

# Pankaj Sharma

314, 305 N.Lincoln Avenue, Urbana, IL, USA

☎ (+1) 217-751-2572

✉ [pankajs2@illinois.edu](mailto:pankajs2@illinois.edu)

📱 Pankaj-UIUC

🌐 [psharma13](#)

## Education

### University of Illinois, Urbana - Champaign

Urbana-Champaign, USA

MASTER OF SCIENCE IN STATISTICS, 4.0/4.0

May 2021

- **Relevant Coursework:** Machine Learning, Applied Regression & Design, Mathematical Statistics, Statistics and Probability, Applied Bayesian Analysis, NLP, Intro to Deep Learning
- **Graduate Teaching Assistant** for IS 590, Practical Health Data Analytics, Spring 2020 at School of Information Sciences

### Indian Institute of Technology, Roorkee

Roorkee, India

BACHELOR OF TECHNOLOGY IN MECHANICAL ENG.

May 2017

- **Relevant Coursework:** Data Structures & Algorithms, Optimization, Numerical Methods, Linear Algebra, Calculus, Hypothesis Testing, Simple Linear Regression

## Experience

### Bayer

UIUC Innovation Centre, USA

DATA SCIENTIST - INTERN

May. 2020 - Jul. 2020

- Developed a clustering model in **Python** to detect similarity between environments in different field locations to optimize testing network.
- Applied algorithms like Dynamic Time Warping and Agglomerative Hierarchical clustering on weather and soil features to find analogous locations.
- Created informative data visualization dashboards using Dash, Matplotlib and Seaborn in Python.

### ZS Associates

Pune, India

DECISION ANALYTICS ASSOCIATE

Jun. 2017 - Jul. 2019

- Automated the process of cleaning and processing the raw noisy data using Python and cut down the operation time to approximately 50%
- Clustered the customers into well-defined segments and developed segment-focused marketing strategies.
- Designed and implemented an Incentive Compensation plan in Excel for a global pharmaceutical client, reduced the client's spending from 150% to 102% of the budget.
- Developed a forecasting tool using Excel and VBA incorporating multi-variable regression with seasonality, added extra functionality to account for effects of market events like new product launch and market size adjustments.

## Project Highlights

### A study of covid-19 spread and impact across counties

UIUC

R, CODE

April. 2020

- Created a predictive model to estimate deaths and cases in the upcoming week using social and demographical data across counties.
- Used Random Forest and GBM to classify high-risk counties and predicting death counts with 83% accuracy for classification and RMSE for regression ranging from 5-10.

### Walmart Sales Prediction in Stormy Weather

UIUC

R, CODE

Dec. 2019

- Predicted Sales of different items in Walmart stores during stormy weather, evaluated and compared the performance of Random Forest, GBM and Stepwise Linear Regression models, achieved a score of 0.10(top 10%) on Kaggle leaderboard.

## Additional Experiences & Skills

### Central Placement Team

IIT Roorkee, India

MANAGER

Jan. 2015 - Dec. 2016

- Led a team of 30 members and managed placements of 3000+ students in 500+ global companies and ensured a successful placement season.

### Technical Skills

PROGRAMMING & DATA ANALYTICS

- Python, R, MATLAB, C, C++, Excel, VBA, SQL, RMarkdown, HTML, PyTorch, Tensorflow, Tableau, GCP