



**Propuesta de modelo predictivo para determinar preferencias y factores que
influyen en el consumo de cerveza artesanal en mujeres**

Marcela Patricia Palencia Argel

Diana Marcela Bustos Astaiza|

Oscar Daniel Roncancio Lucas

Universidad Ean

Facultad de ingeniería

Especialización en gerencia de procesos de calidad e innovación

Especialización en Machine Learning

Bogotá, Colombia

21/07/2024

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres

Marcela Patricia Palencia Argel

Diana Marcela Bustos Astaiza

Oscar Daniel Roncancio Lucas

Trabajo de grado presentado como requisito para optar al título de:

Especialista en gerencia de procesos de calidad e innovación

Especialista en Machine Learning

Tutor (a):

Diana Paola Figueroa Hernandez

Modalidad:

Seminario de investigación

Universidad Ean

Facultad de ingeniería

Especialización en gerencia de procesos de calidad e innovación

Especialización en Machine Learning

Bogotá, Colombia

21/07/2024

Resumen

Dado el notable crecimiento en el consumo de cerveza artesanal, el presente trabajo de investigación propone un modelo para predecir preferencias y factores influyentes en el consumo de cerveza artesanal en la población femenina colombiana, utilizando una de las ramas del *Machine Learning* conocida como aprendizaje supervisado. Basado en la revisión documental se diseñó una encuesta como herramienta de recolección de datos, y se realizó una propuesta de modelo predictivo utilizando algoritmos como *Decision Trees* y *Random Forest* para analizar los datos en una futura implementación. Los hallazgos indican que las mujeres prefieren sabores suaves y afrutados, y enfrentan barreras como la percepción de la cerveza artesanal como una bebida predominantemente masculina. El modelo predictivo desarrollado podría proporcionar a las cervecerías artesanales herramientas basadas en datos para desarrollar productos y estrategias de marketing inclusivas. Con esta propuesta se busca generar datos reales sobre las preferencias de las mujeres y facilitar una mayor participación femenina en el mercado de la cerveza artesanal, contribuyendo al crecimiento y diversificación de este sector en Colombia.

Palabras clave: machine learning, cerveza artesanal, mujeres, inclusión, preferencias.

Abstract

Given the notable growth in the consumption of craft beer, this research proposes a model to predict preferences and influencing factors in its consumption among the Colombian female population, using a branch of Machine Learning known as supervised learning. Based on the literature review, a survey was designed as a data collection tool, and a predictive model proposal was made using algorithms such as Decision Trees and Random Forest to analyze the data in a future implementation. The findings indicate that women prefer light and fruity flavors and face barriers such as the perception of craft beer as a predominantly male beverage. The developed predictive model could provide craft breweries with data-driven tools to develop inclusive products and marketing strategies. This proposal aims to generate real data on women's preferences and facilitate greater female participation in the craft beer market, contributing to the growth and diversification of this sector in Colombia.

Keywords: machine learning, craft beer, women, inclusion, preferences

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres	3
---	---

Contenido

	Pág.
Lista de Figuras	5
Lista de Tablas	6
Problema de Investigación	7
Objetivos	10
<i>Objetivo general</i>	10
<i>Objetivos específicos</i>	10
Justificación	11
Marco Teórico	13
Metodología	20
Hallazgos	22
<i>Análisis sectorial</i>	22
<i>Estado del arte</i>	29
<i>Propuesta de modelo</i>	35
Discusión	52
Conclusiones y Trabajo Futuro	55
<i>Conclusiones</i>	55

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres	4
<i>Trabajo futuro</i>	56
Referencias	57

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres	5
---	---

Lista de Figuras

	Pag.
Figura 1. Mercado de cerveza en América del Sur, 2018-2028.....	22
Figura 2. Crecimiento cerveza premium, 2008-2027.....	23
Figura 3. Importación de las librerías principales.....	37
Figura 4. Código de las simulación de 10000 encuestas aleatorias.....	40
Figura 5. Ejemplo de visualización del df	43
Figura 6. Heatmap de las variables de la encuesta.....	45
Figura 7. Uso de GridSearch.....	47
Figura 8. Matriz de confusión.....	48
Figura 9. Feature importance	50

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres

6

Lista de Tablas

	Pág.
Tabla 1. Análisis de PESTEL sector industrial en Colombia.....	28
Tabla 2. Propuesta de encuesta para generar datos para el modelo predictivo.....	35

Problema de Investigación

El consumo de cerveza artesanal ha experimentado un crecimiento a nivel mundial, transformando los patrones tradicionales de consumo de cerveza. Este movimiento ha influido en cómo los consumidores perciben la cerveza, vinculándola cada vez más a su identidad y experiencias personales. A pesar de que la cerveza es la bebida alcohólica más consumida en términos de volumen a nivel mundial, existen diferencias significativas en los patrones de consumo entre géneros (Betancur, et al., 2020).

Gavilán & Muniesa (2023) identifican cinco tendencias clave en el consumo de cerveza artesanal que pueden guiar las estrategias de las cervecerías: Primero, el consumo ha aumentado especialmente entre menores de 30 años, lo que representa una gran oportunidad; Segundo, los consumidores valoran la procedencia y el carácter sostenible de la cerveza, buscando productos locales y éticos; Tercero, el sabor y la calidad son más importantes que el precio, lo que resalta la percepción premium de la cerveza artesanal; Cuarto, la diversidad es un factor clave, con los consumidores abiertos a experimentar con nuevas marcas; Finalmente, categoría de cerveza artesanal en general. Estas tendencias ofrecen a las cervecerías artesanales la posibilidad de desarrollar productos y estrategias de marketing que respondan a mejor a las demandas del mercado, especialmente entre mujeres, potenciando así su ventaja competitiva.

Numerosos estudios han demostrado que los hombres son más propensos a consumir cerveza que las mujeres (Betancur, et al., 2020). Esta tendencia puede deberse a que los hombres asocian el consumo de cerveza con la asertividad social y percepciones positivas del entorno. Además, el volumen por porción de la cerveza suele ser mayor en comparación con otras opciones alcohólicas, lo cual puede influir en las preferencias de consumo por género. Sin embargo, estas diferencias también pueden estar influenciadas por la manera en que la cerveza es publicitada, y las experiencias subjetivas y emociones

asociadas al sabor de la cerveza pueden variar significativamente entre diferentes grupos de edad y género (Betancur, et al., 2020).

La narrativa dominante en torno a la cerveza ha sido históricamente una que glorifica el consumo como un acto de masculinidad y un ritual social, a menudo excluyendo a las mujeres de manera hostil. No obstante, el auge de la cerveza artesanal ha propiciado una historia más inclusiva, enfatizando la autenticidad, la diversidad y la comunidad local. A pesar de esta perspectiva progresista, las mujeres siguen enfrentando estereotipos de género, sexismo y percepciones de ilegitimidad tanto como consumidoras como productoras de cerveza (Kuehn & Parker, 2018).

En Colombia, el mercado de la cerveza artesanal está en crecimiento, pero el consumo entre mujeres sigue siendo un área poco explorada. Las cervecerías artesanales tienen la oportunidad de capitalizar este significativo segmento del mercado, ya que aproximadamente el 50% de la población mundial está representada por mujeres (Rivaroli, et al., 2020). A pesar de un enfoque creciente en la calidad y la sostenibilidad, persisten barreras relacionadas con estereotipos de género. Esto destaca la necesidad de crear estrategias para identificar los factores que influyen en el consumo de cerveza artesanal por parte de las mujeres y ajustar los productos que ofrece el mercado actualmente para que se alineen con estas preferencias.

Pregunta de investigación.

