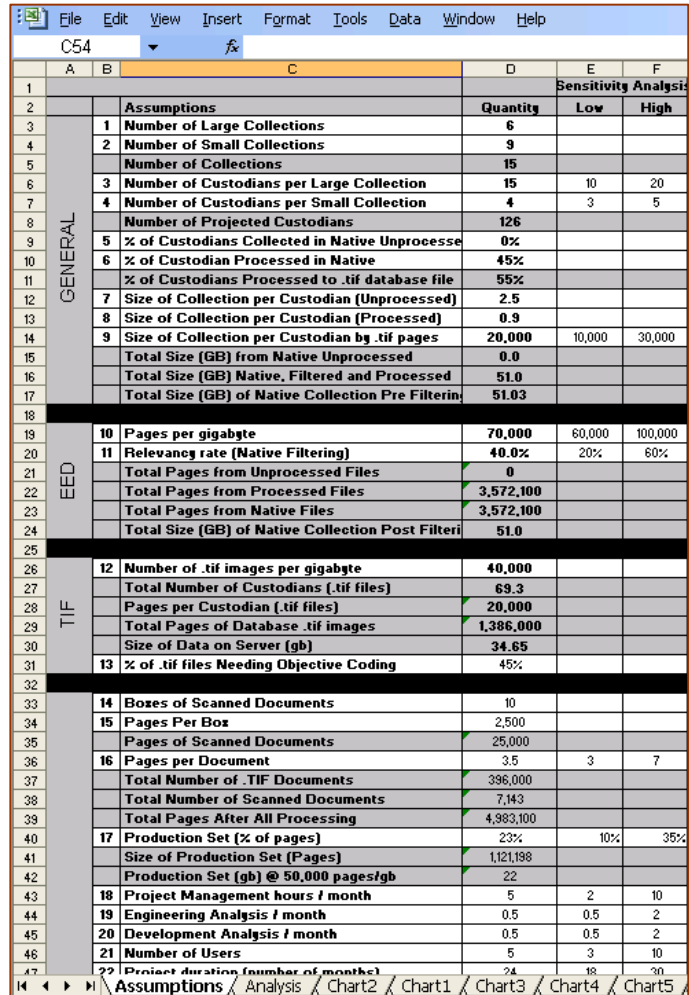


Case Study: Comparing Electronic Discovery Responses to Request For Proposals (RFP)

The exercise of determining proposed pricing in an RFP process is akin to buying a mattress; it is difficult to compare costs because no two bids have similar pricing schemes. In order to most effectively consider the bids it is important to place all proposed pricing against the same set of assumptions. How many gigabytes to be harvested, what will be the size of the resulting database set, what percentage will be deemed responsive, how many months will the data be hosted are just a few of the assumptions often made. Essentially, if you have four proposals, every variable in the pricing must be treated as a separate assumption.

Unfortunately, all too often a low bid provider is selected only to learn that critical assumptions in the winning response were lower vis-à-vis other proposals. In other words, vendor A assumes that there will be 50,000 pages TIFFED per responsive gigabyte while vendor B assumes 70,000. Everything else equal, vendor A will have a lower cost bid and an unfair advantage. To avoid this pitfall, I propose using sensitivity analysis to compare the responses under a controlled set of assumptions.



