**7PAM2002 Data Science Project Module**

**Choosing a Project Form**

**Semester B 2024/2025**

This form will be used by your supervisor to agree your project topic and dataset. Complete as much of the form as possible then submit the form into ‘Assignments’ on the Project Module Canvas site.

**Student Full Name: Yasir Ashfaq**

**Student ID: 21083859**

**Email Address: Yasirashfaqkamboh@gmail.com**

**Course: 1 year / 2 year sandwich / 2 year Advanced Research** *(delete as appropriate)*

**Semester intake to the course: A / B (***delete as appropriate)*

|  |  |
| --- | --- |
| SECTION A |  |
| **What is your proposed project title or topic?** | Supply Chain Management on Retail Store Inventory Forecasting |
| **What is the Research Question for your project?** | How can machine learning models be used in predicting retail store inventory demand through past sales, seasonal trends, and outside factors to manage and maintain supply chain efficiency? |
| **Dataset website address (or organisation and person the dataset is from).** | <https://catalog.data.gov/dataset/warehouse-and-retail-sales> |
| **Where was the data originally collected? (who, when, where)** | Anirudh Chauhan collected the original dataset, which can be found here on Kaggle. The synthetic dataset was created to practise inventory management and demand forecasting. The dataset was created to simulate real world retail scenarios of sales history, inventory levels, etc features. The data set was published on Kaggle on 2024-11-24 and was meant for retail data scientists and analysts to develop and run predictive models to optimise retail inventory. The dataset does not specify Physical location but is widely used for global retail analytics and forecasting applications. |
| **What type of data are you using? (e.g., image/tabular/category/continuous etc)** | It is a tabular structured dataset that is synthetic retail store information. Features such as historical sales, inventory levels, product categories, store IDs as well as timestamps are included in it. It is designed after a real pattern in retail and thus can simulate the real demand patterns, patterns at which inventory turnover and supply chain efficiency analysis are considered. It is suitable for a machine learning model, statistical analysis or time series forecasting as it is in tabular data format. The Kaggle dataset is popularly used for predictive analytics on retail inventory management by businesses to provide efficiency and reduce stock costs. |
| SECTION B |  |
| Can you attend all 6 supervision sessions on-campus? If not state reason. | No, I can’t attend all supervisor meetings due long travelling but I can manage online. |
| What was the subject of your BSc degree and any other Master’s degrees you have taken? | Bachelor of science in Electrical Technology |
| What are your career aims and/or the industry sector you would like to get into? | I aim to work as a data scientist in the tech or healthcare industry, where I can apply machine learning and data analysis to solve real-world problems and make data-driven decisions. |
| A brief account of your programming/data science experience (including work or placement). | I have experience with Python, R, SQL, and tools like TensorFlow. I've worked on several projects. |
| Any work experience (including non-computing related). | I’ve worked in customer service and as a research assistant, which taught me problem-solving, communication, and attention to detail. |
| Do you have any hobbies, activities or interests. | I like Football, Music. |