

Python Data Science, gegevens manipuleren en visualiseren

optionele TOSA®-certificering

Praktijkcursus van 4 dagen - 28h
Ref : IYT - Prijs 2025 : € 2 440 excl. BTW

Data science is een steeds groter wordend multidisciplinair veld. Het is gebaseerd op wetenschappelijke methoden, algoritmen en processen die Python beheerst dankzij een bijzonder rijk ecosysteem. Het is nu de referentietaal geworden voor gegevensanalyse, ongeacht het formaat. Onze training stelt u in staat om de tools, bibliotheken en modules van Python onder de knie te krijgen, zodat u snel vaardigheden op het gebied van data science kunt opdoen met deze taal.

PEDAGOGISCHE DOELSTELLINGEN

Na afloop van de opleiding kan de cursist:

- Een overzicht hebben van het wetenschappelijke ecosysteem van Python
- Leer meer over de essentiële wetenschappelijke bibliotheken voor datawetenschap
- Grote datasets kunnen manipuleren met Python
- De voordelen van datavisualisatie begrijpen
- Gegevens visualiseren met Python

PEDAGOGISCHE METHODEN

De trainer geeft actief les, feedback en demonstraties, zodat de deelnemers ze sneller in de praktijk kunnen toepassen.

HANDS-ON WORK

Er worden talloze oefeningen gebruikt om de onderwerpen te illustreren.

HET PROGRAMMA

laatste update: 02/2025

1) Presentatie van het wetenschappelijke Python-ecosysteem

- Overzicht van het wetenschappelijke ecosysteem van Python: de essentiële bibliotheken.
- Weet waar u nieuwe boekwinkels kunt vinden en beoordeel hun duurzaamheid.
- De belangrijkste open source tools en software voor data science.
- Waarom wetenschappelijke distributie gebruiken, Anaconda.
- De voordelen van een virtuele omgeving begrijpen en weten hoe deze te gebruiken.
- De IPython-interpreter en de Jupyter-server.
- Best practices voor een goede start van uw data science project met Python.
- Wetenschappelijke bestandsformaten en bibliotheken om ze te manipuleren.

De ontwikkelomgeving opzetten: Anaconda installeren, een virtuele omgeving maken, een omgeving exporteren en dupliceren, Jupyter-notebooks gebruiken.

2) De SciPy-stapel

- De basis van essentiële wetenschappelijke bibliotheken waarop alle andere zijn gebaseerd: de SciPy Stack.
- NumPy: numerieke berekening en lineaire algebra (vectoren, matrices, afbeeldingen).
- SciPy, gebaseerd op NumPy voor statistiek, functionele analyse, geospatiale analyse, signaalverwerking, enz.
- Pandas: analyse van gegevens in tabelvorm (CSV, Excel, enz.), statistieken, pivots, filters, zoekopdrachten, enz.

DEELNEMERS

Engineers, ontwikkelaars, onderzoekers, datawetenschappers, data-analisten en iedereen die de wetenschappelijke wereld van Python wil leren kennen.

VOORAFGAANDE VEREISTEN

Praktijk van de Python-taal.

VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vak kennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ... De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

TOEGANGSMODALITEITEN EN -TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

TOEGANKELIJKHEID VOOR MINDERVERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

- Matplotlib: de essentiële bibliotheek voor gegevensvisualisatie.
- Meet de prestaties van de NumPy die door uw Linux is geïnstalleerd en die van Anaconda. Beeldverwerking met NumPy. Eerste plots. Statistische analyse van CSV-bestanden. Eerste mapping elementen. Fouriertransformaties.*

3) Bibliotheken weergeven

- Overzicht van Python visualisatiebibliotheken: 2D/3D, desktop/web, statistiek, cartografie, big data, enz.
 - Desktop-georiënteerde bibliotheken: Matplotlib, Pandas, Seaborn.
 - Webgeoriënteerde bibliotheken: Bokeh, Altair, Plotly, enz.
 - 3D-bibliotheken: Plotly, pythreejs, ipyvolume, enz.
 - Kaartbibliotheken: Cartopy, folium, ipyleaflet, Bokeh, cesiumpy, enz.
 - Big data bibliotheken: datashader, Vaex...
- Een reeks oefeningen met een aantal van de gepresenteerde bibliotheken. Big data, cartografische, 2D- en 3D-visualisatie.*

4) Datavisualisatie

- De voordelen van datavisualisatie
 - PyViz en het HoloViz ecosysteem gebruiken.
 - Presentatie van de tools SuperSet, Mayavi, Paraview en VisIt.
- Visualisatiebibliotheken blijven gebruiken en tools blijven manipuleren.*

5) Wetenschappelijke bestandsformaten en de verwerking van omvangrijke gegevens

- Overzicht van de belangrijkste wetenschappelijke bestandsformaten: NetCDF, HDF5, GRIB, JSON, PARQUET, MATLAB, CGNS, enz.
 - Omgaan met grote gegevens met Dask, Vaex en Xarray.
- Omgaan met gegevens van meer dan GB, lezen en schrijven van NetCDF/HDF5-bestanden. Visualisatie van klimaatgegevens, satellietbeelden, maken van video's/grafische animaties.*

DATA

Neem contact met ons op