```
Cases:
Males 411835 , 62.00 %
Females 249940 , 37.00 %
Not fair
Chi-squared: (4339.7395645571305, 0.0, 3, array([[348214.10812587,  52199.47559971,   7051.49391409,
         4369.92236032],
       [211328.89187413,  31679.52440029,   4279.50608591,
         2652.07763968]]))

Original data set tuples: 1048575
Final data after scrubbing: 661775

The original size of the data contains about a million than the new size.
Which concludes the data scrubbing is complete.
After performing an analysis approach, all analysis use categorical or ordinal values.
The non-parametric techniques will be used.
The statistical analysis base on both gender results, males have higher percentage than females in vehicle
crashes
62% for males, whereas 37% for females.
The ratio for the binomial test of observed frequencies of male versus female drivers in accidents are 1.65
M : 1 F
The comparison of the results between binomial test and ALPHA is not fair.
```

Name	Type	Size	Value
A	str	16	Seating Position
B	str	3	Sex
binomTest	float64	1	0.0
C	str	15	Injury Severity
caseAge	Series	(661775,)	Series object of pandas.core.series module
caseGen	Series	(661775,)	Series object of pandas.core.series module
caseSev	Series	(661775,)	Series object of pandas.core.series module
D	str	3	Age
fem_arr	list	4	[203426, 40268, 3735, 2511]
fem_cal	int	1	37
frame	DataFrame	(661775, 4)	Column names: Seating Position, Gender, Injury Severity, Age
i	int	1	661774
labels	list	4	['Seating Position', 'Sex', 'Injury Severity', 'Age']
male_arr	list	4	[356117, 43611, 7596, 4511]
male_cal	int	1	62
new_size	int	1	661775
num_female	int	1	249940
num_male	int	1	411835
seatPos	Series	(661775,)	Series object of pandas.core.series module
sev_arr	list	2	[[356117, 43611, 7596, 4511], [203426, 40268, 3735, 2511]]
sev_test	tuple	4	(4339.7395645571305, 0.0, 3, Numpy array)
size	int	1	1048575