

Krishna Barfiwala

New York, NY | (669)-243-7714 | krishnabarfiwala2@gmail.com | [LinkedIn](#)

Data Analyst with **4+ years** of experience and dual master's degrees in Statistics and Data Analytics Engineering and optimizing business processes through data-driven insights. Proficient in **SQL, Python, BI tools, and cloud** platforms, combining **statistical rigor** with engineering best practices to deliver **production-grade analytics solutions**.

SKILLS

Languages: Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), SQL, R (dplyr, ggplot2, Shiny), Excel (VLOOKUP, Macros)

Visualization & Analytics: Tableau, Power BI, AWS QuickSight, Alteryx, Qlik Sense

Cloud Platforms: AWS (S3, Glue, Redshift, EC2), GCP (BigQuery, Dataflow), Azure (Data Factory, Synapse Analytics, Blob Storage)

Databases & Data Warehousing: Snowflake, Athena, PostgreSQL, MySQL, MS SQL Server, Hadoop

Data Engineering & Big Data: Apache Spark, Databricks, Apache Airflow, ETL/ELT Pipelines, Parquet/Avro optimization, Git

Statistical & ML Skills: Machine Learning (K-means, Regression), Hypothesis Testing, Time Series Analysis, Multivariate Analysis

EXPERIENCE

Data Analyst

September 2024 - October 2025

Gregory Welteroth Advertising, Pennsylvania, USA

- Designed 7+ **Tableau dashboards** for 5+ major clients, surfacing advertising campaign **KPI insights** (CTR, CPA, conversion rates, ROAS) for monthly performance reviews for executive leadership
- Engineered automated **ETL pipelines** using **Alteryx** (Filter, Join, Formula tools) to process data from **AWS S3**, transform survey, sales datasets, and load results back to **AWS S3** for **Tableau visualization**, saving 10 hours weekly
- Implemented a **Python validation** framework utilizing regex pattern matching and anomaly detection algorithms to resolve multi-source data discrepancies, reducing data reconciliation time by 75%
- Leveraged **agentic AI workflows** (Claude/GPT-4) and **prompt engineering** to accelerate ad-hoc data cleaning and exploratory analysis turnaround time from ~4 hours to 1 hour while optimizing **SQL queries** (indexing, CTEs, join optimization)
- Built **A/B testing framework** in Python with hypothesis testing and confidence interval analysis to validate ad campaign efficiency, achieving statistical significance ($p < 0.05$) and influencing \$500k quarterly budget optimization
- Segmented 100K+ survey respondents using **unsupervised machine learning** (K-means) to identify buyer personas (impulse buyers, luxury seekers), driving \$2M+ **marketing strategies** for fintech clients
- Collaborated with teams to define KPI frameworks, reducing dashboard development time by 30% via standardized models

Data Analyst Intern

May 2023 - September 2023

StudentsLanding Inc., Boston, USA

- Architected end-to-end **AWS S3-to-Snowflake** ETL pipeline by designing star schema for 20K+ real estate properties, implementing data quality checks and materialized views that optimized query performance by 30%
- Analyzed 2M+ customer and listing records using **SQL** (CTEs, joins, aggregations) in **Snowflake** to identify high-value properties and churn drivers (late payments, non-renewals), **reducing churn** by 28% and driving **\$1M revenue increase**
- Built **Power BI dashboard** tracking 10+ tenant **KPIs** (payment trends, occupancy) for proactive outreach to at-risk accounts
- Delivered **weekly reports** to marketing/operations teams, informing **retention strategies** and enabling same-day **decisions**

Data Analyst

January 2021 - August 2022

Kantar Worldpanel, Mumbai, India

- Developed **Excel VBA macros** to automate **segmentation** and **classification** of consumer goods datasets across 50+ product categories, decreasing manual processing from 8 hours to 10 minutes
- Lowered **ETL cycle duration** from ~5 hours to ~2 hours by implementing **AWS Glue** pipelines to process high-volume FMCG product, demographic, household panel datasets stored in S3 and standardize data from 10+ disparate source systems
- Constructed a scalable **AWS Redshift data warehouse** with reusable **data models**, standardizing KPI logic across business units and reducing query latency by 30%
- Enhanced brand performance reporting efficiency by 75% via **Parquet** optimization and **partitioning**, supporting real-time dashboard updates in **Amazon QuickSight** for advertising effectiveness
- Partnered with market research teams across 3 business units to define **KPI frameworks** and data requirements, **reducing** report development time from 2 weeks to 5 days and **eliminating** rework cycles

EDUCATION

Northeastern University, MS in Data Analytics Engineering, Boston, USA

University of Mumbai, MSc in Statistics, Mumbai, India

CERTIFICATIONS & AWARDS

- Alteryx Designer Core
- Alteryx Foundation Micro-Credential
- 2nd prize in Big Data Championship