Introduction to Data Science: Post-Class Assignments

Basic Assignment: Identifying Data Types in AI Services

Title: Data Types in One Hour AI Solution

Objective: Identify and classify different types of data within a business context.

Scenario: One Hour AI Solution connects clients with AI engineers for one-hour problem-solving sessions.

Tasks:

1. Look at these 8 data elements the platform might collect:

   o Client industry (Healthcare, Finance, Retail, etc.)

   o Session date

   o Session duration in minutes

   o Problem type (ML Model Optimization, NLP, etc.)

   o Engineer specialization

   o Client satisfaction rating (1-5)

   o Client's problem description (text)

   o Solution achieved (Yes/No)

2. For each element:

   o Classify it as structured, unstructured, or semi-structured

   o Write a brief reason for your classification (1 sentence)

3. Answer these questions (2-3 sentences each):

   o Which data elements would be most useful for matching clients with engineers?

   o Which data elements might help improve the platform's service quality?

Deliverable: A simple table with your classifications and a short paragraph answering each question (about 1 page total).

Estimated time: 30-45 minutes

Title: Planning a Data Science Approach for One Hour AI Solution

Objective: Apply the data science process steps to a business scenario.

Scenario: One Hour AI Solution wants to use data science to improve their matching of clients with engineers.

Tasks:

1. For THREE of the following data science steps, describe how One Hour AI Solution could apply them (3-4 sentences each):

Choose THREE from:

- o   Finding the Right Problem

- o   Gathering Data

- o   Cleaning the Data

- o   Building Models

- o   Putting Models to Work

2. Choose ONE type of analytics from below and explain how One Hour AI Solution could use it to improve their business (4-5 sentences):

- o   Descriptive Analytics ("What Happened?")

- o   Predictive Analytics ("What Will Happen?")

- o   Prescriptive Analytics ("What Should We Do?")

Deliverable: A one-page write-up with your explanations.

Estimated time: 45-60 minutes

Title: Ethical Data Practices for One Hour AI Solution

Objective: Identify potential ethical concerns in data collection and use.

Scenario: As One Hour AI Solution grows, they need to ensure their data practices are ethical and responsible.

Tasks:

1. Read the following scenario:

One Hour AI Solution collects data including:

- Engineer profiles (experience, specialization, languages spoken)
- Client information (industry, problem descriptions)
- Session feedback and ratings
- Solution details and success rates

They plan to use this data to:

- Match clients with the most suitable engineers
- Set pricing based on problem complexity
- Evaluate engineer performance
- Market their services to potential clients

2. Identify TWO potential ethical concerns that might arise from their data practices, selecting from:

- Privacy issues
- Bias in matching
- Transparency in how data is used
- Data security

3. For each concern you identified:

- Explain specifically why it might be problematic (3-4 sentences)
- Suggest ONE practical solution to address it (2-3 sentences)

Deliverable: A one-page document outlining your ethical analysis and recommendations.

Estimated time: 45-60 minutes