Session 5: Exploring Pig

Assignment 5.2

Implement the use case present in below blog link and share the complete steps along with

screenshot(s) from your end.

<https://acadgild.com/blog/aviation-data-analysis-using-apache-pig/>

Datasets – airports.csv and DelayedFlights.csv

JAR used – piggybank.jar

Created 4 pig scripts for all 4 problem statements –



Now running the scripts and storing the output in system

**Problem Statement 1**

Find out the top 5 most visited destinations.

Solution :-







## Output-

## Problem Statement 2

Which month has seen the most number of cancellations due to bad weather?

Solution –







## Output-

## 

## Problem Statement 3

Top ten origins with the highest AVG departure delay

Solution-







## Output-

## 

## Problem Statement 4

Which route (origin & destination) has seen the maximum diversion?

Solution-







Output-

