

Titulli i Projektit: Manipulim me Big Data, NoSQL dhe Analizë e Grafeve

Lënda: Big Data

Universiteti: Universiteti i Prishtinës “Hasan Prishtina” – FIEK

Studentët: Fjolla Kadriu, Leutrim Morina, Linda Hasanaj, Lum Hoxha

Udhëzime për konfigurim dhe ekzekutim

NoSQL – MongoDB

- **Platformë e përdorur:** Studio 3T + MongoDB Atlas + MongoDB Compass
- **Konfigurim:**
 1. Instaloni MongoDB Community Edition.
 2. Përdorni **Studio 3T** për të importuar **.json** files në një koleksion të ri **mondial**.
 3. Ekzekutoni query-t në dosjet **Query1.js**, **Query2.js**, **Query3.js**.

Big Data – Apache Spark

- **Platformë e përdorur:** Google Colab + PySpark
- **Udhëzime:**
 1. Hapni **BigData.ipynb** në Google Colab.
 2. Kodet ndodhen në fillim të notebook-ut.
 3. Dataset-i **mondial** importohet nga folderi **mondial** (me komanda në notebook).
 4. Dataset-i **usa_county_wise.csv** ngarkohet nga Google Drive ose direkt nga dosja.
 5. Ekzekutoni queries dhe vizualizimet.

Network Analysis – ORA Lite

- **Platformë e përdorur:** ORA-Lite
(<https://www.cmu.edu/casos-center/research/tools/ora-lite.html>)
- **Udhëzime:**
 1. Hapni ORA Lite dhe importoni **SocialNetworkDataset.xlsx** si **Table of Network Links**.
 2. Vizualizimet janë ruajtur në **NetworkAnalysisVisualizations/**.

3. Metrikat janë gjeneruar automatikisht nga ORA dhe ruhen në [All Measures by Category/](#).

Përmbajtja e dosjes .zip

- |— NoSQL Data/
 - | |— JSON collections/ → Tabelat e migruara nga Mondial (.json)
 - | |— MongoDB queries/ → Query .js për MongoDB
- |— Big Data/
 - | |— BigData.ipynb → Notebook Colab me konfigurim Spark, queries dhe vizualizime
 - | |— usa_county_wise.csv → Dataset i ri me >500K rreshta për COVID-19
- |— Network Analysis/
 - | |— SocialNetworkDataset.xlsx → Dataset rrjeti (Facebook edges)
 - | |— NetworkAnalysisVisualizations/ → Skema vizuale (Degree, Closeness, Betweenness, grouped, ungrouped)
 - | |— All Measures by Category/ → Rezultatet e metrikave (.csv dhe .html)
- |— README.pdf → Ky dokument
- |— Dokumentimi_Teknik.pdf → Dokumentimi final i projektit
- |— Aneks A → Kodi