

SAICHOLIK VEMPATI

Mount Pleasant, Michigan | saicholikvempati@gmail.com | +1 (989) 572 8478 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

M.S. Information Systems (Data Analytics) | Central Michigan University, Mount Pleasant, MI | GPA: 3.8/4.0 | Aug 2024 - May 2026

Coursework: Business Data Analytics, Database Systems, Enterprise Systems Management (SAP S/4HANA), Systems Analysis and Design, Statistical Modeling, R for Business Analytics, Tableau Dashboarding, Data Mining, Analytical Decision Modeling, Data Driven Quality Management, Project Management

B.Sc. Mathematics, Statistics and Computer Science | J.K.C. College, India | Aug 2020 - May 2023

Coursework: Stochastic Processes, Probability Theory, Statistical Inference, Numerical Methods, Linear Algebra, Regression Analysis, Optimization Theory, Computational Statistics, Relational Database Management

WORK EXPERIENCE

Student Technical Assistant

Central Michigan University Library, Mount Pleasant, MI | Aug 2024 - Present

- Collect, clean, and validate data from multiple sources; perform statistical analysis identifying trends, patterns, and outliers; develop Power BI and Excel dashboards tracking KPIs reducing reporting errors by 15 percent for cross functional stakeholders.
- Build Python ETL pipelines managing end to end data flows including ingestion, transformation, validation, and delivery; implement error handling, logging, and automated alerting; maintain data models and document pipelines and data dictionaries.
- Develop process documentation, controls, and work instructions supporting standardization, compliance, and audit readiness; provide user support and training; present analytical findings clearly to non technical leadership.
- Support ad hoc analysis and root cause investigation; collaborate with cross functional stakeholders to understand data needs, resolve data issues, and drive continuous improvement of analytical tools and reporting workflows.

Data Analyst Intern

Red Carpet Pvt Ltd, India | Mar 2023 - Nov 2023

- Collected, cleaned, and analyzed large operational datasets using SQL and Python; developed Power BI dashboards and automated reporting pipelines improving reporting accuracy by 25 percent and enabling data driven strategic decision making.
- Designed and implemented statistical and predictive models using regression, correlation, and time series techniques; conducted scenario analysis and ad hoc analyses delivering actionable insights to senior leadership and cross functional stakeholders.
- Built Python ETL pipelines with error handling, logging, and automated alerting; maintained MySQL databases; created process documentation, SOPs, and data dictionaries supporting team knowledge transfer and operational excellence.
- Prepared weekly and monthly status reports and stakeholder presentations; supported regulatory compliance through accurate data management and audit ready reporting; collaborated with Finance, Operations, and Marketing teams.

SAP ERP EXPERIENCE

SAP S/4HANA Enterprise Systems Management

Central Michigan University, BIS647S Enterprise Systems Management | 2024 2025

- Completed hands on SAP S/4HANA rotations across five core modules: Materials Management (MM), Production Planning (PP), Financial Accounting (FI), Managerial Accounting (CO), and Sales and Distribution (SD) in a live enterprise simulation environment.
- Executed the complete Procure to Pay (P2P) process: created material master records across 10 plus views, ran MPS and MRP (transaction MD40), created purchase orders, posted goods receipts, processed vendor invoices, and posted outgoing payments.
- Executed the complete Order to Cash (O2C) process: processed customer sales orders, ran MRP to generate planned production orders, followed orders through manufacturing execution, managed finished goods inventory, and completed customer delivery and billing.
- Managed FI and CO workflows: created general ledger accounts, posted vendor invoices, processed accounts payable, allocated costs across cost centers (Accounting, Marketing, Warehouse, Sales, IT), and performed cost center reposting.
- Produced structured audit ready SAP process documentation and memo reports with screenshot evidence verifying accuracy of all system transactions across all five module rotations.

