

Proyecto 2 - Módulo 2

Andrés de la Rosa

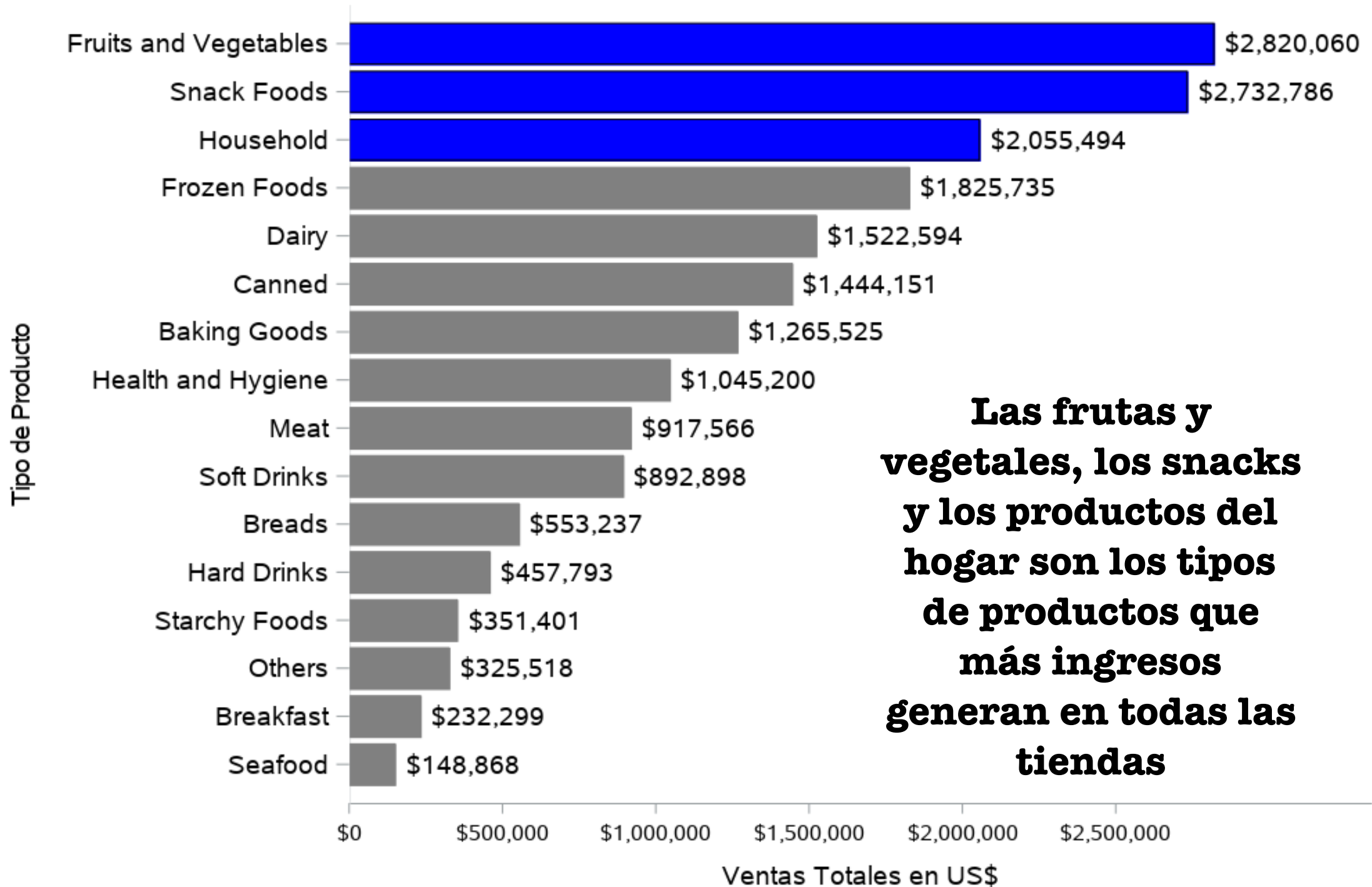
Objetivo del Análisis Exploratorio

Analizar las las ventas reportadas de los productos que se encuentran en el último quintil de las ventas reportadas y a nivel de producto. medir la dependencia de esta variable en otras variables de la base como: el precio del producto, el peso del producto, el nivel de grasa del producto (en los casos que apliquen).

