\*\*📊 Finance & BI Dashboard – AWS vs Azure (Pick n Pay Case Study)\*\*  
  
\*\*📌 Executive Summary\*\*  
  
This project simulates a retail business intelligence pipeline using AWS services (S3, Glue, Athena, QuickSight) to process raw interaction data into actionable dashboards. It compares AWS vs Azure analytics pipelines in cost, flexibility, and ease of use. Simulated for Pick n Pay’s financial and behavioural analytics team.  
  
\*\*🧠 Architecture Overview\*\*  
  
\*\*Figure A.1 – Architecture Diagram\*\*  
  
The pipeline begins by uploading customer interaction and financial data into Amazon S3. AWS Glue crawls and catalogs the data into tables. Amazon Athena enables SQL queries to aggregate, filter, and join datasets. Finally, Amazon QuickSight visualizes KPIs such as event frequency and purchase patterns.  
  
\*\*🛠️ AWS Services Used\*\*  
  
- \*\*Amazon S3\*\* – Stores raw .csv event and financial data  
- \*\*AWS Glue\*\* – Crawls datasets and builds schema in Glue Data Catalog  
- \*\*Amazon Athena\*\* – Runs SQL queries over structured data  
- \*\*Amazon QuickSight\*\* – Creates interactive dashboards  
- \*\*IAM\*\* – Secures access to Athena, S3, and Glue  
- \*\*CloudWatch\*\* – Optional logging for query errors/performance  
  
\*\*💡 Key Features\*\*  
  
- End-to-end BI pipeline built on AWS  
- Auto-schema detection with Glue Crawlers  
- SQL-based querying using Athena  
- QuickSight dashboards with pie/bar/time series  
- Join queries between event and revenue data  
- Forecasting and visual drill-downs  
- Azure Power BI also tested as an alternative  
  
\*\*🧪 Sample Query\*\*  
  
sql  
  
CopyEdit  
  
SELECT event\_type, COUNT(\\*) AS total\_events  
  
FROM interactions\_clean  
  
GROUP BY event\_type;  
  
\_This query aggregates customer events (e.g., VIEW, PURCHASE) from the cleaned dataset using Athena.\_  
  
\*\*📸 Dashboard and Data Pipeline Screenshots\*\*  
  
\*\*Figure 1 – Sample Table View in Athena (Spreadsheet Format)\*\*   
Preview of structured data from interactions\_clean.  
  
\*\*Figure 2 – AWS Glue Crawler Setup for interactions\_clean\*\*   
Shows the Glue crawler configuration used to ingest S3 data.  
  
\*\*Figure 3 – Athena Join Query (financial + interaction)\*\*   
Joining event logs with financial data to show total spend per event type.  
  
\*\*Figure 4 – Aggregating Event Types in Athena (GROUP BY)\*\*   
Displays query to count total events by type (e.g., VIEW, PURCHASE).  
  
\*\*Figure 5 – Athena LIMIT 10 (Preview Raw Data)\*\*   
Simple query to preview the first 10 rows from the event dataset.  
  
\*\*Figure 6 – QuickSight Pie Chart (VIEW vs PURCHASE)\*\*   
Pie chart visualizing distribution of user interactions by type.  
  
\*\*Figure 7 – QuickSight Bar Chart of Events by Date\*\*   
Bar chart showing number of interactions over time.  
  
\*\*Figure 8 – Forecasting or Enhanced KPI Panel (QuickSight)\*\*   
Advanced forecast panel (optional) using QuickSight's ML features.  
  
\*\*🧩 Future Enhancements\*\*  
  
- Integrate anomaly detection (e.g., AWS Lookout for Metrics)  
- Add SageMaker forecasts for retail demand  
- Use \*\*QuickSight Q\*\* for natural language BI insights  
- Compare response time and tuning between Azure and AWS further  
  
\*\*📁 Suggested Repo Structure\*\*  
  
bash  
  
CopyEdit  
  
finance-bi-dashboard/  
  
├── raw\_data/ # S3 data samples (CSV)  
  
├── athena\_queries/ # SQL queries used for analysis  
  
├── images/ # Screenshots and architecture diagrams  
  
├── quicksight/ # Dashboard designs or JSON exports  
  
└── README.md # This file  
  
\*\*🙋 Contact & Credits\*\*  
  
\*\*Joshua Barradas\*\*   
📍 Leeds, UK   
📧 [barradasjoshua48@gmail.com](mailto:barradasjoshua48@gmail.com)   
🔗 GitHub: [@Joshuabarradas234](https://github.com/Joshuabarradas234)