# FAO_black_20

# Food and Agriculture organization of the United Nations

### **Terms of Reference for Consultant Category B \***

|  |  |
| --- | --- |
| **Name:** |  |
| **Job Title\*\*:** | Data Scientist / Methodologist |
| **Division/Department:** | Statistics Division, ESS/ES Department |
| **Programme/Project Number:** |  |
| **Duty Station:** | HQ (Rome, Italy) |
| **Expected Start Date of Assignment:** |  | **Duration:** |  |
| **Reports to:** | ***Name:*** |  | **Title:** | Research and Innovation Team |
| \* Please note: If this TOR is for Consultant / PSA.SBS contract, the minimum relevant experience required **for the assignment** is as follows:

|  |  |  |
| --- | --- | --- |
| 1 year for a category C | 5 years for a category B | 12 years for a category A |

\*\* Please enter a short title (max 25 chars) for this assignment. |
| General Description of task(s) and objectives to be achieved |
| Statistics is a core function of FAO and represents a highly visible area of the Organization’s work. The goal of FAO’s Statistics Division (ESS) is to provide timely and reliable data on hunger, food and agriculture to facilitate the design and monitoring of evidence based policy decisions by member countries, as well as serve as the foremost authoritative source of statistical standards and state-of-the-art methods in agricultural and food statistics. Within the Statistics Division, the Research and Innovation team is responsible for the development of methods, standards, tools and norms used mainly within the organization with the objective to improve the efficiency of FAO’s statistical system and the quality of its output. Rapid technological development requires ESS to innovate in a variety of areas to modernize the statistical business process and meet the increasingly demanding needs for fast, accurate, easy and cost-effective data. Therefore, it is part of FAO’s strategy to engage with non-official, Big Data sources and to rely on data science methods to solve the current data gaps problems. The final objective is to increase the quantity, quality and timeliness of the statistics that the Organization produces. Within the Research and Innovation team, a Data Lab for Statistical Innovation has recently been established to lead the Division’s work related to data science applications (machine learning, webscrapping, crowdsourcing, geospatial analysis) and the use of big data to solve real data problems in agriculture statistics (e.g. crop acreage and yields, and agricultural prices). The data scientist will join this new Data Lab.Under the overall supervision of the Director, ESS, the direct supervision of the Senior Statistician and in collaboration with the IT division (CIO), the primary responsibility of this position is to employ big data and data science techniques across a range of solutions. The specific roles and responsibilities include, but are not limited to the following:* Identify suitable big data sources;
* Develop solutions using big data such as machine learning and statistical techniques to improve current estimates of agricultural prices as well as crop yields in terms of their coverage, timeliness and accuracy;
* Be proficient in quickly developing models, as well as taking prototypes to full production;
* Perform operationally-relevant data analysis ranging from text mining, to webscrapping, machine learning, crowdsourcing, around structure and unstructured data sources; develop and apply new and existing models and algorithms; and publish well documented, reproducible work;
* Assist in developing forecasting and nowcasting models using data from social networks, on-line databases, administrative data, satellite imagery or transaction records;
* Contribute with the ESS data dissemination and data visualization products;
* Any other duty as required.

Education: * Advanced University Degree in Statistics, Data Science, Computer Science and related topics with strong computational elements.

Core skills:* Dynamic data scientist or statistician equipped with highly advanced methodological skills, Big Data processing and analysis techniques, statistical Cloud computing with focus on machine learning, and the capability to turn new developments into production tools in support of regular statistical processes;
* At least five years of professional hands on experience in developing big data solutions and big data technology using different techniques and tools, across multiple projects.
* Advanced programming skills in statistical tools, languages and ML libraries such as R, Python, SQL and Spark ML, Hadoop environment. Proficiency on Docker would be an asset.
* Experience in ETL, Big Data storage.
* Familiar with leading visualization tools (e.g., Shiny Apps, Power BI, Tableau and d3.js)
* Ability to quickly master and apply new methods, to carry out research and and learn on own initiative;
* Ability to make recommendations and advise decision-makers on highly technical issues;
* Focus on delivery and on meeting tight deadlines.
* Highly collaborative and able to work with individuals from many different teams
* Good writing and editing skills, with a strong command of English and an ability to convey complex ideas in a clear, direct, and lively format.
 |
| key performance indicators |
| Expected Outputs: | Required Completion Date: |
| * Full production solutions developed employing big data and data science techniques in areas of agricultural statistics with its set of tested automatic processing routines
* Documents and manuals on the solutions properly published
 |  |