El presente proyecto pretende emplear el modelo de *Machine Learning* (ML) y la revisión de literatura para resolver la siguiente pregunta de investigación: ¿Cuáles son las preferencias y barreras que influyen en el consumo de cerveza artesanal en la población femenina, y cómo puede un modelo predictivo ayudar a identificar estas preferencias? De esta forma, se busca brindar herramientas a las empresas productoras de cerveza

Propuesta de modelo predictivo para determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres

9

artesanal en Colombia para fomentar inclusión de la población femenina en el mercado, así como incentivar el estudio académico en el área de ML.

Objetivos

Objetivo general

Proponer un modelo predictivo que permita determinar preferencias y factores que influyen en el consumo de cerveza artesanal en mujeres, para ofrecer una herramienta de información a las empresas productoras de cerveza artesanal y fomentar la inclusión de la población femenina en este mercado.

Objetivos específicos

- Explorar las preferencias de sabor y estilo de cerveza artesanal en mujeres para desarrollar productos que se alineen mejor con sus gustos y aumentar su participación en el mercado de la cerveza artesanal.
- Determinar las barreras y desafíos percibidos por las mujeres en el consumo de cerveza artesanal para crear estrategias que eliminen estos obstáculos y promuevan una mayor inclusión de las mujeres en este mercado.
- Proponer una guía de implementación de modelo de ML para determinar los factores que influyen en la elección de cerveza artesanal en mujeres y así guiar a las empresas productoras de cerveza artesanal en la creación de productos que se alineen con las preferencias y necesidades de la población femenina.

Justificación

El crecimiento del consumo de cerveza artesanal a nivel mundial ha transformado los patrones tradicionales de consumo, vinculando esta bebida cada vez más con la identidad y experiencias personales de los consumidores (Cunha, et al., 2023). A pesar de su popularidad, existen diferencias significativas en los patrones de consumo entre géneros. Mientras que los hombres suelen asociar la cerveza con asertividad social y percepciones positivas del entorno, las mujeres enfrentan estereotipos y barreras en el consumo de cerveza (Spence, 2019a), situación que se refleja también en el mercado colombiano.

El mercado de la cerveza artesanal en Colombia está en expansión, con un aumento notable en la oferta y demanda de este tipo de cerveza (Forbes, 2023). Sin embargo, el consumo entre mujeres sigue siendo un área poco explorada. Las cervecerías artesanales tienen la oportunidad de capitalizar este segmento de mercado significativo, ya que aproximadamente el 50% de la población mundial está representada por mujeres (Rivaroli, et al., 2020).

Los estudios demuestran que el mercado femenino presenta un potencial considerable para el crecimiento del sector de la cerveza artesanal (Horowiec, 2022). La inclusión de la población femenina en este mercado no solo podría aumentar las ventas, sino también promover una mayor diversidad y equidad de género dentro de la industria cervecera. Además, identificar y comprender las preferencias y factores que influyen en el consumo de cerveza artesanal por parte de las mujeres puede ayudar a las cervecerías a desarrollar productos y estrategias de marketing que respondan mejor a las demandas de este grupo de consumidores.

Este proyecto de investigación propone un modelo predictivo utilizando *Machine Learning* para determinar las preferencias y factores que influyen en el consumo de

cerveza artesanal en la población femenina. Este enfoque permitirá a las cervecerías artesanales obtener datos relevantes y precisos que les ayudarán a crear estrategias más inclusivas y productos alineados con las preferencias de las mujeres. Al hacerlo, se busca fomentar la inclusión de la población femenina en el mercado de la cerveza artesanal, contribuyendo así al desarrollo económico y social de este sector.

Marco Teórico

Machine Learning

La evolución del ser humano a lo largo del tiempo ha venido de la mano con la necesidad de resolver problemas y de cuestionarse, impulsando el desarrollo de herramientas que permiten optimizar labores de la cotidianidad. En esta era, una de esas herramientas importantes ha sido la programación. En términos sencillos, la programación se basa en escribir secuencias de códigos que pueden ayudar a resolver fácilmente tareas o problemas. Estas secuencias de código las interpreta un dispositivo, como puede ser un computador, y tienen que ser dadas en una sintaxis específica para que el dispositivo pueda entenderlas (Navone, 2023).

Python es uno de los medios que tiene el computador para entender estas secuencias de código. Fue creado por Guido Van Rossum entre finales de 1980 e inicios de 1990. La sintaxis manejada por *Python* tiende a ser poco compleja además de poseer de base muchas librerías que son de gran utilidad para resolver casi todos los problemas que pueda tener un programador. Gracias a la disponibilidad de múltiples librerías para procesamiento y visualización de datos matemáticos, así como su manipulación estadística (Challenger, Díaz, & Becerra, 2014), *Python* se ha convertido en una herramienta fundamental para lo que hoy se conoce cómo *Machine Learning*.

La capacidad de las máquinas para aprender y adaptarse es una de las innovaciones más significativas actualmente, pero para darle un uso adecuado es necesario entenderlo conceptualmente. El aprendizaje automático (*Machine Learning*) es un campo de la inteligencia artificial (IA) y la informática que se especializa en utilizar datos y algoritmos para que la IA pueda simular el aprendizaje humano, además de mejorar su precisión gradualmente como explica IBM (s.f.). Su aplicación principal es permitirle a las máquinas realizar tareas que normalmente solo podrían ser realizadas por el ser humano. Antes de continuar, es importante recalcar que ML no es lo mismo que IA. La IA busca desarrollar

máquinas que piensen y tengan habilidades similares a las de los humanos, mientras que ML se trata específicamente de usar conjuntos de datos y algoritmos para enseñarle a las máquinas a adquirir este conocimiento tal como explica Bobadilla (2020).

En la actualidad, existen 4 ramas de ML: el aprendizaje supervisado, el aprendizaje no supervisado, el aprendizaje semi supervisado y el aprendizaje reforzado (IBM, s.f.a). Para el presente trabajo, se hará el enfoque en la rama del aprendizaje supervisado. El aprendizaje supervisado se caracteriza principalmente porque usa información que ya está correctamente clasificada y tiene una etiqueta normalmente conocida con el nombre de la variable objetivo, esto con fines de que el algoritmo aprenda a predecir y clasificar datos con precisión (IBM., s.f.a).

Así mismo, cada una de las ramas que se mencionaron previamente contienen múltiples algoritmos, cada uno con sus particularidades. Para este trabajo se implementarán 2 tipos de algoritmos uno derivado del otro. El primer algoritmo es el *Decision Trees*, un modelo cuyo propósito es la clasificación y la regresión de datos, funciona creando reglas de decisión a partir de las características de una base de datos dada y su propósito es encontrar, con esas reglas, la mejor forma de clasificar un dato. El *Decision Trees* es uno de los algoritmos más simples de entender y no requiere una preparación de datos demasiado compleja. (Scikit-Learn, s.f.). Por su parte, el *Random Forest*, derivado de *Decision Trees*, es un algoritmo que permite generar múltiples árboles de decisión con cierto grado de aleatoriedad (IBM, s.f.b).

Cómo se mencionó previamente, estos 2 algoritmos serán usados en el desarrollo del presente trabajo, como propuesta para predecir y clasificar cuáles son los gustos y factores potenciales de la población femenina a la hora de escoger una cerveza artesanal y brindar así datos relevantes a las productoras para lograr evaluar este segmento poblacional poco explorado.

