A screenshot of a graph

Description automatically generatedA graph on a black background

Description automatically generatedA screenshot of a graph

Description automatically generated

The three distributions you've shared appear to be histograms with overlaid kernel density estimates (KDE) for Payment, Lower Estimate, and Higher Estimate. Let's break down what these plots tell us:

1. Higher Estimate Distribution:

- Central Tendency: The peak is around 22,800, indicating that many of the higher estimates fall in this range.

- Spread: There's a long tail to the right, showing some higher values, but most higher estimates are concentrated below 35,000.

- KDE Insight: The KDE suggests a smooth decrease in frequency as the estimates increase, with few very high estimates.

2. Lower Estimate Distribution:

- Central Tendency: The lower estimates peak around 18,400. This indicates that most lower estimates tend to fall within the range of 18,000-20,000.

- Spread: There’s a sharp drop-off after 20,000, and very few values go beyond 32,000.

- KDE Insight: The KDE here shows a skewed distribution with a long right tail, indicating that while most estimates are on the lower side, some much higher values exist.

3. Payment Distribution:

- Central Tendency: Payments are concentrated around 27,500, with most values falling between 20,000 and 30,000.

- Spread: Like the previous two, there’s a sharp right tail, indicating higher payments for fewer instances.

- KDE Insight: The KDE mirrors the payment's histogram, showing a peak in the range of 20,000 to 30,000, then gradually decreasing with a few high outliers.

Summary of Insights:

- All three distributions have a positive skew, meaning there are more instances of lower or mid-range values, but a few significantly higher values (long right tail).

- The higher estimate values are expectedly larger than the lower estimates, with payments also showing a similar pattern but with a slightly higher concentration around 27,000.

- The KDE overlay helps to smooth out the trends and visualize the likelihood of values, indicating where most data points lie (near the peaks).

These distributions can suggest that while most estimates and payments are concentrated in a mid-range, there are outliers or exceptional cases with significantly higher values. The positive skew in all three suggests infrequent high payments or estimates.

1. General Trends and Concentration:

- Higher Estimate peaks around 22,800, Lower Estimate peaks around 18,400, and Payments peak around 27,500. This shows that the majority of the estimates and payments tend to cluster within these ranges.

- The Higher Estimate has the largest central value, followed by Payments, and then Lower Estimate, indicating that the expected upper range of estimates is predictably higher than the lower estimates and the actual payments.

2. Distribution Skewness:

- All three distributions are positively skewed (right-skewed), meaning that while the majority of values are in the lower to middle ranges, there are fewer but significantly higher estimates and payments.

- This indicates a typical pattern in cost or payment-related data where most transactions occur at a lower or average level, but there are occasional larger outliers.

3. Comparing Estimate Ranges with Actual Payments:

- The Higher Estimate and Payment distributions overlap somewhat, but the higher estimates have a broader range. This suggests that payments are often lower than the upper-end estimates.

- The Lower Estimate distribution has a sharper decline and suggests that most estimates are on the lower end, but actual Payments are distributed slightly above these lower estimates, indicating that final payments tend to be closer to higher estimates rather than the lower estimates.

4. Frequency of Payments and Estimates:

- The Higher Estimate and Payment distributions are relatively wider compared to the Lower Estimate, which is more concentrated. This implies that while higher estimates and payments can vary greatly, lower estimates tend to have less variation and are more predictable.

5. Outliers and Long-Tail Behavior:

- All distributions show a long tail to the right, meaning that there are a few cases with significantly high payments and estimates. These could be outliers or instances of special cases where the costs or payments are much higher than usual.

- This behavior is common in cost data, where a small number of transactions or estimates are disproportionately large.

Potential Insights for Decision-Making:

1. Budget Planning: Since most payments fall between 20,000 and 30,000, budget allocations can be made to cover these typical payment amounts. However, a contingency plan should be considered for the rare, higher-payment cases (long-tail outliers).

2. Estimate Accuracy: The closeness of the Higher Estimate distribution to actual Payments suggests that the higher estimates are often more accurate. However, the lower estimates are more conservative and may not capture the full scope of the payment distribution.

3. Risk Assessment: The positive skew and long tail in the distributions suggest that there are risks of higher-than-expected payments or estimates. A strategy to manage these outliers could involve further analysis to identify patterns that lead to these extreme values.

4. Cost Optimization: The concentration of payments around 27,500 suggests that efforts can be made to optimize payment structures within this range, while analyzing the causes of payments above this level could reveal areas for cost control.

5. Pricing Strategy: If this data is related to healthcare or service costs, the insights suggest that pricing models might be refined to better reflect the median values of payments, with flexibility for handling outliers.