Sin tomar en cuenta las variables agrupadoras de las tiendas únicas: el tamaño de la tienda, la ubicación de la tienda ni tampoco el año en que fue fundada la tienda. De la misma manera, no se hizo un análisis posterior de las multicolinealidades que puedan existir entre variables como el precio del producto y el peso.

Ventas Totales en Dólares

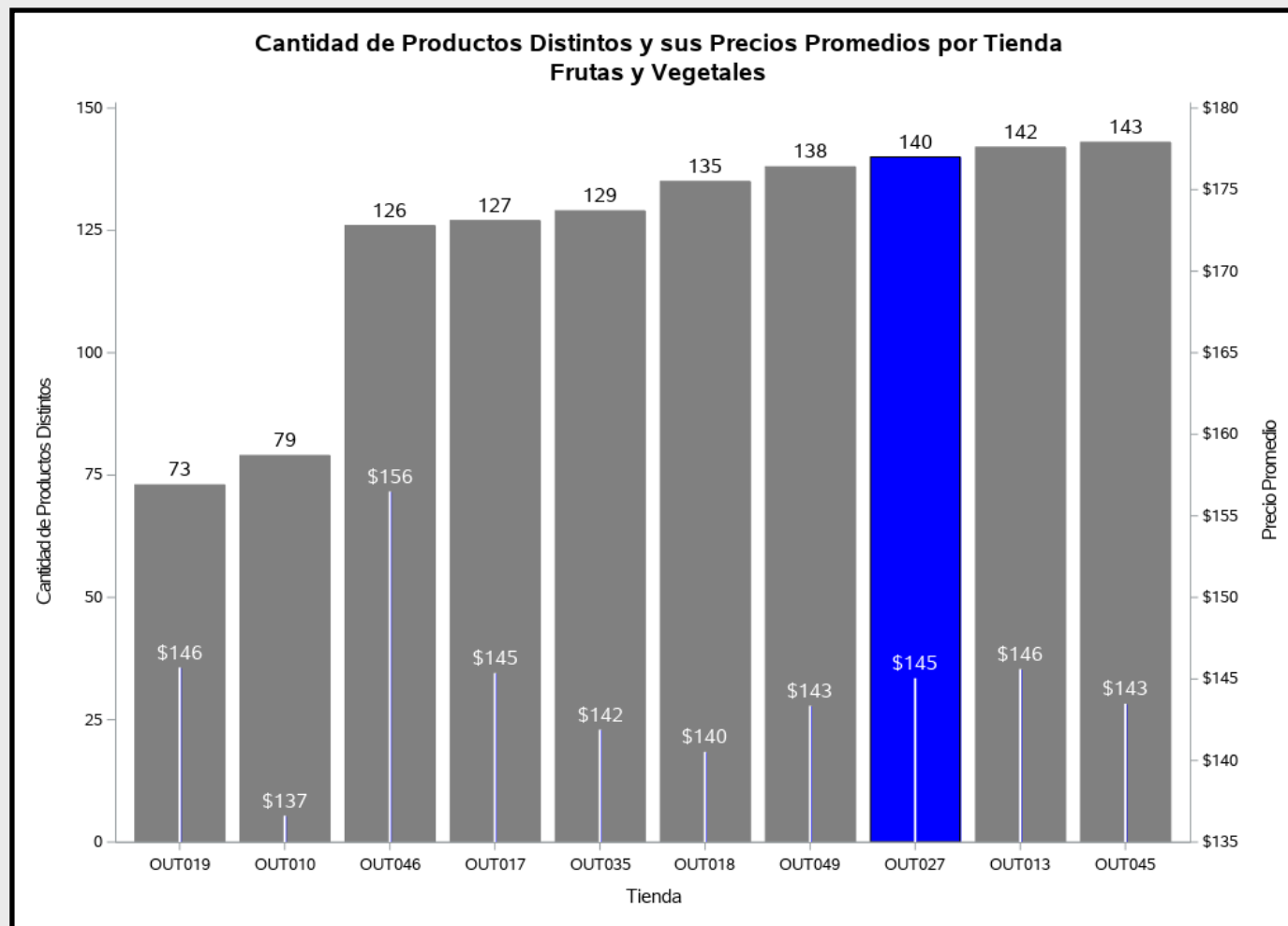
En Todas las Tiendas



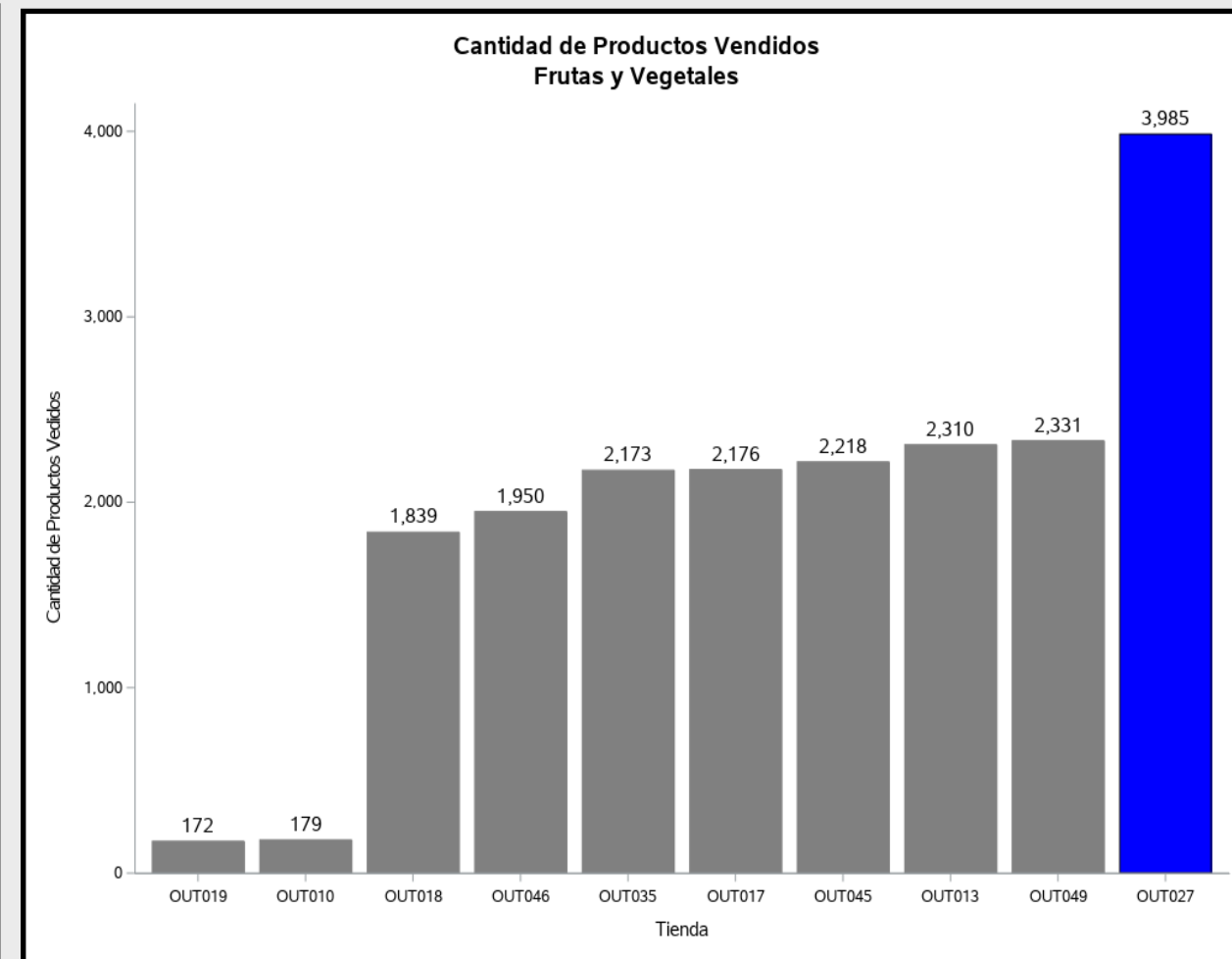
Análisis por Tipo de Producto

Frutas y Vegetales

Cantidad de Productos Distintos y su precio promedio por tienda



Volumen Vendido por Tienda



La tienda que más ingresos tiene por este tipo de producto es **OUT027**. Esta tiene las siguientes características: es una de las tiendas con más productos de este tipo distintos, por lo que tiene una buena diversificación y su precio promedio se encuentra en el tercer quintil de todas las tiendas. También es la **única tienda que es un supermercado del tipo 3** y su tamaño es mediano. Superando en productos vendidos a tiendas similares como OUT013 que es un supermercado grande de tipo 1 que se encuentra en el 3er tier que tiene un precio promedio ligeramente superior.

