

Propuesta de aprendizaje no supervisado para identificar cárteles de manipulación de ofertas en compras públicas ecuatorianas

Andrés Leonardo Lema Rubio

Basándome en el enfoque de aprendizaje automático propuesto por Imhof et al. (2022) para la detección de cárteles de colusión –que emplea estadísticas orientadas a la construcción de modelos de aprendizaje supervisado con datos etiquetados como colusorios en Suiza– planteo la aplicación de un modelo de clustering sobre las mismas métricas, con el fin de identificar patrones de comportamiento anticompetitivo en las compras públicas ecuatorianas. Esta elección responde a la desventaja de que en el Ecuador no existe un conjunto de datos abierto y confiable sobre colusión en el marco del servicio nacional de compras públicas.

Los modelos sobre cárteles completos alcanzaron una capacidad predictiva del 83%, mientras que para cárteles incompletos de entre 2 y 5 licitantes este porcentaje se va reduciendo, debido a la presencia de pujas competitivas que distorsionan las estadísticas de referencia.

En ese sentido, se asume que los cárteles completos no toman en cuenta el riesgo de ser detectados durante los procesos de subasta, por lo que “normalmente” no harían pujas fantasma que simulen competencia.

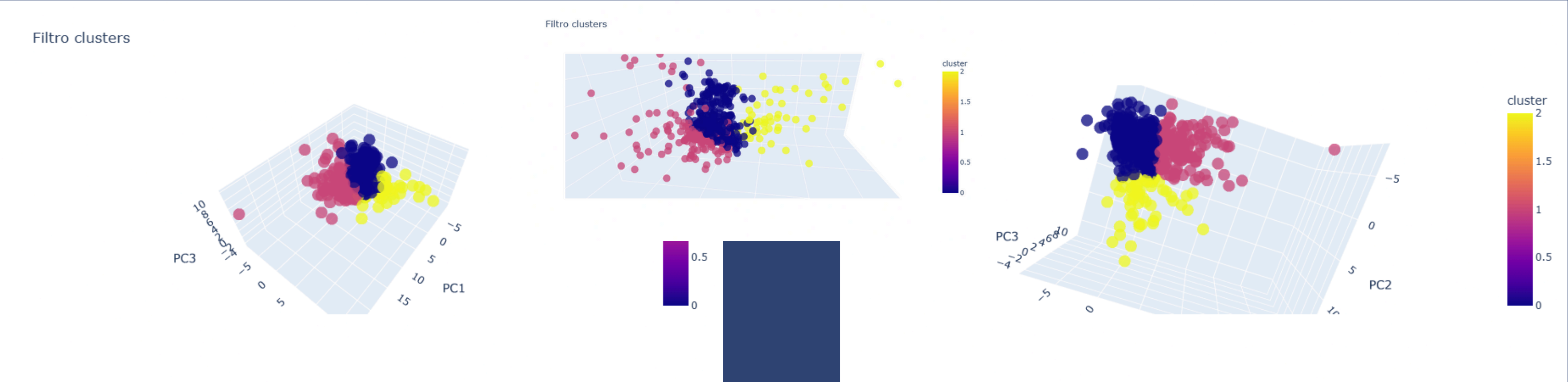
Screens (detectores estadísticos)

Estadísticas derivadas de la distribución de ofertas que permiten detectar anomalías asociadas a posibles prácticas colusorias. La colusión modifica la distribución de las ofertas y los screens capturan esos cambios.