TECHNICAL PROJECTS

NBA 2024 25 Players Analytics Dashboard

Tools: Power BI, DAX, Power Query, REST API, SQL | Business Value: Self Service Sports Performance Analytics

- Business problem: sports analysts and non technical stakeholders needed real time player performance visibility without requiring a data analyst for every report, simulating enterprise sales and operations BI needs.
- Built live enterprise Power BI dashboard integrating NBA Stats REST API via automated Power Query pipelines with daily data refresh; developed semantic data model linking players, teams, games, and season statistics across 500 plus players and 30 teams.
- Developed 20 plus advanced DAX measures tracking KPIs including points per game, field goal percentage, assists, rebounds, shooting efficiency, and Top N rankings; applied statistical benchmarking identifying performance outliers.
- Conducted full validation testing against documented business requirements; enabled 100 percent self service analytics via interactive drill down filters eliminating analyst dependency for recurring performance questions.
- Key metrics: 500 plus players analyzed, 30 teams, 20 plus KPIs tracked, live automated API refresh, 100 percent self service analytics for non technical business users.

Airbnb Market Analysis and Global Productivity Dashboard

Tools: Tableau, SQL, Excel | Business Value: Pricing Optimization, Market Gap Analysis, Neighborhood Benchmarking

- Business problem: fragmented pricing and availability data across global markets made it impossible to identify supply gaps, underpriced neighborhoods, and demand trends to support pricing optimization and operational planning decisions.
- Collected, cleaned, and validated 200,000 plus raw Airbnb records from multiple source systems using SQL and Excel Power Query; resolved data quality issues including missing values, duplicate listings, and currency inconsistencies across 200 plus geographic regions.
- Built interactive Tableau dashboard with Top N most reviewed listings by neighborhood, room type distribution by borough, average price by neighborhood (Manhattan \$198.5, Staten Island \$163.5, Brooklyn \$129.5, Queens \$103.2, Bronx \$94.7), and host growth trends (2008 2015).
- Applied statistical trend analysis and outlier detection to surface pricing gaps across 5 room type categories; dynamic Top N parameter filters enabled stakeholders to self serve pricing and market insights.
- Key metrics: 200,000 plus records processed, 200 plus regions analyzed, 5 neighborhood price benchmarks, 8 years of host growth trend data (2008 2015), 30 percent reduction in manual reporting time.

US Geographic Profit Distribution and Trends Dashboard (2020 2023)

Tools: Tableau | Business Value: Regional Profit Performance Monitoring, Executive Reporting

- Business problem: leadership needed geographic visibility into profit distribution and regional performance trends across the US to support strategic resource allocation and territory management decisions.
- Built an interactive Tableau dashboard visualizing profit distribution across all US states using geographic map visualization with regional color coding (Central, East, South, West); integrated year filter allowing stakeholders to isolate any single year from 2020 to 2023.
- Developed multi year line chart tracking profit trends from 2020 to 2023 across all four regions enabling leadership to identify growth trajectories, declining regions, and seasonal patterns at a glance.
- Embedded numbered state level profit labels directly on the map allowing regional managers to identify top and bottom performing states without any additional navigation or analyst support.
- Key metrics: 4 US regions analyzed, 50 state level profit data points, 4 year trend analysis (2020 2023), interactive year selector, executive ready single view geographic profit monitoring dashboard.

Global Labor Patterns: Work Intensity, Productivity, and Retirement Trends

Tools: Tableau | Business Value: Workforce Analytics, Global HR Benchmarking, Productivity Threshold Analysis

- Business problem: HR and policy leaders needed a global comparative framework to evaluate which countries were overworking their populations, underperforming on productivity, and how retirement age trends varied by gender across a decade.
- Built an interactive Tableau dashboard with 4 dynamic parameter controls: Hours Threshold, Productivity Threshold, Country Parameter (Top N), and Gender selector; enabled stakeholders to define their own benchmarks and instantly see which countries fell into Efficient and Balanced, Moderate, or Overworked and Underproductive quadrants.
- Developed working hours vs. retirement age scatter analysis for top 10 countries (Chile, China, Greece, India, Indonesia, Malta, Mexico, Poland, South Africa, South Korea) with annual working hours ranging from 2,032.8 to 2,251.3 hours; visualized retirement age gender gap trends from 2007 to 2017.
- Applied threshold based quadrant classification enabling leadership to instantly benchmark any country against global productivity and labor intensity standards with a single parameter adjustment.
- Key metrics: 10 plus countries benchmarked, 11 years of retirement age trend data (2007 2017), working hours range 2,032 to 2,251 hours annually, 4 dynamic parameter filters, gender disaggregated retirement trend analysis.

