1.Checking for missing value in the dataset. No missing value found in the dataset.

2.Data Cleaning: (i) Converting the typos ‘ObGyn’ of Specialty column into ‘OBGYN’.

(ii)Converting the typos ‘Unknown’ of Insurance column into ‘unknown’.

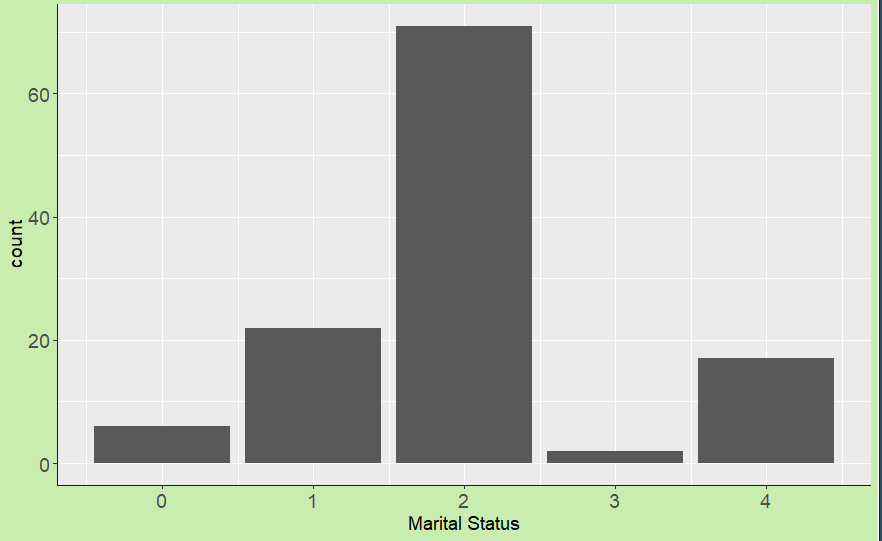
A screenshot of a cell phone

Description automatically generated3.Gender distribution:

Gender statistics (Total : number of cases, pay: Payment ,Med : median ,Mean :mean, SD: Standard Deviation, VAR : Variation )

A screenshot of a cell phone

Description automatically generated

4. Marital Status :

0: divorced , 1:single ,2 :married ,3: widowed ,4 : unknown

60.16% is married and 14.40 % is unknown.

A close up of a sign

Description automatically generated

( Total : number of cases, pay: Payment ,Med : median ,Mean : mean, SD: standard deviation, VAR: variance)

A screenshot of a cell phone

Description automatically generated5.Specialty:

A white sign with black text

Description automatically generated

(Total : number of cases, pay: Payment ,Med : median ,Mean : mean, SD: standard deviation, VAR: variance)

NaN,NA because they consists of 1 value.

The top 5 Specialty consists of 67% of the Payment but they are more number of cases.

6. Insurance:

A screenshot of a cell phone

Description automatically generated

A black sign with white text

Description automatically generated

Workers Compensation have high median but they have there are only 3 cases in the dataset. 30.50% of the insurance is unknown.