Tipo	Indicador	Descripción
Varianza	Coeficiente de Variación (CV)	Desv. estándar / media; detecta convergencia.
	Kurtosis (KURTO)	Alta → convergencia típica de cárteles.
	Margen o Spread (SPD)	(Max – Min) / Min.
Asimetría	Diferencia Porcentual (DIFFP) / Diferencia Absoluta (D)	Diferencia % o absoluta entre 1ª y 2ª oferta más baja.
	Asimetría o Skewness (SKEW)	Mide asimetría de las ofertas.
	Distancia Relativa (RD)	Diferencia relativa entre 1ª y 2ª oferta / desviación estándar.
	Distancia Normalizada (RDNOR)	Diferencia normalizada por promedio de diferencias adyacentes.
	Distancia Relativa Alternativa (RDALT)	Igual a RD pero con solo ofertas perdedoras.
Uniformidad	Estadístico de Kolmogorov-Smirnov (KS)	Prueba no paramétrica de pérdida de uniformidad.

Clusterización Basada en una muestra de 500 procesos de compras publicas

cluster	num_ocids	mean_cvt	mean_kurtosis	mean_spread	mean_diffp	mean_rdt	mean_rdnor	mean_skew	mean_ks
0	303	45	-1.470	125	6	1.447	460	-382	53.381
1	141	165	-1.262	582	16	272	181	-530	7.062
2	56	58	1.629	166	79	384.389	2.517	-1.465	32.698



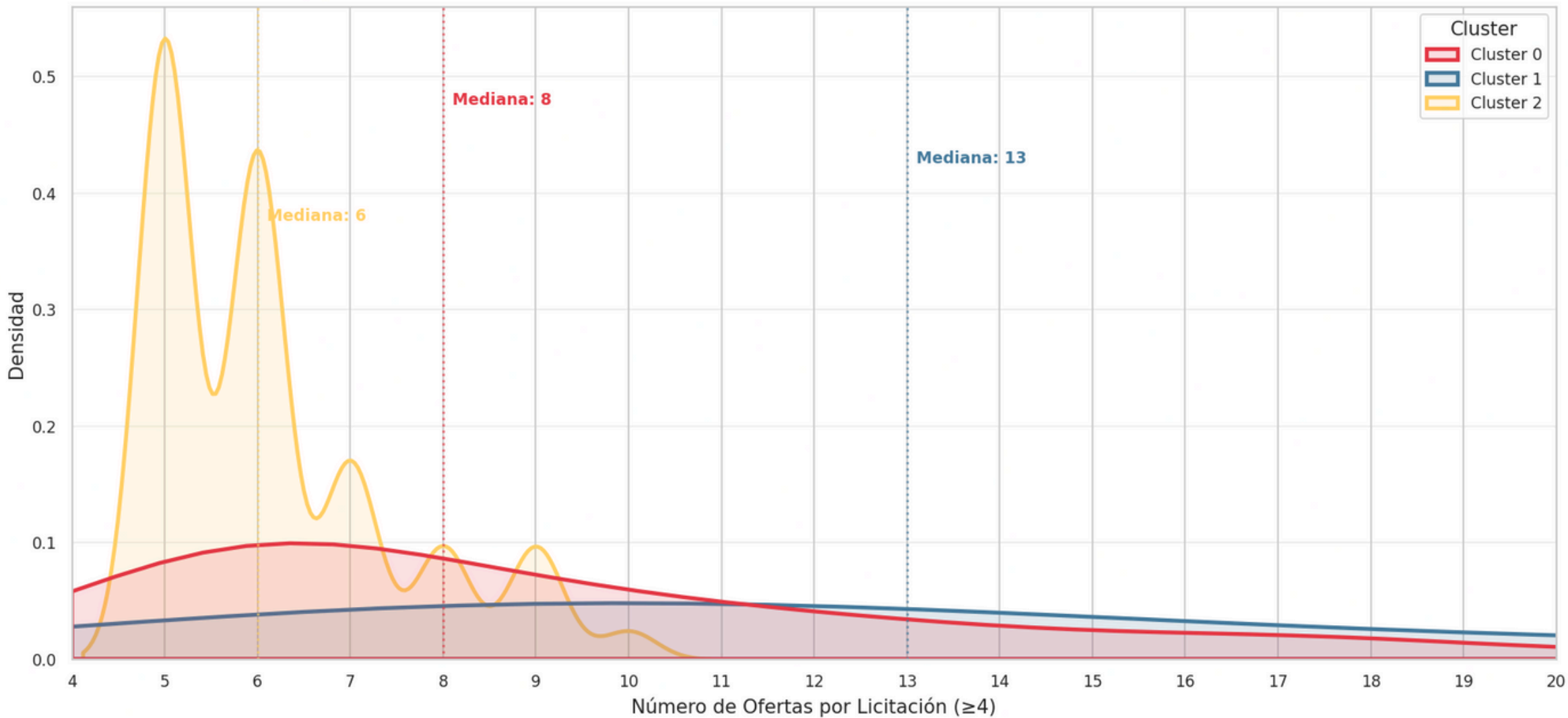
Métrica	Valor	Interpretación
Silhouette Score	2.421	Moderada separación; clusters se solapan un poco.
Calinski-Harabasz Index	907.094	Separación moderada y clusters relativamente densos.
Davies-Bouldin Index	14.884	Dispersión moderada; clusters no demasiado similares.

