Notification of the Requirement – Data Analytics Unit

Government of Andhra Pradesh invites applications for the posts of following positions:

Designation and Requirement

Project Manager – 01 Position
Data Analyst – 01 Position
Data Engineer – 01 Position
Data Scientist – 01 Position

Location: Vijayawada, Andhra Pradesh

Terms of Employment: The above positions are contractual and the period of engagement is initially for two years, extendable for another year based on requirement and performance.

Introduction:

In order to enhance the efficiency and effectiveness in Public Service Delivery, Expenditure optimisation and revenue augmentation of Govt. of Andhra Pradesh, it is decided to establish a Data Analytics Unit within the Finance Department of Andhra Pradesh, which would be tasked to undertake statistical analysis of the transactional data, generate robust models and provide actionable and relevant decision support information. The Data analytics unit shall support individual departments of the state government in generating models bases on the hypothesis formulated by the department and identifying the pattern in the data that point to systematic manipulation.

Therefore, applications are invited from the eligible candidates for the following positions in the Data Analytics unit.

1. **Job Description for Project Manager:**

The role holder is to conduct statistical analysis across multiple government departments. The broad outline of the work includes:

- Maintain strong relationships with senior bureaucrats and mid-level officers and translate their requirements for the technical team members.
- Undertake complete ownership of the Data Analytics Unit (a team of Data Scientists, Analysts, and Data Engineers) and bring-in analytical and problem-solving skills needed to manage multiple factors on a project simultaneously.
- Prepare hypotheses that must be tested to arrive at actionable insights.
- Ensure project activities are on track and support the team in debottlenecking issues.

- Ensure load balancing between the team members and optimum utilization of the resources.

The **Project Manager** will also assist in formulating and implementing best practices for dealing with data at various stages. An indicative set of activities that could be performed is listed below.

- Work with department stakeholders to understand their needs and objectives for each project and serve as the primary point of contact for all data analytics needs
- Translate the analytics needs into project requirements for the technical team members.
- Rapidly develop domain-specific expertise for various government departments.
- Design Data Privacy Protocols and flexible storage protocols.
- Oversee all aspects of each project from start to finish, ensuring that all deliverables are met according to the schedule.
- Develop detailed project plans that outline all deliverables, milestones, and timelines for each project
- Manage the day-to-day operations of data analytics projects, including work allocation, team coordination, and communications with multiple departments and stakeholders
- Develop a strategy for implementing new tools and technologies within the department for analytics purposes
- Conduct regular check-ins with project team members to ensure progress is being made, and identify any potential roadblocks and potential solutions
- Communicate project status and updates to department officials and Senior bureaucrats.
- Prepare reports and presentations on project findings for internal and external audiences
- Continuously look for ways to improve processes and optimize outcomes.

Qualifications:

Applicants holding a bachelor's degree or higher in any field are encouraged to apply. Coursework involving quantitative training, computer programming, statistics, economics, public policy, or data analysis would be given a higher weightage.

Work Experience

- 8+ years of experience in data analytics, business intelligence, project management, tech product manager, or development/business consulting.
- Experience of working with Government officials is highly preferred.
- Experience in developing public policy frameworks.

Skills:

- Excellent client-facing and internal communication skills
- The ability to interact with high-level bureaucrats and coordinate across government departments.
- The ability to structure and design a series of technical and statistical tasks to complete projects.
- Solid organizational skills including attention to detail and multi-tasking skills
- Strong working knowledge of Microsoft Office
- Strong logical reasoning, analytical and problem-solving abilities
- Basic technology and data analysis skills
- Familiarity and interest in one of the following is preferred: data strategy, data quality, business intelligence (BI), data governance, management information (MI), analytics (embedded), visualization and reporting.

Desired Profile:

- Strong interpersonal skills. The ability to work with team members from a non-technical background would be a positive.
- Excellent communication skills and ability to articulate a complex concept with thoughtful, actionable recommendations.
- Strong interest in improving the data capabilities of government systems.
- Exceptional writing/documentation/presentation skills
- Strong analytical and communication skills
- Good interpersonal skills
- Ability to work with strict deadlines
- Deep passion to make an impact on public systems and society

2. **Job Description for Data Analyst**

The role holder is to conduct statistical analysis across multiple government departments. The broad outline of the work includes:

- Interpreting data and analyzing results using statistical techniques
- Developing and implementing data analyses, data collection systems, and other strategies that optimize statistical efficiency and quality
- Acquiring data from primary or secondary data sources and maintaining databases.
- Data analysis in the programming language (R/Python) of a range of datasets, including manipulating datasets, merging data, basic tests of statistical significance

The **Data Analyst** will assist in formulating and implementing best practices for dealing with data at various stages. An indicative set of activities that could be performed is listed below.

