



LISA

Statistical Enhancements to FAEIS in 2010-2011

NIFA and Expert Panel Meetings
Washington, D.C.
April 7-8, 2011

Eric Vance
ervance@vt.edu
Virginia Tech



Laboratory for Interdisciplinary Statistical Analysis

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

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.



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

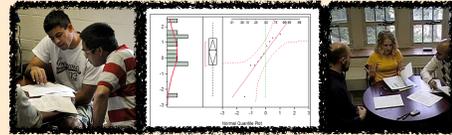


What/Who is LISA?



Laboratory for Interdisciplinary Statistical Analysis

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



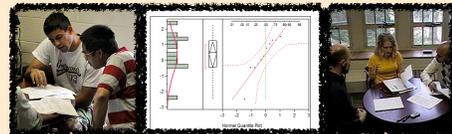
Experimental Design • Data Analysis • Interpreting Results
Grant Proposals • Software (R, SAS, JMP, SPSS...)

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



Laboratory for Interdisciplinary Statistical Analysis

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



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

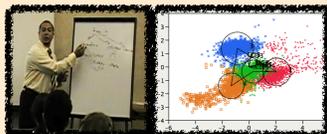
Collaboration

From our website request a meeting for personalized statistical advice

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

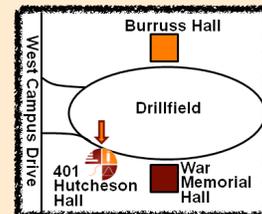
Walk-In Consulting

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



Short Courses

Designed to help graduate students apply statistics in their research





Laboratory for Interdisciplinary Statistical Analysis

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.



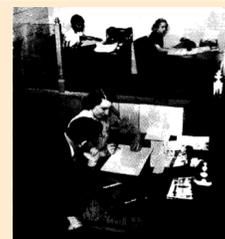
1973: The Statistical Laboratory was re-formed as the Statistical Consulting Center to assist with statistical analyses in every college of Virginia Polytechnic Institute & State University (VPI&SU).

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

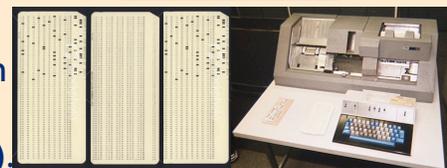


Laboratory for Interdisciplinary Statistical Analysis

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.



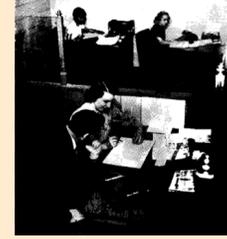
1973: The Statistical Laboratory was re-formed as the Statistical Consulting Center to assist with statistical analyses in every college of Virginia Polytechnic Institute & State University (VPI&SU).



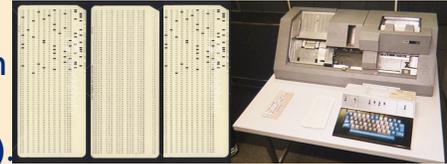
2008: The Statistical Consulting Center was re-organized as the Laboratory for Interdisciplinary Statistical Analysis (LISA) to collaborate with researchers across the Virginia Tech.



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
 Katie Griffin: 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



Katie Griffin, a statistics graduate student, was hired by FAEIS to:

- Support the FAEIS Help Desk with statistical issues to improve data accuracy
- Create reports to compare FAEIS data to IPEDS data
- Mine institutions’ data from their Institutional Research websites

Since being hired as a statistical analyst GRA, Albert Shen:

- 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

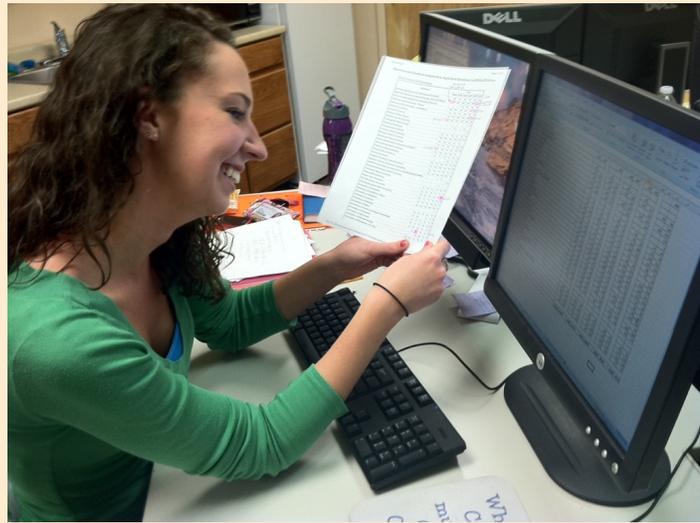


Creating a FAEIS dataset in SAS will provide many benefits:

- 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 GRA, Katie Griffin has:

- Compared FAEIS data to IPEDS data for a range of programs and institutions
- Assisted institutions with data collection and reporting
- Filled in missing FAEIS data with data from IR



Additional enhancements to FAEIS data using statistics:

- SAS algorithm to identify redundant/repeated data entries and misplaced CIP codes
- Automated identification of invalid/problematic data

Discussion:
Your ideas for statistical enhancements to FAEIS?