Cerveza artesanal

Según el Decreto 1686 de 2012 del Ministerio de Justicia de Colombia, la cerveza artesanal se define como una bebida que tiene entre 2.5 y 12 grados de alcohol, obtenida a través de un proceso de fermentación alcohólica con levaduras. El mosto para esta bebida se elabora con cebada malteada o extracto de malta, cereales malteados o no malteados, cebada malteada con frutas, jugos o pulpa de frutas, cebada malteada con granos no cereales, lúpulos, agua potable o microorganismos de uso comercial. La cerveza artesanal se distingue por ser elaborada en pequeñas cantidades con métodos tradicionales y enfocada en la calidad y diversidad de sabores y estilos, en contraste con las producciones industriales más masivas (Rao, 2008).

La cerveza artesanal se compone de cuatro ingredientes principales: malta, levadura, agua y lúpulo (Baiano, 2020). Según la levadura, las cervezas se pueden clasificar en dos estilos: ale en donde las levaduras operan a temperaturas más altas generando un mayor contenido de ésteres (cervezas más afrutadas), una fermentación más rápida (mayor reproducción celular) y una mayor formación de alcoholes superiores (sabor más alcohólico); y lager donde las levaduras funcionan a temperaturas de fermentación más bajas, con una fermentación más lenta y riesgosa y una menor producción de ésteres y alcoholes superiores (Cunha, et al., 2023).

Desde sus inicios en el Reino Unido, la tendencia de la cerveza artesanal se ha expandido globalmente, influenciando mercados en Norteamérica, Europa y más allá. En Estados Unidos, por ejemplo, el movimiento de la cerveza artesanal tomó fuerza en la década de los 80 y ha crecido exponencialmente desde entonces, con miles de cervecerías artesanales operando en todo el país. Este crecimiento ha sido impulsado por una creciente demanda de los consumidores por productos locales, auténticos y de alta calidad, así como por una mayor apreciación de la diversidad y complejidad de los sabores de la cerveza artesanal (Deloitte, 2017).

La cerveza artesanal en Colombia

Según Forbes (2023), la industria de la cerveza artesanal en Colombia ha mostrado un crecimiento significativo en los últimos años, con un crecimiento en ventas del 51% y un notable aumento tanto en la oferta como en la demanda. Actualmente, el país cuenta con al menos 250 cervecerías artesanales. Este fenómeno se ve impulsado por la creciente demanda de productos únicos y de alta calidad, evidenciado en el consumo per cápita de cerveza, que según datos de Euromonitor (2024), alcanzó los 49.9 litros en 2022, frente a los 46.5 litros del año anterior.

De acuerdo con Forbes (2023), el mercado colombiano de cerveza artesanal se distingue por su diversidad en ingredientes y métodos de producción. Existen propuestas que utilizan insumos locales como la yuca, papa, quinua, café, durazno, copoazú y asaí. Además, se elaboran cervezas con agua de fuentes naturales como el volcán Galeras, o con productos autóctonos como la panela de Villeta y el cacao de Tumaco. Esta diversidad no solo satisface las preferencias de los consumidores, sino que también fomenta el desarrollo económico local al utilizar productos y recursos regionales.

Adicionalmente, Forbes (2023) menciona que, aunque la cerveza artesanal representa solo el 1% del mercado cervecero en Colombia, su potencial de crecimiento es considerable. En comparación, en Estados Unidos la cerveza artesanal constituye el 20% del total de ventas de cerveza (Forbes, 2023). Este dato subraya la posibilidad de expansión del sector en el país. Carulla, una de las principales cadenas de supermercados en Colombia, reportó un aumento del 51% en las compras de cervezas artesanales en el 2022. A nivel general, el consumo de cerveza en el país creció un 8% en 2022, pasando de 2,392 millones de litros a 2,581 millones de litros; este crecimiento es más notable en ciudades como Bogotá, Medellín, Cartagena y Cali, con Bogotá concentrando el 44% de las ventas (Euromonitor, 2024).

Factores que afectan las preferencias del consumidor

Las preferencias de consumo de cerveza artesanal están influenciadas por una combinación de factores intrínsecos y extrínsecos, siendo los atributos sensoriales como el sabor y el aroma determinantes cruciales (Spence, 2019^a). Estos atributos no solo definen la calidad percibida del producto, sino que también influyen en la disposición a pagar (WTP) y en la fidelización de los consumidores (Spence, 2019a). Además, la conexión emocional con la autenticidad del producto juega un papel importante en la preferencia y elección de la cerveza por parte de los consumidores (Rao, 2008).

Las cervecerías industriales suelen elaborar cervezas uniformes, mientras que las pequeñas pueden crear productos artesanales con mayor diversidad de sabores. Esta situación ha dado lugar al surgimiento y crecimiento del mercado de la cerveza artesanal, generando lo que se conoce como la "revolución del sabor" como respuesta a la estandarización (Baiano, 2020).

El análisis de las preferencias de los consumidores de cerveza artesanal revela varios factores clave. La diversidad de sabores es un aspecto crucial, valorado por el 60% de los consumidores. La experiencia y la autenticidad de la cerveza artesanal también son altamente apreciadas, con un 31% de los consumidores destacando estos aspectos. Además, la percepción de la cerveza artesanal como un producto elaborado con ingredientes naturales y métodos tradicionales añade valor a la experiencia del consumidor. Otros factores como el grado de alcohol y el estatus asociado al consumo de cervezas artesanales también influyen en la decisión de compra (Deloitte, 2017).

Gabrielyan, et al. (2014) estudiaron los factores sensoriales que afectan la disposición de pago en el mercado de cerveza estadounidense, encontrando que el sabor y el amargor tienen un impacto positivo, siendo el sabor el que tiene el mayor impacto; mientras que la apariencia y el aroma no variaron mucho y no representan un impacto significativo de las

características intrínsecas de las cervezas. Este estudio concluyó que el sabor es la base principal para las expectativas de calidad de los consumidores y sus decisiones de repetir la compra de un producto, siendo el factor diferenciador más importante para las cervezas artesanales.

Adicionalmente, la demanda de la cerveza artesanal está influenciada por su precio, el precio de los productos alternativos, el nivel salarial, las características y el estatus social de los consumidores, así como por la variedad y la experiencia de consumo (Nave, et al., 2021). Según Baiano (2020), los consumidores de cerveza artesanal presentan características comunes en todo el mundo: predominan los hombres, relativamente jóvenes, con un alto nivel educativo, ingresos medios-altos y un interés en productos de calidad y de producción local.

El éxito de la industria de la cerveza artesanal puede deberse a su estrategia de presentarse como un movimiento social que se opone a las grandes empresas cerveceras (Withers, 2017). Según Rao (2008), los movimientos identitarios, impulsados por un sentido de comunidad, surgen para desafiar a las organizaciones o categorías dominantes y buscan crear nuevas identidades colectivas a través de la formación de nuevas organizaciones que promueven la democracia, la participación y el empoderamiento. De esta forma, la inclusión de la población femenina en este mercado en auge puede presentarse como una estrategia innovadora para incentivar un sentido de pertenencia y empoderamiento en un nicho de mercado aun sin explorar.

Consumo de cerveza en las mujeres

Calvo-Porrá et al. (2018) indican que las mujeres son menos asociadas con el consumo de cerveza que los hombres y muestran patrones de consumo diferentes en frecuencia, preferencias de sabor y contexto de consumo. Los estudios demográficos indican que los patrones de consumo de cerveza entre mujeres están influenciados por factores como la

edad, el nivel educativo y el entorno social (Ayres et al., 2011). Las mujeres jóvenes y de mediana edad tienden a preferir bebidas alcohólicas que ofrecen una experiencia sensorial compleja y una conexión emocional con el producto, aspectos que la cerveza artesanal puede satisfacer mediante su variedad de estilos y sabores auténticos (Gómez-Corona et al., 2017).

