Joshua**Grant**

Data Engineer

👤 contact

12918 Sanderling Lane
Knoxville, TN 37922
United States of America
+1 (865) 803 3495 📱
jngrant@live.com ✉
notjustadatum.blogspot.com **B**
github.com/sempervent 🎮
linkedin.com/in/joshuanagrant **in**

◎ objective

to take theories, ideas, and experiments and turn them into production-scalable solutions

</> development

Bash:	■■■■□	
Docker:	■■■■□	
Kubernetes:	■■□□□	
HTML:	■■■□□	
JavaScript:	■■□□□	
LaTeX:	■■■□□	
Perl:	■□□□□	
PHP:	■■□□□	
Python:	■■■■□	
R:	■■■■□	
SQL:	■■■■□	
git:	■■■■□	

🎴 flavors

R:	data.table	shiny
	tidyverse	vegan
	testthat	devtools
	DBI	leaflet
	httr	ggplot2
Python:	pandas	bs4
	PIL	xdgboost
	osgeo	airflow
SQL/DB:	MySQL	PostgreSQL
	MongoDB	Hadoop
	SQLite	Snowflake
LaTeX:	tikz	fancyhdr
	tables	booktabs
JavaScript:	node.js	angular.js
	vue.js	jQuery

📊 Statistics

ANOVA

Experience

Data Engineer
Oak Ridge National Laboratory, Oak Ridge, TN *2018—Present*

- Serves as support staff for the Python project EAGLE-I, where I implemented an ETL workflow from a brittle-based cron system to a more robjust Airflow system deployed in a Kubernetes cluster
- Developed full stack docker images using Python, R, Airflow, and Shiny to automate the acquisition of data for the WSTAMP project for storage in an Hadoop Data Warehouse
- Oversaw the development team for DOE's COVID-19 tracking project and downstream model creation and analysis, for which I received distinguished recognition for the project
- Developed containerized ETL workflows and REST APIs using Swagger for data exchange in the NAERM-RTSA project
- Developed web-scraping and containerized model creation for misinformation classification from a variety of online web sources including social media sites for implementation in advanced statistical models and network analysis
- Serves as Software Quality Assurance board chair for the Geospatial Science and Human Security Division

Data Scientist II Subcontractor
Oak Ridge National Laboratory, Oak Ridge, TN *2018 —2019*

- Automating data ingestion of the WSTAMP webapp via the use of R, Python, Airflow, and Docker
- Authoring an R package to unify geography between online data APIs

Owner & Contractor-For-Hire
Specrabella, Knoxville, TN *2018 —2018*

- Designing, developing, and deploying an interactive energy informatics web application that provides real-time calculation from machine learning derived models
- Consulting with clients on ground-up web application architecture, database and visualization needs
- Developed data ingestion methods using cURL, PHP, and R to scrape energy data from the web
- Streamlined training and development procedures using ansible, vagrant, PHP, and BASH

Bioinformatician, Project Manager, & Technical Sales Representative
Microbial Insights, Knoxville, TN *2016 —2018*

- Established an automated pipeline using R, MySQL, Python, and Bash for NGS analysis, along with intranet website control written in JavaScript and Shiny
- Automated client-bound statistical calculations such as linear models, ANOVA, SOMs, clustering, PCA, and PCoA
- Developed an ETL customer data visualization tool using R, PHP, JQuery, and MySQL to view qPCR results in the context of other samples' and selected parameters
- Optimized data delivery to clients via a custom R package and local shiny applications to quickly generate LaTeX PDF reports
- Constructed, populated, and maintained an intranet wiki based using PostgreSQL and PHP to aid in project management and customer service

Graduate Research Assistant, Dr. Neal Stewart's Plant Biotechnology Lab
University of Tennessee, Knoxville, TN *2014 — 2016*

- Summarized statistical findings of cell suspensions using linear models, ANOVA, and PCA
- Evaluated suspension cultures via chemical and spectral processes for lignin formation and statistically analyzed and summarized my findings for inclusion in DOE reports
- Collaborated on a novel single cell suspension and cryopreservation robotic system

Laboratory Assistant, Dr. Neal Stewart's Plant Biotechnology Lab
University of Tennessee, Knoxville, TN *2012 — 2014*