# Project Idea: UK Road Safety Analysis

**Week 1: Build Data Model, Data Cleaning and Preprocessing**

- **Tasks:**
    - **Data Preprocessing:** Build a data model by merging the three main datasets (Accidents, Vehicles, and Casualties). Clean the data by handling missing values (e.g., in Age, Location, or Time columns) and decoding categorical variables using the Lookup Reference (e.g., converting '1' to 'Sunny', '2' to 'Rainy').
    - **Tools:** SQL, Python (pandas, Matplotlib).
- **Deliverables:**
    - Cleaned master dataset ready for analysis.
    - Data preprocessing notebook documenting the cleaning steps.

**Week 2: Exploratory Analysis Questions Phase**

- **Tasks:**
    - **Determine Data Analysis Questions:** Determine initial descriptive questions to understand the "Big Picture". Focus on identifying general trends such as accident frequency by time of day, day of the week, and monthly variations. Analyze the distribution of accident severity across different cities.
    - **Tools:** SQL, Python (pandas, Matplotlib, Seaborn).
- **Deliverables:**
    - Set of descriptive analysis charts (e.g., "Accidents by Hour", "Accidents by Day of Week").

**Week 3: Deep Dive & Diagnostic Analysis Phase**

- **Tasks:**
    - **Determine Deep Dive Questions:** Instead of predicting the future, dig deeper into "Why" accidents happen.
    - **Correlation Analysis:** Analyze relationships between variables (e.g., "Is there a correlation between Road Surface conditions and Accident Severity?").

    - **Segmentation:** Segment the data to analyze specific high-risk groups (e.g., "Analyze the causes of Fatal accidents specifically involving Motorcycles vs. Cars").
    - **Insight Generation:** Derive actionable recommendations based on historical patterns (e.g., "Identify the top 5 most dangerous road types for pedestrians").
    - **Tools:** Python (pandas, Matplotlib, Seaborn).
- **Deliverables:**
    - Advanced visualization plots (Heatmaps, Correlation Matrices) answering diagnostic questions.
    - List of data-driven recommendations.

**Week 4: Visualization Dashboard and Final Presentation**

- **Tasks:**
    - **Build a Visualization Dashboard:** Build a Tableau (or Power BI) visualization dashboard that visualizes the answers to all answered questions. It should include an interactive map of accident hotspots and filters for weather conditions, vehicle types, and casualty age.
    - **Final Presentation:** Prepare a report and presentation summarizing the project work, including data analysis steps, key insights, and safety recommendations.
    - **Tools:** SQL, Python (pandas, Matplotlib), Tableau/Power BI.
- **Deliverables:**
    - Interactive Visualization dashboard.
    - Final report and presentation