En un estudio realizado entre 300 mujeres consumidoras de cerveza en España, se encontró que todas las participantes consideraron que la cerveza es una bebida inclusiva, no exclusiva de género (Ambar, 2019). Sin embargo, persisten estereotipos en su servicio, pues tres de cada cuatro mujeres han recibido refrescos en lugar de la cerveza que pidieron cuando están acompañadas por hombres, y una de cada tres ha experimentado extrañeza al pedirla estando sola (Ambar, 2019). A pesar de estos desafíos, las mujeres muestran una fuerte disposición hacia la experimentación con cervezas nuevas y locales, con un alto porcentaje dispuesto a probar nuevas variedades tanto en viajes como cuando se les ofrece (Ambar, 2019).

De este mismo estudio, las preferencias de las mujeres encuestadas indicaron una clara preferencia por la cerveza rubia tradicional, aunque las variedades con limón y sin alcohol tienen menos popularidad (Ambar, 2019). El consumo se centra en el gusto por su sabor y la percepción de que es una bebida más apetecible que otras alternativas. Por otro lado, las preferencias en el momento de consumo incluyen la tarde en contextos sociales al aire libre y como acompañamiento de aperitivos, con una preferencia significativa por beberla directamente de la botella frente a usar una copa (Ambar, 2019). Aunque históricamente menos inclinadas hacia el consumo de cerveza, las mujeres están ganando protagonismo en el mercado de la cerveza artesanal debido a las campañas de marketing más inclusivas y productos adaptados a sus preferencias (Spence, 2019b).

Metodología

El presente trabajo de investigación manejó una metodología mixta, uniendo las metodologías cualitativa y cuantitativa con el fin de aprovechar las ventajas que ambas pueden proporcionar (Arias, 2020). La metodología cualitativa no puede ser cuantificada, sino que busca explicar fenómenos a través de análisis, interpretaciones o evaluaciones que se recogen de diversas fuentes, cómo pueden ser entrevistas, revistas y documentos académicos. Por su parte, la metodología cuantitativa es la contraposición de la cualitativa y busca determinar leyes universales para explicar fenómenos netamente a partir de datos numéricos (Sampieri, Collado, & Baptista, 2014).

Se realizó una investigación de tipo descriptiva, la cual busca recopilar, analizar y presentar datos de situaciones que suceden en la vida real, permitiendo posteriormente observarlas y describir posibles características. En palabras resumidas, busca ofrecer una descripción detallada de la situación (Stewart, S.f). Para este caso, se tomó como situación el consumo de cerveza artesanal en la población femenina, usándose fuentes secundarias para analizar las preferencias y factores que tienen mayor influencia.

De esta forma, para la parte cualitativa se utilizó como instrumento de recolección de información la revisión documental de fuentes secundarias, entre ellas bases de datos institucionales (Science Direct, Web of Science, Scopus y Emis), plataformas de artículos científicos (Scielo, Redalyc, Google Scholar y Dialnet) y documentos producidos por diversas instituciones nacionales o asociaciones empresariales. Los estudios más relevantes fueron seleccionados refinando la información según las publicaciones de la última década (2014-2024) y los artículos seleccionados del mayor número de investigaciones basadas en el tema central de este estudio.

Para la parte cuantitativa, se diseñó una encuesta y se desarrolló una guía de implementación de un modelo de Machine Learning con datos simulados para predecir

preferencias y factores de consumo de cerveza artesanal entre mujeres. Esta guía se realizó con la finalidad de brindar una herramienta que permita a productores de cerveza artesanal recolectar datos reales y predecir preferencias y factores que influyen en el consumo de cerveza artesanal entre las mujeres en Colombia, así como una base para futuras investigaciones en esta área.

Hallazgos

Análisis sectorial

Mercado de cerveza en América latina

De acuerdo con Euromonitor (2024), en 2023, América del Sur se posicionó como la cuarta región más grande en el mercado global de la cerveza, con un valor de 18.100 millones de dólares, representando el 8,6% del mercado mundial, precedida por América del Norte con un 24,5% y seguida por Europa del Este con un 5,4%. Ese mismo año, la cerveza constituyó el 71,0% del mercado de bebidas alcohólicas en la región, con un consumo per cápita de 35,4 dólares, superior al promedio mundial de 26,6 dólares per cápita. El mercado cervecero sudamericano representó el 0,38% del PIB de la región. Entre 2018 y 2028, se proyecta un crecimiento interanual constante (ver figura 1), consolidando la importancia de la cerveza en la economía y cultura de América del Sur.

Figura 1. Mercado de cerveza en América del sur, 2018-2028.



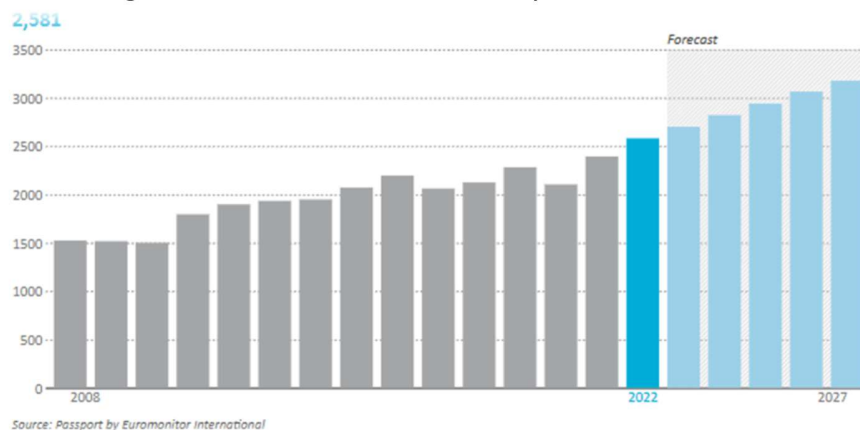
Tomado de Euromonitor (2024)

Mercado de la cerveza en Colombia

Según Euromonitor (2024), en 2022, el mercado de la cerveza en Colombia experimentó un notable aumento del 8% en el volumen total de ventas, alcanzando los 2.600 millones de litros, con la cerveza negra destacándose como la categoría de mayor crecimiento,

incrementando sus ventas en un 23%. Bavaria SA lidera el mercado con una cuota del 88% del volumen total. Se anticipa que el volumen de ventas crecerá a una tasa anual compuesta del 4%, alcanzando los 3.200 millones de litros. A pesar del aumento de costos y precios, el consumo general no se ha visto afectado, con consumidores optando por marcas más asequibles como Costeña Bacana, que ha ganado participación significativa en regiones adicionales tras su expansión. Las cervezas premium, aunque aún un segmento pequeño, continúan creciendo rápidamente (ver figura 2) gracias a presentaciones más pequeñas disponibles en tiendas locales. Sin embargo, la inflación y el aumento del costo de vida pueden afectar negativamente la confianza del consumidor y su poder adquisitivo en 2023, desacelerando el crecimiento del volumen y cambiando ligeramente la demanda hacia el canal off-trade (Euromonitor, 2024).

Figura 2. Crecimiento de cerveza premium, 2008-2027.



Tomado de Euromonitor (2024)

Análisis del sector - Análisis PESTEL

El análisis PESTEL tiene como propósito identificar los principales factores políticos, económicos, sociales, tecnológicos, ecológicos y legales que impactan el sector industrial.

Factores políticos

Estabilidad gubernamental: El actual gobierno está promoviendo la convocatoria de una asamblea constituyente, el cual propone 9 puntos: la implementación del acuerdo de paz del 2016, reforma judicial, reordenamiento territorial, garantizar las condiciones básicas de vida, recuperar los objetivos de la constitución del 91 contra reformados, en materia de priorizar la educación pública y la reforma agraria, garantizar que la política monetaria, manteniendo la independencia del Banco de la República, priorice el empleo y la producción, cambio climático, expedir el estatuto del trabajo y separar la política de la financiación privada y reparación y reconciliación de las víctimas (Murcia, 2024). Estas condiciones están generando desconfianza en la estabilidad del país, ya que por el afán de gobierno en implantar las reformas de salud, pensión y laboral se está buscando diferentes vías para su legalización.