VDI System Usage: IT Risk and Compliance Monitoring

Tools: R, SQL, ggplot2, dplyr, tidyverse | Business Value: IT Risk Management, Compliance Reporting, Infrastructure Optimization

- Business problem: the IT department had no structured visibility into VDI usage patterns, compliance violations, or infrastructure utilization across a large enterprise user base; manual monitoring was slow, inconsistent, and prone to errors.
- Loaded and joined two datasets (VDI server logs and application execution logs) via inner join on VDI_ID; cleaned timestamps, removed invalid stop times, filtered to CMU machines only, and standardized data across 227,386 enterprise operational records.
- Identified 3,013 unique users in 2015 with an average of 139.5 distinct users per day and a peak of 577 users on March 2 aligned with academic cycles; identified top 5 power users (userid5179 with 2,708 logins leading); analyzed remote OS distribution (Windows 179,032 sessions, Mac 46,860 sessions, WinStore 528, iOS 310, Android 65, Linux 65).
- Produced ggplot2 boxplot visualization of shipping time by product category across years; identified compliance risk patterns and power user behavior; delivered audit ready structured reports with documented findings and recommendations for CIO and management.
- Key metrics: 227,386 records analyzed, 3,013 unique users identified, 577 peak daily users, 7 OS categories classified, 5 top power users profiled, 22 percent improvement in compliance reporting accuracy, full audit ready documentation and R code delivered.

US Retail Sales Analysis: Shipping, Profitability, and Customer Insights

Tools: R, tidyverse, dplyr, ggplot2, lubridate | Business Value: Supply Chain Optimization, Regional Profitability, Customer Intelligence

- Business problem: retail operations leadership needed structured analysis of multi region sales data to identify shipping inefficiencies, profitability gaps by product category, underperforming customer segments, and the highest value customers by order size.
- Performed complete data wrangling in R: loaded Central (CSV), West (CSV), and East (tab delimited) region datasets; cleaned and standardized date formats using lubridate (make_date, as.Date, parse_number), resolved duplicate columns, handled NA discounts via coercion, and merged all three into a unified 13,000 plus record allData tibble using bind_rows.
- Identified West region ships fastest (3.90 days average), East at 3.91 days, and Central slowest at 4.06 days; found Technology as the highest profit category in Central (\$33,697) and East (\$45,696), while Office Supplies led in West (\$168,044); identified Consumer segment as lowest profit in Central, Corporate lowest in East, and Home Office lowest in West.
- Produced faceted ggplot2 boxplot visualizing shipping time by product category across 2015 to 2018; identified Raymond Buch as the highest value single order customer chain wide (\$42,157 order CA-2018-140151 in the West region); delivered yearly sales by region showing West growing from \$442,619 (2015) to \$773,739 (2018).

- Key metrics: 3 regional datasets merged (Central, West, East), 13,000 plus records analyzed, 4 years of data (2015-2018), 3 shipping speed rankings, top 3 slowest shipping products identified per region, highest value customer identified (\$42,157 single order), West region annual sales grew 75 percent from 2015 to 2018.

Donation and Charity Management System

Tools: Python, MySQL, CustomTkinter, Matplotlib, ReportLab, bcrypt | Business Value: Nonprofit Operations Automation, Donor Management, Campaign Analytics

- Business problem: a charity organization managed donor records, campaigns, and fund allocations manually using spreadsheets leading to data entry errors, duplicate records, missing audit trails, and inability to generate accurate fund usage reports.
- Designed and built a fully functional desktop application using Python and CustomTkinter with dual role access control (Admin and Donor portals); designed 4 normalized MySQL database tables (users, campaigns, donations, user_preferences) with foreign key constraints, indexed columns, and referential integrity enforcement via InnoDB engine.
- Implemented bcrypt password hashing for secure authentication; built Admin dashboard with real time campaign KPI tracking, donor management, PDF report generation using ReportLab, and Matplotlib embedded analytics charts; built Donor portal with campaign browsing, donation submission, personal history tracking, and notification preference management.
- Developed automated validation scripts with complete error handling and audit ready transaction logging; implemented email notification system, monthly report generation, and impact update distribution; ensured 100 percent automated data entry validation eliminating manual errors.
- Key metrics: 4 normalized database tables with foreign key integrity, dual role user access system (Admin and Donor), bcrypt secured authentication, PDF report generation, Matplotlib embedded analytics, 100 percent automated validation, audit ready transaction history, zero manual data entry required post implementation.

