**David Michael Pressley**

Technologist in the Life Sciences

#### [david@clearcreekanalytics.com](mailto:david@clearcreekanalytics.com)

I am a lifelong learner with 20 years of experience in the Life Sciences spanning medical research, drug discovery and drug development. I have provided competent leadership and individual contribution for biostatistical workflows in the drug development value stream. My expertise is in configuring and enabling analytical systems and ETL workflows and artifacts (primarily using SAS) from protocol development to regulatory post-submission. Engaging users, adapting user requirements, and promoting novel solutions is a soft skill I continually work to improve. My ideal future state is to work with like minded professionals in a high trust, collaborative environment using modern technologies and methods. I am bootstrapping my own analysis platform, utilizing modern alternatives which allow computers to do what they do best, and developers to do what they do best.

EXPERIENCE

Clear Creek Analytics, LLC— ***Owner***

April 2017 - Jan 2018; June 2018 - Current

##### Contractor and Consultant to the pharmaceutical industry

Roivant Sciences Ltd— ***Digital Innovator***

January 2018 - July 2018

Effectively served as a systems architect/product manager for the configuration and enablement of systems of record within the Biostatistics, Clinical Pharmacology, and Nonclinical functional areas. Provided cross functional leadership, subject matter expertise, and individual contribution . Systems, applications and responsibilities include:

* Manual provisioning of infrastructure and applications on a Red Hat Linux 7 terminal server (VirtualBox and AWS configured with ssh and GUI access)
* Windows Server 2016 (jump server and Windows specific application server)
* Creation of bash scripts to:
  + Configure the user’s local working area for project based biostatistical workflows
  + Generate ssh keys for secure access to git and GitHub repositories over ssh
  + Configure Git LFS for efficiently versioning binary files such as SAS datasets
  + Set up synchronization (rsync) and scheduling (cron) of data from vendor SFTP site to user’s local project area
* Utilization of git, GitHub, GitLab, and Git Large file storage to ensure efficiency, traceability of code and data of the biostatistical programming workflow
* Effectively communicated with vendors, colleagues, and management through the use of applications such as Confluence ([HowTo](https://clearcreek.atlassian.net/wiki/spaces/~363221148/pages/596279300/Initializing+a+Project+in+the+Biostatistical+Programming+Workstream?atlOrigin=eyJpIjoiNDA1MDlmYTM0NDMyNDk3N2IwZGYzMmI3YmZiZWEwNzciLCJwIjoiYyJ9), [troubleshooting](https://clearcreek.atlassian.net/wiki/spaces/~363221148/pages/596279300/Initializing+a+Project+in+the+Biostatistical+Programming+Workstream?atlOrigin=eyJpIjoiNDA1MDlmYTM0NDMyNDk3N2IwZGYzMmI3YmZiZWEwNzciLCJwIjoiYyJ9), requirements, knowledge base), Asana (team and project oriented tasks and kanban workflow visualization), Slack (a better alternative to email), and Zoom (video conferencing)
* Negotiated with internal Quality group and external Quality vendors to challenge de facto waterfall approaches to Quality Management
* Signatory and contribution to Quality documentation surrounding the Red Hat based SAS system
* Collected user requirements from functional groups, engaged and directed third party vendors for implementation of infrastructure, and application configuration and enablement
* Contributed to development of a business case for biostatistical and clinical pharmacology functional areas to ascertain the value of each functional group’s application stack as a shared services model
* Led efforts to inventory and roadmap all systems in the clinical development value stream

United Therapeutics Corporation— ***Associate Director, Statistical Programming and Analytical Systems Architect***

March 2006 - March 2017 (Started with title Clinical Data Programmer)

Ensured the needs of internal customers were met through accurate representation of data across the clinical development lifecycle. I led the design and implementation of our analytical environment which utilized virtualization and open source technologies such as Git, Subversion, Python and R.

PROJECTS OF SIGNIFICANCE

***2015: Remodulin Implantable System Supplemental New Drug Application***

**2012 - 2014: *Unituxin Biological Licensing Application and Approval***

**2006 - 2012: *Orenitram New Drug Application and Approval***

**2007 - 2009: *Tyvaso New Drug Application and Approval***

SKILL LEVEL: AREAS OF INTEREST : [EXAMPLE CODE]

**Advanced**:

* SAS Foundation/Graph/Stat/Macro/etc modules : [[readSpec.sas](https://gist.github.com/pressleydavid/5367151583f9a0b29c2c0edb47d7b542), [hashtable.sas](https://gist.github.com/pressleydavid/2fb3e93467dfd652a2507949afd5766b), [addLabels.sas](https://gist.github.com/pressleydavid/4446ee586068bca0e5f5), [createXPT.sas](https://gist.github.com/pressleydavid/f80ca5146a2fb9419127), [unzipXPTLibrary.sas](https://gist.github.com/pressleydavid/1c570253c086f32ee788), [histogramShifts.sas](https://gist.github.com/pressleydavid/eed1aa17b428a210620e), [hemodynamicsGraphs.sas](https://gist.github.com/pressleydavid/ae7ca2cbc20f9f660964), [init.sas](https://gist.github.com/pressleydavid/22bd3ce58ef96d95404b0be4ae08a6eb), [getProjectPath.sas](https://gist.github.com/pressleydavid/1e59ac4936fb5841e1cfa061d39f4ea2), [getdatasetnames.sas](https://gist.github.com/pressleydavid/d5169a590657d1e379736db4d03e77f5)]
* Version Control Systems: Git (fan of [Drieseen](https://nvie.com/posts/a-successful-git-branching-model/) branching model, and GitHub flow model), Subversion

**Intermediate:**

* + Python : [[readFile.py](https://gist.github.com/pressleydavid/939e6871189c6a13c7ca), [readXPT.py](https://gist.github.com/pressleydavid/733eba5b5094d58c4a4e), [rm\_bookMarks.py](https://gist.github.com/pressleydavid/fa8c6437ea9818361d9d), [createFolderTree.py](https://gist.github.com/pressleydavid/98c9d46d2e983aae6969), [sas7bdat.py](https://gist.github.com/pressleydavid/15f5d55d8e576f035bb3)]
  + AWS: familiar with creating instances from AMIs/GUI, accessing instances via ssh and VNC, creating and accessing S3 buckets
  + Bash : [[studysetup.sh](https://gist.github.com/pressleydavid/e0f706ed3bab8c4a8fce3c07a47a7b6f)]
  + Vagrant:
    - RedHat 7 with manual duplication of analytsis platform used in production (SAS, Python, git, GitLab/GitHub, UltraEdit, MobaXterm, TigerVNC server for enabling desktop connections).
    - Ubuntu 16.04 with Neo4j, Apache web server, virtualenv, django, lxml, mod\_wsgi, and openpyxl to surface a running Neo4j instance on localhost

**Beginner/Conceptual Familiarity:**

* + Terraform, Vault, Packer
  + Ansible, Kubernetes, OpenShift (networking is a weakness)
  + C++, R, Apache Spark, JavaScript, Velocity, XML/HTML/CSS
  + Neo4j, Mongo, Redis

PREVIOUS WORK HISTORY

**Amphora Discovery** — ***Programmer/Analyst*** September 2001 - March 2006

**Glaxo Smith Kline** — ***Research Associate (contractor)*** August 2001 - January 2002

**Nortel Networks** — ***Equipment Applications Engineer*** October 2000 - August 2001

**Duke University Medical Center** — ***Research Technician*** December 1998 - October 2000

EDUCATION

## Coursera

29 November 2016

# [HTML, CSS, and Javascript for Web Developers](https://www.coursera.org/account/accomplishments/certificate/LVPFSDYRM8V3)

## EdX

10 July 2015

# [CS100.1x: Introduction to Big Data with Apache Spark](https://s3.amazonaws.com/verify.edx.org/downloads/446c9a2aba924794b17e3e4f3263fd5e/Certificate.pdf)

6 August 2015

# [CS190.1x: Scalable Machine Learning](https://s3.amazonaws.com/verify.edx.org/downloads/0be48224c88540999ed6344243b6a847/Certificate.pdf)

4 September 2015

# [DAT204x: Introduction to R Programming](https://s3.amazonaws.com/verify.edx.org/downloads/eb014507243c4eca90c52e63c63bc911/Certificate.pdf)

6 November 2015

# [6.00.1x: Introduction to Computer Science and Programming Using Python](https://courses.edx.org/certificates/8c77c1a8a4df4d5abed5fe72c0547fef)

**North Carolina State University** — ***Graduate Coursework in Statistics***

Fall Semester 2008 - Spring Semester 2010

# Fall 2008: ST 511 - Experimental Statistics For Biological Sciences

# Spring 2009:ST 512 Experimental Statistics For Biological Sciences

# Spring 2010: ST505 - Applied Nonparametric Statistics

**North Carolina State University** — ***Computer Programming Certificate***

Fall Semester 2004 - Spring Semester 2007

**North Carolina State University** — ***Bachelor of Science in Biochemistry / Bachelor of Arts in Chemistry***

Fall Semester 1993 - Summer Term I 1998