A screenshot of a cell phone

Description automatically generated

A close up of a logo

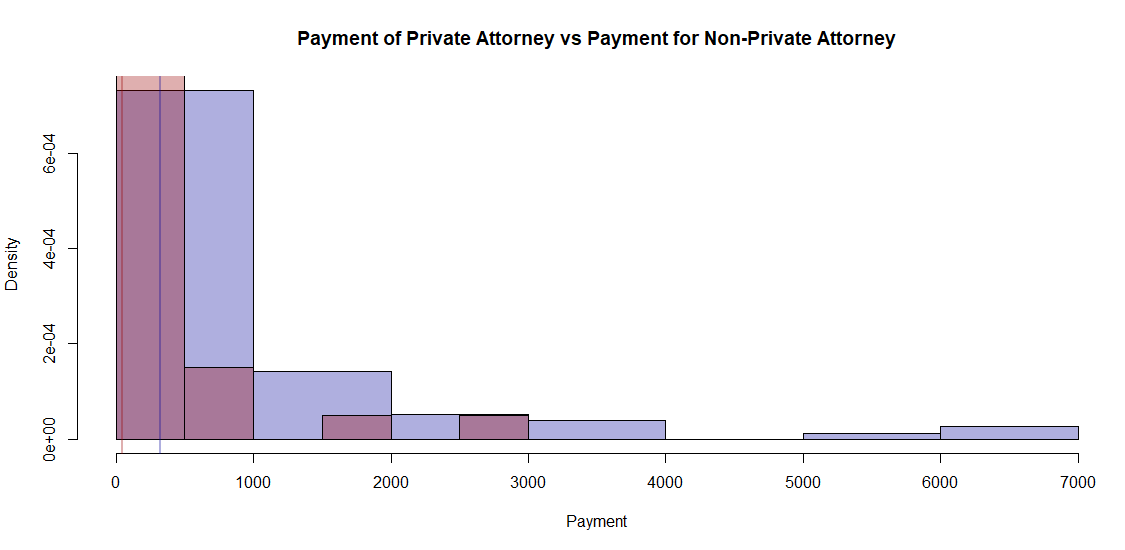
Description automatically generated7.Private Attorney:

A picture containing road, bottle

Description automatically generated

In case of Payment, Private is way more than non-private. If we take all private like Insurance and Private Attorney then it is more than normal.

In case of Payment, Private attorney is influencing more than Private Insurance and Private Insurance have some large values.

Histogram of Private(Private Attorney , Private Insurance) vs all the values in the dataset.

A screenshot of a cell phone

Description automatically generated

The blue vertical line shows the median of the the private while the red vertical line shows the median of the payment of all values.

A screenshot of a cell phone

Description automatically generated8. Severity:

A close up of a keyboard

Description automatically generated

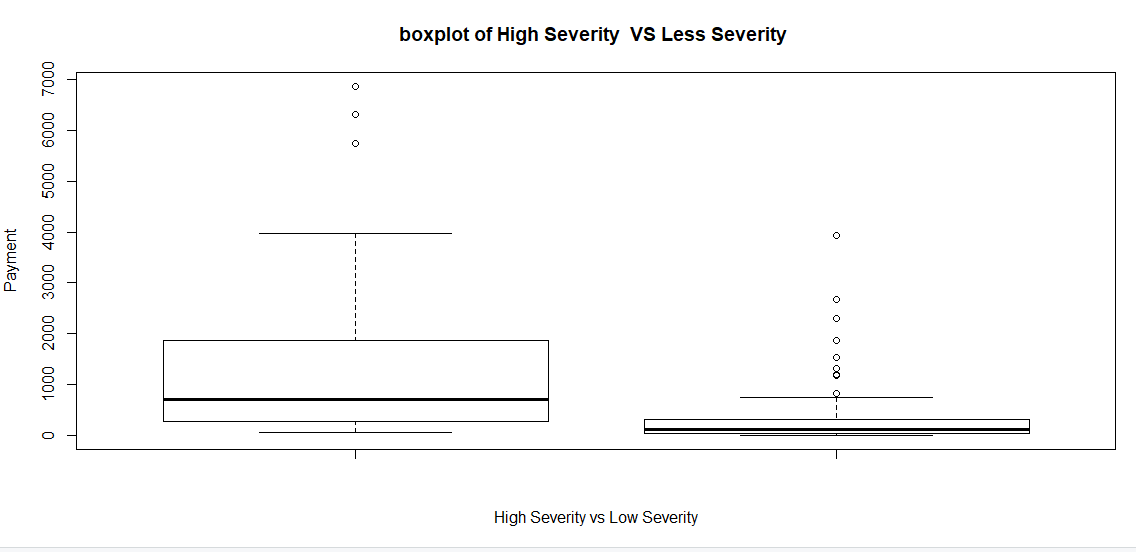
Grouping high severity (from 6-9) (Significant damage to death).

