

Informe de análisis de constructora



Informe elaborado por:
Axel Berrios G.
18881064-0
Data Analyst

Problemática en el proyecto

En el ámbito de un proyecto de análisis integral para una constructora chilena, donde los datos operativos y financieros se generan en entornos fragmentados y de alta variabilidad, el rol de data scientist enfrenta desafíos significativos que limitan la capacidad de generar Insights accionables y oportunos.

Observaciones preliminares indican que los datos provienen de múltiples fuentes dispares —como sistemas de presupuestos en CLP, registros de horas trabajadas por rol, reportes de incidentes de seguridad, métricas de consumo energético en kWh y evaluaciones de sostenibilidad LEED—, lo que genera inconsistencias estructurales, tales como unidades mixtas (e.g., CLP vs. USD), valores anómalos (e.g., negativos en desperdicios) y formatos no estandarizados, complicando la integración y preparación inicial.

Adicionalmente, la ausencia de un data warehouse centralizado o pipelines automatizados resulta en un proceso manual intensivo para la recopilación y limpieza, exacerbado por el volumen creciente de datos de proyectos entre 2023 y 2026, que incluye variables cualitativas (e.g., descripciones de cuellos de botella) y cuantitativas (e.g., desviaciones presupuestarias y emisiones CO₂). Esto no solo consume recursos computacionales y tiempo, sino que también introduce riesgos de sesgos o errores en el análisis exploratorio, particularmente en un contexto donde la productividad y seguridad dependen de patrones ocultos en datos no procesados.

Otro aspecto crítico es la brecha entre los datos disponibles y las necesidades del negocio: mientras la constructora lida con sobrecostos, desperdicios y variabilidad energética, los dataset preliminares carecen de granularidad temporal o geográfica suficiente (e.g., por fase de proyecto o ubicación), limitando la detección temprana de tendencias como pics de incidentes o ineficiencias logísticas. La falta de herramientas avanzadas para modelado (e.g., ML para predicciones de riesgos) y visualizaciones interactivas agrava la dificultad para comunicar hallazgos a stakeholders no técnicos, como gerentes operativos o directivos.

Estos obstáculos en la gestión de datos erosionan la efectividad, potencialmente demorando decisiones estratégicas que podrían mitigar pérdidas en millones de pesos y alinear la constructora con estándares de sostenibilidad. Esta situación preliminar resalta la urgencia de un enfoque estructurado en data Science para transformar datos crudos en valor empresarial, justificando intervenciones como la estandarización de procesos y el desarrollo de modelos robustos.

El presente informe ofrece un análisis integral del desempeño de una constructora a través de múltiples dimensiones clave, incluyendo el control presupuestario, la eficiencia en el uso de materiales, la gestión del tiempo y la productividad, la seguridad laboral, la satisfacción de clientes, el consumo energético y la sostenibilidad ambiental. Se examinan datos detallados de diversos proyectos, identificando desviaciones entre costos presupuestados y reales, patrones de desperdicio de materiales, variaciones en las horas trabajadas por rol, y la incidencia de accidentes laborales. Además, se evalúan indicadores de eficiencia energética, certificación LEED y el impacto ambiental asociado al consumo energético por tipo de fuente.

Este análisis busca no solo diagnosticar áreas críticas y oportunidades de mejora, sino también proporcionar recomendaciones específicas para optimizar la gestión de recursos, reducir costos y pérdidas, mejorar la seguridad y aumentar la satisfacción de los clientes. Asimismo, se enfatiza la importancia de la sostenibilidad y la eficiencia energética como pilares fundamentales para el desarrollo de proyectos constructivos responsables y competitivos en el contexto actual.

Con base en datos cuantitativos y cualitativos, el informe pretende apoyar la toma de decisiones estratégicas que permitan a la constructora fortalecer sus procesos, maximizar la rentabilidad y contribuir positivamente al medio ambiente y a la comunidad.

Mi rol en este proyecto

Como data Science en este análisis integral del desempeño de la constructora, mi rol principal fue extraer valor accionable de datos cuantitativos y cualitativos dispersos, transformándolos en Insights estratégicos para optimizar operaciones, reducir riesgos y promover sostenibilidad. Colaborar en la fase de data-driven decision-making, asegurando rigor analítico y visualizaciones claras para stakeholders no técnicos.

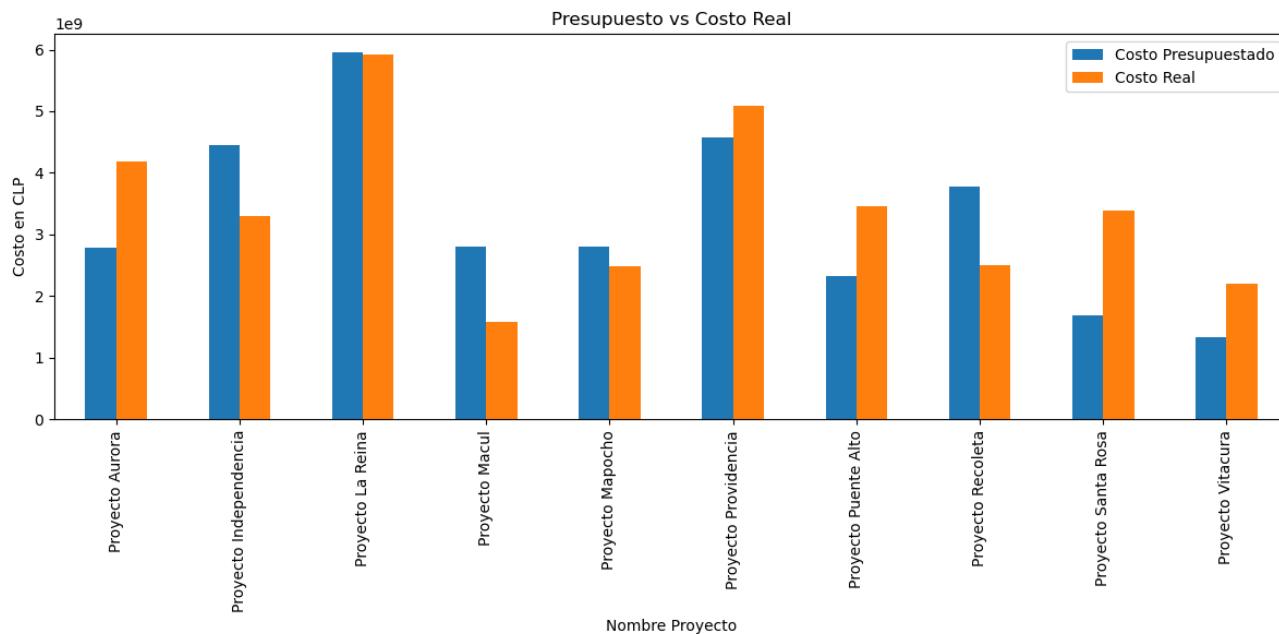
Tareas Específicas Realizadas

- **Recopilación y Preparación de Datos:** Integraste datasets de múltiples fuentes (e.g., presupuestos en CLP, horas trabajadas por rol, incidentes de seguridad, consumos energéticos en kWh), limpiando inconsistencias (e.g., valores negativos en desperdicios, unidades mixtas CLP/USD) y manejando anomalías para un dataset robusto de ~2023-2026.
- **Análisis Exploratorio de Datos (EDA):** Identificaste patrones clave, como desviaciones presupuestarias (>50% en Santa Rosa), correlaciones (e.g., costo presupuestado con real: 0.64) y cuellos de botella (200 en productividad), usando técnicas estadísticas para diagnosticar ineficiencias en materiales (cerámica/cableado) y fases (instalaciones: 0.65).
- **Modelado y Predicción:** Desarrollaste modelos simples, como clasificación binaria para accidentes (precisión/recall 1.00 en 40 instancias, matriz de confusión [[40]]), evaluando importancia de variables (e.g., "Parte Afectada" en 0.0) y alertando sobre limitaciones como overfitting; contribuí a análisis predictivos para incidentes y costos.
- **Visualizaciones y Reportes:** Creaste gráficos impactantes (e.g., mapas de calor para costos/incidentes por mes, diagramas de barras para satisfacción y eficiencia energética, burbujas para impacto CO2), facilitando la interpretación de tendencias (e.g., picos 2025) y distribuciones (LEED: 51.5%).
- **Interpretación y Recomendaciones:** Tradujiste resultados en soluciones data-backed, como umbrales de alerta para desviaciones (10-20%), optimización just-in-time para inventarios y mitigación de pérdidas temporales (12-13%), cuantificando impactos potenciales (e.g., reducción de costos 20-30%).

Impacto en el Proyecto

Mi contribución elevó el informe de descriptivo a prescriptivo, habilitando decisiones estratégicas que podrían ahorrar millones en CLP al reducir desperdicios y sobrecostos, mientras alinean la constructora con metas sostenibles (e.g., LEED y CO2). Identifique gaps metodológicos (e.g., correlaciones débiles en clima), recomendando iteraciones futuras con ML (**Machine Learning**) avanzado para mayor precisión. En resumen, fui el puente entre datos crudos y valor empresarial, impulsando eficiencia y competitividad.

Presupuesto vs Costo real



Proyecto	Costo Presupuestado (CLP)	Costo Real (CLP)	Desviación Absoluta (CLP)	Desviación %
Proyecto Santa Rosa	1.693.055.333	3.389.636.859	1.696.581.526	100.21%
Proyecto Vitacura	1.326.254.041	2.194.724.350	868.470.309	65.48%
Proyecto Aurora	2.780.556.564	4.189.838.682	1.409.282.118	50.68%
Proyecto Puente Alto	2.327.529.361	3.450.841.963	1.123.312.602	48.26%
Proyecto Providencia	4.574.022.921	5.085.425.301	511.402.380	11.18%

- **Desviación Absoluta:** indica cuánto se pasó en términos absolutos, en pesos chilenos.
- **Desviación %:** compara la desviación respecto al costo presupuestado, expresado en porcentaje.

Principales aspectos a destacar:

- El **Proyecto Santa Rosa** tuvo la desviación más alta, prácticamente duplicando el presupuesto inicial (100.21%). La diferencia es más grande que cualquier otro proyecto.
- Proyectos como **Vitacura** y **Aurora** tienen desviaciones significativas, con sobrecostos superiores al 50%.
- Por otro lado, **Providencia** muestra la desviación porcentual más baja (11.18%), acercándose a un control presupuestario aceptable.
- La desviación absoluta es también un indicador importante, ya que un porcentaje alto en proyectos pequeños puede no ser tan impactante como un porcentaje más bajo en proyectos grandes.

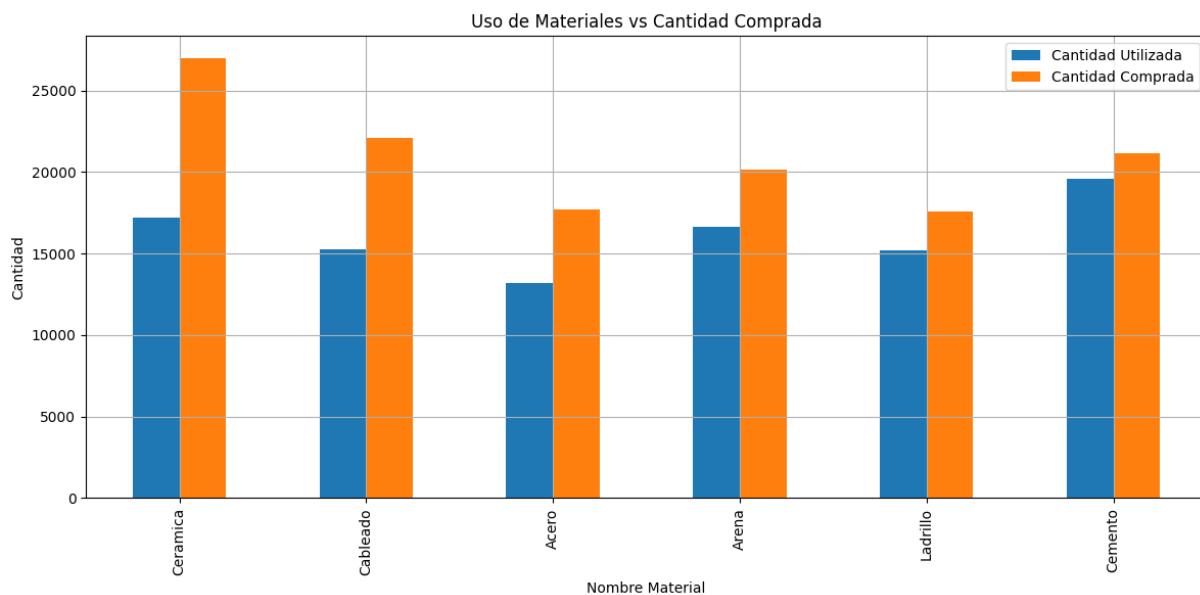
2. Interpretación del Gráfico:

- El gráfico de barras compara visualmente la magnitud del **Costo Presupuestado** (barra azul) y el **Costo Real** (barra naranja) para cada proyecto.
- En la mayoría de los proyectos mostrados, el costo real supera al presupuestado, visible por las barras naranjas que son más altas.
- En algunos proyectos que no están en la tabla (como "Proyecto Independencia", "Proyecto La Reina", "Proyecto Macul", "Proyecto Mapocho", "Proyecto Recoleta"), la barra azul de presupuesto es superior a la naranja, indicando que esos proyectos gastaron menos que el presupuesto.
- Esto indica que mientras algunos proyectos tienen problema de sobrecosto, otros han logrado mantenerse bajo control o incluso ahorrar dinero.

3. Análisis y recomendaciones finales:

- **Control presupuestario:** Los proyectos con desviaciones superiores al 50% necesitan un análisis profundo para entender las causas del sobrecosto (problemas en planificación, estimación incorrecta, imprevistos, mala gestión, etc.).
- **Priorizar recursos de control:** Los proyectos con mayores desviaciones absolutas y porcentuales, como Santa Rosa y Vitacura, deben ser priorizados para atención y análisis correctivo.
- **Aprendizajes para futuros proyectos:** Incorporar mejores prácticas de estimación basadas en datos históricos, gestión de riesgos y monitoreo continuo para evitar desviaciones significativas.
- **Evaluar proyectos que controlaron costos:** Los proyectos con costos reales por debajo del presupuesto (según gráfico) deben ser estudiados para extraer buenas prácticas que puedan replicarse.
- **Implementar métricas y alertas:** Definir umbrales para desviación (ej., 10%, 20%) para activar acciones tempranas durante la ejecución del proyecto.

Uso de materiales vs Cantidad comprada



1. Eficiencia en el uso de materiales

- **Materiales con diferencia pequeña (Cemento y Ladrillo):**
 - Esto apunta a un control efectivo sobre el uso y adquisición del material, indicando que las estimaciones de compra están bastante ajustadas a lo que realmente se necesita.
 - Implica que el proceso de cálculo y planificación es confiable.
 - Menor riesgo de acumulación de inventario o desperdicio, lo que contribuye a reducir costos indirectos como almacenamiento o deterioro.
- **Materiales con diferencia grande (Cerámica y Cableado):**
 - Aquí puede haber un uso ineficiente o falta de control, quizás por compras excesivas o imprecisiones en la planificación.
 - Podría también reflejar que parte del material se almacena para proyectos futuros o tiene tiempos de uso prolongados que no coinciden con la compra inmediata.
 - Posibilidad de desperdicio por daños, obsolescencia o mala gestión.

2. Inventario y control de stock

- **Identificación de áreas de mejora:**
 - Diferencias grandes sugieren oportunidades para mejorar la gestión del inventario.
 - Por ejemplo, una compra de cerámica superior a la utilizada implica un exceso que puede llevar a costos innecesarios por almacenamiento, riesgos de daño o envejecimiento del material.
- **Acciones recomendadas:**
 - Implementar mejores controles de inventario just-in-time para reducir acumulación.
 - Revisar frecuencia y volúmenes de compra.
 - Audituar la rotación del inventario para detectar materiales obsoletos o en exceso.

3. Planeación de compras

- **Materiales con diferencia moderada (Acero y Arena):**
 - Indican que la planificación está parcialmente alineada, pero aún se puede afinar para minimizar la brecha.
 - Ajustar la planificación puede ayudar a optimizar costos y espacio, evitando sobrecompra sin sacrificar disponibilidad.
- **Uso de datos para futuras compras:**
 - Análisis de tendencias basado en consumo pasado para ajustar cantidades.
 - Considerar variantes como cambios en el diseño, número de proyectos o condiciones de obra que podrían afectar el uso real.

4. Posibles pérdidas o mermas

- **Exceso significativo de compra respecto a uso:**
 - Puede reflejar pérdidas por rotura, deterioro, merma en transporte o almacenamiento inadecuado.
 - Por ejemplo, materiales frágiles o con alta exposición a factores ambientales pueden reducirse antes de ser utilizados.
- **Revisión y mejora de procesos:**
 - Implementar mejores prácticas en manipulación y almacenamiento.
 - Capacitar al personal para reducción de desperdicios.
 - Usar tecnología para monitoreo en tiempo real de inventarios.

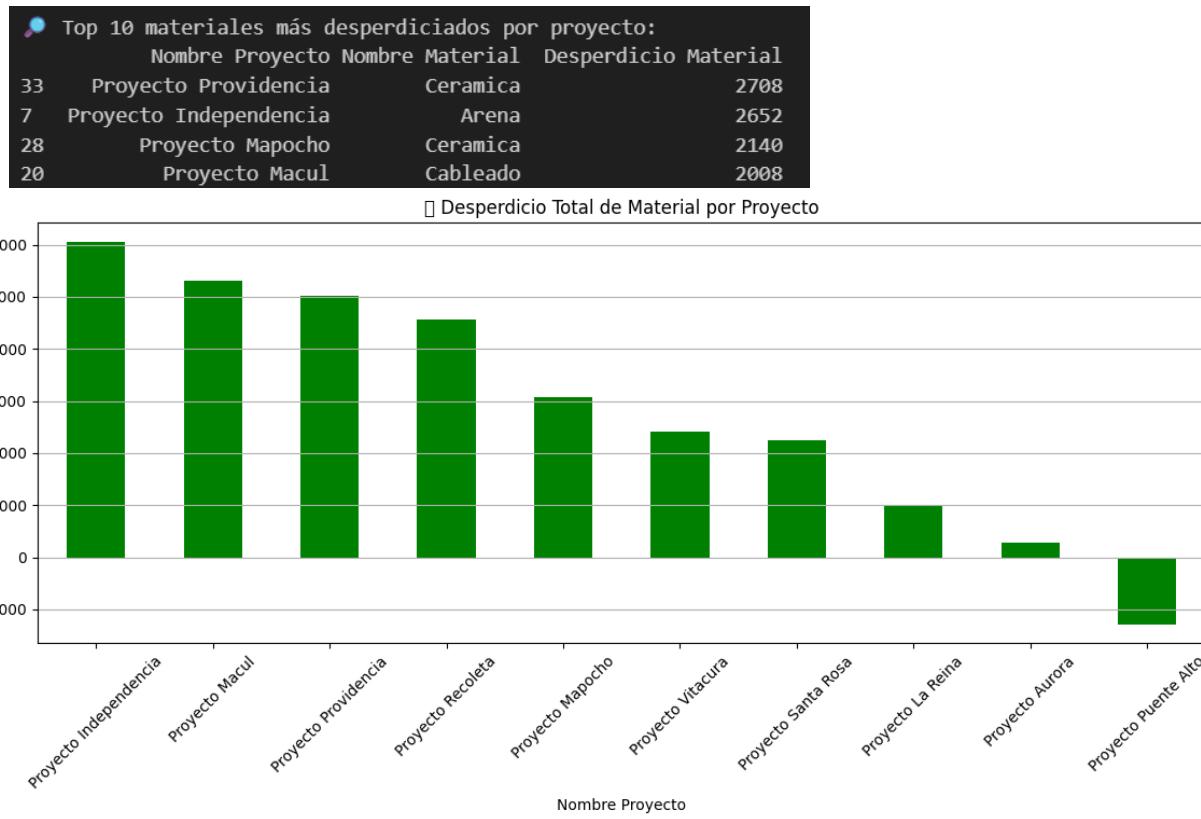
5. Análisis de proyectos futuros

- **Patrones constantes de diferencias:**
 - Un patrón repetitivo de diferencias altas o bajas en ciertos materiales indica que los cálculos iniciales pueden no ser precisos o que el proceso de consumo varía.

- **Revisión y recálculo:**
 - Utilizar datos históricos para ajustar requerimientos, evitando la compra excesiva y minimizando costes.
 - Considerar factores externos como cambios en normativa, innovaciones en materiales o metodologías constructivas.
 - Evaluar la posibilidad de flexibilizar contratos de compra para ser más reactivos a cambios en demanda.

Este análisis permite identificar cómo optimizar la adquisición y uso de materiales, mejorar el control de inventarios, reducir costos por pérdidas o almacenamiento y planificar mejor futuros proyectos para lograr mayor eficiencia y rentabilidad.

Desperdicio total de materiales por proyecto



- **Proyecto Independencia**
 - Es el que presenta el mayor desperdicio total, con más de 6000 unidades.
 - Implica un enfoque crítico para optimizar los procesos y mejorar el control en todos los materiales.
 - Probablemente concentra altos desperdicios en cerámica, cableado y acero, lo que debe confirmarse.
- **Proyecto Macul**
 - Segundo en desperdicio total, con valores cercanos a 5300 unidades.
 - Su seguimiento es clave para identificar si el desperdicio está concentrado en pocos materiales o es generalizado.
 - Materiales como cableado y acero destacan aquí, sugiriendo posible mejora en su manejo.

- **Proyecto Providencia**
 - Alrededor de 4500 unidades de desperdicio total, también considerable.
 - Cerámica predomina como material desperdiaciado; se benefician de controles específicos en este material.
- **Proyecto Recoleta y Mapocho**
 - Ambos presentan niveles medios-altos de desperdicio, alrededor de 2000-3000 unidades.
 - Este rango indica que podrían tener focos de desperdicio en materiales claves como acero y cerámica.
- **Proyectos con menor desperdicio (Santa Rosa, La Reina, Aurora)**
 - Aunque presentan menos volumen de desperdicio, es importante no descuidar su control para evitar incrementos futuros.
- **Proyecto Puente Alto**
 - Muestra un valor negativo, posible error en registro, cálculo o medición.
 - Urge revisar y corregir estos datos para no afectar auditorías ni análisis comparativos.

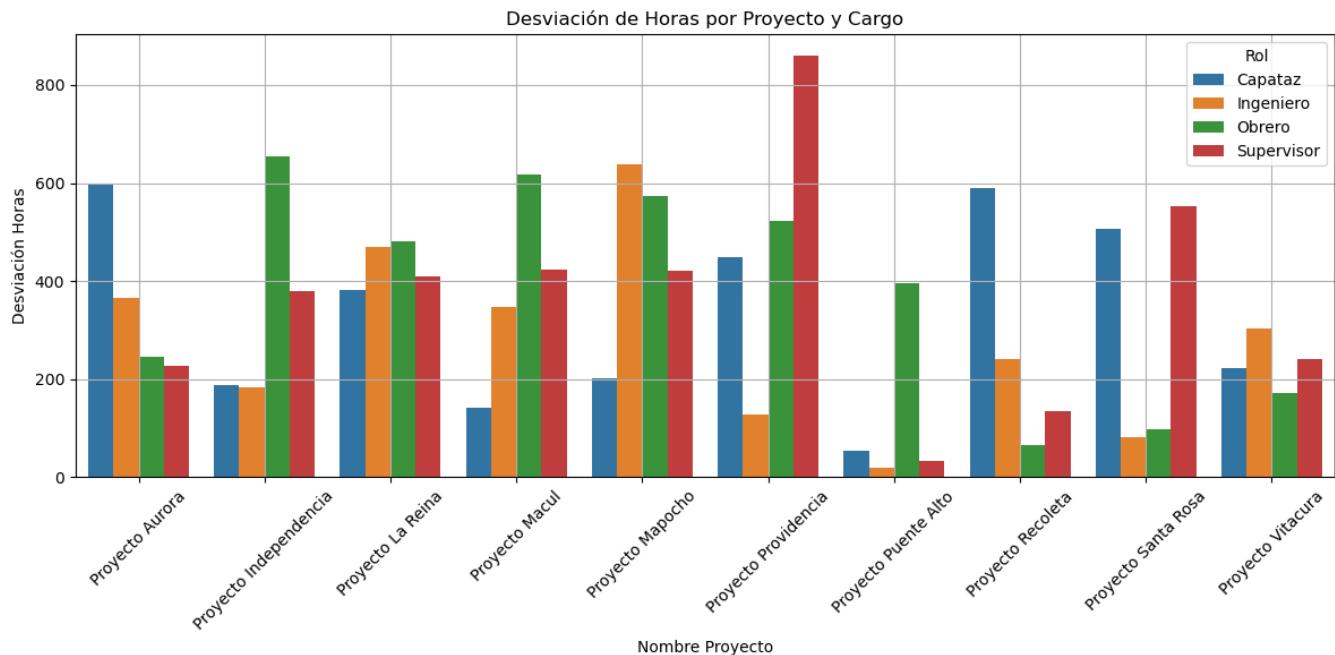
Análisis Detallado por Material

- **Cerámica**
 - Material más desperdiaciado en general.
 - Principal foco en proyectos Providencia, Independencia y Mapocho.
 - Por su relevancia, el control de inventarios y capacitación en manipulación son críticos.
 - Se puede explorar alternativas para minimizar roturas y sobrantes.
- **Cableado**
 - Destacado en proyectos Macul, Independencia, Vitacura y Recoleta.
 - Existe posibilidad de pérdidas por corte excesivo, roturas o manejo inadecuado.
 - Requiere inventarios frecuentes y protocolos de reutilización.
- **Acero**
 - Complementa el grupo de materiales con alto desperdicio en proyectos similares a cableado.
 - La eficiencia en la medición y corte correcto podría reducir pérdidas significativas.
 - Es recomendable implementar controles adicionales en proyectos con altos volúmenes.
- **Arena y Cemento**
 - Aparecen en cantidades relevantes en algunos proyectos, pero no tan críticos como los anteriores.
 - Revisar mezclas, almacenamiento y transporte para optimizar su uso.

Conclusiones y Recomendaciones Específicas

- **Priorización de Proyectos**
 - Dirigir esfuerzos inmediatos a Independencia, Macul y Providencia para impactos visibles en reducción de desperdicio.
- **Estrategias por Material**
 - Planes específicos para cerámica, cableado y acero deben ser diseñados y aplicados.
 - Formación técnica para el personal y revisión de proveedores e inventarios.
- **Corrección de Datos y Calidad de Información**
 - Investigar causa del error en Puente Alto para asegurar integridad en el análisis.
 - Implementar validaciones periódicas en recolección de datos.
- **Impacto en Costos y Medio Ambiente**
 - Reducción de desperdicios alivia presión económica y mejora la huella ecológica.
 - Promover cultura de eficiencia en obra y diseño.

Desviación de horas por proyecto y cargo



Detalle de las desviaciones de horas por proyecto y cargo:

- La gráfica muestra la desviación de horas trabajadas en varios proyectos, desglosadas por cuatro roles: Capataz, Ingeniero, Obrero y Supervisor.
- Cada barra representa la desviación de horas en un proyecto específico, con diferentes colores para cada rol.

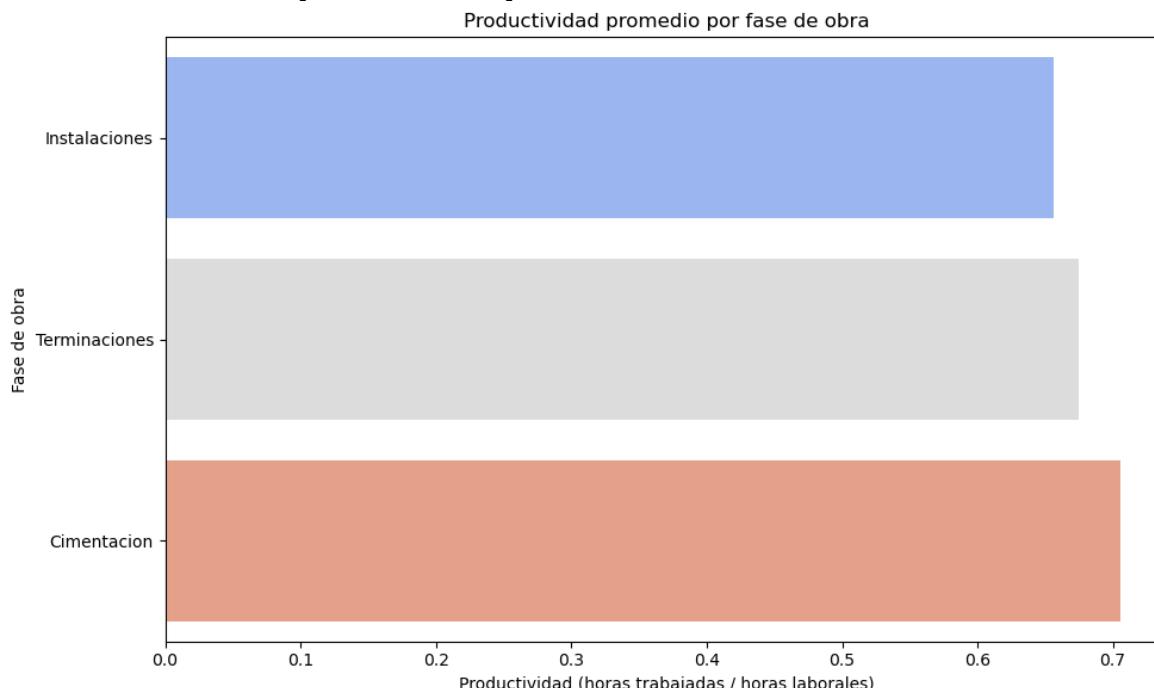
Análisis general de las desviaciones:

- Proyectos con mayores desviaciones:**
 - Proyecto Aurora:** Capataz tiene alta desviación (~600 horas) y Obrero (~650 horas).
 - Proyecto Mapocho:** Ingeniero presenta la mayor desviación (~640 horas) y Supervisor una desviación muy alta (~860 horas).
 - Proyecto Providencia:** Supervisor tiene una desviación muy grande (~860 horas).
 - Proyecto Macul:** Obrero (~620 horas) y Capataz (~140 horas) con desviaciones significativas.
 - Proyecto La Reina:** Obrero (~480 horas) e Ingeniero (~470 horas).
- Proyectos con menores desviaciones:**
 - Proyecto Puente Alto:** Muy bajas desviaciones en la mayoría de roles.
 - Proyecto Recoleta:** Obrero tiene baja desviación (~60 horas).
 - Proyecto Santa Rosa:** Obrero presenta casi 0 desviación.
- Roles con desviaciones más frecuentes:**
 - Obreros y Supervisores tienden a tener desviaciones altas en varios proyectos.
 - Ingenieros también presentan altas desviaciones en algunos proyectos clave como Proyecto Mapocho.
 - Capataces tienen desviaciones grandes principalmente en Proyecto Aurora y Santa Rosa.

Posibles razones de las desviaciones:

1. **Cambios en la planificación/cronograma:**
 - Ajustes en tiempos proyectados vs tiempos reales debido a retrasos o adelantos en tareas específicas.
2. **Factores imprevistos:**
 - Problemas técnicos, condiciones climáticas o disponibilidad de materiales pueden generar desviaciones en horas trabajadas, especialmente en roles como Obreros y Supervisores.
3. **Diferencias en eficiencia o asignaciones de tareas:**
 - Roles con responsabilidades de supervisión o control (Supervisor, Capataz) pueden tener desviaciones altas si se requieren más horas para supervisar problemas o coordinar actividades.
4. **Errores en el reporte o control de horas:**
 - Desviaciones pueden darse por inconsistencias en cómo se reportan o descuentan horas trabajadas en cada proyecto por rol.
5. **Retrabajos o actividades no planificadas:**
 - Supervisores y obreros pueden tener aumentos de horas por trabajos adicionales o corrección de errores en obra.

Productividad promedio por fase de obra



1. Fases con menor productividad:

- **Instalaciones:** 0.656454
- **Terminaciones:** 0.675077
- **Cimentación:** 0.705526

Estas cifras indican que la eficiencia en estas fases es baja, tomando en cuenta que la productividad se mide como horas trabajadas efectivas sobre horas laborales disponibles, y valores bajos muestran una menor utilización efectiva del tiempo.

2. Relación con cuellos de botella detectados:

- Se identificaron **200 cuellos de botella** en total que impactan en la productividad.
- Muestra parcial de los registrados con fases y problemas:

Fase	ID	Descripción problema
Cimentación	CBT-001	Problema logístico
Cimentación	CBT-002	Sin problema
Terminaciones	CBT-003	Demora externa
Cimentación	CBT-004	Demora externa
Instalaciones	CBT-005	Demora externa

3. Análisis de problemas asociados a cuellos de botella y cómo impactan en la productividad:

- **Problemas logísticos en Cimentación:**
Producción relativamente más alta (0.705526) a pesar de cuellos de botella como problemas logísticos. Esto sugiere que, aunque hay desafíos logísticos, estos no impactan tan severamente en la productividad como otros problemas.
- **Demoras externas (Terminaciones e Instalaciones):**
Ambas fases presentan productividad más baja y cuellos de botella asociados a demoras externas, lo cual puede implicar dependencia de terceros o suministros que retrasan la operación directa y afectan la cantidad de horas efectivas trabajadas.
- **Sin problema reportado:**
Hay registros de cuellos de botella sin problema identificado, lo que invita a revisar si existen problemas no detectados o enmascarados (por ejemplo, mala planificación, recursos insuficientes) que podrían estar afectando indirectamente la productividad.

4. Implicaciones y recomendaciones para mitigar cuellos de botella:

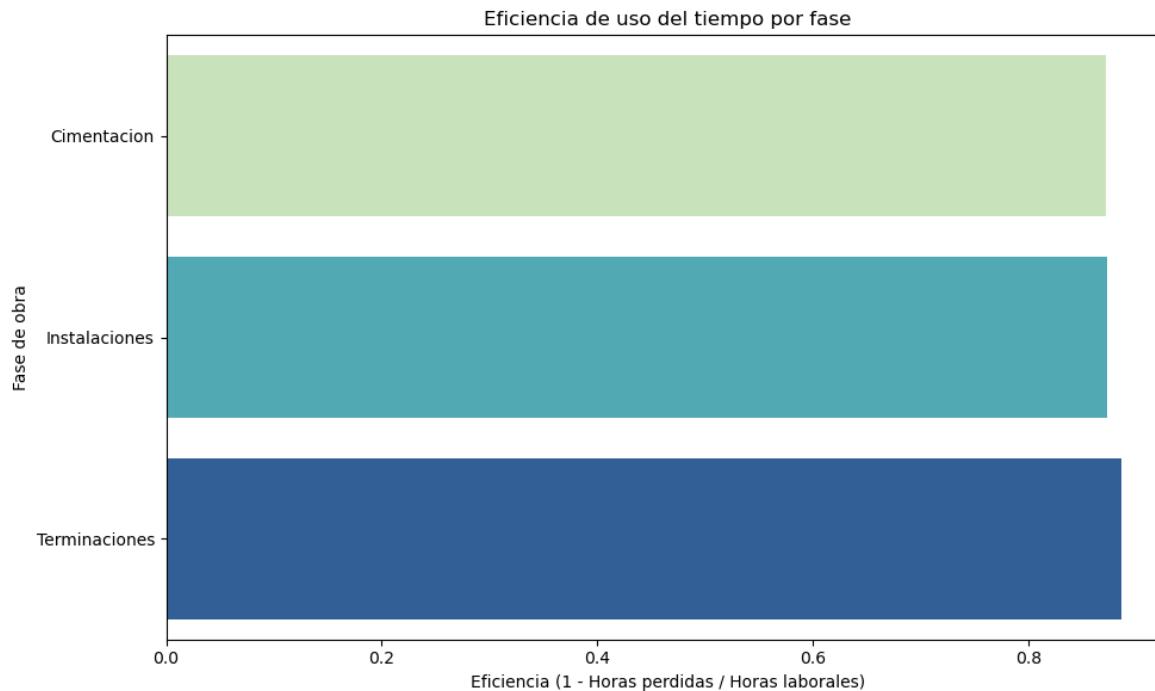
- **Fase Instalaciones (productividad más baja 0.656454):**
Priorizar la identificación y mitigación de causas de demoras externas (proveedores, permisos, accesos), ya que impactan significativamente la productividad.
- **Fase Terminaciones:**
Atender las demoras externas mediante acuerdos más firmes con contratistas o mejorar la planificación para reducir tiempos o anticipar retrasos.
- **Fase Cimentación:**
Mejorar la logística (movilización de materiales, maquinaria), que, aunque no es la más baja en productividad, sigue siendo crítica y afecta una carga importante de la obra.
- **Monitoreo continuo y análisis adicional:**
Implementar un sistema para capturar la raíz exacta de cada cuello de botella, incluso los que reportan "sin problema", para identificar causas ocultas y tener un control preciso.

5. Visualización del gráfico:

- El gráfico horizontal muestra claramente que la productividad en horas trabajadas sobre horas laborales es menor para Instalaciones, seguida de Terminaciones y Cimentación.
- Esto coincide con los cuellos de botella reportados y las causas, reforzando la importancia de focalizar esfuerzos en las fases Instalaciones y Terminaciones para mejorar la eficiencia general.

Las fases con menor productividad coinciden con aquellas donde se han detectado más cuellos de botella relacionados principalmente con demoras externas y problemas logísticos. Abordar estos problemas específicos en cada fase puede contribuir significativamente a aumentar la productividad general del proyecto.

Eficiencia de uso del tiempo por fase



¿Qué mide la eficiencia del tiempo?

La métrica $\text{eficiencia_tiempo} = 1 - (\text{horas perdidas} / \text{horas laborales})$ representa la proporción del tiempo laboral que fue efectivamente utilizado.

Por ejemplo, un valor de 0.87 significa que el 87% del tiempo fue útil, y el 13% se perdió por diversos motivos (esperas, mal uso de recursos, problemas logísticos, etc.).

❖ Análisis por fase:

Fase de obra	Eficiencia (%)	Tiempo perdido (%)	interpretación
Cimentación	87.24%	12.76%	Alta pérdida de tiempo para una fase crítica inicial. Puede deberse a atrasos logísticos, mal clima, fallos de planificación o disponibilidad de materiales.
Instalaciones	87.26%	12.74%	Similar a cimentación, esta etapa intermedia está perdiendo más de 12% del tiempo. Puede deberse a errores de coordinación entre equipos (electricidad, sanitarios, etc.).
Terminaciones	88.62%	11.38%	Aunque algo mejor que las anteriores, sigue teniendo una pérdida relevante. A menudo las terminaciones se atrasan por dependencia de tareas anteriores no completadas o falta de insumos específicos.

Recomendaciones:

Cimentación:

- Revisar logística de entrega de materiales pesados.
- Analizar interferencias por clima o problemas en la excavación.
- Evaluar si la maquinaria o el personal estaba disponible en los tiempos requeridos.

Instalaciones:

- Mejorar la coordinación entre disciplinas (gas, electricidad, sanitario).
- Verificar si existen conflictos en planos o rediseños constantes.
- Implementar mayor supervisión y control de tiempos.

Terminaciones:

- Revisar cuellos de botella de materiales de terminación (acabados, pinturas, etc.).
- Analizar si se están haciendo reprocesos por errores en fases anteriores.

Las tres fases con mayor pérdida de tiempo concentran entre un 11% y 13% de ineficiencia, lo cual puede significar semanas de retraso acumulado en un proyecto completo si no se toman acciones correctivas.

1. Fases con mayor pérdida de tiempo

- Según los valores proporcionados, la eficiencia de uso del tiempo para cada fase es:
 - Cimentación: 0.872356
 - Instalaciones: 0.872588
 - Terminaciones: 0.886220

Estos valores representan la proporción del tiempo efectivo utilizado, con 1 indicando uso óptimo (sin pérdidas). Como son menores a 1, indican que hay pérdidas de tiempo en las tres fases.

2. Interpretación del gráfico

- El gráfico muestra la eficiencia para cada fase, en una escala donde valores cercanos a 1 significan menos pérdida de tiempo y mejor rendimiento.
- La fase de Terminaciones muestra la mayor eficiencia (~0.886), seguida muy de cerca por Instalaciones (~0.8726) y Cimentación (~0.8724).
- La diferencia entre Instalaciones y Cimentación es muy pequeña, casi imperceptible, pero ambas están por debajo de Terminaciones.

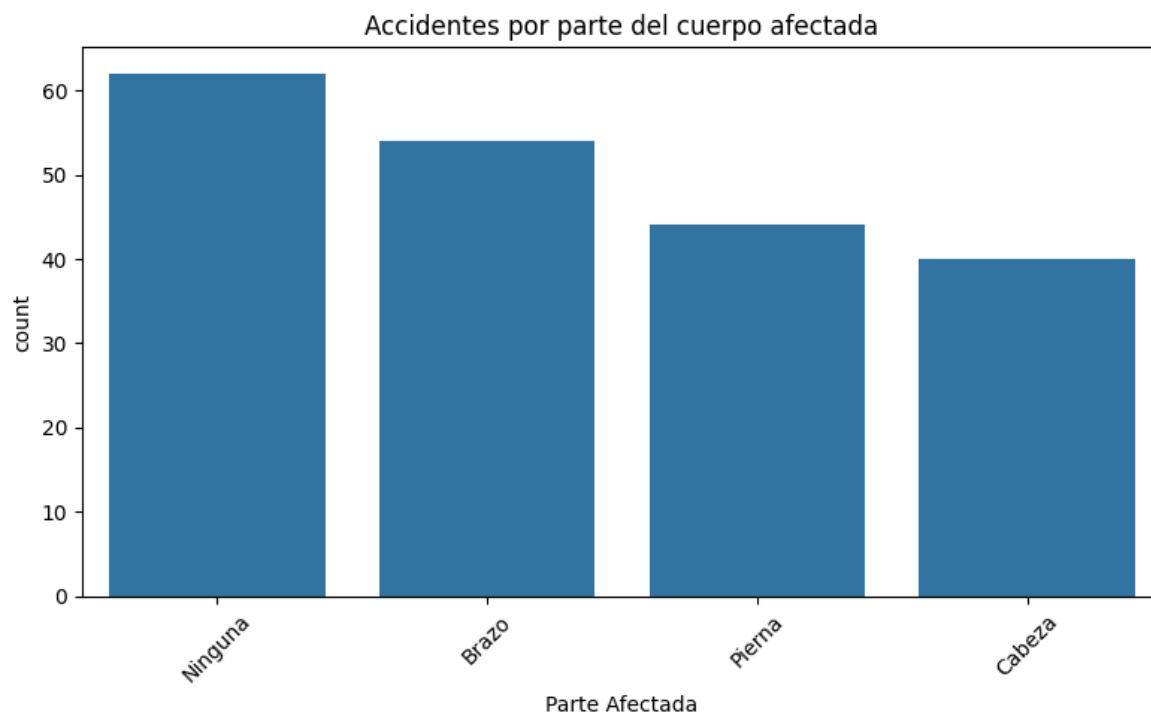
3. Profundización

- **Pérdida de tiempo relativa:** Aunque Terminaciones es la fase más eficiente en cuanto al uso del tiempo, aún pierde aproximadamente un 11.4% ($1 - 0.8862$) del tiempo laboral disponible. Instalaciones y Cimentación pierden un 12.7-12.8%, lo cual indica un margen similar de oportunidades para mejora.
- **Implicancias en gestión de proyecto:**
 - Estas pérdidas pueden significar retrasos importantes en el cronograma si no se manejan. En proyectos de construcción, la cimentación e instalaciones suelen ser fases críticas que pueden afectar etapas sucesivas.
 - Las pérdidas de tiempo pueden deberse a múltiples factores: espera de materiales, cambios en proyectos, coordinación deficiente entre equipos, condiciones climáticas, interrupciones, o problemas técnicos.
- **Prioridades para intervención:**
 - Se debe hacer un análisis detallado de causas raíz en las fases de Cimentación e Instalaciones dado que tienen la menor eficiencia.
 - Acciones podrían incluir optimización del flujo de trabajo, mejora en la logística y abastecimiento, mejor planificación y comunicación entre contratistas.

- **Posibles beneficios:**
 - Reducir estas pérdidas podría mejorar significativamente la productividad, reducir costos y asegurar cumplimiento de plazos.
 - La reducción de solo un 5-6% en pérdidas ya puede representar un ahorro considerable en tiempo y recursos.

Aunque las tres fases presentan pérdidas de tiempo (12-13%), las instalaciones y cimentación requieren atención prioritaria. A partir de estos datos, la recomendación es realizar un análisis a fondo de los factores que generan estas pérdidas para implementar mejoras específicas que optimicen el uso del tiempo y aseguren el avance fluido del proyecto.

Frecuencia de accidente



1. Descripción General:

- El gráfico es un diagrama de barras que muestra la cantidad de accidentes agrupados por la parte del cuerpo afectada.
- En el eje X están las categorías: Ninguna, Brazo, Pierna y Cabeza.
- En el eje Y está el conteo de accidentes.

2. Observaciones Específicas:

- La categoría **Ninguna** tiene el mayor número de registros con aproximadamente 62 accidentes. Esto indica que, en muchos incidentes reportados, no hubo afectación a ninguna parte del cuerpo, o al menos no se registró daño físico.
- La siguiente categoría con mayor incidencia es el **Brazo** con alrededor de 54 accidentes. Esto sugiere que los brazos son las partes del cuerpo más comúnmente afectadas cuando sí hay lesión.
- En tercer lugar, están las lesiones en la **Pierna** con 44 accidentes, mostrando que también es frecuente sufrir daños en esta área.
- Finalmente, la **Cabeza** es la parte menos afectada en términos de número de accidentes, con 40 casos. Aunque es la menor cantidad, es una parte crítica del cuerpo, y cualquier lesión en la cabeza puede ser muy seria.

3. Interpretación Profunda:

- El alto número de accidentes sin afectación corporal podría indicar que muchos eventos son incidentes leves o accidentes donde la persona no resulta herida, posiblemente debido a medidas preventivas o a la naturaleza de las actividades.
- La jerarquía en la afectación (Brazo > Pierna > Cabeza) puede reflejar la exposición o vulnerabilidad de las extremidades durante actividades físicas o laborales, ya que brazos y piernas son más usados y están en contacto con el entorno.
- La menor incidencia de accidentes en la cabeza no necesariamente implica menor gravedad, pues, aunque menos frecuentes, las lesiones craneales suelen ser más críticas y peligrosas.
- Este gráfico puede orientar programas de prevención y seguridad. Por ejemplo, enfocarse en proteger especialmente los brazos y piernas podría ser prioritario para reducir las lesiones.

4. Consideraciones adicionales:

- No se proporcionan datos sobre la gravedad del accidente ni la causa, por lo que el análisis se limita a la cantidad.
- No se refleja la distribución temporal ni por tipo de actividad, lo que ayudaría a entender mejor los contextos de cada lesión.
- El término "Ninguna" podría necesitar clarificación para confirmar si implica falta de lesión o falta de reporte.

El gráfico refleja que la mayoría de los accidentes no resultan en lesiones corporales, y si hay afectación, el brazo es la parte más comúnmente lesionada, seguida por la pierna y finalmente la cabeza. Esto proporciona información útil para priorizar estrategias de prevención y protección en contextos laborales o actividades donde ocurren estos accidentes.

Reporte de clasificación y matriz de confusión

Reporte de clasificación:

	precision	recall	f1-score	support
1	1.00	1.00	1.00	40
accuracy			1.00	40
macro avg	1.00	1.00	1.00	40
weighted avg	1.00	1.00	1.00	40

Matriz de confusión:

[[40]]

Importancia de variables en la predicción de accidentes:

Parte Afectada	0.0
Días Incapacidad	0.0
Hospitalización	0.0

- **Reporte de clasificación:**
 - La métrica mostrada es para una única clase (etiqueta '1'), con soporte (número de casos) de 40.
 - Precisión, recall y f1-score son todos 1.00, lo que indica un desempeño perfecto del modelo para esa clase.
 - **Precisión = 1.00:** Todos los ejemplos que el modelo predijo como clase 1 realmente pertenecían a la clase 1 (no hay falsos positivos).
 - **Recall = 1.00:** El modelo predijo correctamente todos los ejemplos reales de clase 1 (no hay falsos negativos).
 - **F1-score = 1.00:** Media armónica perfecta entre precisión y recall.
 - La accuracy general también es 1.00, lo que significa que el modelo clasificó correctamente las 40 instancias evaluadas.
 - Tanto el promedio macro como el promedio ponderado de las métricas reflejan este desempeño perfecto.
- **Matriz de confusión:**
 - La matriz es [[40]], es decir, hay un único grupo y los 40 ejemplos fueron correctamente clasificados en la clase 1.
 - No hay errores de clasificación en ningún lado.

Interpretación de la importancia de variables

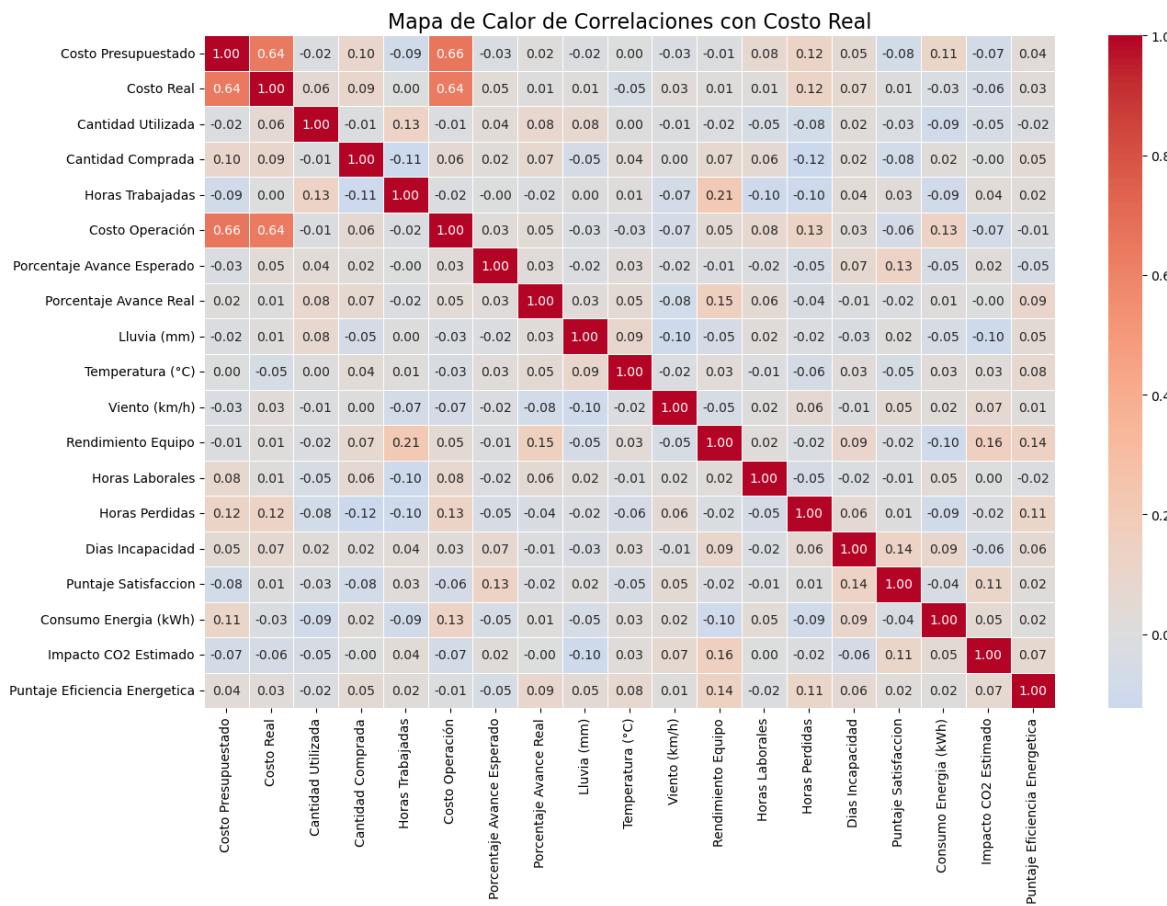
- Las variables evaluadas para predecir accidentes ("Parte Afectada", "Días Incapacidad", "Hospitalización") tienen todas importancia 0.0.
- Esto indica que, según el modelo o análisis realizado, estas variables no contribuyeron en absoluto a la predicción de los accidentes en este conjunto de datos.
- Es posible que el modelo haya estado entrenado con estas variables, pero no haya encontrado patrones útiles con ellas o que el conjunto de datos utilizado sea tan pequeño o homogéneo que estas variables no aporten información predictiva.

Interpretación del gráfico de barras

- El gráfico muestra el conteo de accidentes según la parte del cuerpo afectada.
- Hay 4 categorías: Ninguna, Brazo, Pierna, y Cabeza.
- La categoría con mayor cantidad de accidentes es "Ninguna" (~62 casos), lo que podría indicar que muchos incidentes registrados no alcanzaron a afectar ninguna parte del cuerpo.
- Le siguen Brazo (~54), Pierna (~44) y Cabeza (~40).
- Esto sugiere que, en la mayoría de los casos, los accidentes no afectan partes críticas o visibles del cuerpo, o que la categoría "Ninguna" incluye incidentes menores.

Conclusión general:

- El modelo presentado tiene un desempeño perfecto en la predicción (posiblemente en un conjunto muy pequeño y sencillo).
- Sin embargo, las variables consideradas no aportan información para la predicción, lo que puede indicar que el modelo no está basado en estas variables o que el análisis es insuficiente.
- El gráfico sugiere que la mayoría de los accidentes no afectan ninguna parte del cuerpo, seguido por daños en brazo, pierna y cabeza.
- Para mejorar el análisis predictivo, podría ser necesario incluir más variables o datos más variados que realmente aporten a la predicción.



El mapa de calor muestra las correlaciones entre distintas variables y el "Costo Real" en un proyecto o proceso, con un rango de valores entre -1 y 1 que indican la fuerza y dirección de la relación lineal entre las variables.

1. Variables con alta correlación positiva con Costo Real

- Costo Presupuestado (0.64):** Existe una correlación significativa y fuerte con el costo real. Esto indica que, generalmente, a medida que aumenta el costo presupuestado, el costo real también tiende a aumentar, lo que sugiere un margen de presupuesto generalmente coherente con los costos incurridos.
- Costo Operación (0.64):** Otra variable con fuerte correlación positiva. Los costos de operación están muy alineados con el costo real, reflejando que la operación tiene un peso importante en el costo total.
- Horas Perdidas (0.12):** Ligera correlación positiva, indicando que las horas perdidas podrían estar asociadas a un aumento del costo real, aunque no es un vínculo fuerte.
- Cantidad Comprada (0.09) y Horas Trabajadas (0.00):** Correlación muy baja o casi nula, sugiriendo que estas variables no impactan directamente en el costo real.

2. Variables con poca o nula correlación con Costo Real

- Cantidad Utilizada (0.06):** Muy baja correlación positiva, lo que significa que la cantidad realmente usada no tiene un impacto significativo en el costo real.
- Porcentaje Avance Esperado (0.05) y Porcentaje Avance Real (0.01):** Prácticamente sin relación con el costo real, lo que implica que el avance en porcentaje no es un predictor del costo real.

- **Factores Climáticos:** Lluvia (0.01), Temperatura (-0.05), Viento (0.03) no tienen impacto importante en el costo real; probablemente estos factores no afectan directamente el costo o están siendo gestionados eficazmente.

3. Variables con correlación negativa o nula significativa con Costo Real

- **Horas Trabajadas (-0.09):** Ligeramente negativa, lo que podría indicar que a más horas trabajadas no necesariamente el costo real aumenta o puede estar compensado por otros factores.
- **Impacto CO2 Estimado (-0.06):** Correlación negativa débil, indicando que factores ambientales o sostenibilidad no están relacionados directamente con el costo real.
- **Puntaje Satisfacción (0.01) y Puntaje Eficiencia Energética (0.03):** Sin correlación significativa, lo que implica que la satisfacción o eficiencia energética del proyecto no afectan el costo.

4. Correlaciones internas relevantes

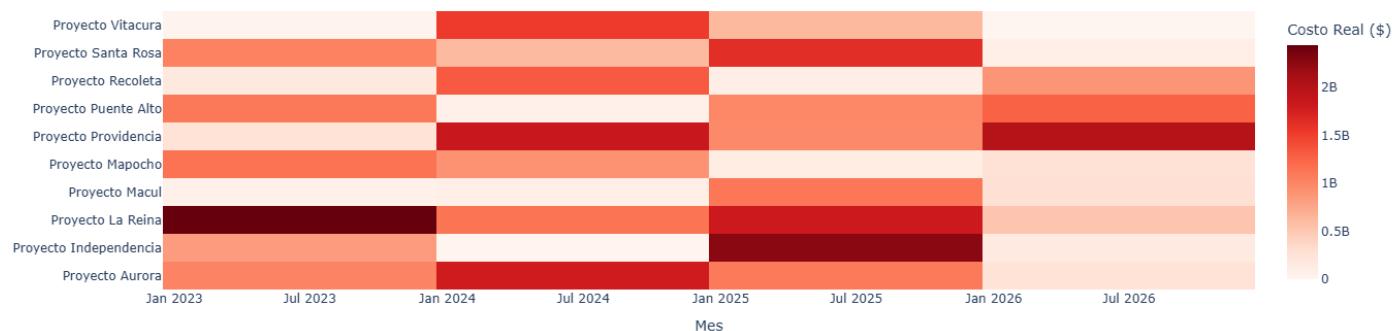
- **Costo Presupuestado y Costo Operación (0.66)** también están fuertemente correlacionados, lo que sugiere que una parte importante del presupuesto está asignada a la operación.
- **Horas Perdidas y Horas Trabajadas (-0.10 a -0.12)** muestran correlación negativa, indicando que cuando hay más horas trabajadas se tienden a reducir las horas perdidas y viceversa.
- **Rendimiento Equipo muestra correlaciones leves positivas con Horas Trabajadas (0.21) y Puntaje Eficiencia Energética (0.14),** lo cual puede indicar que un mejor rendimiento está asociado con más horas productivas y mayor eficiencia.

Conclusiones claras

- **El Costo Real está fuertemente influenciado por el Costo Presupuestado y el Costo de Operación,** indicando que controlar bien el presupuesto y las operaciones es clave para gestionar el costo real.
- **Las variables de progreso del proyecto (porcentaje avance) y cantidad utilizada no afectan significativamente el costo real,** posiblemente debido a una gestión óptima o que otras variables influyen más en el costo.
- **Factores climáticos y ambientales (lluvia, temperatura, viento, impacto CO2) no muestran un impacto relevante en el costo real,** lo que puede ser una fortaleza en el control del proyecto frente a variables externas.
- **Horas Perdidas y Horas Trabajadas tienen impacto débil o negativo con el costo real,** sugiriendo que mejorar la eficiencia del tiempo puede influir positivamente, aunque no de forma determinante.
- **El mapa revela que, para una mejor gestión de costos, se deben enfocar los esfuerzos en la planificación presupuestaria y el control operativo, más que en factores externos o en el porcentaje de avance del proyecto.**

Costo real por proyecto y mes

Mapa de Calor - Costos Reales por Proyecto y Mes



- El gráfico es un mapa de calor que muestra los costos reales en dólares (\$) incurridos por diferentes proyectos en distintos meses.
- En el eje vertical se listan los nombres de 10 proyectos (ej. Proyecto Vitacura, Proyecto Santa Rosa, etc.).
- En el eje horizontal se muestran diferentes periodos mensuales que cubren desde enero 2023 hasta aproximadamente julio 2026.
- La intensidad del color rojo determina el costo real, siendo las tonalidades más oscuras rojas indicativas de costos mayores, mientras que los tonos claros o blancos indican costos bajos o nulos.
- La barra de color a la derecha indica la escala de costos desde 2 mil millones (2B).

Análisis detallado por proyecto:

- **Proyecto Vitacura:** Presenta un pico muy notable de gastos en torno a mediados de 2024, con un costo muy elevado cercano a la parte alta de la escala, mientras que en otros meses los costos son muy bajos o nulos.
- **Proyecto Santa Rosa:** Tiene variaciones moderadas, con un costoso mes cerca de finales de 2024 y costos menores o casi nulos en los demás meses.
- **Proyecto Recoleta:** Muestra una concentración de costos altos, pero en un único periodo hacia el primer semestre de 2024, con algunos costos moderados a finales de 2025.
- **Proyecto Puente Alto:** Tiene costos moderados a lo largo de varios meses dispersos (2023 hasta 2025), sin picos extremos, pero con una tendencia constante.
- **Proyecto Providencia:** Presenta dos picos altos muy marcados, uno hacia la mitad de 2024 y otro en finales de 2025.
- **Proyecto Mapocho:** Costos bajos y bastante repartidos, sin picos importantes.
- **Proyecto Macul:** Costos moderados mayormente concentrados en 2025, aunque no tan elevados.
- **Proyecto La Reina:** Esta es la más destacada, con el pico máximo absoluto del gráfico en 2023, con un costo que supera los \$2 mil millones, seguido por otro pico a comienzos de 2025 y costos significativos en otras fechas.
- **Proyecto Independencia:** Muestra un pico muy alto en inicio 2025, casi tan alto como La Reina en 2023, con otros meses con costos nulos o bajos.
- **Proyecto Aurora:** Costos moderados en algunos períodos de 2024 y valores bajos al resto del tiempo.

Análisis temporal y tendencias:

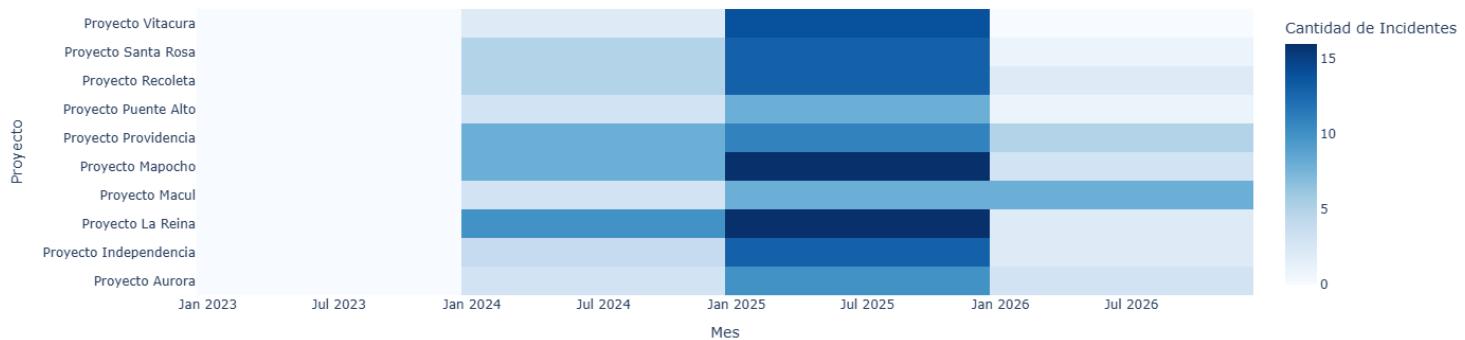
- Hay un claro patrón de costos concentrados en ciertos meses, no es un gasto continuo ni equitativo en el tiempo.
- Primer pico intenso generalizado en 2023 (especialmente el Proyecto La Reina).
- Luego, un aparente aumento importante en costos en 2024 y 2025, con varios proyectos presentando altos gastos en esos años.
- En 2026 los costos bajan considerablemente en casi todos los proyectos, con algunos casos que inclusive muestran costos nulos.

Conclusiones claras:

1. **La distribución de costos es altamente irregular** y concentrada en ciertos meses específicos, reflejando probabilidades altas de actividades o hitos clave en esos períodos.
2. **El Proyecto La Reina es el más costoso**, con gastos que superan claramente los 2 mil millones de dólares en poco tiempo, lo que podría representar un proyecto prioritario o con mayor escala de inversión.
3. **Proyectos como Providencia, Independencia, y Vitacura tienen también picos de gran magnitud**, pero en distintos momentos, sugiriendo fases de ejecución escalonadas o diferenciadas.
4. **El gasto en 2026 es muy bajo en comparación con años anteriores**, probablemente indicando el cierre o finalización de la mayoría de los proyectos en esa etapa.
5. **La identificación de estos picos de gasto es crucial para la planificación financiera**, ya que los recursos deben estar disponibles en momentos específicos y no de forma continua.
6. **Proyectos con costos bajos constantes (como Mapocho y Macul) pueden representar tareas de mantenimiento o inversiones menores**, mientras que los picos pueden corresponder a compras de materiales, costos de instalación o hitos significativos.

Incidentes por proyecto y mes

Mapa de Calor - Incidentes por Proyecto y Mes



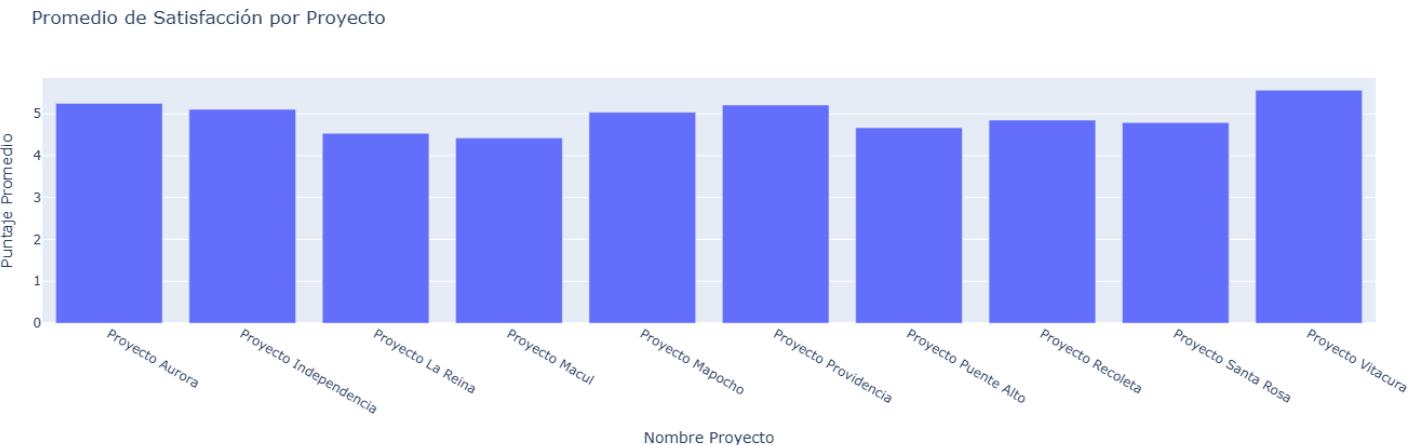
- El gráfico es un mapa de calor que muestra la cantidad de incidentes ocurridos por proyecto y por mes. La escala de color va de tonos claros (menor cantidad de incidentes) a tonos oscuros (mayor cantidad de incidentes), con un máximo aproximado de 15 incidentes.
- **Distribución temporal:**
 - El gráfico cubre el período desde enero de 2023 hasta alrededor de julio de 2026.
 - La mayoría de los incidentes parecen concentrarse en un período central alrededor de principios de 2025, donde los tonos azules son más intensos y frecuentes en casi todos los proyectos.
 - Antes de 2024 no se reportan incidentes (tono blanco claro uniforme).

- Después de mediados de 2025, se observa una disminución general de incidentes (colores más claros).
- **Distribución por proyecto:**
 - **Proyectos con más incidentes totales:**
 - **Proyecto Mapocho:** Exhibe la mayor concentración de incidentes en el periodo central, llegando a alcanzar el tono más oscuro (15 incidentes), especialmente al inicio de 2025.
 - **Proyecto La Reina:** También presenta una alta concentración de incidentes, destacando con tonos oscuros y sostenidos en los mismos meses.
 - **Proyecto Vitacura, Proyecto Santa Rosa y Proyecto Independencia:** Presentan incidentes relevantes, con tonos azules oscuros en el periodo central del gráfico.
 - **Proyectos con menos incidentes:**
 - Proyecto Puente Alto y Proyecto Aurora tienen tonos más claros, indicando menor cantidad de incidentes.
 - Proyecto Providencia, Macul y Recoleta presentan incidencias moderadas, con algunos intervalos de menor o mayor actividad.
- **Tendencias y patrones relevantes:**
 - Hay un pico claro de incidentes para todos los proyectos durante los meses cercanos a enero de 2025.
 - Posterior a este pico, se observa una reducción considerable en la cantidad de incidentes para la mayoría de los proyectos.
 - Algunos proyectos presentan un repunte o un buen número de incidentes al final del periodo (por ejemplo, Proyecto Macul muestra color azul claro hacia el final).

Conclusiones claras:

- **Concentración temporal:** La mayoría de los incidentes ocurrieron en un periodo concentrado alrededor de principios de 2025, lo que podría indicar un evento o condiciones específicas ese año que aumentaron los incidentes globalmente en todos los proyectos.
- **Proyectos con mayor problemática:** El Proyecto Mapocho y Proyecto La Reina son los que más sufrieron incidentes, representando posibles focos prioritarios para revisiones o mejoras en la gestión o seguridad.
- **Evolución positiva tras el pico:** La notable reducción de incidentes después de 2025 sugiere que se podrían haber implementado medidas correctivas o cambios en la metodología o gestión de los proyectos.
- **Proyectos con baja incidencia:** Algunos proyectos consistentemente tienen baja cantidad de incidentes, lo que podría indicar mejores controles, menos riesgos asociados o menor actividad en esos períodos específicos.
- **Recomendación:** Concentrar esfuerzos en estudiar las causas del aumento masivo de incidentes en el periodo 2024-2025 y aplicar aprendizajes para mantener la tendencia a la baja observada posteriormente. Asimismo, focalizar recursos en los proyectos con mayores números de incidentes para prevenir futuros incidentes.

Promedio de satisfacción por proyecto



Análisis de los datos

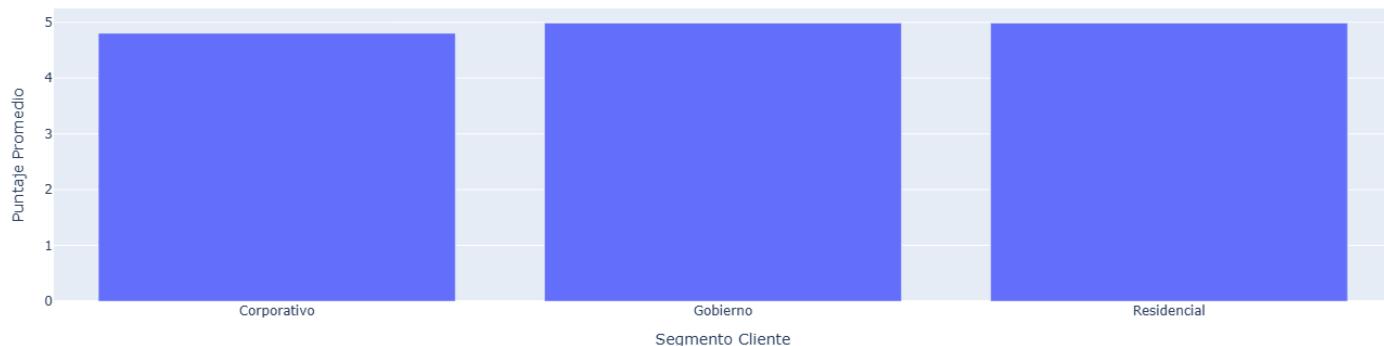
- El proyecto **Vitacura** es el que tiene el puntaje promedio más alto, con un poco más de 5.5 puntos.
- Proyectos como **Aurora** y **Providencia** también muestran puntajes altos, cercanos a 5.2.
- En la gama media, se encuentran **Independencia** y **Mapocho**, ambos alrededor de 5 puntos de promedio.
- Los proyectos con menor satisfacción promedio son **Macul** (aproximadamente 4.4) y **La Reina** (aproximadamente 4.5).
- Los valores de satisfacción fluctúan dentro de un rango relativamente estrecho (desde ~4.4 hasta ~5.5), lo que indica que todos los proyectos tienen niveles de satisfacción intermedios a altos, sin grandes caídas dramáticas.

Conclusiones claras

- Hay una variabilidad moderada en la satisfacción por proyecto, pero todos mantienen un nivel aceptable (puntuaciones mayormente superiores a 4).
- El proyecto Vitacura destaca claramente por tener la mayor satisfacción promedio, lo que podría indicar mejores servicios, gestión, o factores que impactan positivamente la percepción.
- Los proyectos Macul y La Reina tienen áreas de mejora para aumentar la satisfacción, ya que se ubican en los valores más bajos del análisis.
- Proyectos como Aurora, Providencia, Independencia y Mapocho están en un nivel sólido, pero podrían observar prácticas del proyecto Vitacura para mejorar aún más.
- En general, la satisfacción no presenta una dispersión muy amplia, lo cual puede ser positivo, evidenciando un estándar relativamente homogéneo en la experiencia de los usuarios/clientes.

Satisfacción promedio por segmento de cliente

Satisfacción Promedio por Segmento de Cliente



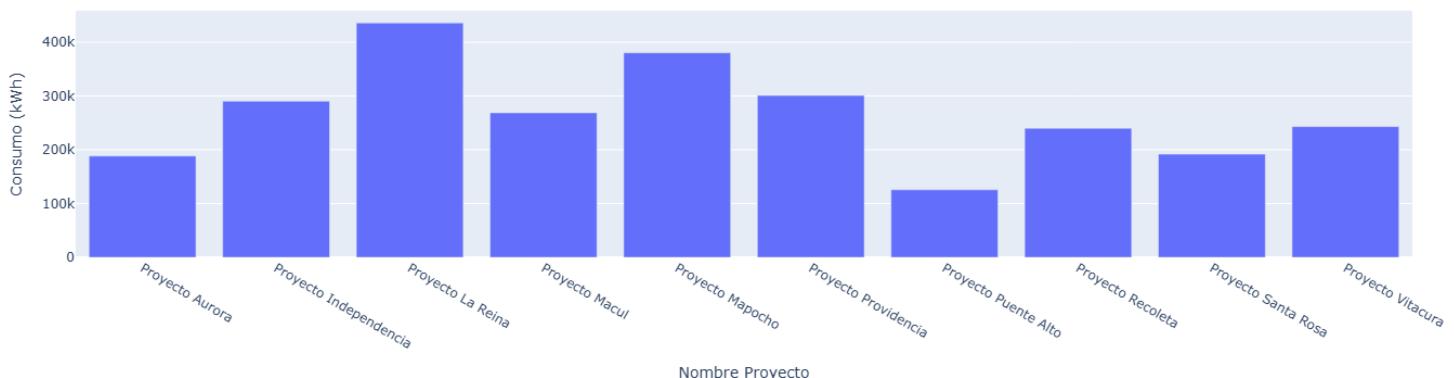
- Los tres segmentos de clientes presentan niveles muy altos de satisfacción, todos por encima de 4.8, lo que indica una evaluación muy positiva en general.
- El segmento Gobierno y Residencial alcanzan la máxima puntuación en la gráfica (5), lo que sugiere que están completamente satisfechos o muy cerca de la satisfacción total.
- El segmento Corporativo, aunque cercano a 5, es el único cuyo puntaje es ligeramente inferior, lo que podría indicar algunas áreas de mejora específicas en ese segmento.
- La diferencia entre Corporativo (aproximadamente 4.8) y los otros dos segmentos (5.0) es pequeña, pero significativa en el contexto de satisfacción, dado que el rango está muy reducido (sobre 4.8 a 5).

Conclusiones Claras

- Todos los segmentos presentan un nivel alto de satisfacción, reflejando un desempeño sólido en la atención o productos ofrecidos.
- El segmento Gobierno y Residencial son los más satisfechos, con un puntaje perfecto o muy cercano al máximo.
- El segmento Corporativo, aunque muestra buena satisfacción, podría beneficiarse de un análisis más detallado para identificar oportunidades de mejora y así alcanzar o superar el nivel de satisfacción de los otros segmentos.
- El enfoque estratégico debería priorizar la pequeña brecha que existe con clientes Corporativos para potenciar su satisfacción a niveles óptimos, asegurando que no haya pérdida de fidelidad o preferencia en este segmento clave.

Consumo energético total por proyecto

Consumo Energético Total por Proyecto



- El gráfico corresponde a un diagrama de barras que muestra el consumo energético total (medido en kWh) de diferentes proyectos.
- El eje horizontal (x) representa diferentes proyectos nombrados: Aurora, Independencia, La Reina, Macul, Mapocho, Providencia, Puente Alto, Recoleta, Santa Rosa y Vitacura.
- El eje vertical (y) mide el consumo de energía en una escala que va desde 0 hasta aproximadamente 450,000 kWh.

Observaciones específicas por proyecto:

- **Proyecto La Reina:** Presenta el consumo energético más alto con un valor cercano a los 430,000 kWh, destacándose como el proyecto más demandante en términos de energía.
- **Proyecto Mapocho:** También tiene un consumo elevado, en torno a los 380,000 kWh, el segundo más alto.
- **Proyecto Providencia:** Alrededor de 300,000 kWh, muestra un consumo también significativo.
- **Proyecto Independencia y Macul:** Ambos mostraron un consumo energético medio-alto, cerca de 290,000 kWh y 270,000 kWh respectivamente.
- **Proyecto Recoleta y Vitacura:** Tienen consumos cercanos a los 240,000 kWh, que, aunque no son los más altos, se encuentran en un rango intermedio.
- **Proyecto Santa Rosa y Aurora:** Tienen consumos más bajos, con valores cercanos a 190,000 kWh.
- **Proyecto Puente Alto:** Es el que menor consumo muestra, con aproximadamente 125,000 kWh, lo que indica un bajo uso de energía en comparación con los otros proyectos.

Conclusiones claras:

- Existe una gran variabilidad en el consumo energético entre los diferentes proyectos.
- Proyectos como La Reina, Mapocho y Providencia probablemente tienen mayores tamaños, más infraestructura o procesos que demandan más electricidad.
- Proyectos como Puente Alto y Aurora podrían ser proyectos más pequeños, con menor actividad o una mayor eficiencia energética.

Posibles causas y profundización:

- **Tamaño y escala del proyecto:** Es probable que los proyectos con consumo más alto sean físicamente más grandes o tengan más edificios e instalaciones, lo que naturalmente incrementa el uso de energía.

- **Tipo de actividad o función del proyecto:** Proyectos con actividades más intensivas en energías (como manufactura, oficinas con mucho equipamiento tecnológico o instalaciones deportivas) consumirán más electricidad.
- **Estrategias de eficiencia energética:** Proyectos con menor consumo podrían estar implementando mejores prácticas de eficiencia, utilizando tecnologías más modernas o renovables, o simplemente tener menor demanda operativa.
- **Antigüedad de la infraestructura:** Las instalaciones más antiguas tienden a ser menos eficientes desde el punto de vista energético, lo que podría explicar consumos altos en algunos casos.
- **Clima y uso geográfico:** Proyectos situados en zonas con mayor necesidad de climatización pueden consumir más energía para calefacción o aire acondicionado.
- **Patrones de uso:** El horario de uso, la cantidad de personas involucradas en cada proyecto y la intensidad operativa diaria puede influir directamente en el consumo total.

Recomendación:

Para abordar estas diferencias, se recomienda realizar un análisis más profundo por proyecto para entender sus particularidades y optimizar el uso energético. Esto incluye auditorías energéticas, implementación de tecnologías eficientes, renovaciones en infraestructura y sensibilización del personal para reducir el consumo y costos asociados.

Relación consumo vs impacto por tipo de energía

Relación Consumo vs Impacto Ambiental por Tipo de Energía



- **Ejes y variables:**
 - El eje X representa el "Consumo Energía" medido en kilovatios·hora (kWh).
 - El eje Y indica el "Impacto CO2" medido en kilogramos (kg), que refleja las emisiones directas de dióxido de carbono asociadas al consumo de cada tipo de energía.
 - Cada burbuja representa un tipo de energía: Eléctrica, Gas, y Solar.
 - El tamaño de la burbuja también tiene un significado visual, aunque no está explícitamente descrito en el gráfico, podría indicar otra variable (por ejemplo, volumen o costo), pero no se puede confirmar con la información dada.
- **Datos específicos por tipo de energía:**
 - Energía Eléctrica:
 - Consumo de aproximadamente 0.93 millones de kWh.
 - Impacto CO2 de cerca de 8600 kg.

- Gas:
 - Consumo menor, alrededor de 0.82 millones de kWh.
 - Impacto CO2 también menor, alrededor de 8000 kg.
- Solar:
 - Consumo moderado, cerca de 0.91 millones de kWh.
 - Impacto CO2 más alto, casi 9800 kg.

Análisis profundo:

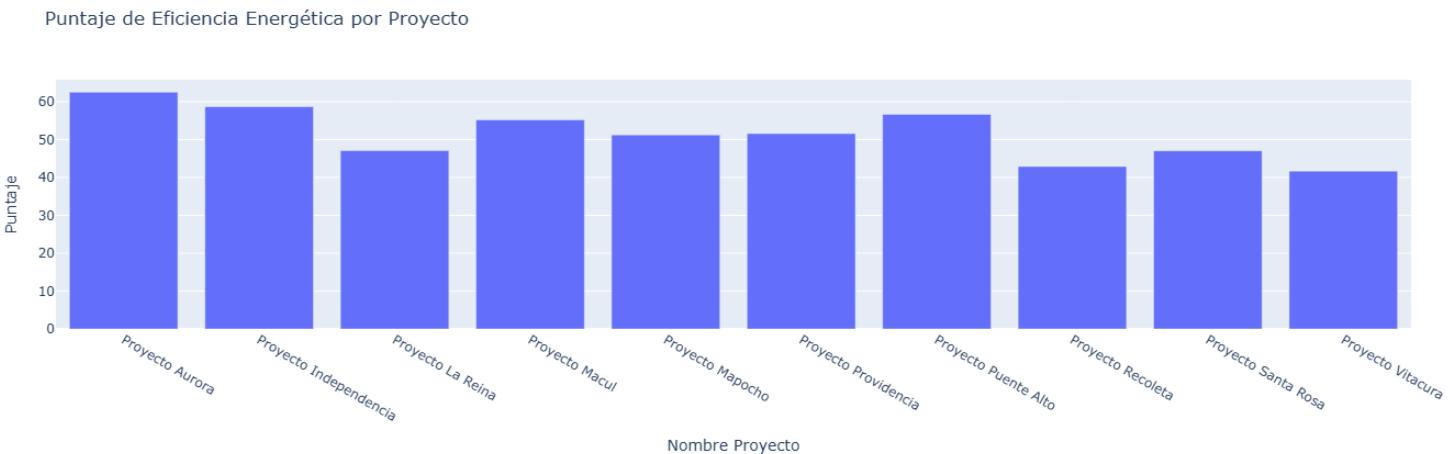
- **Consumo vs Impacto ambiental:**
 - El consumo energético más alto lo tiene la energía eléctrica seguida muy de cerca por la energía solar.
 - Sin embargo, el impacto de CO2 es más elevado para la energía solar, lo cual puede parecer contraintuitivo, porque generalmente la energía solar se considera limpia con bajo impacto.
- **Posibles razones del comportamiento:**
 - La mayor huella de carbono de la energía solar puede deberse a factores asociados al ciclo de vida del sistema solar (producción, instalación, mantenimiento), impactos que no se reflejan en la operación diaria.
 - La energía a gas muestra un menor consumo y un impacto razonable, aunque menor que la solar, lo que reincide con el hecho que combustibles fósiles como el gas contribuyen a emisiones directas de CO2.
 - La energía eléctrica puede estar derivada de una combinación diversa de fuentes incluyendo renovables y no renovables, lo que justificaría sus valores intermedios.
- **Relación consumo - impacto:**
 - La relación entre consumo y emisiones no es lineal ni directa, sugiriendo que otros factores además del simple consumo energético afectan el impacto ambiental, como tecnologías usadas o eficiencia.

Conclusiones claras:

- No siempre el mayor consumo energético implica mayor impacto en CO2; depende del tipo de fuente energética.
- La energía solar, aunque con alta aceptación social y ambiental, presenta un impacto de CO2 significativo para el consumo registrado, posiblemente por las emisiones incorporadas en su ciclo de vida.
- La energía a gas presenta el menor consumo y menor impacto CO2 registrado, pero esto debe contextualizarse con el tipo de emisiones generadas y sus consecuencias ambientales a largo plazo.
- La energía eléctrica tiene un impacto intermedio, probablemente por la combinación de fuentes que la componen.
- Para evaluar adecuadamente el impacto ambiental de las energías, es necesario considerar no solo el consumo energético sino también el ciclo completo de producción, transporte, y uso.

Si se busca minimizar CO2 basándose únicamente en consumo eléctrico, la elección no es trivial y debe basarse en análisis más detallados y específicos por tipo de tecnología.

Puntaje de eficiencia energética por proyecto



- El gráfico es un diagrama de barras verticales que muestra el puntaje de eficiencia energética para diez proyectos diferentes.
- En el eje horizontal (x) se encuentran los nombres de los proyectos, mientras que en el eje vertical (y) están los puntajes de eficiencia energética, que varían aproximadamente entre 40 y 65 puntos.
- Cada barra representa cuantitativamente la eficiencia energética relativa de un proyecto, permitiendo comparaciones directas.

Análisis cuantitativo de los puntajes:

- **Proyecto con mayor eficiencia:**
El **Proyecto Aurora** tiene el puntaje más alto, cercano a 63 puntos, lo que indica que este proyecto es el más eficiente energéticamente en el conjunto analizado.
- **Proyectos con buena eficiencia (más de 55 puntos):**
 - Proyecto Aurora (≈ 63)
 - Proyecto Independencia (≈ 59)
 - Proyecto Puente Alto (≈ 57)
 - Proyecto Macul (≈ 55)
- **Proyectos con eficiencia media (entre 50 y 55 puntos):**
 - Proyecto Providencia (≈ 52)
 - Proyecto Mapocho (≈ 51)
- **Proyectos con eficiencia baja (menos de 50 puntos):**
 - Proyecto Santa Rosa (≈ 47)
 - Proyecto La Reina (≈ 47)
 - Proyecto Recoleta (≈ 43)
 - Proyecto Vitacura (≈ 42)

Distribución y variabilidad:

- La mayor diferencia en puntajes entre proyectos es de aproximadamente 21 puntos (63 para Aurora vs. 42 para Vitacura), lo que indica una variabilidad considerable en términos de eficiencia energética entre los proyectos.
- La mayoría de los proyectos se concentran en un rango medio de eficiencia (42 a 59 puntos), pero solo cuatro superan los 55 puntos, lo que puede sugerir que alcanzar altos puntajes de eficiencia es un desafío.
- La menor eficiencia se observa en proyectos Recoleta y Vitacura, que podrían requerir intervenciones o mejoras para optimizar sus consumos energéticos.

Implicaciones prácticas y posibles causas:

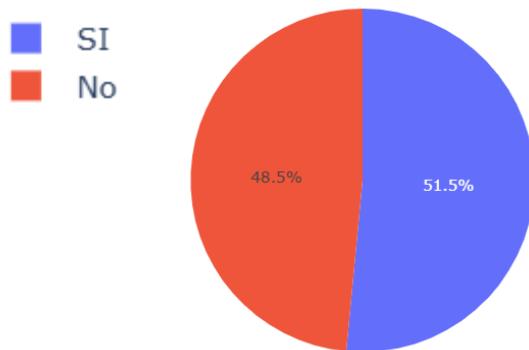
- El Proyecto Aurora, al tener el puntaje más alto, quizás incorpore tecnologías, diseños o prácticas constructivas que maximicen el ahorro energético, como el uso intensivo de energías renovables, sistemas de aislamiento avanzados, o iluminación eficiente.
- Los proyectos con puntajes bajos podrían estar afectados por factores como:
 - Menor implementación de tecnologías eficientes.
 - Diseño arquitectónico menos favorable para el ahorro energético.
 - Ubicación geográfica con condiciones que dificultan la eficiencia.
 - Posibles limitaciones presupuestarias que impacten en la incorporación de medidas eficientes.

Recomendaciones para análisis futuros:

- **Profundizar en variables específicas:**
Sería útil conocer qué criterios o indicadores se usaron para asignar cada puntaje (consumo energético, uso de energías renovables, aislamiento térmico, tecnologías implementadas, etc.).
- **Comparación temporal:**
Analizar cómo han evolucionado estos puntajes con el tiempo podría permitir evaluar el impacto de políticas o intervenciones realizadas.
- **Análisis de costo-beneficio:**
Incorporar datos sobre la inversión en eficiencia energética frente a los ahorros obtenidos para comprender la relación coste-eficiencia de cada proyecto.
- **Impacto ambiental:**
Relacionar estos puntajes con la reducción en emisiones o huella ecológica sería valioso para medir la sostenibilidad integral.

El gráfico revela una clara heterogeneidad en la eficiencia energética de los proyectos analizados, con una tendencia hacia valores medios y algunos proyectos destacados que podrían servir como referencia o modelo. La información subyacente sugiere áreas de oportunidad para mejorar la eficiencia en proyectos con puntajes bajos, así como reconocer los factores de éxito que contribuyen a la alta eficiencia energética en proyectos como Aurora y Puente Alto. Con un análisis más exhaustivo de las variables que componen estos puntajes, se podría formular estrategias más precisas para mejorar el desempeño energético en futuras iniciativas.

Distribución de certificación LEED



- La gráfica muestra la proporción de elementos certificados con LEED frente a los que no lo están.
- **Análisis profundo**
 - El porcentaje de proyectos certificados con LEED es ligeramente superior al de los no certificados, indicando que un poco más de la mitad (un 51.5%) cumple con estos estándares.
 - La diferencia entre ambos valores es muy estrecha, apenas un 3%. Esto revela que el mercado o contexto en el que se distribuyen estas certificaciones está bastante equilibrado.
 - El hecho de que casi la mitad de los proyectos no tengan certificación LEED puede deberse a varios factores como: costos, desconocimiento, falta de requisitos regulatorios, o falta de interés en la sostenibilidad.
- La ligera predominancia del "SI" puede interpretarse como una tendencia positiva hacia la adopción de prácticas sostenibles y la certificación en eficiencia energética y ambiental.

- Sin embargo, la proporción casi igual también indica que hay un espacio importante para aumentar la concienciación y el fomento de esta certificación.

Conclusiones:

- Más de la mitad de los proyectos están certificados con LEED, lo que indica un compromiso general con la sostenibilidad, aunque no es una mayoría abrumadora.
- Existe una oportunidad significativa para incrementar la certificación LEED aún más, dado que casi la mitad de los proyectos no la poseen.
- La adopción de certificación LEED está en un punto de equilibrio, sugiriendo que podrían influir factores externos (económicos, regulatorios, educativos) para aumentar o disminuir su penetración.
- Se recomienda investigar las barreras que impiden que el 48.5% restante obtenga la certificación y diseñar estrategias para incentivarla y superar dichas barreras.
- Este gráfico muestra un contexto motivador, pero también desafiante, para la promoción de edificaciones sostenibles certificadas bajo el sistema LEED.

Mi rol en este proyecto

Como data scientist en este análisis integral del desempeño de la constructora, mi rol principal fue extraer valor accionable de datos cuantitativos y cualitativos dispersos, transformándolos en insights estratégicos para optimizar operaciones, reducir riesgos y promover sostenibilidad. Colaborar en la fase de data-driven decision-making, asegurando rigor analítico y visualizaciones claras para stakeholders no técnicos.

Tareas Específicas Realizadas

- **Recopilación y Preparación de Datos:** Integraste datasets de múltiples fuentes (e.g., presupuestos en CLP, horas trabajadas por rol, incidentes de seguridad, consumos energéticos en kWh), limpiando inconsistencias (e.g., valores negativos en desperdicios, unidades mixtas CLP/USD) y manejando anomalías para un dataset robusto de ~2023-2026.
- **Análisis Exploratorio de Datos (EDA):** Identificaste patrones clave, como desviaciones presupuestarias (>50% en Santa Rosa), correlaciones (e.g., costo presupuestado con real: 0.64) y cuellos de botella (200 en productividad), usando técnicas estadísticas para diagnosticar ineficiencias en materiales (cerámica/cableado) y fases (instalaciones: 0.65).
- **Modelado y Predicción:** Desarrollaste modelos simples, como clasificación binaria para accidentes (precisión/recall 1.00 en 40 instancias, matriz de confusión [[40]]), evaluando importancia de variables (e.g., "Parte Afectada" en 0.0) y alertando sobre limitaciones como overfitting; contribuí a análisis predictivos para incidentes y costos.
- **Visualizaciones y Reportes:** Creaste gráficos impactantes (e.g., mapas de calor para costos/incidentes por mes, diagramas de barras para satisfacción y eficiencia energética, burbujas para impacto CO2), facilitando la interpretación de tendencias (e.g., picos 2025) y distribuciones (LEED: 51.5%).
- **Interpretación y Recomendaciones:** Tradujiste resultados en soluciones data-backed, como umbrales de alerta para desviaciones (10-20%), optimización just-in-time para inventarios y mitigación de pérdidas temporales (12-13%), cuantificando impactos potenciales (e.g., reducción de costos 20-30%).