- Using automated tools to extract data from primary and secondary sources
- Removing corrupted data and fixing coding errors and related problems
- Developing and maintaining databases, data systems – reorganizing data in a readable format
- Performing analysis to assess quality and meaning of data
- Filter and "clean" data by reviewing computer reports, printouts, and performance indicators to locate and correct code problems
- Using statistical tools to identify, analyze, and interpret patterns and trends in complex data sets that could be helpful for the diagnosis and prediction.
- Assigning numerical value to essential KPIs so that the performance can be assessed and compared over periods of time.
- Analyzing trends that impact both the departments and the state
- Preparing reports for the officials stating trends, patterns, and predictions using relevant data
- Working with programmers, engineers, and management heads to identify process improvement opportunities, propose system modifications, and devise data governance strategies
- Preparing final analysis reports for the stakeholders to understand the data-analysis steps, enabling them to take important decisions based on various facts and trends

Qualifications:

Applicants holding a bachelor's degree or higher in any field are encouraged to apply. Coursework involving training in Computer Science, Statistics, Informatics, Information Systems or another quantitative field would be given a higher weightage

Work Experience

- At least 2+ of work experience (including full-time or part-time internships) in a similar role
- Prior experience working with large datasets is preferable

Skills

- Strong mathematical skills to help collect, measure, organize and analyze data
- Knowledge of more than one programming language like SQL, Oracle, R, MATLAB, and Python.
- Technical proficiency in database design development, data models, techniques for data mining, and segmentation.
- Experience in handling reporting packages like Business Objects, programming (Javascript, XML, or ETL frameworks), databases
- Proficiency in statistics and statistical packages like Excel, SPSS, SAS to be used for data set analyzing
- Knowledge of data processing platforms like Hadoop and Apache Spark is a positive.
- Knowledge of data visualization software like Tableau and QlikView is preferred
- Knowledge creating and applying the most accurate algorithms to datasets to find solutions
- Adept at queries, writing reports, and making presentations.
- Problem-solving skills.
- Accuracy and attention to detail.

Desired Profile:

- Strong interpersonal skills. The ability to work with team members from a non-technical background would be a positive.
- Strong interest in developing statistical capabilities in public policy.
- Excellent communication skills and ability to articulate a complex concept with thoughtful, actionable recommendations.
- Exceptional writing/documentation/presentation skills

3. Job Description – Data Engineer:

The role holder is to conduct statistical analysis across multiple government departments. The broad outline of the work includes:

- Regular scraping of data from government databases, which is to be linked
- Creating pipelines to gather and store the data, setting up API linkages
- Integrate the databases using common identifiers.
- Development of databases and associated protocols for efficient storage and processing of data
- Assisting business analysts and data scientists with their data needs

The **Data Engineer** will assist in formulating and implementing best practices for dealing with data at various stages. An indicative set of activities that could be performed is listed below.

- Collection
    - Create and maintain optimal data pipeline architecture
    - Develop and deploy APIs for data linkages
    - Identify, design, and implement internal process improvements: automating manual processes, optimizing data delivery, re-designing infrastructure for greater scalability, etc
    - Build the infrastructure required for optimal extraction, transformation, and loading of data from a wide variety of data sources using SQL, python and 'big data' technologies
    - Work with stakeholders to assist with data-related technical issues and support their data infrastructure needs
    - Scraping sites that employ various data dropdown and data entry fields
- Organization
    - Assemble large, complex data sets that meet functional/non-functional business requirements.
    - Create standardized datasets out of the data from multiple government departments.
    - Integrate multiple datasets using common identifiers
    - Develop and enforce metadata protocols.
    - Ensure data consistency in various databases.
    - Develop version control protocol for data and create data logs for back-tracking

- Cleaning
    - Set standards for rigorous data quality checks
    - Develop semi-automated data-cleaning mechanisms
- Storage
    - Explore various solutions for the storage of data and metadata
    - Assist team members in ensuring all data is properly stored, adhering to all data privacy protocols
- Analysis
    - Build analytics tools that utilize the data pipeline to provide actionable insights into various KPIs
    - Create data tools for analysts and data scientist team members to assist them in analytics.
- Visualization/Presentations
    - Develop or assist in making presentations to those with less technical backgrounds

Qualifications:

Applicants holding a bachelor's degree or higher in any field related to Computer Science, Statistics, Informatics, Information Systems or another quantitative field would be given a higher weightage.

Work Experience

- At least 5+ years of work experience in Data Engineer role.
- Prior experience working in a policy setting and government data
- Significant coursework in data architecture and database structures, as well as basic statistical inference
- Coursework in economics, data mining, and machine learning is desirable

Skills

- Experience with one or more general-purpose programming languages (C++, Python, Ruby).
- Experience with big data tools: Hadoop, Spark, Kafka, etc.
- Strong SQL/PLSQL knowledge, data-modeling
- Experience with data pipeline and workflow management tools: Azkaban, Luigi, Airflow, etc.

- Experience with stream-processing systems: Storm, Spark-Streaming, etc.
- Experience with object-oriented/object function scripting languages: Python, Java, C++, Scala, etc.

