

SHWETA YADAV

DATA ANALYST

CONTACT

✉ datascientist.shweta@gmail.com
☎ (720) 345-4211
📍 San Jose, CA
🌐 linkedin.com/in/shwetayadav15/
🐙 github.com/shweta-yadav15/
🌐 shweta-yadav15.github.io/

PROFILE

A Data Analyst passionate about combining data, storytelling, and kindness as a catalyst for social change and to make a positive difference in the community.

EDUCATION

UNIVERSITY OF COLORADO, Denver | 2020
MS in Computer Science

G B PANT ENG. COLLEGE, India | 2018
B. Tech in Computer Science

TECHNICAL SKILLS

- **Programming:** Python, SQL
- **Database Management:** MySQL, PostgreSQL, Oracle
- **Data Visualization:** Tableau, Matplotlib, Seaborn, Power BI
- **Data Wrangling:** Pandas, NumPy, Scikit-Learn, SciPy, Excel
- **Cloud Computing:** AWS
- **Deployment:** Flask, Heroku, HTML, CSS, Bootstrap
- **Statistics:** Hypothesis testing, Central Limit Theorem, Statistical Modeling, A/B testing

MANAGEMENT SKILLS

- Leadership & Growth mindset
- Project & Team Management
- Communication & Problem Solving

WORK EXPERIENCE

Data Scientist | EmployHer, Seattle 09/2020 – Present

- Working on a **Speech Recognition System** that first extracts speech from a video and then convert it into a text document that will be further used to pull out candidate's skills.

Data Analyst | University of Colorado, Denver 02/2019 – 05/2020

- Evaluated performance of professors based on students' comments on ratemyprofessor.com by developing a **sentiment analysis system** that classify them into three categories positive, negative and neutral.
- Transformed, vectorized text data, used lemmatization, n-gram, and tf-idf. Implemented model using **Naive Bayes** with 75% accuracy. Integrated the data and trained model on **AWS**.

Junior Data Analyst | Apex Health, India 01/2018 – 07/2018

- Identified trends related to conditions and treatment costs in different populations using **large medical and dental claims data**. Performed **statistical analysis** and visualizations on 30+ providers that was used to mine association rules.

Business Analyst Intern | BHEL, India 05/2017 – 08/2017

- Performed ad-hoc analysis with **SQL on products sales data** for a financial forecasting model to improve manufacturing process.
- Visualized monthly sales growth using Tableau to help leaders to act on sales revenue trends resulting in **30% increased sales**.

DATA ANALYTICS PROJECTS

Disaster Messages Classification 🔄

- Deployed a **web-application** that classifies messages in **real-time** for emergency workers into 36 categories to detect important messages that need immediate attention.
- Developed the model with ETL process using **random forest, machine learning pipeline and NLP** algorithms with 75% precision.

E-Commerce Data Analysis 🔄

- Performed detailed **exploratory data analysis** using visualization libraries to understand key metrics affecting customer's review score. Predicted **customer lifetime value** after segmenting customers for next 6 months using **XGBoost** with 90% accuracy.

Mental health in Tech Industry Analysis 🔄

- Analyzed employers' trends towards mental health and developed a classifier model to identify how likely it is that an employee will seek treatment with **AdaBoost classifier and decision trees** resulting in 82.18% accuracy