STEPS FOR INSTALLATION (WINDOWS)

HADOOP : https://www.youtube.com/watch?v=kUX6dCbdU3Q&t=470s

Java version : 1.8.0_202

Hadoop version : 3.2.4 (https://archive.apache.org/dist/hadoop/common/hadoop-3.2.4/hadoop-3.2.4.tar.gz)

INSTALLATIONS :

Hadoop Installation Guidelines

1. Uninstall all the previous versions of java and install -

   https://www.oracle.com/in/java/technologies/javase/javase8-archive-downloads.html
   (**Java SE Development Kit 8u202 – Windows x64)**

2. Create file ` (e.g., `**C:\Java**`), Extract and Copy the extracted folder to `C:\Java\jdk`

3. Followed step includes , extracting jre as `**C:\Java\jre**`

4. Set env variables as

   User variable : JAVA_HOME : **C:\Java\jdk**

   System variable path : **C:\Java\jdk\bin**

5. Confirm java installation from cmd : **java -version**

6. Download and extract **Hadoop - 3.2.4** as (**C:\Hadoop**)

7. Go to etc folder and make changes in necessary files:

   a) hadoop-env.cmd :

   ```
   set JAVA_HOME=C:\Java\jdk
   ```

   b) core-site.xml :

   ```
   <configuration>
   <property>
   <name>fs.defaultFS</name>
   <value>hdfs://localhost:9000</value>
   </property>
   </configuration>
   ```

   c) Create a folder in **C:\hadoop named data such that C:\hadoop\data** and two sub-folders named **datanode** and **namenode** i.e. **C:\hadoop\data\datanode** and **C:\hadoop\data\namenode**

   d) Paste the below content in hdfs-site.xml

```
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
</property>

<property>
<name>dfs.namenode.name.dir</name>
<value>C:\hadoop\data\namenode</value>
</property>

<property>
<name>dfs.datanode.name.dir</name>
<value>C:\hadoop\data\datanode</value>
</property>
</configuration>
```

e) Mapred-site.xml

```
<configuration>
<property>
<name>mapreduce.framework.name</name>
<value>local</value>
</property>
</configuration>
```

f) Yarn-site.xml

```
<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>
<value>org.apache.hadoop.mapred.shuffleHandler</value>
</property>
</configuration>
```

8. Delete the existing bin from Hadoop folder and replace it with new bin

   (https://drive.google.com/drive/folders/1iURNbow2IglhAhSy3sfY5xxVfAg33NBW)

9. Try running winutils.exe from this new bin folder , if it does not run and throws error,

10. Go to https://www.dll-files.com/msvcr120.dll.html

Version 12.0.40664.0	**MD5:** ❓ b70474fe249402e251a94753b742788c	
Architecture 64		
File size 0.92 MB	**SHA-1:** ❓ f53b3c21adf75dc84977067869253e207f1b9795	
Language U.S. English		
Company Microsoft Corporation	⬇ Download	
Description Microsoft® C Runtime Library	**Zip file size:** 0.44 MB	

Download the above-mentioned file

Extract it and copy **mscvr120.dll** file and paste it in **C:\Windows\System32** such that

C:\Windows\System32\mscvr120.dll exists

11. Setup env variables

    a) User variables:

    HADOOP_HOME : **C:\hadoop**

    HADOOP_USER_CLASSPATH_FIRST : true

    b) System variable path:

    C:\hadoop\bin

    C:\hadoop\sbin

12. Open command prompt in admin mode and

    ```
    >  hdfs namenode -format
    >  jps
    >  start-all.cmd
    >  jps
    ```

13. Check for localhost if working (localhost:9870) & (localhost:8088)

14. Done!!