Tests de diferencias entre clusters para estadísticos de interés

Métrica	ANOVA (F, p-valor)	Kruskal-Wallis (H, p-valor)	Tamaño del Efecto (η^2)
CVt	F=393.2, p<0.0001	H=288.97, p<0.0001	0.62 ()
KS Statistic	F=20.42, p<0.0001	H=288.80, p<0.0001	0.18 ()
RDt	F=6.51, p=0.0016	H=170.34, p<0.0001	0.09 ()
DIFFP	F=134.20, p<0.0001	H=130.43, p<0.0001	0.45 ()

	cvt	ks_statistic	rdt	diffp
0	0.044712	53.380961	1.446801	0.006198
1	0.164841	7.062497	0.272174	0.015941
2	0.057957	32.698101	384.389034	0.078720

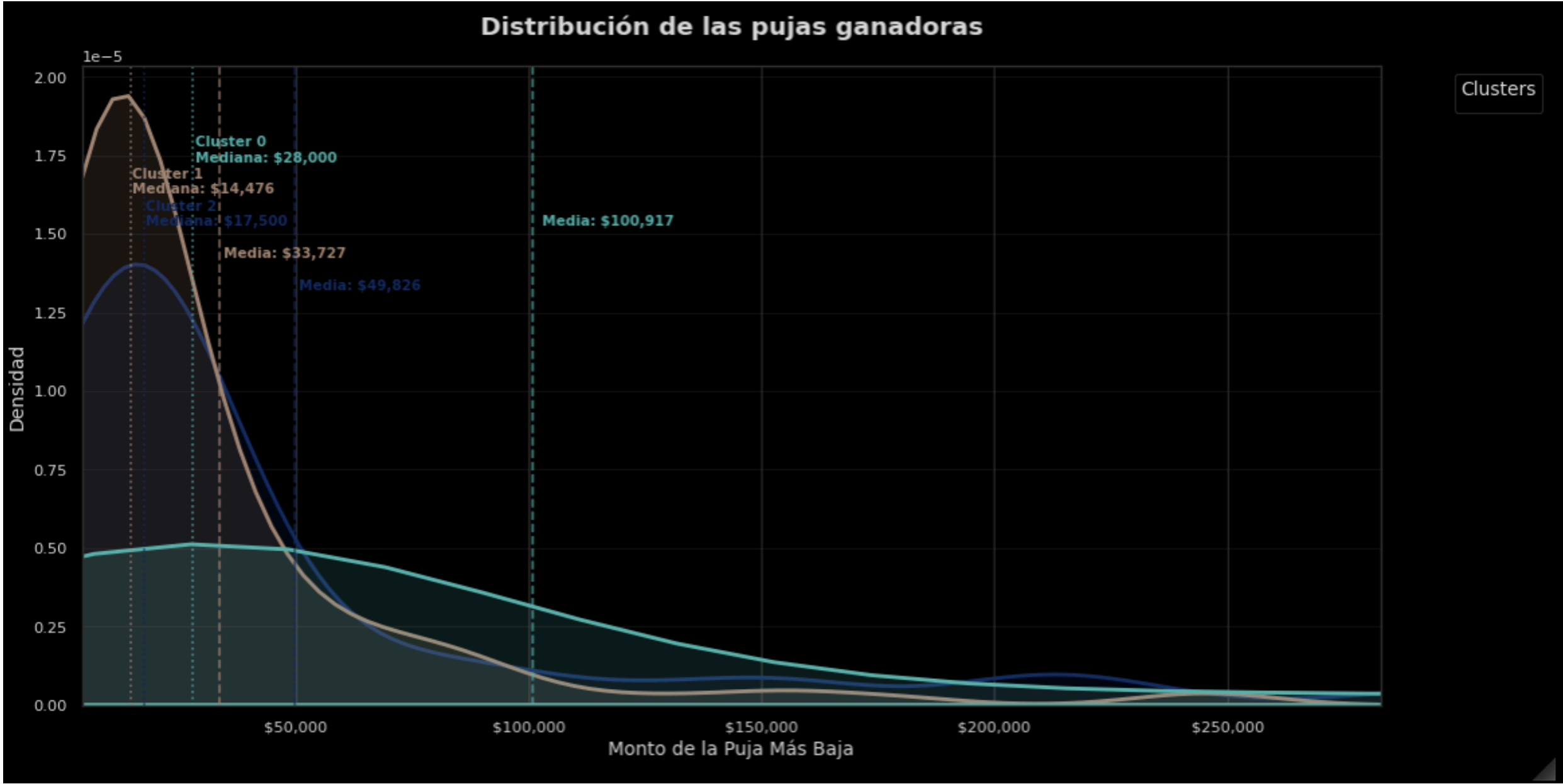
Distribución de Cantidad de Ofertas por OCID según Cluster



Tal como Imhof et.al (2022) sospechan, un coeficiente de variación bajo es indicador de convergencia en ofertas llevadas a cabo por cárteles, en este caso los clústeres 0 y 2 presentan dicho comportamiento junto con un kolmogorov smirnov statistic muy alto con relación al cartel 1, lo cual no indica una distribución de ofertas normal.

Además de acuerdo al gráfico de densidad, la cantidad de pujas en el cluster 1 sigue una distribución más uniforme, teniendo en cuenta que las subastas realmente competitivas tardan más en cerrarse debido a la participación activa de los oferentes. No es el caso del cluster 2, donde dichos procesos no superan la 11ava puja.

Discusión



La colusión no solo beneficia a los oferentes involucrados, sino que perjudica al estado, ya que como podemos observar, la distribución de las pujas ganadoras en términos de mediana y media para un cluster “competitivo” como el 1 indica que los valores adjudicados tienden a ser menores que aquellos donde existe colusión, porque por definición, la manipulación de ofertas en una subasta inversa busca maximizar el ingreso del oferente ganador, y este ingreso es estrictamente mayor que aquel generado si se hubiese adjudicado un proceso “competitivo”

Comparación	Test Estadístico	Hipótesis Nula (H ₀)	Hipótesis Alternativa (H ₁)	U	p-valor	Interpretación
Cluster 1 < Cluster 0	Mann–Whitney U	La mediana del Cluster 1 ≥ mediana del Cluster 0	La mediana del Cluster 1 < mediana del Cluster 0	145.890	0	Significativamente menor (p < 0.05)
Cluster 1 < Cluster 2	Mann–Whitney U	La mediana del Cluster 1 ≥ mediana del Cluster 2	La mediana del Cluster 1 < mediana del Cluster 2	33.805	(0,0581)	No significativamente menor (p > 0.05)

GRACIAS