

README - Projekti Big Data

Përmbajtja e folderit .zip:

- |— README.pdf
- |— Dokumentimi.pdf
- |— Prezantimi.pptx
- |— AneksA.pdf
- |— Imported datasets/
 - | |— city_temperature/
 - | |— COVID-19/
 - | |— student_habits
- |— XQueries/
 - | |— XQuery1.xq
 - | |— XQuery2.xq
 - | |— XQuery3.xq
 - | |— XQuery4.xq
 - | |— XQuery5.xq
- |— XPaths/
 - | |— XPath1.xq
 - | |— XPath2.xq
 - | |— XPath3.xq
 - | |— XPath4.xq
 - | |— XPath5.xq
- |— XML_Data/
 - | |— mondial.xml
 - | |— mondial.xsd
- |— xmllint_validation/
 - | |— validation_output.png
- |— Visualizations/

```
| | — PowerBIVisualization.pbix
| | — LookerStudioVisualization.pdf
| | — PowerBIVisualization.pdf
| — View_sql/
| | — View2.sql
```

Si të përdoret projekti?

1. Validimi i XML me xmllint

- Hap terminalin dhe navigo në folderin `xmllint_validation`
- Komanda për validim: `xmllint --noout --schema mondial.xsd mondial.xml`.

2. Ekzekutimi i pyetjeve në BaseX

- Hap `BaseX GUI` ose `BaseX Editor`
- Importo `mondial.xml` si një databazë e re
- Ngarko dhe ekzekuto query-t nga folderi `XQueries` dhe `XPaths`.

3. Vizualizimet në Power BI

- Hap dosjen `Visualizations` më pas skedarin `PowerBIVisualization.pbix`
- Raportet përfshijnë:
 - Temperaturat vjetore.
 - Liqenet për shtet.
 - Lumenjtë nga një view i filtruar.

4. Vizualizimet në Looker Studio

- Vizualizimi gjendet në dosjen `Visualizations` kurse skedarin `LookerStudioVisualization.pdf` e kemi eksportuar si PDF

- Linku i këtij vizualizimi është:

<https://lookerstudio.google.com/reporting/edbea0f8-c86b-4e02-bb1e-15fc7c5a0cab>

Vërejtje

- Për të përdorur skedarin `.pbix`, duhet të keni të instaluar **Power BI Desktop**.
- Për ekzekutimin e query-ve është përdorur **BaseX 10+**.
- Validimi me `xmllint` kërkon që të jetë instaluar paketa `libxml2`.