

Python - Úvod do ETL

Kód kurzu: PYTHON_ETL

ETL znamená Extract-Transform-Load - databázové funkce, které jsou integrovány do jednoho nástroje, který extrahuje data z jedné databáze a umístí je do jiné databáze. Extrakt je proces čtení dat z databáze. Transformace je proces převodu získaných dat z původní formy do formy, ve které se očekává jejich umístění v jiné databázi. Načíst je proces zápisu dat do cílové databáze. ETL se používá k migraci dat z jedné databáze do druhé za účelem vytvoření datových trhů (DataMart) a datových skladů (DataWarehouse) a také k převodu databáze z jednoho formátu nebo typu do jiného.

Pro koho je kurz určen

Kurz je určen pro každého adepta, ale také pro zkušeného datového inženýra, který má zájem v budoucnu používat programovací jazyk Python pro získávání, zpracování a přípravu dat. Výcvik probíhá extrémně intenzivně s praktickými příklady z reálného světa, a proto se znalost programovacího jazyka Python doporučuje alespoň na střední a vyšší úrovni (PYTHON_INTRO a PYTHON_ADV) a také alespoň na střední úrovni znalostí relačních databází.

Požadavky na účastníka

- Znalost programování v Pythonu na úrovni kurzu PYTHON_INTRO a PYTHON_ADV

Metody výuky

- Odborný výklad s praktickými příklady, cvičení na počítačích.

Studijní materiály

- Online prezentace učiva a cvičení.

Osnova kurzu

- Úvod (Python, ETL, datové inženýrství)
- Zdroje dat (CSV, JSON, RESTapi, MySQL, PostgreSQL, Oracle 12c, NoSQL)
- Datové cíle (WareHouse, DataMart)
- Pandy
- ETLAchémie
- Bonobo
- Luigi
- Proud vzduchu
- Talend
- Distribuované systémy (SQOOP, PIG a SPARK / Impala vs Presto /, HIVE)
- Závěr

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved