

REPORT BCG:

Instruction:

- 1) Provide necessary input to config file under 'Config folder'. File contain details on data, data path and result path info.
- 2) Run main_analysis.py program. Based on the configuration and the data provided, it will do required analysis
- 3) All analysis will be saved in result.txt file, under Output folder.

Source Data:

No	Table Name	Count
1	Charges_use	116110
2	Damages_use	24950
3	Endorse_use	159818
4	Primary_Person_use	156954
5	Restrict_use	159852
6	Units_use	173499

Analysis:

1. Analytics 1: Find the number of crashes (accidents) in which number of persons killed are male?

Number of males killed in crashes: 182

2. Analysis 2: How many two wheelers are booked for crashes?

a) Total 2-wheeler crashes count: 911

b) Distinct 2-wheeler crashes count (based on VIN): 745

3. Analysis 3: Which state has highest number of accidents in which females are involved?

Analysis shows Texas state with the highest number of female accidents.

4. Analysis 4: Which is the Top 5th to 15th VEH_MAKE_IDs that contribute to a largest number of injuries including death?

Top 5th to 15th VEH_MAKE_IDs: NISSAN, HONDA, GMC, HYUNDAI, KIA, JEEP, CHRYSLER, MAZDA, VOLKSWAGEN, PONTIAC, LEXUS

5. Analysis 5: For all the body styles involved in crashes, mention the top ethnic user group of each unique body style?
- a) For body BUS style max occurrence in BLACK ethnicity.
 - b) For body PASSENGER CAR, 4-DOOR style max occurrence in WHITE ethnicity.
 - c) For body FIRE TRUCK style max occurrence in WHITE ethnic
 - d) For body VAN style max occurrence in WHITE ethnicity.
 - e) For body PICKUP style max occurrence in WHITE ethnicity.
 - f) For body SPORT UTILITY VEHICLE style max occurrence in WHITE ethnicity.
 - g) For body FIRE TRUCK style max occurrence in WHITE ethnicity.
 - h) For body TRUCK style max occurrence in WHITE ethnicity.

6. Analysis 6: Among the crashed cars, what are the Top 5 Zip Codes with highest number crashes with alcohols as the contributing factor to a crash (Use Driver Zip Code)

Top 5 zip codes are 78521, 76010, 79936, 79938, 78741

7. Analysis 7: Count of Distinct Crash IDs where No Damaged Property was observed and Damage Level (VEH_DMAG_SCL~) is above 4 and car avails Insurance?

Count of Distinct Crash IDs: 6138

8. Analysis 8: Determine the Top 5 Vehicle Makes where drivers are charged with speeding related offences, has licensed Drivers, uses top 10 used vehicle colours and has car licensed with the Top 25 states with highest number of offences (to be deduced from the data)
- Top 5 Vehicle Makes based on the data analysis are – FORD, CHEVROLET, TOTOTA, DODGE, HONDA

Note:

Name	File Location
Configuration	\BCG\Config\config.json
Raw output data	\BCG\Output\report.txt
Data	\BCG\Data*.csv