Leyes fiscales: Se propone para el 2025 una nueva reforma tributaria, que de acuerdo con lo expresado por el presidente Gustavo Petro, tiene como objetivo principal bajar la tasa de renta de las empresas del 35% al 30% y subir la renta de personas naturales (Lombo Delgado, 2024).

Factores económicos

Tasas de interés: En julio de 2024 se presentó una disminución en la tasa de interés, la cual se encuentra en 11.25%. La decisión refleja una disminución equivalente en la tasa de interés real, tanto anticipada como retrospectivamente, dada la estabilidad de la inflación y las expectativas inflacionarias en mayo. Una reducción más amplia de la tasa de interés podría no ser sostenible a largo plazo, debido a los riesgos que persisten para alcanzar la meta de inflación anual en el plazo establecido. Entre estos riesgos, destacan la resistencia a la baja de la inflación en ciertos servicios, como arrendamientos y tarifas de energía. También señalan el riesgo asociado con los precios de los alimentos, que

vieron un aumento en mayo. Además, advierten que, aunque las expectativas de inflación muestran una tendencia a la baja, todavía están por encima de la meta para los plazos en los que la Junta Directiva del Banco de la República pretende alcanzarla (Banco de la República, 2024).

Inflación: En mayo de 2024, la inflación anual se mantuvo en 7,2%, mientras que la inflación subyacente bajó de 6,4% a 6,1%. La inflación de los alimentos aumentó de 3,0% a 4,4% debido al impacto de El Niño en ciertos productos perecederos. A excepción de los alimentos, la inflación en otros segmentos del IPC disminuyó. En cuanto a las expectativas, la encuesta de junio del Banco de la República mostró una reducción de la inflación esperada a un año, de 4,6% a 4,3%, y estabilidad en 3,8% para fin de 2025 (Banco de la República, 2024).

Tasa de desempleo: Para el mes de mayo de 2024, la tasa de desocupación del total nacional fue 10,3%, mientras que en el mismo mes de 2023 fue 10,5%. La tasa global de participación se ubicó en 64,1% y la tasa de ocupación en 57,5%, estas tasas en mayo de 2023 fueron 63,9% y 57,2%, respectivamente (DANE, 2024).

Tendencias del PIB: El crecimiento del PIB del 0,9% en el primer trimestre de 2024 superó las expectativas del Banco de la república, que eran del 0,3%. Este aumento se debió principalmente a una mejora en la demanda externa neta, con una reducción en las importaciones y un incremento en las exportaciones. En el segundo trimestre, se prevé que la economía continuara recuperándose, respaldada por indicadores como el ISE, el crédito al consumo, los pagos electrónicos, y las estadísticas de exportaciones, importaciones, transporte de carga y pasajeros aéreos (Banco de la República, 2024).

Factores Socio culturales

En su mayoría, las investigaciones señalan que los hombres son los más propensos al consumo, pues mientras ellos lo hacen al menos una vez a la semana, las mujeres

consumen una vez al mes. Así mismo, los lugares constituyen otro factor que influye en el consumo de alcohol; entre los que se identifican se encuentran discotecas, casas, bares y parques (Calle & Giraldo, 2020). Los factores socio culturales será abordados en mayor profundidad en la sección de estado del arte.

Factores Tecnológicos

Nivel de desarrollo tecnológico y uso de software: Colombia tiene disponibilidad en universidades para aprender *Python*, *numpy* para operaciones matemáticas y estadísticas más complejas, *matplotlib* y *seaborn* para visualización de datos y *sklearn*, para procesamiento de datos y creación de modelos de ML (Las Dos Orillas, 2024).

Infraestructura telecomunicaciones – Internet: En Colombia, la cobertura de Internet fijo en los hogares es del 83,2% en áreas urbanas, mientras que en zonas rurales solo llega al 46,3%. Por otro lado, las zonas rurales tienen una mayor proporción de conectividad móvil en los hogares, con un 62,1%, en comparación con el 40,7% en las zonas urbanas. Además, los hogares que disponen de ambos tipos de conexión son más frecuentes en las áreas urbanas, con un 23,9%, frente al 8,5% en las zonas rurales (Suárez Bernal, 2024).

Factores Ambientales

Los factores ambientales son especialmente relevantes para las empresas cerveceras, dado el impacto significativo que la producción de cerveza puede tener sobre el medio ambiente. La industria cervecera debe cumplir con una serie de normativas que regulan el uso de recursos naturales, la gestión de residuos y la protección del entorno. En Colombia, diversas leyes y regulaciones establecen estándares para la gestión sostenible de recursos hídricos, la disposición de residuos y el manejo de suelos contaminados. Estas leyes no solo buscan minimizar el impacto ambiental de las actividades cerveceras, sino también

fomentar prácticas sostenibles y responsables. A continuación, se detallan las principales normativas:

- **Ley 1333 de 2009:** Conocida como el Código de Régimen Sancionatorio Ambiental y de los Recursos Naturales, esta ley establece las normativas para la sanción de infracciones ambientales y la protección de los recursos naturales.
- **Ley 373 de 1997:** La Ley de Aguas regula el uso, manejo y conservación de los recursos hídricos, estableciendo directrices para la gestión sostenible del agua.
- **Ley 1259 de 2008:** La Ley de Residuos y Suelos Contaminados se enfoca en la gestión de residuos y la rehabilitación de suelos afectados por contaminación.
- **Ley 1672 de 2013:** Esta ley, denominada Ley de Gestión Integral de Residuos y Residuos Peligrosos, establece un marco para la gestión de residuos, incluyendo aquellos considerados peligrosos.
- **Ley 99 de 1993:** La Ley de Medio Ambiente crea la estructura básica para la protección ambiental y el manejo sostenible de los recursos naturales en Colombia.
- **Ley 1733 de 2014:** Introduce el Sistema Nacional de Cambio Climático (SISCLIMA), que coordina y regula las políticas relacionadas con el cambio climático en el país.

Factores Políticos

En Colombia, la regulación de la industria de bebidas alcohólicas se rige por varias leyes y decretos que establecen los parámetros para su fabricación, distribución y control. A continuación, se presentan las normativas clave:

- **Constitución Política de Colombia, Artículo 333:** Este artículo establece la libertad económica y la iniciativa privada como principios fundamentales en Colombia, permitiendo el ejercicio de la actividad económica dentro de los límites del bien común. También destaca la función social de las empresas y el compromiso del Estado para fortalecer las organizaciones solidarias y promover el desarrollo empresarial.
- **Decreto 3192 de 1983:** Este decreto reglamenta parcialmente el Título V de la Ley 9 de 1979, que abarca la regulación de fábricas de alcohol y bebidas alcohólicas. Establece directrices para la fabricación, hidratación, envasado, distribución,

exportación, importación y venta de estos productos, y define mecanismos de control en todo el territorio nacional.

- **Ley de Financiamiento, 1943 de 2018:** La Ley de Financiamiento elimina las restricciones sobre el periodo fiscal en el que se puede solicitar un descuento, permitiendo que este se impute en cualquier periodo fiscal posterior al pago del impuesto, siempre que no exceda el límite máximo establecido en el artículo 259 del Estatuto Tributario.
- **Decreto 1366 de 2020:** Este decreto establece las disposiciones para otorgar el registro sanitario a las bebidas alcohólicas fabricadas y comercializadas por microempresas. Además, define los requisitos para obtener la certificación en buenas prácticas de manufactura, asegurando así estándares adecuados de calidad y seguridad en la producción de estas bebidas.