Desired Profile:

- Strong interpersonal skills. The ability to work with team members from a non-technical background would be a positive.
- Excellent communication skills and ability to articulate a complex concept with thoughtful, actionable recommendations.
- Strong interest in developing statistical capabilities in public policy.
- Exceptional writing/documentation/presentation skills

4. **Job Description – Data Science.**

The role holder is to conduct statistical analysis across multiple government departments. The broad outline of the work includes:

- Perform descriptive, predictive, and prescriptive analytics
- Basic summary statistics of the data collated
- Develop, integrate, manage, and monitor statistical / machine learning models to assist in various government functions
- Setting up monitoring and dashboarding infrastructure.

The **Data Scientist** will assist in formulating and implementing best practices for dealing with data at various stages. An indicative set of activities that could be performed is listed below.

- Collection
    - Develop an automated system for real-time data pipeline and batch processing
    - Scrap sites that employ various data dropdown and data entry fields
    - Identify valuable data sources and automate collection processes
    - Assist in establishing links across datasets with common identifiers
- Organization
    - Develop and enforce metadata protocols

- Cleaning
    - Undertake pre-processing of structured and unstructured data
- Analysis
    - Analyze large amounts of information to discover trends and patterns
    - Develop, integrate, manage, and monitor predictive and machine learning/statistical model
    - Combine models through ensemble modeling
    - Develop a suite of automated tests for machine learning microservices
    - Promote best practices within the team to improve quality, reduce risk, increase security, promote efficiency, and minimize costs of machine learning models.
- Visualization/Presentations
    - Present information using data visualization techniques
    - Propose solutions and strategies to business challenges
    - Develop or assist in making presentations to those with less technical backgrounds

Qualifications:

Applicants holding a bachelor's degree or higher in any field are encouraged to apply. Coursework involving quantitative training including computer programming, statistics, mathematics, economics, or data analysis would be given a higher weightage.

Work Experience

- 4+ years of experience in working on Data Science Projects in a statistical setting.
- Demonstrated experience working across multiple databases and understanding data architecture.
- Prior experience working with government databases and public policy frameworks would be a positive.

Necessary Skills

- Hands-on experience with Python and R.
- A strong understanding of statistical modeling/machine learning/data mining.
- The candidate must be comfortable with tests of statistical significance, categorical and regression analysis.

- Proficiency in Python and AWS machine-learning packages.
- Experience in a cloud environment for software development and deployment
- Proficiency in Data Visualization and Data Wrangling
- Knowledge of MS Office, especially Excel and PowerPoint is desired.
- Knowledge of SQL; experience using business intelligence tools (e.g. Tableau) and data frameworks (e.g. Hadoop) is an asset.
- Ability to document protocols and statistical codes.

Secondary Skills

- Familiarity with more than one cloud computing environment (AWS, Azure, GCP)
- Working experience with Tensorflow, Pytorch, Scikitlearn, etc.
- Experience with Python web framework, e.g. Flask, Django, or FastAPI is a plus
- Familiarity with distributed data processing tools like Spark, Snowflake

Desired Profile:

- Strong interpersonal skills. The ability to work with team members from a non-technical background would be a positive.
- Excellent communication skills and ability to articulate a complex concept with thoughtful, actionable recommendations.
- Strong interest in developing statistical capabilities in public policy.

Application Process:

- Curriculum Vitae with Covering Letter describing your interest and suitability for the position (Maximum of 3 Pages)
- Description of one or more previous projects which you have worked on relevant to the position, not exceeding two pages. Describe the objective of the project, your role (if you were working in a team), and the skills or technologies you used.

Applicants will be shortlisted based on the merit of their application. Shortlisted applicants will be invited for an interview. The interview slots will be scheduled in consultation with the shortlisted applicants and the availability of the interview panel members.

The Government of Andhra Pradesh reserves the right to defer the process of filling the positions if it does not find suitable candidates during the selection process.

How to Apply

- Interested candidates can e-mail a Cover Letter, describing their suitability along with their CV, Covering Letter and Projects Description to **careers@apcfss.in**. The Subject Line shall be as follows: **'Application for DAU Positions – <Position Name>'**

Please feel free to send your queries, to **careers@apcfss.in**. Generally, all queries will be answered within two working days.

All the above three documents (Curriculum Vitae, Cover Letter and Project Description, in the same order) should be uploaded separately.

Timelines:

The tentative timelines are as follows, which may change subject to administrative exigencies and number of applications received:

1. Deadline for Receipt of Applications: 10 days from the date issue of the advertisement
2. Interview Call for Shortlisted Candidates: Within 7 days from the last date of receipt of application.
3. Interviews for Shortlisted Candidates: Within 14 days from the last date of receipt of application.
4. Announcement of Selected Candidates: Within 21 days from the last date receipt of application.
5. Joining Date: Within one or two months from the issue of letter of selection.