**COM618 – Data Science – Lab Activities**

**Week 1 Practical Lab Activities**

**Tools Required:** Python (Jupyter Notebook or Google Colab), Excel, Internet Access, Miro or Jamboard, Presentation Software

**Lab 2: Big Data in Action – The 5 V’s Challenge**

**Objective:** To explore the five dimensions of Big Data—Volume, Velocity, Variety, Veracity, and Value—through sector-specific case studies.

**Group Setup:** Divide students into groups. Each group is assigned a different sector:

* **Group A:** Healthcare
* **Group B:** Retail and E-commerce
* **Group C:** Transport and Smart Cities

**Tasks:**

1. **Research & Identify Examples**
   * Find real-world examples of each of the 5 V’s within your assigned sector.
   * Use online datasets (e.g., NHS, Kaggle, Transport for London) to support your findings.
2. **Create a Sector Infographic**
   * Design a one-slide infographic summarising your examples.
   * Include brief definitions and visuals for each V.
3. **Present Your Findings**
   * Each group presents their infographic to the class (5 minutes per group).
   * Highlight how Big Data is transforming your sector.

**Expected Outcomes:**

* Clear understanding of the 5 V’s of Big Data
* Ability to apply theoretical concepts to real-world contexts
* Improved presentation and teamwork skills

**Lab 3: Role Play – Build Your Data Science Team**

**Objective:** To understand the interdisciplinary roles within a data science project and simulate collaborative planning.

**Group Setup:** Divide students into three project teams. Each team will role-play the following positions:

* **Data Engineer**
* **Data Scientist**
* **Data Analyst**
* **Domain Expert**

**Scenario:** Your team is tasked with designing a data-driven solution to **optimise energy usage in smart homes**.

**Tasks:**

1. **Define Role Responsibilities**
   * Each member outlines their role’s contribution to the project.
   * Discuss tools, skills, and data requirements.
2. **Draft a Mini Project Plan**
   * Identify the data sources (e.g., smart meter readings, weather data).
   * Choose appropriate tools (e.g., Python, SQL, Power BI).
   * Outline the workflow: data collection → cleaning → analysis → reporting.
3. **Visualise the Workflow**
   * Use Miro or Jamboard to create a visual map of your team’s process.
   * Include timelines, dependencies, and expected outcomes.
4. **Team Reflection**
   * Briefly discuss challenges and how cross-functional collaboration improves outcomes.

**Expected Outcomes:**

* Deeper understanding of data science roles
* Practical planning and workflow mapping
* Enhanced communication and project design skills

**Lab 5: Mini Case Study – Data Science in Daily Life**

**Objective:** To connect data science principles to everyday experiences and reflect on their impact.

**Group Setup:** Divide students into three groups. Each group selects a different everyday application:

* **Group A:** Streaming services (e.g., Netflix, Spotify)
* **Group B:** Online shopping and recommendation engines
* **Group C:** Fitness tracking and wearable tech

**Tasks:**

1. **Map the Data Journey**
   * Identify how data is collected, cleaned, analysed, and used to make decisions.
   * Highlight user interactions and feedback loops.
2. **Identify Roles and Tools**
   * Determine which data science roles are involved.
   * List tools and technologies used (e.g., machine learning models, cloud platforms).
3. **Create a Case Study Slide**
   * Summarise your findings in a one-slide case study.
   * Include diagrams or flowcharts to illustrate the data journey.
4. **Group Reflection**
   * Discuss how data science improves the user experience.
   * Consider ethical implications (e.g., privacy, bias).

**Expected Outcomes:**

* Ability to relate data science to real-world applications
* Improved analytical and storytelling skills
* Awareness of ethical considerations in data use