

Varun Patnana

Tampa, FL | (813)-705-1997 | varunpatnana@usf.edu | [Linkedin](#)

- Currently pursuing M.S. in Business Analytics (Data Science) at the University of South Florida.
- Earned B.Tech degree in Mechanical engineering from Gitam University.
- 5.5+ years experience in sourcing, studying, transforming interpreting, and analyzing data to drive business decisions.
- Currently seeking an internship in the field of Data Engineering, Data Analytics, Data Science and Machine learning.

EDUCATION

University of South Florida, Tampa, FL

Expected Graduation date:December 2022

Masters in Business Analytics and Information Systems.

TECHNICAL SKILLS

Programming & Tools: Python, R, Tableau, PowerBI, MS Excel, Github, R-Studio, Ab-Initio, Informatica, Unix, JIL, Jenkins, HP ALM

Data Science & Machine Learning: SQL, SAS Miner, Linear Regression, Decision Trees, Neural Networks, A/B Testing, ETL, NumPy

Databases: MySQL, SQL Server, PostgreSQL, Oracle, Mainframe,

Others: PVCS, DevOps

WORK EXPERIENCE

Data Engineer, TATA Consultancy Services, India

August 2015-July 2020

ETL Development (Client: American Insurance Group)

- Developed the ETL process for insurance policy reconciling the data of Massachusetts state for Pilgrim insurance company that processes 330 Gigabytes of data everyday.
- Designed ETL process for merger between Glatfelter insurance group and AIG which processed 400 Gigabytes of data daily.
- Devised ETL process to merge the data from different data bases such as Mainframe, AS400, Oracle and DB2 to access data of multiple applications which reduced the manual work of policy validation by 20 hours every week.
- Made xml - text conversion process and the vice-versa using Ab-initio for Florida DMV data processing.
- Created ETL process using Ab-Initio to merge the data of Black Board Specialty Insurance company with AIG for 23 states.
- Book of business is constructed using Ab-Initio to send active policies to state DMV's on monthly and weekly basis which eliminated the manual intervention and cut the effort by 5 Hours every week.
- Automated data cleaning via UNIX scripting and corresponding Ab-initio job management to handle multiple xml's.
- Established ETL business process activity and status ledger using Ab-Initio for periodic policy reporting for 17 states.
- Handled the redesign of claim handling for suppressed claims for various Third-Party Administrator's.
- Optimised deployment process to deploy code through DevOps which reduce the effort by 15 hours for every deployment.
- Set up GitHub Repository to move code from PVCS Version Manager to GitHub that lowered deployment time by more than 24 hours.
- Implemented DevOps lite to implement DML's and DDL's which reduce the manual effort by 20 hours every week.

ETL Development (Client: Key Bank)

- Migrated Key Bank's on-premise datastore of 580 Giga Bytes from IBM server infrastructure to cloud based HADOOP.
- Executed critical data integration initiative via dataset parameterization of 450 Parametrized sets.
- Automated job execution for both QA/Testing and Production layers via mainframe and JIL.

Data Analyst, TATA Consultancy Services, India

Software defects prediction: (Client: Microsoft)

- Collected daily, bi-weekly, and monthly work items data from Azure DevOps using OData connector.
- Creating data models, transforming data using Ab-initio ETL tool and generating reports in **Power BI** for work items
- Optimized the testing strategy to decrease the average man-hours by 2hrs/week.
- Constructed metrics and **KPIs** for senior management to give an overview of the testing team performance.
- Created reports and dashboards for high impact and high-risk areas of the product by analyzing the feature usage data.
- Prioritized the test cases by predicting the test case failure using traditional testing metrics
- Developed a predictive model with logistic regression to detect the failed test cases which helped the client to do testing efficiently, increasing the fault test cases detection by 15%

Monitoring app behaviour from the user's perspective using analytics.

- Cleaned, pre-processed, and organized the telemetry and user survey data (1Million records) using **Python & PySpark**.
 - Curated interesting dashboards about user's activity, most used features, crashes, and engagement.
 - Analyzed the user's survey data (10000+ records) to know about user preferences to better serve them.
 - Created service **KPIs** to help the engineers to improve the app performance.
 - Created a clustering model to cluster users to know the user base and fabricated a dashboard with **Tableau** for clients.
- Impact:** Decreased the app screens loading time by 5ms and targeted users with custom features.