# Ericka B. Smith

(425) 308-7120

smith.ericka.b@gmail.com
erickabsmith.github.io
linkedin.com/in/erickabsmith

## SUMMARY

Statistician experienced with statistical modeling, data exploration, communication, coding, and collaborative work.
Expertise in: R, GitHub, Shiny, Tidyverse

## EDUCATION

**M.S. Statistics** — **Oregon State University** — **03/2021**
Teaching Assistant: Methods of Data Analysis (Graduate Level), Intro to Statistical Methods
Leadership: Vice President for OSU Student Statistics Org., Mentor for Statistics Dept.
Coursework: Probability, Computing, & Simulation; Data Visualization; Command Line Analysis; Python; Sampling; Survey Methods; Generalized Regression Models; Time Series; Unix/Linux

**B.S. Math/Biology** — **Western Washington University** — **06/2015**
Employment: Database Assistant using Excel & Access; STEM Tutor teaching R & Excel
Coursework: Computer Programming & Data Structures; Differential Equations; Linear Algebra

## SKILLS

**Programming:** R, Python, SQL, HTML, Shiny, Tidyverse
**Toolkit:** GitHub, Atom, Excel, EC2, Jupyter Notebook, PyCharm, PuTTY, GIS

## WORK EXPERIENCE

**Research Assistant** *06/2020 – 12/2020*
*Cooperative Institute for Marine Resource Studies*
- Leveraged data manipulation skills & discipline knowledge for analysis of large, messy datasets.
- Conducted end-to-end exploratory data analysis.
- Created publication-quality visualizations in R with ggplot2.
- Researched & implemented clustering algorithm to solve a data grouping problem.

**Intern** *08/2020 – 11/2020*
*National Oceanic & Atmospheric Administration*
- Automated data management processes in R, resulting in a dramatic decrease in turnaround time & a reduction in errors.
- Presented to an audience of 50+

**Technician** *03/2019 – 09/2019*
**Intern** *10/2018 – 03/2019*
*Quantum Spatial Inc.*
- Employed R, Python, Excel, Shiny, & GIS for machine learning modeling & predictive analytics of large spatial datasets.
- Led a training for 15 people on a software transition, saving on labor & licensing costs.
- Collaborated on multiple programming projects that effectively reached goal of streamlining data pipeline.

## PROJECTS

**Classification using Satellite Imagery and Random Forests**
*Master's Thesis* *12/2020 – current*
- Evaluated different machine learning algorithms.
- Built simulations in R training many random forests to compare accuracy between different tuning parameters & input data.

**A Comparison of Generalized Regression Models for Interpretation of COVID-19 Case Counts**
*Coursework* *11/2020 – 12/2020*
- Analyzed, fit, & evaluated negative binomial, quasi-Poisson, & Poisson regression models to reveal patterns in count data.

**Bootstrapping the Pacific Crest Trail in Parallel**
*Coursework* *10/2020 – 12/2020*
- Tested use of socket method of parallel processing workflow (R, Amazon Elastic Compute Cloud, AWS) to efficiently manage bootstrap resampling & simulations for parameter estimation & hypothesis testing.

**Interactive & Explorative Web-Based Tool for Communication of Large Ensemble Models**
*Independent Team Project* *01/2020 – 07/2020*
- Collaborated via GitHub & R to develop Shiny app that consolidates & analyzes multiple high-volume datasets.
- **Awarded Honorable Mention (2nd place) nationwide from the American Statistical Association.**
- Independently designed visualization to convey intricacies of ensemble models to technical & non-technical audiences.