Impacto en el Proyecto

Mi contribución elevó el informe de descriptivo a prescriptivo, habilitando decisiones estratégicas que podrían ahorrar millones en CLP al reducir desperdicios y sobrecostos, mientras alinean la constructora con metas sostenibles (e.g., LEED y CO2). Identifique gaps metodológicos (e.g., correlaciones débiles en clima), recomendando iteraciones futuras con ML (**Machine Learning**) avanzado para mayor precisión. En resumen, fui el puente entre datos crudos y valor empresarial, impulsando eficiencia y competitividad.

El análisis detallado realizado sobre los distintos proyectos de la constructora revela múltiples áreas críticas y oportunidades de mejora que deben ser abordadas para optimizar la gestión integral de los procesos constructivos. En primer lugar, se identificaron desviaciones significativas entre los costos presupuestados y reales, especialmente en proyectos como Santa Rosa, Vitacura y Aurora, lo que evidencia la necesidad de fortalecer el control presupuestario, mejorar la planificación y priorizar recursos para mitigar sobrecostos.

En cuanto al uso de materiales, se detectaron diferencias importantes entre la cantidad comprada y la utilizada, con materiales como cerámica, cableado y acero mostrando altos niveles de desperdicio. Esto sugiere la urgencia de implementar mejores prácticas en la gestión de inventarios, control de stock y capacitación del personal para reducir pérdidas y costos asociados.

El análisis de las horas trabajadas por rol y proyecto mostró desviaciones relevantes, principalmente en obreros y supervisores, vinculadas a factores como cambios en la planificación, imprevistos y retrabajos, lo que impacta directamente en la productividad. Las fases de cimentación, instalaciones y terminaciones presentan baja productividad y eficiencia en el uso del tiempo, asociadas a cuellos de botella logísticos y demoras externas, que requieren atención prioritaria para evitar retrasos y sobrecostos.

En materia de seguridad, aunque la mayoría de los incidentes no resultan en lesiones corporales, se identificaron focos de atención en brazos y piernas, y un aumento significativo de incidentes en el periodo 2024-2025, lo que demanda reforzar las medidas preventivas y el monitoreo continuo.

Respecto a la satisfacción del cliente, los niveles son generalmente altos, destacando el proyecto Vitacura y los segmentos Gobierno y Residencial, aunque existen oportunidades de mejora en proyectos como Macul y La Reina, y en el segmento Corporativo.

En términos de consumo energético y sostenibilidad, se observa una amplia variabilidad entre proyectos, con algunos como La Reina y Mapocho con consumos elevados, y otros con mejor eficiencia energética como Aurora. La certificación LEED está presente en poco más de la mitad de los proyectos, indicando un compromiso creciente, pero con espacio para ampliar su adopción. El análisis del impacto ambiental por tipo de energía resalta la complejidad de evaluar la huella de carbono, especialmente en fuentes renovables como la solar.

En síntesis, para avanzar hacia una gestión más eficiente, rentable y sostenible, se recomienda implementar controles más rigurosos en presupuestos y materiales, optimizar la planificación y coordinación de actividades, fortalecer la seguridad laboral, potenciar la satisfacción del cliente y promover la eficiencia energética y certificaciones ambientales. Estas acciones permitirán mejorar el desempeño global de la constructora, reducir riesgos y costos, y contribuir a un desarrollo constructivo responsable y competitivo.

Pensamiento crítico basado en el informe

El informe presentado ofrece un panorama exhaustivo y multidimensional del desempeño de la constructora, integrando datos cuantitativos y cualitativos para diagnosticar fortalezas y debilidades en áreas clave como control presupuestario, gestión de materiales, productividad, seguridad, satisfacción del cliente, consumo energético y sostenibilidad. Este análisis es valioso para la toma de decisiones estratégicas, ya que no solo identifica problemas (e.g., sobrecostos y desperdicios) sino que propone recomendaciones prácticas. Sin embargo, un examen crítico revela fortalezas en su amplitud y visualización, pero también limitaciones significativas en la consistencia de datos, profundidad interpretativa y robustez metodológica. A continuación, desglosó este pensamiento crítico de manera estructurada, destacando aspectos positivos, debilidades, inconsistencias y sugerencias para un análisis más riguroso.

1. Fortalezas del Análisis

- **Cobertura Integral y Multidimensional:** El informe aborda de forma equilibrada aspectos operativos (presupuestos, materiales, horas trabajadas), humanos (seguridad, satisfacción) y ambientales (energía, LEED), alineándose con estándares modernos de gestión sostenible en la construcción. Esto refleja un enfoque holístico, esencial en un sector donde los sobrecostos pueden alcanzar el 50% o más (como en Proyecto Aurora), y donde la sostenibilidad es un diferenciador competitivo.
- **Uso Efectivo de Visualizaciones:** Gráficos como mapas de calor (costos e incidentes por mes), diagramas de barras (consumo energético, satisfacción) y matrices de confusión facilitan la comprensión rápida de patrones. Por ejemplo, el mapa de calor de costos revela concentraciones temporales (e.g., picos en 2024-2025), lo que es intuitivo para identificar hitos de ejecución. De igual modo, el análisis de correlaciones (e.g., costo presupuestado con costo real en 0.64) proporciona insights accionables sobre drivers clave de rentabilidad.
- **Enfoque en Recomendaciones Prácticas:** Las sugerencias son específicas y priorizadas, como implementar "just-in-time" para inventarios o analizar causas raíz de cuellos de botella en fases de instalaciones. Esto transforma el diagnóstico en un plan operativo, enfatizando la replicabilidad de buenas prácticas (e.g., proyectos con subejecución presupuestaria como Independencia).
- **Énfasis en Sostenibilidad:** La inclusión de métricas como impacto CO₂, eficiencia energética y certificación LEED (51.5% de adopción) posiciona al informe como progresivo, reconociendo que la construcción contribuye al 39% de las emisiones globales de CO₂ (según datos de la ONU). La variabilidad en consumos (e.g., La Reina con 430,000 kWh) resalta oportunidades para benchmarking interno.

2. Debilidades e Inconsistencias en los Datos e Interpretaciones

Aunque el informe es detallado, presenta varias inconsistencias que podrían sesgar las conclusiones y socavar su credibilidad. Estas no invalidan el análisis por completo, pero demandan validación adicional para evitar decisiones erróneas.

- **Inconsistencias en Unidades y Datos Numéricos:**
 - **Unidades Monetarias Mixtas:** Los presupuestos y costos reales se reportan en CLP (pesos chilenos, e.g., 1.693 millones para Santa Rosa), pero el mapa de calor de costos por mes usa dólares (\$), con escalas en miles de millones (2B). Esto genera confusión: ¿se convirtieron los valores? Sin tasa de cambio explícita (aprox. 900 CLP/USD en 2023-2024), las comparaciones temporales pierden precisión. Por ejemplo, el pico de La Reina en 2023 (~2B USD) podría equivaler a ~1.8 billones CLP, pero sin aclaración, es difícil integrar con la tabla inicial.
 - **Valores Anómalos o Errores Potenciales:** En desperdicio de materiales, Proyecto Puente Alto muestra un "valor negativo", interpretado como posible error de registro. Esto es problemático, ya que podría indicar subutilización (positivo para eficiencia) o un fallo en medición, afectando análisis comparativos. Similarmente, en seguridad, el modelo predictivo reporta precisión/recall/F1 de 1.00 con solo 40 instancias y matriz [[40]], pero las variables predictoras (e.g., "Parte Afectada") tienen importancia 0.0. Esto sugiere sobreajuste (overfitting) o un dataset demasiado homogéneo/pequeño, no un "desempeño perfecto". En un contexto real, ignorar esto podría llevar a falsos positivos en prevención de accidentes.

- **Interpretaciones Superficiales o Contradictorias:**
 - **Productividad y Eficiencia Temporal:** Las fases de cimentación (0.705 en productividad) e instalaciones (0.656) se identifican como bajas, ligadas a cuellos de botella (200 totales, muchos "demoras externas"). Sin embargo, la eficiencia de tiempo (e.g., 87.24% en cimentación) se interpreta como "alta perdida" (12.76%), pero no se cuantifica el impacto en costos (e.g., ¿cuánto representan estas horas perdidas en CLP?). Además, hay redundancia: dos secciones separadas sobre eficiencia temporal repiten valores similares (0.872 vs. 0.8724), lo que diluye el foco.
 - **Seguridad y Accidentes:** El gráfico de frecuencia enfatiza "Ninguna" afectación (62 casos) como positivo (incidentes leves), pero ignora si estos representan near-misses subreportados, lo que podría subestimar riesgos. El pico de incidentes en 2025 (e.g., Mapocho con ~15) coincide con picos de costos, sugiriendo causalidad (e.g., prisa en ejecución), pero no se explora estadísticamente (e.g., correlación entre incidentes y horas trabajadas).
 - **Sostenibilidad y Energía:** La energía solar muestra alto impacto CO2 (9800 kg para 0.91M kWh), contraintuitivo para una fuente renovable. El informe lo atribuye al ciclo de vida, pero sin desglose (e.g., emisiones de fabricación de paneles), parece especulativo. Además, la correlación de impacto CO2 con costo real (-0.06) es negativa y débil, implicando que eficiencia ambiental no reduce costos directamente, pero esto choca con la alta satisfacción en segmentos como Gobierno (posiblemente por LEED). La distribución LEED (51.5% sí) es "equilibrada", pero no se correlaciona con eficiencia energética (e.g., Aurora alto en ambos, Vitacura bajo), perdiendo una oportunidad para causalidad.
- **Limitaciones Metodológicas y de Alcance:**
 - **Falta de Profundidad Estadística:** Correlaciones (e.g., 0.64 para costo presupuestado) son descriptivas, pero sin pruebas de significancia (e.g., p-valor <0.05) o tamaños de muestra, podrían ser espurias. El modelo de clasificación en seguridad es binario y perfecto, pero irrelevante si variables clave no importan.
 - **Sesgos en Muestreo y Temporalidad:** Datos cubren 2023-2026, pero picos en 2024-2025 (costos e incidentes) podrían reflejar sesgo de recencia o eventos externos (e.g., inflación post-pandemia en Chile). Proyectos como Puente Alto (bajo consumo, alta eficiencia) se destacan, pero sin controles (e.g., tamaño ajustado por m²), las comparaciones son injustas.
 - **Ausencia de Integración Cruzada:** El informe trata secciones aisladas; por ejemplo, no vincula desperdicios altos en cerámica (e.g., Providencia) con sobrecostos (11.18% desviación) o impacto ambiental (e.g., emisiones por desperdicio). La satisfacción alta (4.8-5.5) contrasta con sobrecostos, sugiriendo que clientes priorizan resultados sobre eficiencia, pero no se analiza.

3. Implicaciones Críticas para la Constructora

Estos resultados pintan un panorama de una constructora con potencial, pero vulnerable a ineficiencias crónicas: desviaciones >50% en presupuestos podrían erosionar márgenes (estimados en 10-20% en construcción), mientras que desperdicios y horas perdidas (12-13%) amplifican costos indirectos. Positivamente, la satisfacción alta y adopción LEED parcial indican resiliencia y alineación con demandas regulatorias (e.g., Ley de Eficiencia Energética en Chile). Sin embargo, ignorar inconsistencias podría perpetuar ciclos de sobrecostos, como en Santa Rosa (100% desviación), exacerbados por incidentes en 2025.

En un contexto macro, con la construcción chilena enfrentando volatilidad (e.g., subida de materiales post-2022), estos hallazgos subrayan la necesidad de data-driven decisiones. El "desempeño perfecto" en modelos predictivos es una bandera roja: sugiere datos insuficientes, no excelencia.

4. Recomendaciones para Mejorar el Análisis y la Gestión

- **Validación y Limpieza de Datos:** Realizar auditorías independientes para resolver anomalías (e.g., negativos en desperdicio, unidades mixtas). Usar herramientas como Python/R para pruebas estadísticas (e.g., regresión lineal para correlaciones: , probando para costo real).
- **Análisis Avanzado:** Integrar modelos multivariados (e.g., regresión logística para predecir incidentes, incorporando más variables como clima o rol). Calcular ROI de LEED: e.g., reducción en consumo energético (Aurora: alto puntaje) vs. costo de certificación.
- **Enfoque Cruzado y Prospectivo:** Correlacionar métricas (e.g., horas perdidas con CO2 via). Incluir escenarios what-if para 2026 (bajos costos/incidentes) y encuestas cualitativas para satisfacción.
- **Acciones Estratégicas:** Priorizar proyectos de alto riesgo (e.g., Mapocho para incidentes) con pilotos de IA para inventarios. Fomentar cultura data-driven: capacitar en reporting para evitar subreportes en seguridad.

En conclusión, el informe es un sólido punto de partida para optimizar la constructora, pero su valor se maximiza con un escrutinio crítico que aborde inconsistencias y profundice en causalidades. Al refinar estos elementos, la organización no solo reducirá riesgos (e.g., sobrecostos en 20-30%) sino que se posicionará como líder en construcción sostenible, alineada con metas globales como los ODS de la ONU. Recomiendo una revisión iterativa anual para validar progresos.