

Terms of Reference: Data Scientist / Data Curator

Background / General description:

WORLD BANK GROUP CONTEXT

The World Bank's Data Council coordinates the World Bank Group's approach to data, in support of the objectives of eradicating extreme poverty and promoting shared prosperity. The Data Council has identified the production of high quality statistical indicators and data products as a key priority area to help improve development outcomes. The World Development Indicators (WDI) is one of the longest running and most widely accessed data product of the Bank and also is the backbone of the Bank's open data site.

POSITION CONTEXT

This position will be based in the Data Management and Services team of the Development Data Group (DECDG). DECDG houses the World Bank's main data team, together with the Secretariat for the Data Council. Its mission is to put data to work for development by increasing country capacity to produce, disseminate, and use data of all kinds; improving accessibility, management, and dissemination of the WBG's own data assets; and by producing high quality statistical indicators and data products to help improve development outcomes. The work of the department calls for extensive partnership with other international organizations, with member countries' statistical offices, and with regions, global practices, and other units.

Within DECDG, the Data Management and Services team produces high quality data outputs, including the World Development Indicators database, various subnational databases; and the World Bank's Open data catalog. The team also manages business processes and data quality procedures related to time series data collection and dissemination, and a portfolio of key data management and dissemination platforms. The team also produces <http://data.worldbank.org>.

Duties and Responsibilities

DECDG is seeking to recruit a short term consultant to work with the Data Management and Services team, within the Development Indicators and Data Services cluster. Specific duties and accountabilities include the following:

Data Management and Quality assurance

- Create, document, and maintain software tools (Python and R packages) that others will use.
- Work with topical data administrators of the WDI database to maintain efficient data quality procedures based on international standards and recognized best practice.
- Assist in ingesting data from multiple sources (relational databases, APIs, web, etc.)
- Perform functional tasks related to the World Development Indicators database and production processes, and support dissemination of related data and data products.
- Solve any routine data-related problem you encounter using modern programming languages, libraries and tools
- Be the liaison between Bank groups and DECDG team members in adding, documenting and maintaining new datasets to the Open Data Catalog

- Work with standard codes, country classifications, indicators definition, source, and methodology metadata. Participate in improving the metadata compiled and disseminated by DECDG.
- Contribute to DECDG's efforts to enhance data dissemination through blogs and data visualizations.
- Participate in regular team activities aimed at renewing and improving data products, systems, access facilities, and web-based tools.
- Communicating with the team leader work status, foreseeable problems, and processes.

Selection Criteria

Master's Degree with a minimum of 2 years' experience or Bachelor's Degree with a minimum of 3 years' experience in economics, international relations, statistics, mathematics, social sciences, data curation or related field.

Essential skills

- **Statistical literacy:** You don't need a PhD in Quantitative Methods (no problem if you do!) but you need to be good at responsibly analyzing and presenting data using the right tools and methods. Expertise in time-series data management would be an advantage.
- **Coding:** You should be able to solve any routine data-related problem you encounter using modern programming languages, libraries and tools, and you should be comfortable developing robust, well documented software tools that others will use and maintain. It is also a big plus if you are familiar with version control (Git) and collaborative development (Github).
- **Data wrangling:** You should possess outstanding data wrangling skills, including experience working with relational databases (SQL), semi-structured data (JSON, XML), and unstructured data (text)
- **Communications:** written, oral, & visual. Whether it's documenting where exactly a piece of data comes from and how it's been produced, or creating charts and visualizations that explain an issue clearly, you've got to be good at engaging others. If selected, you will have to interact effectively with a range of clients and colleagues.
- **Problem solving:** You should be a creative problem solver with strong skills in design and problem formulation. We expect you to rapidly analyze, interpret and integrate information from varied sources to formulate clear conclusions and recommendations.
- **Positive attitude:** You are enthusiastic about development work, and you are able to take initiatives.

Other skills & experiences it would be great to see

- **Languages and tools:** Experience with Python, R, their most common libraries are especially useful for working with data and statistics. For visualization we use a lot of ggplot2, Tableau, Datawrapper and Excel, along with custom javascript and libraries such as D3 and Highcharts.
- **Topical knowledge:** Knowledge of one or more domains of development data such as economic statistics, demographic data, gender and social statistics, or other topics would be an advantage.
- **Open data:** Knowledge of open data standards and technological solutions a plus
- **Data curation:** Experience with cataloging and metadata standards a plus.

The consultant will work for 150 days within the period of July 1st, 2017 – June 30th 2018 in the offices of the World Bank in Washington, D.C.

Interested candidates should send their resume to Tony Fujs: tfujs@worldbank.org