Software Skills for Researchers

- Software Applications are an integral part of any type of research.
- Today's researcher needs a wide range of software use and management skills:
  - Software evaluation and selection
  - Hardware device evaluation and selection
  - Advanced user skills in general use software
  - Advanced user skills in selected specialized software packages
- Also needs Data Management Skills:
  - File Management
  - Data/Date Conversion
  - Data Management
  - Data Archiving

DataOne Position Description

**Climate Scientist**

**Required Qualifications:** Candidates should have . . . expertise in advanced interactive visualization techniques . . . The candidate must have knowledge and practical experience in developing and using visualization software such as VisTrails or UV CDAT or other advanced Visualization packages. . . Experience with acquiring and managing spatial data . . . The candidate must have experience in using Python, PERL, or other languages for managing high-volume complex data. The candidate should have familiarity with UNIX and Windows operating systems.

US Government Position Description

**Social Scientist**

For GS-11 Grade Level:
A. One year of specialized experience equivalent to the GS-9 grade level in Federal service, providing analytic support for policy analysis and social science research related to evaluating legislative, regulatory and/or the delivery of public health or human services programs; performing quantitative analyses related to health and human services programs and policy utilizing statistical software such as STATA, SPSS, SAS or equivalent statistical software.

Research Support Applications

- **Project Management Software**
  - Basecamp, MS Project, others
  - Wikis, Blogs
  - Lab Notebook Software

- **Workflow Software**
  - General Workflow Management (Visio, Dia)
  - Experimental/Scientific Workflow (MyExperiment, Kepler)

- **Administrative Management**
  - Office Applications, Email, Scheduling
  - MS Office, OpenOffice, LibreOffice

Data Analysis Software Categories

- **Qualitative Analysis**
  - Enables collection and interpretation of behavioral data - Atlas.ti

- **Quantitative/Statistical Analysis**
  - Provides tools that allow the relationships between data elements to be expressed in mathematical terms - SAS/SPSS

- **Content Analysis**
  - Provides search, comparison and analysis tools for large collections of text-based documents - LightSide

- **Data Acquisition Software**
  - Often paired with hardware devices, captures data from sensors, experimental devices - LabView

- **Geographic/Mapping (GIS)**
  - Provides geospatial context to data - ArcGIS
Data Analysis Software Categories

- Mathematical
  - Performs advanced, abstract mathematical functions - MATLAB
- Modeling & Simulation
  - Related to visualization, adds time and space factors to data analysis
- Analysis Programming Languages
  - Enable programmers and skilled researchers to write customized analysis functions and scenarios – R, Fortran
- Visualization
  - Transforms data into visually identifiable scales and relationships
- Specialty
  - Performs functions that are unique to a field of study or analysis
- Others...

Selecting and Purchasing Data Analysis Software

- Does your data analysis software meet your needs for all stages of the data lifecycle?

Selecting and Purchasing Data Analysis Software

- Use the right tool for the right job
- Core and specialty applications are expensive
- Most applications are hybrid – serving multiple purposes
- Look for open source
- Look for free/web-based tools
- UNM Licensed software is available
- Student/Educator pricing

Qualitative Software Features

- Codebook management
- Point-and-click coding
- Auto Coding
- Margin notes
- Weighting values
- Content linking
- Multimedia Analysis
- Transcription tools (for AV data)
- Survey import/management
- Reporting and summarization

Qualitative Analysis Software

- NVIVO [http://www.qsrinternational.com]
- QDA Miner [http://provalisresearch.com]
- Content Analysis
Atlas.ti

- **Summary**: “The purpose of ATLAS.ti is to help researchers uncover and systematically analyze complex phenomena hidden in text and multimedia data. The program provides tools that let the user locate, code, and annotate findings in primary data material, to weight and evaluate their importance, and to visualize complex relations between them.”

- **Used in**: Anthropology, Sociology, Psychology, Business, Marketing, Computer Use Studies

- **Product Information**: [http://www.atlasti.com](http://www.atlasti.com)

- **Product Pricing**: $99 student rate with ID

- **Special Function Add-ons**:
  - Geospatial Data Plotting
  - Online Survey Management
  - Data Visualization

- **Platform Availability**:
  - Windows
  - Mac OSX

---

**Quantitative Analysis Software**

- **SAS** ([http://www.sas.com](http://www.sas.com))
- **STATA** ([http://www.stata.com](http://www.stata.com))
- Microsoft Excel
- Many others

---

**Quantitative Software Features**

- Analysis of variance
- Regression
- Categorical data analysis
- Multivariate analysis
- Survival analysis
- Psychometric analysis
- Cluster analysis
- Nonparametric analysis
- Survey data analysis
- Compare data against common distributions
- Imputation for missing values
SPSS (UNM Licensed)

