Anusha Anand Karve

San Diego, CA | anushakarve@gmail.com | 206-423-1565 | linkedin.com/in/anusha-k-057a28329

F-1 Student | CPT Eligible (Summer 2026)

SUMMARY

Data Science graduate student at San Diego State University with experience designing scalable data warehouses, building end-to-end machine learning pipelines, and implementing time-series analytics systems. Strong foundation in SQL, Python, statistical modeling, and data engineering principles with a focus on translating complex datasets into actionable business insights. Seeking Data Science / Data Analytics internship opportunities.

EDUCATION

San Diego State University (SDSU), San Diego, CA   Aug 2025 – Present

M.S., Computational Science (Data Science Emphasis)

GPA: 3.8 / 4.0

Relevant Coursework: Computational Database Fundamentals, Mathematical Modeling, Computational Methods for Scientists

P.E.S's Modern College of Engineering
(Savitribai Phule Pune University), India   Aug 2021 – May 2025

B.E., Electronics & Computer Engineering

Ranked 1st in Department

Relevant Coursework: Database Management Systems, Data Mining & Warehousing, Data Analytics using Python, Data Structures & Algorithms, Business Intelligence

PROJECTS

Instacart Retail Analytics Warehouse (30M+ Records)   Sep 2025 – Dec 2025

Tech: PostgreSQL, Python, SQL, Docker, Power BI

- Built a scalable PostgreSQL data warehouse ingesting 30M+ Instacart order-item records using bulk `COPY` loading and incremental batch processing.
- Designed a partitioned star schema (`fact_orders`, `fact_order_items`, `dim_user`, `dim_product`, `dim_date`) to support performant behavioral retail analytics.
- Implemented idempotent incremental load logic and automated data quality checks to maintain referential integrity and consistency.
- Developed optimized SQL queries and BI dashboards analyzing customer retention, reorder rates, cohort behavior, and basket size trends.

Customer Churn Risk Analytics Pipeline (IBM Telco)   Jan 2026 – Apr 2026

Tech: Python, SQL, Pandas, scikit-learn

- Built an end-to-end churn analytics pipeline to identify high-risk customers and support retention targeting.
- Engineered predictive features including tenure buckets and service engagement proxies; optimized models using PR-AUC and threshold tuning.
- Addressed class imbalance using class weights and evaluation via Precision–Recall tradeoffs.
- Implemented batch scoring workflow producing churn probabilities, risk bands, and deployable artifacts.
- Added model interpretability using feature importance and customer-level driver summaries.

Event-Driven Market Volatility & Sector Sensitivity Platform (Ongoing)   May 2026 – Present

Tech: Python, SQL, PostgreSQL, Pandas, NumPy, Time-Series Analysis

- Designed and implemented a time-series analytics pipeline integrating SPDR sector ETF price data with macroeconomic event calendars.

- Engineered rolling risk metrics including daily returns, 21-day and 63-day annualized volatility, and maximum drawdown.

- Developed event-window analysis (-5 to +5 trading days) to measure cumulative sector returns surrounding macro releases (e.g., CPI).

- Currently extending the system with automated API ingestion (Yahoo Finance, FRED) and volatility regime classification modules.

PROFESSIONAL EXPERIENCE

Data Science Intern May 2024 – Aug 2024
S.M. Enterprises

- Built and evaluated predictive forecasting models on structured operational datasets to support business planning and demand estimation.

- Performed end-to-end data preprocessing including missing value handling, feature engineering, and validation using Python and SQL.

- Conducted exploratory analysis to identify key drivers influencing operational performance metrics.

- Developed stakeholder-facing analytical reports and visualizations translating model outputs into actionable insights.

- Collaborated with business teams to refine data requirements and improve reporting workflows for recurring analytics tasks.

Junior Analyst Intern Jan 2024 – Apr 2024
TechR Business Solutions

- Conducted SQL-based performance analysis on operational datasets to monitor KPIs and identify emerging performance trends.

- Designed and improved structured data transformation workflows to enhance reporting accuracy and consistency.

- Developed interactive dashboards enabling leadership visibility into customer and operational metrics.

- Automated recurring reporting queries to reduce manual analysis effort.

Software Developer Intern Oct 2023 – Dec 2023
TechR Business Solutions

- Developed backend automation scripts to streamline internal data handling and reporting processes.

- Optimized data extraction and preprocessing modules to improve reliability of downstream analytics workflows.

- Worked cross-functionally with analytics teams to structure data outputs for reporting and modeling use cases.

TECHNICAL SKILLS

Programming: Python, SQL, R

Data Engineering & Warehousing: PostgreSQL, Dimensional Modeling, Star Schema Design, Partitioning, Incremental ETL, Bulk Loading (COPY), Data Quality Validation

Machine Learning & Statistical Modeling: Logistic Regression, Tree-Based Models, Feature Engineering, Class Imbalance Handling, PR-AUC Optimization, Model Interpretability

Analytics & Time-Series: Exploratory Data Analysis, Cohort Analysis, Rolling Volatility, Event-Window Analysis, Drawdown Metrics

Business Intelligence & Visualization: Power BI, KPI Development, Dashboard Design, SQL Analytical Views

Tools: Docker, Git, Jupyter, Linux