

# VANESSA COLLIER

vanessacollier.datascience@gmail.com | 206-707-5505 | www.data-alchemy.studio | github.com/  
vanessacollier.datascience

## PROFESSIONAL SUMMARY

- Data analyst and research professional with experience delivering actionable insights across public policy, health, and economic domains — with a strong commitment to equity-centered, transparent, and reproducible analysis.
- Skilled in building end-to-end reproducible data pipelines, cleaning complex multi-source datasets, automating workflows, and maintaining data quality standards across recurring data processes.
- Experienced communicating findings to technical and non-technical audiences — including elected officials and agency stakeholders — with clear, interpretable visualizations and narrative summaries.
- Graduate-level foundation in statistical analysis, quantitative methods, program evaluation, and public policy applied to real-world research environments.

## TECHNICAL SKILLS

**Languages:** R, Python, SQL, HTML, CSS

**R Ecosystem:** tidyverse, sf, ggplot2, gganimate, Shiny, Plotly, rmarkdown, knitr, Quarto, data.table, tidymodels, caret, randomForest, plm, reticulate, haven, foreign, jsonlite, httr, rvest, DataExplorer, broom, ExPanDar

**Python:** Pandas, NumPy, SciPy, Matplotlib, Seaborn, Plotly, Selenium, os

**Data Engineering:** Automated pipelines (R + Python), web scraping, API integration, scheduled workflows, JSON and survey data handling

**Data Visualization & BI:** R Shiny, ggplot2, Plotly — dashboard design, production reporting, and stakeholder-facing tools

**Spatial Analysis:** sf (R), ArcGIS Pro — choropleth and animated spatial visualizations

**Version Control:** GitHub — repository architecture, workflow documentation, team onboarding and training

**Tools:** Positron IDE, RStudio, Quarto, Posit Cloud, Qualtrics, Adobe Illustrator, Google Drive, Microsoft Office

## PROFESSIONAL EXPERIENCE

**Research Professional III — Data Analyst / Analytics Engineer** — Institute for Social and Economic Research (ISER), University of Alaska Anchorage *Nov 2022 – Dec 2025*

- Built end-to-end web scraping pipeline in Python and R to collect and clean grocery price data from 50+ stores across Alaska, enabling cost-of-living research previously unavailable at this scale.
- Developed automated data cleaning and integrity validation workflows, reducing manual QA burden and ensuring consistency and accuracy across recurring data pulls — directly analogous to HMIS data quality oversight.
- Created interactive dashboards in R Shiny for economic and migration analysis, translating complex longitudinal datasets into accessible tools for non-technical stakeholders.

- Presented analytical findings directly to policymakers including Rep. Mary Peltola, translating statistical outputs into actionable policy narratives for varied audiences.
- Managed full project lifecycle across multiple concurrent research initiatives — from data acquisition and cleaning through analysis, visualization, and stakeholder delivery — with attention to reproducibility and documentation.
- Established ISER's first GitHub repository and version control infrastructure; trained research staff on Git workflows, improving reproducibility and cross-team collaboration.

## EDUCATION

**Master of Science — Data Analytics and Public Policy** | Johns Hopkins University, Advanced Academic Programs | 2019 – 2022

*Concentration: Statistical Analysis | GPA: 3.7 | Relevant courses: Program Evaluation, Quantitative Methods, Statistics for Policy Analysis, Data Science, Python for Data Analysis, R Programming, Data Visualization, Predictive Analytics*

**Bachelor of Arts — Political Science** | University of Washington | 2004 – 2007

*GPA: 3.6 | Political Science Honors Program | Dean's List | Washington State Legislative Internship (2006)*

## CERTIFICATIONS & TRAINING

- ArcGIS Foundations — Udemy / University of Alaska Fairbanks
- Human Subjects Research — CITI Program
- SQL Associate Certification — DataCamp (expected June 2026)

## PORTFOLIO PROJECT

**Healthcare Access & Provider Shortage Analysis — King County, WA** (*In Progress*)

- Independent geographic supply-side analysis using ACS, CMS/NPI, and HRSA data to identify ZIP code areas with low primary care provider density and high Medicaid enrollment. Stack: R, sf, ggplot2, Quarto, GitHub.