Double-click (or enter) to edit

```
import numpy as np
import pandas as pd

file_path = '/content/RS_246.csv'

road_data = pd.read_csv(file_path)
road_data.head()
```

	Unnamed: 0	Unnamed: 1	Distribution of road traffic deaths by type of road user (%)	Distribution of road traffic deaths by type of road user (%).1	Distribution of road traffic deaths by type of road user (%).2	Distribution of road traffic deaths by type of road user (%).3	Distri o t dea type o user
0	Countries, territories and areas	Year	Drivers/ passengers of 4-wheeled vehicles	Drivers/ passengers of motorized 2- or 3-wheelers	Cyclists	Pedestrians	uns roa
1	Albania	2016	39.4	11.9	7.8	38.7	
2	Andorra	2013	NaN	50.0	NaN	50.0	

Next steps: View recommended plots

```
#Check the missing datas
road_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149 entries, 0 to 148
Data columns (total 7 columns):
 #   Column                                                        Non-Null Count
---  ------                                                        --------------
 0   Unnamed: 0                                                    149 non-null
 1   Unnamed: 1                                                    149 non-null
 2   Distribution of road traffic deaths by type of road user (%)    135 non-null
 3   Distribution of road traffic deaths by type of road user (%).1  129 non-null
 4   Distribution of road traffic deaths by type of road user (%).2  127 non-null
 5   Distribution of road traffic deaths by type of road user (%).3  137 non-null
 6   Distribution of road traffic deaths by type of road user (%).4  132 non-null
dtypes: object(7)
memory usage: 8.3+ KB
```

```
road_data.isnull().sum()
```

```
Unnamed: 0                                                      0
Unnamed: 1                                                      0
Distribution of road traffic deaths by type of road user (%)   14
Distribution of road traffic deaths by type of road user (%).1  20
```

```
Distribution of road traffic deaths by type of road user (%).2    22
Distribution of road traffic deaths by type of road user (%).3    12
Distribution of road traffic deaths by type of road user (%).4    17
dtype: int64
```

```
np.mean(road_data)
```

```
/usr/local/lib/python3.10/dist-packages/numpy/core/fromnumeric.py:3502: FutureWarning
  return mean(axis=axis, dtype=dtype, out=out, **kwargs)
/usr/local/lib/python3.10/dist-packages/numpy/core/fromnumeric.py:3502: FutureWarning
  return mean(axis=axis, dtype=dtype, out=out, **kwargs)
Series([], dtype: float64)
```

```
road_data.drop([0], axis = 0, inplace = True)
road_data
```

```
updated_df = road_data
updated_df['Drivers/passengers of 4-wheeled vehicles']=updated_df['Drivers/passengers of
```

```
road_data.fillna(road_data.mean(), inplace = True)
road_data
```

Next steps: View recommended plots

```
#road_data.fillna(road_data.select_dtypes(np.number).mean(), inplace = True)
road_data.head()
```

Next steps: View recommended plots

```
updated_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 1 to 148
Data columns (total 7 columns):
 #   Column                                                          Non-Null Count
---  ------                                                          --------------
 0   Unnamed: 0                                                      148 non-null
 1   Unnamed: 1                                                      148 non-null
 2   Distribution of road traffic deaths by type of road user (%)    134 non-null
 3   Distribution of road traffic deaths by type of road user (%).1  128 non-null
 4   Distribution of road traffic deaths by type of road user (%).2  126 non-null
 5   Distribution of road traffic deaths by type of road user (%).3  136 non-null
 6   Distribution of road traffic deaths by type of road user (%).4  131 non-null
dtypes: object(7)
memory usage: 8.2+ KB
```

```
updated_df = road_data
updated_df['Distribution of road traffic deaths by type of road user (%)']=updated_df['Di
updated_df.info()
```

```
#value_1 = {'Distribution of road traffic deaths by type of road user (%)':148}

updated_df = road_data.fillna(road_data['Distribution of road traffic deaths by type of r
updated_df
```

```
road_data.mean()
```

```
<ipython-input-48-c3a1fcaf6caf>:1: FutureWarning: The default value of numeric_only i
  road_data.mean()
Series([], dtype: float64)
```

```
road_data.fillna(0.)
```

```
road_data.mean()
```

```
<ipython-input-51-c3a1fcaf6caf>:1: FutureWarning: The default value of numeric_only i
  road_data.mean()
Series([], dtype: float64)
```

```
road_data.drop([], axis=1, inplace=True)
road_data
```

```
column_mean = road_data['Distribution of road traffic deaths by type of road user (%)'].m
road_data['Distribution of road traffic deaths by type of road user (%)'].fillna(column_m
```

```
road_data
```

Next steps: View recommended plots

```
road_data.drop(road_data.index[0], axis = 0, inplace=True)
road_data
```

```
column_mean = road_data['Distribution of road traffic deaths by type of road user (%).3']
road_data['Distribution of road traffic deaths by type of road user (%).3'].fillna(column
```

```
road_data
```

```
road_data['Distribution of road traffic deaths by type of road user (%)']=pd.to_numeric(r
road_data
```

```
road_data[40:50]
```