- **Summary:** "With SPSS predictive analytics software, you can predict with confidence what will happen next so that you can make smarter decisions, solve problems and improve outcomes."
- **Used In:** Business, Anthropology, Sociology, Psychology, Business, Marketing, Computer Use Studies
- **UNM Product Pricing:** $79
  - [http://it.unm.edu/software/faculty-staff/windows/index.html](http://it.unm.edu/software/faculty-staff/windows/index.html)
- **Add-On Features**
  - Collaboration
  - Excel Interface
  - Data Collection
- **Platform Availability:**
  - Windows
  - Mac OSX

SAS (UNM Licensed)

- **Summary:** "From traditional analysis of variance and predictive modeling to exact methods and statistical visualization techniques, SAS/STAT software provides tools for both specialized and enterprise-wide analytical needs."
- **Used In:** Business, Economics, Finance, Natural/Physical Sciences
- **UNM Product Pricing:** $120-170 yearly
  - [http://it.unm.edu/software/faculty-staff/windows/index.html](http://it.unm.edu/software/faculty-staff/windows/index.html)
- **Add-On Features**
  - Scripting Language
  - Data Visualization
  - Advanced Analytics Module
  - Mapping/GIS
- **Platform Availability:**
  - Windows
  - Mac OSX

**DATA VISUALIZATION SOFTWARE**

- "I use SAS to analyze prevalence data to determine the burden of diabetes in the state of Texas. I also use SAS to analyze data to report to CTC for grant requirements. I also use SAS to help my customers (the public)."
Visualization Feature Sets

- Mapping data sets (down to level of US states and counties).
- Broad range of charts and plots:
  - Scatter, line, area, bubble, multiple axis, overlay.
  - Bar, pie, donut, star, block.
- Customized colors, line styles, symbols.
- 2-D and 3-D plots with tilting and rotation.
- Generate static or dynamic interactive (Java or ActiveX) charts and graphs with drill-down capabilities.
- Link graphs to Web pages.
- Embed interactive graphics in Web pages or Microsoft documents.
- Support for common types of printers and plotters.

Data Visualization

- Microsoft Excel (!)
- MATLAB (UNM Licensed)
- Tableau Desktop
- TrendAnalyzer
- VisTrails
- Visual.ly (http://visual.ly)
- Many Specialized Applications
  - Climate Visualization
    - UV-CDAT
    - NCAR

MS Excel

- Summary: Excel provides a unified container for collecting, storing and visualizing any form of data.
- Used by: Engineering, Natural/Physical Sciences, Social Sciences.
- UNM Licensed: http://it.unm.edu/download/
- Special Function Add-ons:
  - DataUp
- Platform Availability:
  - Windows
  - Mac OSX

MATLAB (UNM Licensed)

- Summary: "MATLAB is a high-level language and interactive environment for numerical computation, visualization, and programming"
- Used in: Engineering, Mathematics, Natural/Physical Sciences, Statistics.
- UNM License: http://it.unm.edu/software/matlab
- Special Function Add-ons:
  - Data Acquisition
  - Database connectivity
  - Signal Processing
  - Image Processing
- Platform Availability:
  - Windows
  - Mac OSX
  - Linux
  - Mobile

Features
- Mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, numerical integration, and solving ordinary differential equations.
- Functions for integrating MATLAB-based algorithms with external applications and languages such as C, Java, .NET, and Microsoft Excel.
- Development tools for improving code quality and maintainability and maximizing performance.
- Tools for building applications with custom graphical interfaces.
Tableau Desktop

- **Summary**: “Tableau Desktop is based on breakthrough technology from Stanford University that lets you drag & drop to analyze data. You can connect to data in a few clicks, then visualize and create interactive dashboards with a few more. Shift fluidly between views, following your natural train of thought.”
- **Used in**: Business, Economics, Finance, Social Sciences
- **Free Trial Available**
- **Platform Availability**:
  - Windows
  - Mac OS X
  - Mobile

DATA ANALYSIS PROGRAMMING LANGUAGES

R

- Programming Language for statistical data analysis and Graphics
- Extremely popular for Quantitative Data Visualization
- Programming Tools are free, open source
- R Project website: [http://www.r-project.org/](http://www.r-project.org/)
- R Website includes Tutorials, Manuals, Training
- Available on Windows, Mac, Unix

Programming Languages

- C/C++
- FORTRAN (still around)
- Python (open source)
- R (open source)
- S
- Embedded Languages
  - SAS
  - MATLAB

R Example Code
Poster developed from R