



## Proyectos Análisis de datos S1-2025

Selecciones 3 por orden de preferencia y envíe un correo a [mabvidal@inf.udec.cl](mailto:mabvidal@inf.udec.cl); [annaismolina201@udec.cl](mailto:annaismolina201@udec.cl) y [caquijada@inf.udec.cl](mailto:caquijada@inf.udec.cl)

### 1. Análisis de proteínas entre peces infectados y sanos. (Benjamín Espinoza, Ricardo Charris, Nicolás López, Sofía López)

Este proyecto tiene como objetivo analizar datos de expresión de proteínas en salmón del Atlántico (*Salmo salar*) para comparar individuos infectados con el parásito *Caligus* y aquellos sanos, utilizando técnicas de ciencia de datos y aprendizaje automático. Se espera identificar diferencias significativas en la expresión de proteínas, además de trabajar con bases de datos biológicas para determinar la función a la que pertenecen estas proteínas, seguido de pruebas de hipótesis y modelado predictivo.

### 2. Análisis de accidentes en Chile (Vicente Cuello Diban, Iván Zapata Ávila, Maximiliano López Becerra, Franco Vidal Carvajal)

Carabineros de Chile, reporta los siniestros viales y sus consecuencias mensualmente, sin embargo para este estudio nos interesa analizar si existen diferencias significativas y correlaciones en el tipo de accidentes y perfil de las personas que cometen infracciones, como también sus consecuencias (muertos, graves, m/grave, leves, ilesos). En cuanto a la conducta de la persona se cuenta con un pequeño perfil para saber si era un conductor, pasajero o peatón y sus consecuencias. También se espera explorar si hay una relación con la fecha en que ocurren los accidentes, entre otros.

### 3. Análisis de la carga parasitaria reportada por Sernapesca (Nicolas Pino, Nicolás Jarpa e Ignacio Barría)

Sernapesca, a través de sus informes sanitarios, monitorea las cargas parasitarias en la salmonicultura, especialmente en centros de alta vigilancia, y ha detectado aumentos en algunas regiones y disminuciones en otras, destacando la importancia de la vigilancia y el control de la caligidosis. Se espera analizar además, si variables ambientales influyen en esta carga parasitaria, implementar predicciones, correlaciones, etc.

### 4. Análisis de precios al consumidor de la canasta básica familiar (Jorge Fernandez, Sebastian Avila, Diego Gutierrez y Mirko Peñailillo)

Conjuntos de datos de precios al consumidor de los principales alimentos de la canasta familiar que son capturados en supermercados, ferias libres, carnicerías y panaderías del país, localizados en las regiones de Arica y Parinacota, Coquimbo, Valparaíso, Metropolitana, Maule, Ñuble, Biobío, La

Araucanía y Los Lagos. La información disponible comprende frutas y hortalizas, carnes, legumbres, cereales, lácteos, huevos, pan, entre otros alimentos.

## 5. Detección de brechas y tendencias en Ciencias de la Computación

DBLP computer science bibliography, o simplemente dblp, es una base de datos bibliográfica - no contiene el texto completo de los trabajos - y cuya característica principal es que es de acceso abierto. Su objetivo es cubrir todas las áreas de la computación (desde algoritmos, inteligencia artificial, compiladores, data mining, bioinformática, redes, robótica, seguridad, virtualización, etc), eso sí, desde un enfoque internacional, por lo que se centran en la producción científica en lengua inglesa. El objetivo es detectar las últimas tendencias y brechas que puedan existir en la academia en el área de computación.

## 6. Análisis de la encuesta nacional urbana de seguridad ciudadana (Mario Salgado Orellana, Cristobal Figueroa Burgos, Vicente Rios Adasme, Joaquin San Martin Vargas)

La Encuesta Nacional Urbana de Seguridad Ciudadana (ENUSC) es un estudio realizado por el INE en Chile, que busca conocer la percepción de inseguridad, la victimización de personas y hogares, y la reacción frente al delito en zonas urbanas, tanto a nivel nacional como regional. Se espera conocer cuáles son las percepciones de la ciudadanía y en qué inciden de manera de comprender cómo deberían enfocarse las políticas públicas correspondientes.

## 7. Curiosidades de Airbnb en Santiago, Chile (Sergio Opazo Zamorano, Jorge Avendaño Flores, Vicente Wiederhold)

Airbnb es una plataforma en línea que conecta a viajeros con anfitriones que ofrecen alojamientos para estancias cortas, como habitaciones, apartamentos o casas, permitiendo a las personas reservar y hospedarse en diferentes lugares del mundo. En este proyecto se espera que explore relaciones entre huéspedes, anfitriones, características de las habitaciones/departamentos/casa, ubicación, precio, etc. ¿existen correlaciones? También explorar comentarios.

## 8. Análisis financiero de los casinos de juegos (Bastían Ceballos, Carlos Alvarez, Guillermo Oliva, Martín Llanos)

La Superintendencia pone a su disposición publicaciones referida a la industria de casinos de juego, tales como el Informe Anual de la Industria de Casinos de Juego, el que entrega antecedentes respecto de las características de los casinos de juego y de sus resultados operacionales en forma anual; y el Boletín Estadístico, en el que mensualmente se dan a conocer los principales resultados referidos a los ingresos brutos del juego, impuestos, número de visitas, gasto promedio por visita y porcentaje de retorno real a clientes en máquinas de azar. Se debe analizar el comportamiento anual, tendencias, correlaciones, entre otros.

## 9. Análisis de Atenciones de urgencia en el sistema de Salud (Ignacio Diaz, Martin Figueroa, Alex Blanchard y Benjamín Diaz)

El MINSAL reporta anualmente las atenciones de urgencia de la red pública de salud. Este conjunto de datos contiene información agregada de las consultas de urgencias de la red pública de Salud obtenidas de las atenciones realizadas en los centros de salud y establecimientos hospitalarios en el país, originadas a partir del registro del DAU (Datos de Atención de Urgencia). Se entiende por urgencia o emergencia vital toda condición clínica que implique riesgo de muerte o de secuela funcional grave, la atención de urgencia o emergencia vital ante un hecho de tal envergadura, debe ser inmediata e impostergable. Se espera encontrar relaciones, y comparar bajo distintos criterios.

#### 10. Análisis de matrículas en pregrado y oferta académica

El consejo nacional de educación superior, reporta anualmente tanto la matrícula en pregrado como la oferta académica por universidad. En este proyecto, se espera que analice cada dataset por sí mismo y luego correlacionar entre sí, explore si existen tendencias y/o patrones por valor, región, carrera, etc.

11. Gabriel Huerta, Diego Oyarzo y Javier Cadagán. Base de datos llamada ChatCoder2.0 es del año 2006 aproximadamente, hay aproximadamente 300 archivos .txt con conversaciones que duran algunos días.