Assignment-1

Learning Outcomes

Understand the concepts of parallel and distributed databases including the map-reduce parallelism framework, NoSQL, Spark, data replication and distributed query processing.

This assignment provides you with hands-on exercise with xpath.

A. Practice the xpath queries stated in

<https://codebeautify.org/Xpath-Tester>

with the example XML document on the page. Take screen shots [10pts]

A screenshot of a computer

Description automatically generated with medium confidence

B. Use xslplanes2.xml on xpather.com

The following link is useful to open xml files.

<https://www.freecodecamp.org/news/what-is-an-xml-file-how-to-open-xml-files-and-the-best-xml-viewers/>

1. Write the command that extracts the color of the plane? [2pts]

/planes/plane[1]/color

1. What are the colors of the each plane? [2pts]

/planes/plane/color

1. Add <pilot> Joe1</pilot> at the end of the </color> for the first plane. Save the xml file. Write the command to extract pilot. [2pt]

/planes/plane[1]/pilot

C. Use the following link to download the XML data from US Congress.

<https://xml.house.gov/MemberData/MemberData.xml>

1. Can you extract the names of the congress members? If yes, do so and save the file and upload. [2pts]

yes see file list\_of\_names.txt

2. State the number of state districts? Write the xpath command that you used. [2pts]  
441.0  
count(//member)