Ollscoil Teicneolaíochta an Atlantaigh Dún na nGall
Atlantic Technological University Donegal

Assignment Outline

Assignment:	Assignment 3: Data Analytics Plan
Objective:	The objective of this assignment is to enable students to develop an analysis plan for a dataset and analysis question. You are given a real-world big data analytics dataset, and a choice of analysis questions. Your task is to describe the dataset, choose a question, and develop an analysis plan that will answer the question.
Learning Outcomes:	• Defend the selection of an appropriate scripting language for the analysis of data • Identify, explore and make informed judgement relating to the design, modelling and implementation of data analytics systems paying particular attention to legal and ethical issues • Reflect upon their personal professional practice demonstrating appropriate and effective communication and organisational skills in the development of an application as part of a team • Analyse complex problems resultant from the use of large-scale data sources for data analysis and defend appropriate approaches to solving the business problem • Apply research, information gathering, critical analysis, design and implementation techniques appropriately and effectively • Design and implement standard implementation techniques for the communication of analysis in a business appropriate manner.
Due Date:	Thursday 5th Feb 2026 @ 11.30pm
Submission:	Assignment should be submitted to Blackboard

Notes

This is an individual assignment and there should be no plagiarism of online resources or malpractice. You should not use Generative AI resources such as ChatGPT to prepare your assessment. Students have a responsibility to keep their own work secure and maintain the integrity and validity of their academic output. **This assessment is worth 20% of the total marks** for the Big Data Analytics module.

Ollscoil Teicneolaíochta an Atlantaigh Dún na nGall
Atlantic Technological University Donegal

Assignment Overview

Assignment Title: Big Data Analytics Planning for a Business

This assignment focuses on planning and design of a big data analytics solution. You are not required to implement any code or dashboards in this assignment. You will analyse a given case study, review the provided datasets, and develop a data analytics plan that will later be implemented in final Assignment 4.

Company Background:

FreshNest Retail is a medium-sized retail company operating multiple stores and warehouses across different regions. The company sells fast-moving consumer goods and relies on several suppliers and third-party logistics providers. Over the last year, FreshNest has experienced Frequent stockouts of popular products, Excess inventory of slow-moving items, Unreliable supplier deliveries, Rising logistics delays and costs. Management believes that these problems are linked to poor data integration and lack of analytics-driven decision making.

Business Problem:

The company collects large volumes of data from different operational systems, but this data is stored in separate files or systems, not analysed together, and not used for proactive decision making. Senior management has asked the analytics team to design a big data analytics solution that can Integrate multiple datasets, provide operational visibility, support forecasting and risk monitoring, enable dashboard-based decision support. Before implementation, management requires a detailed analytics plan.

Data Availability:

You will work with the datasets available with this assignment (.zip).

Based on business problem, top management developed following **Research Questions**. You must select **any four questions** and design an analytics plan to address them.

1. How do sales demand patterns vary across time, products, and regions?
2. Which products and locations are at the highest risk of stockouts?
3. How efficiently is inventory being managed across warehouses and stores?
4. Which suppliers show consistent delivery delays or poor reliability?
5. What are the main factors contributing to logistics delays and increased delivery times?
6. Can historical data be used to forecast future demand and identify high-risk periods?

Task Requirements

You must develop a **Big Data Analytics Plan** covering the following sections.

Task 1: Understanding the Case and Selected Research Questions (Marks 10%)

Briefly summarise the business problem in your own words. Clearly list the **four research questions** you have selected and explain **why they are important** for the company.

Task 2: Data Description and Understanding (Marks 20%)

For each dataset you plan to use, explain 1) Datasets and purpose 2) Key variables and their meaning 3) Data types required to answer research questions 4) Potential data quality issues. Also explain how datasets will be linked together to answer the selected research questions.

Task 3: Data Ingestion and Processing Plan (Marks 20%)

Describe how the data will be handled in a **big data analytics environment**. Your plan should explain 1) Where data will be stored 2) How data will be loaded 3) Cleaning and preprocessing steps like handling missing, duplicate, or inconsistent data 4) Planned transformations such as Joins Aggregations, Feature engineering, etc. Justify each step based on the selected research questions.

Task 4: Data Analytics Plan (Marks 40%)

For each selected research question, explain the type of analytics required (Descriptive, diagnostic, predictive, or Prescriptive Analytics).

Explain analytics techniques you may use (Trend analysis, Aggregation and KPIs development, Correlation or hypothesis testing, Forecasting or regression modelling).

You must also explain 1) Which Spark SQL functions or DataFrame operations may be required 2) Which Python libraries may be used for analysis. No code is required. Focus on explanation and justification.

Optional: You may plan for streaming data analytics for higher grades.

Task 5: Data Visualisation and Dashboard Plan (Marks 10%)

Design a **conceptual dashboard** to present insights to management. Explain 1) Key KPIs to be shown 2) Types of visualisations required (e.g., bar/pie/line plots, heatmaps, etc) 3) Why each visualisation is suitable. Explain how the dashboard will help answer the selected research questions.

You may use any platform for visualisation and Dashboards (Python libraries or Databricks or PowerBI or Tableau).

Deliverables

Students must submit a **single written report (2000 words +/- 200)** including all tasks. No code, dashboards, or screenshots are required in this assignment.

Note:

The datasets provided for this assignment are **synthetic,** created solely for this assignment only. They do not represent, replicate, or belong to any real organisation, company, or individual. No personal data, confidential information, or commercially sensitive data is contained in these datasets. Any resemblance to real organisations, brands, or operations is purely coincidental.