



UNIVERSIDAD ARTURO PRAT
DEL ESTADO DE CHILE

CONOCIMIENTO Y TERRITORIO

FACULTAD DE INGENIERÍA Y ARQUITECTURA

INGENIERÍA CIVIL EN COMPUTACIÓN E INFORMÁTICA



**ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE DATAMART PARA
LICITACIONES DE INSTITUCIONES UNIVERSITARIAS CHILENAS
INTELIGENCIA DE NEGOCIOS**

INTEGRANTES:

Gianfranco Pallero
David Neira
Juan González

IQUIQUE-CHILE

2024

Contenido

1.	Introducción	5
2.	Definición de la problemática	6
3.	Presupuesto	7
4.	Especificación de los requisitos estratégicos	8
4.1	Descripción de Requerimientos Estratégicos	8
4.2	Tabla de temas analíticos.....	8
4.3	Matriz de procesos/dimensiones	9
4.4	KPI y especificaciones	9
5.	Descripción de las fuentes de datos de origen	17
	Sistemas Operacionales Involucrados (SOI)	17
5.1	Modelo Entidad-Relación de los SOI.....	18
5.2	Modelo Físico Asociado	18
5.3	Fuentes de datos externas	18
5.4	Detalle de la fuente de información	18
6.	Diseño del Datamart	23
6.1	Descripción de la Etapa	23
6.2	Descripción de Herramientas Utilizadas	23
6.3	Descripción de componentes del Modelo Multidimensional.....	24
6.4	Modelo Multidimensional.....	29
6.5	Justificación del Modelo.....	29
6.7	Modelo Físico Multidimensional.....	33
6.8	Script DDL.....	33
7.	Análisis exploratorio de datos.....	33
7.1	Descripción de la etapa	33
7.2	Descripción de Herramientas Utilizadas	33
7.3	Estadísticos Descriptivos	34
7.4	Gráficos	45
8.	PROCESO ETL	58
8.1	Descripción del Proceso y Datos.	58
8.2	Descripción del modelo	59
8.3	Descripción de Problemas Encontrados	60
	Criterios Adoptados y Lógica de Transformación	61

Estadísticas Básicas de Respaldo de Criterios	¡Error! Marcador no definido.
Descripción de Herramientas Utilizadas.	62
Carga y Análisis del Proceso ETL (<i>Por cada Dimensión, Hecho y Métrica</i>)	62
Modelo ETL (Indicar Herramienta Utilizada)	71
Problemas en la Carga de Datos	71
Rendimiento de Herramientas y Tiempos de Carga.....	71
9. Explotación de Datos CON OLAP.....	72
9.1 Descripción de la Etapa	72
9.2 Descripción de Herramientas Utilizadas.....	72
9.3 Indicadores obtenidos y Validación del Indicador.....	72
Dashboards.....	72
KPI 1.....	75
KPI 2.....	77
KPI 3.....	79
KPI 4.....	80
KPI 5.....	83
KPI 6.....	86
KPI 7.....	88
KPI 8.....	90
KPI 9.....	93
KPI 10.....	95
KPI 11.....	97
KPI 12.....	100
10. Conclusiones	103
Conclusiones sobre el Proyecto	103
Conclusiones sobre la Tecnología Utilizada	103
Conclusiones sobre de los Resultados Obtenido.....	103
11. Anexos.....	103

<i>Ilustración 1 - Dimensión Proveedor.....</i>	24
<i>Ilustración 2- Dimensión Rubro.....</i>	24
<i>Ilustración 3 - Dimensión Oferta.....</i>	24
<i>Ilustración 4 - Dimensión Producto</i>	25
<i>Ilustración 5 - Dimensión Licitacion</i>	25
<i>Ilustración 6 - Dimensión Region.....</i>	25
<i>Ilustración 7 - Dimensión Contrato.....</i>	26
<i>Ilustración 8 - Dimensión Comprador.....</i>	26
<i>Ilustración 9 - Dimensión Tiempo.....</i>	27
<i>Ilustración 10 - Tabla Hecho hecho_licitacion</i>	28
<i>Ilustración 11 - Modelo físico Multidimensional</i>	29
<i>Ilustración 12 - Dashboard requerimiento 1</i>	73
<i>Ilustración 13- Dashboard requerimiento 3.....</i>	74
<i>Ilustración 14- Dashboard requerimiento 4.....</i>	74
<i>Ilustración 15- Dashboard requerimiento 5.....</i>	75

1. Introducción

Considerando la complejidad creciente en la gestión de información en el entorno empresarial moderno, PyGo SPA se enfrenta a la necesidad apremiante de interpretar adecuadamente los conjuntos de datos públicos proporcionados por el gobierno de Chile. Este desafío se traduce en la dificultad para evaluar con precisión las necesidades del mercado y las oportunidades derivadas de las licitaciones universitarias, poniendo en riesgo su competitividad y éxito en estas nuevas iniciativas. Para abordar esta problemática, PyGo SPA ha optado por implementar estrategias específicas de Inteligencia de Negocios (BI). Este enfoque permitirá a la empresa transformar los datos brutos en información estructurada y relevante, mejorando así su capacidad de toma de decisiones y estrategias de mercado. A través del análisis de datos, la visualización de información y el modelado predictivo, PyGo SPA busca obtener una comprensión profunda del entorno del mercado público y las expectativas de las universidades chilenas, identificando áreas clave para la expansión futura de sucursales y consolidando su posición en el mercado.

2. Definición de la problemática

PyGo SPA es una empresa privada dedicada a la venta de productos no especializados y actualmente se encuentra en proceso de transición hacia la integración en el mercado público, particularmente a través de la participación en licitaciones ofrecidas por universidades chilenas, se busca edificar una futura sucursal en un punto estratégico como inversión a largo plazo. Este movimiento estratégico se fundamenta en el aprovechamiento de los conjuntos de datos públicos habilitados por el gobierno de Chile, que proporcionan información valiosa sobre diversos aspectos del mercado y las oportunidades disponibles en el sector público.

No obstante, la empresa enfrenta desafíos significativos relacionados con la interpretación adecuada de estos datos. El principal obstáculo es la falta de capacidad para analizar y comprender correctamente la información disponible, lo que impide una evaluación precisa de las necesidades del mercado y las oportunidades de negocio en el ámbito de las licitaciones universitarias. Esta dificultad en la interpretación puede llevar a decisiones empresariales mal informadas y, por ende, afectar negativamente la competitividad y el éxito de PyGo SPA en estas nuevas iniciativas.

Para superar estos retos, PyGo SPA ha decidido adoptar estrategias específicas de Inteligencia de Negocios (BI). El uso de estas técnicas permitirá a la empresa transformar los datos brutos en información estructurada y relevante, mejorando así su capacidad de toma de decisiones y estrategia de mercado. Mediante la implementación de herramientas de BI, como el análisis de datos, la visualización de información y el modelado predictivo, PyGo SPA espera obtener una comprensión profunda y detallada del entorno del mercado público y de las expectativas de las universidades chilenas respecto a sus proveedores, además de puntos clave donde elaborar futuras sucursales.

3. Presupuesto

Para llevar a cabo este proyecto se realizó una estimación de costos a lo largo de 5 años teniendo en consideración costos de infraestructura y recursos humanos para la realización del datamart. Considerando un costo total de 17.145 USD

Costos directos						
Detalle	Año					Total
	Año 1 (USD)	Año 2 (USD)	Año 3 (USD)	Año 4 (USD)	Año 5 (USD)	
Costo de infraestructura						
Servidor	\$911,00	\$911,00	\$911,00	\$911,00	\$911,00	\$4.555,00
Mantenimiento	-	\$100,00	\$100,00	\$100,00	\$100,00	\$400,00
Licencia powerBI premium	\$250,00	\$250,00	\$250,00	\$250,00	\$250,00	\$1.250,00
Licencia Knime	\$1.188,00	\$1.188,00	\$1.188,00	\$1.188,00	\$1.188,00	\$5.940,00
Licencia PowerDesigner	\$1.000,00	\$1.000,00	\$1.000,00	\$1.000,00	\$1.000,00	\$5.000,00
Costo humano						
Recursos Humanos	\$3000	-	-	-	-	\$3000
Total						\$17.145,00

Tabla 1 Presupuesto

4. Especificación de los requisitos estratégicos

4.1 Descripción de Requerimientos Estratégicos

REQ - 1	ANALIZAR EL COMPORTAMIENTO DE LA COMPETENCIA A LO LARGO DEL TIEMPO
OBJETIVO	IDENTIFICAR LOS COMPETIDORES PRINCIPALES PARA COMPRENDER EL MERCADO Y DILUCIDAR OPORTUNIDADES

Tabla 2 Requerimiento estratégico nº1

REQ - 2	ANALIZAR EL COMPORTAMIENTO DE LOS COMPRADORES A LO LARGO DEL TIEMPO
OBJETIVO	ADQUIRIR CONOCIMIENTO ACERCA DE COMPRADORES FRECUENTES Y SEGUN QUE REGIONES

Tabla 3 Requerimiento estratégico nº2

REQ - 3	ANALIZAR LA DEMANDA DE PRODUCTOS EN EL MERCADO A LO LARGO DEL TIEMPO
OBJETIVO	IDENTIFICAR LOS PRINCIPALES PRODUCTOS DEMANDADOS POR LAS INSTITUCIONES Y SEGUN QUE REGIONES

Tabla 4 Requerimiento estratégico nº3

REQ - 4	ANALIZAR EL ÉXITO DE LAS OFERTAS DE LOS PROVEEDORES A LO LARGO DEL TIEMPO
OBJETIVO	IDENTIFICAR PATRONES EN LAS OFERTAS EXITOSAS PROVENIENTES DE PROVEEDOR

Tabla 5 Requerimiento estratégico nº4

REQ - 5	EVALUAR EL IMPACTO DE LA REGIONALIZACIÓN EN LAS DECISIONES DE ADJUDICACIÓN
OBJETIVO	COMPRENDER CÓMO LA LOCALIZACIÓN GEOGRÁFICA AFECTA LAS OPORTUNIDADES DE ADJUDICACIÓN PARA MEJORAR LA COBERTURA GEOGRÁFICA DE LAS OFERTAS

Tabla 6 Requerimiento estratégico nº5 Tabla de temas analíticos

4.2 Tabla de temas analíticos

Temas analíticos	Análisis o requisitos inferidos	Proceso de negocio de soporte	Comentario
Comportamiento de la competencia	Análisis del comportamiento de la competencia dentro del mercado a lo largo del tiempo	Identificar las principales amenazas presentes en el mercado	Este tema busca comprender el panorama competitivo para identificar oportunidades y amenazas.
Comportamiento de los compradores	Análisis del comportamiento de los compradores a lo largo del tiempo	Hallar los principales compradores del mercado de licitaciones	Este tema busca entender las necesidades y preferencias de los compradores enfocar los esfuerzos en

			compradores de mayor potencial.
Demanda de los productos en el mercado	Analizar la demanda de productos a lo largo del tiempo	Identificar los productos más demandados en el mercado	Este tema busca entender las necesidades y preferencias de los compradores enfocar los esfuerzos en productos de mayor potencial
Éxito de las ofertas de proveedores	Análisis de las ofertas exitosas a lo largo del tiempo	Detectar tendencias en el éxito de las ofertas en los procesos licitativos	Este tema busca mejorar la tasa de éxito en las licitaciones.
Impacto de la regionalización	Ánalisis del impacto de la regionalización en las adjudicaciones	Comprender como el factor de la región incide en el éxito de las ofertas	Este tema busca entender cómo afecta la región en los procesos licitativos.

Tabla 7 Temas analíticos

4.3 Matriz de procesos/dimensiones

	Dimensiones						
Proceso de negocio	Tiempo	Región	Licitación	Oferta	Proveedor	Producto	Comprador
Identificar las principales amenazas presentes en el mercado	X			X	X		X
Hallar los principales compradores del mercado de licitaciones	X	X	X			X	X
Identificar los productos más demandados en el mercado	X	X				X	X
Detectar tendencias en el éxito de las ofertas en los procesos licitativos	X	X	X	X			
Comprender como el factor de la región incide en el éxito de las ofertas	X	X	X	X	X		

Tabla 8: Matriz de procesos Kimball

4.4 KPI y especificaciones

Se opta por separar las siguientes métricas para no sobrecargar las fórmulas y que sean legibles.

Variable 1 Conteo total de licitaciones

Símbolo	Tl
Objetivo	Obtener la cantidad total de licitaciones realizadas para ser utilizada en distintos contextos

Tabla 9 - Métrica Conteo Total de licitaciones

Variable 2 Conteo total de ofertas

Símbolo	To
Objetivo	Obtener la cantidad total de ofertas realizadas para ser utilizadas en distintos contextos

Tabla 10 - Métrica Conteo Total de ofertas

Variable 3 Conteo total de proveedores

Símbolo	Tp
Objetivo	Obtener la cantidad total de proveedores para ser utilizados en distintos contextos

Tabla 11- Métrica Conteo Total de proveedores

Variable 4 Conteo total de montos estimados

Símbolo	Cm
Objetivo	Obtener la suma total de montos estimados por licitación para ser utilizado en distintos contextos

Tabla 12 - Métrica Conteo Total de montos estimados

Variable 5 Conteo total de adjudicaciones

Símbolo	Ta
Objetivo	Obtener la cantidad total de adjudicaciones por licitación para ser utilizada en distintos contextos

Tabla 13 - Métrica Conteo Total de adjudicaciones

Dimensión	Símbolo
Tiempo	t
Producto	p
Comprador	c
Proveedor	pr
Región	r
Licitación	l
Oferta	o

Tabla 14 Gránulo

REQ-3 Análisis de la demanda

KPI - 01	<i>Porcentaje de licitaciones por categoría de producto en los periodos 2019 – 2023 para cada región.</i>
Objetivo	<i>Identificar la frecuencia con la que se licitan ciertos productos a lo largo del tiempo.</i>
Meta	<i>Siendo este un porcentaje respecto al total, se busca el mayor porcentaje presente como meta, mientras más cercano a 100%, mayor posibilidad de que licitaciones a futuro sean para la misma categoría de productos.</i>
Periodo	<i>Mensual – Semestral - Anual</i>
Dimensiones	<i>Producto - Tiempo – Comprador – Región</i>
Criterios	<i>Se considera el número de licitaciones por categoría de producto hechas por un comprador</i>
Observación	<i>Este KPI responde tanto al REQ-3 como al REQ-5</i>
Fórmula	$= \left(\frac{\sum \text{Licitación}}{Tl} \times 100 \right)_{t*p*r}$

Tabla 15 KPI - 01 Porcentaje de licitaciones por categoría de producto en los periodos 2019 - 2023 para cada región

REQ-1 Análisis competencia

KPI - 02	<i>Tasa de crecimiento de ofertas por proveedor por año.</i>
Objetivo	<i>Indicar el crecimiento de los distintos proveedores según el paso de los años, útil para evaluar la intensidad del mercado y los proveedores emergentes.</i>
Meta	<i>Se desea un mercado estable que no contenga empresas altamente crecientes en participación, cualquier empresa con una tasa de crecimiento de participación mayor a 10% significa una posible competencia perdida.</i>
Periodo	<i>Anual</i>
Dimensiones	<i>Tiempo - Proveedor - Oferta</i>
Criterios	<i>Se considera el total de ofertas hechas por un proveedor</i>
Observación	<i>Es importante adjuntar a este KPI la métrica de cantidad de ofertas por proveedor en un año para mejor contextualización de la realidad.</i>
Fórmula	$= \left(\left(\frac{T_{O(t)} - T_{O(t-1)}}{T_{O(t-1)}} \right) * 100 \right)_{pr}$

Tabla 16 KPI - 05 Tasa de crecimiento de ofertas por proveedor por año

KPI - 03	<i>Probabilidad de que un proveedor dentro del ranking de la licitación se adjudique una licitación de cierto comprador a lo largo del tiempo.</i>
-----------------	--

Objetivo	<i>Descubrir patrones de asociación históricos según ofertantes y compradores para medir y evaluar tendencias de adjudicación a proveedores.</i>
Meta	<i>Siendo esta una probabilidad histórica, se desea no encontrar probabilidades mayores al 50% pues ello significaría que un proveedor domina el mercado de cierto comprador en su mayoría.</i>
Periodo	<i>Mensual – Semestral - Anual</i>
Dimensiones	<i>Tiempo – Proveedor - Comprador - Oferta</i>
Criterios	Se consideran solo las licitaciones adjudicadas a un proveedor del ranking
Observación	Total de licitaciones adjudicadas a el proveedor por cada institución en un tiempo determinado. Este KPI no es un modelo predictivo, es una probabilidad histórica de adjudicaciones según ofertas, se puede interpretar como una tasa de éxito que se puede proyectar a futuro, no obstante, no sigue un orden predictivo
Fórmula	$= \left(\left(\frac{\sum (LictAdjudicada)_{pr*c*t}}{To_{pr*c*t}} \right) \times 100\% \right)$

Tabla 17 KPI – 03 Probabilidad de que un proveedor dentro del ranking de la licitación se adjudique una licitación de cierto comprador a lo largo del tiempo

KPI - 04	<i>Índice de participación competitiva a lo largo del tiempo.</i>
Objetivo	<i>Evaluuar el nivel de actividad y éxito de los competidores en las licitaciones a lo largo del tiempo, identificando quienes son los principales actores en el mercado.</i>
Meta	<i>Cualquier empresa con índice mayor a 0.5 contiene una elevada competitividad en el mercado.</i>
Periodo	<i>Mensual – Semestral - Anual</i>
Dimensiones	<i>Tiempo - Oferta – Proveedor</i>
Criterios	Se consideran las ofertas que resultaron en una adjudicación y las ofertas ganadoras, pero que no resultaron en una adjudicación

Observación	<p>Es importante adjuntar a este KPI la cantidad de ofertas realizadas para comprender de mejor forma la competitividad de cada empresa.</p> <p>Usamos una multiplicación entre los índices de victorias según proveedor, aceptación de ofertas según proveedor y participación según proveedor, mientras más cercano sea a 1, más competitivo y exitoso es el proveedor.</p>
Fórmula	$IPC_t = \left(\frac{\text{ConteoOfertas(adjudicadas)}_{pr}}{To_{pr}} \times \frac{\text{ConteoOfertas(aceptadas)}_{pr}}{To_{pr}} \times \frac{To_{pr}}{To} \right)_t$

Tabla 18 KPI - 04 Índice de participación competitiva a lo largo del tiempo

REQ-2 Análisis compradores

KPI - 05	Tasa de crecimiento del monto promedio estimado por categoría de producto por institución por tipo de licitación a lo largo del tiempo.
Objetivo	Analizar el crecimiento económico de los presupuestos estimados para las categorías de los productos en distintas instituciones a lo largo del tiempo.
Meta	Cualquier tasa de crecimiento que sea positiva implica un mercado creciente y beneficioso, no obstante, se pone un objetivo idóneo de 5%.
Periodo	Mensual – anual
Dimensiones	Tiempo - Licitación - Comprador - Producto
Criterios	Se considera el monto estimado de licitación por institución y tipo de licitación.
Observación	No se considera la cantidad de ítems a comprar pues se busca generalizar los montos.
Fórmula	$TDC = \left(\left(\frac{\left(\left(\frac{(\Sigma \text{MontoEstLict})}{Tl} \right)_t - \left(\left(\frac{(\Sigma \text{MontoEstLict})}{Tl} \right)_{p*c*(t-1)} \right)}{\left(\left(\frac{(\Sigma \text{MontoEstLict})}{Tl} \right)_{(t-1)} \right)} \right)_{p*c*l} * 100 \right)$

Tabla 19: KPI - 05 Tasa de crecimiento del monto promedio estimado por categoría de producto por institución por tipo de licitación a lo largo del tiempo

KPI - 06	Promedio de montos estimados por región de comprador por categoría de producto.
Objetivo	Comprender que montos estimados poseen ciertos compradores para ciertos tipos de productos despreciando la cantidad de productos.
Meta	No se considera una meta específica para este KPI, no obstante, a mayor sea el promedio, mayor margen de beneficios puede significar para la empresa integrarse al mercado.
Periodo	Mensual – Semestral - Anual

Dimensiones	Tiempo - Licitación - Comprador – Producto - Región
Criterios	Se consideran los montos estimados de licitación por institución compradora dentro de la región
Observación	Este KPI responde tanto al REQ-2 como al REQ-5.
Fórmula	$= \left(\frac{\sum (MontEstLict)}{Cm} \right)_{c*p*r*t*l}$

Tabla 20 KPI – 06 Promedio de montos estimados por región de comprador por categoría de producto

KPI - 07	Porcentaje de licitaciones por institución por año.
Objetivo	<i>Evaluuar de forma superficial que instituciones desarrollan más licitaciones a lo largo de los años.</i>
Meta	<i>En consideración de la alta cantidad de instituciones con la capacidad de licitar, cualquier institución con un porcentaje mayor a 1% debe ser considerada de gran magnitud.</i>
Periodo	Mensual - Anual
Dimensiones	Tiempo - Licitación - Comprador
Criterios	Se considera el promedio por institución compradora
Observación	La agrupación de este KPI en conjunto al volumen de presupuesto estimado da paso a la clasificación de compradores en la dimensión comprador.
Fórmula	$= \left(\frac{\sum (Licitacion)}{Tl_t} \right)_c$

Tabla 21: KPI – 07 Porcentaje de licitaciones por institución por año

REQ-4 Análisis Oferta

KPI - 08	Análisis de ranking de proveedores que realizan las ofertas más exitosas por región por año.
Objetivo	<i>Proporcionar visión clara de la efectividad de los proveedores al obtener adjudicaciones en diferentes regiones y años, ponderando las métricas MontoTotalOfertada, To y ratio de éxito para seleccionar ofertas con tendencia en común y evaluarlas en carácter individual.</i>
Meta	<i>Este KPI no cuenta con una meta clara, pues busca generar un ranking entre proveedores mediante una ponderación o exageración</i>

	<i>de sus ganancias y éxitos y en base a este éxito apoyar el análisis para los otros KPI.</i>
Periodo	Anual
Dimensiones	<i>Tiempo – Oferta – Proveedor – Región</i>
Criterios	Se consideran solo las ofertas que resultaron en una adjudicación
Observación	El área de interés de este KPI se centra en los proveedores dentro del top 10, estos son los proveedores que han realizado las ofertas más exitosas en distintas regiones. <i>Este KPI responde tanto al REQ-4 como al REQ-1 Y REQ-5</i>
Fórmula	$= (Vol_{adj} \times conteo(ofertas)_{adj} \times Ratio_{exito\ ofertas})_{pr*r*t}$ $Vol_{adj} = \left(\sum_{o*pr*r*t} Monto\ TotalOferta(ganadora) \right)$ $conteo(ofertas)_{adj} = (To(ganadora))_{pr*r*t}$ $Ratio_{exito\ ofertas} = \left(\frac{To(ganadora)}{To} \right)_{pr*r*t}$

Tabla 22: KPI –08 Análisis de ofertas exitosas del ranking de proveedores por región por año

KPI - 09	Relación promedio entre monto ofertado respecto al monto total estimado, según ofertas ganadoras y perdedoras
Objetivo	<i>Medir comparativamente si el monto ofertado influye en el éxito de la oferta.</i>
Meta	<i>Al ser una relación promedio para analizar las otras ofertas, no se cuenta con una meta exacta, sin embargo, cualquier valor sobre 1 para ofertas ganadoras puede significar una holgura en ciertas ofertas respecto al presupuesto estimado por el comprador, lo cual es beneficioso a larga escala pues no limita el presupuesto de mayor forma.</i>
Periodo	Mensual – Anual
Dimensiones	<i>Tiempo – Oferta</i>
Criterios	Se consideran los montos ofertados ganadores y perdedores, al igual que el monto estimado de la licitación a la que se oferta
Observación	Es importante recalcar que una oferta ganadora es inherentemente adjudicada, pero una oferta perdedora pudo ganar y rechazar el contrato por conveniencia, por lo tanto, esta medida tiene un sesgo innato e indeterminable.
Fórmula	$= \left(\frac{\sum \left(\frac{MontoOfertaLicitacion}{MontoEstLicitacion} \right)}{To} \right)_{t*o}$

Tabla 23 KPI - 09 Relación promedio entre monto ofertado respecto al monto total estimado, según ofertas ganadoras y perdedoras

KPI - 10	Razones entre ganadores y perdedores según cantidad de ofertantes.
-----------------	--

Objetivo	Analizar como la presencia y el número de competidores impactan en el éxito de los proveedores en distintas ofertas. Marcar posibles tendencias de éxito en distintas cantidades de ofertas.
Meta	Un valor sobre 0.5 implicaría que la mayoría de postulantes gana las ofertas, lo cual representa una estadística positiva para cualquier empresa.
Periodo	
Dimensiones	Tiempo – Oferta – Licitación
Criterios	Se consideran tanto las ofertas adjudicadas y las ofertas perdedoras
Observación	Mientras más cercano sea el valor a 1, mayor probabilidad de éxito para cualquier ofertante. Para cada cantidad de ofertantes debe existir una misma razón, esta no necesariamente representará cada ejemplo, pero sí un aproximado global.
Fórmula	$R_i = \frac{\text{Conteo}(\text{Licitaciones ganadas con } i \text{ cantidad de ofertas})}{\text{Conteo}(\text{Licitaciones con } i \text{ cantidad de ofertas})}$

Tabla:24 KPI - 10 Razón entre ganadores y perdedores según cantidad de ofertantes a lo largo del tiempo

REQ-5 Análisis impacto Regionalización

KPI - 11	Índice de concentración de adjudicaciones por región
Objetivo	Medir la concentración del mercado en base a cómo las licitaciones son adjudicadas a los principales proveedores en diferentes regiones.
Meta	La meta ideal dependerá de la estrategia empresarial utilizada, sin embargo, en primera instancia se opta por un mercado moderado entre 1.500 y 2.500.
Periodo	Anual

Dimensiones	Tiempo – Oferta – Licitación – Región - Proveedores
Criterios	Se consideran las ofertas que resultaron en una adjudicación
Observación	Calculamos el índice de concentración mediante el índice de <i>Herfindahl - Hirschman (HHI)</i> . <i>Bajo estándar de HHI, el límite inferior es 0 y el superior 10.000, la tendencia a números bajos implica un mercado poco concentrado y altamente competitivo, mientras que un número mayor a 2.500 implica un mercado altamente concentrado</i>
Fórmula	$HHI = \sum_{(i=1)}^n (s_i)^2 ; s_i = \left(\frac{\sum \text{OfertasGanadoras}}{Ta} \right)_{p*t*l*o*pr}$

Tabla 25 KPI - 11 Índice de concentración de adjudicaciones por región y sector

KPI - 12	Porcentaje de adjudicaciones a proveedores locales vs foráneos
Objetivo	Determinar la preferencia o ventaja competitiva de los proveedores locales frente a los externos según distintas regiones.
Meta	Por el bien de la apertura de mercado se esperan valores no superiores a 25%, pues valores bajo 25% significarían una disponibilidad a mercados nacionales en general sin importar la región.
Periodo	Mensual – anual
Dimensiones	Tiempo – Licitación – Proveedor - Región
Criterios	Se consideran las licitaciones adjudicadas dentro de una región para un proveedor
Observación	Un alto porcentaje local puede significar una regionalización relevante, donde proveedores locales tienden a ganar de forma constante a proveedores foráneos.
Fórmula	$= \left(\frac{\text{Conteo(Licitaciones ganadas proveedores locales)}}{To_r} \times 100 \right)_r$

Tabla 26 KPI - 12 Porcentaje de adjudicaciones a proveedores locales en relación a foráneos

5. Descripción de las fuentes de datos de origen

Sistemas Operacionales Involucrados (SOI)

El conjunto de datos utilizado para la elaboración del Datamart fue extraído del sitio de Datos Abiertos Dirección-ChileCompra ([Mercado Público - Datos abiertos \(chilecompra.cl\)](http://Mercado Público - Datos abiertos (chilecompra.cl))), publicado por el mismo organismo Dirección ChileCompra, Ministerio de Hacienda, Gobierno de Chile, bajo leyes de transparencia.

Esta página cubre un conjunto de datos de todas las compras públicas afectas a la Ley N°19.886 a lo largo de todo Chile, se seleccionaron las compras realizadas por Sector del estado de Gob. Central/Universidades, realizando una unión entre órdenes de compra y

licitaciones para concluir con un conjunto de datos que abarca de forma completa el proceso licitativo del *Gob. Central y Universidades* de Chile. Los datos se recopilan desde enero del 2019 hasta diciembre del 2023, el formato de los datos recopilados se encuentra en un modelo .csv semestral para licitaciones y órdenes de compra de forma excluyente.

Herramienta	<i>Knime</i>
Función	<i>Herramienta para proceso de ETL</i>
Justificación	<i>Es gratuito, stand-alone y altamente visual</i>
Versión	5.2.2

Tabla 27 Knime

Herramienta	<i>PowerBI</i>
Función	<i>Herramienta de visualización de datos</i>
Justificación	<i>Incluye una versión gratuita con alta personalización de gráficos</i>
Versión	PowerBI Desktop

Tabla 28 PowerBI

5.1 Modelo Entidad-Relación de los SOI

No se tiene un modelo entidad-relación de los SOI

5.2 Modelo Físico Asociado

No se tiene un modelo físico asociado a las bases de datos.

5.3 Fuentes de datos externas

No se consideran fuentes de datos externas para la aplicación del proyecto.

