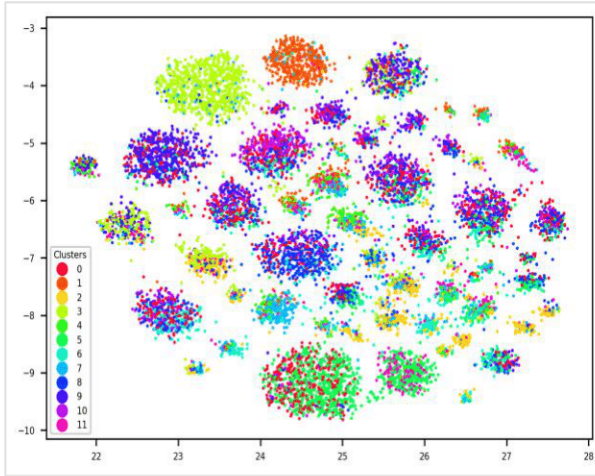
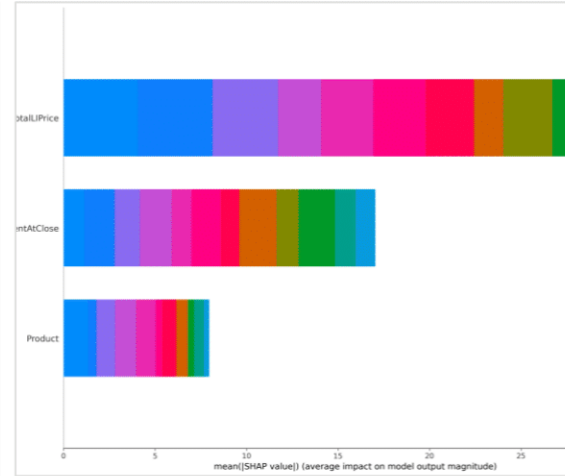


Price *K Clustering Analysis* with Python

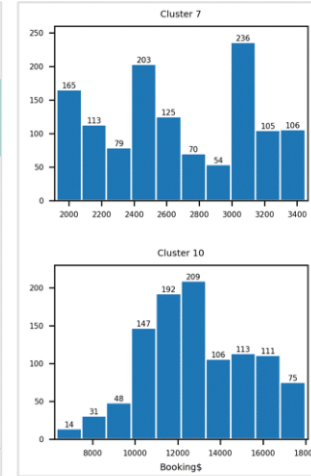
Old Methodology



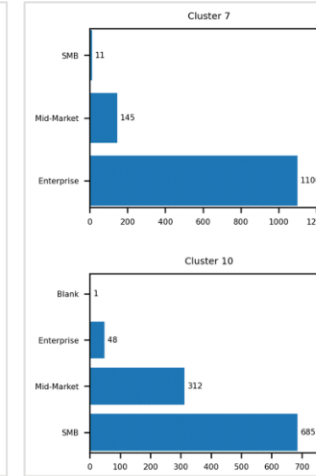
2D embedding of dataset



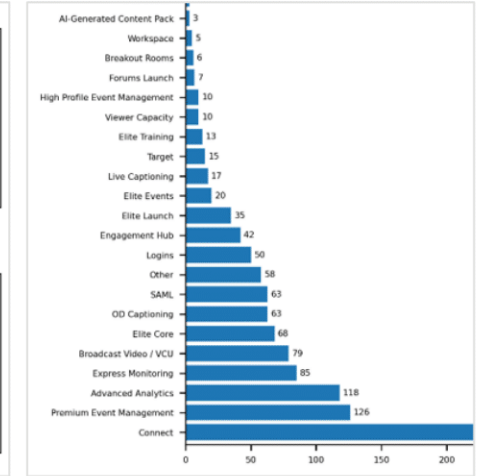
significance of each field in determining clusters



histogram of pricing field per cluster



bar graph of segment and product counts per cluster



- All contracts grouped into 12 clusters.
- Output reveals which product and segment combinations have an underlying pricing structure.
- We state which fields we used (e.g., pricing field, booking price, etc.)
- Note that pricing outliers were removed per iteration.

New Methodology

- K-modes clustering:** Groups categorical data based on the number of matching fields (column values) between data points.
- Helps us identify which combinations of products tend to be found in the same contract.
 - E.g., the dataset in Figure 1 would cluster contracts a, b, c together due to their matching 'Elite Core' and 'Connect' values.