Frutas y Vegetales

Qué Explican las Ventas de este tipo de producto, a nivel de producto

$$\begin{array}{ccccccccc} \text{Ventas} = & 290.86280 & + & 10.66224 & \text{Grasa} & - & 4.25392 & \text{Peso} & - & 4702.18252 & \text{Visibilidad} & + & 16.51983 & \text{Precio} \\ & (0.1274) & & (0.9020) & & & (0.6777) & & & (0.000) & & & (0.000) \end{array}$$

Este modelo explica un 33.90% de la variabilidad de las ventas de este tipo de producto a nivel de producto y es estadísticamente significativo al 5% del nivel de significancia. Los residuales se distribuyen normal pero hay indicios de heterocedasticidad (ver diagnósticos del modelo adjunto).

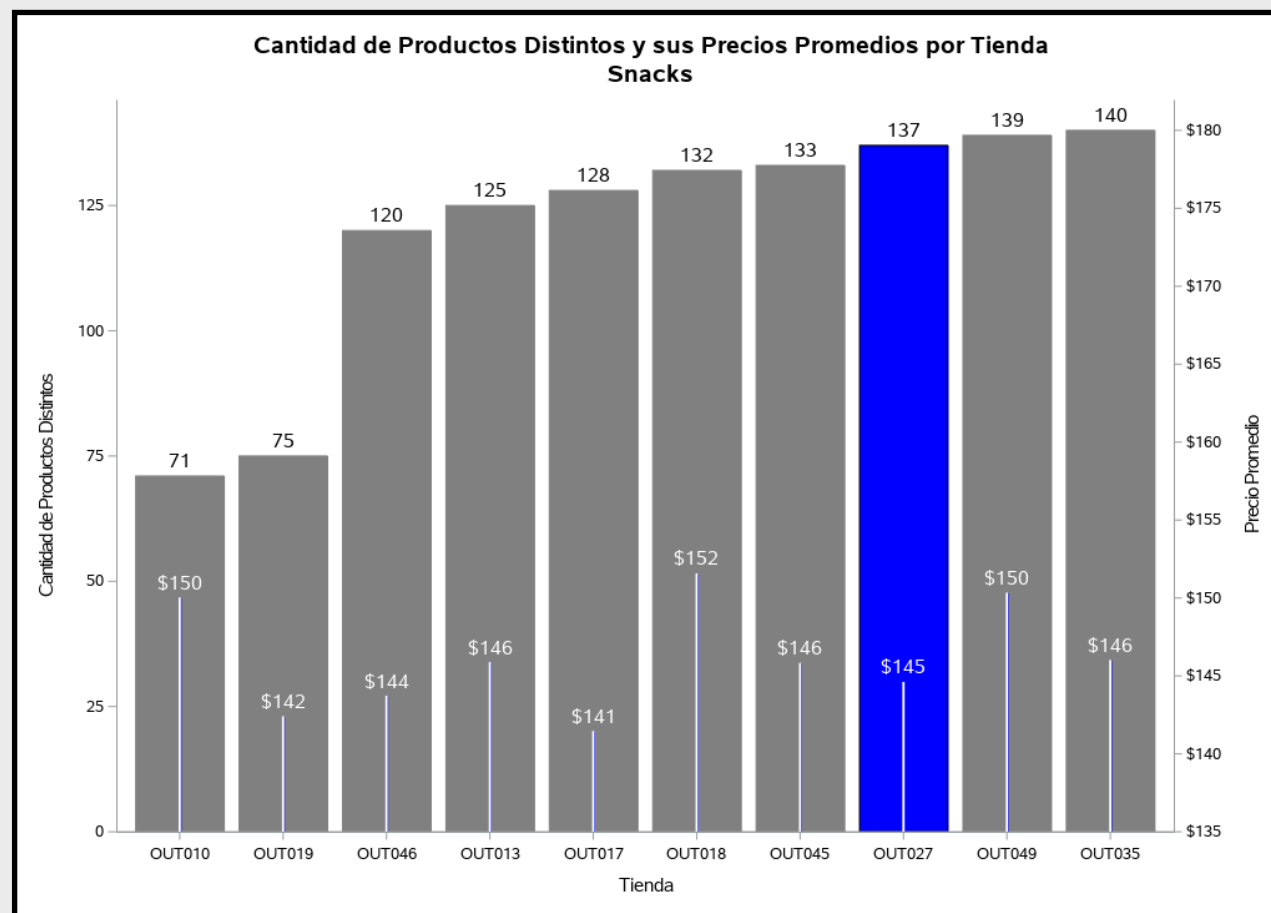
Conclusiones

Las frutas y vegetales con grasa tienen en promedio mayores ventas que los que no tienen grasa. Aunque este parámetro no es estadísticamente significativo. El modelo sugiere que **si aumenta la visibilidad** de los productos para este tipo de producto **entonces las ventas disminuyen 4,720**. Esto es porque la mayoría de estos productos tienen poca visibilidad y tienden a ser los más baratos y los mas comprados. También, si esto se agrupa por tipo de producto se puede ver una relación directa entre el precio del producto y la visibilidad que indica que a mayor precio mayor visibilidad. Por lo que se sospecha que exista multicolinealidad entre la visibilidad y el precio.

