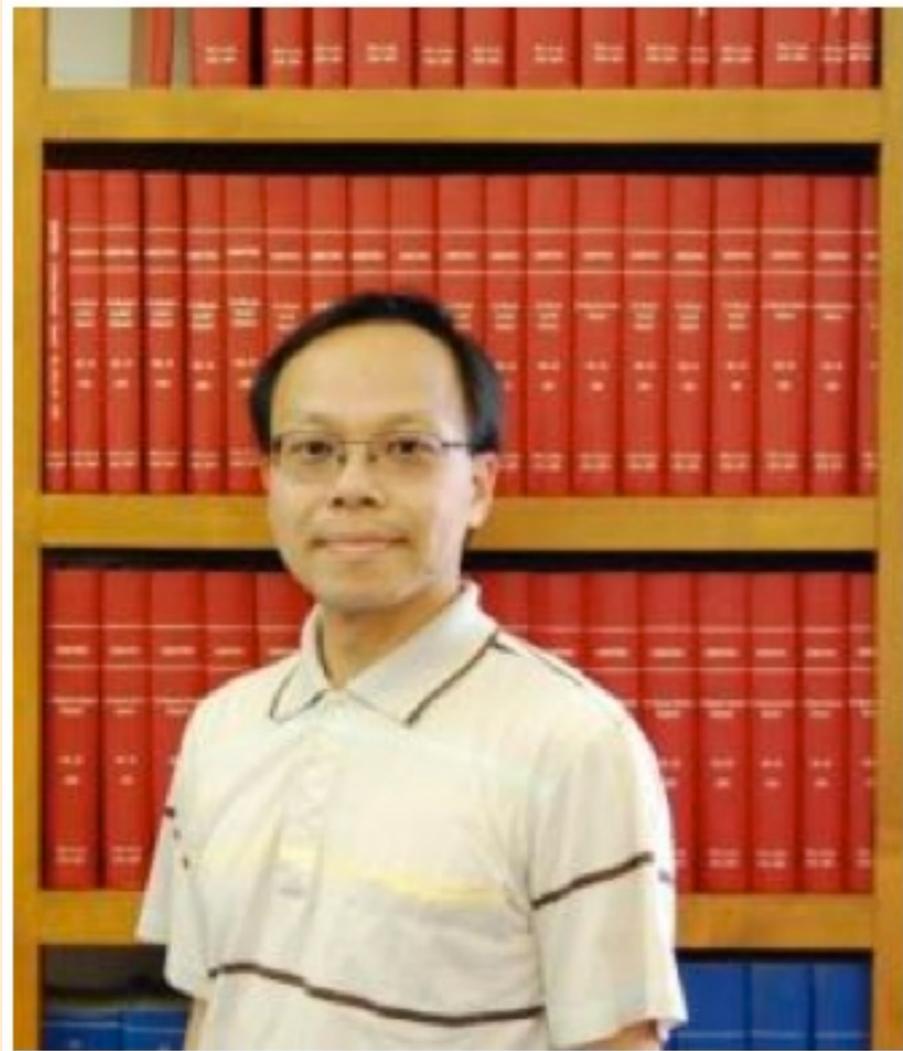
Albert Shen: 20 hours/wk

Eric Smith: ~2 hours/wk

Eric Vance: ~4 hours/wk

Additional LISA
collaborators as needed





Albert Shen, a statistics graduate student and LISA associate collaborator, was hired by FAEIS to:

- Create a SAS dataset from the FAEIS database currently in Oracle
- Develop algorithms and procedures for detecting outliers
- Deal with other statistical issues such as missing values and “gaps” in the data

Creating a FAEIS dataset in SAS will provide many benefits:

- Flexibility to analyze the data using SAS (or any other statistical programs R, JMP, SPSS, ...)
- Visualizing the data
- Detecting outliers
- Analyzing trends in the data, such as graduation rates
- Consistency of analyses
- Portability of reports

Since being hired as a statistical analyst GRA:

- Created SAS datasets
- Verified that SAS reports are identical to Report Builder
- Developed algorithms to identify outliers
- Developed algorithms to identify zeros and missing data



Additional things that could be done using statistics:

- SAS algorithm to identify redundant/repeated data entries and misplaced CIP codes
- Automated identification of invalid/problematic data
- Comparing IPEDs data with FAEIS data
- Your ideas?

Timeline for deliverables:

- ✓ Creation of SAS dataset and report verification 10/2010
- ✓ Identification of outliers and missing data 11/2010
- Identify redundant entries and misplaced CIPs 12/2010
- Statistical update quarterly report 12/2010
- Automated identification of problematic data 01/2011
- Statistical Expert Panel meeting in D.C. 03/2011

The FAEIS team and LISA are working together to incorporate statistical thinking into FAEIS processes, thereby adding long-term value to FAEIS.



Discussion:

What are your concerns with FAEIS data
or creative ideas for their future use?



Photo Credit: Betty Higginbotham