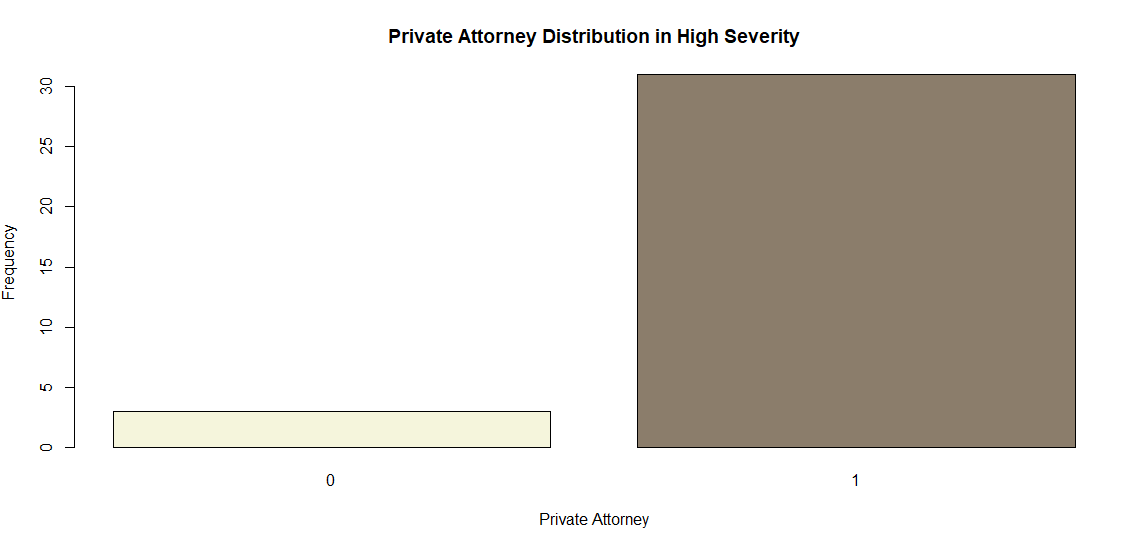
Majority of the high severity (from 6-9) are done by Private Attorney.

A picture containing electronics

Description automatically generated

The high severity (from significant damage to death) constitute 63.92 % of the payment even though low severity is more in numbers. High severity leads to more payment.

91.17 % of the cases of high severity is done by Private attorney.



9.Age:

A close up of a map

Description automatically generated

We are grouping the age according to particular age.

Age<18:



Age>=18 and Age <40:

Age>=40 and Age <60:

A picture containing outdoor, road

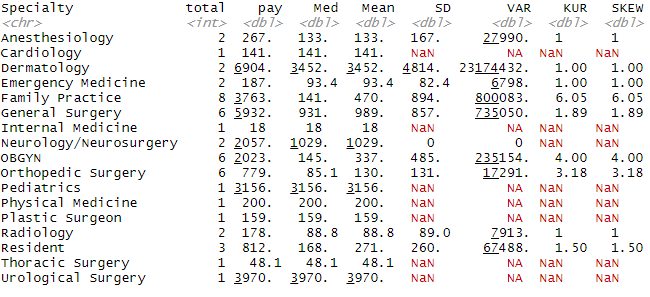
Description automatically generatedAge >=60:

Age with Specialty:

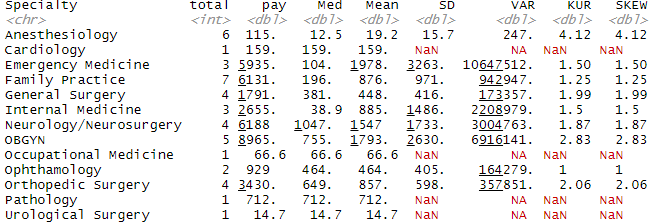
A black sign with white text

Description automatically generatedAge <18:

Age>=18 and Age <40:



Age >=40 and Age <60:

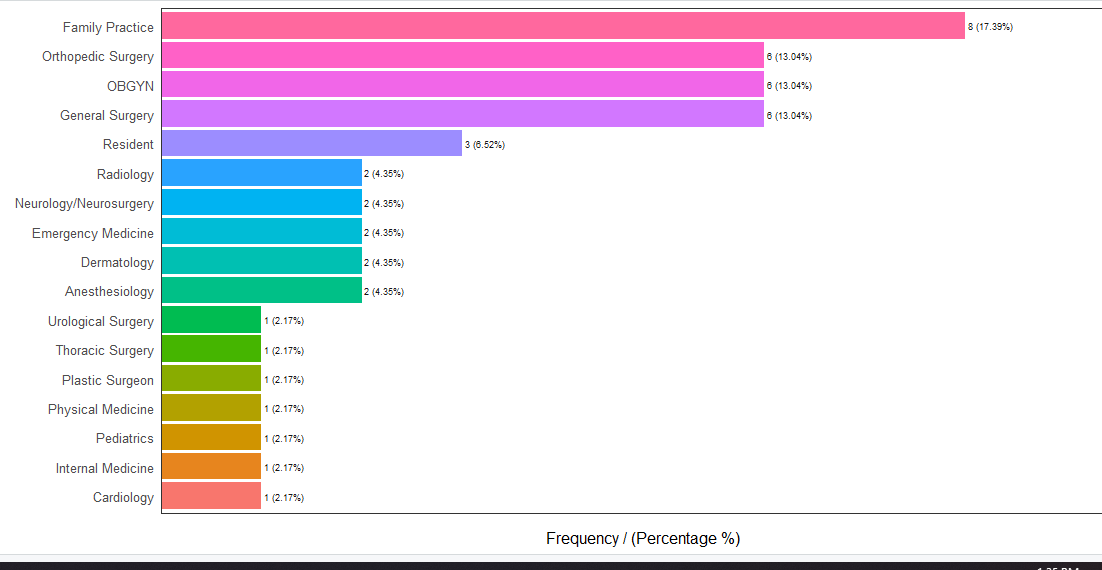


Age>=60:

A black and silver text on a screen

Description automatically generated

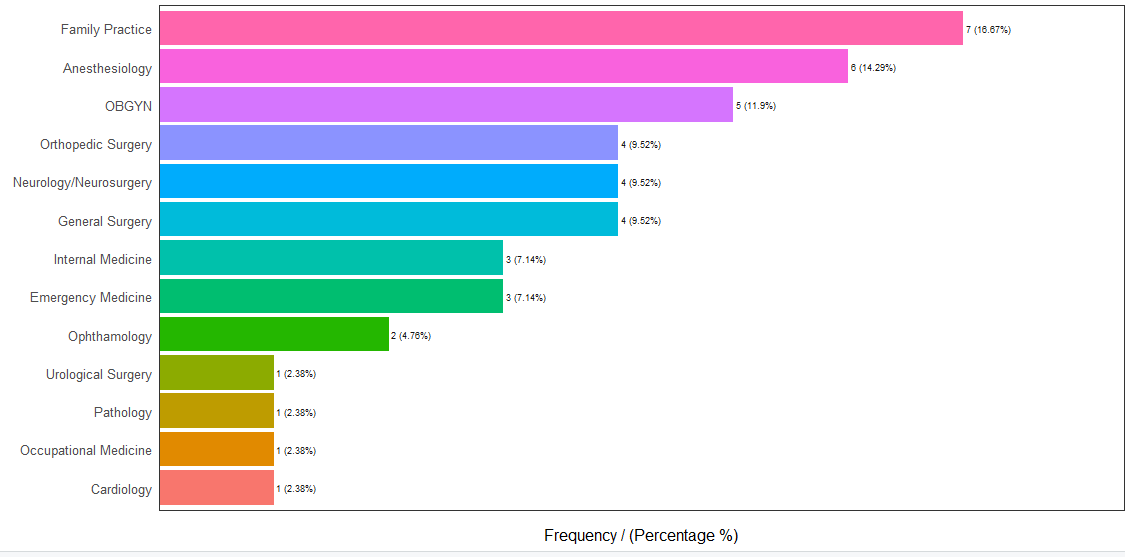
Age>=18 and Age <40



A screenshot of a cell phone

Description automatically generatedAge<18:

Age >=40 and Age <=60:



Age >60:

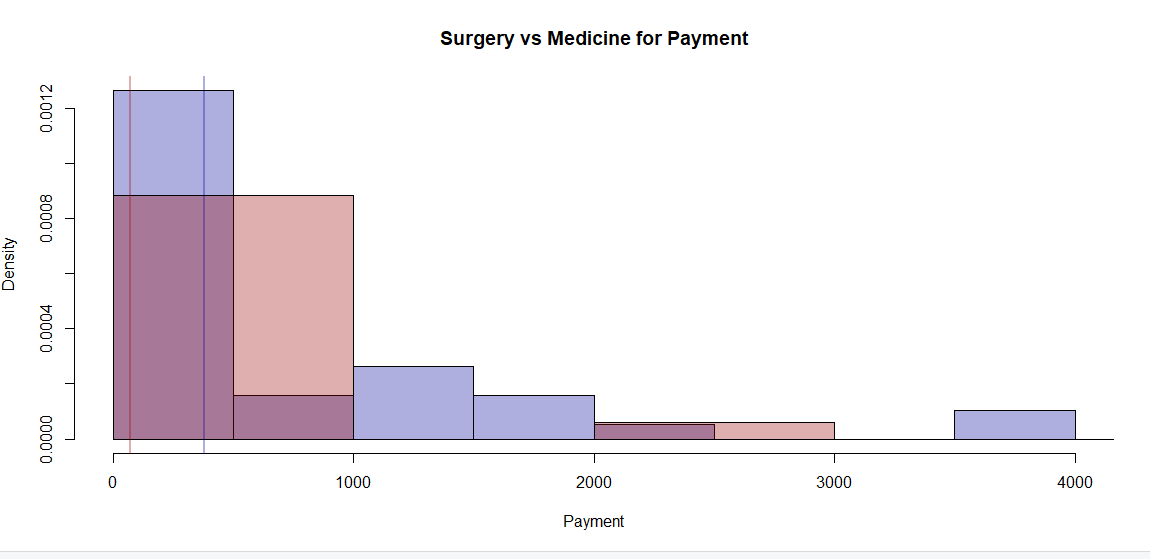
A screenshot of a cell phone

Description automatically generated

10. In Specialty , Specialties like Surgery(General Surgery, Orthopedic Surgery, Urological Surgery etc) have higher payment while Specialties like Medicine (Internal Medicine, Physical Medicine, Occupational Medicine etc) have low payment.

