

**Módulo: ADM1929 - Business understanding: Pensamiento analítico basado en datos. - (A51)**

**Actividad: Reto de aprendizaje 19. Importación y entendimiento de data**

**Nombre: Roberto Mora Balderas**

**Asesor: José Carlos Soto Monterrubio**

**Fecha: 31 de julio de 2023**

## Objetivo:

Analizar la data proporcionada con la intención de plantear un proyecto con sentido de minería de datos.

## Instrucciones:

Importa la [base de datos](#) y analiza los atributos asociados. Realiza las visualizaciones que consideres adecuadas y explica en un documento la data proporcionada y contestando las siguientes preguntas:

1. ¿Qué tipo de datos presenta el problema? ¿Numérico o categórico?
2. ¿A qué corresponde la primera columna? ¿Es relevante para nuestro problema?
3. ¿A qué se refieren los diferentes atributos? Explica brevemente cada uno en un renglón. De ser necesario busca el significado en internet o pregunta en el foro con tus compañeros.
4. ¿Qué visualizaciones permiten entender los datos de mejor manera? Coloca en el documento al menos 2 visualizaciones.
5. ¿Cómo crees que ayudaría el análisis de clúster a al problema que nos planteamos?

## Desarrollo:

1. Viendo de manera rápida el data set que tenemos podemos observar que la mayoría de los datos son del tipo numérico, salvo la primera columna que es el id del producto el cual posee un valor alfanumérico.

Name	Type
✓ <b>Product ID</b>	Nominal
✓ <b>Fixed Costs ratio</b>	Real
✓ <b>equated margins</b>	Real
✓ <b>Process costs</b>	Real
✓ <b>Utilization rate</b>	Real
✓ <b>Demand_growth</b>	Real
✓ <b>Revenue</b>	Real
✓ <b>RoboticAssy</b>	Real
✓ <b>Rawmatl Costs</b>	Real

2. La primera columna posee el id de los productos, en nuestro caso no es relevante debido a que no es una propiedad que podemos utilizar para agrupar, sin embargo, su utilidad sería para en un análisis posterior para identificar cada producto en la categoría que corresponde.

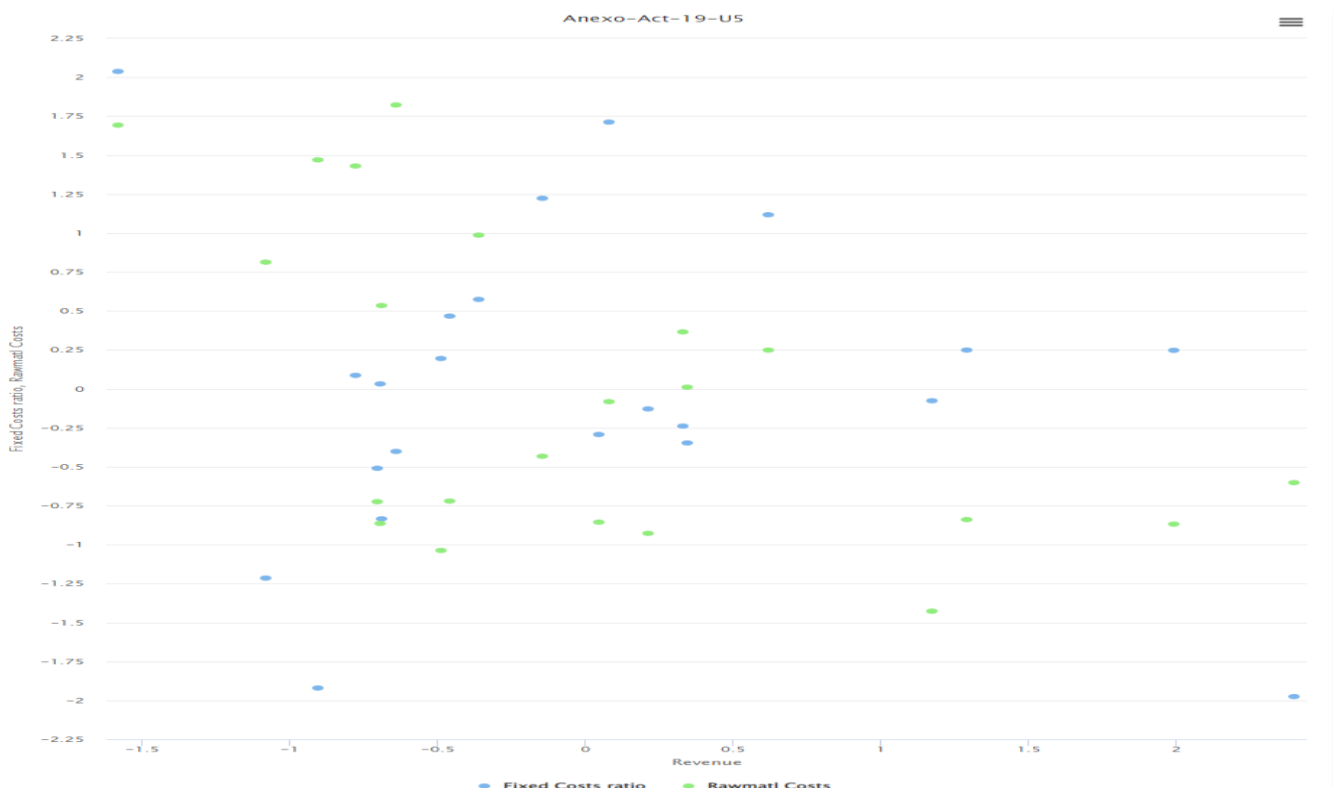
Name	Type	Missing	Statistics	Filter (9 / 9 attributes):
Product ID	Nominal	0	Least strimfm215370 (1) Most Flrpn12457 (1)	Search for Attributes