Steps:

1. Pivot dataset so that each row represents a contract and column represents a product.

Row Labels	ACE	Advanced Analytics	AI-Generated Content Pack	Breakout Rooms	Broadcast Video / VCU	Connect	Dedicated Resource	eHub Launch	Elite Core
10X012207MS1725						1			1
10X012306MS7249						1			1
110012307AM3195									

2. Identify k value via Silhouette Method by identifying the k value corresponding to the max silhouette score. (Figure 2)
3. Perform k modes clustering via Python and plot a bar graph of the counts of each product per cluster. (Figure 3)
4. We ran 3 iterations of 14 clusters (42 clusters total), noting the top 3 products in each cluster and the frequency of product combinations.

Figures:

Contract ID	Elite Core	Connect	Logins	etc
a	1	1	1	
b	1	1		
c	1	1		
d		1	1	

Figure 1

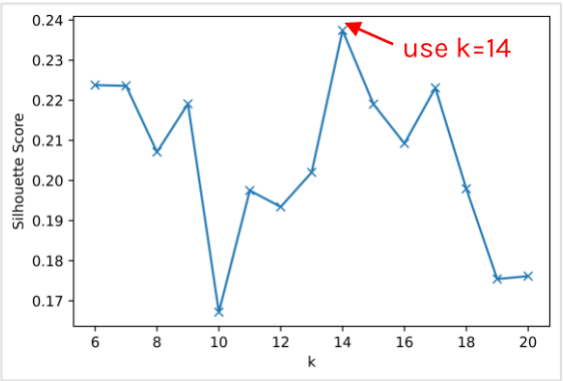


Figure 2: Silhouette Method

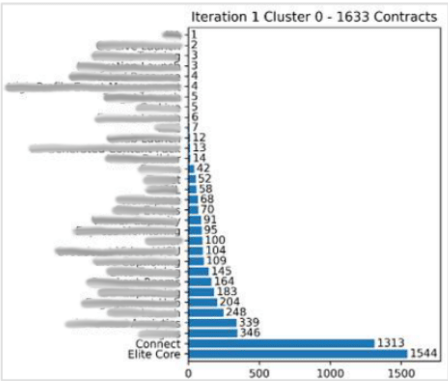


Figure 3: Cluster bar graph

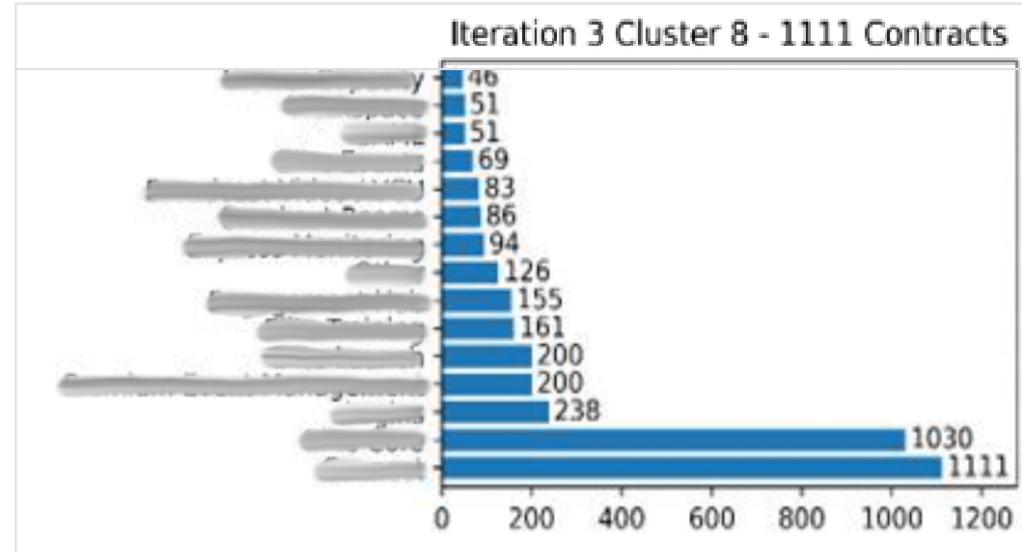
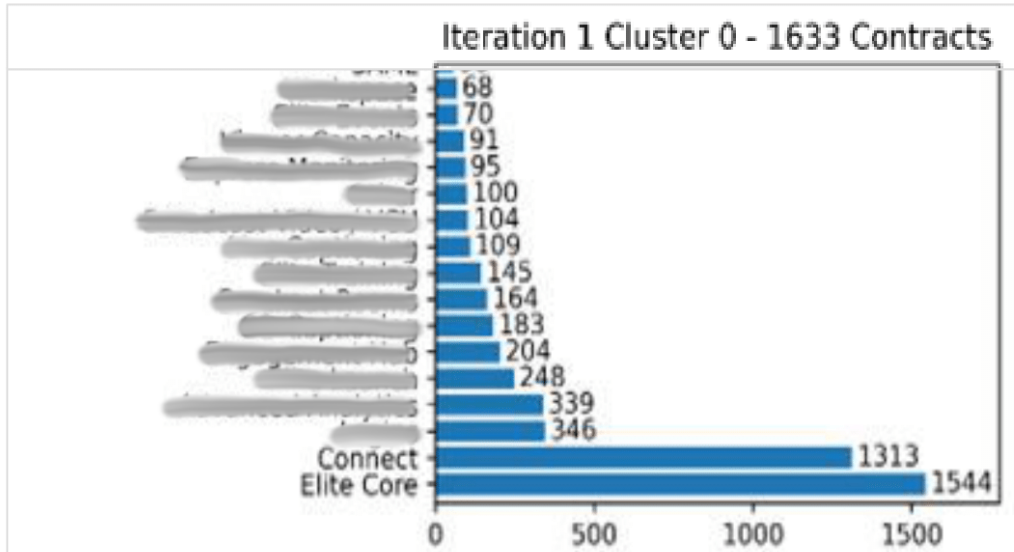
*Data is blurred or altered for confidentiality