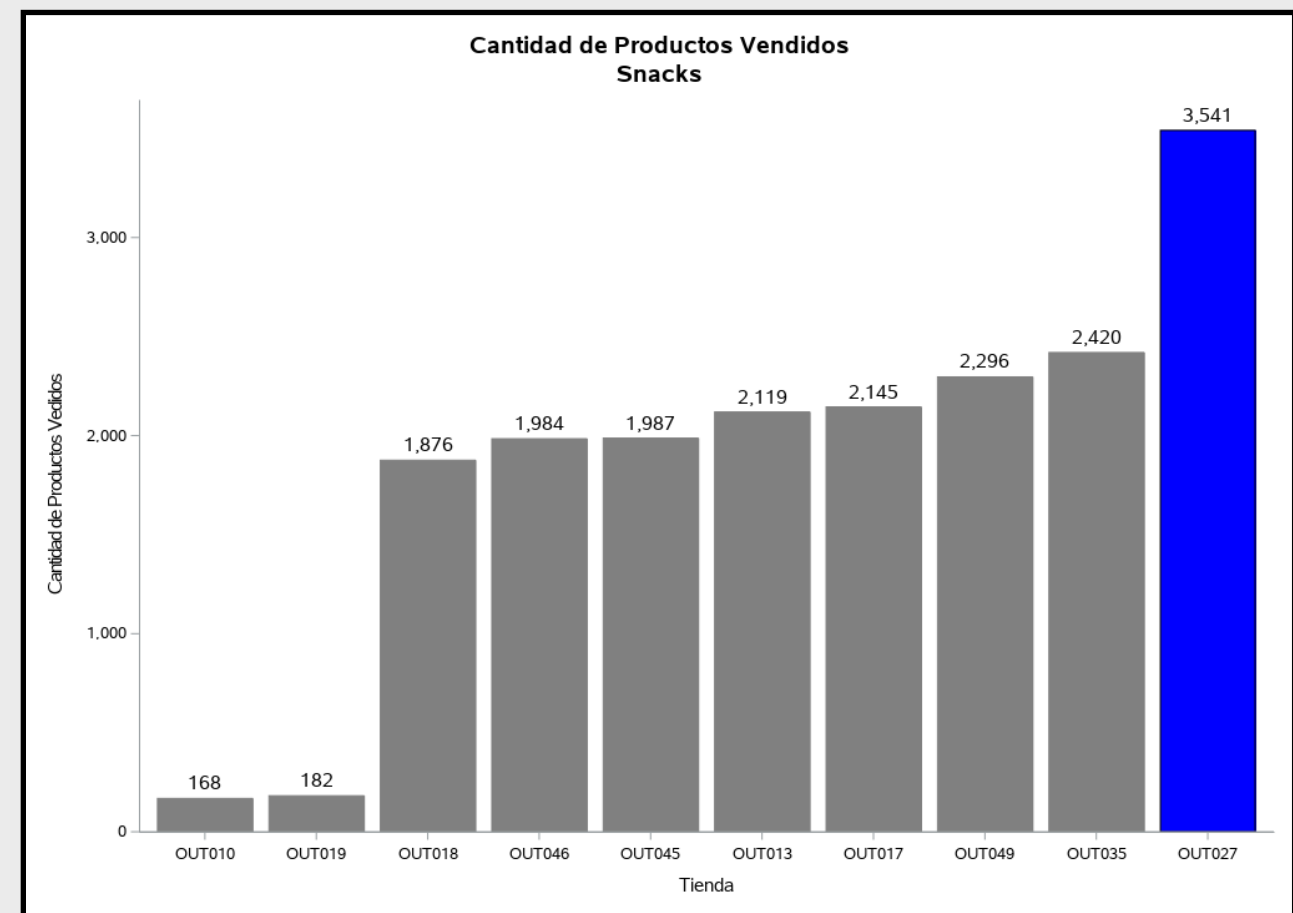
Si el precio aumenta en una unidad para este tipo de producto entonces las ventas aumentan 16.51. Este parámetro es estadísticamente significativo al 1%

