



Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación
IIC2026 – Visualización de Información

Proyecto

Diseño de una herramienta de visualización

Fecha de inicio: **28 de abril** a las **20:00** hrs

Fecha de entrega del **Avance**: **22 de mayo** a las **20:00** hrs

Fecha de entrega: **28 de junio** a las **20:00** hrs

Evaluación en el contexto del curso

Esta evaluación es de naturaleza **sumativa**, y pretende rescatar evidencias del desarrollo de ciertos resultados de aprendizaje. Por eso, tras su realización y entrega, recibirá retroalimentación sobre su desempeño y una nota que lo refleja.

Específicamente, esta evaluación se relaciona con el resultado de aprendizaje indicado en el programa del curso: **Diseñar una herramienta de visualización de información mediante un proceso de análisis y diseño, en una situación escogida.**

Todo el material del curso es relevante para esta evaluación.

Durante las últimas sesiones del curso, se realizarán espacios de trabajo asistido para esta evaluación en horario de clase.

Este proyecto consiste en 2 entregas: una propuesta de proyecto detallada en la sección [Propuesta de proyecto](#), y la entrega final de este que incluye un informe y la herramienta de visualización. Luego de la entrega final, se incluirá una evaluación de pares, la cual puede tener directa repercusión en la nota final de los integrantes.

1. Diseño de una herramienta de visualización

La intención de esta evaluación es vivir una experiencia que integre y resuma el modelo de trabajo revisado en el curso Visualización de Información. Es decir, se busca que se aplique el modelo anidado de Tamara Munzner revisado en clase, en sus cuatro distintos niveles, para resolver un problema de visualización.

Como resultado de esta experiencia, deben entregar **dos resultados principales**: un **informe en formato HTML** donde se explique el proceso realizado y las decisiones tomadas en el camino; y **una herramienta de visualización de información** implementada por ustedes que sea el reflejo de su proceso de diseño.

Tanto el tema del problema de visualización como el *dataset* que usarán como base para su experiencia y creación, **puede ser escogido por ustedes**. Pueden dedicarlo a un tema que les interese, que los apasione, que los entretenga, o del cual simplemente quieran aprender. En **Anexo: Ejemplos de exámenes de semestres pasados** se presentan algunos ejemplos de exámenes con temáticas únicas que pueden servirte de inspiración.

El primer requisito mínimo de esta evaluación es la originalidad. **No se permite replicar una herramienta confeccionada por terceros**. Se permite utilizar los mismos datos que otra herramienta, pero las visualizaciones a crear deben ser diferentes. Un aspecto vital de esta evaluación es que **la herramienta a confeccionar debe tener su propia contribución substancial a la visualización de información**.

Junto a la originalidad, el segundo requisito mínimo de esta evaluación es confeccionar al menos **tres visualizaciones diferentes**. Estas visualizaciones deben cumplir ciertas condiciones detalladas en la sección **Tercer nivel: Codificación visual y de interacción**.

El resto de este enunciado explicita los pasos a seguir en esta experiencia, requisitos a cumplir y lo que se espera entregues como resultado final.

1.1. Elección de tema y datos a utilizar

El primer paso es escoger la temática a tratar y el (o los) *dataset(s)* a utilizar.

La única restricción a considerar sobre la elección de tema y datos, es que no es posible elegir un *dataset* o tema ya trabajado **con D3** en otra instancia del curso (Demos de clase, Ayudantías o Tareas). Al momento de elegir el *dataset*, deben registrar este en el formulario indicado a continuación. De este modo, si en una ayudantía **posterior** a haber respondido el formulario se ocupa el mismo *dataset* registrado, se permitirá su uso de todas formas. Pueden responder las veces que quieran el formulario para actualizar el *dataset* a ocupar.

Link al formulario: <https://forms.gle/PbFM4jAq9LCLanrM8>

Fuera de lo anterior, **no hay límites en los temas a abordar**, mientras se traten y presenten con respeto. Como hemos visto a lo largo del semestre, ¡las visualizaciones de información permiten contar historias de cualquier tipo!

Alternativamente se proveen en **Anexo: Temas y datasets predeterminados** varios *datasets* temáticos pre-establecidos, de los cuales pueden escoger uno para trabajar.

Si tienen alguna duda sobre qué tema puedes trabajar, pregunten en las [discussions del Syllabus](#) o durante las sesiones de trabajo en clases.

1.2. Informe

Uno de los resultados finales a entregar es un informe. Este debe ser un documento en formato **HTML**, que se pueda abrir y leer directamente en un navegador web. Se invita a que se agregue estilo a dicho informe mediante un archivo externo de extensión **CSS**, con el fin de hacerlo visualmente más agradable y legible, pero no es completamente necesario.

Este informe debe contar sobre el proceso generando la herramienta de visualización final. Específicamente, se espera que se incluya los resultados que se obtienen al llevar a cabo los primeros niveles del modelo anidado de Tamara Munzner. En las siguientes sub-secciones se aclara lo que se espera de cada nivel.

1.2.1. Primer nivel: Caracterización de dominio

Se debe desarrollar y aplicar el primer nivel del modelo anidado: **Caracterización de dominio**. La caracterización conceptual debe ser parte del **informe** a entregar y debe tener su propia sección en este.

Como parte del primer nivel, es necesario dar **contexto y describir la situación a modelar**. Específicamente:

- Es importante dar suficiente contexto para que personas que no conozcan del tema puedan continuar leyendo.
- Dar a entender la naturaleza de los datos a trabajar.
- Explicitar la fuente de donde se obtienen los datos a trabajar y si hubo algún tipo de pre-procesamiento previo.
- Dar a conocer al potencial usuario objetivo de la herramienta.
- Explicitar el tipo de uso que tendría la herramienta.

No es necesario un nivel de detalle enorme para esta sección, pero **sí lo suficiente para articular el trabajo de los siguientes niveles**.

1.2.2. Segundo nivel: Abstracción de datos y tareas

También se debe aplicar el segundo nivel del modelo anidado: **Abstracción de datos y tareas**. Los resultados del trabajo de abstracción deben ser parte del **informe** a entregar y también debe tener su propia sección en este.

Se espera que se realice el **análisis de datos y tareas** de la situación propuesta. En base a las ideas plasmada en la caracterización del dominio, hay que realizar un análisis consecuente y que utilice el vocabulario e ideas introducidas en el curso.

Para la abstracción de tareas, se espera se identifiquen **tres tareas principales**, junto con explicaciones breves que aclaren su relación con la caracterización de dominio. Estas tres tareas pueden considerar acciones de distintos niveles de abstracción o todas pertenecer un mismo nivel, pero deben ser **distinguibles entre sí**. Es decir, no deben ser idénticas o ser distintas formas de expresar la misma tarea.

Es posible incluir más de tres tareas si así lo desean, pero todo lo adicional no se considerará dentro de la corrección base. Por eso **es imperante explicitar en el informe cuales son las tres tareas principales a considerar en la corrección¹**.

¹Más detalles de esto en **Corrección y pauta de evaluación**.

Por otro lado, en la abstracción de datos se espera que determinen los tipos singulares de datos (ítem, atributo, enlace, posición) y describan el tipo de atributos para cada dato que corresponda. **No es necesario determinar tipo de dataset** ni justificar todas las categorizaciones realizadas.

Eso sí, **solo es necesario considerar datos propios o derivados del dataset que efectivamente sean parte de las tareas de interés y que se ocuparán finalmente**. Es decir, si el conjunto de datos original incluye atributos, ítems o información que no están en el espectro de interés, no es necesario trabajarlos o analizarlos en el informe.

1.2.3. Tercer nivel: Codificación visual y de interacción

También se debe aplicar el proceso del tercer nivel de: **Codificación visual y de interacción**. El desarrollo de este nivel también debe entregarse como parte del **informe** y debe tener su propia sección en este.

Específicamente, se espera un **listado de elecciones** de marcas, canales, *idioms* visuales, *idioms* de interacción y decisiones de diseño realizadas para crear una solución de visualización, que se base en la abstracción de datos y tareas. Se esperan **justificaciones coherentes y breves** para cada elección y decisión realizada, en base a los distintos principios, criterios o guías revisadas en el curso.

Finalmente, hay un listado de decisiones mínimas que su herramienta **deberá desarrollar** de forma justificada. Este listado es:

1. Uso de navegación por medio de `d3.Zoom` o de selección de múltiples elementos por medio de `d3.Brush`.
2. Incluir al menos un filtro en la visualización para reducir la cantidad de información desplegada. Este puede ser por medio de elementos de *input* de HTML (`<input/>`) o interactuando directamente con la visualización.
3. Incluir una interacción con el *mouse* para **seleccionar un elemento (ítem, atributo, categoría, etc.) en la visualización**. El efecto tras efectuar dicho evento queda a tu criterio.
4. Implementar vistas coordinadas entre dos o más visualizaciones, es decir, la acción sobre una visualización tiene repercusión en otra visualización. Como referencia, en la tarea 3 ya se implementó vistas coordinadas cuando se solicitó que la visualización de películas se actualizara según el trébol seleccionado.
5. Utilizar correctamente el *Data Join* cuando es necesario actualizar y eliminar elementos del SVG. Es importante mencionar que eliminar y crear un nuevo SVG para actualizar la visualización no es una correcta aplicación del *Data Join*.
6. Seleccionar al menos **3 visualizaciones** para resolver las tareas visuales. Estas deben cumplir las siguientes condiciones:
 - (a) Una visualización **NO puede pertenecer al siguiente conjunto**: gráfico de barra, gráfico de barra apilado, gráfico de área apilado, *hexbin*, histograma, gráfico de dispersión, gráfico de burbuja o un *span chart*. En otras palabras, debe ser una visualización donde no hemos entregado el código ya armado.
 - (b) Otra visualización **DEBE ser alguna del siguiente conjunto**: mapa de coropleta, mapa de puntos/brubujas, cartograma, matriz de adyacencia, grafo, *treemap*, *Circular Packing*, mapa de calor, gráfico de multilínea, gráfico de caja o gráfico de violín.
 - (c) La última visualización no tiene ninguna restricción mientras sea diferente a las demás.

En caso de implementar menos de 3 visualizaciones, el proyecto será penalizado. Además, se aplicará una penalización adicional si no se respetan las siguientes condiciones:

- En caso de implementar únicamente 2 visualizaciones, estas deben respetar los puntos (a) y (b).
- En caso de implementar únicamente 1 visualización, esta debe respetar el punto (a).

1.3. Herramienta de visualización

El otro resultado final a entregar de esta evaluación es una herramienta de visualización. Esta también debe estar contenida en un documento HTML, que se extiende mediante hojas de estilo CSS, y se implementa mediante un programa escrito en **JavaScript** y se base únicamente en [D3.js](#).

Este documento se considerará el resultado del aplicar último nivel del modelo anidado:

1.3.1. Cuarto nivel: Implementación algorítmica

En base a las decisiones de codificación y diseño realizadas, se debe trabajar el cuarto nivel de: **Implementación algorítmica**. Es decir, **hacer realidad la herramienta de visualización mediante programación**.

Se espera que la herramienta entregada refleje y sea consecuente con el trabajo expuesto en el informe, especialmente con las decisiones y codificaciones seleccionadas para la situación.

También, como una buena interfaz humano-computador, se espera que no presente errores de funcionamiento o en consola, además de ser una herramienta auto-contenida con los elementos contextuales necesarios.

Al momento de implementar la herramienta, se debe tener en cuenta las siguientes consideraciones técnicas:

- Se debe utilizar D3.js para el desarrollo de las visualizaciones, y solo las versiones 6 o 7 de D3.js.
- Las visualizaciones deben ser desarrolladas con elementos propios de un SVG.
- No utilizar *loops* (`for`, `while`, `forEach`, `each`) para construir la visualización. Solo si necesitan calcular algún dato adicional, se permitirá utilizar *loops*. Se recomienda confirmar su uso en el foro antes de entregar.
- Se espera estilamiento nativo mediante CSS para esta evaluación, es decir, no se permite importar *frameworks* o herramientas de estilamiento ya construidas.
- El programa solo puede hacer uso de funciones nativas de JavaScript o provistas por D3.js. Pueden existir excepciones a esta regla, como la librería de manejo de TopoJSON o la librería de D3 para hacer Hexbin. De tener la duda sobre cierta librería, pregunta en las [discussions del Syllabus](#).
- Implementar las visualizaciones indicadas en el informe y no otras.

De no seguir alguna de estas especificaciones, la parte de la evaluación que revisa los aspectos de implementación no serán corregidos.

1.4. Consejos importantes

Aquí se listan consejos importantes a considerar cuando se entable a realizar esta evaluación:

- Aprovechar plazo de la evaluación y comenzar con anticipación. Al menos comiencen leyendo bien el enunciado y pensando qué tipo de situación se desea trabajar. También se adjunta en **Anexo: Ejemplos de exámenes de semestres pasados** algunos ejemplos de exámenes de semestres pasados que servirán para entender lo que se espera de la evaluación. Se recomienda revise algunos de estos ejemplos.
- Como son múltiples niveles a desarrollar, es posible separar el trabajo por hacer en base a esos niveles. Así se contará con múltiples momentos acotados de avance, en vez de pocos largos periodos de trabajo.
- Se recomienda que durante la primera semana de plazo, al menos se elija el tema y *dataset(s)* a trabajar y comenzar redactando la Caracterización de Dominio. De no encontrar un tema después de intentar bastante, se sugiere elegir uno de los temas en **Anexo: Temas y datasets predeterminados**.
- El modelo anidado presentado no es necesariamente lineal en el tiempo. Puede ocurrir que uno se de cuenta que algo no calza durante el nivel de implementación, que ya se había decidido previamente en el nivel de codificación. Está bien volver, corregir o reconsiderar aspectos de niveles anteriores.
- Se recomienda tomar nota del proceso a medida que se trabaje y usar eso como base para el informe. Una vez que se esté cerca de completar la evaluación, comenzar completando el informe. Aún que también es completamente válido comenzar por el informe y luego avanzar en la implementación.
- Se recomienda pre-procesar los datos a trabajar. Filtrar aspectos que no se vayan a utilizar, y transformar el resultado a un formato y organización que sea ventajoso en el programa de la herramienta.

2. Propuesta de proyecto

Para esta evaluación, deberán hacer una entrega parcial del proyecto. En esta entrega deberán:

- Indicar de forma definitiva el/los *datasets* a utilizar.
- Incluir un *link* para ver todos los datos o al menos un subconjunto de ellos.
- Entregar una pequeña descripción de los datos a utilizar para que el cuerpo docente entienda qué representan estos datos. Basta con nombrar los diferentes tipos de atributos y su significado. Pueden inspirarse en la explicación de los datos de la Tarea 2 y 3.
- Proponer al menos **2 ideas de visualizaciones** que desean desarrollar con estos datos. Estas ideas pueden cambiar durante el desarrollo del proyecto, esto solamente es para poder entregar retroalimentación de la dificultad que implican dichas ideas.

Esta entrega equivale a **5 décimas de la nota final del proyecto**. Por lo cual, no entregar lo solicitado equivale a optar a un 6.5 como máximo en el proyecto (sin considerar el *bonus*).

Cercano a la fecha de entrega, se abrirá un buzón por Canvas para que realicen esta entrega en **formato PDF**.

3. Corrección y pauta de evaluación

Para la corrección de esta evaluación, se revisará el resultado de la herramienta entregada y del informe. Además, se usará una pauta como guía. Cada ítem será calificado de forma ternaria: Cumple totalmente el ítem, cumple parcialmente el ítem o no cumple el ítem, o de forma binaria: cumple totalmente o no.

Además de determinar el nivel de logro alcanzado, el equipo docente adjuntará retroalimentación escrita que complemente la corrección.

Como el proyecto es un trabajo que puede tomar muchas formas y direcciones, se articulará la distribución de puntaje de forma que sea posible contar con una base mínima de alcance a lograr en la entrega. Esta sección del enunciado intenta explicar cómo se distribuye el puntaje en la corrección final.

En la sección **Segundo nivel: Abstracción de datos y tareas**, se solicitan **tres tareas principales**. Este requerimiento es la forma de asegurar un alcance mínimo para la evaluación. Será en función de esas tres tareas cómo se articula puntaje en el resto de los niveles.

Es decir, se repartirá el puntaje en la sección de abstracción de tareas en base a estas tres tareas principales. Por eso es importante que **se explicita en el informe cuales son las tres tareas principales mínimas a considerar**. Por lo tanto, no indicar las tareas principales en el informe repercutirá en que la mayoría de los ítems no serán corregidos.

De forma similar, la corrección de codificaciones se realizará organizándolas como grupos de elecciones relacionadas a las tres tareas principales identificadas. Lo mismo ocurre con puntaje de implementación, que se organiza en parte por aspectos funcionales relacionados a las tareas a cubrir y las decisiones indicadas en el informe.

Por otro lado y como fue mencionado anteriormente, si bien es posible considerar más tareas en una evaluación si así se desea, eso no tiene una recompensa en nota y solo se entregará retroalimentación si el corrector lo encuentra oportuno.

4. Entregables

Se espera como entregables dos resultados principales: un **informe** y una **herramienta de visualización de información**. Ambos resultados debe entregarse como documentos **HTML**, extendidos mediante archivos **CSS** y **JavaScript**. No se aceptarán entregas en cualquier otro formato distinto al indicado anteriormente (archivos PDF, TXT, DOC, etc...). De no entregar o entregar un formato diferente al especificado, no se revisará la entrega y se colocará nota mínima.

La entrega se realizará mediante un archivo comprimido de extensión **ZIP** que contenga la herramienta e informe, mediante la plataforma Canvas en la evaluación correspondiente. Cualquier otro archivo (imágenes, hojas de estilo, *scripts*, **datos**, etc.) que sea necesario para la visualización correcta de tu entrega **debe también ir incluido en el archivo comprimido**.

5. Dudas

Cualquier duda que tengas sobre esta evaluación, prefieran publicarla en el **Syllabus del curso** correspondiente a esta evaluación. También, sientan la libertad de responder dudas de tus pares si crees que conoces la respuesta. En caso de tener dudas que impliquen mostrar su solución o partes de ella, no utilicen este medio de consulta. Para estos casos, envíe un correo al cuerpo docente o muestre su solución solo en reunión personal (remota o presencial) cuando se reúna con algún miembro del cuerpo docente.

6. Política de atraso

Existe la posibilidad de entregar esta evaluación con hasta **3 días de atraso** a partir de la fecha de entrega definida en el enunciado. En la eventualidad de entregar pasada la fecha de entrega, se aplicará una **reducción** a la nota máxima que podrás obtener en esta evaluación.

De haber atraso, **la nota máxima a obtener** se reduce en **0.5 puntos (5 décimas)** por cada día de atraso. Cada día de atraso se determina como el techo de días de atraso. Por lo tanto, en caso de entregas atrasas, la nota final de la tarea se calcula mediante la siguiente fórmula:

$$\min(7 - (0.5 \times \text{días_atraso}), \text{nota_obtenida}) + \text{bonus}$$

Por otro lado, entregas con más de 3 días (72 hrs) de atraso no serán recibidas y serán evaluadas con la **calificación mínima (1.0)**.

7. Bonus de la evaluación (10 décimas)

El día 6 de julio se utilizará el módulo de examen asignado por Ingeniería para rendir un examen *bonus*. El rendimiento en este examen *bonus* bonificará la nota final del examen y/o permitirá recuperar puntos RC. Esta evaluación consistirá en 30 preguntas en Canvas cuya corrección será automatizada.

Los términos y condiciones de esta evaluación son:

- Esta evaluación es de carácter opcional y no incluye opción de recuperativa en caso de ausentarse. Solo se tendrá consideración con quienes justifiquen debidamente su ausencia con su Unidad Académica y dicha justificación sea aceptada.
- Esta evaluación *bonus* se realiza de forma **individual**.
- Deben traer computador o *tablet* para poder contestar la evaluación. El computador solo será el medio para responder, pero no se permitirá utilizarlo para estudiar los contenidos. Cualquier estudiante que se sorprenda utilizando el computador para fines distintos a responder el cuestionario, no se le entregará la bonificación y se le solicitará abandonar la sala.
- Pueden traer los apuntes que estimen convenientes. No obstante, estos deben ser escritos a mano (en papel o tablet).² No se aceptará impresiones como apuntes para esta evaluación.
- La evaluación será por Canvas y solo permitirá ingresar las respuestas una vez. Finalizado el plazo, se entregará el puntaje final y cada estudiante podrá canjear el puntaje obtenido siguiendo la siguientes reglas:
 - Puede canjear 3 puntos para tener 1 décima adicional a la nota del proyecto.
 - Puede canjear 10 puntos para tener 1 punto RC adicional.
 - Luego de canjear todo el puntaje posible en los dos ítems anteriores, si quedan puntos por canjear, puede solicitar *stickers*.