	A	B	C	D	E	F
1						
2			Assumptions	Quantity	Low	High
3		1	Number of Large Collections	6		
4		2	Number of Small Collections	9		
5			Number of Collections	15		
6		3	Number of Custodians per Large Collection	15	10	20
7		4	Number of Custodians per Small Collection	4	3	5
8			Number of Projected Custodians	126		
9		5	% of Custodians Collected in Native Unprocesse	0%		
10		6	% of Custodian Processed in Native	45%		
11			% of Custodians Processed to .tif database file	55%		
12		7	Size of Collection per Custodian (Unprocessed)	2.5		
13		8	Size of Collection per Custodian (Processed)	0.9		
14		9	Size of Collection per Custodian by .tif pages	20,000	10,000	30,000
15			Total Size (GB) from Native Unprocessed	0.0		
16			Total Size (GB) Native, Filtered and Processed	51.0		
17			Total Size (GB) of Native Collection Pre Filterin	51.03		
18						
19		10	Pages per gigabyte	70,000	60,000	100,000
20		11	Relevancy rate (Native Filtering)	40.0%	20%	60%
21			Total Pages from Unprocessed Files	0		
22			Total Pages from Processed Files	3,572,100		
23			Total Pages from Native Files	3,572,100		
24			Total Size (GB) of Native Collection Post Filteri	51.0		
25						
26		12	Number of .tif images per gigabyte	40,000		
27			Total Number of Custodians (.tif files)	69.3		
28			Pages per Custodian (.tif files)	20,000		
29			Total Pages of Database .tif images	1,386,000		
30			Size of Data on Server (gb)	34.65		
31		13	% of .tif files Needing Objective Coding	45%		
32						
33		14	Boxes of Scanned Documents	10		
34		15	Pages Per Box	2,500		
35			Pages of Scanned Documents	25,000		
36		16	Pages per Document	3.5	3	7
37			Total Number of .TIF Documents	396,000		
38			Total Number of Scanned Documents	7,143		
39			Total Pages After All Processing	4,983,100		
40		17	Production Set (% of pages)	23%	10%	35%
41			Size of Production Set (Pages)	1,121,198		
42			Production Set (gb) @ 50,000 pages/gb	22		
43		18	Project Management hours / month	5	2	10
44		19	Engineering Analysis / month	0.5	0.5	2
45		20	Development Analysis / month	0.5	0.5	2
46		21	Number of Users	5	3	10
47		22	Project duration (number of months)	24	18	30

1.0 Project Description

Law firm XYZ, (hereinafter Law Firm) represents a consortium of 15 financial institutions in a civil law suit. Each of the banks will be conducting its own collection and document review. The law firm representing the consortium will be responsible for reviewing the responsive documents from all 15 institutions. The project includes assumptions detailing the type and volume of electronic documents collected by the various parties. Law firm believes that some banks will be primary participants and will have large collections, while others will have a secondary role and small collections of data.

Due to the large number of unknowns, Law Firm requires an RFP be sent to four electronic discovery vendors for consideration. The winning bid must include pricing for data filtering, electronic processing, scanning, coding, promotion to TIFF, online review and document production. To handle the disparate data collections, the winning provider must demonstrate superior project management capability. An analysis of the pricing schemes is needed in order to best prepare for a final analysis.

1.1 Constraints

Bids will be received and each responding vendor will be given 60 minutes to present their solution. Attorneys for the consortium will view each presentation. No other coordination meetings are planned nor will there be any meetings between the 15 financial institutions regarding data collections. The data to be managed will be known only upon receipt.

1.2 Initial Analysis

Four responses to the RFP were received. Along with the bids, each vendor provided an estimated cost for work to be performed. On their face, the proposed costs for all the bids were within general proximity, with no one bid standing out.

In an effort to place the bids on equal footing, a set of assumptions were developed and the costs estimated accordingly. Twenty two assumptions (47 variables from assumptions) were determined as follows:

	Assumptions		Assumptions
1	Number of Large Collections	12	Number of .tif images per gigabyte
2	Number of Small Collections		Total Number of Custodians (.tif files)
	Number of Collections		Pages per Custodian (.tif files)
3	Number of Custodians per Large Collection		Total Pages of Database .tif images
4	Number of Custodians per Small Collection		Size of Data on Server (gb)
	Number of Projected Custodians	13	% of .tif files Needing Objective Coding
5	% of Custodians Collected in Native Unprocessed	14	Boxes of Scanned Documents
6	% of Custodian Processed in Native	15	Pages Per Box
	% of Custodians Processed to .tif database file		Pages of Scanned Documents
7	Size of Collection per Custodian (Unprocessed)	16	Pages per Document
8	Size of Collection per Custodian (Processed)		Total Number of .TIF Documents
9	Size of Collection per Custodian by .tif pages		Total Number of Scanned Documents
	Total Size (GB) from Native Unprocessed		Total Pages After All Processing
	Total Size (GB) Native, Filtered and Processed	17	Production Set (% of pages)
	Total Size (GB) of Native Collection Pre Filtering		Size of Production Set (Pages)
10	Pages per gigabyte		Production Set (gb)
11	Relevancy rate (Native Filtering)	18	Project Management hours / month
	Total Pages from Unprocessed Files	19	Engineering Analysis / month
	Total Pages from Processed Files	20	Development Analysis / month
	Total Pages from Native Files	21	Number of Users
	Total Size (GB) of Native Collection Post Filtering	22	Project duration (number of months)

*Variables in grey boxes above are calculated from multiple assumptions.