5.4 Detalle de la fuente de información

Para la fuente de datos se utilizaron 2 datasets complementarios para conformar el dataset final, esto producto de una ausencia de datos pertenecientes a la región de tanto la unidad compradora como la proveedora. Se extrajo únicamente la columna “Región Unidad de Compra” y “RegiónProveedor” de la sección *Datos complementarios* de chilecompra.

“dataset_lic”

Filas	1.906.649
Columnas	74
Peso (GB)	2.65
Periodo	2019 - 2023
Escala temporal	Segundos

Tabla 29 Dataset licitaciones original

A continuación, se presenta el detalle de las columnas obtenidas del dataset.

Campo	Descripción	TipoDatos
ProveedorRUT	Rut del proveedor	Varchar
RegionProveedor	Región del proveedor	Str
FechaPublicacion	Fecha de publicación de la licitación	DateTime
FechaInicioPreguntas	Apertura periodo para consultas de la licitación	DateTime
FechaFinalPreguntas	Cierre periodo para consultas de la licitación	DateTime
FechaPublicacionRespuestas	Fecha de publicación de respuestas a las consultas	DateTime
FechaActoAperturaTecnica	Se abre un acta para licitaciones que busquen productos muy específicos, no se hace una oferta económica, solo las especificaciones del producto	DateTime
FechaActoAperturaEconomica	Se abre un acta para que puedan ofertar proveedores que cumplan con especificaciones técnicas	DateTime
FechaCierre	Fecha de cierre de la licitación	DateTime
FechaAdjudicacion	Fecha de adjudicación de la licitación	DateTime
FechaActaAprobacion	Fecha en la cual es certificado el acta que formaliza al ganador de la licitación	DateTime
SubContratacion	Indica si la licitación permite subcontratación (Si/No)	Varchar
Contrato	Usa o no contrato en la licitación	Varchar
TiempoDuracionContrato	Tiempo de duración del contrato	Int
UnidadTiempoDuracionContrato	Contrato Inmediato o en tiempo	Varchar
TipoEjecucion	Ejecución inmediata o ejecución en el tiempo	Varchar
PlazoPagoContrato	Plazo establecido para el pago	Varchar
TipoPagoContrato	Método de pago (Transferencia electrónica, cheque o ambas)	Varchar

ExtensionPlazo	SI/NO si el contrato puede extenderse en plazo	VARCHAR
UnidadCompra	Unidad de compra de la institución	VARCHAR
entCode	Identificador de la entidad de compra	VARCHAR
RubroN1	Rubro del proveedor	VARCHAR
RubroN2	Rubro del proveedor	VARCHAR
RubroN3	Rubro del proveedor	VARCHAR
CodigoProductoONU	Identificador del producto	VARCHAR
ONUProducto	Exactamente igual a NombreItem	VARCHAR
NombreItem	Nombre del producto	VARCHAR
DescripcionItem	Descripción del producto	VARCHAR
UnidadMedida	Unidad con la que se mide la cada CantidadItem	VARCHAR
Proveedor	Nombre de la empresa proveedora	VARCHAR
ActividadProveedor	Similar a rubro, pero con la descripción propia por proveedor (puede repetirse)	VARCHAR
TamanoProveedor	Micro, pequeña, mediana, grande.	VARCHAR
NombreOferta	Nombre de la oferta que responde a la licitación	VARCHAR
EstadoOferta	Se acepta o rechaza la oferta que fue realizada	VARCHAR
ResultadoOferta	Ganadora o perdedora (En este caso ganar significa que se adjudicó la licitación)	VARCHAR
Institución	Nombre de la institución	VARCHAR
UnidadCompraRut	Rut del comprador	VARCHAR
Región Unidad de Compra	Región de la institución que compra	VARCHAR
NroLicitacion	Identificador del proveedor	VARCHAR
NombreLicitacion	El nombre de la licitación	VARCHAR
TipoLicitacion	Categorías de la licitación (privada/publica y rango del monto)	VARCHAR
Descripcion	Descripción de lo que solicita la licitación	VARCHAR
MonedaLicitacion	Moneda de la licitación (CLP, ARG, etc.)	VARCHAR
ContemplaObrasPúblicas	Si la licitación contempla obras públicas	VARCHAR
BaseEstimacionMontoLicitacion	Indica el presupuesto disponible, precio referencial o si no es posible estimar.	VARCHAR
FuenteFinanciamiento	Origen de los fondos que se utilizarán para el proceso licitativo.	VARCHAR

JustificacionMontoEstimado	La justificación del monto estimado en la licitación	Varchar
UnidadTiempoEvaluacion	Indica la unidad de tiempo en relación al periodo de evaluación, esto es horas, días, semanas o meses	Varchar
LicitacionInformada	Las licitaciones “Informadas” corresponden a las licitaciones que, no obstante encontrarse excluidas de la aplicación de la ley N°19.886, son igualmente informadas a través del www.mercadopublico.cl por aplicación del Artículo 21º de la misma.	Varchar
LicitacionBaseTipo	Las Bases Tipo son una modalidad que busca, por una parte, apoyar la gestión de los organismos públicos en la elaboración de sus licitaciones desde las 100 UTM y, por otra parte, facilitar la participación de los proveedores del Estado en procesos de un mismo rubro a través de cláusulas administrativas estandarizadas. Si o no.	Varchar
TipoAdjudicacion	Adjudicación múltiple, es decir varias adjudicadas a un solo proveedor o múltiples proveedores	Varchar
TipoAprobacionAdjudicacion	Distintas formas legales mediante las cuales se formaliza la adjudicación en el proceso licitativo. (Resolución, acuerdo, etc.)	Varchar
NumeroActaAprobacion	Identificador del acta de aprobación	Varchar
TipoConvocatoria	Indica si las convocatorias son abiertas o cerradas	Varchar
NroEtapasLicitacion	Por 1 o 2 etapas	Varchar
ProhibicionSubContratacion	Muestra las diversas políticas o condiciones estipuladas en las bases de la licitación que regulan si los adjudicatarios pueden o no subcontratar a otros para realizar cualquier parte del trabajo que les ha sido conferido	Varchar
TomaRazonContraloria	Si un proceso de compra debe someterse a un control de juricidad o no producto de contratos que superan cierta cantidad de	Varchar

	UTM dependiendo de la región y el tipo de licitación.	
ContratoRenovable	Indica si el periodo del contrato es renovable	Int
ValorTiempoRenovacion	Indica el valor numérico en relación al periodo de renovación de contrato	Int
UnidadTiempoRenovacion	La unidad de tiempo asociada a la renovación del contrato	Varchar
FechaEstimadaFirmaContrato	Fecha estimada en la que se consolidará el contrato entre ambas partes proveedora y compradora	DateTime
ObservacionContrato	Describe observaciones dentro del mismo contrato	Varchar
Sector	Gobierno central/ Universidades	Varchar
EspecificacionesProveedor	Descripción detallada de la propuesta que cada proveedor presenta en respuesta a la solicitud del comprador.	Varchar
MontoNetoOferta	Monto de la oferta del proveedor sin IVA	
MontoEstimadoVisible	¿Es visible el monto estimado al momento de publicar la licitación?	Int
EstadoLicitacion	Estado final de la licitación (Adjudicada, Desierta, Revocada, etc.)	Varchar
MontoTotalOferta	Monto ofertado por el proveedor	Int
CantidadOferta	Cantidad de ítem ofertado por el proveedor	Int
CantidadItem	Cantidad de cierto producto que se pide en la licitación	Int
MontoEstimadoLicitacion	Estimación del presupuesto del comprador	Int
PublicidadOfertasTecnicas	La publicidad de ofertas técnicas es una opción que tendrán los organismos públicos de decidir si los antecedentes técnicos enviados por los proveedores serán de conocimiento público o mantendrán su carácter privado. Esto con el fin de contribuir a la transparencia de los procesos otorgando la posibilidad	Varchar

	de acceder a información relevante de la licitación.	
RazonPublicidadOfertasTecnicas	Motivo de la visualización de la oferta técnica	Varchar

Tabla 30 Columnas dataset licitaciones original

6. Diseño del Datamart

6.1 Descripción de la Etapa

En esta fase se desarrolla la creación del modelo físico de la base de datos multidimensional que contiene los datos recolectados. La meta de esta etapa es modelar una base de datos ajustada capaz de apoyar al cumplimiento del objetivo de la problemática, utilizando las tablas hecho y dimensiones como punto de partida para construir el Datamart.

6.2 Descripción de Herramientas Utilizadas

Herramienta	Dbeaver
Función	Gestor de base de datos
Justificación	Es gratuito, interfaz amigable y compatibilidad con postgresQL
Versión	24.0.4

Tabla 31 - Herramienta Dbeaver

Herramienta	Power Designer
Función	Modelamiento de base de datos
Justificación	Permite generar un script DLL y realizar el modelo de base de datos multidimensional
Versión	24.0.4

Tabla 32 - Herramienta Power Designer

Herramienta	PostgreSQL
Función	Motor de base de datos
Justificación	Motor de base de datos gratuito, confiable y robusto.
Versión	15.2

Tabla 33 - Herramienta PostgreSQL

6.3 Descripción de componentes del Modelo Multidimensional

proveedor	
idProveedor	INT GENERATED BY
rutProveedor	varchar
nombreProveedor	varchar
tamanoProveedor	varchar
actividadProveedor	varchar
regionProveedor	varchar

Ilustración 1 - Dimensión Proveedor

- Aquí se almacenan los datos relacionados con los proveedores que participan en las licitaciones.

rubro	
idRubro	INT GENERATED BY
rubro1	varchar
rubro2	varchar
macroRubro1	varchar
macroRubro2	varchar

Ilustración 2- Dimensión Rubro

- Esta dimensión se refiere a las categorías o áreas temáticas a las que se especializan los proveedores.

oferta	
idOferta	INT GENERATED BY
nombreOferta	varchar
estadoOferta	varchar
resultadoOferta	varchar

Ilustración 3 - Dimensión Oferta

- Esta dimensión almacena información sobre las ofertas presentadas por los proveedores en las licitaciones.

producto	
idProducto	INT GENERATED BY
codigoProducto	varchar
nombreItem	varchar
descripcionItem	varchar
unidadMedidaItem	varchar
categoria	varchar

Ilustración 4 - Dimensión Producto

- En esta dimensión almacenamos toda la información relacionada al producto que se solicita en la licitación.

licitacion	
nroLicitacion	varchar <pk>
tipoLicitacion	varchar
nombreLicitacion	varchar
estadoLicitacion	varchar
descripcion	varchar
contemplaObrasPublicas	varchar
licitacionInformada	varchar

Ilustración 5 - Dimensión Licitacion

- En esta dimensión abarcamos detalles de la propia licitación lo cual no debe ser confundido con el proceso licitativo.

region	
idRegion	INT GENERATED BY
nombreRegion	varchar
macroZona	varchar
nombreRegionCo	varchar

Ilustración 6 - Dimensión Region

- Se integra información relacionada a la región en términos amplios.

contrato	
idContrato	INT GENERATED BY
contrato	varchar
subContrato	varchar
unidadTiempoDuracion	varchar
plazoPagoContrato	varchar
tipoEjecucionContrato	varchar
tipoPagoContrato	varchar
extensionPlazo	varchar

Ilustración 7 - Dimensión Contrato

- En esta dimensión se contempla información referente al contrato (si es que existe) posterior a una adjudicación.

comprador	
idComprador	INT GENERATED BY
rutUnidadComprador	varchar
nombreUnidadCompradora	varchar
codigoInstitucionCompradora	varchar
nombreInstitucionCompradora	varchar
regionUnidadCompradora	varchar

Ilustración 8 - Dimensión Comprador

- En esta dimensión incluye información relevante relacionada al organismo comprador.

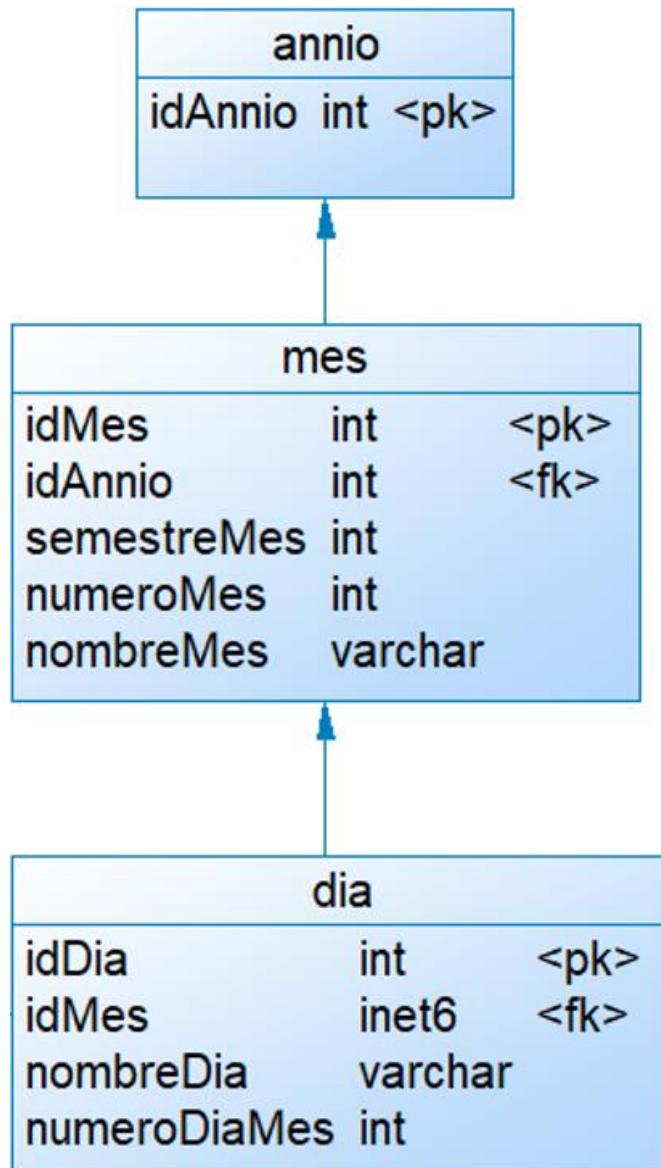


Ilustración 9 - Dimensión Tiempo

- La dimensión tiempo se compone de varias subdimensiones, siendo estas año, mes y día del registro.

hecho_licitacion		
id	int	
idLicitacion	varchar	<fk1>
idComprador	int	<fk2>
idProveedor	int	<fk3>
idOferta	int	<fk4>
idProducto	int	<fk5>
idContrato	int	<fk6>
idDia	int	<fk7>
idRubro	int	<fk8>
idRegionComprador	int	<fk9>
idRegionProveedor	int	<fk10>
fechaPublicacion	date	
fechaCierre	date	
fechaAdjudicacion	date	
cantidadItem	int	
cantidadOferta	int	
montoEstimadoLicitacion	int8	
montoNetoOferta	int8	
montoTotalOferta	int8	

Ilustración 10 - Tabla Hecho hecho_licitacion

- Se obtiene la correspondiente tabla hecho donde se alojan las distintas métricas utilizadas.

6.4 Modelo Multidimensional

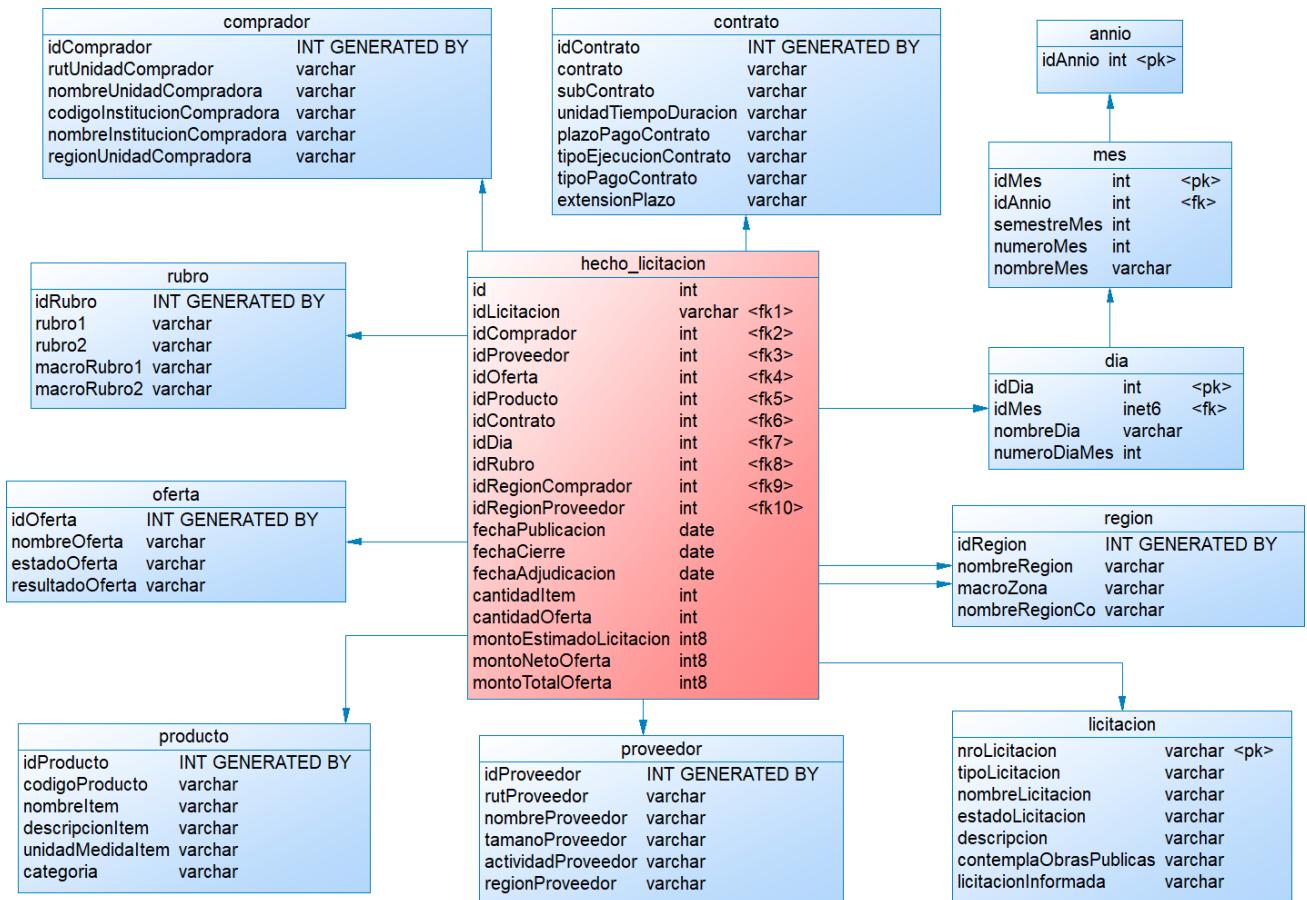


Ilustración 11 - Modelo físico Multidimensional

6.5 Justificación del Modelo

A continuación, se detallará la tabla “hecho” y sus dimensiones

Tabla 34 - Detalle de justificación dimensión Comprador

Dimensión comprador

Tabla la cual registra los distintos compradores que realizan las licitaciones.

Nombre columna	Descripción columna	Tipo de atributo
<i>idComprador</i>	Identificador de comprador	Cuantitativo
<i>rutUnidadComprador</i>	Rut de la unidad compradora	Cualitativo
<i>nombreUnidadCompradora</i>	Nombre de la unidad compradora	Cualitativo

<i>codigoInstitucionCompradora</i>	Código único de la institución que realiza la licitación.	Cualitativo
<i>nombreInstitucionCompradora</i>	Nombre de la institución compradora.	Cualitativo
<i>regionUnidadCompradora</i>	Región a la que pertenece la unidad compradora	Cualitativa

Tabla 35 Detalle dimensión comprador

Dimensión rubro

Tabla que registra los rubros del proveedor, es decir, de quien está postulando al proceso licitativo.

Nombre columna	Descripción columna	Tipo de atributo
<i>idRubro</i>	Identificador de rubro del proveedor	Cuantitativo
<i>rubro1</i>	Primer rubro en el que se especializa un proveedor	Cualitativo
<i>rubro2</i>	Segundo rubro en el que se especializa un proveedor	Cualitativo
<i>macroRubro1</i>	Primer macrorubro en el que se especializa el proveedor	Cualitativo
<i>macroRubro2</i>	Segundo macrorubro en el que se especializa el proveedor	Cualitativo

Tabla 36 Detalle dimensión rubro

Dimensión oferta

Tabla que almacena datos respectivos a la oferta realizada en el proceso licitativo.

Nombre columna	Descripción columna	Tipo de atributo
<i>idOferta</i>	Identificador de la oferta del proveedor	Cuantitativo
<i>nombreOferta</i>	Nombre con el que se presenta la oferta	Cualitativo
<i>estadoOferta</i>	Representa si la oferta fue aceptada	Cualitativo
<i>resultadoOferta</i>	Indica si la oferta fue adjudicada	Cualitativo

Tabla 37 Detalle dimensión oferta

Dimensión producto

Tabla respectiva a los datos del producto licitado.

Nombre columna	Descripción columna	Tipo de atributo
<i>idProducto</i>	Identificador de producto	Cuantitativo
<i>codigoProducto</i>	Un código único que identifica a cada producto	Cualitativo
<i>nombreItem</i>	Nombre del producto	Cualitativo

<i>descripciónItem</i>	Describe al producto correspondiente	Cualitativo
<i>unidadMedidaItem</i>	La unidad de medida con la que se mide el ítem licitado	Cualitativo
<i>categoria</i>	Categoría de productos a la que pertenece un determinado ítem	Cualitativo

Tabla 38 Detalle dimensión producto

Dimensión proveedor

Tabla que almacena los datos del proveedor que postula a la licitación.

Nombre columna	Descripción columna	Tipo de atributo
<i>idProveedor</i>	Identificador del proveedor	Cuantitativo
<i>rutProveedor</i>	Rut del proveedor	Cualitativo
<i>nombreProveedor</i>	Nombre del proveedor	Cualitativo
<i>tamanoProveedor</i>	Tamaño del proveedor	Cualitativo
<i>actividadProveedor</i>		Cualitativo
<i>regionProveedor</i>	Región a la que pertenece el proveedor	Cualitativo

Tabla 39 Detalle dimensión proveedor

Dimensión contrato

Tabla que almacena datos relacionados al contrato posterior al proceso de adjudicación

Nombre columna	Descripción columna	Tipo de atributo
<i>idContrato</i>	Identificador del contrato	Cuantitativo
<i>contrato</i>	Si lleva contrato o no la licitación	Cualitativo
<i>SubContrato</i>		Cualitativo
<i>unidadTiempoDuracion</i>	Unidad de tiempo en la que se maneja la duración del contrato	Cualitativo
<i>plazoPagoContrato</i>	Plazo determinado para el pago según lo estipulado en el contrato	Cualitativo
<i>tipoEjecucionContrato</i>		Cualitativo
<i>tipoPagoContrato</i>	Medio por el cual se realiza el pago	Cualitativo

Tabla 40 Detalle dimensión contrato

Dimensión licitacion

Dimensión que almacena información respecto a la propia licitación

Nombre columna	Descripción columna	Tipo de atributo
<i>nroLicitacion</i>	Identificador de la licitacion	Cualitativo
<i>tipoLicitacion</i>		Cualitativo

<i>nombreLicitacion</i>		Cualitativo
<i>estadoLicitacion</i>	Tamaño del proveedor	Cualitativo
<i>descripcion</i>	Descripción de licitación	Cualitativo
<i>contemplaObrasPublicas</i>	Indica si la licitación contempla obras públicas	Cualitativo
<i>licitacionInformada</i>	Indica si la licitación es de tipo informada	Cualitativo

Tabla 41 Detalle dimensión licitación

Dimensión Tiempo

Dimensión que nos permite realizar análisis de licitaciones de manera temporal

Nombre columna	Descripción columna	Subdimensión
<i>idAnno</i>	Identificador del año	anno
<i>idMes</i>	Identificador del mes	mes
<i>semestreMes</i>		mes
<i>numeroMes</i>	Numero con el que se corresponde el mes	mes
<i>nombreMes</i>	Nombre del mes	mes
<i>idDia</i>	Identificador del día	dia
<i>nombreDia</i>	Nombre del día	dia
<i>numeroDiaMes</i>		dia

Tabla 42 Detalle dimensión tiempo

Hecho hecho_licitacion

Tabla que almacena datos relacionados al proceso licitativo.

Nombre columna	Descripción columna	Tipo de atributo
<i>id</i>	Identificador de la tabla hecho	Cuantitativo
<i>fechaPublicacion</i>	Fecha de publicación de la licitación	Cualitativo
<i>fechaCierre</i>	Fecha de cierre de la licitación	Cualitativo
<i>fechaAdjudicacion</i>	Fecha en la que se adjudico la licitación	
<i>cantidadItem</i>	Cantidad del producto buscado en la licitación	Cuantitativo
<i>cantidadOferta</i>	Cantidad del producto ofertado por el proveedor	Cuantitativo
<i>montoEstimadoLicitacion</i>	Presupuesto estimado de la licitación	Cuantitativo
<i>montoNetoOferta</i>	Monto neto del total ofertado por el proveedor	Cuantitativo
<i>montoTotalOferta</i>	Monto total ofertado por el proveedor en la licitación	Cuantitativo

Tabla 43 Detalle tabla hecho

6.7 Modelo Físico Multidimensional

6.8 Script DDL

Se adjuntará el script DLL en el anexo.

7. Análisis exploratorio de datos

7.1 Descripción de la etapa

En la siguiente sección, presentamos un análisis exploratorio detallado de los datos contenidos en el dataset seleccionado. Este proceso se estructuró en varias etapas clave para mejorar la calidad y la interpretabilidad de la información:

1. **Exploración inicial:** Se realizó un análisis superficial de los datos, identificando su estructura y sus tipos de valores para prepararlos para la siguiente etapa de limpieza.
2. **Limpieza:** Este es el proceso más largo de las etapas. Se aplicaron criterios para normalización de datos, como, por ejemplo, nombres de regiones, borrado de columnas, corrección de errores, borrado de valores atípicos, etc. Luego se aplican una serie de criterios para filtrar los datos válidos.
3. **Complementarios:** Luego de dilucidar el tipo de datos obtenidos en la exploración inicial y su posterior limpieza, se agregan columnas que ayuden a englobar datos que pueden ser demasiado específicos o generalizar más ciertos datos.
4. **Gráficos:** Finalmente se realiza la generación de gráficos para identificar información relevante o valiosa de una forma visual, incluyendo distintos tipos de gráficos dependiendo del tipo de dato.

7.2 Descripción de Herramientas Utilizadas.

Se hizo uso de diversidad de herramientas para facilitar el análisis exploratorio, entre ellas Python, debido a la gran cantidad de librerías que provee para la ciencia de datos, como, por ejemplo, Pandas, Numpy, Matplotlib, entre otras.

Herramienta	Descripción
Visual Studio Code	Editor de código abierto y gratuito, altamente personalizable, compatible con gran cantidad de lenguajes de programación,

	además de ser compatible con varias plataformas
Jupyter notebook	Herramienta para visualizar datos en big data y ciencia de datos que es altamente compatible con <i>Python</i>
Python	Lenguaje de programación altamente versátil para la ciencia de datos gracias a sus librerías y sintaxis simple
Pandas – Numpy – Matplotlib - Seaborn	Librerías de python para el análisis y visualización de datos

Tabla 44 - herramientas utilizadas OLAP

7.3 Estadísticos Descriptivos

Exploración inicial

Descripción de variables cuantitativas

Estadístico	MontoEstimadoLicitacion	CantidadItem
Count	1906649.0	1906647.0
Mean	383827922.996278	18354.559275
Std	18835947899.824329	621521.800588
Min	0.0	0.0
25%	2420800.0	1.0
50%	8755183.0	2.0
75%	37710321.0	16.0
Max	5856793744545.0	140400000.0
Kurtosis	23209.742801	28152.767374
Skewness	121.206983	134.881237
Mode	0.0	1.0
Nulls	0.0	2.0
% Nulls	0.00%	0.00%

Tabla 45 Descripción variables cuantitativas n°1

Estadístico	MontoNetoOferta	MontoTotalOferta	CantidadOferta
Count	1906649.0	1906649.0	1906649.0
Mean	4725135576.154398	10801324719.277279	7664.495692
Std	6517891622463.025391	8071539790401.210938	448424.147851
Min	-1.0	-1.0	0.0
25%	0.0	0.0	0.0
50%	0.0	0.0	0.0
75%	22679.0	186400.0	1.0
Max	9000000000000000.0	9000000000000000.0	140400000.0
Kurtosis	1906643.988357	972520.53038	63103.11869
Skewness	1380.813522	954.015283	215.61119
Mode	0.0	0.0	0.0
Nulls	1380.813522	0.0	0.0
% Nulls	0.00%	0.00%	0.00%

Tabla 46 Descripción variables cuantitativas n°2

Dentro de la tabla se puede observar un dato extraño tanto en “MontoTotalOferta” como en “MontoNetoOferta” en donde su valor mínimo es de -1.0. Esto es ilógico tratándose de montos de pesos chilenos, por lo que en el proceso de búsqueda de información en la fuente de datos se descubrió que se tratan de montos que la unidad de compra falló en registrar correctamente.

Además, se observa que muchas de las modas corresponden a el valor 0, implicando que puede distorsionarse los análisis producto de esto.

Estadístico	MontoEstimadoLicitacion
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	1843128.0

Tabla 47 Visualización MontoEstimadoLicitacion

Se consideran válidos montos estimados mayores a 100 CLP.

