**Data Science Analyst- job post**

[PepsiCo](https://in.indeed.com/cmp/PepsiCo?campaignid=mobvjcmp&from=mobviewjob&tk=1ho2r0gkbkjgk801&fromjk=aa03c0708a874f94)

[19,735 reviews](https://in.indeed.com/cmp/PepsiCo/reviews?campaignid=mobvjcmp&cmpratingc=mobviewjob&from=mobviewjob&tk=1ho2r0gkbkjgk801&fromjk=aa03c0708a874f94&jt=Data+Science+Analyst)

Hyderabad, Telangana•Remote

[PepsiCo](https://in.indeed.com/cmp/PepsiCo?campaignid=mobvjcmp&from=mobviewjob&tk=1ho2r0gkbkjgk801&fromjk=aa03c0708a874f94)

Hyderabad, Telangana•Remote

Apply now

**Full job description**

Overview:

**Main Purpose:**

The Data Science will work in developing Machine Learning (ML) and Artificial Intelligence (AI) projects. Specific scope of this role is to develop ML solution in support of ML/AI projects using big analytics toolsets in a CI/CD environment. Analytics toolsets may include DS tools/Spark/Databricks, and other technologies offered by Microsoft Azure or open-source toolsets. This role will also help automate the end-to-end cycle with Azure Machine Learning Services and Pipelines.

You will be part of a collaborative interdisciplinary team around data, where you will be responsible of our continuous delivery of statistical/ML models. You will work closely with process owners, product owners and final business users. This will provide you the correct visibility and understanding of criticality of your developments.

Please note that this role will be based ONLY in India. The role does not involve any movement to other PepsiCo offices outside India in future.

Responsibilities:

* Delivery of key Advanced Analytics/Data Science projects within time and budget, particularly around DevOps/MLOps and Machine Learning models in scope
* Collaborate with data engineers and ML engineers to understand data and models and leverage various advanced analytics capabilities
* Ensure on time and on budget delivery which satisfies project requirements, while adhering to enterprise architecture standards
* Use big data technologies to help process data and build scaled data pipelines (batch to real time)
* Automate the end-to-end ML lifecycle with Azure Machine Learning and Azure Pipelines
* Setup cloud alerts, monitors, dashboards, and logging and troubleshoot machine learning infrastructure
* Automate ML models deployments

Qualifications:

* 5+ years of overall experience that includes at least 4+ years of hands-on work experience Data Science / Machine learning
* Minimum 4+ year of SQL experience
* Minimum 2+ years of Python and Pyspark experience
* Experience in DevOps and 2+ yrs in Machine Learning (ML) with hands-on experience with one or more cloud service providers AWS, GCP, (Azure preferred) is preferred
* BE/B.Tech in Computer Science, Maths, technical fields.
* Stakeholder engagement-BU, Vendors.

**Skills, Abilities, Knowledge:**

* Data Science – Hands on experience and strong knowledge of building machine learning models – supervised and unsupervised models. Knowledge of Time series/Demand Forecast models is a plus
* Programming Skills – Hands-on experience in statistical programming languages like Python, Pyspark and database query languages like SQL
* Statistics – Good applied statistical skills, including knowledge of statistical tests, distributions, regression, maximum likelihood estimators
* Cloud (Azure) – Experience in Databricks and ADF is desirable
* Familiarity with Spark, Hive, Pig is an added advantage
* Model deployment experience will be a plus
* Experience with version control systems like GitHub and CI/CD tools
* Experience is Exploratory data Analysis
* Knowledge of ML Ops / DevOps and deploying ML models is required
* Experience using MLFlow, Kubeflow etc. will be preferred
* Experience executing and contributing to ML OPS automation infrastructure is good to have
* Exceptional analytical and problem-solving skills
* Experience building statistical models in the Commercial, Net revenue Management or Supply chain space is a plus

**Differentiating Competencies Required**

* Ability to work with virtual teams (remote work locations); collaborate with technical resources (employees and contractors) based in multiple locations across geographies
* Participate in technical discussions, driving clarity of complex issues/requirements to build robust solutions
* Strong communication skills to meet with business, understand sometimes ambiguous, needs, and translate to clear, aligned requirements
* Able to work independently with business partners to understand requirements quickly, perform analysis and lead the design review sessions
* Highly influential and having the ability to educate challenging stakeholders on the role of data and its purpose in the business
* Places the user in the centre of decision making
* Teams up and collaborates for speed, agility, and innovation
* Experience with and embraces agile methodologies
* Strong negotiation and decision-making skill
* Experience managing and working with globally distributed teams