Based upon the assumptions, the following estimated costs were tallied:

Vendor	Low Range	High Range
A	\$591,586	\$649,546
B	\$503,457	\$606,035
C	\$673,705	\$938,357
D	\$409,086	\$473,626

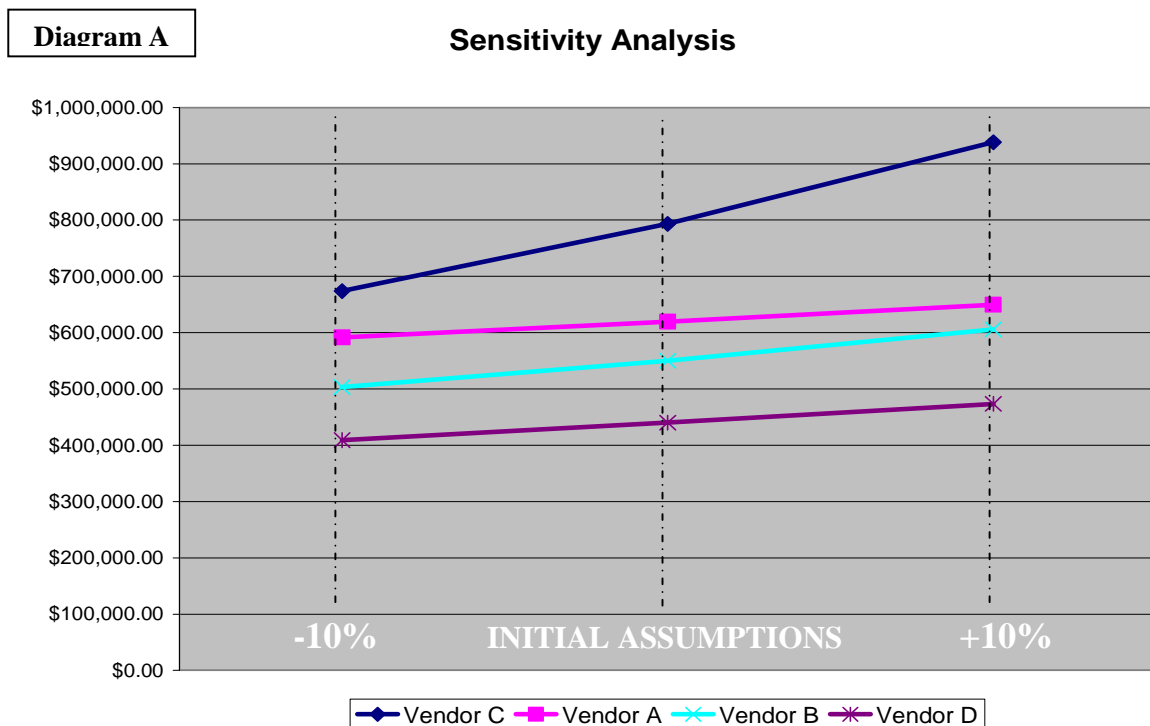
With the bids on equal footing, we immediately glean that Vendor D is the low and Vendor C the high cost providers. Vendors A and B are competitive to each other.

1.3 Executive Summary

Due to many unknowns, an analysis was developed to test the pricing schemes under different scenarios. Sensitivity analysis was run against the 22 assumptions with the resulting costs being measured and presented in the charts below. In summary, we concluded the following:

1. Vendor A is a high cost provider with potential pricing strength in native processing and production. If either native processing or document productions run towards the high end of our assumptions, vendor A becomes very attractive (See Diagrams B and C).
2. Vendor B performs well but potentially runs out of favor if the percentage of documents ultimately produced increases beyond 30% of the collection (See Diagram B).
3. Vendor C is the highest cost provider. In virtually every scenario, vendor C is significantly higher cost than all other providers. Of note, Vendor C's bid quickly becomes unreasonably high, if the size of the collection increases or if the percentage of files collected favors Native format. This is particularly troublesome because native formatting is often preferred.
4. Vendor D was the best price under all scenarios.

Diagram A shows the costs per proposal with all assumptions plus or minus 10%. This diagram gives the best overall picture of the likely project costs per proposal. Here we conclude that vendor D has a fairly significant cost advantage while vendor C must be considered the high cost provider. Vendor C's bid ranges from a low of \$670,000 to a high of \$940,000, a spread of \$250,000. The other three vendors maintain a degree of certainty as the assumptions are considered, vendor A approximately \$600,000, vendor B \$550,000 and vendor D \$450,000.