En la Tabla 1, se realiza un análisis cuantitativo de PESTEL para el sector industrial en Colombia, donde cada elemento evaluado tiene una ponderación de 0.07. La convención definida para calificación del impacto está dada de la siguiente manera: 1 para calificación insignificante, 2 para leve, 3 para medio, 4 para mayor y 5 para importante.

Tabla 1. Análisis de PESTEL sector industrial en Colombia.

VARIABLE	FACTOR	PONDERACIÓN	IMPACTO	CALIFICACIÓN PONDERADA
Político	Estabilidad gubernamental	0,07	5	0,36
	Leyes Fiscales (Impuestos)	0,07	5	0,36
Económicos	Tasas de interés	0,07	4	0,29
	Inflación	0,07	4	0,29
	Tasa de desempleo	0,07	2	0,14
	Tendencias del PIB	0,07	4	0,29
Socio - Cultural	Frecuencia de consumo por genero	0,07	5	0,36
Tecnológicas	Nivel de desarrollo tecnológico y uso de software	0,07	4	0,29
	Infraestructura telecomunicaciones - Internet	0,07	3	0,21
Ambientales	Legislación medioambiental	0,07	4	0,29
Legales	Constitución Política de Colombia artículo 333	0,07	5	0,36
	Decreto 3192 de 1983	0,07	5	0,36
	Ley de financiamiento, 1943 de 2018	0,07	5	0,36
	El Decreto 1366 de 2020	0,07	4	0,29
Total				4,21

De acuerdo con Wheelen & Hunger (2013), el sector cervecero se encuentra en un ambiente de buenas oportunidades, pues la calificación ponderada total es superior a 3 (4.21).

Estado del arte

Spence (2019a) encontró que existe una serie de diferencias individuales que genéticamente modifican y diferencian la percepción del gusto, el olfato, la estimulación trigeminal y del color entre hombres y mujeres, relevantes para la percepción de alimentos y bebidas. A pesar de que el mismo autor resalta que las mujeres pueden llegar a ser más sensibles a la identificación de sabores y olores, también identifica que los gustos y asociaciones alimentarias se adquieren en realidad como resultado de la experiencia, lo que puede tener un impacto mayor entre las preferencias alimentarias que presentan hombres y mujeres.

En este sentido, este apartado pretende exponer las preferencias que presentan las mujeres frente a las cervezas artesanales, en cuanto a sabor y estilo. Además, busca resaltar las experiencias en común entre las mujeres, identificadas como barreras y desafíos en un ámbito predominantemente masculino como lo es la industria de la cerveza artesanal; desde el punto de vista cómo consumidoras, productoras o participes de esta industria.

Preferencias de sabor y estilo de cerveza artesanal en mujeres

Actualmente se encuentran pocos estudios sobre las preferencias reales de cerveza de las mujeres, especialmente en países como Colombia. Sin embargo, el hecho de que el interés y consumo de cerveza artesanal por parte de las mujeres esté aumentando sugiere que los estereotipos de género en la cultura de la cerveza artesanal están cambiando hacia

un estado más equitativo (Horowiec, 2022). Lo anterior deja en evidencia la importancia de estudiar y definir estas preferencias de manera real y no sólo como conclusión de los estereotipos establecidos en este mercado.

Diversos estudios han encontrado una percepción generalizada de que la cerveza femenina es ligera, dulce, frutal y refrescante, mientras que la cerveza masculina es pesada, fuerte, amarga y áspera (Chapman et al. 2018; Darwin, 2017; Kappelé, 2015). Estas descripciones no solo caracterizan la masculinidad y feminidad de la cerveza, sino que también describen los atributos de la masculinidad tradicional y la feminidad ideal, lo que sugiere que el consumo de cerveza según el género puede confirmar o refutar los estereotipos de género prevalecientes (Darwin, 2017).

Muggah & McSweeney (2017a) realizaron un estudio para crear descriptores en cervezas canadienses, encontrando como descriptores comunes tanto en hombres como en mujeres: amargor, regusto, intensidad del sabor, carbonatación y dulzura; por su parte, las mujeres identificaron 2 descriptores adicionales: frutal y refrescante. Este estudio también identificó que el amargor es uno de los atributos más importantes que genera un desagrado hacia las cervezas por parte de la población femenina.

En general, las mujeres tienen un paladar diferente a los hombres, prefiriendo cervezas más complejas, menos amargas y con sabores bien balanceados (Donadini et al., 2016). Los resultados obtenidos por Lazzari, et al. (2021) muestran que hombres y mujeres tienen diferentes preferencias respecto al estilo de cerveza. Las mujeres suelen buscar cervezas novedosas y con nuevos sabores y son más críticas que los hombres, quienes son más tradicionales y están influenciados por sus amigos. Este estudio también identificó que cervezas con sustitutos de lúpulo que disminuyen el amargor de la cerveza tenían un impacto positivo en la aceptabilidad del producto por parte de la población femenina.

Horowiec (2022) encontró que varias mujeres estadounidenses prefieren cervezas ácidas, las cuales son consideradas femeninas, pero una cantidad igual prefirieron las

cervezas tipo IPA (Indian Pale Ale), clasificadas como cervezas masculinas. Aunque persiste la percepción de que las mujeres prefieren cervezas más ligeras, dulces y afrutadas, las preferencias reales en este estudio muestran que los estilos de cerveza artesanal considerados masculinos están volviéndose más equitativos en términos de género. Y tal vez, esta percepción del gusto en cervezas de las mujeres refleja los estereotipos de aquellos que hacen recomendaciones a personas con poco conocimiento de la cultura de la cerveza artesanal, al ser estas características las que generalmente se le recomiendan a personas que están empezando a incursionar en este mercado (Chapman et al., 2018). Lo anterior deja en evidencia la poca participación que ha tenido históricamente la mujer en este mercado.

Por otro lado, los consumidores de cerveza toman sus decisiones de compra basándose en factores extrínsecos (por ejemplo, marca y precio), características demográficas y factores intrínsecos (por ejemplo, aroma y sabor) (Gabrielyan et al., 2014). Muggah & McSweeney (2017b) identificaron como indicadores importantes de la intención de compra el sabor, la marca, la apariencia, el estilo de la cerveza, el método de producción y el empaque. Este estudio encontró que el amargor genera disgusto hacia la cerveza, mientras que las consumidoras femeninas prefieren la dulzura. En cuanto al estilo de cerveza, la ale y la IPA tuvieron el efecto positivo más significativo en la elección de las consumidoras femeninas, sin embargo, los autores destacan que los estilos de cerveza que se indicaron como impulsores negativos del gusto, como la stout y la de trigo, no son tan populares en Canadá como las ales y las lagers, y la menor exposición a estos tipos de cervezas puede afectar la aceptabilidad del consumidor.

Muggah & McSweeney (2017b) también evaluaron el impacto del color en la intención de compra, encontrando que la cerveza de color negro tuvo un efecto negativo significativo en la elección de las mujeres, y los demás colores (dorado, rojo, ámbar y café) presentaron un efecto positivo. Este resultado puede deberse a que las consumidoras de cerveza

pueden asociar las cervezas de color negro con intensidades de sabor altas, especialmente amargura (Delwiche, 2004). Finalmente, Muggah & McSweeney (2017b) encontraron que los productos locales tuvieron un impacto positivo en la intención de compra e identificaron que los participantes del estudio creían firmemente que comprar productos locales beneficiaba a la economía local.

