

# Prince Kumar

## Aspiring Data Scientist

princekr301@gmail.com | +918873276380 | South Delhi

[LINKEDIN](#) | [GITHUB](#) | [HACKERRANK](#) | [BLOG](#)

## EDUCATION

### Diploma in Data Science

| Madrid Software | (Mar 2022 - 2023) |

#### Skills Acquired:

- Python | R-Programming | SQL | Advance Excel
- Power BI | Tableau | Statistics
- Machine Learning | Hypothesis Testing |
- Natural Language Processing (NLP) |
- Statistical Analysis | Predictive Analysis |
- Data Mining | Data Storytelling | ETL Process |
- Data Analysis Tools | Data Visualization Tools |

### Bachelor of Science - Mathematics

| Lalit Narayan Mithila University | (2017 – 2020) |

| Percentage: 58.4 %

#### Skills Acquired:

- Mathematical Problem Solving |
- Statistics | Probability | Linear Algebra |
- Discrete Mathematics | Differential Equations |

### 12H SCIENCE

RBS COLLAGE TEYAI BEGUSARAI

Jun 2015 - Feb 17 || Percentage: 54.4%

### 10TH

BN + 2 HIGH SCHOOL TEYAI, BEGUSARAI

MAR 2014 - MAR 2015 || Percentage: 60.2

## KEY SKILLS

- ☐ Data Analysis & Visualization.
- ☐ Descriptive & Inferential Statistics.
- ☐ Machine Learning.
- ☐ Natural Language Processing (NLP).
- ☐ Descriptive & Diagnostic Analysis.
- ☐ Regression & Classification.
- ☐ Proficient in Excel & Advance Excel.
- ☐ Proficient in SQL Scripting & Windows function.
- ☐ Proficient in Report Building.

## PROFILE SUMMARY

- ◆ Looking for a challenging role in a reputable organization to utilize my technical skills for the growth of the organization as well as to enhance my knowledge about new emerging trends in the IT sector
- ◆ Over 1+ years of experience as a Data Science trainee, with a well-rounded understanding of data science principles.
- ◆ Interest in analyzing raw data and statistics, which brings excellent communication abilities with hands on experience in building machine learning models and creating dashboards with a story to tell.

## WORK EXPERIENCE

### MADRID SOFTWARE | DATA SCIENCE INTERNSHIP DELHI (MAR 2022 – AUG 2023)

#### Role & Responsibilities

- Experienced in **Supervised & Unsupervised** Machine Learning Projects.
- Collect, clean, and preprocess data from diverse sources for analysis.
- Report Building using **Power BI & Tableau**.
- Data Manipulation using **SQL & Pandas**.
- Data Cleaning & Data Modeling using **Excel & Pandas**.

### Data Science Internship

|| Oasis Infobyte || (Aug 2023 – Sep 2023) ||

#### Roles & Responsibilities:

- ◆ Conducted comprehensive **data cleaning, transformation and validation** processes to prepare datasets for analysis.
- ◆ Created visually appealing **dashboards and reports** to present findings and communicate data-driven insights.
- ◆ Developed email spam detection model using text classification and natural language processing techniques.
- ◆ Hands-on experience with various tools, technologies and methods

## MY PROJECTS

### Predicting Car's Ex-Showroom Sales Price.

| Domain: **Automotive Industry** | Type: **Regression project** |

- ◆ **Tech Stack:** Python, Pandas, NumPy, Matplotlib, Seaborn, SciPy, Machine Learning, Sklearn & Jupiter Notebook.
- ◆ **Objective:** Develop a robust **predictive model** to estimate car ex-showroom prices based on a dataset containing **140 car attributes**.
- ◆ **Methodology:** Utilize **data mining techniques** and supervised **machine learning algorithms** to analyze the dataset, identify **key factors** impacting price variations, and perform thorough feature selection for accurate price prediction.
- ◆ **Results:** The developed predictive model achieved **high accuracy** and **low error rates**, with an **R-squared of 97%**, **Adjusted R-squared of 96%** and **RMSE of 0.0001506**. These results validate its effectiveness in estimating car ex-showroom prices

# TECHNICAL SKILLS

- ◆ **Programming:** Python | R | SQL
- ◆ **Databases:** MySQL | PostgreSQL
- ◆ **Report Building:** Power BI | Tableau
- ◆ **Python Libraries:** NumPy | Pandas  
Matplotlib | Seaborn | Sklearn | SciPy
- ◆ **R Packages:** GGplot2 | DPLYR
- ◆ **Tools:** MS Excel | Google Sheets | MYSQL -  
Workbench | Sublime | Jupiter Notebook | GitHub.
- ◆ **Soft Skills:** Communication, Collaboration  
| Adaptability, Problem Solving | Time-  
Management, Critical Thinking

## ACHIEVEMENTS

☐ Passed: IBM Machine Learning Essentials Quiz

- 🏆 HackerRank 5th Star Badge in Python  
I have got 93k rank out of 20 million ranks
- 🏆 HackerRank 5th Star Badge in SQL  
I have got 98k rank out of 20 million ranks

## CERTIFICATION

**Advance Diploma in Computer Application**  
**Lakshi Institute of Advance Technologies Private limited**  
Bihar || **Grade :- A+ ||** (April 2015 to May 2016)  
1. Microsoft Excel 2. MS Word 3. MS Power Point  
4. Adobe Photoshop 5. Basic part of computer

### PYTHON AND HIS LIBRARY

**Data Flair with 90% test Marks**  
**NumPy, Pandas, Matplotlib, Seaborn, SciPy**

### Power BI

**Skill Nation Certificate**  
Dashboard, Dax Function, M Language and  
Power Query

## LANGUAGE

1. HINDI 2. ENGLISH

## Hobbies

Cricket, Travelling, Creating Dashboard in power BI



## MOVIE RECOMMENDATION PROJECT

**Domain:** Movie Industry || **Type:** NLP ||  
**Python, Machine Learning, Statistics, sklearn, nltk**

- Building a **movie recommendation system** involves a combination of **data processing, feature engineering, model training, and evaluation**. By leveraging machine learning techniques, these systems can provide users **with personalized movie suggestions**, enhancing their movie-watching experience and facilitating the discovery of new movies they might enjoy.
- For More information, please visit my **GitHub to refer my notebook**



## STOCK MARKET ANALYSIS

**Domain:** Stock Sales || **Type:** Data Analysis project  
**Power Bi, Google studio, Excel**

- I have created **Power Bi Dashboard** Here I have Analyzed that which **industry Grow** more and who is coming Down in Market Price with the Reason entire insights is very useful.
- In industry has **stock** price high Market price Last one Year price and current Price. **Extracted Meaningful Insights**
- For More information, please visit my GitHub to refer my notebook

## SMS SPAM DETECTOR PROJECT

**Python, statistics, machine Learning, nltk**

- Here I have Build a **Machine Learning perfect Model** with **96%** Accuracy on **test data** with minimum error
- In this project I have classified the **spam messages** and **Ham messages** with the help **supervised learning classification**.
- With the help **Natural language toolkit** I found the top 30 spam word that is used in spam messages and top 30 Ham word that is used in Normal messages.
- For more information, please watch my Notebook that's given on **GitHub**

## Nike Sales Analysis Project

**Power BI, Excel, Python and MySQL**

- I have created **power Bi Dashboard of Nike sales Analysis** and fond lots of Insights
- Same work done in Python and visualized with matplotlib and seaborn and perform on **MySQL**.
- Please visit to see the complete project of Nike sales given Link or LinkedIn also.