Estadístico	MontoTotalOferta
Total de datos	1906649.0

Datos nulos	0.0
Datos válidos	981753.0

Tabla 48 Visualización MontoTotalOferta

Se consideran válidas ofertas mayores o iguales a 0 CLP, pero solo si en la columna “ResultadoOferta” es “Perdedora”, es decir, no adjudicada.

Estadístico	MontoNetoOferta
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	981753.0

Tabla 49 Visualización MontoNetoOferta

Se consideran válidas ofertas mayores o iguales a 0 CLP, pero solo si en la columna “ResultadoOferta” es “Perdedora”, es decir, no adjudicada.

Estadístico	CantidadItem
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	1906637.0

Tabla 50 Visualización CantidadItem

Se consideran válidos la cantidad de ítem mayores a 0.

Estadístico	CantidadOferta
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	981756.0

Tabla 51 Visualizacion CantidadOferta

Se consideran válidos valores en donde “CantidadOferta” mayores o iguales a 0, en donde la columna “ResultadoOferta” deberá ser “Perdedora”, es decir, no adjudicada

Descripción de variables cualitativas

	NroLicitacion	NombreLicitacion	TipoLicitacion	Descripcion	MonedaLicitacion
count	1906649	1906649	1906649	1906649	1906649
unique	71618	68321	12	68309	5
top	45-22-LR23	Adquisición de libros para Programa Bibliometro	Licitación Pública Entre 100 y 1000 UTM (LE)	La presente licitación tiene por objeto regula...	CLP
freq	25182	52233	888603	25504	1878049
nulos	0	0	0	0	0
porcentaje de nulos	0.00%	0.00%	0.00%	0.00%	0.00%

Tabla 52 - Descripción variables cualitativas n°1

	MontoEstimadoVisible	BaseEstimacion	MontoLicitacion	FuenteFinanciamiento
count	1906649		1906525	747427
unique	2		3	6791
top	Si	Presupuesto disponible		SNBP
freq	966681		1184934	48935
nulos	0		124	1159222
porcentaje de nulos	0.00%		0.01%	60.80%

Tabla 53 Descripción variables cualitativas n°2

	JustificacionMontoEstimado	FechaPublicacion	FechaInicioPreguntas	FechaFinalPreguntas
count	891410	1906649	1906649	1906649
unique	12524	69799	58552	32065
top	0 06-01-2022 12:26:05	06-01-2022 13:00:00	17-01-2022 20:00:00	
freq	268243	25504	25504	25513
nulos	1015239	0	0	0
porcentaje de nulos	53.25%	0.00%	0.00%	0.00%

Tabla 54 Descripción variables cualitativas n°3

	FechaPublicacionRespuestas	FechaActoAperturaTecnica	FechaActoAperturaEconomica	FechaCierre
count	1906649	1906649	1906649	1906649
unique	29676	34408	34584	24472
top	20-01-2022 18:00:00	07-02-2022 15:01:00	07-02-2022 15:01:00	07-02-2022 15:01:00
freq	25537	25809	25806	25523
nulos	0	0	0	0
porcentaje de nulos	0.00%	0.00%	0.00%	0.00%

Tabla 55 Descripción variables cualitativas nº4

	FechaAdjudicacion	FechaEntregaEnSoporteFisico	FechaEstimadaEvaluacionOfertas
count	1905892	52049	579459
unique	71509	1086	74
top	06-10-2023 10:35:17	09-12-2021 0:00:00	01-01-1900 0:00:00
freq	25182	2291	455946
nulos	757	1854600	1327190
porcentaje de nulos	0.04%	97.27%	69.61%

Tabla 56 Descripción variables cualitativas nº4

	UnidadTiempoEvaluacion	EstadoLicitacion	ContemplaObrasPublicas	LicitacionInformada
count	1877192	1906649	1345848	1906649
unique	4	5	2	1
top	Horas	Adjudicada	NO	NO
freq	1291961	1898090	1340495	1906649
nulos	29457	0	560801	0
porcentaje de nulos	1.54%	0.00%	29.41%	0.00%

Tabla 57 Descripción variables cualitativas nº5

	TipoAdjudicacion	TipoAprobacionAdjudicacion	NumeroActaAprobacion	FechaActaAprobacion
count	1906649	1906649	1906203	1906649
unique	1	6	21854	1500
top	Multiple con Orden de Compra	Resolucion	6548	06-11-2018 0:00:00
freq	1906649	1686814	27218	31466
nulos	0	0	446	0
porcentaje de nulos	0.00%	0.00%	0.02%	0.00%

Tabla 58 Descripción variables cualitativas nº6

	TipoConvocatoria	NroEtapasLicitacion	SubContratacion	ProhibicionSubContratacion
count	1906649	1906649	1906649	427422
unique	2	2	2	6498
top	Abierta	Una etapa	Si	Remitirse a las bases
freq	1904085	1849018	966681	42303
nulos	0	0	0	1479227
porcentaje de nulos	0.00%	0.00%	0.00%	77.58%

Tabla 59 Descripción variables cualitativas n°7

	TomaRazonContraloria	PublicidadOfertasTecnicas	RazonPublicidadOfertasTecnicas	Contrato
count	1769392	1686136		353789 1769392
unique	2	2		888 3
top	No	Si	Todas las ofertas técnicas serán visibles al p...	No
freq	1715059	1670955		338446 926591
nulos	137257	220513		1552860 137257
porcentaje de nulos	7.20%	11.57%		81.44% 7.20%

Tabla 60 Descripción variables cualitativas n°8

	UnidadTiempoDuracionContrato	FechaEstimadaFirmaContrato	TipoEjecucion	PlazoPagoContrato
count	1450396	99804	1330257	1450396
unique	4	2453	2	3
top	Horas	26-12-2019 0:00:00	Contrato de Ejecucion Inmediata	A 30 dias
freq	904297	5927	838739	1433949
nulos	456253	1806845	576392	456253
porcentaje de nulos	23.93%	94.77%	30.23%	23.93%

Tabla 61 Descripción variables cualitativas n°9

	TipoPago	ObservacionContrato	ExtensionPlazo	UnidadCompra	UnidadCompraRUT
count	1345724	442786	1768113	1906649	1906649
unique	4	6787	2	1665	528
top	Cheque	La oferta económica debe incluir todo tipo de ...	NO	Servicio Nacional de Capacitación y Empleo	60.905.000-4
freq	776124	64093	1498016	74612	175384
nulos	560925	1463863	138536	0	0
porcentaje de nulos	29.42%	76.78%	7.27%	0.00%	0.00%

Tabla 62 Descripción variables cualitativas n°10

	Institucion	Sector	RubroN1	RubroN2	RubroN3
count	1906649	1906649	1906647	1906647	1906647
unique	257	1	56	345	1730
top	DIRECCION GENERAL DE GENDARMERIA DE CHIL	GOB. CENTRAL, UNIVERSIDADES	Productos impresos y publicaciones	Medios impresos	Publicaciones impresas
freq	301935	1906649	283918	255043	255043
nulos	0	0	2	2	2
porcentaje de nulos	0.00%	0.00%	0.00%	0.00%	0.00%

Tabla 63 Descripción variables cualitativas n°11

	ONUProducto	Nombrelitem	DescripcionItem	UnidadMedida	Proveedor
count	1906647	1906647	1900347	1906649	1906649
unique	8730	8805	237622	97	39909
top	Libros de referencias para bibliotecas	Libros de referencias para bibliotecas	Convenio marco para la adquisición de servicio...	Unidad	CENTRO DE EVALUACION Y CERTIFICACION DE COMPET...
freq	105522	105522	715	1515665	39129
nulos	2	2	6302	0	0
porcentaje de nulos	0.00%	0.00%	0.33%	0.00%	0.00%

Tabla 64 Descripción variables cualitativas n°12

	ProveedorRUT	ActividadProveedor	TamanoProveedor	NombreOferta
count	1906648	1831608	1906649	1906649
unique	39947	19074	5	209257
top	76.177.214-7	OTRAS ACTIVIDADES EMPRESARIALES N.C.P.	Micro	COTIZACION DE LIBROS
freq	39129	58146	653237	9272
nulos	1	75041	0	0
porcentaje de nulos	0.00%	3.94%	0.00%	0.00%

Tabla 65 Descripción variables cualitativas n°13

	EspecificacionesProveedor	EstadoOferta	MonedaOferta	ResultadoOferta
count	882640	1906649	882667	1906649
unique	593698	2	5	2
top	Adjunto anexo económico.	Aceptada	CLP	Perdedora
freq	3470	1758574	860074	981756
nulos	1024009	0	1023982	0
porcentaje de nulos	53.71%	0.00%	53.71%	0.00%

Tabla 66 Descripción variables cualitativas n°14

Limpieza

Normalización:

La normalización fue aplicada principalmente a los nombres de las regiones provenientes de la unidad de compra y del proveedor. Esto debido a que muchas de las regiones variaban tanto en formato de mayúsculas y minúsculas, abreviaciones de nombres y uso de números romanos para identificar la región.

Se siguieron los siguientes pasos para lograr la normalización:

1. Conversión de la totalidad del texto a minúsculas y se eliminación de los espacios al principio y al final.
2. Se crea un diccionario para mapear varias regiones de formas de nombres de regiones a una forma estandarizadas. Las claves de este diccionario son expresiones regulares que coinciden con diferentes formas del nombre de una región, y los valores son los nombres estandarizados.
3. Finalmente, el dataframe filtrado es almacenado para su posterior uso.

Proceso de limpieza:

1. Se eliminaron todas las filas que contienen valores “NaN”.

MontoEstimadoLicitacion	LicitacionBaseTipo	TiempoDuracionContrato	ContratoRenovable
1906649.0	0.0	1906649.0	0.0
383827922.996278	NaN	119.365489	NaN
18835947899.824329	NaN	65023.399277	NaN
0.0	NaN	0.0	NaN
2420800.0	NaN	0.0	NaN
8755183.0	NaN	0.0	NaN
37710321.0	NaN	4.0	NaN
5856793744545.0	NaN	44227650.0	NaN

Tabla 67 Visualización de datos NaN

2. Se realiza un filtro para incluir datos que abarquen montos de licitaciones superiores a 100 CLP.
3. En variables cuantitativas como “MontoTotalOferta” y “MontoNetoOferta” se presenta un dato extraño, donde su mínimo de -1. Como se mencionó anteriormente en la exploración inicial, esto pertenece a un error humano, además de presentar gran cantidad de valores 0 y 1 como se muestra en la siguiente ilustración. Para emendar estos datos inválidos dada su inadmisibilidad en contexto de licitaciones, se ha igualado el monto total de la oferta (en donde esta resulta en adjudicación) al monto estimado de la licitación.

MontoTotalOferta
0.0
1.0
120000.0
60000.0
30000.0
...

Tabla 68 Cantidad de cada variables en "MontoTotalOferta"

MontoNetoOferta	
0.0	1023800
1.0	29271
15000.0	2752
12000.0	2475
10000.0	2394
...	

Tabla 69 Cantidad de cada variable en "MontoNetoOferta"

4. En la columna “CantidadOferta”, que corresponde a la cantidad de producto o servicio ofertado por el proveedor, se obtenían casos en donde la cantidad ofertada por el proveedor era 0 y aun así se adjudicaba una licitación. Para este caso se realizó una modificación similar al proceso anterior, igualando la cantidad ofertada por el proveedor a la cantidad de ítems pedidos en la licitación adjudicada.

CantidadOferta	
0.0	1023988
1.0	420924
2.0	65330
3.0	38315
10.0	35973
...	

Tabla 70 Cantidad de cada variable en "CantidadOferta"

5. Después, en la cantidad de oferta, se aplica un filtro en donde el valor si el valor es distinto de 0 se calcula el cociente entre el “MontoTotalOferta” y “CantidadOferta” cambiando los valores para que coincidan la cantidad de ítems de la licitación y luego se reemplazan los valores en “MontoNetoOferta” con los valores de las filas donde la “CantidadOferta” es igual a 1
6. Luego filtramos la nueva columna por valores mayores a 0, dado que es aceptable que un monto oferta pueda ser 0, pero no que sea equivalente a -1.
7. Se realiza un reemplazo de la columna “MontoTotalOferta” original por la nueva columna, asignándole el mismo nombre.

8. Antes de finalizar se realiza un filtrado por “MontoTotalOferta” en donde los valores sean inferiores o iguales a 100.000.000 CLP.

9. Se realiza una limpieza final de outliers en los montos estimados de licitación.

Complementarios:

Se crean columnas para complementar el análisis exploratorio, entre ellas se encuentran:

- **MacroRubro1:** Se engloban los rubros de los proveedores presentes en la columna RubroN1.
- **MacroRubro2:** Se engloban los rubros de los proveedores presentes en la columna RubroN2 debido a su gran cantidad.
- **Categoría:** Se crean categorías para la gran variedad de productos que existen dentro del dataset.
- **MacroZona:** Se crean macrozonas para englobar los datos relacionados a la región.
- **NombreRegionCo:** Para agregar de prefijo “Región” al nombre de las regiones.

A continuación, se presentan las variables cuantitativas posterior al proceso de limpieza

Estadístico	MontoEstimadoLicitacion
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	818707.0

Tabla 71 Visualización MontoEstimadoLicitacion - Limpio

Estadístico	MontoNetoOferta
Total de datos	1906649.0
Datos nulos	0.0
Datos válidos	818707.0

Tabla 72 Visualización MontoNetoOferta - Limpio

Estadístico	MontoTotalOferta
Total de datos	1906649.0

Datos nulos	0.0
Datos válidos	818707.0

Tabla 73 Visualización MontoTotalOferta - Limpio

Estadístico	CantidadItem
Total de datos	1906647.0
Datos nulos	0.0
Datos válidos	818707.0

Tabla 74 Visualización CantidadItem - Limpio

Estadístico	CantidadOferta
Total de datos	1906647.0
Datos nulos	0.0
Datos válidos	818707.0

Tabla 75 Visualización CantidadOferta - Limpio

A continuación, agrupamos todos los datos con una cantidad alta de nulos en una sola tabla para mejor visualización.

Posteriormente se eliminan columnas dado que:

- No son relevantes para nuestro análisis
- Poseen una alta cantidad de nulos
- Presentan un único valor repetido
- Presentan alta cantidad de valores NaN

7.4 Gráficos

A continuación, se presenta la removida de outliers presentes en la columna “MontoEstimadoLicitacion”

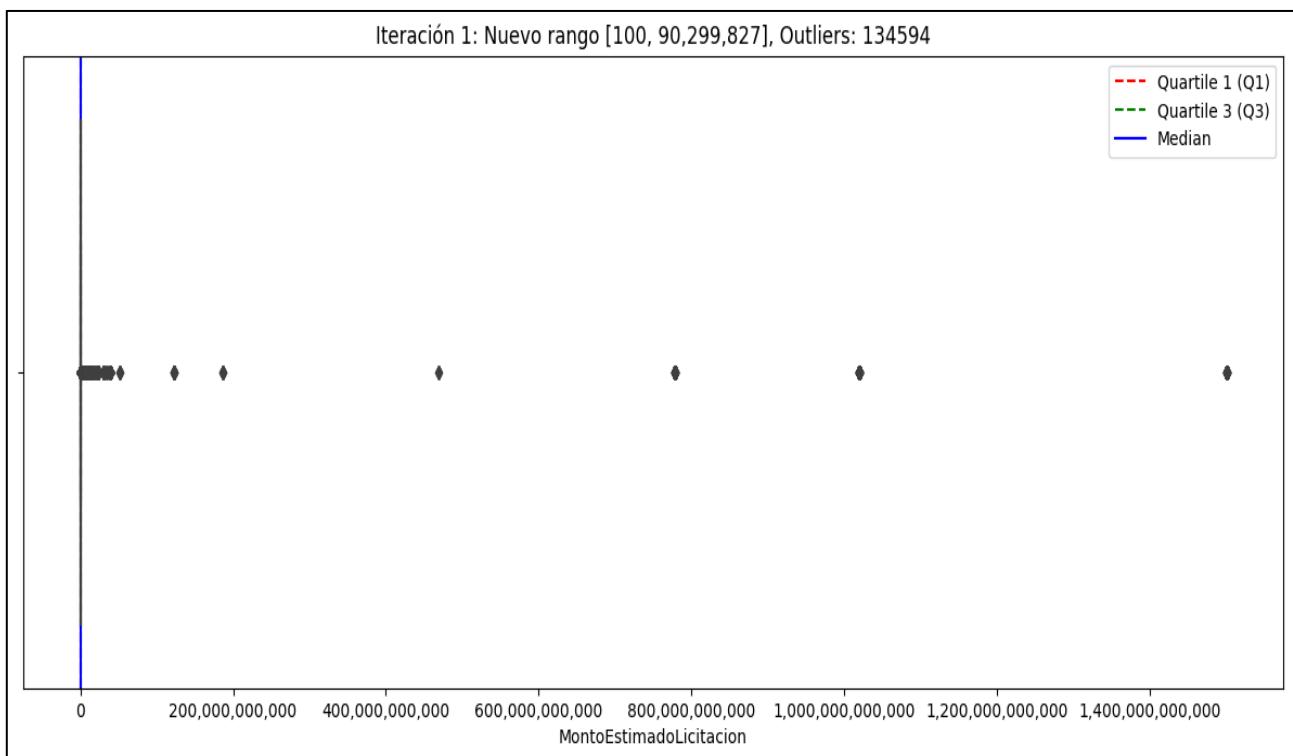


Gráfico 1 Boxplot "MontoEstimadoLicitación" - Con outliers

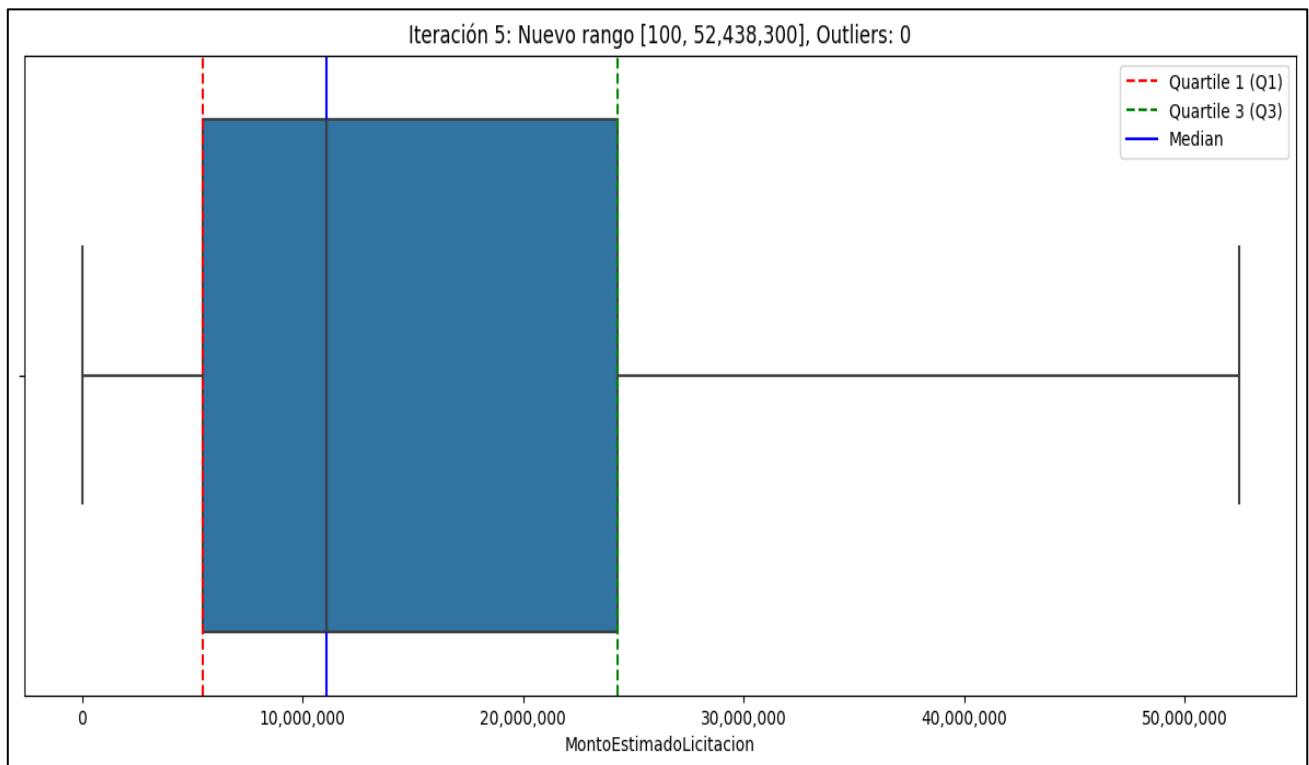


Gráfico 2 Boxplot "MontoEstimadoLicitación" - Sin outliers

- En el gráfico se puede observar que hay un rango estrecho entre el cuartil 1 y el cuartil 3 en comparación con el límite superior, indicando una mayor concentración de montos.
- Por otro lado, se observa que la mediana está bastante cerca del cuartil 1, lo que sugiere un sesgo hacia valores más bajos
- Se observa como producto de la limpieza de outliers el límite máximo ahora es 52.438.300.

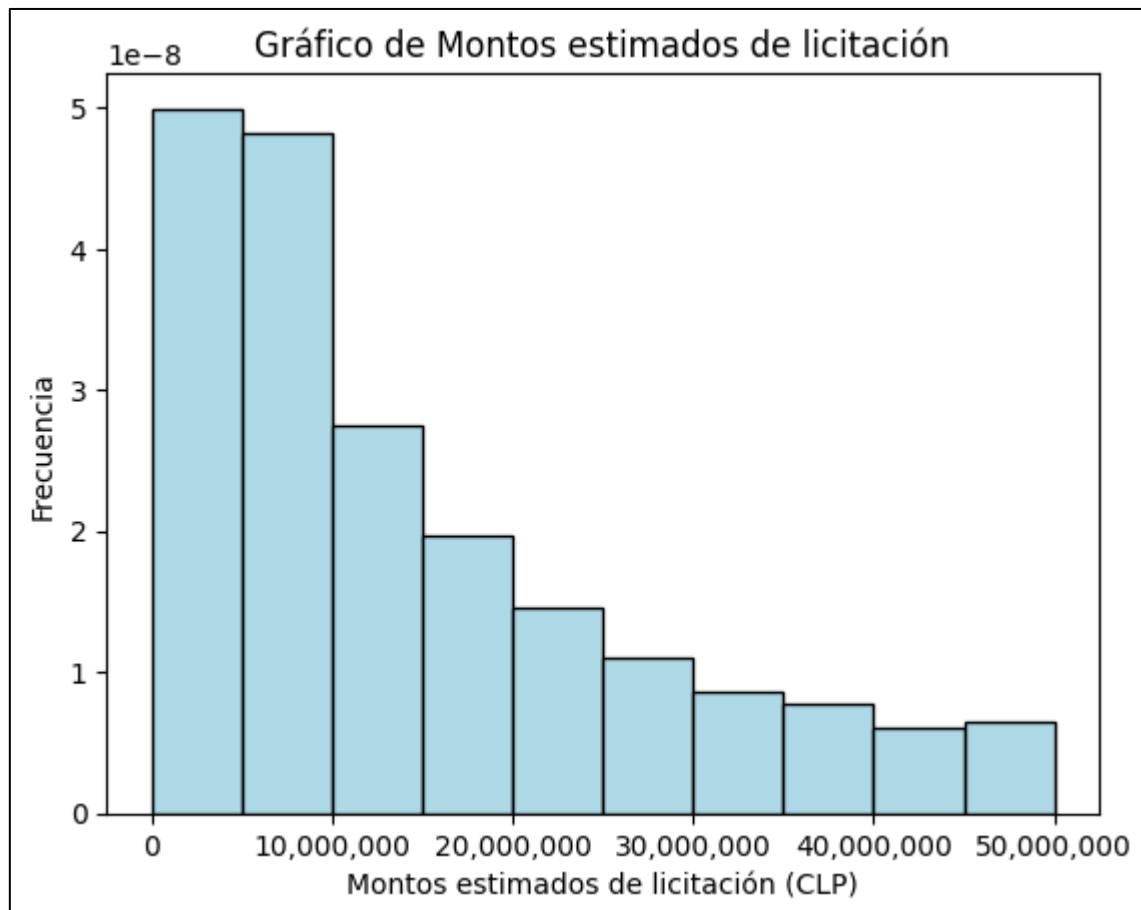


Gráfico 3 Montos estimados de licitación

- Se puede observar como la gran mayoría de montos se concentran entre los rangos 0 y 10.000.000.
- Por otro lado, se observa una tendencia al decrecimiento de la cantidad de montos estimados a medida que aumenta el valor de este.

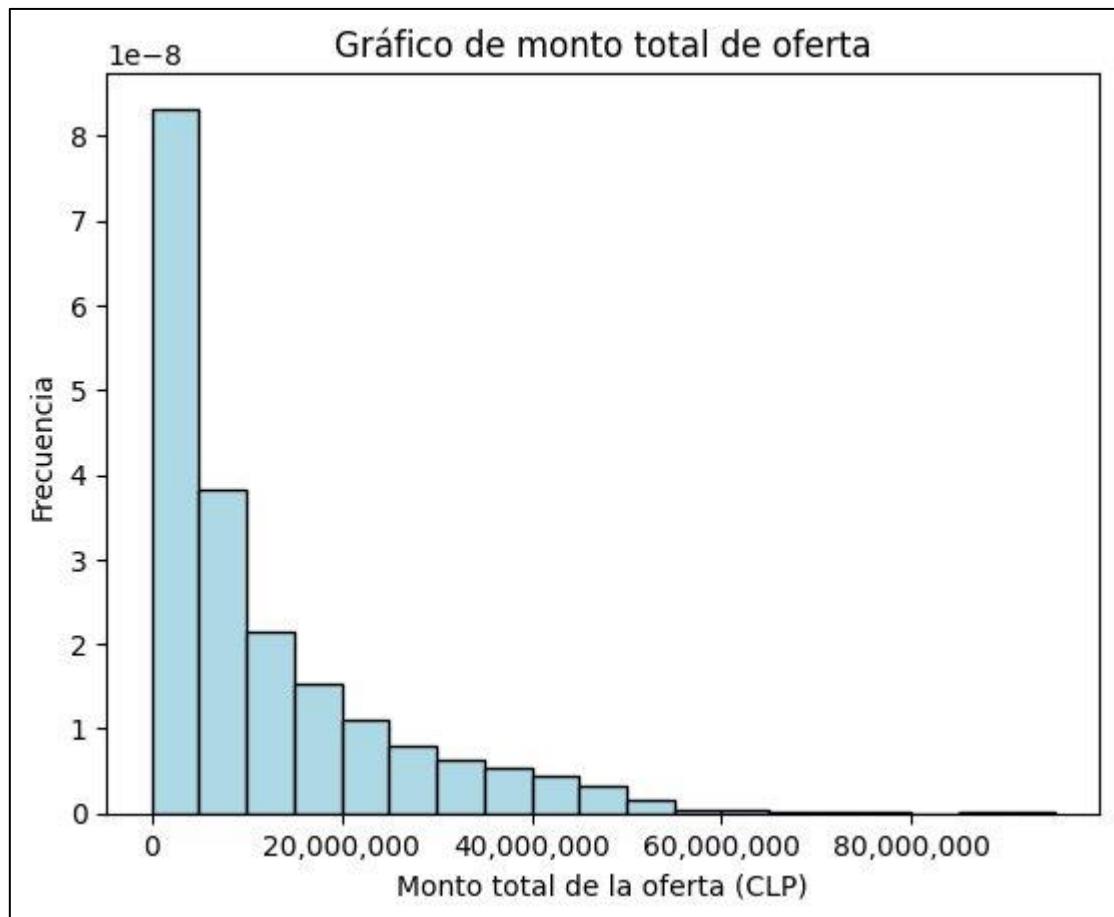


Gráfico 4 Monto total de oferta

- Se puede observar una alta concentración de los montos de oferta entre el rango de 0 y 20.000.000.
 - Se puede observar un decrecimiento bastante pronunciado de los montos de oferta durante los rangos 0 y 20.000.000 para luego estabilizarse a medida que aumentan los montos de oferta.

