Swanand Vijay Vidwans

Jersey City, New Jersey | swanandvidwans33@gmail.com | PORTFOLIO |LINKEDIN | TABLEAU| GITHUB|+1(862)373-8327

EDUCATION

Master of Science in Software Development & Engineering, Pace University, New York, NY	**GPA:3.89**	Dec 2021

Courses: Database Management Systems, Requirements Analysis, Project, and Change Management

Bachelor of Engineering in Electronics Engineering , University of Mumbai, Mumbai, India	**GPA:3.2**	May 2019

Courses: Object Oriented Programming, Microcontroller Programming, Operating Systems, Data Encryption

PROFESSIONAL EXPERIENCE

Research Assistant, Pace University, New York, NY   **Sept 2021 – December 2021**
- Assist professor with administrative tasks , organizing events, and conduct various research tasks.
- Gather research data about 'how SDLC practices can be implemented in ETL for Data & AI'.
- Analyze findings, prepare reports for presenting the objectives derived from google analytics and formulate Tableau dashboards for user engagement.

Junior Data Analyst, SITH – Solutions for IT, Mumbai, India   **Dec 2018 – Dec 2019**
- Co-ordinated with Sr. Data Analysts as a consultant to generate capacity planning reports to generate data-driven solutions reducing operation time by 40%.
- Authored SQL queries to extract KPIs from dimensions and facts tables.
- Participated in the phases of data mining such as data collection, data cleaning, developing models, validation, visualization and performed Gap analysis.
- Proposed Tableau dashboards to monitor KPIs: total tasks, completion rates, improving operator efficiency by 12%.
- Used MS Excel, Access to create reports, implement verification procedures as per client requirements.
- Converted 15+ sales & financial reports to Tableau dashboards, increasing readability for stakeholders.
- Implemented MS SQL Server data warehouse using SQL queries to create fact & dimension tables for business metrics.

ACADEMIC PROJECTS

AWS Project – Build an ETL Data Pipeline on AWS Cluster   **May 2022**
- Built a fully working scalable, reliable, and secure AWS EMR complex data pipeline (ETL) from scratch that provides support for all data stages from data collection to data analysis and visualization.
- Extracted the raw sales data into AWS S3 and created an external Hive table on top of S3 for staging table purposes.
- Performed various ETLs in Hive to store processed data into Hive managed table and connected EMR Hive to Tableau desktop.
- Visualized various KPIs in the sales data using Tableau and extracted refreshes in tableau server discussion.

Web Server log processing using Hadoop in Azure   **April 2022**
- Used Hadoop, Flume, Spark, and Hive to process the Web Server and logs dataset to glean more insights on log data.
- Created VM and connected data to VM using Putty application. Ingested log data using Flume and Processed it using Spark.
- Performed analysis on log data using Hive.

Walmart Sales Forecasting   **Mar 2022**
- Predicted sales for each department using historical data from Walmart dataset containing data of 45 Walmart stores.
- Performed basic EDA followed by univariate and bi-variate analysis. Plotted trend for each predictor and did in-depth weekly and monthly analysis.
- Prepared data for modelling and made predictions using statistical techniques. Created time series ARIMA models and compared them using multiple metrics.

Cyclistic Bike Share Analysis | Python, Data Analysis   **Sept 2021**
- Extracted, cleaned, and analyzed historic bike data of casual riders and annual members using Python libraries to gain insights.
- Summarized the analysis by supporting visualizations and key findings for the period of May through September using plotnine & geopy and found twice the number of bikes are rented compared to other months.
- Designed marketing strategies aiming to convert casual riders into members by providing introductory discounts, semi-annual memberships, and referral bonuses through mobile app.

TECHNICAL SKILLS

Programming Languages:	Python, SQL, R, VBA, HTML, CSS
Database Programming:	MySQL, MS SQL Server, SSMS, Oracle, MongoDB
Libraries:	Pandas, Matplotlib, sk-learn, Seaborn, NumPy, Beautiful Soup, geopy
Tools:	TensorFlow, SAP, Tableau, Salesforce, Google Cloud, MS Excel, MS Office, MS Access, Git, GitHub, JIRA, Google Analytics, Jupyter Notebook, SharePoint, Azure, Hadoop, MapReduce, Hive, Apache, Sqoop, Flume, Oozie
Data Science:	Combinatorics, Bayesian Inference, Distributions, Descriptive Statistics, Hypothesis Testing, Linear regression, Logistic regression, Cluster Analysis, Software Integration, Forecasting, Neural Networks Artificial Intelligence

CERTIFICATIONS

Azure Fundamental Certifications (AZ-900, DP-900), Azure Data Engineer Associate (DP-203), Google Data Analytics Professional, Tableau Associate, Excel for Business Analytics, SQL Essentials, Python for Data Science, Salesforce Ranger