

**Department of Artificial Intelligence & Data Science****Vision of the Department***To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.***Mission of the Department***To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.***Session 2025-2026****Vision:** Dream of where you want.**Mission:** Means to achieve Vision**Program Educational Objectives of the program (PEO):** (broad statements that describe the professional and career accomplishments)

PEO1	Preparation	P: Preparation	Pep-CL abbreviation pronounce as Pep-si-IL easy to recall
PEO2	Core Competence	E: Environment (Learning Environment)	
PEO3	Breadth	P: Professionalism	
PEO4	Professionalism	C: Core Competence	
PEO5	Learning Environment	L: Breadth (Learning in diverse areas)	

Program Outcomes (PO): (statements that describe what a student should be able to do and know by the end of a program)**Keywords of POs:**

Engineering knowledge, Problem analysis, Design/development of solutions, Conduct Investigations of Complex Problems, Engineering Tool Usage, The Engineer and The World, Ethics, Individual and Collaborative Team work, Communication, Project Management and Finance, Life-Long Learning

PSO Keywords: Cutting edge technologies, Research

“I am an engineer, and I know how to apply engineering knowledge to investigate, analyse and design solutions to complex problems using tools for entire world following all ethics in a collaborative way with proper management skills throughout my life.” to contribute to the development of cutting-edge technologies and Research.

Integrity: I will adhere to the Laboratory Code of Conduct and ethics in its entirety.**Name and Signature of Student and Date**

(Signature and Date in Handwritten)



Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Session	2025-26 (ODD)	Course Name	BDH Lab
Semester	7 AIDS	Course Code	22ADS703
Roll No	21	Name of Student	Sanskriti.Paunikar

Practical Number	7
Course Outcome	CO1:- 1. Understand big data analytics and its business applications. CO2:- Analyze the HADOOP and Map Reduce technologies associated with big data analytics. CO3:- Apply Big Data analytics Using Pig and Hive.
Aim	Installation of Apache Pig on Linux with Hadoop Integration
Problem Definition	
Theory (100 words)	Apache Pig is a high-level platform for processing large datasets in Hadoop using Pig Latin, a scripting language that simplifies writing MapReduce programs. Pig translates scripts into MapReduce jobs, enabling parallel data processing on HDFS. Installing Pig on Linux involves downloading the Pig binary, configuring environment variables, and integrating it with an existing Hadoop installation. This integration allows Pig to access HDFS data and leverage Hadoop's distributed computing framework. Pig is ideal for ETL operations, data transformations, and iterative data processing, offering simplicity over traditional Java MapReduce while maintaining Hadoop's scalability, fault tolerance, and parallel processing capabilities.
Procedure and Execution (100 Words)	Steps of Implementation:- 1. Ensure Hadoop is installed and running. 2. Download and extract Pig. 3. Move Pig folder to `/usr/local/pig`. 4. Set environment variables ('PIG_HOME', 'HADOOP_HOME', update 'PATH'). 5. Source `~/bashrc` to apply changes. 6. Start Pig Grunt shell using `pig`. 7. Verify Hadoop integration by running Pig commands to load/store data in HDFS.



Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Code:

Pig Releases

Please make sure you're downloading from a [nearby mirror site](#), not from www.apache.org.

Older releases are available from the [archives](#).

Name	Last modified	Size	Description
Parent Directory		-	
latest/	2016-06-07 22:38	-	
pig-0.15.0/	2015-06-05 23:01	-	
pig-0.16.0/	2016-06-07 22:38	-	
HEADER.html	2015-01-13 00:20	317	