Holding all assumptions constant and testing the sensitivity of one specific variable, we find that there are three assumptions worth looking at more closely. The first assumption tested was the percentage of Custodians Processed in native format. This analysis shows the total cost per vendor as the database moves from a collection of 100% native files to one made up entirely of image files. The initial assumption in the spreadsheet is that 50% of all files received from consortium members would be in native format. As the makeup of the data changes from native to processed image files, vendor D and vendor B have projected costs running on a parallel slope while vendor A has a slightly higher cost. Vendor C proposal heavily taxes native productions and only becomes truly competitive if we receive significantly more image files than native files from the associated banks. Without available information regarding the data to be collected, vendor C should be removed from consideration.

Diagram B

Sensitivity Analysis on Native Processing

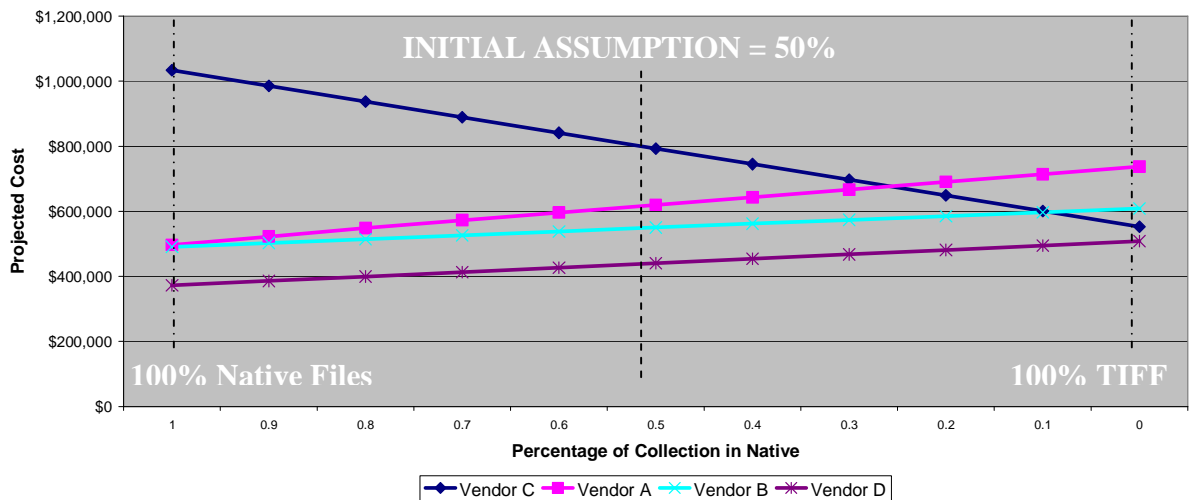
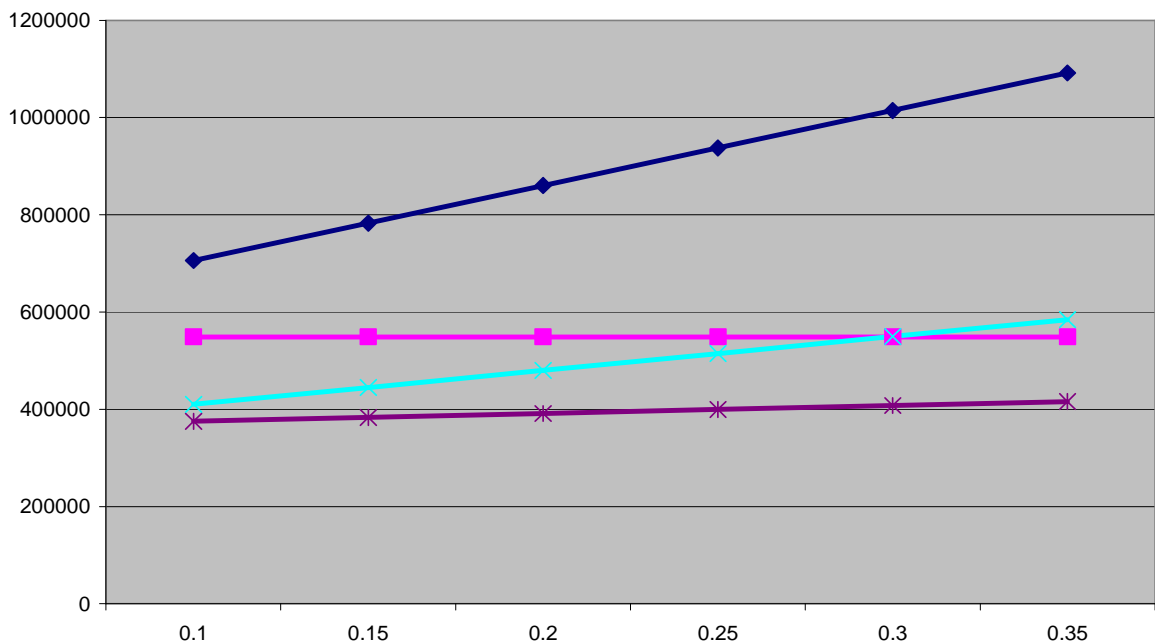


