

**Terms of References (TORs)**  
**Data Scientist**  
**Federal Board of Revenue**

**Background:**

Federal Board of Revenue (FBR) is implementing the Pakistan Raises Revenue Project (PRRP) [P165982] with the assistance of the World Bank to contribute to a sustainable increase in domestic revenue by broadening the tax base and facilitating compliance. The project comprises following two basic components: (A) result-based component (Component 1) of US\$ 320 million and; (B) the traditional Investment Project Financing (IPF) component (Component 2) of US\$ 80 million .

The PRR project includes support for strengthening audit and compliance improvement measures. Under the project, the FBR has conducted risk-based audit and achieved relevant targets. FBR has also established a Compliance Risk Management (CRM) unit, and developed a CRM framework/plan (with the support of IMF). The CRM unit is established to identify, assess, and prioritize compliance risks within FBR. The CRM unit functions involve collecting and analyzing data from various sources, developing risk profiles, creating and updating a risk register and planning and implementing risk treatment strategies.

FBR is the lead government agency for revenue collection and has huge data of tax payer population and their tax declarations along with third party Data/Information from various government agencies and banks which needs to be used in a manner that can result in increased revenue for FBR. CRM Wing of FBR indents to make use of this data with the help of advanced analytic techniques and Machine Learning (ML) for multiple activities of risk profiling of taxpayer's population, selection of cases for audit, issuance of pre and post declaration alerts and devising of intelligent and human free risk parameters to be made part of an interactive dashboard for tax officials containing collated and meaningful data for discharge of their duties. At the moment CRM Wing in FBR does not have a data scientist to help meaningfully use the available data and information, and is dependent on PRAL's team for data analysis to inform decision making.

For reasons mentioned above, the FBR intends to hire services of an experienced and qualified professional for the position of 'Data Scientist' having experience in data analytics and use of machine learning and /or computer-assisted language learning tools for FBR' s Audit/CRM Wing.

**Duties and responsibilities:**

- Application of modern data analytics for Compliance Risk Management
- Reporting to the Member (Audit / CRM), the occurrence, progress and status of all risks and appropriate actions taken thereof;
- Revalidation of business processes on ongoing basis, identify potential risk factors within and outside of the organization;

- Assisting FBR to populate identified risks in CRM risk register. Maintaining the Risk Register by ensuring that all the current and future material risk exposures of the FBR are identified, assessed, quantified, appropriately mitigated and managed;
- Liaison with internal and external data source through an integrated system with Information Technology Wing of FBR for the purpose of creation of an indigenous data bank
- Collaborate with domain expert teams to identify complex business problems
- Extraction and processing of existing Oracle based data at PRAL using ETL (Extract, Transform, Load) process
- Develop predictive and descriptive models for financial forecasting
- Conduct exploratory data analysis to uncover insights from historical data
- Identification of risk prone taxpayer's using historical data through supervised and unsupervised ML techniques
- Deployment of user-friendly Dashboards and presenting data insights using data visualization
- Communicate findings to non-technical stakeholders
- Collaboration with PRAL team for deployment of approved models using ML Ops
- Stay up-to-date with industry trends and emerging technologies
- Any other relevant task assigned by the Member (Audit/CRM), Federal Board of Revenue

**Deliverables:**

- Identify risks based on patterns of declaration filing of the taxpayers with the help statistical techniques
- Developing Risk Register based on risk identified
- Selection of cases for statutory audit based on CRM framework
- Development of User-friendly interactive Dashboard duly embedding unsupervised risk parameters and third-party data for use of officers in the field.
- Deployment of End to End fully automated ML based risk identification system integrated with database available in PRAL

**Qualification and experience:**

- Masters' / Bachelor Degree (at least sixteen years of education) in Data Sciences of Data Analytics or equivalent from HEC recognized university.
- Four (04) years of post-qualification experience in data analytics and use of machine learning and /or computer-assisted language learning tools. Hands on experience of working with python, XG Boost, Decision Tree, Shapex or other such advanced tools for data analytics.
- Knowledge of Pakistan's taxation laws, and key tax administration issue may add preference

**Duration:** February 2025 to 30<sup>th</sup> June 2025

**Selection Process:**

An individual consultant will be selected in accordance with the process stipulated in accordance with “World Bank: Procurement Regulations for Investment Project Financing Goods, Works, Non-Consulting and Consulting Services” (July 2016) revised November 2017 and August 2018.