Apache/2.4.10 (Debian) Server at mirror.fibergrid.in Port 80

```
hduser@hadoop:/home/subipalaniappan$ sudo wget https://apachemirror.wuchna.com/pig/pig-0.16.0.tar.gz
[sudo] password for hduser:
--2021-05-12 04:52:13-- https://apachemirror.wuchna.com/pig/pig-0.16.0/pig-0.16.0.tar.gz
Resolving apachemirror.wuchna.com (apachemirror.wuchna.com)... 143.110.177.196
Connecting to apachemirror.wuchna.com (apachemirror.wuchna.com)|143.110.177.196|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 177279333 (169M) [application/x-gzip]
Saving to: 'pig-0.16.0.tar.gz'

pig-0.16.0.tar.gz      100%[=====>] 169.07M  12.9MB/s   in 15s
2021-05-12 04:52:28 (11.6 MB/s) - 'pig-0.16.0.tar.gz' saved [177279333/177279333]

hduser@hadoop:/home/subipalaniappan$ sudo mkdir -p /usr/local/pig
hduser@hadoop:/home/subipalaniappan$ ls
pig-0.16.0.tar.gz
hduser@hadoop:/home/subipalaniappan$ sudo tar -zxvf pig-0.16.0.tar.gz
```

```
hduser@hadoop:/home/subipalaniappan$ ls
pig-0.16.0  pig-0.16.0.tar.gz
hduser@hadoop:/home/subipalaniappan$ cd pig-0.16.0/
hduser@hadoop:/home/subipalaniappan/pig-0.16.0$ ls
CHANGES.txt  RELEASE_NOTES.txt  contrib  legacy  pig-0.16.0-core-h1.jar  src
LICENSE.txt   bin                docs     lib     pig-0.16.0-core-h2.jar  test
NOTICE.txt    build.xml          ivy      lib-src  scripts                tutorial
README.txt    conf               ivy.xml  license  shims

hduser@hadoop:/home/subipalaniappan/pig-0.16.0$ sudo mv * /usr/local/pig
mv: target 'pig' is not a directory
hduser@hadoop:/home/subipalaniappan/pig-0.16.0$ sudo mv * /usr/local/pig
hduser@hadoop:/home/subipalaniappan/pig-0.16.0$ ls
hduser@hadoop:/home/subipalaniappan/pig-0.16.0$ cd /usr/local/pig/
hduser@hadoop:/usr/local/pig$ ls
CHANGES.txt  RELEASE_NOTES.txt  contrib  legacy  pig-0.16.0-core-h1.jar  src
LICENSE.txt   bin                docs     lib     pig-0.16.0-core-h2.jar  test
NOTICE.txt    build.xml          ivy      lib-src  scripts                tutorial
README.txt    conf               ivy.xml  license  shims
hduser@hadoop:/usr/local/pig$
```

```
hduser@hadoop:/$ sudo chown -R hduser:hadoop /usr/local/pig
hduser@hadoop:/$ sudo vim ~/.bashrc
hduser@hadoop:/$ source ~/.bashrc
hduser@hadoop:/$ pig -x local
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.0.jar) method org.apache.hadoop.security.authentication.util.KerberosUtil.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further reflective access
WARNING: Please use the following command to configure future tool versions:
jlink --add-modules java.se.ee --illegal-access=warn
```



Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Output Analysis	After installation, running Pig in Grunt shell allows execution of Pig Latin scripts. Commands like LOAD, FILTER, GROUP, and FOREACH process data and produce intermediate and final results. The output confirms proper Hadoop integration when Pig successfully reads from and writes to HDFS, executes jobs via MapReduce, and displays aggregated or transformed results. Verification using HDFS commands (hdfs dfs -ls /path) ensures that Pig scripts correctly manipulate data in the Hadoop ecosystem. This demonstrates Pig's ability to simplify large-scale data processing while maintaining efficiency and scalability..
Link of student Github profile where lab assignment has been uploaded	https://github.com/sanskriti-1234/BDH.git
Conclusion	Apache Pig was successfully installed and integrated with Hadoop on Linux. Pig provides a high-level, script-based approach for processing large datasets, reducing the complexity of writing MapReduce programs. The experiment confirms Pig's capability to read, transform, and store data on HDFS efficiently, making it a practical tool for ETL, data analysis, and big data processing tasks in Hadoop environments.
Plag Report (Similarity index < 12%)	
Date	30/10/2025