3. A continuación, se proporcionan los significados encontrados:
  - a) Product ID: Es un identificador único o código que se asigna a cada producto en el sistema de la empresa. Se utiliza para distinguir y rastrear individualmente los diferentes productos que se fabrican o comercializan.
  - b) Fixed Costs ratio: Es la proporción o porcentaje de costos fijos en relación con los costos totales de producción u operación de un producto o proceso. Los costos fijos son aquellos que no varían con la cantidad producida o vendida, como alquiler, salarios fijos, entre otros.
  - c) Equated Margins: También conocido como "margen equitativo" o "margen igualado", se refiere a igualar o ajustar los márgenes de beneficio entre diferentes productos o servicios para obtener una distribución más uniforme de las ganancias.
  - d) Process Costs: Son los costos asociados con el proceso de fabricación o producción de un producto. Esto incluye el costo de materias primas, mano de obra directa, costos de maquinaria y otros costos directos relacionados con el proceso de producción.
  - e) Utilization Rate: Es la proporción del tiempo o capacidad de producción utilizada en relación con la capacidad total disponible. Por lo general, se expresa en porcentaje y representa la eficiencia con la que se está utilizando la capacidad de producción.
  - f) Demand\_growth: Es el crecimiento de la demanda o la tasa de cambio en la cantidad requerida de un producto o servicio durante un período específico. Indica cómo la demanda está aumentando o disminuyendo con el tiempo.