Software Project Management Simulation

Tools: SimProject, EVM, Excel, PowerPoint | Business Value: Enterprise IT Project Controls, Financial Performance Modeling, Executive Reporting

- Business problem: simulated a real world enterprise IT project management engagement requiring full financial tracking, cost and schedule variance analysis, and executive SteerCo reporting across a 13-week \$44,460 project.
- Built Excel based EVM financial models tracking Budget at Completion (BAC), Earned Value (EV), Actual Cost (AC), Schedule Performance Index (SPI), Cost Performance Index (CPI), Schedule Variance (SV), and Cost Variance (CV) at each project milestone.
- Identified budget and schedule risks early through variance analysis; recommended and implemented corrective actions realigning the project to baseline targets; prepared and delivered SteerCo style executive presentations communicating project status, risks, and recommendations to non technical stakeholders.
- Achieved 98 percent quality adherence across all project deliverables; reduced schedule variance by 12 percent through proactive intervention; delivered complete EVM documentation including weekly status reports and milestone performance summaries.
- Key metrics: \$44,460 budget managed, 98 percent quality adherence, 12 percent schedule variance reduction, 13 weeks of weekly SteerCo reporting, full EVM documentation (BAC, EV, AC, SPI, CPI, SV, CV) delivered.

TECHNICAL SKILLS

SAP ERP: SAP S/4HANA, Materials Management (MM), Production Planning (PP), Financial Accounting (FI), Managerial Accounting (CO), Sales and Distribution (SD), Procure to Pay (P2P), Order to Cash (O2C), Material Master Creation, MRP and MPS (MD40), Cost Center Management, General Ledger Accounting

Programming: Python (Pandas, NumPy, SciPy, Matplotlib, Scikit learn, CustomTkinter, ReportLab, bcrypt, openpyxl), R (tidyverse, dplyr, ggplot2, lubridate, lm, glm), SQL (MySQL, Oracle SQL, SQL Server), MATLAB, C++ (learning)

Business Intelligence: Power BI (DAX, Power Query, Semantic Models, Drill Down Visualizations, Validation Testing), Tableau (Geographic Maps, Parameter Controls, Top N Filters, Trend Analysis, Quadrant Analytics), Excel (Advanced: Pivot Tables, VLOOKUP, XLOOKUP, SUMIFS, Power Query), Looker Studio

Data Engineering: Python ETL Pipelines, REST API Integration, MySQL Database Design, Foreign Key Constraints, InnoDB Engine, Error Handling and Logging, Data Validation and QA, GCP, BigQuery, AWS, Snowflake (familiar), Databricks (learning), Git and GitHub

Statistical and Quantitative: Statistical Modeling, Regression Analysis, Time Series Forecasting, Stochastic Processes, Numerical Methods, Optimization Theory, EVM Methodology, Anomaly Detection, Threshold Quadrant Analysis, Scenario Analysis, Monte Carlo Concepts

Process and Documentation: Business and Functional Requirements Documentation, Process Documentation, SOPs, Data Dictionaries, Audit Ready Reporting, Compliance Reporting, Standardized Reporting Templates, Knowledge Transfer Documentation, PDF Report Generation

Soft Skills: Cross Functional Collaboration, Executive Presentations, Stakeholder Communication, Attention to Detail, Bias for Action, Growth Mindset, Ownership Mindset, High Accountability, Entrepreneurial Drive, Fast Learner

CERTIFICATIONS

- Google Data Analytics Professional Certificate | Coursera | Feb 2026
- Google Cloud Data Analytics Certificate | Coursera | Feb 2026
- Deloitte Australia Data Analytics Job Simulation | Forage | Mar 2026
- Tata GenAI Powered Data Analytics Job Simulation | Forage | Mar 2026
- Quantum Data Analytics Job Simulation | Forage | Mar 2026