Homework 1

Answers:

1. Amount: Continuous numerical

Severity: Ordinal categorical

Age: Discrete numerical

Private Attorney: Nominal categorical

Marital Status: Nominal categorical

Specialty: Nominal categorical

Insurance: Nominal categorical

Gender: Nominal categorical

1. For Amount:
2. The distribution is right skewed.
3. The mean is 91044.915, the standard deviation is 164501.80.
4. The quartiles are 7500, 22750, 92670.
5. The IQR is 92670-7500=85170. It is obvious to discover about a dozen extreme values above 200000.
6. For Age:
7. The distribution is roughly symmetric.
8. The mean is 42.830508, the standard deviation is 18.004418
9. The quartiles are 31, 41.5, 56.25.
10. The IQR is 56.25-31=25.25. The reasonable range covers all age data and there is no extreme values.
11. It seems that there is no linear relationship between age and amount. Many of observations are far from fitted line, and the fitted line hardly slopes. But the observations are symmetrically distributed, most gathering between 20 and 60. High amounts gather in the middle. We guess that people between 20 and 60 have a higher possibility to have higher amounts.
12. The observations are left skewed, and the fitted line also slopes. It indicates that severity has a positive relationship with amount.
13. Private Attorney: Private, Not Private

Marital Status: Married, Single, Unknown, Divorced, Widowed

Specialty: Anesthesiology, Cardiology, Dermatology…, exactly as displayed in Frequency, Exhibit 1.8

Insurance: Medicare/Medicaid, Private, Unknown, No Insurance, Workers Compensation

1. 60.17% of data is female.
2. They are Medicare/Medicaid, Private, Unknown, No Insurance, Workers Compensation, counting 16(13.56%), 51(43.22%), 36(30.51%), 12(10.17%), 3(2.54%).
3. Anesthesiology is 11.0%, Dermatology is 1.7%, Orthopedic Surgery is 9.3%.
4. Divorced is 5.08%, Widowed is 1.69%.
5. Private Attorney and Marital Status seem to be related to the amount.
6. When claimant is Married the amount doesn’t tend to be higher.
7. The amounts are no higher for female than male.
8. Obviously, it’s related because the mean and the quartiles for Private Attorney is much higher than those of Not Private Attorney.
9. There is still a positive relationship between Severity and Amount, because we can see that fitted lines still slope up for male and female.
10. When controlling gender, there is still a relationship between Private Attorney and Amount. Private Attorney overtakes Not Private in mean and quartiles for both male and female. For a general term, when we include both variables in the same analysis, it is possible we could draw different conclusion, because big difference of statistical data caused by one variable (like there was no difference between Private Attorney for women, but big difference for men) will be potentially counted into the influence of other variables if we only look at single variables without removing the difference brought by this statistical variable.
11. It appears that Marital Status, Private Attorney, Severity and Age have relationships with Amount, and Private Attorney and Severity are the most related to the Amount. (Age tends to show that age between 20 and 60 gets higher amount. Marital Status tends to show that Divorced has far higher amount.) Private Attorney shows that Private Attorney has higher amount than Not Private Attorney, and Severity has a positive relationship with Amount. Private Attorney and Severity still have obvious relationships with Amount even after controlling Gender, so each is a univariate relationship with Amount. Because both influence Amount, the relationship between Private Attorney, Severity and Amount is a multivariate relationship. I