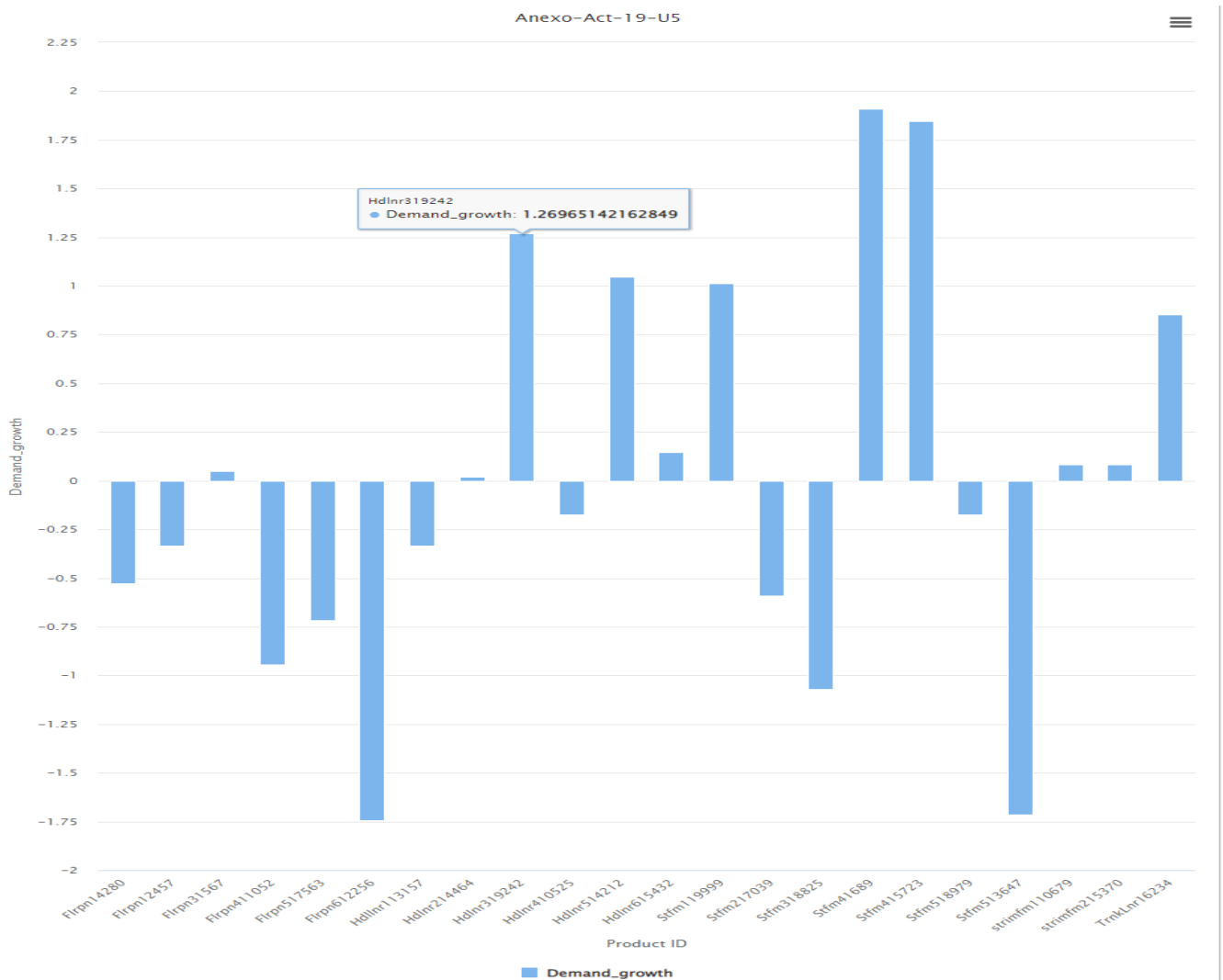
- g) **Revenue:** Es el ingreso total generado por la venta de productos o servicios durante un período específico. Representa la cantidad de dinero que ingresa a la empresa por sus ventas.
- h) **RoboticAssy:** Es un término que sugiere la presencia o uso de sistemas robóticos o automatizados en el ensamblaje o fabricación de productos. Indica que el proceso de ensamblaje puede estar parcial o totalmente automatizado mediante el uso de robots o maquinaria automatizada.
- i) **Rawmatl Costs:** Son los costos asociados con la adquisición o compra de materias primas utilizadas en el proceso de producción. Estos costos representan el gasto en los insumos básicos necesarios para fabricar un producto o proporcionar un servicio.

4. A continuación, se muestran dos visualizaciones:

- a) **Gráfico de dispersión (Scatter Plot):** Permite visualizar la relación entre dos variables, lo que ayudaría a entender si existe alguna correlación o patrón entre ellas. Por ejemplo, se podría usar para ver cómo se relaciona el ingreso (Revenue) con los costos fijos (Fixed Costs ratio) o con los costos de materias primas (Rawmatl Costs).



b) Gráfico de barras (Bar Chart): Puede ser útil para comparar los diferentes valores de una variable categórica, como la presencia de RoboticAssy en los productos (valores binarios 0 y 1). También se podría utilizar para comparar las demandas de crecimiento (Demand\_growth) de diferentes productos.



5. A continuación, se mencionan algunos de los puntos favorables que nos brinda el análisis de cluster:

- Identificación de patrones de comportamiento: El análisis de clúster agruparía los productos que tienen comportamientos de costos similares. Esto permitiría identificar grupos de productos que están sujetos a fluctuaciones de costos de insumos de manera similar, lo que podría indicar factores comunes que afectan sus costos.

- Segmentación de productos según la variación de costos: Al agrupar los productos en clústeres, se podría establecer una segmentación de productos basada en la variación de costos. De esta manera, la empresa podría comprender qué grupos de productos enfrentan mayores variaciones de costos y cuáles tienen costos más estables.
- Planificación de estrategias de compra: Con la segmentación de productos basada en el análisis de clúster, la empresa podría desarrollar estrategias de compra específicas para cada grupo de productos, teniendo en cuenta las fluctuaciones de costos asociadas con cada clúster.
- Optimización de la producción y asignación de recursos: Al agrupar productos con características y comportamientos de costos similares, la empresa podría optimizar la producción y asignación de recursos para satisfacer las demandas de cada grupo de manera más eficiente.

## Referencias

- Anáhuac Online. (2019). *Problema de negocio* [Contenido creado para Anáhuac Online].
- Vohra, G. (2018). *Cluster Analysis For Business*. <https://bit.ly/3pICgmq>
- Whittaker, C. (2019). *7 Innovative Uses of Clustering Algorithms in the Real World*. <https://bit.ly/2KVXP0u>