En Colombia no se han realizado estudios para determinar las preferencias de sabores y estilos en la población femenina, pero se puede decir que la percepción generalizada de que las mujeres prefieren cervezas ligera, dulce, frutal y refrescante también aplica a las mujeres colombianas basado en los productos en el mercado que pretenden atender este segmento. Una de las cervezas más conocidas en Colombia enfocada a la población femenina es la *Redd's* caracterizada por ser una cerveza ligera, sin amargor, cítrica y refrescante (El tiempo, 2007). Otro ejemplo, relativamente más nuevo en el mercado de cerveza “chic”, es la 3 Cordilleras rosada, la cual es una cerveza dulce, con toques cítricos y aroma y sabor a frutos rojos (3cordilleras, s.f). También se encuentran en el mercado colombiano la cerveza rosada de la cervecería BBC, rosé, la cual presenta un aroma a fresa, sabor dulce a fresa, con cuerpo ligero y refrescante (BBC cervecería, 2024).

Barreras y desafíos percibidos por las mujeres en el consumo de cerveza artesanal

En espacios públicos como bares, pubs y cervecerías, se suele considerar a los hombres como los expertos en cervezas artesanales y se asume que tienen conocimiento sin importar lo que pidan o cuánto sepan, mientras que las mujeres son juzgadas o criticadas por lo que ordenan, especialmente si piden un tipo de cerveza que no se considera una “verdadera cerveza” (Nanney et al., 2020). Las mujeres también sufren de sexualización y objetificación como parte de los anuncios de cerveza, siendo vistas como un objeto que puede disfrutarse con amigos (Poon, 2021, p.20). Un ejemplo claro de esta situación en Colombia fueron los populares anuncios de las tradicionales “Chicas Águila”

representadas por modelos colombianas hipersexualizadas para atraer al consumidor masculino. Sin embargo, esta imagen de las cervezas Águila ha cambiado actualmente. Ahora, las “Chicas Águila” están representadas por las jugadoras de la selección femenina de Colombia, lo que da indicios de un cambio hacia un mercado más equitativo e incluyente.

Las mujeres también han experimentado una considerable presión social para no beber o para limitar su consumo a formas de consumo de alcohol culturalmente restringidas (Thurnell-Read, 2022), lo que contribuye a la percepción de que las mujeres no son consumidoras legítimas de cerveza. Además del estereotipo de que la cerveza es una bebida masculina, también existe la creencia de que ciertos estilos de cerveza son masculinos mientras que otros son femeninos. En la cultura de la cerveza artesanal, hay una amplia gama de estilos de cerveza, de las cuales, las cervezas oscuras y/o nacionales se consideran masculinas y son vistas como las cervezas auténticas (Chapman et al., 2018; Darwin, 2018). Estas incluyen cervezas más pesadas y con sabores robustos, como las porters y stouts, así como aquellas con un fuerte sabor a lúpulo, como las IPAs y American pale ales (APAs) (Horowiec, 2022). Por otro lado, las cervezas ácidas, afrutadas y de trigo se consideran femeninas (Chapman et al., 2018; Darwin, 2018). Estas cervezas son más ligeras, dulces y afrutadas en sabor a las cuales, frecuentemente, se les llama cervezas "para chicas", "rosas" o "chics" y suelen ser infravaloradas, no siendo consideradas como verdaderas cervezas (Nanney et al., 2020).

Chapman et al. (2018) consideran que las prácticas relacionadas con el consumo de alcohol de hombres y mujeres están relacionadas con el entendimiento culturalmente arraigados de lo que significa ser hombre o mujer; por lo que los anuncios relacionados con el alcohol en las revistas femeninas presentan principalmente champán o vino espumoso para representar la feminidad en armonía con el glamour, la sofisticación, el elitismo social y una delicadeza similar a la de una celebridad, mientras que las revistas

masculinas solo presentan cerveza en espacios masculinizados como pubs o eventos deportivos.

Las percepciones sobre las preferencias de sabor son una de las principales formas en que las divisiones de género se han manifestado en la cultura emergente y ahora, en muchos países, bien establecida de la cerveza artesanal (Thurnell-Read, 2022). La percepción de una cerveza como masculina o femenina depende en parte de los significados incorporados en su etiquetado, marca y otros discursos creados por productores que, en el caso de la cerveza artesanal, son mayoritariamente hombres. Si bien las tácticas de marketing basadas en las percepciones de género pueden atraer a algunas mujeres al movimiento de la cerveza artesanal, la mayoría de las mujeres ya inmersas en la cultura rechazan las nociones estereotipadas femeninas de que, por ser mujeres, deben gustarles la cerveza rosa y sus rasgos femeninos asociados (Chapman et al., 2018). De esta forma la manera de promocionar las cervezas, o el *Storytelling*, para mujeres juega un papel importante. Kuehn & Parker (2018) describen que las mujeres cerveceras en Nueva Zelanda emplean atributos de la cerveza artesanal como procedencia, lugar, herencia, sostenibilidad, localismo y comunidad como principales autenticadores de legitimidad y promotores de sus productos.

Otra de las barreras que encuentran las mujeres en el mercado de cervezas son las diversas situaciones en las que no son tomadas en serio por sus compañeros (hombres) bebedores de cerveza o son encasilladas como alguien que no sabe nada sobre cerveza (Chapman et al., 2018). Más aun, Tak et al. (2019) encontraron en un estudio realizado con 150 voluntarios estadounidenses que una etiqueta de cerveza producida con el nombre de una mujer presenta peores resultados que la misma etiqueta producida con el nombre de un hombre, los participantes afirmando que pagarían menos por una cerveza producida por mujeres y tendrían menores expectativas sobre la calidad y el sabor. Sin embargo, los mismos autores descubrieron que si la cerveza estaba etiquetada como

ganadora de un premio o si la cervecera artesanal era una "experta en cerveza", los participantes no mostraban niveles similares de discriminación o sexismo sobre el producto. Para afrontar estos escenarios, las mujeres buscan certificaciones o premios como validación, desarrollan empatía con su comunidad o se sienten empoderadas por sus contrapartes femeninas y las sociedades de cerveza artesanal centradas en las mujeres (Frankel, et al., 2021).

Propuesta de modelo predictivo

Teniendo en cuenta el estado del arte de las preferencias y barreras que presentan las mujeres como consumidoras de cerveza artesanal; y la falta de datos en Colombia para el estudio de las preferencias de este segmento poblacional, se propone la encuesta expuesta en la tabla 2 para recolectar datos en Colombia sobre esta temática y poder tener datos necesarios para poder predecir los factores que influyen en la elección de cerveza artesanal en mujeres colombianas, lo que permitirá guiar a las empresas productoras de cerveza artesanal en la creación de productos que se alineen con las preferencias y necesidades de la población femenina.

