AE1 ASSESSMENT RE-EXPLAINED

Part A – What You Need to Do:

1. Find Healthcare Datasets

- Go online and look for **free, public datasets** related to **health or medicine**. You can use websites like **Google Dataset Search**, **WHO**, or **Data.gov.uk**.
- Choose one dataset that interests you (for example, COVID data, heart disease stats, etc.).
- Use a **visualisation tool** (like Excel, Python, or Tableau) to create graphs or charts that help explain the data.
- In your **report**, explain:
    - Why you picked that tool (e.g., Python because it's powerful and flexible).
    - Why the visualisation is useful in the **healthcare field** (e.g., showing trends in illness can help hospitals prepare better).

2. Clean and Prepare the Data

- Look through your dataset and check if it's **messy**:
    - Are there any **missing values**?
    - Are there any **errors or duplicates**?
- Use techniques to **clean it up** (like removing rows with missing data or fixing errors).
- In your report, explain **what you did to clean it** and **why** (back it up with some quick research – like why it's important to remove missing data for accurate results). Use appropriate citations or references.

3. Explore the Data

- Use tools like **SPSS**, **Python**, Power BI or Tableau to **analyse the data**.
- Try to find any **interesting patterns** or **trends**. For example, "Heart disease is more common in people over 50."
- Write about the patterns you found and explain what they might mean.

Part B – What You Need to Do (in plain English):

4. Build a Predictive Model + Visualise It

- **Find a problem in your dataset** – for example, maybe you noticed that hospital readmission rates are high, or that some age groups have higher risk for certain illnesses.
- Think about how you could **predict something** based on the data:
    - Like: "Can we predict if a patient is likely to get diabetes?"
    - Or: "Can we forecast how many hospital visits there will be next month?"
- Use a tool you are **comfortable with** (like Python, Tableau, or Power BI or WEKA) to:
    - **Visualise the data** again (charts, graphs, etc.)
    - **Build an excellent model** that makes predictions or shows future trends.
- In your report, explain:
    - What tool, models and techniques you used (e.g., Python with scikit-learn, SVP, Randon forest, K- Means clustering, Ensemble Model, Tableau with forecast tools).
    - Why you chose them.
    - What **patterns or trends** you discovered.

5. Evaluate Your Work

- Look at the **results of your model**:
    - Were the predictions accurate?
    - What went well, and what could be better?
- **Compare your findings** with other studies or reports.
    - For example: "My results are similar to what the WHO found in their 2022 report on diabetes."
- This shows that your work is based on real-world research and not just guesses with excellent latest sources/references including 2024 & 2025.

Important Things to Include in Your Report:

- **Screenshots** of your data, graphs, analysis, etc.
- **Code snippets** if you used something like Python (just the important parts).

- **Explain everything clearly**: what you did, why you did it, and what you found.
- Use **examples from research** or articles that show how tools like Python or Tableau have been used in healthcare.

Submit This: FINAL SUBLISSION MUST INCLUDE

What You Need to Submit

You are required to submit **two files** for this project:

1. **Your Report**
    - This should be a **Word document or PDF**.
    - Upload it **separately** from the rest of your files.
2. **Your Artefacts (Project Files)**
    - These include your:
        - **Dataset**
        - **Source code** (e.g., Python scripts, Jupyter notebooks)
        - **Tableau or Microsoft Power BI or WEKA**
        - **Screenshots or visualisation images**
        - **Recorded demo/video**
        - **Any other relevant files**
    - All of these should be placed into a **single compressed ZIP folder** and uploaded together.
    - You will be expected to **demonstrate your artefacts** through a **short, recorded video** (maximum **7 minutes**). Upload the video and submit it.