**Propuestas de Trabajo Big Data UNT 2024**

**Grupo I:**

**Pilar Valdez**

**Jose López**

**Propuesta 1: El método de clúster como herramienta para la asignación de programas educativos**

Con frecuencia los gobiernos desarrollan programas que tienen como objetivo alcanzar mayores tasas de alfabetización, mejorar la calidad de la educación de las instituciones y/o estudiantes. Sin embargo, usualmente cuentan con presupuestos limitados para llevarlos a cabo, por lo que se plantea la necesidad de desarrollar métodos de asignación eficientes. La idea de este trabajo es construir una base de datos a nivel estudiante que contenga variables relevantes para la determinación de asignación a un programa específico a partir del método de clúster.

Dependiendo del tipo de programa y su principal foco de interés, podríamos agrupar a los estudiantes por rendimiento (bajo, medio, alto), por facilidad para dado tipo de aprendizaje (visual, auditivo, kinestésico), entre otras agrupaciones, con la intención de que cada programa maximice el impacto positivo en los niveles de educación de los estudiantes de acuerdo con las características del grupo al que pertenecen.

Consideramos como predictores potencialmente útiles a este propósito los siguientes factores: contexto familiar (número de hermanos, con quien vive), algún indicador del nivel de ingresos de la familia (por ejemplo, ¿el estudiante recibe un subsidio de transporte o alimentación, su familia y/o el estudiante reciben al menos un plan del estado?), número de faltas durante el último año escolar, educación de los padres, apoyo extracurricular en alguna materia, actividades extracurriculares relacionadas al ocio, actividad física, acceso a la tecnología, aprendizaje digital, sueño, tiempo dedicado al estudio en casa, participación en clase, apoyo terapéutico, ambiente de estudio y condiciones de salud físicas o mentales limitantes.

**Propuesta 2: Comercio Mundial de Productos Citrícolas**

**Motivación para estudiar el tema. ¿cuál es el contexto?**

La industria citrícola en Tucumán es una importante actividad económica de la provincia de Tucumán, que genera dos tipos principales de productos exportables, por un lado la exportación de fruta fresca y por otro la exportación de productos industriales derivados de la fruta fresca (aceite, jugo y cascara deshidratada).

En lo que a Tucumán se refiere la industria citrícola actualmente esta dominada por el Limón, hace 40 años era la naranja (devastada por una plaga sobre la fruta dulce). Es posible que a futuro se reintroduzca la naranja.

En lo que a producción de fruta fresca se refiere, el mundo se puede dividir en hemisferio Norte y Sur, por la contra-estación entre uno y otro. Cuando el hemisferio Norte se queda sin cítricos, se encuentra en producción el hemisferio sur. De manera que los principales competidores para nuestra provincia son Sudáfrica y la misma América del sur.

Ahora, bien los productos industrializados pueden ser guardados durante 1 año a 3 años aproximadamente, por lo que la competencia ya no depende de la estación del año.

Se observó en Tucumán en los últimos tres años una baja en la rentabilidad de la exportación de fruta fresca, lo que derivó en desponte de superficie productiva. Se torna interesante poder entender a nivel mundial el mercado.

**Pregunta de Investigación**

¿Como podemos caracterizar la oferta – demanda en el mercado de la Industria Citrícola a Nivel Mundial? ¿Podemos caracterizar a los Competidores Oferentes? Generar perfiles de precios pagados y barreras aduaneras.

**Posibles datos a usar.**

La FAO (Food and Agriculture Organization) publica datos de hectáreas y producción de diferentes Países de prácticamente todas las frutas.

La OMC (Organización Mundial del Comercio) tiene una Agencia llamada el Centro Internacional de Comercio (ITC) la cual publica una herramienta de análisis de mercado gratuita llamada Trade Map. Es una base de datos en línea que proporciona estadísticas comerciales internacionales. Ofrece una serie de indicadores útiles sobre el rendimiento de las exportaciones, la demanda internacional, los mercados alternativos y el papel de los competidores[1](https://intracen.org/es/recursos/herramientas/trade-map)[2](https://www.diariodelexportador.com/2014/10/herramientas-de-analisis-de-mercado-del_22.html). Trade Map cubre 220 países y territorios y 5300 productos del Sistema Armonizado.

Tenemos acceso a una Base Comercial Privada denominada Softrade. Que cuenta con datos tomados no solamente de oficinas estadísticas oficiales sino también de investigaciones de empresas privadas como ser exportadores e importadores y las empresas de transporte.

**Métodos para la etapa explorativa de cada idea con explicación del razonamiento para usar cada método en esta etapa**

La idea es poder armar base de datos que nos permitan:

Analizar Precios.

Estimar ofertas y demandas por productos a nivel mundial.

Analizar destino de las exportaciones.

Identificar el perfil de nuestros competidores por producto y destino.

Anticipar la oferta mundial en base a datos de hectáreas cultivadas en diferentes países.

Estimar sotck de productos industriales a nivel mundial.

Los Métodos serían:

**Limpieza de datos**. Normalizar los datos, quitar duplicados y dar tratamiento a los datos inexistentes.

**Análisis Descriptivos**. Histogramas, Kernells, PCA, Clusters.

Todo aplicado sobre Oferentes, demandantes, PRECIOS y CANTIDADES.

Analizar también en lo posible los usos de los productos y sus posibles sustitutos. Por ejemplo. El Aceite y jugo se utilizan para saborizar, las grandes compañías saborizan enormes volúmenes de bebidas. A medida que la ciencia permite saborizar con diferentes fórmulas cambian las demandas. A veces usted puede comprar una bebida sabor naranja cuyo sabor fue generado con parte de Lima!

Lo mismo sucede con los perfumes y medicamentos.

**Predecir.** Predecir precios y cantidades en el corto plazo sería un objetivo muy ambicioso! Depende del buen manejo de todos los puntos previos.