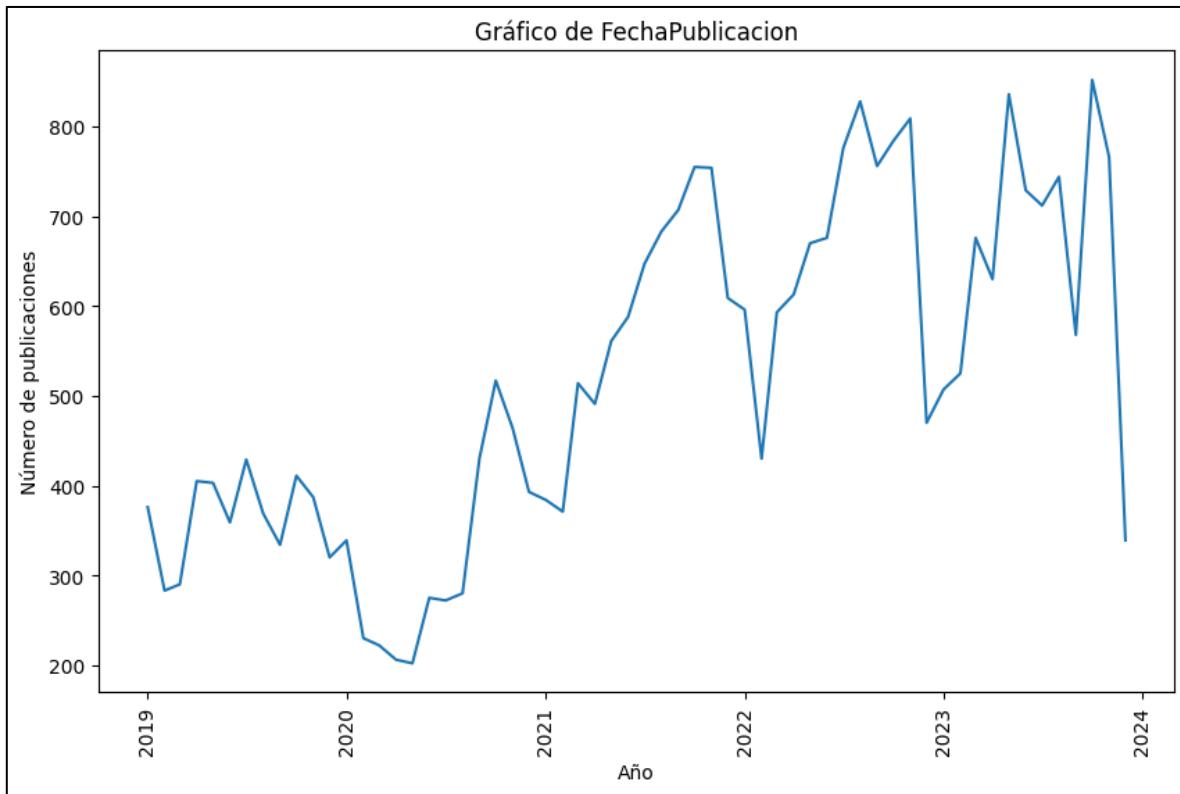


Gráfico 5 Fecha de publicación

- En el gráfico se puede apreciar como había una tendencia a la baja desde 2019 encontrando su valle más bajo a mediados del año 2020 disminuyendo considerablemente las publicaciones.
- En el gráfico anterior se puede observar que, en un periodo de 5 años, enero 2019 a diciembre 2023, existe una tendencia al alza en el número de licitaciones por año.

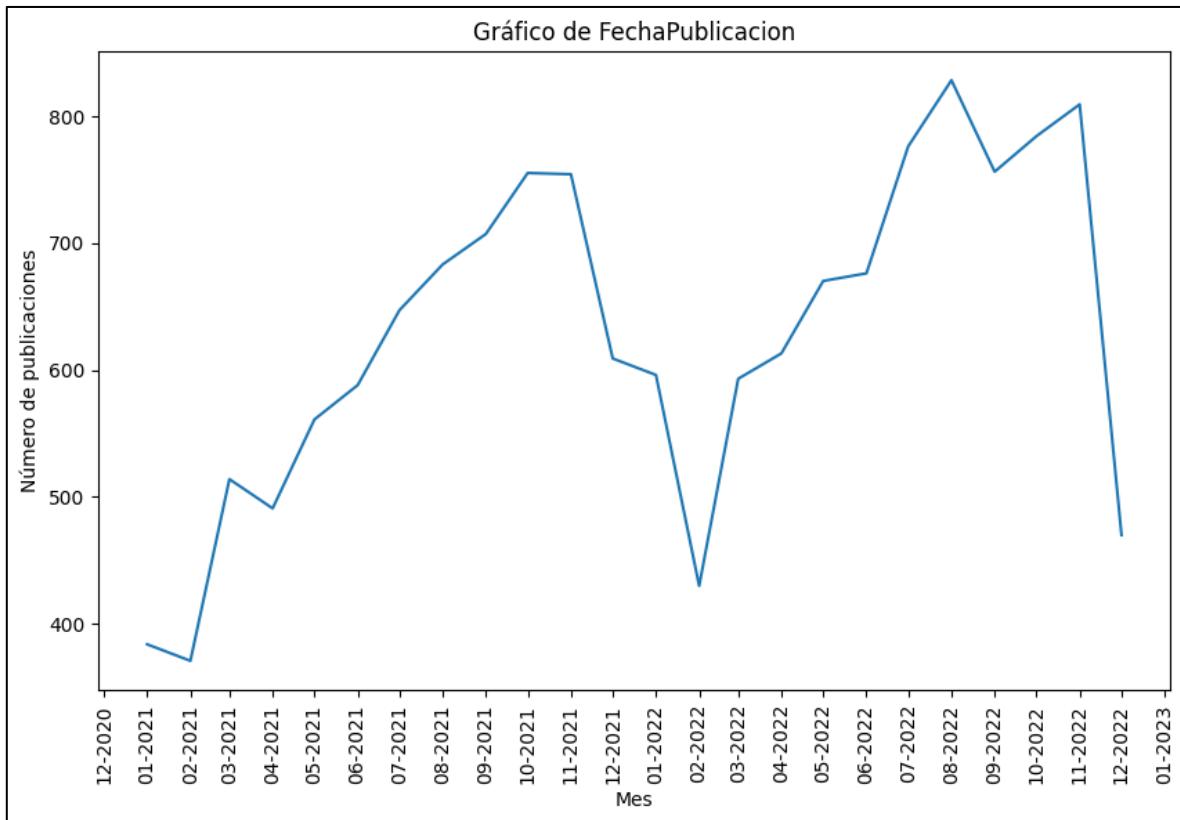


Gráfico 6 Fecha de publicación - Zoom

- Por su parte, hay un patrón en el número de licitaciones por mes, se observa, por ejemplo. Que los meses de enero, febrero, marzo son meses con muy pocas licitaciones y a partir de abril existe un gran número de licitaciones en comparación a los meses anteriores. Lo anterior podría tener relación con la llegada primer decreto presupuestario del año para la mayoría de las instituciones públicas.

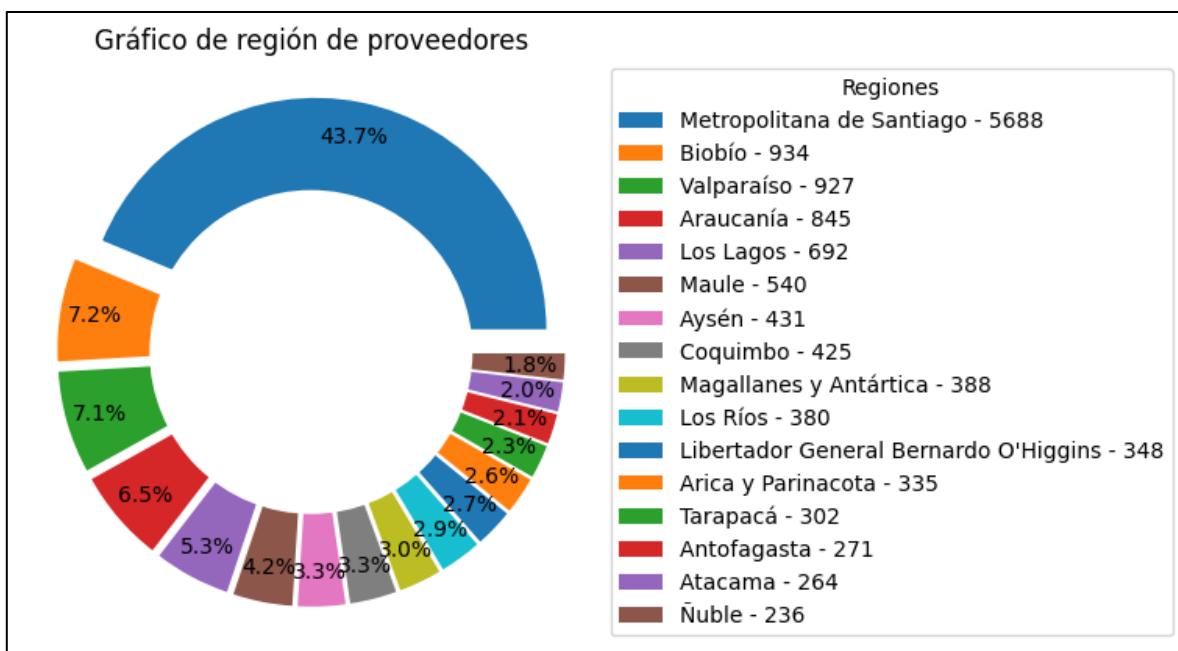


Gráfico 7 Región de proveedores

- En el gráfico se puede observar la región los proveedores en Chile, en donde se puede apreciar una gran concentración de proveedores en la región metropolitana de Santiago con un 43.7%.
- Por otro lado, se puede observar como el top 4 de regiones (Metropolitana de Santiago, Biobío, Valparaíso y Araucanía) siendo un 25% del total de regiones que conforman un 64.5% de los proveedores.
- Se puede observar cómo la región del Ñuble es la que contiene el menor porcentaje de proveedores con solo 1.8%

Gráfico de región de unidad de compra

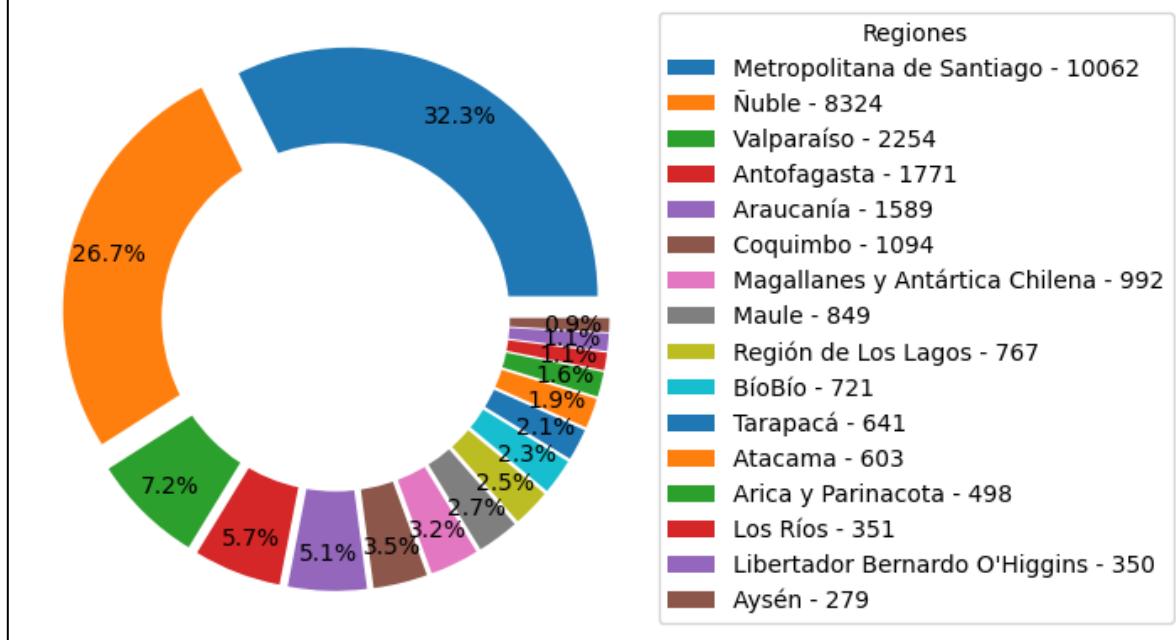


Gráfico 8 Región de unidad de compra

- Se puede observar en el gráfico que las regiones “Metropolitana de Santiago” y “El Ñuble” siendo el 12.5% de las regiones, abarcan un 59% de los proveedores.
- Además, se observa la región Metropolitana de Santiago abarca la mayor cantidad de unidades compradoras con un 32.3%.
- Por otro lado, la región de Aysén es la que menor concentración posee con un 0.9%.

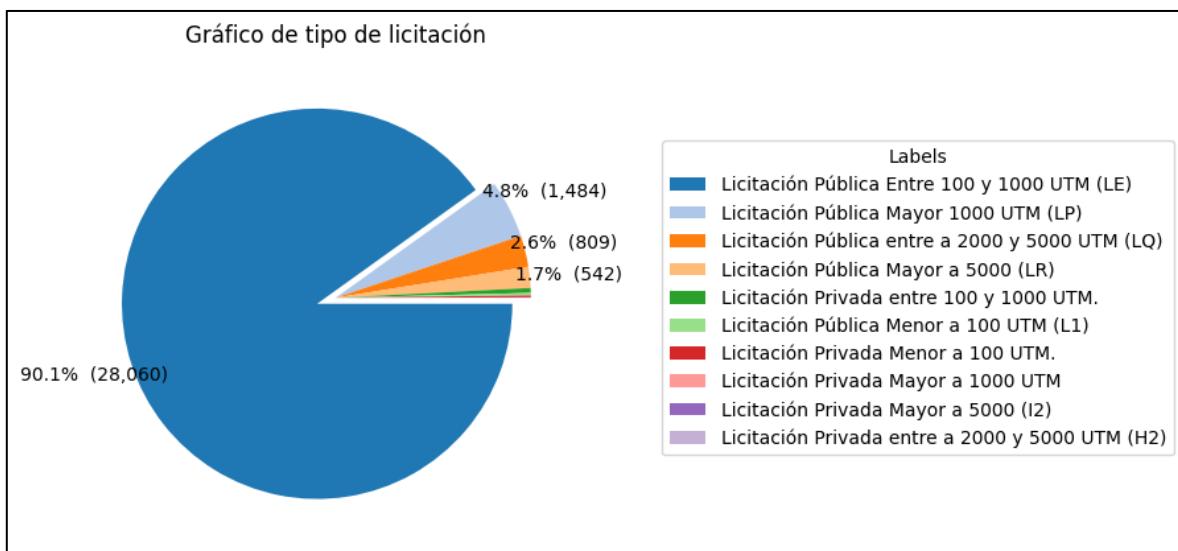


Gráfico 9 Tipo de licitación nº1

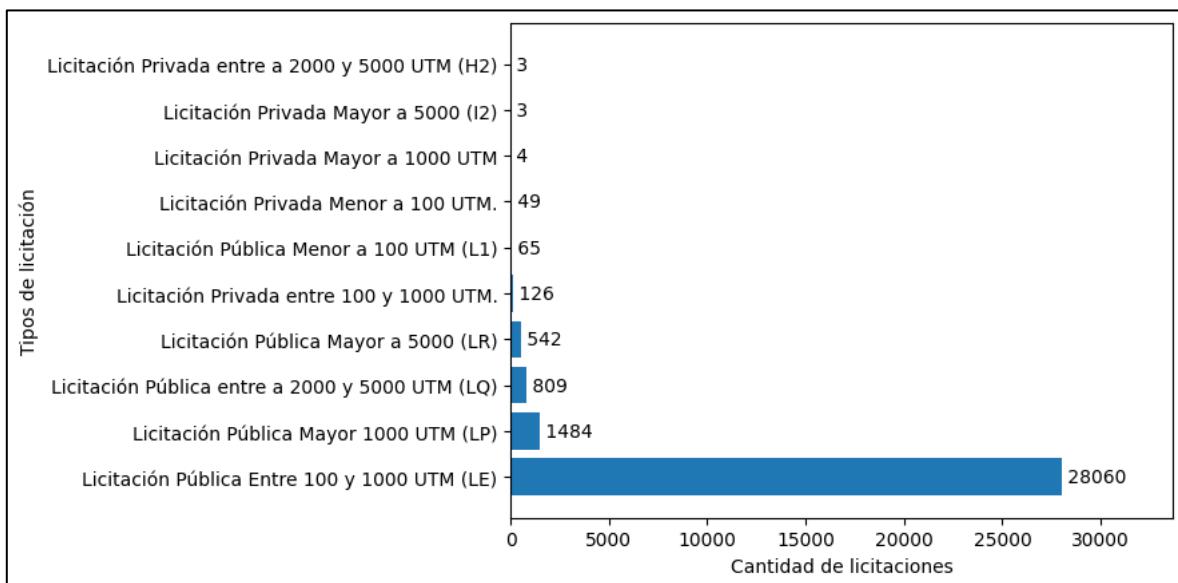


Gráfico 10 Tipo de licitación nº2

- En el gráfico se observa una gran concentración de licitaciones de entre 100–1000 UTM y 90.1%.
- Los proyectos de mayor envergadura, superiores a 2000 UTM, son escasos siendo un 1.7% solo en licitaciones públicas entre montos de 2000 y 5000 UTM.

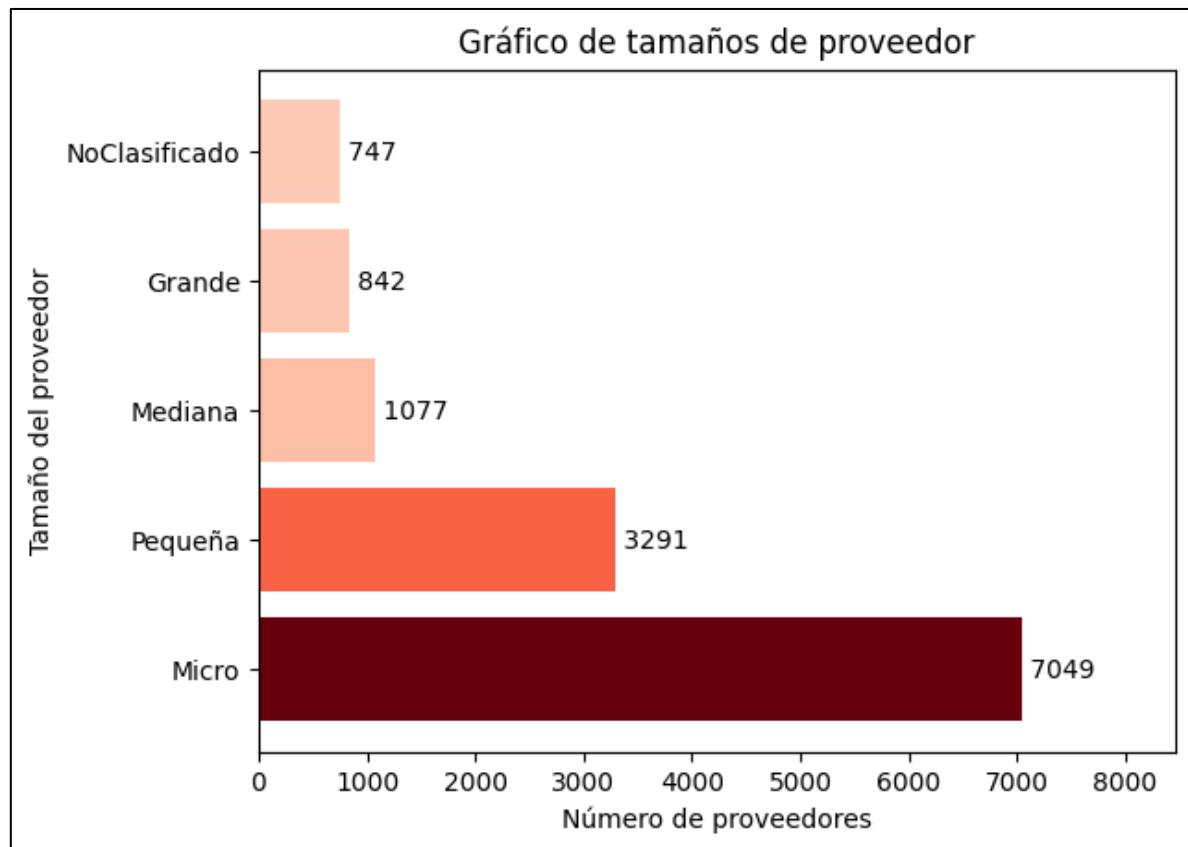


Gráfico 11 Tamaños de proveedor

- Se puede observar cómo en gran medida el mercado de las licitaciones se compone de microempresas con una concentración del 54.1%.

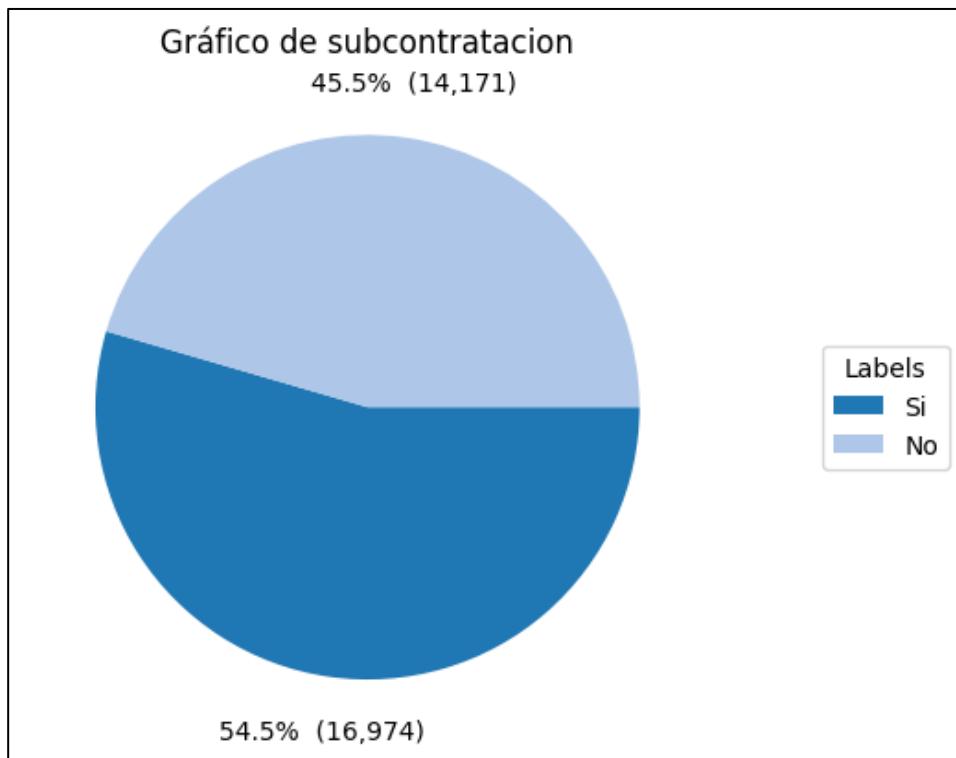


Gráfico 12 Subcontratación

- Se puede observar como la mayoría de las licitaciones admiten subcontratación con un 54.5%.

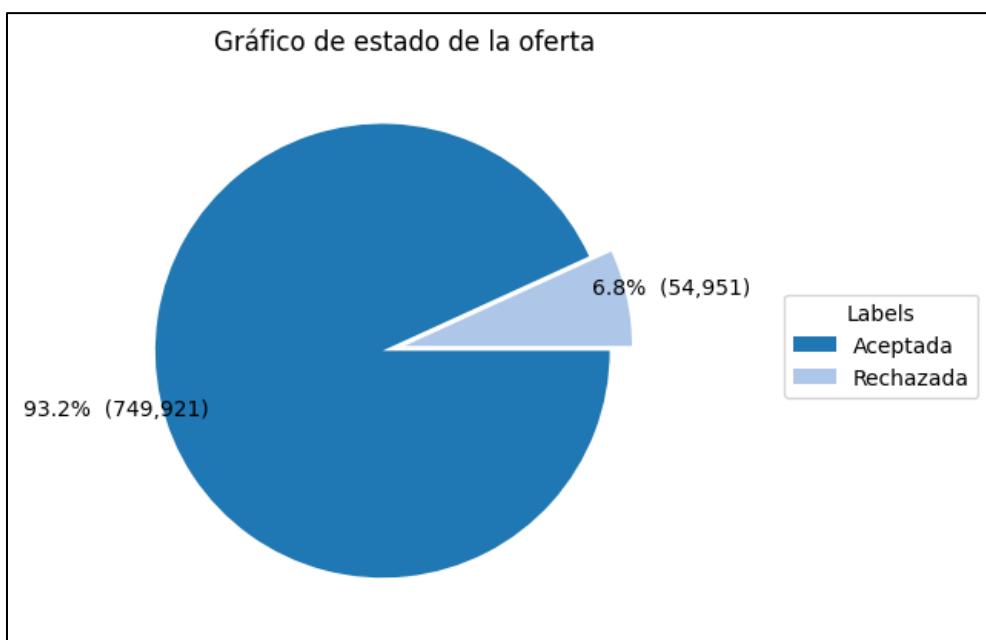


Gráfico 13 Estado de la oferta

- Se puede observar como la gran mayoría de las ofertas en el proceso licitatorio son aceptadas, con un porcentaje de 93.2% de aceptación.

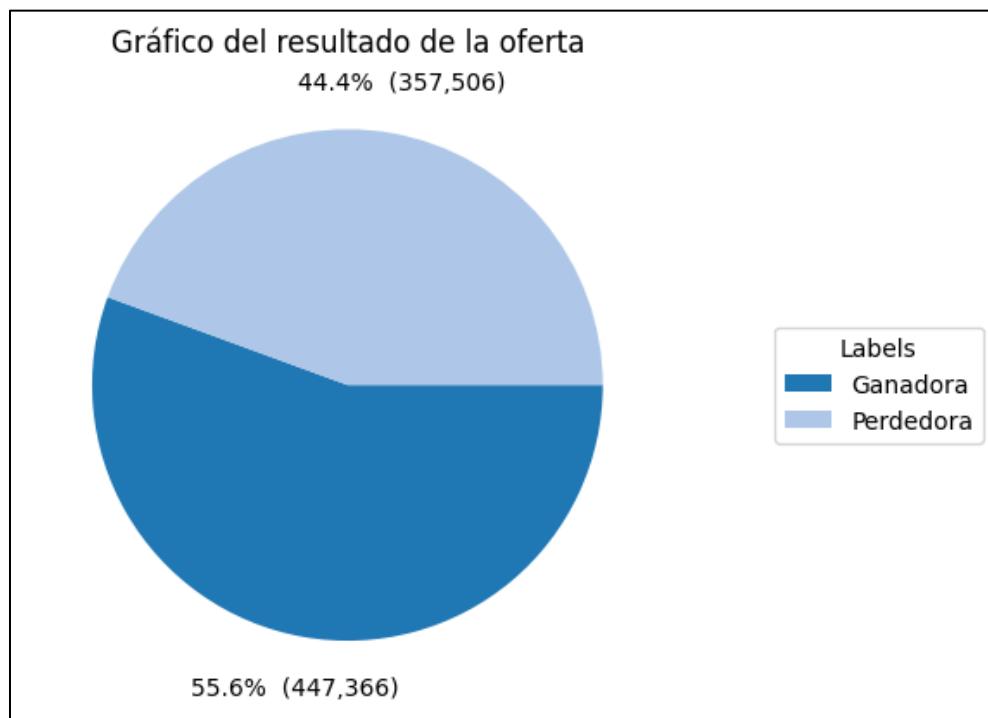


Gráfico 14 Resultado de la oferta

- Se puede observar como del total de ofertas solo el 55.6% son adjudicadas.

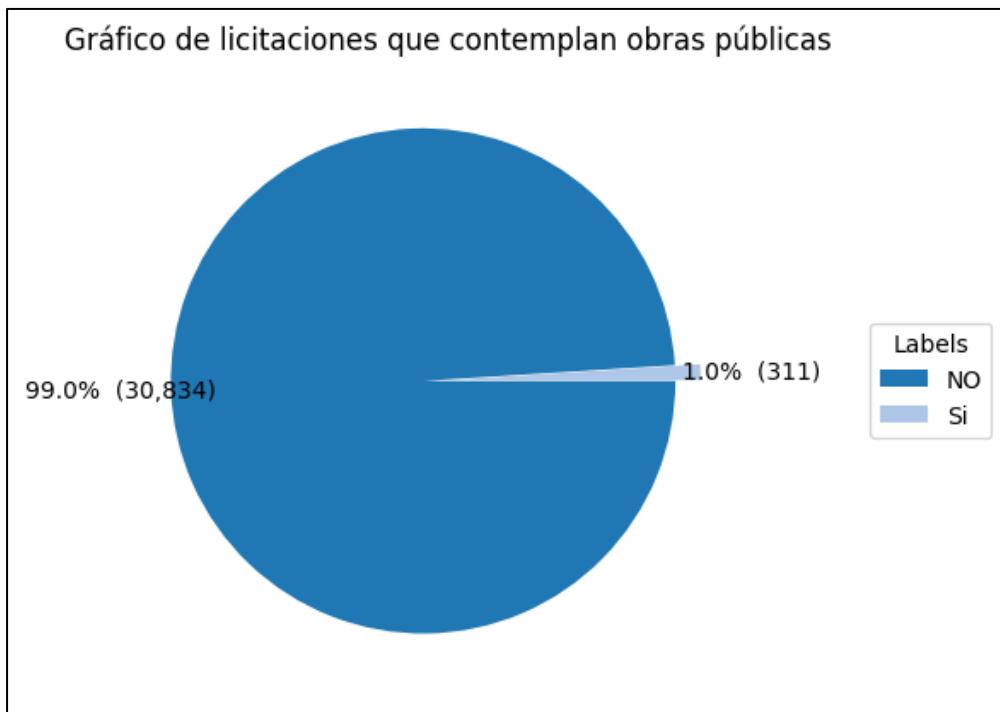


Gráfico 15 Contempla obras públicas

- Se puede observar cómo casi la totalidad de las licitaciones no contemplan obras públicas con un 99.0%.

8. PROCESO ETL

8.1 Descripción del Proceso y Datos.

El proceso ETL (Extract, Transform, Load) es una metodología esencial para la integración de datos, que consiste en extraer información de diversas fuentes, transformarla para hacerla coherente y útil, y finalmente cargarla en un destino específico. En el caso de PyGo SPA, este proceso implica la extracción de datos brutos de conjuntos de datos públicos facilitados por el gobierno de Chile.

