

GET ME
THE DATABASE

////// START //////////////////////////////////////

☒ Column Types

☒ Number of Rows

☒ Character Types

String/Char
- INT, NUM
- BOOLEAN

CSV!!!

////// END //////////////////////////////////////

for
[B] or [2]

Two Ways to....

Analyze the
following
DATA SET

[B]

PERCENTAGE of predictor for possible injuries

also RATE of total % that people are likely to encounter injuries
or EACH of their Speed limit zones BASED on this database sheet

[2]

Make a list

Split columns

Two things

①

Representation of
Speed Limit Zone

(Int)

②

Calculate mean & %
of total amount of injuries
for each Speed limit zone

Sum
+
Value_counts

①

[A]

INJURY & Speed Limit Zones

Question: What are the Top 5 injury types for each Speed limit zone?

CLEAN the
DATASET

①

Check for any
duplicate strings

②

if duplicate
strings replace

NO DATA

Group the following
total populated incidents
for each Speed limit zone

③

MAKE A STATEMENT

Value_counts()

& Sum()

Make a new chart
that lists two columns
Summarize

dataframe

from here

UPPER
case

for each
string type