Diagram C

Project Costs As Size of Production Increase

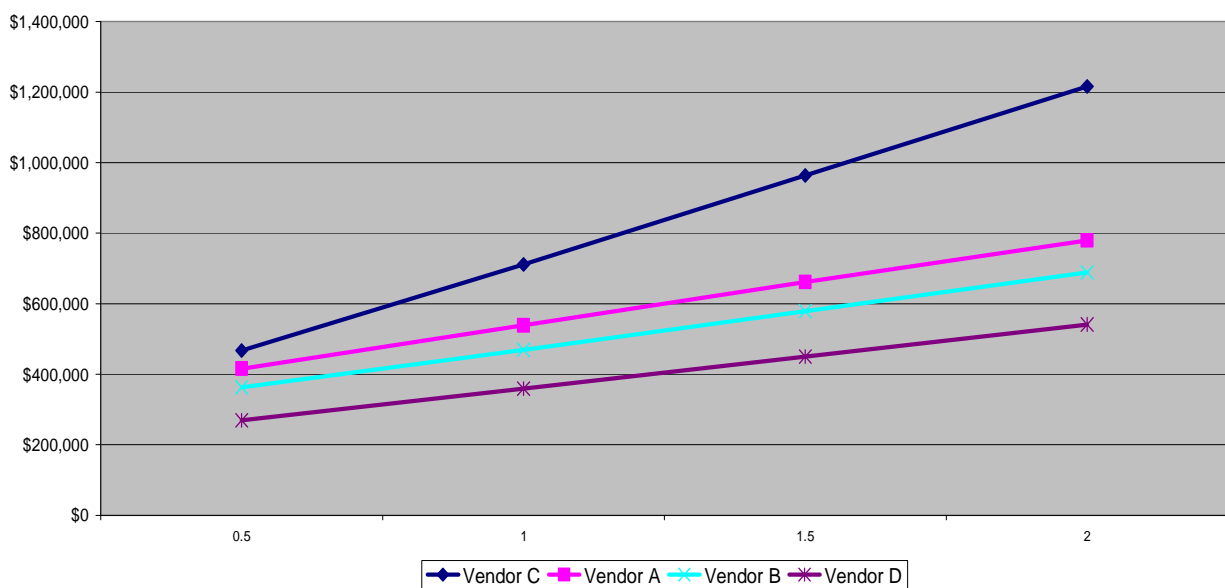


The next assumption considered (Diagram C) was Size of Production. Here vendor A does not have a cost associated with production because it relies on Law firm to perform this function. We therefore conclude that as size of production increases vendor A gains favorable consideration. After Vendor A, Vendor D provides a pricing scheme that is modestly affected by the size of the production while Vendor B and Vendor C demonstrate more sensitivity to the number of documents produced.

The final assumption manipulated was the average size of collection per custodian. Again, vendor C's proposal was the most sensitive. The remaining three proposals generally have parallel slopes. From this we understand a risk should the size of data per custodian be likely increases, vendor C becomes less competitive while the relationship between the other three bids remain constant.

Diagram D

Sensitivity Analysis on Changing Size of Collection per Custodian (gb/custodian)



In conclusion, this case study provides an appropriate analysis for gaining a more complete understanding of proposed project costs. The exercise is very important because it allows the litigation team to look closely at how a project is likely to proceed on paper before committing to an electronic discovery provider. In this case, the litigation team had four talking points going into the various vendor presentations.

1. How can we protect ourselves if the size of native files to be collected significantly increases? Diagram B demonstrates that vendor C becomes very risky as this variable increases.
2. In general terms, is vendor C's solution significantly better than the other options to justify the premium sought?
3. In comparing vendor A against vendor B, the two closest proposals in cost, is it likely that the size of production or the percentage received in TIFF format will likely favor A or B.
4. Finally, does vendor D adequately meet all requirements in light of its low bid?