Estos datos luego son sometidos a un proceso de transformación mediante técnicas de Inteligencia de Negocios (BI), utilizando la herramienta gratuita KNIME. El objetivo es convertir los datos en información estructurada y relevante para mejorar la toma de decisiones y la estrategia de mercado de PyGo SPA.

Una vez transformados, los datos se encuentran listos para su análisis y aplicación en la toma de decisiones empresariales:

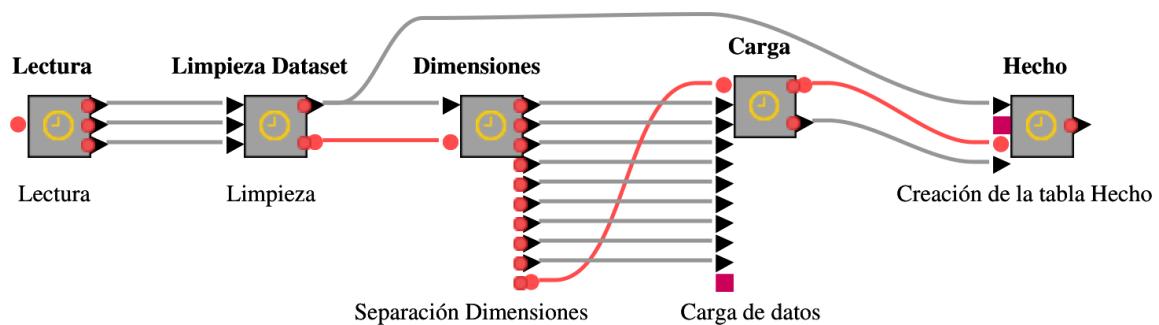


Imagen 2 Proceso ETL general

8.2 Descripción del modelo

Nodo	Descripción
Benchmark Start	Inicia el proceso ETL. Puede estar relacionado con un temporizador o una acción específica.
Lectura	Lee un conjunto de datos desde una fuente externa (por ejemplo, una base de datos o un archivo CSV).
Limpieza del Dataset	Realiza procesos de limpieza y transformación en los datos para prepararlos para su análisis posterior.
Separación de dimensiones	Divide los datos en dimensiones relevantes (por ejemplo, tiempo, producto, ubicación).
DB Connector	Se conecta a una base de datos para cargar o actualizar datos.

Carga	Carga los datos transformados en una tabla o almacén de datos final.
Creación de la tabla Hecho	En un esquema de estrella o copo de nieve, esta tabla contiene los datos de hechos (por ejemplo, ventas totales).
Benchmark End	Finaliza el proceso ETL. Puede estar relacionado con un temporizador o una acción específica.
CSV Writer	Escribe los datos en un archivo CSV como parte del flujo

Tabla 76 Descripción del modelo

8.3 Descripción de Problemas Encontrados

- **Problemas al importar datos de diferentes fuentes:**
Las fuentes variaban desde bases de datos en línea hasta archivos locales y servicios web, cada una con su propio esquema y formato de datos. Esta diversidad obligó a implementar múltiples nodos de lectura y configuraciones específicas para cada tipo de fuente, complicando el diseño y la gestión del flujo de trabajo. Parte de la solución del mismo fue descartar ciertos tipos de datos de ingreso, dígase de la API de ChileCompra que se intentó implementar para corregir los datos erróneos.
- **Problemas con la codificación del texto:**
En la manipulación de archivos CSV dentro de KNIME, se encontraron varios problemas de codificación de texto. Las diferencias en la codificación entre los archivos de entrada causaban errores en la interpretación de caracteres especiales y acentos. Para solucionarlo, fue necesario identificar y estandarizar la codificación de cada archivo a un formato común (en este caso, Latin-1) antes de procesarlos.
- **Valores de montos incorrectos o muy elevados:**
Se detectaron valores anormalmente altos o erróneos en los montos registrados, lo que podía conducir a conclusiones incorrectas. Para resolver este problema, se implementaron nodos de limpieza y algunos de validación.
- **Importación con diferentes formatos:**
Al trabajar con distintos archivos se debieron normalizar ciertos archivos, columnas o incluso tipo de archivo en un conjunto similar. Para el tratado de estos

formatos se utilizó una combinación de nodos para transformar y homogeneizar las estructuras de datos.

- **Columnas descartables por valores perdidos**
- **Atributos en formato de horas innecesario**

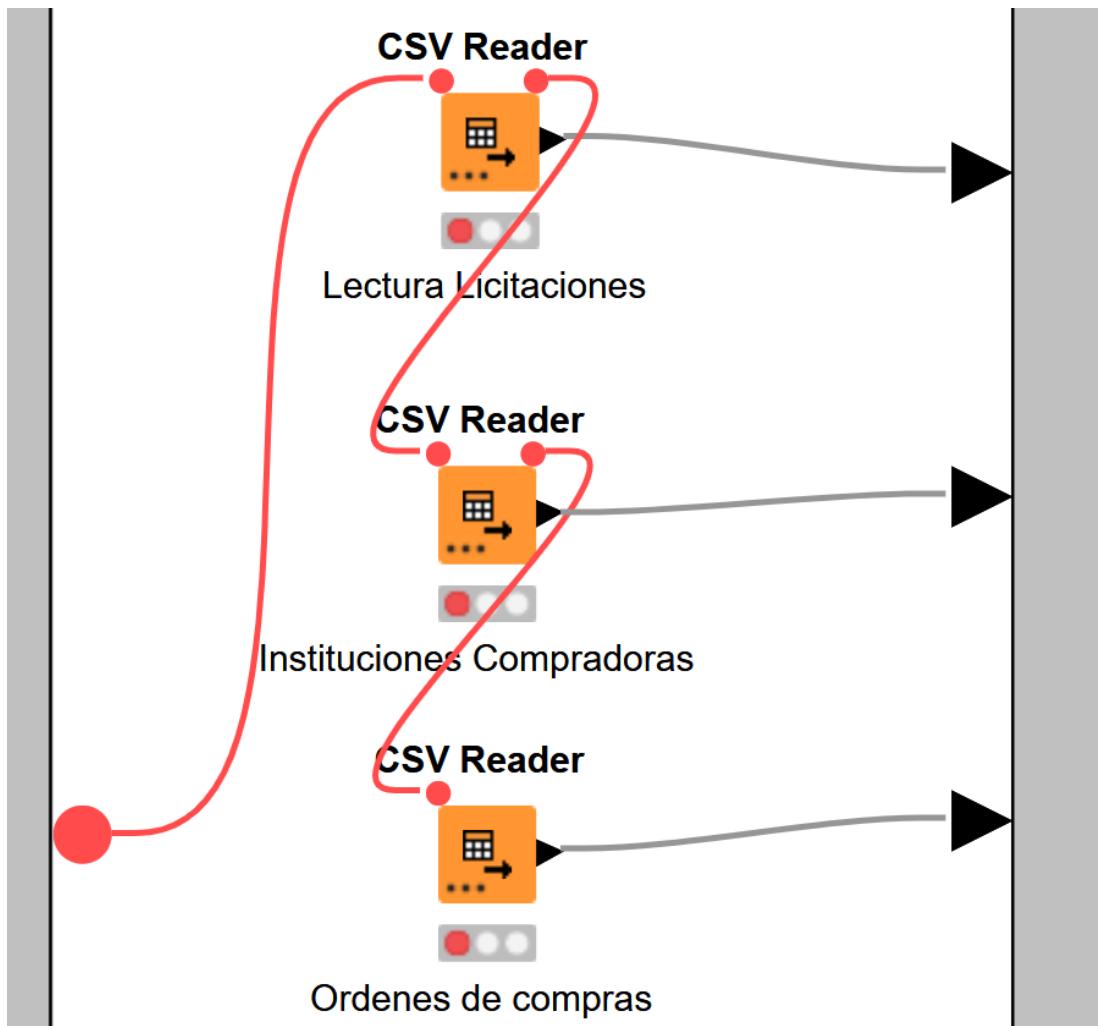
Criterios Adoptados y Lógica de Transformación

(detallado; fórmulas, redondeos, cálculos, aproximaciones, tipo de cambio, etc.)

1. Se eliminaron todas las filas con valores faltantes (NaN).
2. Se filtraron los datos para incluir aquellos con montos de licitación mayores a 100 CLP.
3. Se corrigieron valores anómalos en las variables “MontoTotalOferta” y “MontoNetoOferta”, cuyos mínimos mostraban valores negativos (-1), atribuidos a errores humanos, así como valores de 0 y 1. Para solucionarlo, se ajustó el monto total de la oferta al monto estimado de la licitación en los casos adjudicados.
4. Se ajustó la “CantidadOferta” donde había casos de ofertas adjudicadas con cantidad cero, igualando la cantidad ofertada a la cantidad solicitada en la licitación adjudicada.
5. Se aplicó un filtro a la “CantidadOferta” para eliminar valores de cero y luego se recalcularó el “MontoTotalOferta” dividiéndolo por la “CantidadOferta” ajustada. Los nuevos valores se usaron para actualizar “MontoNetoOferta” en las filas donde “CantidadOferta” era igual a 1.
6. Se filtró la columna “MontoTotalOferta” para eliminar valores negativos, asegurando que los montos fueran mayores a cero.
7. Se reemplazó la columna “MontoTotalOferta” original por la columna corregida, conservando el mismo nombre.
8. Se realizó un filtrado final en “MontoTotalOferta” para excluir valores superiores a 100.000.000 CLP.
9. Se llevó a cabo una limpieza final para eliminar valores atípicos en los montos estimados de las licitaciones.

Descripción de Herramientas Utilizadas.

Carga y Análisis del Proceso ETL (*Por cada Dimensión, Hecho y Métrica*)



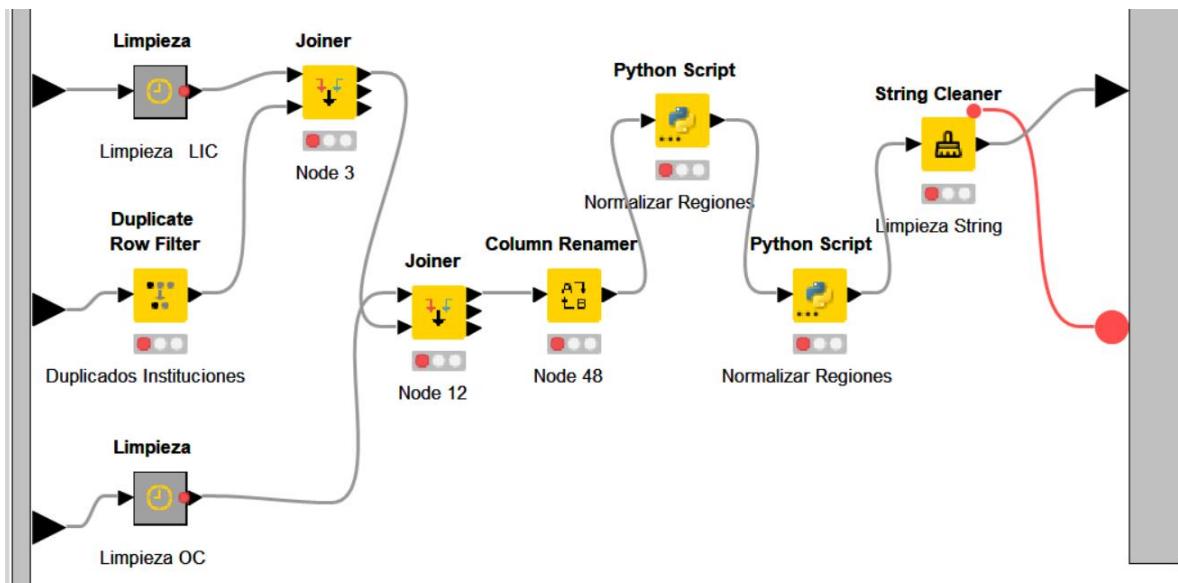
CSV Reader (Lector de archivos CSV): Estos tres nodos representan la lectura de archivos CSV (Comma-Separated Values). Cada uno de ellos lee datos de diferentes fuentes:

Lectura Licitaciones: Este nodo lee datos relacionados con licitaciones o procesos de adquisición.

Instituciones Compradoras: Aquí se leen datos sobre las instituciones que realizan compras.

Ordenes de compras: Este nodo lee datos específicos sobre las órdenes de compra.

Los datos leídos por los nodos "CSV Reader" probablemente se procesan o almacenan en algún lugar posteriormente. Las líneas rojas indican la dirección del flujo de datos.



Limpieza: El nodo “Limpieza” realiza tareas de limpieza y preprocesamiento de datos. Esto puede incluir la eliminación de valores nulos, la normalización de datos o la corrección de errores en los datos de entrada.

Joiner: El nodo “Joiner” se utiliza para combinar datos de diferentes fuentes o tablas. Puede realizar operaciones de unión, como uniones internas o externas, para combinar datos relacionados.

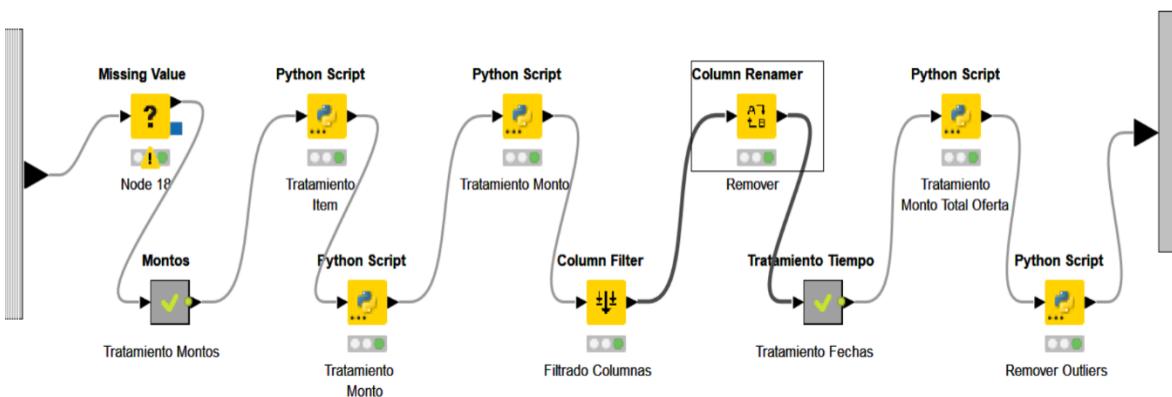
Python Script: El nodo “Python Script” permite ejecutar scripts de Python dentro del flujo de trabajo. Esto es útil para realizar cálculos personalizados o aplicar funciones específicas a los datos.

String Cleaner: El nodo “String Cleaner” se utiliza para limpiar y estandarizar cadenas de texto en los datos. Puede eliminar caracteres no deseados o realizar otras transformaciones en las cadenas.

Column Renamer: Este nodo cambia los nombres de las columnas en el conjunto de datos. Es útil para dar nombres más descriptivos o para estandarizar los nombres de las columnas.

Normalizar Regiones: El nodo “Normalizar Regiones” probablemente normaliza los valores en una columna específica, como regiones geográficas. Esto podría implicar la estandarización de nombres de regiones o la asignación de códigos únicos.

Duplicados Instituciones: El nodo “Duplicados Instituciones” filtra filas duplicadas en el conjunto de datos. Esto asegura que no haya registros repetidos en la tabla.



Identificación de Valores Faltantes (Missing Value): En esta etapa, se identifican los valores faltantes en el conjunto de datos.

Tratamiento de Datos con Scripts de Python: En particular, se realizan tratamientos específicos para los “Items” y los “Montos”.

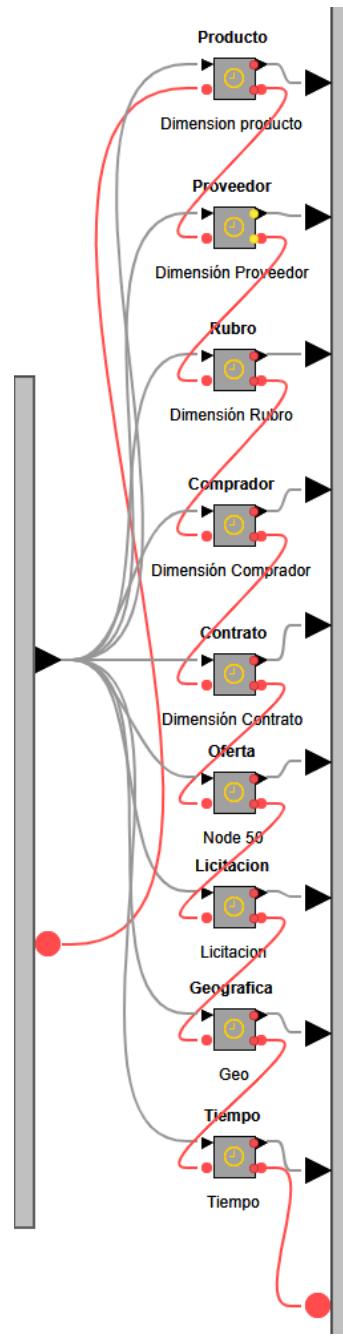
Filtrado de Columnas (Column Filter): Se filtran las columnas relevantes para el análisis posterior. esto con el fin de implicar seleccionar solo las columnas necesarias o eliminar columnas innecesarias.

Renombrar Columnas (Column Renamer): Se cambian los nombres de las columnas para que sean más descriptivos o estandarizados. esto facilita la comprensión y el análisis de los datos.

Normalización de Fechas y Tiempos: Se aplican transformaciones a las columnas relacionadas con fechas y horas. Esto puede incluir la conversión de formatos, la extracción de componentes específicos

Tratamiento del Monto Total de la Oferta: Se realiza un procesamiento específico en la columna que contiene el monto total ofertado. Esto podría incluir la corrección de errores, la normalización o la detección de valores atípicos.

Remoción de Valores Atípicos (Outliers): Finalmente, se eliminan los valores atípicos del conjunto de datos. esto ayuda a garantizar que los datos analizados sean representativos y no estén sesgados por valores extremos.



Producto (Dimensión Producto) Esta dimensión se refiere a los diferentes productos o artículos que se están analizando. Incluye diferente información sobre los productos su nombre , códigos de productos o categorías de productos.

Proveedor (Dimensión Proveedor):

La dimensión del proveedor se relaciona con las entidades o empresas que suministran los productos o servicios. Este incluye información sobre los proveedores, como nombres, identificadores o ubicaciones.

Rubro (Dimensión Rubro):

El rubro se refiere a las categorías o sectores a los que pertenecen los productos o servicios.

Por ejemplo, si estamos analizando datos de compras, el rubro podría ser “Tecnología”, “Alimentación”, “Construcción”, etc.

Comprador (Dimensión Comprador):

Esta dimensión representa a las personas o entidades que realizan las compras.

Incluye detalles sobre los compradores, como nombres, su rut único y region.

Contrato (Dimensión Contrato):

La dimensión del contrato se relaciona con los acuerdos formales entre compradores y proveedores, existen diferentes tipos de contratos y pagos para estos formatos

Oferta (Dimensión Oferta):

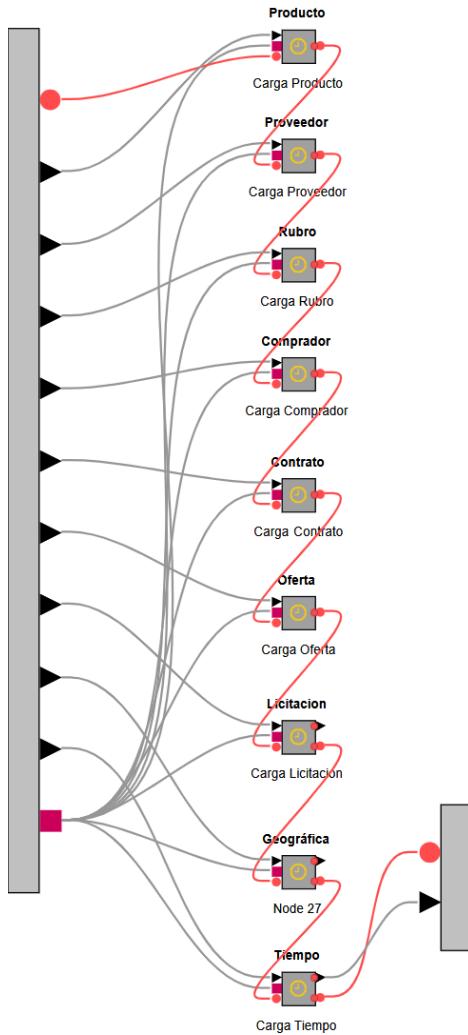
Esta dimensión se refiere a las propuestas o licitaciones presentadas por los proveedores, esto además incluyen el estado del oferta, como puede ser adjudicata o si fue aceptada o no

Geográfica (Dimensión Geográfica):

La dimensión geográfica se relaciona con la ubicación geográfica de las transacciones. Estas incluyen las regiones dentro del territorio chileno, tanto para proveedores como compradores

Tiempo (Dimensión Tiempo):

Esta dimensión representa la marca de tiempo asociada con cada transacción. esta incluye detalles sobre fechas, meses o años



Selección de Tablas de la Base de Datos (DB Table Selector):

Los nodos etiquetados como “DB Table Selector” representan la selección de tablas específicas de la base de datos.

Cada tabla corresponde a una dimensión específica, como “Producto”, “Proveedor”, “Rubro”, “Comprador”, “Contrato”, “Oferta” y “Licitación”.

Lectura de Datos desde la Base de Datos (DB Reader):

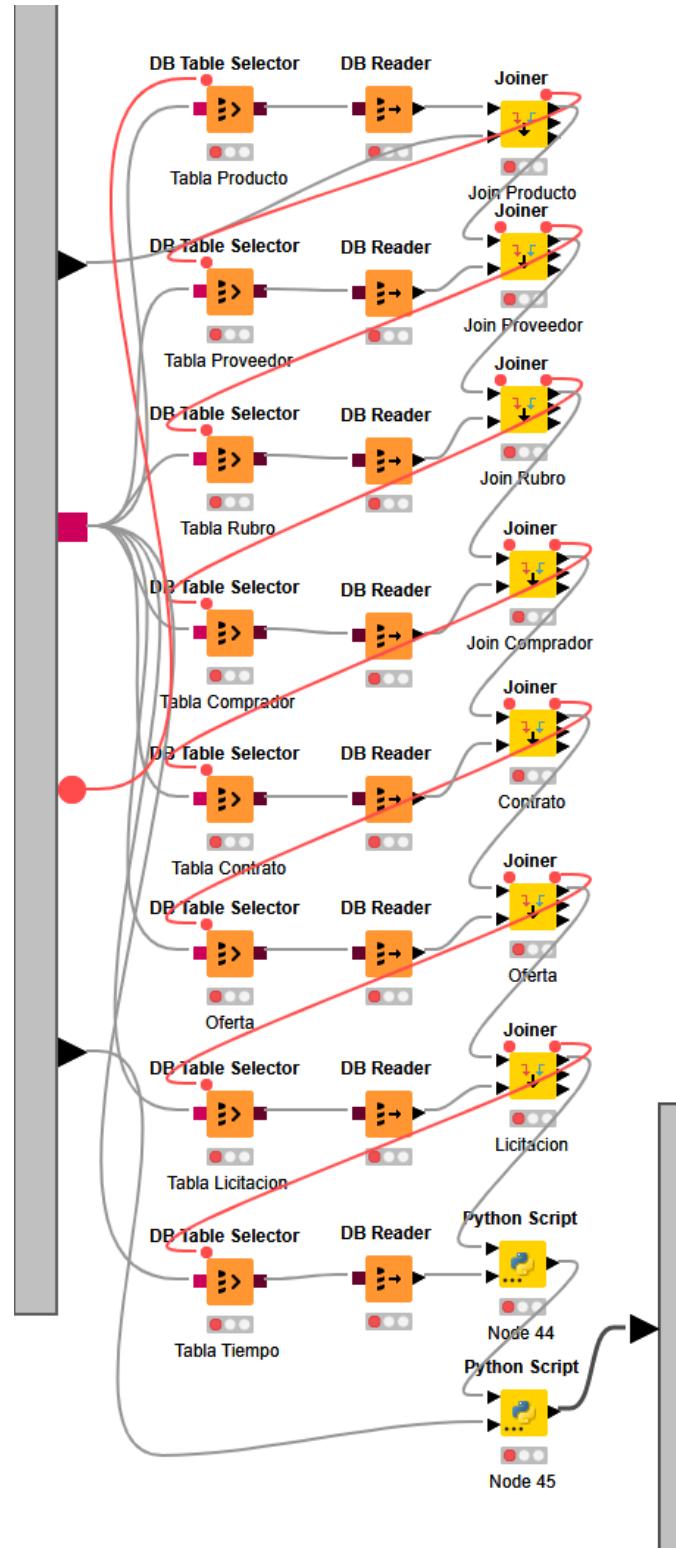
Los nodos “DB Reader” están conectados a los “DB Table Selector”. Estos nodos leen los datos de las tablas seleccionadas desde la base de datos.

Unión de Datos (Joiner):

Los nodos “Joiner” son donde se realiza la unión entre los datos limpios y las dimensiones. Cada “Joiner” combina los datos de una tabla específica (por ejemplo, “Producto”) con los datos limpios previamente procesados.

Variables de Flujo (Flow Variables):

Las variables de flujo son dinámicas y flexibles en KNIME. Se utilizan para almacenar valores como números, cadenas de texto o incluso tablas completas. En este flujo, las variables de flujo se utilizan para controlar el flujo de datos y asegurar que las dimensiones se carguen en el **orden correcto** en la base de datos.



Selección de Tablas de la Base de Datos (DB Table Selector): Los nodos etiquetados como “DB Table Selector” representan la selección de tablas específicas de la base de datos.

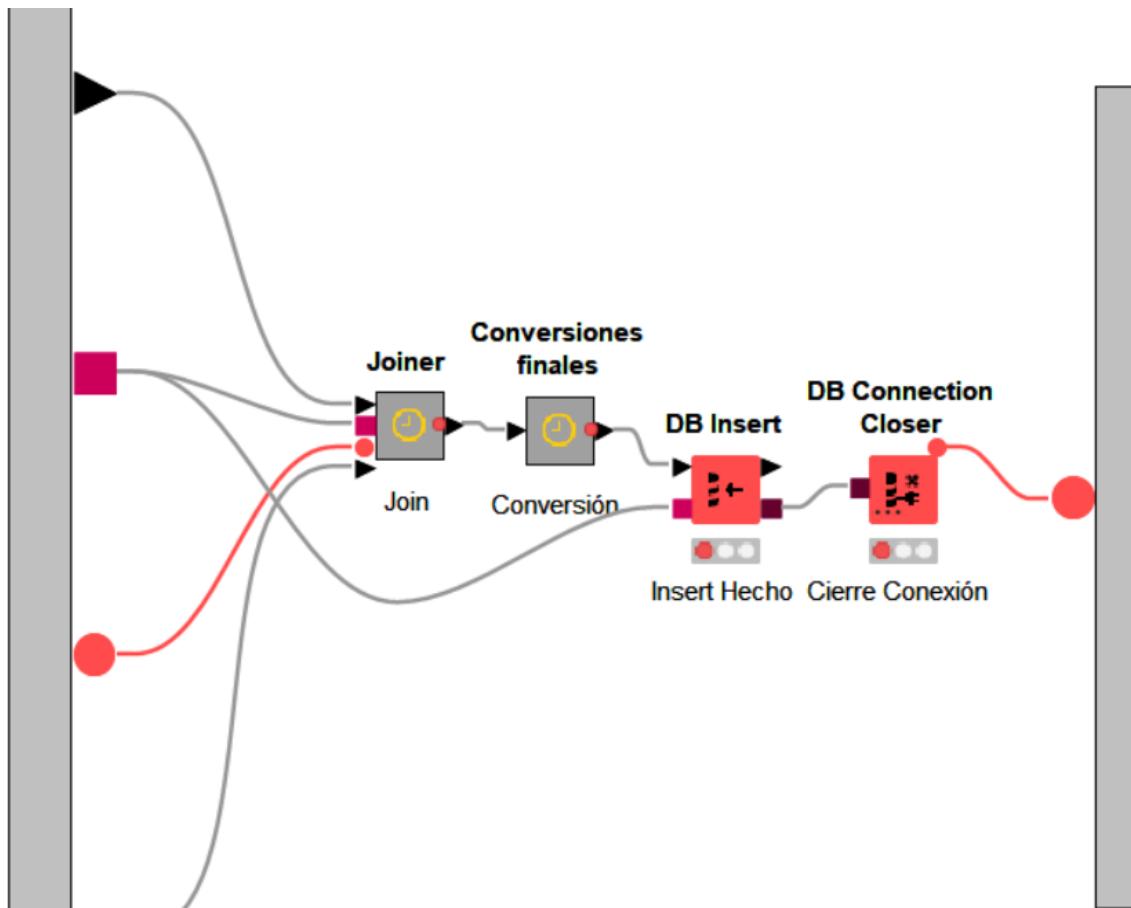
Cada tabla corresponde a una dimensión específica, como “Producto”, “Proveedor”, “Rubro”, “Comprador”, “Contrato”, “Oferta” y “Licitación”.

Lectura de Datos desde la Base de Datos (DB Reader): Los nodos “DB Reader” están conectados a los “DB Table Selector”. Estos nodos leen los datos de las tablas seleccionadas desde la base de datos.

Unión de Datos (Joiner): Los nodos “Joiner” son donde se realiza la unión entre los datos limpios y las dimensiones. Cada “Joiner” combina los datos de una tabla específica (por ejemplo, “Producto”) con los datos limpios previamente procesados.

