# Ribhay Singh

[singhribhay@gmail.com] | [925-315-0834] | [https://www.linkedin.com/in/ribhaysingh/] | [https://ribhay.vercel.app/]

## EDUCATION

**Arizona State University** | Tempe, AZ **December, 2025**

*Bachelor of Science in Data Science (Computer Science Track)*

- Awarded **Dean's List** for academic excellence in Fall 2024 and Spring 2025

## INDUSTRY EXPERIENCE

***Initialyze - San Francisco Bay Area***

*Software Engineer Intern* **May 2024 – August 2024**

- Led a Data ingestion project to acquire and integrate data from internal web services into Adobe Experience Platform using Dataflow UI and REST APIs, improving data integration **throughput by 20%**.
- Utilized Databricks notebooks for ETL development, enabling data transformation, validation and logging across distributed datasets.
- Performed data profiling and querying using MySQL and AEP Query Service to enhance reporting accuracy and validate data integrity.
- Contributed to schema creation and data governance processes within AEP's Experience Data Model.

## PERSONAL PROJECTS

**FitPro - Fitness Analytics Platform | Python, PostgreSQL, scikit-learn, Databricks, Tableau, Streamlit**

- Developed a full-stack fitness tracking and analytics web application that predicts calories burned with **80% $R^2$ score**, analyzes workout trends and forecasts goal achievement using tree-based ML models (XGBoost, Random Forest).
- Designed a PostgreSQL data pipeline with Databricks for real time analytics and goal tracking,
- Developed Tableau dashboards and Streamlit modules to visualize trends and calorie predictions.

**EchoMind AI - Mental Health Journaling Tool | Python, PostgreSQL, Streamlit, FastAPI , GPT-3.5**

- Developed a mental health journaling tool that summarizes and analyzes user entries using an LLM powered by GPT-3.5, achieving **90% sentiment alignment** during internal validation.
- Designed and deployed a FastAPI REST API backend by SQLite to process **100+ journal entries**.
- Implemented an intuitive Streamlit frontend for real-time journaling and instant feedback.

**Real Estate Price Prediction Model | Python, scikit-learn, pandas, Flask, HTML/CSS**

- Built a machine learning model using multiple linear regression to predict housing prices, achieving **85% $R^2$ score.**
- Cleaned and analyzed a Kaggle housing dataset, applying EDA and feature engineering to identify key factors.
- Deployed a Flask web application with HTML/CSS for live user inputs and real-time price predictions.

**World Happiness Report | R (dplyr, ggplot2), Python (pandas, Numpy, Matplotlib)**

- Conducted multivariate analysis in R on the 2019 World Happiness Report covering **150 countries** to model socio-economic impacts on happiness scores.
- Built regression models and visualization pipelines to uncover global trends and key happiness factors.
- Developed a Python analysis workflow to validate results and enable cross-validation of findings.

## TECHNICAL SKILLS

**Languages:** Python, R, SQL, Java, C++, HTML/CSS/ JavaScript
**Libraries & Frameworks:** scikit-learn, PyTorch, TensorFlow, pandas, NumPy, FastAPI, Flask, Streamlit
**Data & ML:** Data wrangling, feature engineering, predictive modeling, NLP, statistics & probability, A/B Testing, Excel
**Tools:** Tableau, Power BI, VS Code, Databricks, Spark, Git/Github, Adobe Experience Manager, Adobe Experience Platform
**Certifications:** Complete Data Science Bootcamp, Complete SQL Bootcamp, HTML/CSS/Javascript/React Online Certification Course, Adobe Experience Manager Masterclass