UNITED NATIONS SYSTEM STAFF COLLEGE

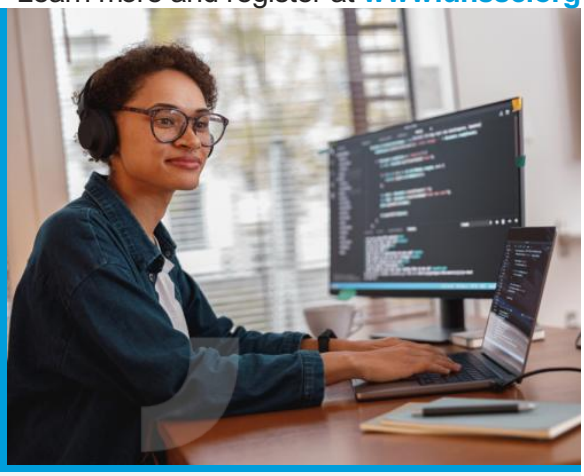
Learn more and register at **www.unssc.org**

ONLINE

01 JUN 2026 - 12 JUN 2026

Data Quality for Impact, with Python

Data and EvaluationDigital and Artificial Intelligence (AI)



LANGUAGE	DURATION	ENROLL BY	PRICE	LOCATION	TARGET
English	2 weeks	01 Jun 2026	550 USD	ONLINE	Everyone

CONTACT elp@unssc.org

This 2-week online course equips UN analysts and data practitioners with practical skills to systematically assess and improve data quality using Python.

Introduction

Poor-quality data remains one of the most costly and persistent challenges facing organizations today, undermining analysis, weakening evidence, and eroding trust in decision-making. This course equips UN personnel - analysts and data practitioners - with practical skills to systematically assess, diagnose, and improve data quality before it reaches dashboards, models, or senior decision-makers. Through hands-on exercises, participants will learn how to identify common data issues, apply structured quality checks, and implement corrective actions using the Python programming language. No prior experience with Python is required.

Objectives

By the end of the course, participants will be able to:

- Explain data quality dimensions in UN contexts (i.e. accuracy, completeness, timeliness, consistency)
- Use Python to profile, clean, and assess UN real-world datasets
- Define transparent, auditable data quality rules
- Document and communicate data quality decisions for evaluation, reporting, and policy use
- Apply data quality practices to UN data sets, survey, or programme data

Course methodology

This is an online-led instructor course.

Participants will get access to the UNSSC UNKampus30 platform, where they will find the asynchronous learning material. They will also have the opportunity to practice with optional exercises between the webinars, reflect in their personal blog and interact in the asynchronous discussion forum with the UNSSC instructor and team.

The weekly instructor-led webinars are conducted on Zoom. The webinars will take place every Tuesday and Thursday from 3:00 PM to 4:30 PM CEST. Participants need a computer (or mobile device), a reliable internet connection and either a headset with a microphone to connect to the audio through a computer, or a telephone. We recommend accessing audio through your computer. No special software is required, but participants must be able to access Zoom. We will send access instructions to registered participants, and we recommend that you download the application and test your setup in advance.

Python tools:

- Any Python Integrated Development Environment (such as Jupyter Notebook, which can be downloaded for free as part of the Anaconda Distribution)
- Google Colab (accessible through any web browser)

Course contents

This course offers a dynamic and engaging virtual learning experience, guiding participants through practical exercises.

The course is organized in 2 weekly modules, as follows:

WEEK 1

Data Quality foundations and Python introduction

- What “good data” means in the UN system
- Common data quality issues in datasets, survey, and administrative data
- First hands-on Python experience (no heavy coding)
- Use of Large Language Models (LLMs) to help with coding challenges, along with avoiding LLM-related pitfalls

Data profiling with Python: Systematically assessing data quality

- Data profiling as a foundation for data quality
- Moving from manual inspection to structured, repeatable data quality assessment
- Interpreting profiling results before any cleaning or analysis

WEEK 2

Data cleaning with Python

- Applying rule-based data cleaning using Python based on profiling results
- Making explicit, transparent decisions about how data quality issues are handled
- Building reproducible cleaning workflows suitable for UN operational and evaluation contexts

Documenting and applying data quality in UN work

- Documenting data quality decisions in Python
- Applying Python-based data quality checks in real UN workflows

In the final activity, participants will apply profiling, cleaning techniques, and data quality rules to document either their own dataset or a provided UN dataset.

Target audience

This course aims to equip UN personnel at all levels with practical skills to assess, improve, and responsibly use data in their day-to-day work.

Cost of participation

$ 550