

# Surabhi Chavan

☎ 4699272529 ✉ surabhichavan15@gmail.com in www.linkedin.com/in/surabhi-chavan25 🏠 Dallas, TX

## EDUCATION

### University of Texas at Dallas

M.S Business Analytics, Concentration: Data Science | GPA: 3.84/ 4.0

Coursework: Database foundations, Business analysis, Advanced Statistics, Data visualization, Applied Machine Learning, Applied Econometrics and Time series analysis, Marketing and Web Analytics

Dallas, Texas

Aug '21 – Present

### University of Mumbai

B.E Electronics and Telecommunications Engineering | GPA: 8.23/10

Coursework: Statistics

Mumbai, India

Aug' 14 – Jun' 18

## TECHNICAL SKILLS

- **Programming languages:** Python, R, SQL
- **Data Science Packages/Libraries:** Scikit-Learn, Pandas, NumPy, NLTK, Matplotlib
- **Machine Learning:** Supervised learning: Regression (Linear regression, Polynomial regression, Support Vector regression, Decision Tree Regression, Random Forest Regression), Classification (Logistic Regression, K-Nearest neighbors (KNN), Support Vector machine (SVM), Naive Bayes, Decision Tree, Random Forest); Unsupervised learning: Clustering (K-means & Hierarchical clustering)
- **Time Series analysis:** Auto Regression, Moving-Average, Autoregressive-Moving-average, ARIMA
- **Natural Language Processing:** Bag of words model
- **Database:** MySQL, MongoDB
- **BI & Visualization:** Tableau
- **Web Analytics:** Google Analytics, Adobe Analytics
- **Data Analysis:** Python, R programming, Microsoft Excel
- **Statistics:** Hypothesis testing, Significance testing

## WORK EXPERIENCE

### Atos Syntel

Associate Consultant (GCM 2)

Pune, India

Mar' 19 – Jul' 21

- Conducted data analysis on incident data using Python and achieved 30% reduction in SLA breach
- Built 2 interactive dashboards and visualizations to understand the incident creation statistics in Tableau
- Developed database views using SQL by joining multiple tables which resulted in creation of 10 new reports
- Acted as a liaison between 2 internal teams and clients to review and finalize process solutions

## PROJECTS

### Greenhouse gas emissions over the world throughout the years (Python, Tableau)

Feb 2022

- Conducted data cleaning and analysis using Python on dataset containing 968 records, 30 features and formed hypothesis using the data
- Built a dynamic and interactive dashboard to create visualizations to accept/ reject the hypothesis

### Texas Environment Quality Complaints (Python, Tableau)

Feb 2022

- Performed data cleaning and analysis using Python on dataset containing 136560 records, 33 features and formed 3 hypothesis using the data
- Built a data story to visualize the hypothesis and accepted/ rejected the hypothesis based on the visualization results

### IMDB review-Sentiment Analysis (Python, Natural language processing)

Jan 2022

- Performed stop word removal using Python from training (1125 records), test data (375 records) and extracted clean data
- Conducted sentiment analysis on the test data using SVM classification algorithm and achieved accuracy of 79%

### Titanic Survival Prediction using Logistic regression (Python, Machine Learning)

Dec 2021

- Performed missing values imputation, variable and feature encoding using Python on the training (889 records) and test (418 records) records
- Predicted passenger survival by building a predictive model using logistic regression and achieved accuracy of 75.83%

### Data Modeling for Airbnb homestay activity using MySQL Workbench (SQL)

Nov 2021

- Designed database using Entity relationship diagrams and transformed the data to 3rd Normalization form
- Created schema, table and loaded 10,000 records for listing activity of Airbnb homestays in Seattle, WA

### Effect of COVID-19 vaccination in US (Tableau, Python)

Oct 2021

- Conducted data cleaning, analysis using Python on primary (15,461 records) and secondary (63,404 records) dataset
- Visualized effect of vaccination on the state wise mortality rates and confirmed cases in the USA before and after vaccination

## ACHIEVEMENT

- Ranked 4<sup>th</sup> amongst 60 teams in Data Visualization Challenge'22 conducted by University of Texas at Dallas in Spring 2022
- Received a SPOT Recognition award from Atos Syntel appreciating my work towards the project
- Led an onboarding project training for 3 new employees in Atos Syntel team