Surgery : Median : 379.25 , Mean : 734.53

Medicine: Median : 77.7 , Mean : 594.8

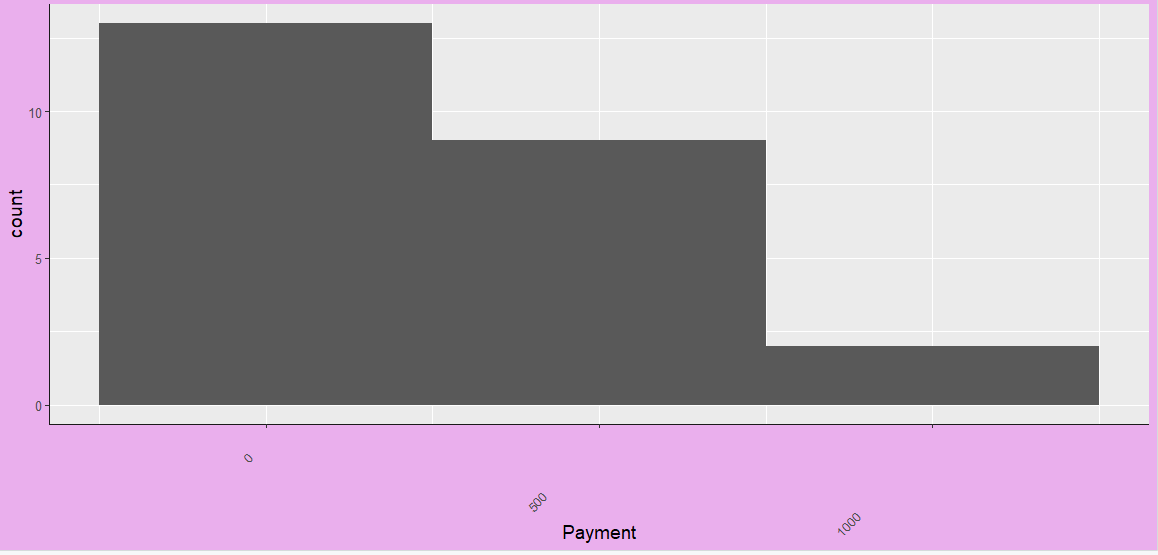


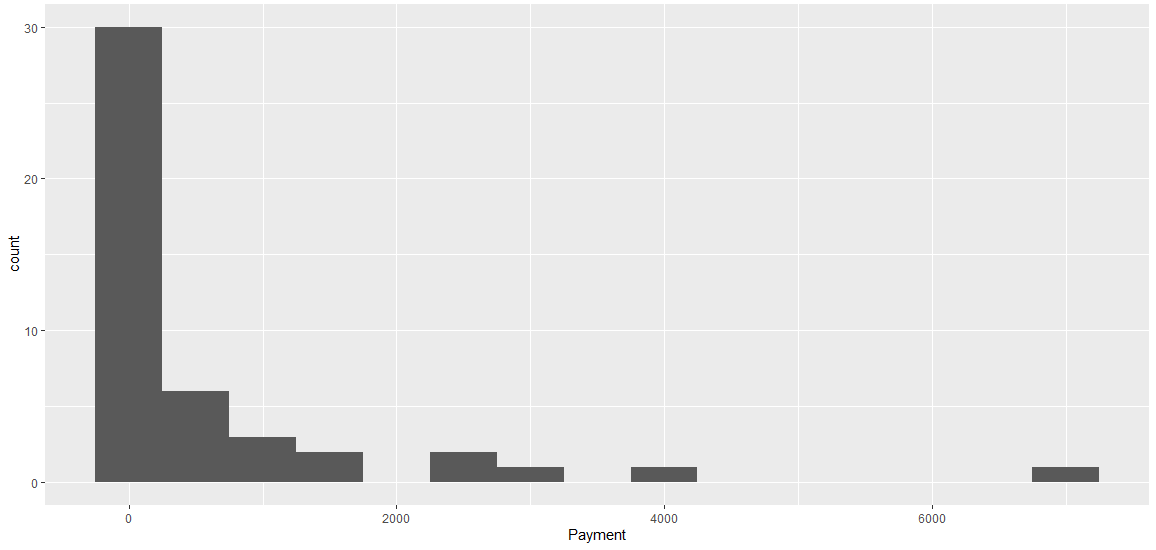
A picture containing map, text

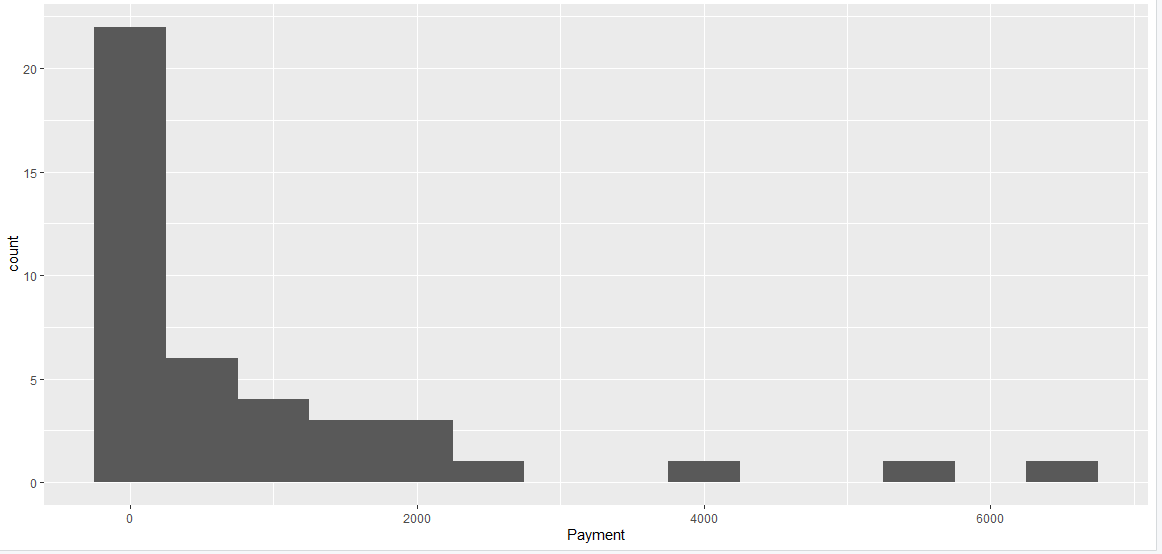
Description automatically generated11. Correlation plot :

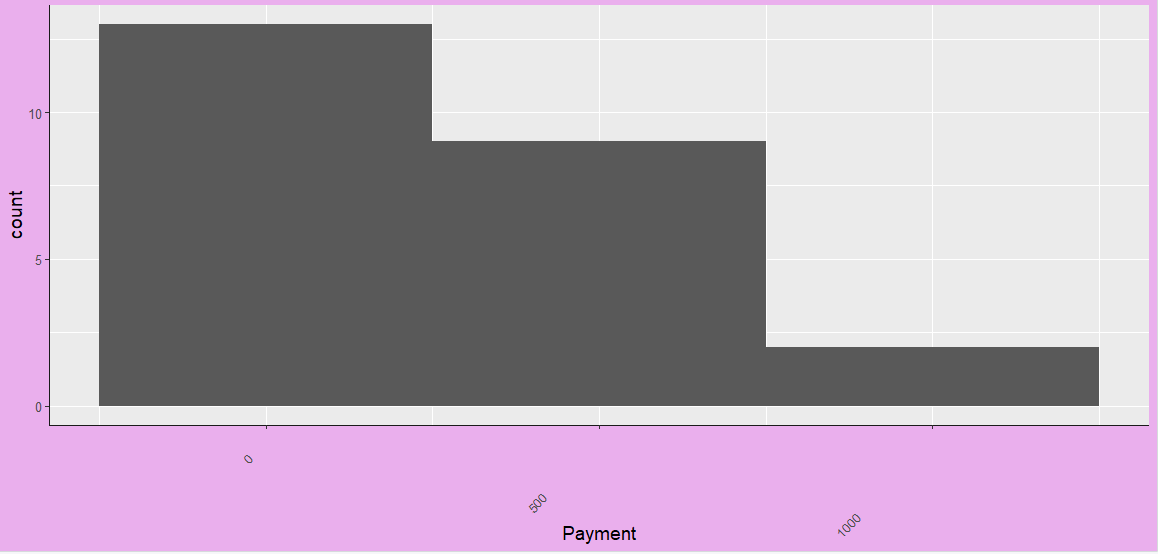
12. High severity is found in upper age groups and most of the higher severity cases are carried by Private Attorney .It also leads to high payment. In upper age group, Specialties like General Surgery and Anesthesiology are most common. In case of General Surgery , the severity is more in upper age groups and it also leads to high payment while in case of Anesthesiology, the severity is less and majority is from Non-Private and Payment is also less.

13. Age groups with Payment:

Age <18 :

Age >=18 and Age <40:

Age>=40 and Age <60:

Age >=60: