

# CC5208 Visualización de Información

## Tarea 3: Visualizaciones con D3

Prof. Benjamin Bustos

Aux. Pablo Estefó

7 de mayo de 2018

Fecha de entrega: domingo 20 de mayo a las 23:59

El objetivo de esta tarea es utilizar D3.js para crear visualizaciones Web interactivas y con ellas analizar datos de condición de vulnerabilidad estudiantil en colegios de Chile. Los datos fueron recolectados desde <http://www.junaeb.cl/ive>.

1. Construya un diagrama de dispersión (*scatter plot*) de los colegios de educación media. Para esto, cargue los datos del archivo `IVE-SINAE-MEDIA.csv` y agregue las siguientes características a su visualización:
  - Eje x: total de matriculados en el año 2015.
  - Eje y: porcentaje de vulnerabilidad IVE-SINAE Color: dependencia del colegio.
  - Área: porcentaje de alumnos con primera prioridad de vulnerabilidad.
  - Incluya una leyenda de los colores
  - Incluya un tooltip interactivo que al pasar el mouse sobre cada colegio muestre información acerca de éste.
2. Diseñe una visualización (que no sea un diagrama de dispersión) que permita observar la distribución de alumnos vulnerables por región con respecto a su dependencia. Tenga en consideración lo siguiente:
  - Los datos de alumnos de educación básica y media en los archivos `IVE-SINAE-BASICA.csv` y `IVE-SINAE-MEDIA.csv`, respectivamente.
  - La función `nest()` de D3 le puede ser útil para agrupar datos por región.
  - Visite la página de D3.js [d3js.org](http://d3js.org) para explorar opciones de visualizaciones posibles a crear con D3.

- La visualización no necesariamente debe mostrar todos los datos al mismo tiempo; puede utilizar selectores interactivos que carguen un set de datos por vez.
- Considere agregar un *tooltip* interactivo que al pasar el mouse entregue un resumen de los datos.

Para la evaluación de esta pregunta se considerará tanto que funcione correctamente y la visualización creada cumpla con los requisitos anteriores, como que su diseño esté bien justificado.

3. Realice un análisis de los datos usando las visualizaciones creadas y entréguelo en un informe de máximo cinco hojas (incluyendo portada y figuras). ¿Qué puede decir de los datos? ¿Qué patrones se observan? Además, para el punto 2, agregue una descripción detallada de por qué eligió la visualización usada y justifique su diseño.

## Descripción de los datos

El Sistema Nacional de Asignación con Equidad (SINAE) es un conjunto de criterios que permite identificar distintos grupos dentro de la población de estudiantes de educación básica y media de establecimientos municipales o particulares subvencionados del país, de acuerdo al nivel de vulnerabilidad que presentan. Así, la población identificada como “vulnerable” queda clasificada en 3 prioridades: primera, segunda y tercera prioridad.

A continuación se presenta la descripción las prioridades[1] y de los demás datos[2] en los archivos a utilizar:

- RBD: Rol base de datos del establecimiento (nominal)categorico.
- DV\_RBD: Dígito verificador del RBD (nominal)categorico.
- NOMBRE\_ESTABLECIMIENTO: Nombre del establecimiento (nominal categorico)
- DEPENDENCIA: Dependencia del establecimiento, {MUNICIPAL - DAEM, PARTICULAR SUBVENCIONADO, MUNICIPAL - CORPORACIÓN, ADMINISTRACIÓN DELEGADA} (nominal categorico)
- AREA: Tipo de área donde se ubica el establecimiento {URBANO, RURAL} (nominal categorico)

- **COD\_REGION:** Código de la región donde se ubica el establecimiento (nominal categórico)
- **COD\_PROV:** Código de provincia donde se ubica el establecimiento (nominal categórico)
- **COD\_COM\_EST:** Nombre de comuna donde se ubica el establecimiento (nominal categórico)
- **COMUNA:** Nombre de comuna donde se ubica el establecimiento (nominal categórico)
- **PRIMERA PRIORIDAD:** estudiantes con riesgos socioeconómicos (ordinal cero absoluto)
- **SEGUNDA PRIORIDAD:** estudiantes con menor riesgo socioeconómico que en 1a Prioridad, pero que además presenta riesgos socio-educativos (problemas de rendimiento escolar, asistencia o deserción) (ordinal cero absoluto)
- **TERCERA PRIORIDAD:** estudiantes con igual riesgo que 2a Prioridad pero sin los problemas mencionados (ordinal cero absoluto)

FUENTE:

[1] <http://junaebabierta.junaeb.cl/catalogo-de-datos/indicadores-de-vulnerabilidad/>

[2] [https://www.junaeb.cl/wp-content/uploads/2013/02/libro\\_junaeb.pdf](https://www.junaeb.cl/wp-content/uploads/2013/02/libro_junaeb.pdf)