

# KOUSHIK KHAN

+91 750 192 3251 ◇ Bangalore, India

✉ koushik.khan.ds@gmail.com ◇ in linkedin.com/in/koushikkhan ◇ koushikkhan.github.io

## OBJECTIVE

A seasoned data scientist with close to nine years of industry expertise in developing data-driven solutions. Skilled in applying data wrangling techniques, statistical modeling & machine learning algorithms on large datasets. Having proven track record of aligning data insights with business objectives, collaborating cross-functionally, and innovating with new modeling techniques. Seeking to leverage skills to address diverse business challenges and drive actionable resolutions in a dynamic professional environment.

## PROFESSIONAL EXPERIENCE

### Principal Data Scientist

S&P Global Inc.

Aug 2021 - Till Date

Bangalore, India

- Directed a team in building essential devops pipelines and APIs for the core functionalities of maritime supply chain products.
- Boosted team efficiency by 70% with the development of automation tools for data extraction, processing and reporting.
- Migrated legacy code to AWS that boosted product performance, and led to attracting new clients with 15% higher revenue.
- Reduced trade financing risks for banking clients by developing ML models to detect dual-use goods in bill-of-lading descriptions.
- Leveraged analytical skills to develop predictive models for forecasting ship emissions and its environmental impacts.

Methods & Tech stacks used: Machine Learning, Python (PySpark), SQL, PowerBI, DevOps, AWS

### Data Scientist

IBM India Pvt. Ltd.

Oct 2016 - Jul 2021

Bangalore, India

- Orchestrated an automated system streamlining tech support ticket analysis, significantly improving team efficiency.
- Accomplished an average 10% sales growth, across markets, using Market-Mix-Model derived from sales and ad data.
- Envisioned, built and deployed a banking FAQ retrieval system for clients that outperformed existing term-based models.
- Was instrumental in crafting text-cleaning and text-labelling processes for IBM Watson text classification API.

Methods & Tech stacks used: Text Mining with ML, Regression (MMM), Latent Semantic Indexing, Python

### Data Analyst

BCausE Enterprise Pvt. Ltd.

Oct 2015 - Aug 2016

Noida, India

- Collaborated on identifying the consumer base for an in-house product by developing a customer segmentation model.
- Facilitated my team by developing automation tools for periodically scraping data from web pages.

Methods & Tech stacks used: Clustering, SQL, Python (BeautifulSoup)

## KEY SKILLS

- **Programming Languages:** Python, R and SQL
- **Tools & Technologies:** PySpark, AWS (EC2, S3, Athena, Lambda), Docker, REST API, CI/CD, DevOps, Linux, Power BI
- **Analytics:** Statistical Modeling (Regression, Classification, Time Series), Machine Learning, Text Mining & NLP

## EDUCATION

M.Sc, Statistics: (CGPA 8.43), Visva Bharati University, West Bengal, India

2013 - 2015

B.Sc, Statistics: (CGPA 8.26), Visva Bharati University, West Bengal, India

2010 - 2013

## CERTIFICATIONS

- Microsoft Certified: Azure Fundamentals
- Microsoft Certified: Azure Data Scientist Associate
- Microsoft Certified: Azure AI Fundamentals

## EXTRA-CURRICULAR ACTIVITIES

- Authoring technical articles for my blog as well as other platforms.
- Exploring and learning new methodologies, tools and technologies to keep myself up-to-date.