



+ Code + Text

Connect ▼

Colab AI



```
age_band_of_casualty      0
casualty_severity         0
pedestrian_location       0
pedestrian_movement      0
car_passenger            0
bus_or_coach_passenger    0
pedestrian_road_maintenance_worker 0
casualty_type             0
casualty_home_area_type   0
casualty_imd_decile       0
lsoa_of_casualty         0
dtype: int64
```

```
[ ] df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
collision_year	62674.0	2023.000000	0.000000	2023.0	2023.0	2023.0	2023.0	2023.0
vehicle_reference	62674.0	1.467562	4.849047	1.0	1.0	1.0	2.0	992.0
casualty_reference	62674.0	1.374972	1.551044	1.0	1.0	1.0	1.0	70.0
casualty_class	62674.0	1.491177	0.737688	1.0	1.0	1.0	2.0	3.0

df.head()

	collision_index	collision_year	collision_reference	vehicle_reference	casualty_reference	casualty_class	sex_of_casualty	age_of_casualty	age
0	2023010419171	2023	010419171	1	1	3	2	20	
1	2023010419183	2023	010419183	2	1	1	1	25	
2	2023010419183	2023	010419183	3	2	2	2	38	
3	2023010419189	2023	010419189	1	1	1	1	50	
4	2023010419191	2023	010419191	2	1	1	1	34	

[] df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 62674 entries, 0 to 62673
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   collision_index        62674 non-null  object
1   collision_year         62674 non-null  int64
```

1	2023010419183	2023	010419183	2	1	1	1	25	
2	2023010419183	2023	010419183	3	2	2	2	38	
3	2023010419189	2023	010419189	1	1	1	1	50	
4	2023010419191	2023	010419191	2	1	1	1	34	

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 62674 entries, 0 to 62673
Data columns (total 19 columns):
#   Column              Non-Null Count  Dtype
---  -
0   collision_index      62674 non-null  object
1   collision_year       62674 non-null  int64
2   collision_reference  62674 non-null  object
3   vehicle_reference   62674 non-null  int64
4   casualty_reference  62674 non-null  int64
5   casualty_class      62674 non-null  int64
6   sex_of_casualty     62674 non-null  int64
7   age_of_casualty     62674 non-null  int64
8   age_band_of_casualty 62674 non-null  int64
9   casualty_severity   62674 non-null  int64
```



+ Code + Text

Connect ▾

Colab AI



pedestrian_road_maintenance_worker	62674.0	0.037208	0.284849	-1.0	0.0	0.0	0.0	2.0
casualty_type	62674.0	9.507643	16.671019	-1.0	1.0	9.0	9.0	98.0
casualty_home_area_type	62674.0	1.049925	0.932578	-1.0	1.0	1.0	1.0	3.0
casualty_imd_decile	62674.0	4.259980	3.221955	-1.0	2.0	4.0	7.0	10.0

```
df[df.columns[1]].dtype == 'int64'
```

```
True
```

```
[ ] for col in df.columns:
    if df[col].dtype == 'object':
        print(col)
```

```
collision_index
collision_reference
lsoa_of_casualty
```

```
[ ] for col in df.columns:
    print(f'{col} : {df[col].value_counts().size}')
```