Tablas de Dimensiones: Cada “Joiner” está conectado a una tabla de dimensión específica (por ejemplo, “Tabla Producto”, “Tabla Proveedor”, etc.). Estas tablas de dimensiones contienen información relevante sobre los elementos específicos de cada dimensión.

Al final del flujo, hay dos nodos de Python estos nodos se usan para realizar el join entre las diferentes regiones, entre región de proveedor y región de comprado



Joiner (Unión): Este nodo se encarga de unir datos. Puede combinar tablas o conjuntos de datos basándose en columnas comunes. Por ejemplo, si tienes una tabla de dimensiones y

una tabla de hechos, el nodo “Joiner” te permite unir estas dos tablas para crear una vista combinada.

Conversiones finales: Este nodo probablemente realice transformaciones o conversiones en los datos antes de cargarlos en la tabla de hechos. Puede incluir cambios de formato, cálculos o limpieza de datos.

DB Insert (Inserción en la base de datos): Una vez que los datos están listos, este nodo los inserta en la tabla de hechos en una base de datos. Esencialmente, es el paso final antes de que los datos se almacenen permanentemente.

DB Connection Closer (Cierre de conexión a la base de datos): Después de cargar los datos, este nodo cierra la conexión con la base de datos. Es importante cerrar la conexión correctamente para liberar recursos y evitar problemas de rendimiento.

Modelo ETL (Indicar Herramienta Utilizada)

La elección de KNIME como herramienta ETL se debe a su interfaz gráfica intuitiva que permite crear flujos de trabajo sin necesidad de programación, su amplia variedad de funciones predefinidas para tareas de ETL, su flexibilidad para integrar datos de múltiples fuentes y su escalabilidad para manejar conjuntos de datos de diferentes tamaños. Además, la comunidad activa de usuarios y desarrolladores que contribuyen con extensiones y nodos personalizados también influye en su elección.

Problemas en la Carga de Datos

La implementación de KNIME para el proceso ETL no presentó mayores complicaciones en la etapa de carga de datos, especialmente en lo referente a problemas en la carga de datos. Esto sugiere una integración efectiva y una capacidad de manejo de datos sólida por parte de la herramienta.

Rendimiento de Herramientas y Tiempos de Carga

Se adjunta como benchmark.xlsx

9. Explotación de Datos CON OLAP

9.1 Descripción de la Etapa

La etapa de explotación de datos con OLAP fue realizada utilizando la herramienta PowerBI, esta representa una fase crucial en el procesamiento y presentación de la información extraída de las licitaciones Gubernamentales Chilenas.

Este proceso consta de varias etapas fundamentales que incluyen: Extracción, limpieza, análisis y visualización de datos. Cada una de estas etapas se realiza con el objetivo de obtener indicadores (KPI) útiles para revelar patrones y tendencias significativas, al mismo tiempo que proporcionar una visualización clara y comprensible para la correcta toma de decisiones.

9.2 Descripción de Herramientas Utilizadas.

Nombre de la aplicación	PowerBI
Tipo de herramienta	Herramienta OLAP
Justificación de uso	Capacidad de visualizar datos de diversas maneras como gráficos o tablas, utilizando una conexión a base de datos multidimensional, además de realizar cálculos entre datos específicamente soportados para relaciones en la propia base de datos.
Versión	2.128.1380.0 64-bit (abril de 2024)

Tabla 77 Herramienta OLAP

9.3 Indicadores obtenidos y Validación del Indicador.

Dashboards

Para la elaboración de los indicadores se ha optado por la agrupación de los mismos según requerimiento, se opta por mostrar una vista previa general antes de la especificación individual.

Requerimiento 2: Analizar el comportamiento de los compradores a lo largo del tiempo

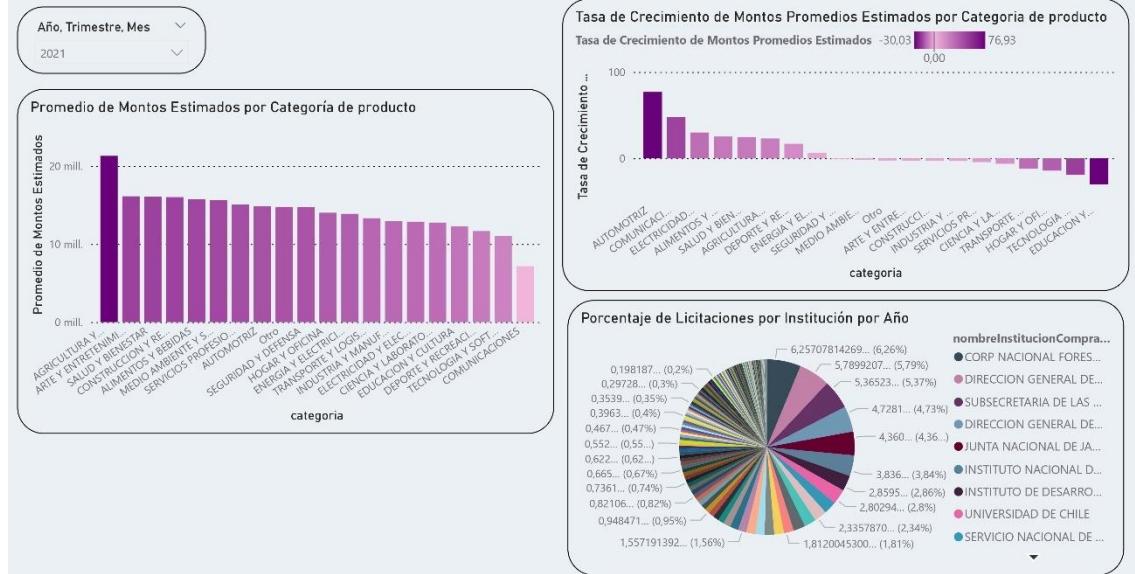


Imagen 3 Dashboard requerimiento 2

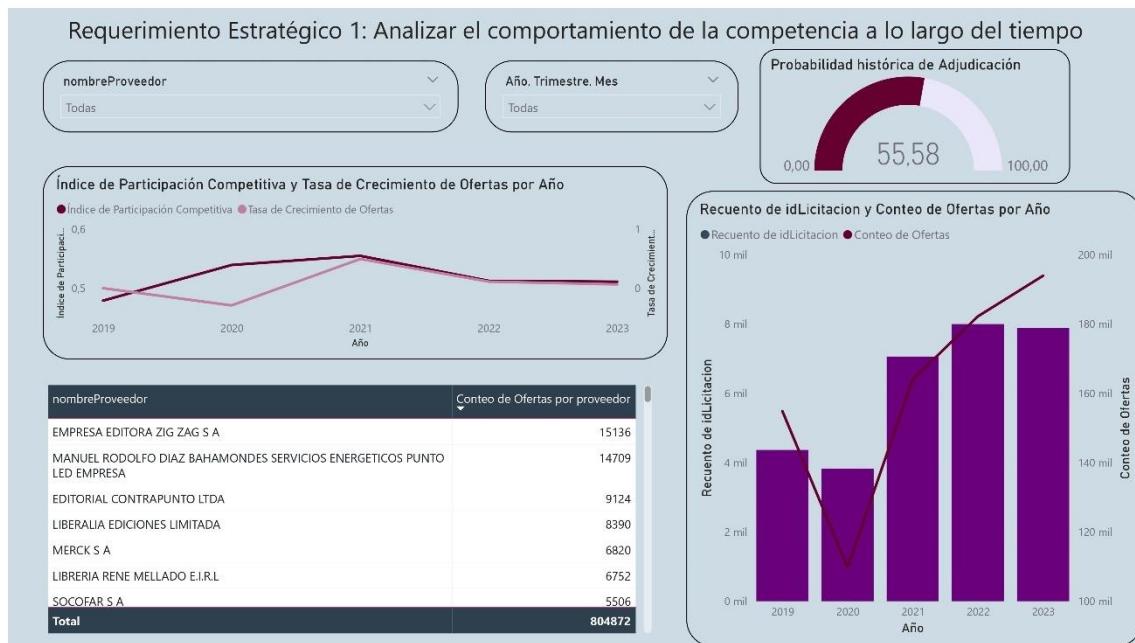


Ilustración 12 - Dashboard requerimiento 1

REQ-3: Analizar la demanda de productos en el mercado a lo largo del tiempo

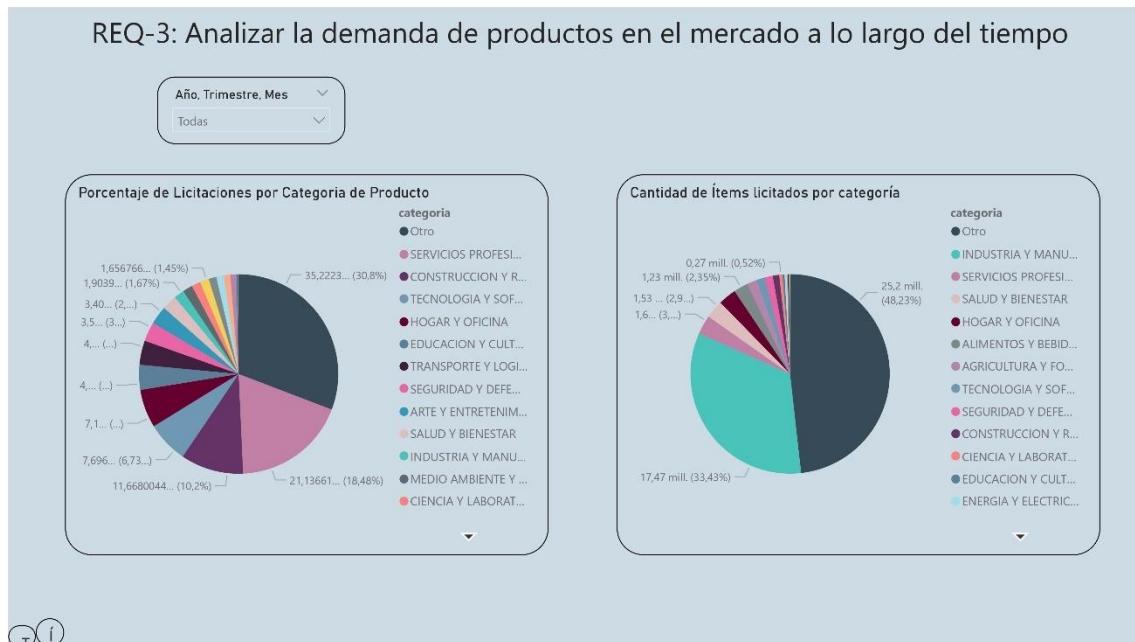


Ilustración 13- Dashboard requerimiento 3

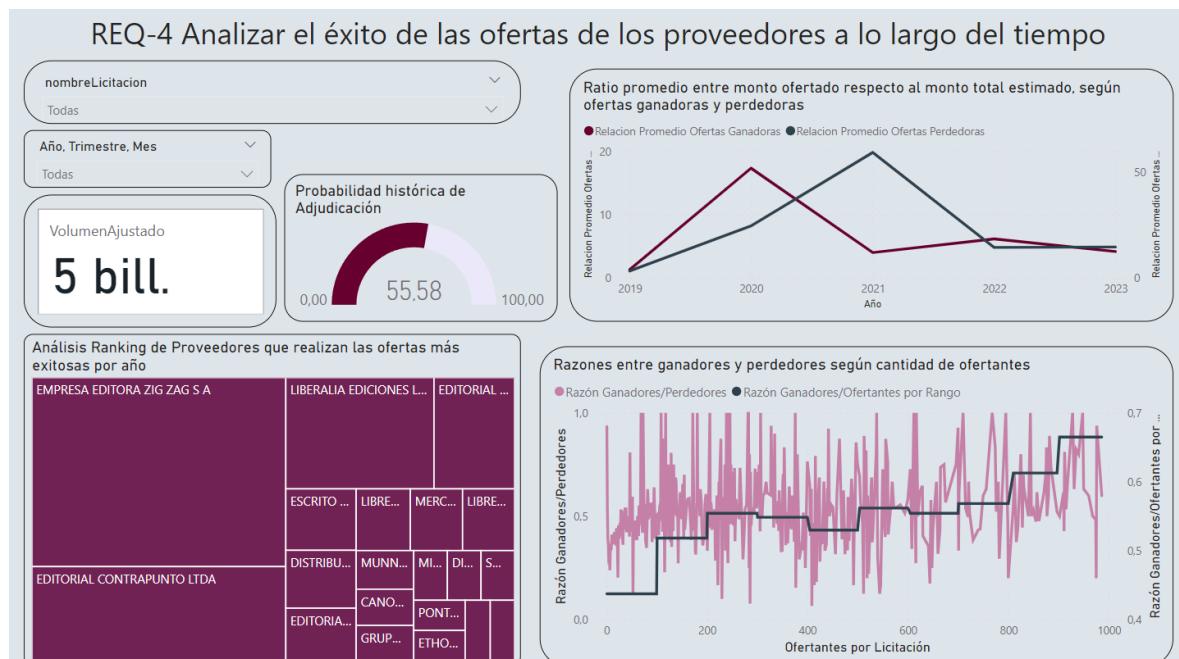


Ilustración 14- Dashboard requerimiento 4

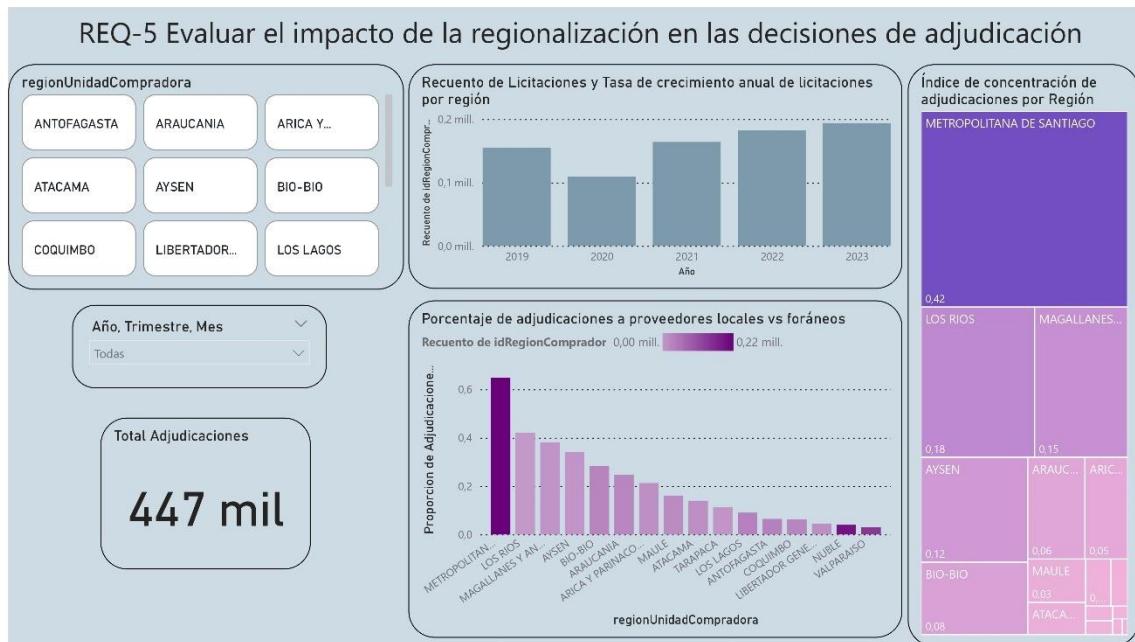


Ilustración 15- Dashboard requerimiento 5

Cada uno de estos dashboard posee un mínimo de 3 KPI visibles, como se puede observar, algunos de estos kpi tienen un indicador de color propio para resaltar la diferencia entre valores y otros que no requieren de esta comparativa mantienen un color plano, esto se verá a más detalle a continuación en la descripción individual de cada uno.

KPI 1 Porcentaje de licitaciones por categoría de producto en los períodos

2019 – 2023 para cada región:

Interfaz OLAP.



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Porcentaje de Licitaciones por Categoría =
VAR TotalLicitaciones =
    CALCULATE(
        DISTINCTCOUNT('public_hecho_licitacion'[idLicitacion]),
        REMOVEFILTERS('public_hecho_licitacion')
    )
VAR TotalLicitacionesPorCategoria =
    CALCULATE(
        DISTINCTCOUNT('public_hecho_licitacion'[idLicitacion]),
        ALLEXCEPT('public_hecho_licitacion', 'public_producto'[categoria], 'public_hecho_licitacion'[fechaPublicacion].[Año], 'public_hecho_licitacion'[fechaPublicacion].[Mes], 'public_hecho_licitacion'[fechaPublicacion].[Trimestre])
    )
RETURN
    DIVIDE(
        TotalLicitacionesPorCategoria,
        TotalLicitaciones,
        0
    ) * 100
WITH total_rows_cte AS (
    SELECT
        COUNT(DISTINCT "idLicitacion") AS total_rows,
        p."categoria" AS Categoria
    FROM

```

```

    public.hecho_licitacion hl
  JOIN
    public.producto p ON hl."idProducto" = p."idProducto"
  WHERE
    EXTRACT(YEAR FROM hl."fechaPublicacion") = 2022
  GROUP BY
    p."categoria"
)
SELECT
  total_rows AS total_filas,
  total_rows_cte.Categoría AS nombre_categoria,
  (total_rows_cte.total_rows * 100.0) / SUM(total_rows_cte.total_rows)
OVER () AS porcentaje_total
FROM
  total_rows_cte
ORDER BY
  total_rows_cte.total_rows DESC;

```

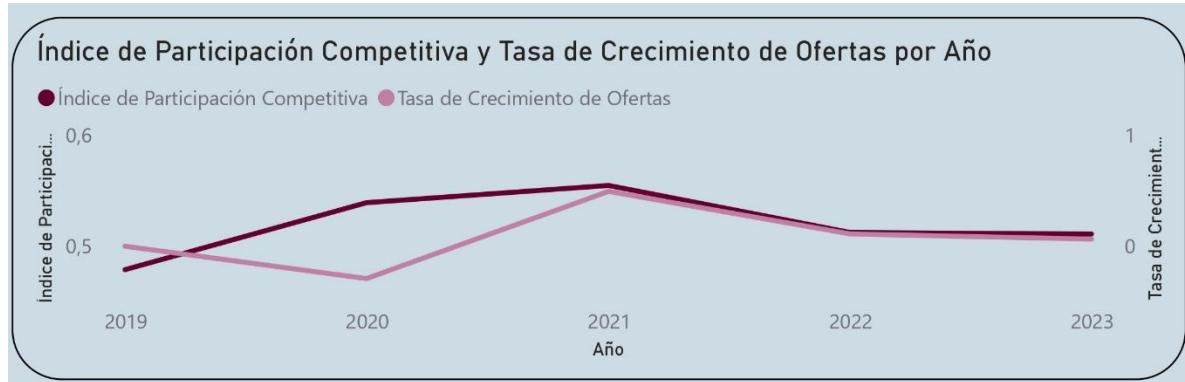
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

2022, SERVICIOS PROFESIONALES Y DE NEGOCIOS = 5.27 (17.67%)

KPI 2 Tasa de crecimiento de ofertas por proveedor por año:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

Tasa de Crecimiento de Ofertas =

VAR AnoActual = MAX('public hecho_licitacion'[fechaPublicacion].[Año])

VAR OfertasAnioActual =

CALCULATE(

```

        COUNTROWS('public hecho_licitacion'),
        FILTER(
            ALL('public hecho_licitacion'[fechaPublicacion].[Año]),
            'public hecho_licitacion'[fechaPublicacion].[Año] = AnoActual
        )
    )

VAR AnoAnterior = AnoActual - 1

VAR OfertasAnioAnterior =
CALCULATE(
    COUNTROWS('public hecho_licitacion'),
    FILTER(
        ALL('public hecho_licitacion'[fechaPublicacion].[Año]),
        'public hecho_licitacion'[fechaPublicacion].[Año] = AnoAnterior
    )
)

RETURN
DIVIDE(OfertasAnioActual - OfertasAnioAnterior, OfertasAnioAnterior, 0)

```

```

WITH OfertasAnioActual AS (
    SELECT
        COUNT(*)::FLOAT AS ofertas_anio_actual -- Convertir a flotante
    FROM
        public."hecho_licitacion"
    WHERE
        EXTRACT(YEAR FROM "fechaPublicacion") = 2021
), OfertasAnioAnterior AS (
    SELECT
        COUNT(*)::FLOAT AS ofertas_anio_anterior -- Convertir a flotante
    FROM
        public."hecho_licitacion"
    WHERE
        EXTRACT(YEAR FROM "fechaPublicacion") = 2020 -- Año anterior a
2021
)

SELECT
    ofertas_anio_actual AS ofertas_2021,
    ofertas_anio_anterior AS ofertas_2020,
    CASE
        WHEN ofertas_anio_anterior = 0 THEN 1.0 -- Establecer un valor
predeterminado si el divisor es 0 (como flotante)

```

```

        ELSE ((ofertas_anio_actual - ofertas_anio_anterior) /
NULLIF(ofertas_anio_anterior, 0)) * 100
    END AS TasaCrecimientoOfertas
FROM
    OfertasAnioActual
CROSS JOIN
    OfertasAnioAnterior;

```

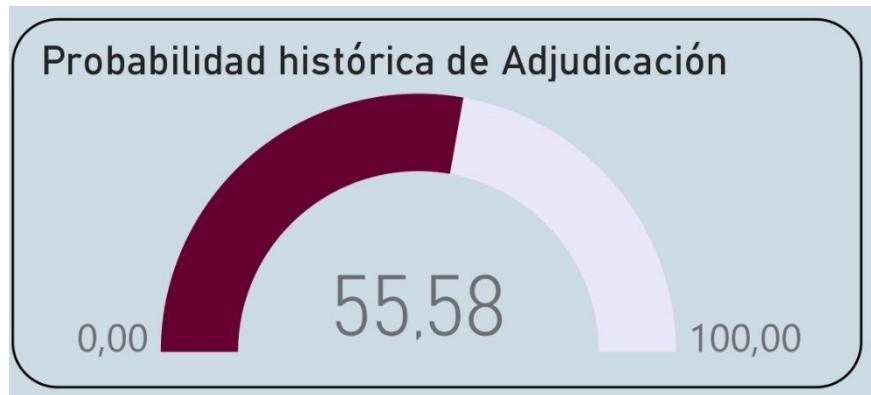
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

tasa crecimiento ofertas 2021 (2020 -> 2021) = 49.40066309614894

KPI 3 Probabilidad de que un proveedor dentro del ranking de la licitación se adjudique una licitación de cierto comprador a lo largo del tiempo:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Probabilidad histórica de Adjudicación1 =
(DIVIDE(
    CALCULATE(
        COUNTROWS('public hecho_licitacion'),
        ALLEXCEPT('public hecho_licitacion', 'public proveedor', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre], 'public comprador'),
        'public oferta'[resultadoOferta] = "GANADORA"
    ),
    CALCULATE(
        COUNTROWS('public hecho_licitacion'),

```

```

    ALLEXCEPT('public.hecho_licitacion', 'public.proveedor', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre])
),
0
)
* 100)

```

```

WITH total_rows_cte AS (
    -- Calcula el número total de filas
    SELECT COUNT(*) AS total_rows
    FROM public.hecho_licitacion hl
),
ganadora_rows_cte AS (
    -- Calcula el número de filas donde el resultado de la oferta es
    "GANADORA"
    SELECT COUNT(*) AS ganadora_rows
    FROM public.hecho_licitacion hl
    JOIN public.oferta o ON hl."idOferta" = o."idOferta"
    WHERE o."resultadoOferta" = 'GANADORA'
)
SELECT
CASE
    WHEN total_rows > 0 THEN
        (ganadora_rows * 100.0) / total_rows
    ELSE
        0
END AS Probabilidad_historica_de_Adjudicacion
FROM
    ganadora_rows_cte,
    total_rows_cte;

```

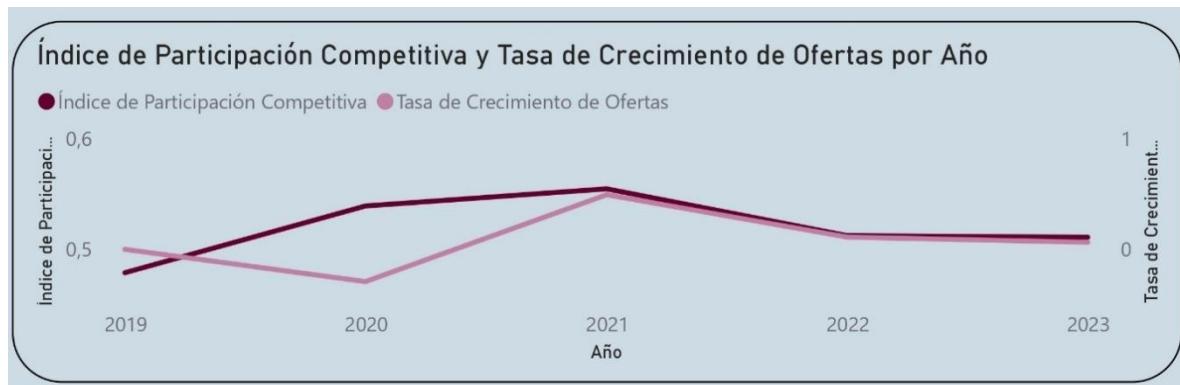
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL, en este caso uno general y globalizado:

```
PROBABILIDAD_HISTORICA = 55.5822540726972736
```

KPI 4 Índice de participación competitiva a lo largo del tiempo:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Índice de Participación Competitiva =
    DIVIDE([Conteo de Ofertas Adjudicadas] , [Conteo de Ofertas], 0) *
    DIVIDE([Conteo de Ofertas Aceptadas], [Conteo de Ofertas], 0) *
    DIVIDE([Conteo de Ofertas], COUNTROWS('public hecho_licitacion'), 0)
-----
Conteo de Ofertas Adjudicadas =
CALCULATE(
    COUNT('public hecho_licitacion'[id]),
    'public oferta'[resultadoOferta] = "GANADORA",
    ALLEXCEPT('public hecho_licitacion', 'public proveedor', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre])
)
-----
Conteo de Ofertas Aceptadas =
CALCULATE(
    COUNT('public hecho_licitacion'[id]),
    'public oferta'[estadoOferta] = "ACEPTADA",
    ALLEXCEPT('public hecho_licitacion', 'public proveedor', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre])
)
-----
Conteo de Ofertas =
CALCULATE(
    COUNTROWS('public hecho_licitacion'),
    ALLEXCEPT('public hecho_licitacion', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public hecho_licitacion'[idProveedor]
))
-----
```

WITH Ofertas_Adjudicadas AS (

```

SELECT
    COUNT(hl.id) AS Conteo
FROM
    public.hecho_licitacion hl
JOIN
    public.oferta o ON hl."idOferta" = o."idOferta"
JOIN
    public.proveedor p ON hl."idProveedor" = p."idProveedor"
WHERE
    o."resultadoOferta" = 'GANADORA'
    AND EXTRACT(YEAR FROM hl."fechaPublicacion") = 2021
),
Ofertas_Aceptadas AS (
    SELECT
        COUNT(hl.id) AS Conteo
    FROM
        public.hecho_licitacion hl
    JOIN
        public.oferta o ON hl."idOferta" = o."idOferta"
    JOIN
        public.proveedor p ON hl."idProveedor" = p."idProveedor"
    WHERE
        o."estadoOferta" = 'ACEPTADA'
        AND EXTRACT(YEAR FROM hl."fechaPublicacion") = 2021
),
Total_Ofertas AS (
    SELECT
        COUNT(*) AS Conteo
    FROM
        public.hecho_licitacion hl
    WHERE
        EXISTS (
            SELECT 1
            FROM public.hecho_licitacion hl_inner
            WHERE EXTRACT(YEAR FROM hl_inner."fechaPublicacion") = 2021
            AND hl.id = hl_inner.id
        )
)
SELECT
    COALESCE((SELECT Conteo FROM Ofertas_Adjudicadas), 0) AS
Conteo_Adjudicadas,
    COALESCE((SELECT Conteo FROM Ofertas_Aceptadas), 0) AS
Conteo_Aceptadas,
    COALESCE((SELECT Conteo FROM Total_Ofertas), 0) AS Conteo_Total,
    CASE

```