Snacks

Cantidad de Productos Distintos y su precio promedio por tienda



Volumen Vendido por Tienda



El mismo patrón con la tienda OUT027 se repite, pero este vez el precio se encuentra en el segundo quintil. En este tipo de producto es interesante como el precio promedio de la tienda OUT010, que es un Grocery Store de tamaño pequeño se encuentra en el penúltimo quintil. Aunque la tienda OUT019, teniendo las mismas características que el OUT010 tiene un precio promedio que se encuentra en el primer quintil. Sugiriendo otros factores como la ubicación de la tienda que explican este comportamiento. OUT010 es Tier 3 y OUT019 es Tier 1.

Snacks

Qué Explican las Ventas de este tipo de producto, a nivel de producto

$$\begin{array}{ccccccccc} \textit{Ventas} = & 387.64983 & + & 2.21045\textit{Grasa} & - & 2.12395\textit{Peso} & - & 4524.89054\textit{Visibilidad} & + & 15.26951\textit{Precio} \\ & (0.0588) & & (0.9805) & & (0.8406) & & (0.000) & & (0.000) \end{array}$$

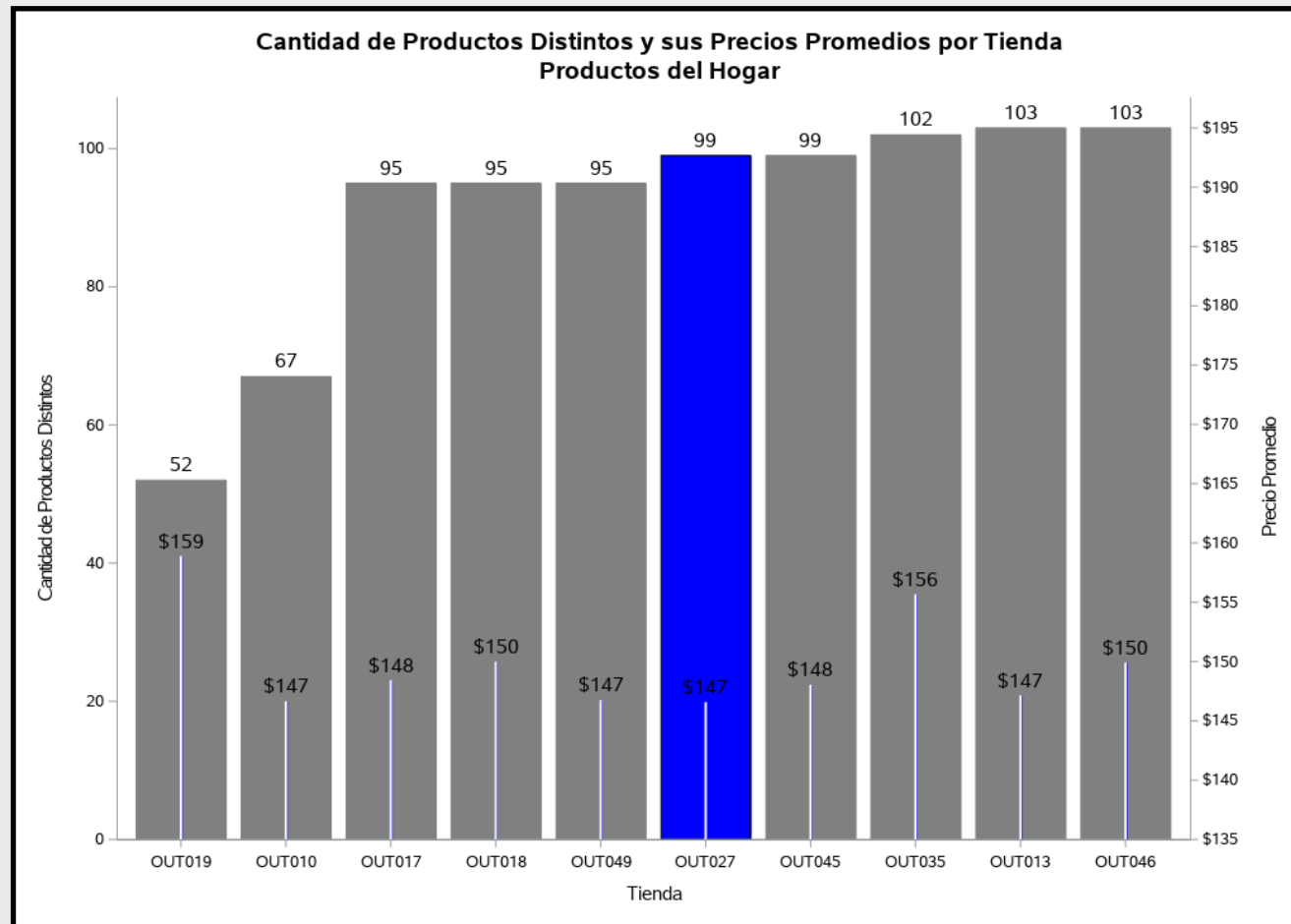
Este modelo explica un 28.82% de la variabilidad de las ventas de este tipo de producto y es estadísticamente significativo al 5% del nivel de significancia. Los residuales se distribuyen normal pero hay indicios de heterocedasticidad (ver diagnósticos del modelo adjunto).

