# IAN JOHNSON

## BIOMEDICAL ENGINEER

(830)-688-2069 ian.james.johnson.01@gmail.com /in/ianjohnson123/ github.com/ian-james-johnson

Veteran with a background in scientific investigation and data analysis. My research background has focused on the intersection between biology and material science. This interdisciplinary background has demonstrated to me the need for new approaches to understand complex data.

## TECHNICAL SKILLS

SQL - Python - Pandas - Matplotlib - Seaborn - Plotly - Machine Learning - Natural Language Processing - Apache Spark - Data Storytelling - Git - Jupyter - Anaconda - Tableau - Statistics

## EDUCATION

**University of California at Riverside**
PhD in Bioengineering: 2010 – 2015
GPA: 3.6

**Texas Tech University Health Sciences Center**
MS in Biotechnology: 2006 – 2008
GPA: 3.4

**Texas A&M University at Galveston**
BS in Marine Science: 1998 – 2002
GPA: 2.7

**Codeup**
2021
Students develop expertise across the full data science pipeline (planning, acquisition, preparation, exploration, modeling, delivery), and become comfortable working with real, messy data to deliver actionable insights to diverse stakeholders.

## PROFESSIONAL EXPERIENCE

**Army Institute of Surgical Research**
Post-Doctoral Fellow: 2016 - 2021
Investigated bone regeneration treatments in animal models using image and data analysis of uCT volumes.

**Texas Tech University Health Science Center**
Research Technician IV: 2008 - 2010
Performed lab and animal experiments.

**US Army**
Chemical Operations Specialist: 2003 - 2006
Maintained company level CBRNE room.
Team leader for Biological Integrated Detection System (BIDS).

## DATA SCIENCE PROJECTS

**Drug Discovery**
Decision Tree: 2021
Team project to estimate the effectiveness of chemical compounds at inhibiting a disease target. This model can be used for pre-screening of chemical compounds so that only the most likely candidates proceed to expensive, time-consuming, and risky lab testing. This is important because it takes an average of $2.6 billion to develop a new drug. The model calculates Lipinski parameters and chemical fingerprints of the chemical compounds, which are used as data for the decision tree. Future work will involve improving the model accuracy.

**Investigation of Covid-19 in India**
Regression Analysis: 2021
Investigated dataset of Covid-19 and vaccinations in India using regression analysis. Demonstrated that the state in India greatly affected the relationships between vaccination, Covid mortality rates, and adverse vaccine responses. Surprisingly, the plots of Covid survivors and Covid casualties across states were near mirrors of each other. This implies that once infected, the treatment across states had similar impact on survival. This analysis will be useful for identifying best-practices in the response to Covid.

**Determine Programming Language From Readme**
Natural Language Processing (NLP): 2021
Team project to determine the programming language in Github repositories by examining the readme using NLP. It was demonstrated that the top four most commonly used programming languages had traits that differentiated them. These traits included word usage, topics implied by words used, and readme length.

**Examination of web server logs**
Anomaly Detection: 2021
Discovered anomalous patterns among web page requests and users from web server logs.