```

        WHEN (SELECT Conteo FROM Total_Ofertas) > 0 THEN
            (COALESCE((SELECT Conteo FROM Ofertas_Adjudicadas), 0) * 1.0 /
COALESCE((SELECT Conteo FROM Total_Ofertas), 0)) *
            (COALESCE((SELECT Conteo FROM Ofertas_Aceptadas), 0) * 1.0 /
COALESCE((SELECT Conteo FROM Total_Ofertas), 0)) *
            (COALESCE((SELECT Conteo FROM Total_Ofertas), 0) * 1.0 /
COALESCE((SELECT Conteo FROM Total_Ofertas), 0))
        ELSE
            0
    END AS Indice_Participacion_Competitiva;

```

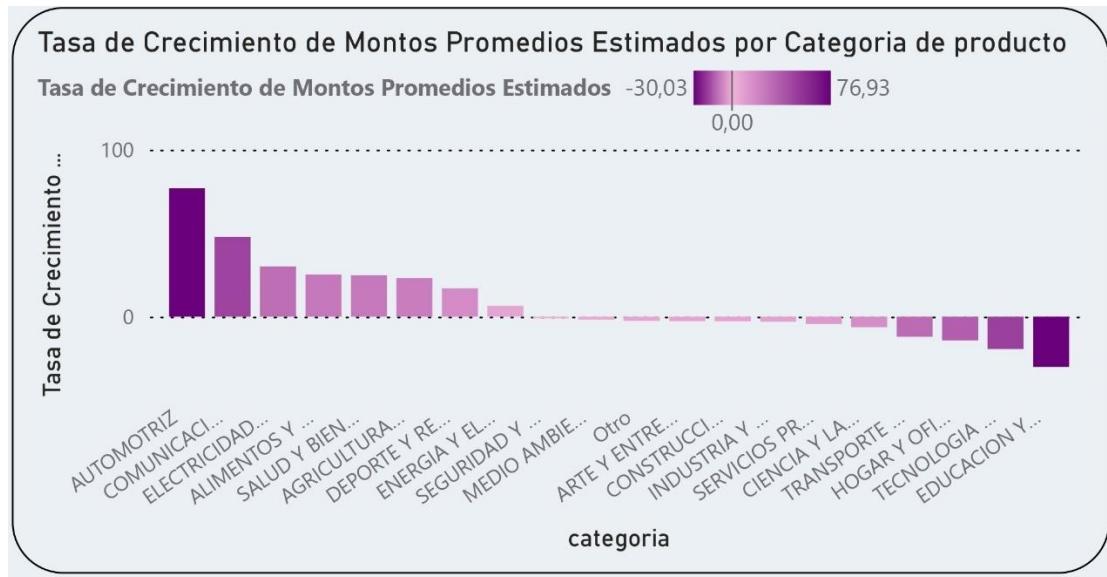
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

2021, IPC = 0.5547024

KPI 5 Tasa de crecimiento del monto promedio estimado por categoría de producto por institución por tipo de licitación a lo largo del tiempo:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Tasa de Crecimiento de Montos Promedios Estimados =
VAR TotalEstimado =
CALCULATE(
[Promedio de Montos Estimados],
FILTER(

```

```

        ALLEXCEPT('public hecho_licitacion', 'public producto'[categoria]),
        'public hecho_licitacion'[fechaPublicacion].[Año] = MAX('public
hecho_licitacion'[fechaPublicacion].[Año])
    )
)
VAR TotalEstimadoAnterior =
CALCULATE(
[Promedio de Montos Estimados],
FILTER(
    ALLEXCEPT('public hecho_licitacion', 'public producto'[categoria]),
    'public hecho_licitacion'[fechaPublicacion].[Año] = MAX('public
hecho_licitacion'[fechaPublicacion].[Año]) - 1
)
)
RETURN
IF(
    ISBLANK(TotalEstimadoAnterior) || TotalEstimadoAnterior = 0,
    BLANK(),
    ((TotalEstimado - TotalEstimadoAnterior) / TotalEstimadoAnterior) * 100
)

```

```

Promedio de Montos Estimados =
DIVIDE(
    CALCULATE(
        SUMX(
            VALUES('public hecho_licitacion'[idLicitacion]),
            CALCULATE(
                MAX('public hecho_licitacion'[montoEstimadoLicitacion])
            )
        ),
        ALLEXCEPT(
            'public hecho_licitacion',
            'public hecho_licitacion'[idRegionComprador],
            'public producto'[categoria],
            'public hecho_licitacion'[fechaPublicacion].[Año],
            'public hecho_licitacion'[fechaPublicacion].[Mes],
            'public hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    ),
    CALCULATE(
        DISTINCTCOUNT('public hecho_licitacion'[idLicitacion]),
        ALLEXCEPT(
            'public hecho_licitacion',
            'public hecho_licitacion'[idRegionComprador],
            'public producto'[categoria],
            'public hecho_licitacion'[fechaPublicacion].[Año],

```

```

    'public.hecho_licitacion'.[fechaPublicacion].[Mes],
    'public.hecho_licitacion'.[fechaPublicacion].[Trimestre]
)
),
0
)
WITH Promedio2022 AS (
SELECT
    AVG(max_monto_estimado) AS monto_promedio_actual_2022
FROM
(
    SELECT
        hl."idLicitacion",
        MAX(hl."montoEstimadoLicitacion") AS max_monto_estimado
    FROM
        public."hecho_licitacion" hl
    JOIN
        public."producto" pd ON hl."idProducto" = pd."idProducto"
    JOIN
        public."licitacion" l ON hl."idLicitacion" = l."nroLicitacion"
    WHERE
        pd."categoria" = 'AGRICULTURA Y FORESTAL'
        AND EXTRACT(YEAR FROM hl."fechaPublicacion") = 2022
    GROUP BY
        hl."idLicitacion"
) AS subquery
), Promedio2021 AS (
SELECT
    AVG(max_monto_estimado) AS monto_promedio_actual_2021
FROM
(
    SELECT
        hl."idLicitacion",
        MAX(hl."montoEstimadoLicitacion") AS max_monto_estimado
    FROM
        public."hecho_licitacion" hl
    JOIN
        public."producto" pd ON hl."idProducto" = pd."idProducto"
    JOIN
        public."licitacion" l ON hl."idLicitacion" = l."nroLicitacion"
    WHERE
        pd."categoria" = 'AGRICULTURA Y FORESTAL'
        AND EXTRACT(YEAR FROM hl."fechaPublicacion") = 2021
    GROUP BY
        hl."idLicitacion"
)
)
```

```

        ) AS subquery2
    )

SELECT
    Promedio2022.monto_promedio_actual_2022,
    Promedio2021.monto_promedio_actual_2021,
    CASE
        WHEN Promedio2021.monto_promedio_actual_2021 IS NULL THEN 1
        ELSE ((Promedio2022.monto_promedio_actual_2022 -
    Promedio2021.monto_promedio_actual_2021) /
    Promedio2021.monto_promedio_actual_2021) * 100
    END AS TasaCrecimientoMontoPromedioEstimado
FROM
    Promedio2022
CROSS JOIN
    Promedio2021;

```

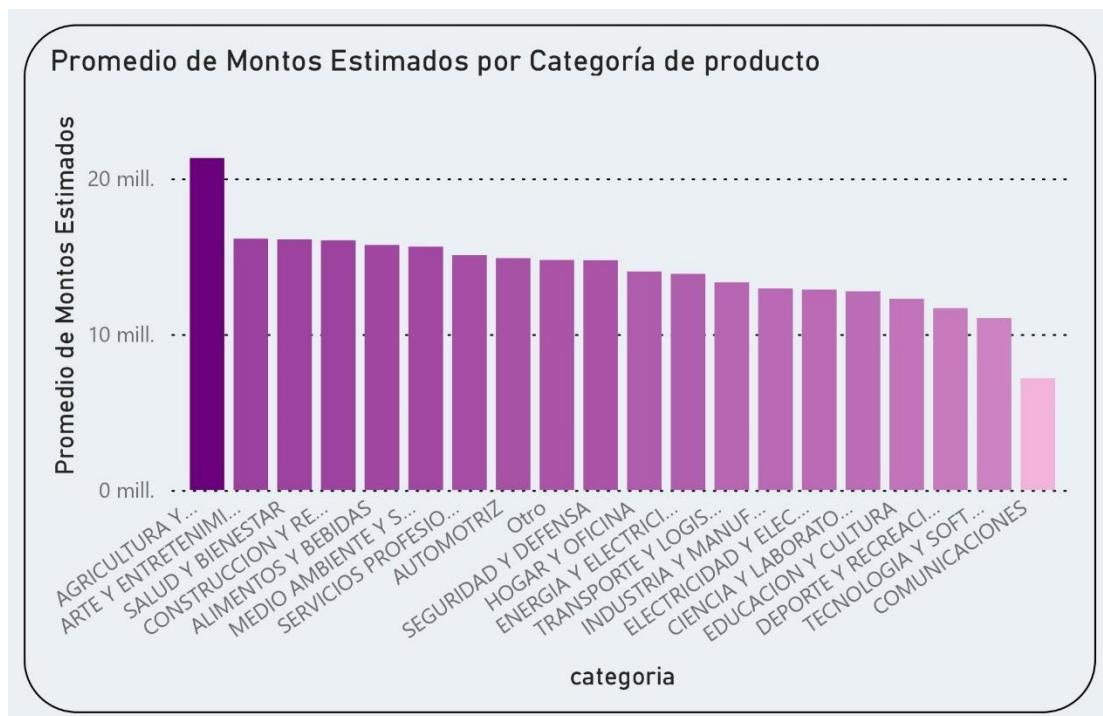
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

tasa crecimiento 2022 (2021 -> 2022) = 5.11

KPI 6 Promedio de montos estimados por región de comprador por categoría de producto.:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Promedio de Montos Estimados =
DIVIDE(
    CALCULATE(
        SUMX(
            VALUES('public_hecho_licitacion'[idLicitacion]),
            CALCULATE(
                MAX('public_hecho_licitacion'[montoEstimadoLicitacion])
            )
        ),
        ALLEXCEPT(
            'public_hecho_licitacion',
            'public_hecho_licitacion'[idRegionComprador],
            'public_producto'[categoria],
            'public_hecho_licitacion'[fechaPublicacion].[Año],
            'public_hecho_licitacion'[fechaPublicacion].[Mes],
            'public_hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    ),
    CALCULATE(
        DISTINCTCOUNT('public_hecho_licitacion'[idLicitacion]),
        ALLEXCEPT(
            'public_hecho_licitacion',
            'public_hecho_licitacion'[idRegionComprador],
            'public_producto'[categoria],
            'public_hecho_licitacion'[fechaPublicacion].[Año],
            'public_hecho_licitacion'[fechaPublicacion].[Mes]
        )
    )
)

```

```

        'public.hecho_licitacion'[fechaPublicacion].[Mes],
        'public.hecho_licitacion'[fechaPublicacion].[Trimestre]
    )
),
0
)

WITH TempTable AS (
    SELECT
        hl."idLicitacion",
        MAX(hl."montoEstimadoLicitacion") AS max_monto_estimado
    FROM
        public."hecho_licitacion" hl
    JOIN
        public."producto" pd ON hl."idProducto" = pd."idProducto"
    WHERE
        pd."categoria" = 'AGRICULTURA Y FORESTAL'
        AND EXTRACT(YEAR FROM hl."fechaPublicacion") = 2022
    GROUP BY
        hl."idLicitacion"
)
SELECT
    SUM(max_monto_estimado) / NULLIF(COUNT(DISTINCT "idLicitacion"), 0) AS
PromedioMontosEstimados
FROM
    TempTable;

```

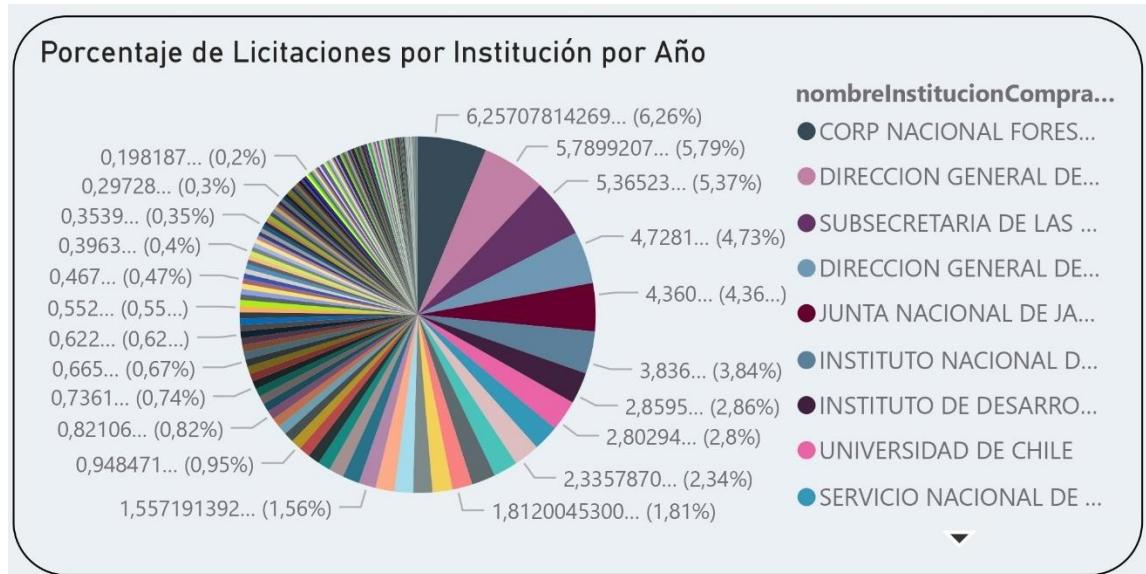
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

2022, AGRICULTURA Y FORESTAL = 22395268

KPI 7 Porcentaje de licitaciones por institución por año:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Porcentaje de Licitaciones por Institución =
DIVIDE(
    CALCULATE(
        DISTINCTCOUNT('public_hecho_licitacion'[idLicitacion]),
        ALLEXCEPT(
            'public_hecho_licitacion',
            'public_comprador'[nombreInstitucionCompradora],
            'public_producto'[categoria],
            'public_hecho_licitacion'[fechaPublicacion].[Año],
            'public_hecho_licitacion'[fechaPublicacion].[Mes],
            'public_hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    ),
    CALCULATE(
        DISTINCTCOUNT('public_hecho_licitacion'[idLicitacion]),
        ALLEXCEPT(
            'public_hecho_licitacion',
            'public_producto'[categoria],
            'public_hecho_licitacion'[fechaPublicacion].[Año],
            'public_hecho_licitacion'[fechaPublicacion].[Mes],
            'public_hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    ),
    0
) * 100

WITH cte_licitaciones_unicas AS (
    SELECT COUNT(DISTINCT hl."idLicitacion") AS ConteoDeLicitaciones
    FROM hecho_licitacion hl

```

```

JOIN public.comprador c ON hl."idComprador" = c."idComprador"
WHERE EXTRACT(YEAR FROM hl."fechaPublicacion") = 2021
AND c."nombreInstitucionCompradora" LIKE 'CORP NACIONAL FORESTAL'
),
cte_total_licitaciones AS (
  SELECT COUNT(distinct hl."idLicitacion") AS ConteoDeLicitacionesTotal
  FROM hecho_licitacion hl
  JOIN public.comprador c ON hl."idComprador" = c."idComprador"
  WHERE EXTRACT(YEAR FROM hl."fechaPublicacion") = 2021

)
SELECT
  cte_licitaciones_unicas.CuentoDeLicitaciones,
  cte_total_licitaciones.CuentoDeLicitacionesTotal,
  (cte_licitaciones_unicas.CuentoDeLicitaciones::float /
  cte_total_licitaciones.CuentoDeLicitacionesTotal) * 100 AS Porcentaje
FROM cte_licitaciones_unicas, cte_total_licitaciones;

```

Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

2021 CORP NACIONAL FORESTAL = 6.26%

KPI 8 Análisis de ranking de proveedores que realizan las ofertas más exitosas por región por año:

Interfaz OLAP

Análisis Ranking de Proveedores que realizan las ofertas más exitosas por año				
EMPRESA EDITORA ZIG ZAG S A	LIBERALIA EDICIONES L...		EDITORIAL ...	
	ESCRITO ...	LIBRE...	MERC...	LIBRE...
	DISTRIBU...	MUNN...	MI...	DI...
		CANO...		PONT...
	EDITORIA...	GRUP...	ETHO...	
EDITORIAL CONTRAPUNTO LTDA				

Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```
Análisis Ranking de Proveedores = [VolumenAjustado] * [Conteo Ofertas Ajustado] * [RatioExitoOfertas]
```

```
VolumenAjustado =
VAR MontoTotalOfertaGanadora =
    CALCULATE(
        SUM('public hecho_licitacion'[montoTotalOferta]),
        'public oferta'[resultadoOferta] = "GANADORA",
        ALLEXCEPT(
            'public hecho_licitacion',
            'public proveedor'[nombreProveedor], 'public proveedor'[regionProveedor],
            'public comprador'[regionUnidadCompradora], 'public
            hecho_licitacion'[fechaPublicacion].[Año], 'public
            hecho_licitacion'[fechaPublicacion].[Mes], 'public
            hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    )
RETURN
    IF(ISBLANK(MontoTotalOfertaGanadora), 0, MontoTotalOfertaGanadora)
```

```
Conteo Ofertas Ajustado =
VAR TotalOfertasGanadoras =
    CALCULATE(
        COUNTROWS('public hecho_licitacion'),
```

```

'public oferta'[resultadoOferta] = "GANADORA",
ALLEXCEPT(
    'public hecho_licitacion',
    'public proveedor'[nombreProveedor], 'public proveedor'[regionProveedor],
'public comprador'[regionUnidadCompradora], 'public
hecho_licitacion'[fechaPublicacion].[Año]
)
)
RETURN
IF(ISBLANK(TotalOfertasGanadoras), 0, TotalOfertasGanadoras)

-----
RatioExitoOfertas =
VAR TotalOfertasGanadorasPorRegionPorAno =
CALCULATE(
    COUNTROWS('public hecho_licitacion'),
    'public oferta'[resultadoOferta] = "GANADORA",
    ALLEXCEPT(
        'public hecho_licitacion',
        'public proveedor'[nombreProveedor], 'public proveedor'[regionProveedor],
'public comprador'[regionUnidadCompradora], 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre]
)
)
VAR TotalOfertasPorRegionPorAno =
CALCULATE(
    COUNTROWS('public hecho_licitacion'),
    ALLEXCEPT(
        'public hecho_licitacion',
        'public proveedor'[nombreProveedor], 'public proveedor'[regionProveedor],
'public comprador'[regionUnidadCompradora], 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre]
)
)
RETURN
IF(ISBLANK(TotalOfertasPorRegionPorAno), 0,
DIVIDE(TotalOfertasGanadorasPorRegionPorAno, TotalOfertasPorRegionPorAno, 0))

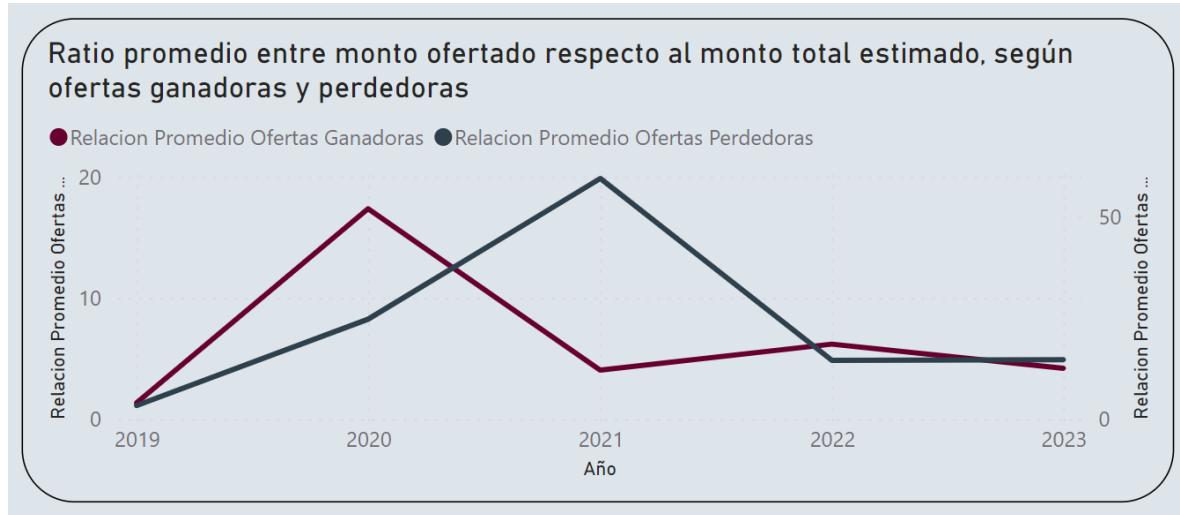
```

Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

KPI 9 Ratio promedio entre monto ofertado respecto al monto total estimado, según ofertas ganadoras y perdedoras:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Relacion Promedio Ofertas Ganadoras =
VAR MontoOfertadoMontoEstimado =
    CALCULATE(
        SUMX(
            'public hecho_licitacion',
            'public hecho_licitacion'[montoTotalOferta] / 'public
hecho_licitacion'[montoEstimadoLicitacion]
        ),
        'public oferta'[resultadoOferta] = "GANADORA",
        ALLEXCEPT(
            'public hecho_licitacion',
            'public proveedor'[nombreProveedor], 'public licitacion', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public
hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre]
        )
    )

VAR TotalOfertasGanadoras =
    CALCULATE(
        COUNTROWS('public hecho_licitacion'),
        'public oferta'[resultadoOferta] = "GANADORA",
        ALLEXCEPT(
            'public hecho_licitacion',
            'public proveedor'[nombreProveedor], 'public licitacion', 'public
hecho_licitacion'[fechaPublicacion].[Año], 'public

```

```

hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre]
)
)

RETURN
DIVIDE(
    MontoOfertadoMontoEstimado,
    TotalOfertasGanadoras,
    0
)

```

```

WITH Ganadores AS (
    SELECT
        hl."montoTotalOferta" / hl."montoEstimadoLicitacion" AS ratio
    FROM
        hecho_licitacion hl
    JOIN
        oferta o ON hl."idOferta" = o."idOferta"
    WHERE
        o."resultadoOferta" = 'GANADORA' AND EXTRACT(YEAR FROM
hl."fechaPublicacion") = 2021
),
Perdedores AS (
    SELECT
        hl."montoTotalOferta" / "montoEstimadoLicitacion" AS ratio
    FROM
        hecho_licitacion hl
    JOIN
        oferta o ON hl."idOferta" = o."idOferta"
    WHERE
        o."resultadoOferta" = 'PERDEDORA' AND EXTRACT(YEAR FROM
"fechaPublicacion") = 2021
)
SELECT
    AVG(ratio) AS promedio_ganadores,
    (SELECT AVG(ratio) FROM Perdedores) AS promedio_perdedores
FROM
    Ganadores;

```

Resultados Obtenidos

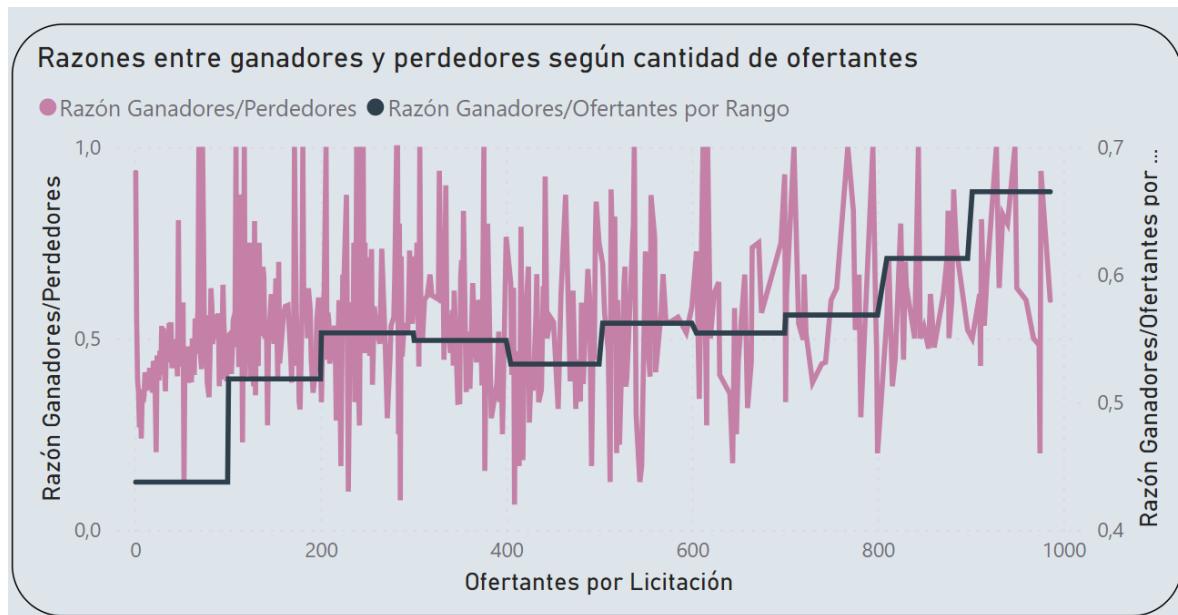
La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL (aunque con un pequeño margen de error, posiblemente por tratamiento de floats):

2021, Ganadores = 4.03

2021, Perdedores = 59.53

KPI 10 Razones entre ganadores y perdedores según cantidad de ofertantes:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

Razón Ganadores/Ofertantes por Rango =

```
DIVIDE(  
    SUM([Ganadores por Rango]),  
    SUM([Ofertantes por Rango]),  
    0  
)
```

Ganadores por Rango =

```
CALCULATE(  
    COUNTROWS('public hecho_licitacion'),  
    'public oferta'[resultadoOferta] = "GANADORA",  
    ALLEXCEPT('public hecho_licitacion', 'public hecho_licitacion'[Rango Ofertantes])  
)
```

Ofertantes por Rango =

```
CALCULATE(  
    COUNTROWS('public hecho_licitacion'),  
    ALLEXCEPT('public hecho_licitacion', 'public hecho_licitacion'[Rango Ofertantes])  
)
```

```

-----  

Rango Ofertantes =  

VAR CountOfertantes = CALCULATE(COUNTROWS('public hecho_licitacion'), ALLEXCEPT('public  

hecho_licitacion', 'public licitacion'[nroLicitacion], 'public  

licitacion'[nombreLicitacion]))  

VAR LowerBound = FLOOR((CountOfertantes - 1) / 100, 1) * 100  

VAR UpperBound = LowerBound + 99  

RETURN  

FORMAT(LowerBound, "0") & "-" & FORMAT(UpperBound, "0") & " ofertantes"  

-----  

WITH cte_rangos AS (  

    SELECT  

        CASE  

            WHEN COUNT(*) BETWEEN 0 AND 99 THEN '0-99'  

            WHEN COUNT(*) BETWEEN 100 AND 199 THEN '100-199'  

            -- Agrega más casos según sea necesario para otros rangos  

            ELSE 'Otros'  

        END AS rango_ofertantes,  

        hl."idLicitacion",  

        o."resultadoOferta"  

    FROM  

        public.hecho_licitacion hl  

        JOIN public.oferta o ON hl."idOferta" = o."idOferta"  

    WHERE  

        EXTRACT(YEAR FROM hl."fechaPublicacion") = 2020  

    GROUP BY  

        hl."idLicitacion", o."resultadoOferta"  

),  

cte_ganadores_por_rango AS (  

    SELECT  

        rango_ofertantes,  

        COUNT(*) AS ganadores_por_rango  

    FROM  

        cte_rangos  

    WHERE  

        "resultadoOferta" = 'GANADORA'  

    GROUP BY  

        rango_ofertantes  

),  

cte_ofertantes_por_rango AS (  

    SELECT  

        rango_ofertantes,  

        COUNT(*) AS ofertantes_por_rango  

    FROM  

        cte_rangos  

    GROUP BY

```

```

        rango_ofertantes
)
SELECT
    opr.rango_ofertantes,
    CASE
        WHEN opr.ofertantes_por_rango = 0 THEN NULL
        ELSE gpr.ganadores_por_rango * 1.0 / opr.ofertantes_por_rango
    END AS razon_ganadores_ofertantes
FROM
    cte_ofertantes_por_rango opr
LEFT JOIN
    cte_ganadores_por_rango gpr ON opr.rango_ofertantes =
gpr.rango_ofertantes;

```

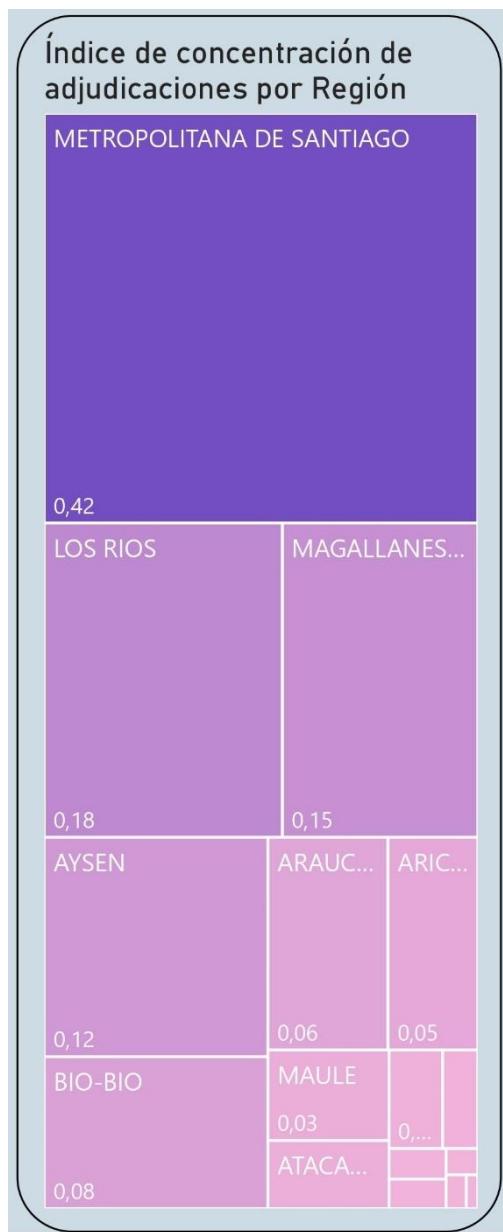
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

2020, 100-199 rango razon ganadores/ofertantes = 0.52

KPI 11 Índice de concentración de adjudicaciones por región:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Indice de Concentracion de Adjudicaciones por Region =
VAR SumaProporcionCuadrado =
    SUMX(
        VALUES('public comprador'[regionUnidadCompradora]),
        POWER([Proporcion de Adjudicaciones a Proveedores locales por Region], 2)
    )