Conclusiones

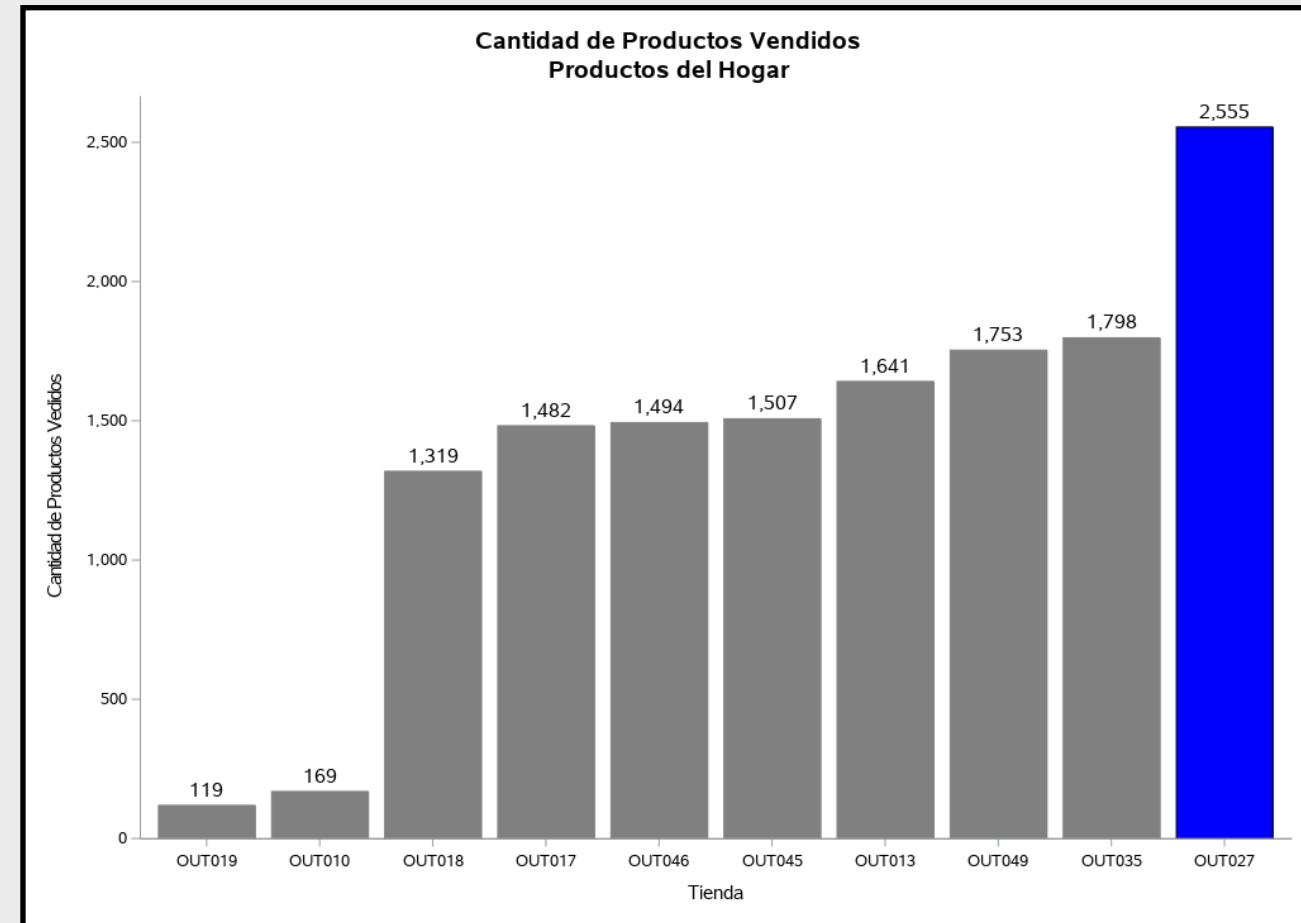
Los coeficientes para los snacks tienen el mismo comportamiento que para las frutas y los vegetales en cuanto a su signo. Solo la visibilidad y el precio son estadísticamente significativos para explicar la variabilidad de las ventas.

