

Documentación Dashborad Frescura Cervecera

Realizado por: Santiago Pulido, Miller Patiño y Diego Peñaloza

Grupo: Los cuenteros de la MIAD

LINK DASHBOARD: <https://miad-visualizacion.streamlit.app/>

A continuación, encontrará la documentación para generación del dashboard de frescura de cerveza en el mercado colombiano. La medida de frescura es importante ya que indica cuánto dura un producto en el mercado y por consiguiente sirve por una parte para indicarle al consumidor cuales son los productos más frescos y por otra parte indicarles a los productores como se encuentra la saturación del mercado de acuerdo a los productos que produce. Esta segunda parte es muy relevante de cara a las empresas productoras ya que influencia directamente la decisión de producción. Productos con mayor frescura (menos días en el mercado) cuentan con una alta demanda, por lo que debe producirse más.

Datos utilizados

La base de datos tiene 9 variables, de las cuales 5 son categóricas, 2 son fechas y 2 son numéricas.

Adicionalmente a esta información inicial se utilizará el campo 'Referencia' para extraer las variables:

- Marca
- Tipo de envase
- Tamaño o capacidad del envase
- Retornabilidad

También se calcularán las siguientes variables:

- Fecha de producción
- Mes de producción
- Mes de medición

Finalmente se crea una variable categórica de nivel de frescura a partir de las reglas:

Tabla 1. Parámetros de construcción de la variable de clasificación de frescura

Marca	Edad de producto	Clasificación
1	< 90	Ideal
2		
3		
1	90 <= Edad_producto <= 120	Debe mejorar
2		
3		
1	> 120	Inaceptable
2		
3		
4	< 60	Ideal

5		
6		
7		
4	60 <= Edad_producto <= 100	Debe mejorar
5		
6		
7		
4	> 100	Inaceptable
5		
6		
7		

Que nos daría un total de 11 variables categóricas, 4 son fechas y 3 son numéricas.
A continuación, se realiza una breve descripción de las variables disponibles.

Tabla 2. Descripción de los campos incluidos en la base final

Columnas	Fuente	Tipo de dato	Descripción
Referencia	Original en la base	Categórica	Referencia de la marca, compuesta por la marca, el envase, el tamaño y el tipo
Establecimiento	Original en la base	Categórica	Lugar donde se realizó la encuesta
Ciudad	Original en la base	Categórica	Ciudad
Canal	Original en la base	Categórica	Grupo de lugar donde se realizó la encuesta
Fecha_vencimiento	Original en la base	Fecha	Fecha de vencimiento de la muestra evaluada
Edad_producto	Original en la base	Numérica	Se obtiene a partir: Fecha_encuesta - (Fecha_vencimiento - Vida_Util)
Lote	Original en la base	Categórica	Lote del producto evaluado
Vida_util	Original en la base	Numérica	Vida útil del producto evaluado
Fecha_encuesta	Original en la base	Fecha	Fecha en la que se realizó la evaluación
Marca	Construcción	Categórica	Marca del producto
Envase	Construcción	Categórica	Tipo de envase (lata / botella)
Tamaño	Construcción	Numérica	Tamaño del envase en ml

Retornabilidad	Construcción	Categórica	Tipo de retornabilidad del envase (retornable / no retornable)
Fecha de producción	Construcción	Fecha	Fecha vencimiento – vida útil
Mes de producción	Construcción	Categórica	Mes de la fecha de producción
Mes de medición	Construcción	Categórica	Mes de la fecha de medición
Nivel_Frescura	Construcción	Categórica	Construcción de acuerdo con las reglas indicadas arriba (Ideal, debe mejorar, inaceptable)

Procesamiento de los datos

A continuación, se indica cual es el proceso de obtención y transformación de datos para alimentar la visualización:

1. Fuente de Información: La fuente de información se obtiene desde el departamento de calidad de la empresa productora. Cada registro de esta base constituye una medida de frescura de un producto en específico. La información se diligencia por miembros del equipo de calidad.
2. La información se obtiene semanalmente de parte del grupo de calidad en un archivo llamado calidad.
3. Se debe modificar el nombre a 'Datos.xlsx', asegurarse que la información se encuentra en la hoja 'Export' y guardar la información en la ruta del reporte.
4. El proceso del proyecto toma el notebook de Python "ETL" con el cual se depura la información y se construye la estructura necesaria para el reporte.
5. Como resultado de este proceso se genera el dataframe sobre el cual se va a construir la visualización del reporte en la plataforma Streamlit.
6. Una vez actualizada la base de datos y generado el reporte se deben verificar las visualizaciones y estar atento a cambios abruptos en los datos que ameriten una revisión profunda del proceso.