RETURN
    SumaProporcionCuadrado

```

```

-----  

Proporcion de Adjudicaciones a Proveedores locales por Region =  

VAR Total_de_adjudicaciones_en_la_Region =  

    CALCULATE(  

        COUNTROWS('public hecho_licitacion'),  

        'public oferta'[resultadoOferta] = "GANADORA",  

        ALLEXCEPT('public hecho_licitacion', 'public  

comprador'[regionUnidadCompradora], 'public hecho_licitacion'[fechaPublicacion].[Año],  

'public hecho_licitacion'[fechaPublicacion].[Mes], 'public  

hecho_licitacion'[fechaPublicacion].[Trimestre])  

    )  

VAR Adjudicaciones_a_empresas_locales =  

    CALCULATE(  

        COUNTROWS(  

            FILTER(  

                'public hecho_licitacion',  

                RELATED('public comprador'[regionUnidadCompradora]) = RELATED('public  

proveedor'[regionProveedor]) &&  

                RELATED('public oferta'[resultadoOferta]) = "GANADORA"  

            )  

        ),  

        ALLEXCEPT('public hecho_licitacion', 'public  

comprador'[regionUnidadCompradora], 'public hecho_licitacion'[fechaPublicacion].[Año],  

'public hecho_licitacion'[fechaPublicacion].[Mes], 'public  

hecho_licitacion'[fechaPublicacion].[Trimestre])  

    )  

RETURN  

    DIVIDE(Adjudicaciones_a_empresas_locales, Total_de_adjudicaciones_en_la_Region, 0)  

WITH total_rows_cte AS (  

    -- Calcula el número total de filas donde la región de unidad de  

    compra del comprador es "METROPOLITANA" y el resultado de la oferta es  

    "GANADORA"  

    SELECT COUNT(*) AS total_rows  

    FROM public.hecho_licitacion hl  

    JOIN public.comprador c ON hl."idComprador" = c."idComprador"  

    JOIN public.oferta o ON hl."idOferta" = o."idOferta"  

    WHERE c."regionUnidadCompradora" = 'METROPOLITANA DE SANTIAGO' AND  

o."resultadoOferta" = 'GANADORA'  

),  

ganadora_same_region_rows_cte AS (  

    -- Calcula el número de filas donde el resultado de la oferta es  

    "GANADORA" y la región de unidad de compra del comprador es igual a la  

    región del proveedor  

    SELECT COUNT(*) AS ganadora_same_region_rows  

    FROM public.hecho_licitacion hl  

    JOIN public.comprador c ON hl."idComprador" = c."idComprador"

```

```

    JOIN public.oferta o ON hl."idOferta" = o."idOferta"
    JOIN public.proveedor p ON hl."idProveedor" = p."idProveedor"
    WHERE c."regionUnidadCompradora" = 'METROPOLITANA DE SANTIAGO' AND
o."resultadoOferta" = 'GANADORA'
        AND c."regionUnidadCompradora" = p."regionProveedor"
),
proporcion_cuadrado_cte AS (
    -- Calcula La proporción al cuadrado de las filas donde la región del
comprador y del proveedor coinciden
    SELECT
        CASE WHEN total_rows > 0 THEN
            POWER(ganadora_same_region_rows * 1.0 / total_rows, 2)
        ELSE
            0
        END AS proporcion_cuadrado
    FROM
        ganadora_same_region_rows_cte,
        total_rows_cte
)
-- Calcula el porcentaje y la proporción al cuadrado
SELECT
    CASE WHEN total_rows > 0 THEN
        (ganadora_same_region_rows * 100.0) / total_rows
    ELSE
        0
    END AS percentage,
    proporcion_cuadrado
FROM
    ganadora_same_region_rows_cte,
    total_rows_cte,
    proporcion_cuadrado_cte;

```

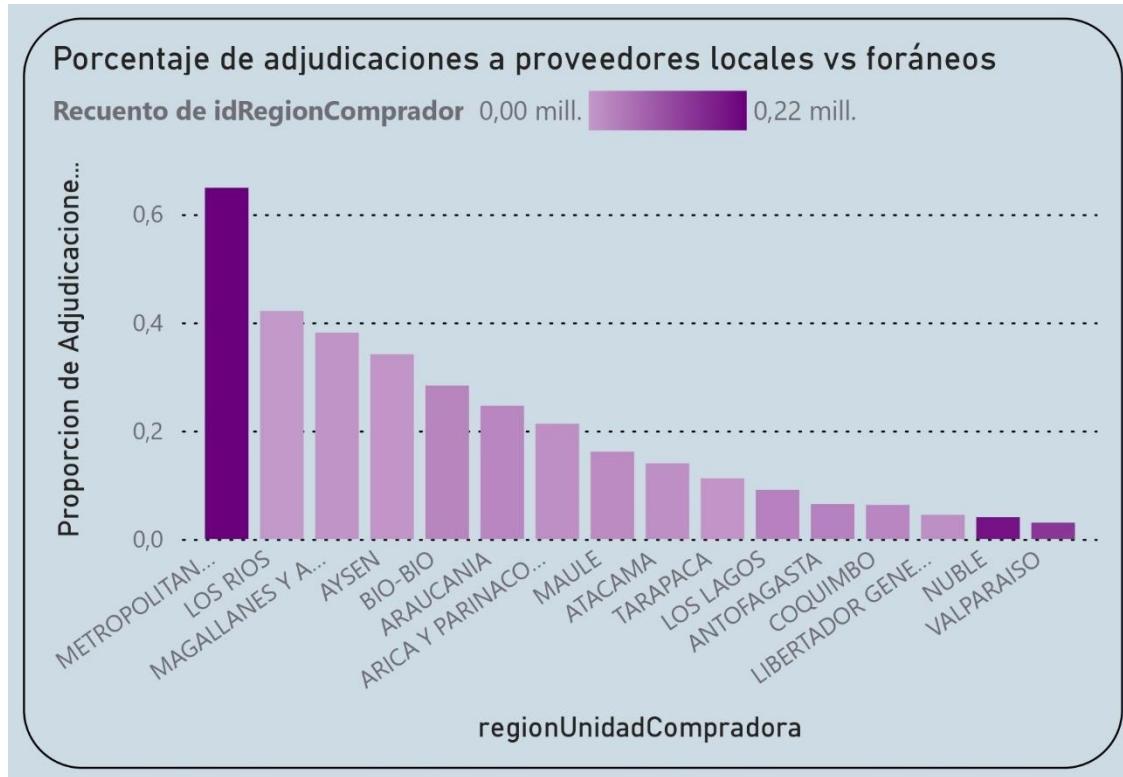
Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

total de años, Santiago = 0.42

KPI 12 Porcentaje de adjudicaciones a proveedores locales vs foráneos:

Interfaz OLAP



Sentencia SQL de Validación [Data Mart] y Query de Herramienta Olap.

```

Proporcion de Adjudicaciones a Proveedores locales por Region =
VAR Total_de_adjudicaciones_en_la_Región =
    CALCULATE(
        COUNTROWS('public hecho_licitacion'),
        'public oferta'[resultadoOferta] = "GANADORA",
        ALLEXCEPT('public hecho_licitacion', 'public
comprador'[regionUnidadCompradora], 'public hecho_licitacion'[fechaPublicacion].[Año],
        'public hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre])
    )

VAR Adjudicaciones_a_empresas_locales =
    CALCULATE(
        COUNTROWS(
            FILTER(
                'public hecho_licitacion',
                RELATED('public comprador'[regionUnidadCompradora]) = RELATED('public
proveedor'[regionProveedor]) &&
                RELATED('public oferta'[resultadoOferta]) = "GANADORA"
            )
        ),
        ALLEXCEPT('public hecho_licitacion', 'public
comprador'[regionUnidadCompradora], 'public hecho_licitacion'[fechaPublicacion].[Año],
        'public hecho_licitacion'[fechaPublicacion].[Mes], 'public
hecho_licitacion'[fechaPublicacion].[Trimestre])
    )

```

```

'public.hecho_licitacion'.[fechaPublicacion].[Mes], 'public
hecho_licitacion'.[fechaPublicacion].[Trimestre])
)

RETURN
    DIVIDE(Adjudicaciones_a_empresas_locales, Total_de_adjudicaciones_en_la_Región, 0)
WITH total_rows_cte AS (
    -- Calcula el número total de filas donde la región de unidad de compra del comprador es "LOS RIOS" y el resultado de la oferta es "GANADORA"
    SELECT COUNT(*) AS total_rows
    FROM public.hecho_licitacion hl
    JOIN public.comprador c ON hl."idComprador" = c."idComprador"
    JOIN public.oferta o ON hl."idOferta" = o."idOferta"
    WHERE c."regionUnidadCompradora" = 'LOS RIOS' AND o."resultadoOferta" = 'GANADORA'
),
ganadora_same_region_rows_cte AS (
    -- Calcula el número de filas donde el resultado de la oferta es "GANADORA" y la región de unidad de compra del comprador es igual a la región del proveedor
    SELECT COUNT(*) AS ganadora_same_region_rows
    FROM public.hecho_licitacion hl
    JOIN public.comprador c ON hl."idComprador" = c."idComprador"
    JOIN public.oferta o ON hl."idOferta" = o."idOferta"
    JOIN public.proveedor p ON hl."idProveedor" = p."idProveedor"
    WHERE c."regionUnidadCompradora" = 'LOS RIOS' AND o."resultadoOferta" = 'GANADORA'
        AND c."regionUnidadCompradora" = p."regionProveedor"
)
-- Calcula el porcentaje
SELECT (ganadora_same_region_rows * 100.0) / total_rows AS percentage
FROM ganadora_same_region_rows_cte, total_rows_cte;

```

Resultados Obtenidos

La validación y el dashboard comparten los valores asignados, el elegido para mostrar es el mismo que se selecciona en la sentencia SQL:

total de años, Los Rios = 0.42

10. Conclusiones

Conclusiones sobre el Proyecto

La implementación del datamart ha sido crucial para mejorar el proceso de toma de decisiones. Al facilitar el acceso y la extracción de información crítica, ha mejorado significativamente la capacidad para analizar y visualizar las tendencias del mercado de licitaciones para instituciones públicas.

Conclusiones sobre la Tecnología Utilizada

También es crucial la elección de herramientas que faciliten las tareas para el proyecto que se esté realizando. En el contexto del proceso para la creación del datamart, el uso de herramientas adecuadas fue vital dada la gran cantidad de datos que requirieron ser manipulados. En estos apartados, se puede destacar el gran enfoque visual de la herramienta *Knime* para el proceso de ETL facilitando bastante el aprendizaje y uso adecuado. Y también, por otro lado, se puede nombrar la relevancia de *Python*, que, gracias a su diversidad de librerías orientadas a la ciencia de datos, facilitaron procesos como el análisis exploratorio.

Conclusiones sobre los Resultados Obtenidos

Entre los principales hallazgos a mayor cantidad de ofertantes por licitación mayor índice de éxito, en su gran mayoría las ofertas realizadas tienen un presupuesto mayor estimado por la empresa y aun así son adjudicadas. Otro hallazgo interesante es, que, si bien la concentración de adjudicaciones coincide con la región metropolitana de Santiago, la segunda región con más adjudicaciones se corresponde con la región de los ríos.

11. Anexos

```
12. -- DROP SCHEMA public;
13.
14. CREATE SCHEMA public AUTHORIZATION postgres;
15.
16. COMMENT ON SCHEMA public IS 'standard public schema';
17.
18. -- DROP SEQUENCE public."comprador_idComprador_seq";
19.
20. CREATE SEQUENCE public."comprador_idComprador_seq"
21.     INCREMENT BY 1
22.     MINVALUE 1
23.     MAXVALUE 9223372036854775807
24.     START 1
25.     NO CYCLE;
26.
```

```

27. -- Permissions
28.
29. ALTER SEQUENCE public."comprador_idComprador_seq" OWNER TO gpallero;
30. GRANT ALL ON SEQUENCE public."comprador_idComprador_seq" TO
    gpallero;
31.
32. -- DROP SEQUENCE public."contrato_idContrato_seq";
33.
34. CREATE SEQUENCE public."contrato_idContrato_seq"
35.     INCREMENT BY 1
36.     MINVALUE 1
37.     MAXVALUE 9223372036854775807
38.     START 1
39.     NO CYCLE;
40.
41. -- Permissions
42.
43. ALTER SEQUENCE public."contrato_idContrato_seq" OWNER TO gpallero;
44. GRANT ALL ON SEQUENCE public."contrato_idContrato_seq" TO gpallero;
45.
46. -- DROP SEQUENCE public."hecho_licitacion_idProducto_seq";
47.
48. CREATE SEQUENCE public."hecho_licitacion_idProducto_seq"
49.     INCREMENT BY 1
50.     MINVALUE 1
51.     MAXVALUE 9223372036854775807
52.     START 1
53.     NO CYCLE;
54.
55. -- Permissions
56.
57. ALTER SEQUENCE public."hecho_licitacion_idProducto_seq" OWNER TO
    gpallero;
58. GRANT ALL ON SEQUENCE public."hecho_licitacion_idProducto_seq" TO
    gpallero;
59.
60. -- DROP SEQUENCE public.hecho_licitacion_id_seq;
61.
62. CREATE SEQUENCE public.hecho_licitacion_id_seq
63.     INCREMENT BY 1
64.     MINVALUE 1
65.     MAXVALUE 9223372036854775807
66.     START 1
67.     NO CYCLE;
68.

```

```

69. -- Permissions
70.
71. ALTER SEQUENCE public.hecho_licitacion_id_seq OWNER TO gpallero;
72. GRANT ALL ON SEQUENCE public.hecho_licitacion_id_seq TO gpallero;
73.
74. -- DROP SEQUENCE public."oferta_idOferta_seq";
75.
76. CREATE SEQUENCE public."oferta_idOferta_seq"
77.     INCREMENT BY 1
78.     MINVALUE 1
79.     MAXVALUE 9223372036854775807
80.     START 1
81.     NO CYCLE;
82.
83. -- Permissions
84.
85. ALTER SEQUENCE public."oferta_idOferta_seq" OWNER TO gpallero;
86. GRANT ALL ON SEQUENCE public."oferta_idOferta_seq" TO gpallero;
87.
88. -- DROP SEQUENCE public."producto_idProducto_seq";
89.
90. CREATE SEQUENCE public."producto_idProducto_seq"
91.     INCREMENT BY 1
92.     MINVALUE 1
93.     MAXVALUE 9223372036854775807
94.     START 1
95.     NO CYCLE;
96.
97. -- Permissions
98.
99. ALTER SEQUENCE public."producto_idProducto_seq" OWNER TO gpallero;
100. GRANT ALL ON SEQUENCE public."producto_idProducto_seq" TO gpallero;
101.
102. -- DROP SEQUENCE public."proveedor_idProveedor_seq";
103.
104. CREATE SEQUENCE public."proveedor_idProveedor_seq"
105.     INCREMENT BY 1
106.     MINVALUE 1
107.     MAXVALUE 9223372036854775807
108.     START 1
109.     NO CYCLE;
110.
111. -- Permissions
112.
113. ALTER SEQUENCE public."proveedor_idProveedor_seq" OWNER TO gpallero;

```

```

114. GRANT ALL ON SEQUENCE public."proveedor_idProveedor_seq" TO
    gpallero;
115.
116. -- DROP SEQUENCE public."region_idRegion_seq";
117.
118. CREATE SEQUENCE public."region_idRegion_seq"
119.     INCREMENT BY 1
120.     MINVALUE 1
121.     MAXVALUE 9223372036854775807
122.     START 1
123.     NO CYCLE;
124.
125. -- Permissions
126.
127. ALTER SEQUENCE public."region_idRegion_seq" OWNER TO gpallero;
128. GRANT ALL ON SEQUENCE public."region_idRegion_seq" TO gpallero;
129.
130. -- DROP SEQUENCE public."rubro_idRubro_seq";
131.
132. CREATE SEQUENCE public."rubro_idRubro_seq"
133.     INCREMENT BY 1
134.     MINVALUE 1
135.     MAXVALUE 9223372036854775807
136.     START 1
137.     NO CYCLE;
138.
139. -- Permissions
140.
141. ALTER SEQUENCE public."rubro_idRubro_seq" OWNER TO gpallero;
142. GRANT ALL ON SEQUENCE public."rubro_idRubro_seq" TO gpallero;
143. -- public.annio definition
144.
145. -- Drop table
146.
147. -- DROP TABLE public.annio;
148.
149. CREATE TABLE public.annio (
150.     "idAnnio" int4 NOT NULL,
151.     CONSTRAINT annio_pkey PRIMARY KEY ("idAnnio")
152. );
153.
154. -- Permissions
155.
156. ALTER TABLE public.annio OWNER TO gpallero;
157. GRANT ALL ON TABLE public.annio TO gpallero;

```

```

158.
159. -- public.comprador definition
160.
161. -- Drop table
162.
163. -- DROP TABLE public.comprador;
164.
165. CREATE TABLE public.comprador (
166.     "idComprador" serial4 NOT NULL,
167.     "rutUnidadComprador" varchar NULL,
168.     "nombreUnidadCompradora" varchar NULL,
169.     "codigoInstitucionCompradora" varchar NULL,
170.     "nombreInstitucionCompradora" varchar NULL,
171.     "regionUnidadCompradora" varchar NULL,
172.     CONSTRAINT comprador_pkey PRIMARY KEY ("idComprador")
173. );
174. CREATE INDEX "idx_nombreUnidadCompradora" ON public.comprador USING
    btree ("nombreUnidadCompradora");
175. CREATE INDEX "idx_rutUnidadComprador" ON public.comprador USING
    btree ("rutUnidadComprador");
176.
177. -- Permissions
178.
179. ALTER TABLE public.comprador OWNER TO gpallero;
180. GRANT ALL ON TABLE public.comprador TO gpallero;
181.
182. -- public.contrato definition
183.
184. -- Drop table
185.
186. -- DROP TABLE public.contrato;
187.
188. CREATE TABLE public.contrato (
189.     "idContrato" serial4 NOT NULL,
190.     contrato varchar NULL,
191.     "subContrato" varchar NULL,
192.     "unidadTiempoDuracion" varchar NULL,
193.     "plazoPagoContrato" varchar NULL,
194.     "tipoEjecucionContrato" varchar NULL,
195.     "tipoPagoContrato" varchar NULL,
196.     "extensionPlazo" varchar NULL,
197.     CONSTRAINT contrato_pkey PRIMARY KEY ("idContrato")
198. );
199.
200. -- Permissions

```

```

201.
202. ALTER TABLE public.contrato OWNER TO gpallero;
203. GRANT ALL ON TABLE public.contrato TO gpallero;
204.
205. -- public.licitacion definition
206.
207. -- Drop table
208.
209. -- DROP TABLE public.licitacion;
210.
211. CREATE TABLE public.licitacion (
212.     "nroLicitacion" varchar NOT NULL,
213.     "tipoLicitacion" varchar NULL,
214.     "nombreLicitacion" varchar NULL,
215.     "estadoLicitacion" varchar NULL,
216.     descripcion varchar NULL,
217.     "contemplaObrasPublicas" varchar NULL,
218.     "licitacionInformada" varchar NULL,
219.     CONSTRAINT licitacion_pkey PRIMARY KEY ("nroLicitacion")
220. );
221.
222. -- Permissions
223.
224. ALTER TABLE public.licitacion OWNER TO gpallero;
225. GRANT ALL ON TABLE public.licitacion TO gpallero;
226.
227. -- public.oferta definition
228.
229. -- Drop table
230.
231. -- DROP TABLE public.oferta;
232.
233. CREATE TABLE public.oferta (
234.     "idOferta" serial4 NOT NULL,
235.     "nombreOferta" varchar NULL,
236.     "estadoOferta" varchar NULL,
237.     "resultadoOferta" varchar NULL,
238.     CONSTRAINT oferta_pkey PRIMARY KEY ("idOferta")
239. );
240.
241. -- Permissions
242.
243. ALTER TABLE public.oferta OWNER TO gpallero;
244. GRANT ALL ON TABLE public.oferta TO gpallero;
245.

```

```

246. -- public.producto definition
247.
248. -- Drop table
249.
250. -- DROP TABLE public.producto;
251.
252. CREATE TABLE public.producto (
253.     "idProducto" serial4 NOT NULL,
254.     "codigoProducto" varchar NULL,
255.     "nombreItem" varchar NULL,
256.     "descripcionItem" varchar NULL,
257.     "unidadMedidaItem" varchar NULL,
258.     categoria varchar NULL,
259.     CONSTRAINT producto_pkey PRIMARY KEY ("idProducto")
260. );
261.
262. -- Permissions
263.
264. ALTER TABLE public.producto OWNER TO gpallero;
265. GRANT ALL ON TABLE public.producto TO gpallero;
266.
267. -- public.proveedor definition
268.
269. -- Drop table
270.
271. -- DROP TABLE public.proveedor;
272.
273. CREATE TABLE public.proveedor (
274.     "idProveedor" serial4 NOT NULL,
275.     "rutProveedor" varchar NULL,
276.     "nombreProveedor" varchar NULL,
277.     "tamanoProveedor" varchar NULL,
278.     "actividadProveedor" varchar NULL,
279.     "regionProveedor" varchar NULL,
280.     CONSTRAINT proveedor_pkey PRIMARY KEY ("idProveedor")
281. );
282. CREATE INDEX "idx_nombreProveedor" ON public.proveedor USING btree
    ("nombreProveedor");
283. CREATE INDEX "idx_rutProveedor" ON public.proveedor USING btree
    ("rutProveedor");
284.
285. -- Permissions
286.
287. ALTER TABLE public.proveedor OWNER TO gpallero;
288. GRANT ALL ON TABLE public.proveedor TO gpallero;

```

```

289.
290. -- public.region definition
291.
292. -- Drop table
293.
294. -- DROP TABLE public.region;
295.
296. CREATE TABLE public.region (
297.     "idRegion" serial4 NOT NULL,
298.     "nombreRegion" varchar NULL,
299.     "macroZona" varchar NULL,
300.     "nombreRegionCo" varchar NULL,
301.     CONSTRAINT region_pkey PRIMARY KEY ("idRegion")
302. );
303.
304. -- Permissions
305.
306. ALTER TABLE public.region OWNER TO gpallero;
307. GRANT ALL ON TABLE public.region TO gpallero;
308.
309. -- public.rubro definition
310.
311. -- Drop table
312.
313. -- DROP TABLE public.rubro;
314.
315. CREATE TABLE public.rubro (
316.     "idRubro" serial4 NOT NULL,
317.     rubro1 varchar NULL,
318.     rubro2 varchar NULL,
319.     "macroRubro1" varchar NULL,
320.     "macroRubro2" varchar NULL,
321.     CONSTRAINT rubro_pkey PRIMARY KEY ("idRubro")
322. );
323.
324. -- Permissions
325.
326. ALTER TABLE public.rubro OWNER TO gpallero;
327. GRANT ALL ON TABLE public.rubro TO gpallero;
328.
329. -- public.mes definition
330.
331. -- Drop table
332.
333. -- DROP TABLE public.mes;

```

```

334.
335. CREATE TABLE public.mes (
336.     "idMes" int4 NOT NULL,
337.     "idAnnio" int4 NULL,
338.     "semestreMes" int4 NULL,
339.     "numeroMes" int4 NULL,
340.     "nombreMes" varchar NULL,
341.     CONSTRAINT mes_pkey PRIMARY KEY ("idMes"),
342.     CONSTRAINT "mes_idAnnio_fkey" FOREIGN KEY ("idAnnio") REFERENCES
343.         public.annio("idAnnio")
344. );
345. CREATE INDEX "idx_idAnnio" ON public.mes USING btree ("idAnnio");
346. -- Permissions
347.
348. ALTER TABLE public.mes OWNER TO gpallero;
349. GRANT ALL ON TABLE public.mes TO gpallero;
350.
351. -- public.dia definition
352.
353. -- Drop table
354.
355. -- DROP TABLE public.dia;
356.
357. CREATE TABLE public.dia (
358.     "idDia" int4 NOT NULL,
359.     "idMes" int4 NULL,
360.     "numeroDiaAnnio" int4 NULL,
361.     "nombreDia" varchar NULL,
362.     "numeroDiaMes" int4 NULL,
363.     CONSTRAINT dia_pkey PRIMARY KEY ("idDia"),
364.     CONSTRAINT "dia_idMes_fkey" FOREIGN KEY ("idMes") REFERENCES
365.         public.mes("idMes")
366. );
367. CREATE INDEX "idx_idMes" ON public.dia USING btree ("idMes");
368. -- Permissions
369.
370. ALTER TABLE public.dia OWNER TO gpallero;
371. GRANT ALL ON TABLE public.dia TO gpallero;
372.
373. -- public.hecho_licitacion definition
374.
375. -- Drop table
376.

```

```

377. -- DROP TABLE public.hecho_licitacion;
378.
379. CREATE TABLE public.hecho_licitacion (
380.     id serial4 NOT NULL,
381.     "idLicitacion" varchar NOT NULL,
382.     "idComprador" int4 NOT NULL,
383.     "idProveedor" int4 NOT NULL,
384.     "idOferta" int4 NOT NULL,
385.     "idProducto" serial4 NOT NULL,
386.     "idContrato" int4 NOT NULL,
387.     "idDia" int4 NOT NULL,
388.     "idRubro" int4 NOT NULL,
389.     "idRegionComprador" int4 NOT NULL,
390.     "idRegionProveedor" int4 NOT NULL,
391.     "fechaPublicacion" date NULL,
392.     "fechaCierre" date NULL,
393.     "fechaAdjudicacion" date NULL,
394.     "cantidadItem" int4 NULL,
395.     "cantidadOferta" int4 NULL,
396.     "montoEstimadoLicitacion" int8 NULL,
397.     "montoNetoOferta" int8 NULL,
398.     "montoTotalOferta" int8 NULL,
399.     CONSTRAINT hecho_licitacion_pkey PRIMARY KEY (id),
400.     CONSTRAINT "hecho_licitacion_idComprador_fkey" FOREIGN KEY
        ("idComprador") REFERENCES public.comprador("idComprador"),
401.     CONSTRAINT "hecho_licitacion_idContrato_fkey" FOREIGN KEY
        ("idContrato") REFERENCES public.contrato("idContrato"),
402.     CONSTRAINT "hecho_licitacion_idDia_fkey" FOREIGN KEY ("idDia")
        REFERENCES public.dia("idDia"),
403.     CONSTRAINT "hecho_licitacion_idLicitacion_fkey" FOREIGN KEY
        ("idLicitacion") REFERENCES public.licitacion("nroLicitacion"),
404.     CONSTRAINT "hecho_licitacion_idOferta_fkey" FOREIGN KEY
        ("idOferta") REFERENCES public.oferta("idOferta"),
405.     CONSTRAINT "hecho_licitacion_idProducto_fkey" FOREIGN KEY
        ("idProducto") REFERENCES public.producto("idProducto"),
406.     CONSTRAINT "hecho_licitacion_idProveedor_fkey" FOREIGN KEY
        ("idProveedor") REFERENCES public.proveedor("idProveedor"),
407.     CONSTRAINT "hecho_licitacion_idRegionComprador_fkey" FOREIGN KEY
        ("idRegionComprador") REFERENCES public.region("idRegion"),
408.     CONSTRAINT "hecho_licitacion_idRegionProveedor_fkey" FOREIGN KEY
        ("idRegionProveedor") REFERENCES public.region("idRegion"),
409.     CONSTRAINT "hecho_licitacion_idRubro_fkey" FOREIGN KEY
        ("idRubro") REFERENCES public.rubro("idRubro")
410. );

```

```

411. CREATE INDEX "idx_fechaAdjudicacion" ON public.hecho_licitacion
      USING btree ("fechaAdjudicacion");
412. CREATE INDEX "idx_fechaCierre" ON public.hecho_licitacion USING
      btree ("fechaCierre");
413. CREATE INDEX "idx_fechaPublicacion" ON public.hecho_licitacion USING
      btree ("fechaPublicacion");
414. CREATE INDEX "idx_idComprador" ON public.hecho_licitacion USING
      btree ("idComprador");
415. CREATE INDEX "idx_idContrato" ON public.hecho_licitacion USING btree
      ("idContrato");
416. CREATE INDEX "idx_idDia" ON public.hecho_licitacion USING btree
      ("idDia");
417. CREATE INDEX "idx_idLicitacion" ON public.hecho_licitacion USING
      btree ("idLicitacion");
418. CREATE INDEX "idx_idOferta" ON public.hecho_licitacion USING btree
      ("idOferta");
419. CREATE INDEX "idx_idProducto" ON public.hecho_licitacion USING btree
      ("idProducto");
420. CREATE INDEX "idx_idProveedor" ON public.hecho_licitacion USING
      btree ("idProveedor");
421. CREATE INDEX "idx_idRegionComprador" ON public.hecho_licitacion
      USING btree ("idRegionComprador");
422. CREATE INDEX "idx_idRegionProveedor" ON public.hecho_licitacion
      USING btree ("idRegionProveedor");
423.
424. -- Permissions
425.
426. ALTER TABLE public.hecho_licitacion OWNER TO gpallero;
427. GRANT ALL ON TABLE public.hecho_licitacion TO gpallero;
428.

429. -- Permissions
430.
431. GRANT ALL ON SCHEMA public TO postgres;
432. GRANT ALL ON SCHEMA public TO public;

```