**RETO 1: LOS DATOS CUENTAN MUCHAS HISTORIAS**

EL abandono escolar es un problema grave en nuestro país y en nuestro estado. Los niños y niñas abandonan la escuela y no regresan, causando un daño irreparable en sus vidas y en su futuro económico. Entender el problema del abandono escolar no es sencillo – este es un problema multidimensional y los factores que inciden en el abandono no son los mismos en todas las escuelas. Por ejemplo, uno esperaría que las escuelas con mayor abandono estuvieran en los municipios más pobres, no cual no es necesariamente cierto. Lo que sí sabemos es que las razones de deserción no solo tienen que ver con factores escolares, sino con condiciones socioeconómicas, de salud, de criminalidad, y otras.

El análisis se centrará en las **escuelas secundarias de Jalisco**

En este reto, se les darán datos acerca de las escuelas, y de factores varios (educativos, socioeconómicos, salud, crimen, marginación) en archivos .xls (que son fácilmente convertibles a formato .csv)

Estos archivos no son “reales” en el sentido de que no son consistentes en el tiempo (hay archivos con información de diferentes años), pero contienen datos reales en algún momento del tiempo y son datos abiertos y/o públicos que provienen del INEGI, Datos abiertos federales, datos abiertos de Jalisco etc.

Por favor noten que los archivos tienen diferentes grados de granularidad geográfica: municipio (todos), localidad, y colonia (solo ZMG y ciudades medianas).

**ARCHIVOS**

AA1 1 listado de escuelas Jalisco

* Listado de **todas** las escuelas secundarias de Jalisco, dirección y geo-localizacion. Noten que una escuela puede tener varios turnos

AA2 indicadores de eficiencia por escuela de Jalisco

* **Todas** las escuelas de Jalisco
* Los principales indicadores de eficiencia en los que estamos interesados son:
  + Deserción intra-curricular (el abandono una vez que los alumnos entraron a secundaria: de primero a segundo, de segundo a tercero, y los que egresan de tercero). En este ejercicio no estamos interesados en el abandono de primaria a secundaria, o de secundaria a prepa.
  + Eficiencia terminal (cuantos alumnos salieron / cuantos alumnos entraron hace 3 años)

AA3 planea\_basica por escuela Jalisco

* **Todas** las escuelas de Jalisco
* Los resultados del examen Planea (que los alumnos hacen en tercero de secundaria) por escuela, matemáticas y español

AA4 delitos por **municipio** Jalisco

AA5 1 Índice de marginación por **municipio** Jalisco

AA6 1 madres adolescentes por **municipio** Jalisco

AA6 2 población indígena **municipio** Jalisco

AA7 1 socio económica por **colonia** ZMG Jalisco

* El Salto, Guadalajara, Tlajomulco, Tonalá, Zapopán, Tlaquepaque

AA7 2 socio económica por **colonia** ciudades medias Jalisco

* Arandas, Lagos de Moreno, Tepatitlán, Zapotlán el Grande, Puerto Vallarta, Ocotlán

Para los verdaderamente valientes, tenemos datos del Censo Nacional de Población 2020 Jalisco

AA8 1 Censo Jalisco 2020

* El censo completo por población

AA8 2 diccionario Censo Jalisco 2020

* El diccionario de datos para interpretar las variables del censo 2020

**¿ QUÉ ESPERAMOS QUE HAGAN ?**

Queremos que busquen una o más historias acerca del abandono escolar en Jalisco que se generen en los datos que tienen disponibles. Estas historias son en realidad hipótesis que tienen que ser probadas. Pueden centrarse en un municipio, en una colonia, e inclusive en una escuela individual. Pueden hacer analítica de datos y visualización para mostrar el problema. O pueden usar técnicas de aprendizaje de máquinas (*machine learning*) para buscar patrones, clasificar o predecir. Lo más importante es: usen su imaginación, pero todo tiene que estar sustentado en datos y modelos. Si encuentran otros datos (abiertos, públicos) que sean relevantes, identifíquenlos y usenlos.

Por ejemplo:

* Pueden escoger un municipio donde el abandono (deserción intracurricular y/o eficiencia terminal) es particularmente bueno (o malo) y tratar de entender porque.
* Pueden escoger un factor o múltiples factores (por ejemplo educativo, socieconómico o de marginación) y ver qué efecto tiene en abandono escolar, etc.

El resultado del hackathon es que cuenten una o varias historias, pero esa historias debe estar completamente sustentada en datos y modelos. Pueden usar cualquier herramienta (e.g. MS Azure ML studio) o lenguaje de programación (Python , R) o librerías.