

PAYAL PRAVIN GORE

☎ +1-812-327-3770 ✉ payal.gore6433@gmail.com 🔗 [LinkedIn: Payal Gore](#) 🌐 [Portfolio](#) 📍 Open to Relocation

EDUCATION

Master of Science in Data Science	Indiana University Bloomington	3.47	2023 - 2025
Bachelor of Engineering in Information Technology	Mumbai University	9.55	2019 - 2023

SKILLS

Data Analytics & Visualization: SQL (CTEs, Joins, Window Functions, Optimization), Python (Pandas, NumPy, Matplotlib, Scikit-learn), R (ggplot2, plotly), Power BI (DAX, Power Query), Tableau, Excel (Pivot Tables, VLOOKUP)
Data Processing & Statistical Analysis: Data Wrangling, Cleaning, Preprocessing, Exploratory Data Analysis (EDA), KPI Tracking, A/B Testing, Statistical Modeling, Data Storytelling, Decision-Making Insights
Database & Cloud Technologies: MySQL, PostgreSQL, AWS (S3, EC2, SageMaker), GCP, NoSQL (MongoDB)
Machine Learning: Customer Segmentation, Time Series Analysis, Predictive Modeling, Market Basket Analysis, ETL

PROFESSIONAL WORK EXPERIENCE

Data Research Analyst

Dec 2024 - Present

Indiana University Bloomington – School of Public Health, IN, USA

Nutritional Impact Study – Protein vs. Carb Bars

- Wrangled & cleaned data for 250+ participants using Python (Pandas, NumPy) & Excel (Pivot Tables, VLOOKUP)
- Uncovered preference for whey protein & low-sugar bars through EDA, guiding product formulation
- Ran A/B tests comparing protein & carb bars; protein bars cut post-meal calories by 20%, shaping product strategy

Student Health Engagement Study

- Built a GCP-based segmentation pipeline integrating CRM & Google Analytics data for 1,200+ students; used BigQuery & DataFlow to cluster users into 4 engagement profiles, optimizing outreach strategies
- Streamlined campaign targeting by processing engagement data with Cloud Functions & BigQuery, & visualizing insights in Looker Studio; increased participation by 25%

Research Analyst

May 2024 - Jul 2024

Indiana University Bloomington – Luddy School of Informatics, IN, USA

- Queried & visualized with SQL & Tableau to extract KPIs from 1,000+ student logs - tracking logins, collaboration, & edits; found 35% had no peer interactions despite high usage.
- Used time series analysis to track login peaks, session lengths, & engagement drop-offs, enabling targeted support
- Worked with Product Managers to define goals in requirement docs, presenting insights to leadership via PowerPoint

Data Science Intern

Jun 2022 – Sep 2022

Vidyalankar Institute of Technology Mumbai, MH, India

- Applied regression on Qualtrics data to identify 5 key stress drivers, with academic pressure impacting 80% of students
- Created geospatial Power BI dashboards using map visuals, anxiety-level heatmaps, & campus-type slicers to reveal 40% higher anxiety in urban areas, guiding wellness actions in 3 regions

PROJECT EXPERIENCE

CartInsights: Decoding Customer Purchasing Behavior - [Github](#)

Feb 2025

- Segmented customers with K-Means & PCA into 5 shopper profiles, boosting retention by 15% via targeted marketing.
- Applied Apriori algorithm on 3M+ transactions, uncovering 28 high-lift product pairs, boosting cross-sell sales by 10%.
- Built XGBoost & Neural Network models (AUC-ROC: 0.85) to predict reorders, improving demand forecasting & cutting inventory overstock by 20%.

HealthPulse: US Healthcare Analytics - [Github](#)

Jan 2025

- Cleaned & transformed U.S. healthcare data (2019-2020) using Power Query & M Language, improving data accuracy & readiness for analysis by 20%.
- Designed custom DAX measures & interactive dashboards with drill-throughs, bookmarks, & slicers to visualize hospital performance, patient outcomes, & payer-provider dynamics.
- Identified cost-saving opportunities by analyzing expense trends, improving resource allocation by 15%.

DataFlowX: Scalable Big Data Pipeline for Sales Analytics - [Github](#)

Dec 2024

- Designed & implemented a scalable data pipeline using AWS (S3, EC2) & PySpark, optimizing data ingestion, transformation, & analytics for 10,000+ sales records.
- Built ETL workflows & automated data processing with Spark SQL, AWS CLI, & SageMaker, enabling predictive modeling ($R^2 = 78.5\%$) & interactive visualizations in QuickSight.

Excel Finance Insight - [Github](#)

Dec 2024

- Built automated dashboard, tracking \$1M in revenue, margins, & expenses using Power Query, Power Pivot, & DAX
- Used vlookup for dynamic variance analysis for budget, prior period, & prior year comparisons
- Created interactive visuals (doughnut charts, KPI boxes, cost breakdowns) to streamline reporting for 10+ stakeholders

A/B Test Facebook vs. AdWords Performance with Python - [Github](#)

Nov 2024

- Performed T-test to validate conversion rate differences b/w ad platforms, enabling 20% ad spend reallocation
- Analyzed click-through rates (CTR), conversion rates, & cost-per-conversion (CPC) for both platforms
- Developed regression model to predict Facebook ad conversions from clicks, achieving an R^2 score of 76.35%

RESEARCH PUBLICATION EXPERIENCE

Authored the research paper [Text to Image Synthesis using Generative Adversarial Networks](#)