**Grupo Idea 12:** Ximena Catalán, Roberto Cantillan, Kevin Carrasco, Valentina González

Pregunta:

¿Cómo es la relación que existe entre nivel de contaminación al que está expuesto el establecimiento educacional (a 1 km a la redonda) y el desempeño académico en el colegio (SIMCE)?

Hipótesis:

* En establecimientos expuestos a mayor contaminación, tendrán peor rendimiento en ambas pruebas.

Qué esperamos:

Visualizar de manera interactiva asociaciones entre variables.

Visualizar el efecto de la contaminación.

Datos:

Se utilizará las bases de datos administrativos del SIMCE para todos los establecimientos del país desde el año 1998 al 2021. De ella se obtienen también variables asociadas a los colegios.

Y los datos de contaminación son extraídos del “Atmospheric Composition Analysis Group”. A través de la medida de contaminación: PM2.5.

Variables a utilizar:

* RBD
* SIMCE (Mat y Leng) promedio por colegio + por año
* N° alumnos por año
* N° rinde mat
* N° rinde leng
* Valor matrícula o copago en pesos
* Nivel de contaminación por ubicación
* Curso en el que se aplica la prueba
* Efecto contaminación (añadir como variable)
* Geolocalización.

Pasos siguientes:

* Unir bases de datos.
* Decidir manera de visualización.
* Calcular efecto contaminación.
* Código Shiny para visualizar u otra herramienta.