²[Estudios demuestran que escribir a mano ayuda a la memoria](#)

8. Anexo: Temas y *datasets* predeterminados

En esta sección se listan una serie de temas y *datasets* diferentes. Son posibles elecciones a trabajar en la evaluación.

Es importante notar que si se escoge un tema de este listado, **no es necesario usar el *dataset* completo**. Perfectamente puede que atributos, datos o secciones de ciertos conjuntos queden fuera de la situación de interés a trabajar. Similar a como aconsejado en [Consejos importantes](#), en dichos casos se recomienda pre-procesar y filtrar los datos, para así utilizar un conjunto más liviano y flexible.

También, es posible combinar datos de distintas fuentes. Por ejemplo, trabajar algún conjunto de esta lista que tiene datos geométricos y combinarlo con un archivo GeoJSON que apoye esos datos.

8.1. Datos de participación electoral en Chile

Con el plebiscito 2020, y época electoral que existe en Chile, un tema de discusión frecuente es la participación electoral: cuántas personas efectivamente votan.

Un tema de interés posible es entender cómo fue la participación electoral durante el 2017, otro año cuando Chile vivió múltiples elecciones. SERVEL publica datos sobre este ámbito, por ejemplo, para las [Elecciones Primarias 2017](#). La página muestra por separado la participación electoral según varias segmentaciones, como por edad y partido político, pero no muestra cruces entre segmentaciones.

También provee los [datos originales](#) utilizados para realizar tales segmentaciones. Tiene información de votantes específicos: edad, género, nacionalidad, comuna, partido y si sufragó o no. **Este archivo es bastante pesado, por lo que se recomienda pre-procesarlo.** [Aquí](#) entregamos una versión ya filtrada y dividida por regiones de Chile.

Esta es sola una mirada a considerar de datos electorales. El [sitio de SERVEL](#) provee estadísticas de [participación electoral en varias elecciones](#), e incluso [datos abiertos electorales en general](#).

8.2. Otros datos procesados por el Profesor

En el siguiente [Spreadsheet](#) se recopilaron diferentes *dataset* utilizados por el profesor Hernán Valdivieso en diferentes instancias relacionadas con la visualización.

8.3. Otros *datasets* sin preprocesar

A continuación se incluye un listado de *datasets* de interés, pero que no han sido procesados y/o analizado por el cuerpo docente:

- [Dataset de libros](#).
- [Dataset de League of Legend](#) o [página oficial con datos del juego](#).
- [Dataset de ventas de videojuegos](#).
- [Dataset con datos de aplicaciones de Google Play Store](#).
- [Dataset de astronautas de NASA](#)
- [Dataset de Animal Crossing New Horizons](#)

9. Anexo: Ejemplos de exámenes de semestres pasados

En los siguientes enlaces: [salón de la fama 2020](#), [salón de la fama 2021](#) y [salón de la fama 2022](#) podrás encontrar las herramientas funcionales de varios exámenes del año 2020 al 2022 que siguieron instrucciones de evaluación similares al propuesto.

Son libres de revisar estos ejemplos como inspiración para encontrar su propio tema o *dataset* a trabajar. Nota que es completamente posible usar el mismo tema o *dataset* que alguna de estas evaluaciones, ya que las tareas o contexto a resolver puede ser muy distinto. Pero también ojo que el replicar o usar como base una de estas evaluaciones (código o de diseño) caería como falta de integridad académica, como es indicado en el [programa del curso](#).

El enlace mencionado solo incluye herramientas funcionales y no informes de dichos exámenes. Como referencia de cómo estructurar el informe del proyecto, te presentamos a continuación adaptaciones de los informes de dos ex-estudiantes del curso:

- [Informe de Daniela Concha](#), sobre el tema “Lenguas extintas y en peligro”. Puedes acceder a la herramienta correspondiente [aquí](#).
- [Informe de Francisco Guiñez](#), sobre el tema “Minneapolis: Violencia Policial y Racismo”. Puedes acceder a la herramienta correspondiente [aquí](#).