DATA SCIENCE ASSIGNMENT – SMART CITY TRAFFIC & SAFETY ANALYTICS

Instructions: Complete all tasks in a Jupyter/Colab Notebook and upload the notebook on Google Classroom (GCR). Make sure your code is clean, commented, and readable.

CASE STUDY:

The Lahore Smart City Authority is implementing a new analytics system to improve traffic flow, reduce congestion, and identify high■risk accident zones.

You are hired as a Data Analyst to process raw IoT and surveillance-based traffic data.

DATASET:

traffic_logs = [

"id:501,zone:A1,vehicle:Car,speed:62,time:08:30,violations:[None],status:Smooth",

"id:502,zone:A1,vehicle:Bike,speed:85,time:09:10,violations:[Helmet],status:Busy",

"id:503,zone:B2,vehicle:Bus,speed:45,time:17:25,violations:[None],status:Smooth",

"id:504,zone:C3,vehicle:Car,speed:110,time:14:15,violations:[Overspeed],status:Congested",

"id:505,zone:A1,vehicle:Truck,speed:40,time:18:50,violations:[None],status:Smooth"

]

QUESTIONS:

Q1: Convert each traffic log into structured Python data with validation.

Q2: Calculate average speed per zone.

Q3: Determine peak hour (hour with highest traffic entries).

Q4: Find vehicles with speed > 80 km/h.

Q5: Count occurrences of each violation type.

Q6: Compute safety index for each zone.

Q7: Create a summary for each vehicle category.

Q8: Identify high-congestion zones.

Q9: Classify each log into time windows (Morning, Afternoon, Evening, Night).

Q10: Generate final zone-level report (vehicles, avg speed, violations, common vehicle type, safety category).