Aaron Ma

Vancouver, BC | LinkedIn | (647) 336-9078 | aaronardenma@gmail.com | Github | Portfolio

EDUCATION

University of British Columbia — **Vancouver, BC**
***BS**c in Integrated Science* | *Computer Science, Data Science, Statistics (GPA: 3.8)* — *Sep 2023 – May 2026*
***BA** in Psychology, Minor in Commerce* with *Distinction* — *May 2022*

Relevant Courses: *Software Construction* | *Statistical Inference for Data Science* | *Basic Algorithms and Data Structures (Upcoming)*

SKILLS

Languages: HTML5, CSS3, JavaScript, Python, R, SQLite, Java
Tools & Frameworks: SAP Analytics Cloud | pandas | NumPy | matplotlib | pytest | Git | tidyverse | ggplot2 | DBI | dplyr | plotly | nltk | Tableau | Swing | JSON
Certifications: Google Data Analytics – *Google Career Certificates*

WORK EXPERIENCE

SAP — **Vancouver, BC**
iXp Customer Success Marketing & Communications Specialist — *Dec 2022 – Aug 2023*

- Automated 3 unique customer success (CS) events' data cleansing & manipulation workflows using **Python** libraries **pandas** & **NumPy**, increasing efficiencies by on average **98%**
- Spearheaded a CS **data dashboard** for 6 products for SAP North America's 50+ internal pre/post sales teams with **SAP Analytics Cloud**
- Decreased **time** lag of CS event data availability to internal teams by **50%** through **workflow automation**
- Wrote **64 self-contained unit tests** for data validation using **pytest** reflecting full code coverage with documentation for technical & non-technical users
- Delivered quarterly business insights using **Excel Pivot Tables & Charts** involving customer acquisition, engagement, and breakdown
- Crafted **impact analysis decks** for 5 CS events using **Excel**, **Tableau** and **Python**, analyzing redundancies & alignment with business objectives

PROJECTS

Netflix Wrapped Data Analysis — *March 2023 | Repository*

- Utilized **Python** libraries **pandas** and **nltk,** to perform **data manipulation** and create **automated workflows** with Netflix viewing activity
- Used **Kaggle Netflix Dataset** as a Netflix API alternative for **data validation** and **enrichment** for media type, rating, and in the future genre analysis
- Executed **token analysis matching** for **data validation** between different database sources for future enrichment
- Performed **data visualization** with **plotly** illustrating watch time patterns across media types, unique titles, ratings, and time

Game Party Finder — *Feb – Mar 2023 | Repository*

- Used **Java** with **OOP** (Object-Oriented Programing) to create a party finder for video games amongst individuals
- Conducted **unit testing** with **JUnit** in non-UI packages with **TDD** (Test Driven Development) principles
- Creating graphical user interface with **Swing** library with **exception handling** in model and ui packages
- Added data persistence with **JSON libraries**, covering all aspects of package objects

Maternal Health Risk Classifier Model — *Nov 2023 – Dec 2023 | Repository*

- Used **R** in **Jupyter** to perform **k-nearest neighbors classification** on UCI Maternal Health Risk Dataset
- Investigated 6 health factors that have been shown to impact maternal health risk levels in a Bangladesh population
- Cleaned and manipulated dataset using **R libraries**: **tidyverse**, **dpylr**, **kknn**
- Created **data visualizations** with **ggplot2** to show risk level distributions, risk factor breakdown, and classifier accuracy metrics
- Achieved **82% accuracy**, **88% precision**, **86% recall** post-optimization through **cross-validation**
- Utilized **git**, and **Github** for version control and merge conflicts