

# Vivek Kumar

<https://sites.google.com/view/vivekdata>

+91 7909089293, +91 8441870339

Vivekrohtas06@gmail.com

Analytically minded Data Scientist with more than 6 years of experience. Worked extensively in Dashboarding, Exploratory Data analysis, Machine Learning, SQL & Python. Looking for an opportunity to leverage my skills to bring Actionable Insights from Data.

## Experience

### 1. Upstox

#### Data Scientist (Mar-2024 - Present)

- Design & Analyse A/B Tests, Build ML Models, Analyse Data and Maintain QuickSight/Tableau Dashboards.
- Automated Analysis & Dashboards to Save 40+. Hours of Monthly bandwidth.

### 2. Paytm

#### Data Scientist (Jan – 2023 to Feb 2024)

- Clean, Explore, Analyse data using SQL & Python libraries like Pandas.
- Build Machine learning Models for different business problems.

#### Major Impacts:

Built ML model to reduce leakages in Install to signup funnel which resulted in 25% Increase in signups & onboard 100 thousand extra customers annually.

### 3. PWC SDC, Bangalore

#### Data Science Associate - II (Nov - 2021 to Jan - 2023)

- Collect data and bring Insights from it to solve complex business problems.

#### Major Impact:

Developed efficient, reusable Python web scraper, Alteryx & PowerBI frameworks, saving 50+ hours per project and ensuring timely delivery with enhanced billability.

### 4. Infosys LTD., Hyderabad

#### Digital Specialist Engineer - Data Analyst (Jul - 2019 to Nov - 2021)

- Proficiently write SQL & Python code to migrate large volume data.

## Key Projects

- Customer Retention Model** (Skills used- ML, Data Cleaning, XGBoost)
  - i) Built Long-term & short term retention model which was used to predict severity of drop in Txn. & App activity of customers.
  - ii) pilot was done on 30% customers and a 20% difference was observed between pilot & control after intervention by marketing team
- Install to signup Model** (Skills used-Data Cleaning, Logistic Regression)
  - i) Built Logistic Regression model to predict Customer's signup after App Installation. Marketing teams used the model to design campaigns which helped reduce leakages in the install to signup funnel.
- Location Based Bank Linking** (Skills used - Pandas, Excel, Data Analysis)
  - i) Built a bank predictor analytics framework using location and recent UPI registrations, plugging leaks in UPI registration funnel and achieving a 12% lift over control. recommends accurate 3-5 banks for ~98% customers, particularly impactful in Tier-2 and Tier-3 cities, enhancing the UPI registration experience.
- Affluence Score / CLTV** (Skills used- ML, Data Cleaning, Seaborn, Clustering)
  - i) Build Model to assign affluence score of every Customer. CLTV score was used for New customers to determine their Affluence. Score was used as an alternate to Income documents in Insurance registration. This Uplifted sell of Insurance easing documentation. It also helped identify potential customers.

## Skills

- SQL
- Python – Pandas, Seaborn, SkLearn etc.
- Machine Learning
- Tableau
- PySpark
- Hypothesis Testing
- AWS – S3, QuickSight, Athena etc.

## Education

### 1. Bachelor of Engineering

University Institute of Engr. & Tech. (UIET)  
Panjab University, Chandigarh  
Bachelor of Engineering  
Electronics & Communication Engineering  
2015 – 2019

### 2. Scaler Academy

Data Science & Machine Learning  
2021-2022

## Work Portfolio

<https://sites.google.com/view/vivekdata>

## Certifications

1. HackerRank SQL Advanced
2. Google Data Analytics
3. Python Data Structures - Coursera
4. Machine learning – Coursera
5. Alteryx Core Certification
6. Statistical Learning – Stanford Online
7. Financial Accounting & Analysis - IIMB

## Academic Achievements

- City topper of National Science Olympiad(Thrice)

## Volunteering

- Fund Raising at MAD
- Art Of Living