

Capstone Project

PySpark

Please find the AIRLINES data set (airlines.csv)

Year

Quarter

Average revenue per seat

Total number of booked seats

- 1) What was the highest number of people travelled in which year?
- 2) Identifying the highest revenue generation for which year
- 3) Identifying the highest revenue generation for which year and quarter (Common group)