Productos del Hogar

Cantidad de Productos Distintos y su precio promedio por tienda



Volumen Vendido por Tienda



El mismo patrón con la tienda OUT027 se repite, pero este vez el precio se encuentra en el primer quintil.

Artículos del Hogar

Qué Explican las Ventas de este tipo de producto, a nivel de producto

$$\begin{array}{ccccccc} \textit{Ventas} = & 422.63976 & + & 5.19512\textit{Peso} & - & 5426.72084\textit{Visibilidad} & + & 14.22702\textit{Precio} \\ & (0.025) & & (0.6319) & & (0.000) & & (0.000) \end{array}$$

Este modelo explica un 33.90% de la variabilidad de las ventas de este tipo de producto a nivel de producto y es estadísticamente significativo al 5% del nivel de significancia. Los residuales se distribuyen normal pero hay indicios de heterocedasticidad (ver diagnósticos del modelo adjunto).

Conclusiones

Los coeficientes para los Artículos del Hogar tienen el mismo comportamiento que para los Snacks en cuanto a su signo. **Solo la visibilidad y el precio son estadísticamente significativos** para explicar la variabilidad de las ventas.

Comparando con los otros grupos, en este tipo de producto, el peso es el que tiene el coeficiente más grande y aunque no es estadísticamente significativo su valor P es menor que para los otros productos.

Conclusiones

La tienda OUT017 es una tienda de tamaño mediana que esta bien ubicada y dentro de esta ubicación mantiene una excelente combinación de precios ya que nunca es la que tiene el mayor ni el menor precio en estos tres productos analizados. Otras tiendas, que comparten esta buena ubicación (Tier 3) como **OUT018 y OUT010** pudieran beneficiarse de esta estrategia **reduciendo los precios de sus productos.**

A nivel de producto, solo la **visibilidad y el precio del producto** son estadísticamente significativas para explicar la variabilidad de las ventas.

Otras consideraciones

Valores vacíos

Sabiendo que el peso de no tiene un efecto (comprobado por un ANOVA $F=$). Para las 2 empresas fundadas en 1985, id_outlet -> y id_outlet -> Se imputo el peso de los productos utilizando la media del peso de los productos de todos

Modificación de los niveles de la variable categórica Item_Fat.

Otros Modelos.

(1)

$$CantidadVendida = Bo + Peso + Visibilidad + Precio$$

El Precio nunca es estadísticamente significativo, sugiriendo que el cambio en el precio no aumenta ni disminuye la cantidad vendida. La variable es la única que siempre es estadísticamente significativa y posee un coeficiente negativo, similar al obtenido en los modelos anteriores.

(2)

Agrupando por Outlet_Establishment_Year

$$\begin{aligned} \text{Item_Outlet_Sales} = & \text{Baking_Good} + \text{Bread} + \text{Breakfast} + \text{Canned} + \text{DairyFrozen} + \text{Fruits_Vegetables} + \\ & \text{Frozen} + \text{Hard_Drinks} + \text{Health_Hygiene} + \text{Household} + \text{Meat} + \text{Others} + \text{Seafood} + \text{Snack_Foods} + \\ & \text{Soft_Drinks} + \text{Starchy_Foods} + \text{fat1} + \text{item_weight_2} + \text{Item_Visibility} + \text{Item_MRP} \end{aligned}$$

(cada producto es una variable binaria)

Este modelo por grupo dio muchos parámetros sesgados ya que existen tiendas que no tienen las clasificaciones agrupadoras (mejoraría si se aumentara la muestra y si disminuyera la cantidad de informaciones incompletas utilizando métodos más avanzados): A partir de este se determino el objetivo y se construyó lo que muestra en las siguientes notas.