Summary of Results (partial screenshot)

Combinations	Frequency	% of contracts with combo
Connect + Elite Core + Logins	7	16.61%
[blurred]	3	14.63%
[blurred]	3	9.31%
C [blurred]	3	7.43%
E [blurred]	3	6.70%
[blurred]	1	5.11%
[blurred]	1	5.04%
[blurred]	3	4.86%
[blurred]	1	3.56%
[blurred]	1	2.81%
[blurred]	2	2.38%
[blurred]	1	2.33%

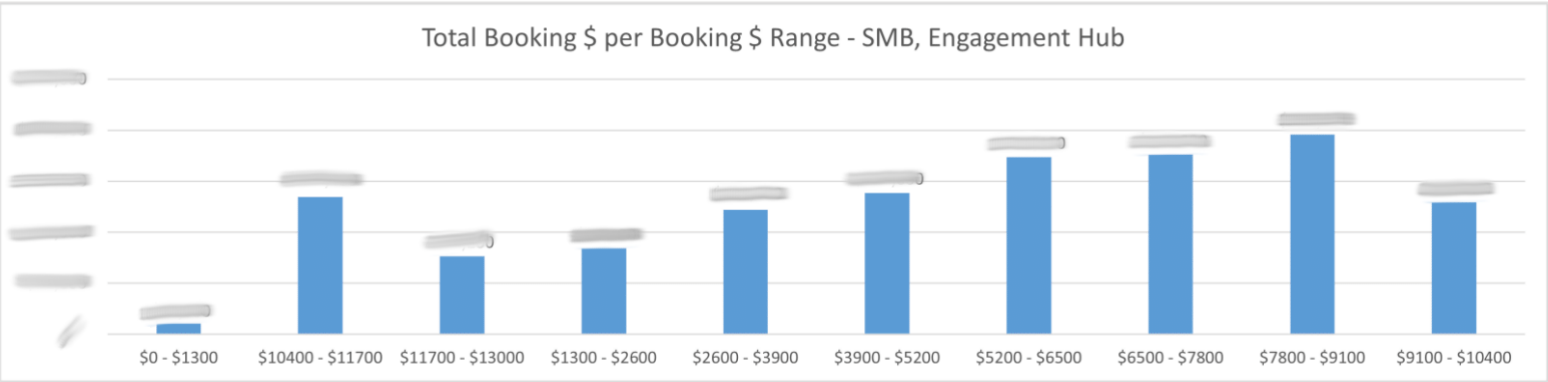
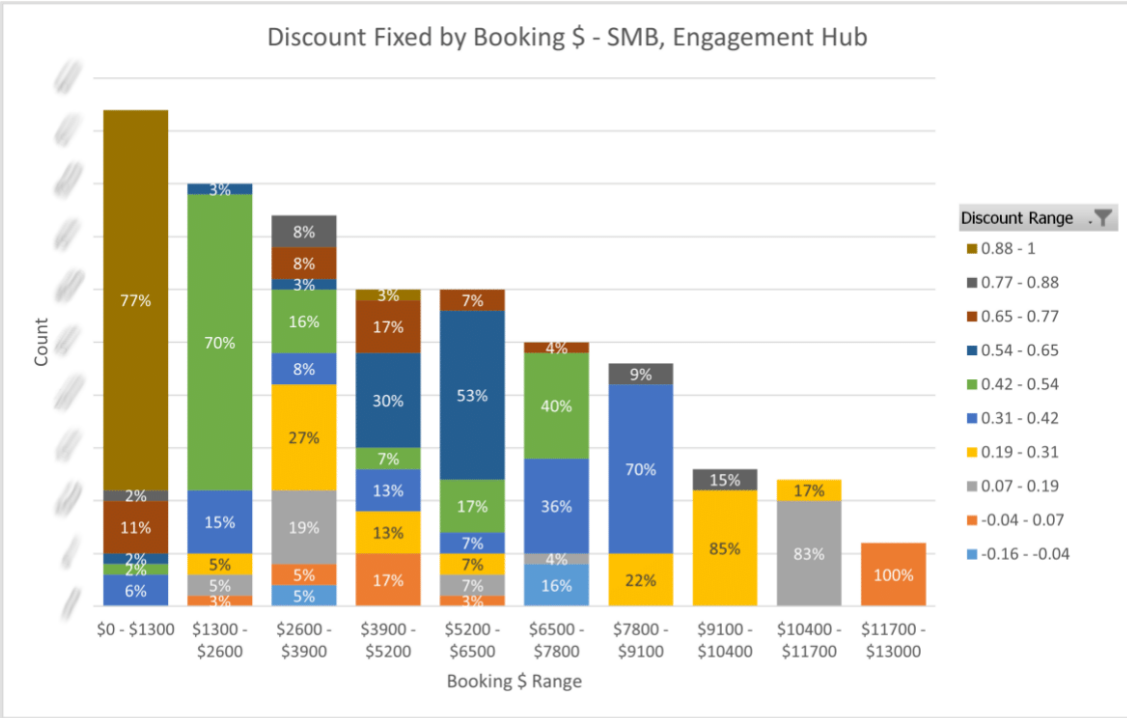
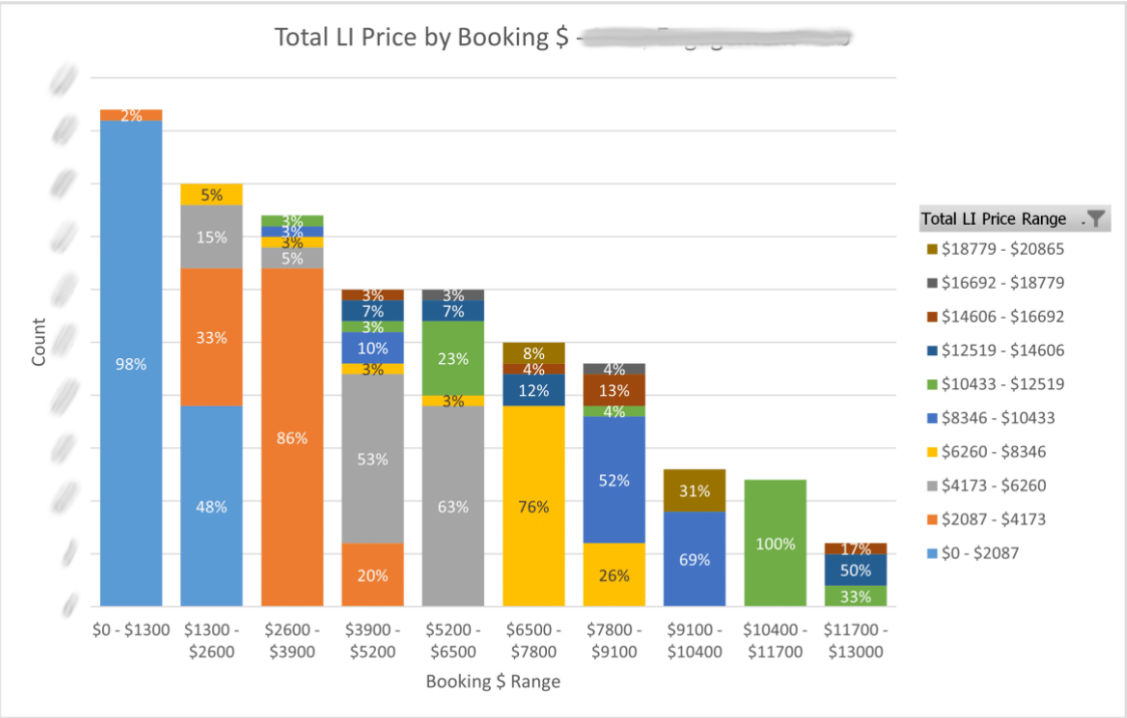
- **Combinations:** All combinations composed of the top 3 products from each cluster.
- **Frequency:** Count of clusters containing the combination.
- **% of contracts with combo:** The percentage of all contracts (from original dataset) containing the combination.
- E.g., Connect + Elite Core + Logins were the top 3 products in 7 clusters and 16.61% of all contracts contain the combination.

Notable Product Combinations



- Most notable bundles featured Connect & Elite Core together.
- These 2 graphs represent approx. 47% of total contracts while other cluster appearances were <5% of total contracts.

Identify Cluster Trends



Look at the relationship between each pricing field per product combination to determine why the products were clustered together.

*Data is blurred or altered for confidentiality