

# ANTHONY JOSEPH MONTAJES

☎ 604-831-0670 ✉ [montajes@student.ubc.ca](mailto:montajes@student.ubc.ca) [in](#) [anthony-joseph-montajes](#) [montajesaj1](#) 📍 BC, Canada

## EDUCATION

University of British Columbia, Bachelor of Arts

Expected 2026

Major in Cognitive Science, Minor in Computer Science

**Coursework** Algorithm Design and Analysis, Software Engineering, Cloud Computing, Machine Learning, Databases

**Honors/Awards** WL International Undergraduate Research Award (\$6,000)

## EXPERIENCE

**Data Scientist Intern**

September 2024 - Present

EPAM Systems

Vancouver, BC

- Working in a project-based team focused on implementing frameworks for LLM bias detection.
- Leveraged spaCY and pandas to conduct NLP analysis of proprietary LLM outputs, using seaborn to visualize insights and support the development of a bias detection framework.
- Investigated state-of-the-art LLM bias detection methods and initiated a proof of concept using knowledge graphs.

**Teaching Assistant**

September 2024 - Present

University of British Columbia

Vancouver, BC

- Collaborated with a Computer Science professor to teach Cognitive Science research methods, emphasizing human-centered AI in user-system development.
- Delivered personalized academic support through one-on-one consultations and online forums.
- Used Excel to visualize aggregate class statistics; revised syllabus to better meet student needs.

**Data Analyst Research Assistant (Co-op)**

March 2024 – September 2024

Strategy and Business Economics Division, Sauder School of Business

Vancouver, BC

- Secured WLIURA ☞ (\$6000) to conduct research on the relationship between market concentration and innovation pipelines in the pharmaceutical industry; cross-collaboration with researchers from the University of Virginia and MINES ParisTech.
- Implemented fuzzy string matching algorithms in Python using Levenshtein Distance to identify matches in disjoint data.
- Engineered a Selenium based webscraper automating the data collection for 3000+ molecules to be used as training data. Employed stratified and clustering sampling techniques in conjunction with black box tests to verify outputs.
- Created ETL scripts to perform data-cleaning making efficient use of Python's anonymous lambda functions and regex.
- Leveraged Pandas to clean data, perform EDA, and build merge pipelines for csv files with 100,000+ observations.

**Data Science Coordinator**

October 2023 – Present

Visual Cognition Lab, UBC

Vancouver, BC

- Engaged project leads to develop workshops targeting pain points in statistical intuition, visualization, and reproducibility.
- Streamlined analysis by engineering internal tools and reproducible workflows using Python and Firebase.
- Utilized JavaScript, NodeJS, and React to construct experiment conditions; maintained CI/CD using GitHub Actions.

## PROJECTS

**Audit Risk Classifier** | *R, TidyVerse, Jupyter*

December 2023

- Utilized R programming to enhance fraud detection, creating a KNN classifier identifying high-risk firms with an accuracy of 87.7%, precision of 86.2%, and recall of 94.9% within a dataset of 777 firms.
- Leveraged TidyVerse to streamline data preprocessing, feature engineering, visualization, and model fitting.

**Matrix Calculator** | *Java, JUnit, Swing*

December 2022

- Developed an application to perform matrix operations with a GUI in Swing and persistence using JSON.
- Implemented a robust back-end leveraging OOP design principles (Singleton, Observer Pattern, Interfaces).
- Achieved 100% code coverage across all branches using JUnit following blackbox and whitebox testing methods.

## SKILLS

**Programming Languages** Python, R, SQL, Java, C, C++, TypeScript, Scala.

**Tools/Environments** Tableau, Excel, Conda, Jupyter, Linux, Git, MongoDB, AWS, Bash, Docker.

**Data Science** ETL, Scrapers, Pipelines (Cleaning, Wrangling, Modelling), Hypothesis Testing, Bayesian Analysis.

**Libraries** selenium, pandas, numpy, scipy, matplotlib, seaborn, sklearn, PyMC, PyTorch, TidyVerse, spaCy.