Tabla 1. *Propuesta de encuesta para generar datos para el modelo predictivo*

1. ¿Se identifica usted como mujer? Si__ No__
2. ¿Qué edad tiene? 18-25 años__ 26-35 años__ 36-45 años__ 46-60 años __ más de 60 años__
3. ¿En qué ciudad vive? _____
4. ¿Tiene hijos? Si__ No__

<p>5. ¿Tiene pareja?</p> <p>Si_ No_</p> <p>6. ¿Con qué frecuencia fuma cigarrillos?</p> <p>Menos de 1 al día__ Entre 1 y 4 al día__ Entre 5 y 9 al día__ Más de 10 al día__</p> <p>7. ¿De cuánto son sus ingresos mensuales?</p> <p>1-2smlv__ 3-4 smlv__ 5-6 smlv__ más de 7 smlv__</p> <p>8. ¿Qué nivel educativo tiene?</p> <p>Bachiller__ Pregrado__ Posgrado__ Ninguno__</p> <p>9. Se considera usted consumidora de cerveza:</p> <p>Frecuente__ Ocasional__ No consume__</p> <p>10. ¿Qué tipo de cerveza prefiere?</p> <p>Cerveza artesanal__ Cerveza industrial__</p> <p>11. ¿Se considera usted como conocedora de la cerveza artesanal?</p> <p>Si__ No__</p> <p>12. ¿Se ha sentido discriminada en un entorno cervecero?</p> <p>Si__ No__</p> <p>13. ¿Se siente usted cómoda y segura cuando bebe cerveza en lugares públicos?</p> <p>Si__ No__</p> <p>¿Por qué? _____</p> <p>14. ¿Qué atributos considera usted debe tener una cerveza de su gusto?</p> <p>Ligera__ Pesada__ Fuerte__ Refrescante__ Dulce__ Amarga__ Ácida__ Robusta__</p> <p>Oscura__ Frutal__ Otra_____</p> <p>15. ¿Qué color de cerveza considera usted más agradable?</p> <p>Dorada__ Ámbar__ Roja__ Café__ Negra__</p>
--

16. Ordene de mayor a menor (siendo 1 el mayor y 7 el menor) la influencia de los siguientes factores sobre su preferencia por una cerveza:

Precio__ Origen__ Sostenibilidad__ Comunidad__ Sentido de pertenencia__
Representación femenina__ Empaque__

17. Prefiere usted consumir cerveza en presentación en:

Lata__ Botella__ Barril/Jarra/Vaso__

18. ¿Qué tipo de marca prefiere?

Local__ Internacional__

19. ¿Qué estilo de cerveza prefiere?

Ale__ IPA__ Lager__ Stout__ Trigo__ Pilsner__ Otra__ No conoce o no sabe __

Debido a las restricciones de tiempo en esta investigación no es posible recopilar suficientes encuestas para dar soporte a un modelo de ML, por lo que el modelo propuesto se probó con base en una simulación de datos generados en *Python* a partir de la encuesta propuesta en la tabla 2.

En primer lugar, se realizó la importación de las librerías necesarias para el desarrollo de la simulación en *Python* y algunas otras necesarias posteriormente en el modelo, cómo se puede apreciar en la figura 3.

Figura 3. Importación de las librerías principales

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Nota. Adaptado de Roncancio, O, 2024.

https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

La librería de *pandas* es una de las herramientas más usadas en el análisis de datos, esto debido a su facilidad de uso, su excelente rendimiento, flexibilidad, entre otras virtudes (pandas, 2024). Esta librería es fundamental y es base para el desarrollo de casi cualquier proceso que incluya tratamiento de datos en *Python*. Por otra parte, la librería *Matplotlib* se utiliza para la generación de gráficas basadas en datos almacenados en las diferentes estructuras trabajadas en *Python* (Unipyhton, 2017). *Seaborn* es una librería fundamentada en *Matplotlib*, y así como su antecesora se utiliza principalmente para crear visualizaciones gráficas, aunque con una mayor variedad para satisfacer las necesidades del programador (Eyre, 2024). Además de estas tres librerías también se hace uso de *numpy*, que es sumamente importante para realizar cálculos matemáticos y estadísticos avanzados dentro de *Python*, y por supuesto esto lo hace una herramienta invaluable para la manipulación de datos (Roch, s.f.).

Una vez se importan las librerías necesarias se empieza con el proceso de simulación de datos, para esto se leyó el archivo de departamentos y municipios de Colombia recopilado del Ministerio de Tecnologías de la Información y las Comunicaciones (2024) con la librería *pandas* y se guardó en un *dataframe* llamado *df_ciudades*. Cabe destacar que un *dataframe* (*df*) es una de las unidades básicas que usa *pandas* para almacenar sus datos y mantiene una estructura matricial de filas por columnas (pandas, 2024). Después de almacenar el *df* se empieza a parametrizar la encuesta de modo que se pueda generar la simulación con valores aleatorios, primero se usa la librería *numpy* nombrada como *np*, se inicializa la colección de funciones *random* de la librería para generar números aleatorios y finalmente se establece el parámetro de la función *seed* con un valor de 0 para que se puedan replicar los valores aleatorios generados. Es importante aclarar que una función es un grupo de líneas de código que realizan funciones específicas cuando se invocan (Llerena Izquierdo, 2020) así mismo una colección de funciones es simplemente un grupo de funciones.

Posteriormente se creó una variable llamada "N" que se estableció con el valor de las muestras que se querían simular. Una variable en programación es una manera de almacenar información en la memoria del computador o dispositivo que se esté usando y así cómo su nombre lo indica esta puede cambiar su valor por otro (Llerena Izquierdo, 2020). Existen múltiples tipos de variables en programación, dentro del desarrollo de la simulación se usará recurrentemente el tipo de variable llamado listas. Las listas son estructuras de datos que pueden almacenar valores ya sean números, texto o incluso ambas (Aprende Python, s.f.). Para efectos de una gran muestra de valores se escogió la variable "N" con un valor de 10000, por lo que se simuló el equivalente a 10000 encuestas.

En la figura 4 se muestra una imagen del código usado para la simulación con propósito de explicarlo detalladamente.

Cada una de las líneas que empieza con el símbolo "#" corresponde a un comentario realizado dentro del código creado cómo se puede observar en la Figura 4. En cuanto a la generación de características de los encuestados, se simuló si el encuestado se identifica como mujer (0) o no (1), permitiendo así un enfoque específico en el consumo de cerveza artesanal entre mujeres. Previamente en la propuesta de encuesta realizada se puede observar que las respuestas para la pregunta número 1 son (Si) y (No) respectivamente, no obstante, para propósitos de eficiencia y generar menos carga posteriormente a los modelos de ML se hace uso de una técnica llamada *OneHot Encoding*. Este concepto será explicado en los siguientes párrafos. Cuando se trabaja con modelos de ML se emplean 2 tipos de variables, las numéricas y las categóricas (Brownlee, Machine Learning Mastery, 2020).

Figura 4. Código de la simulación de 10000 encuestas aleatorias.

```
#Primero las personas van a responder si se identifican o no como mujeres esto sera en un dataset y esta sera la primera
Mujer = np.random.choice([0, 1], N)
#La segunda columna sera la edad separada en estos grupos: 18-25 años__ 26-35 años__ 36-45 años__ 46-60 años __ más de 60 años
edad = np.random.choice([1, 2, 3, 4, 5], N)
#La tercera sera en que ciudad de Colombia viven, para esto vamos a usar la columna de municipio del df_ciudades
ciudad = np.random.choice(df_ciudades['MUNICIPIO'], N)
#La cuarta columna sera si la persona tiene hijos o no
hijos = np.random.choice([0, 1], N)
#La quinta columna sera si la persona tiene pareja o no
pareja = np.random.choice([0, 1], N)
#La sexta sera con que frecuencia fuma la persona en estos rangos: Menos de 1 al día__ Entre 1 y 4 al día__ Entre 5 y 9 al día__ más de 9 al día
fuma = np.random.choice([1, 2, 3, 4], N)
#La septima sera ¿Cuánto son sus ingresos mensuales? 1-2smlv__ 3-4 smlv__ 5-6 smlv__ más de 7 smlv
ingresos = np.random.choice([1, 2, 3, 4], N)
#La octava sera ¿Qué nivel educativo tiene? Bachiller__ pregrado__ posgrado__ ninguno__
educacion = np.random.choice([1, 2, 3, 0], N)
#La novena sera Se considera usted consumidora de cerveza: frecuente__ ocasional__ no consume__
cerveza = np.random.choice(['Frecuente', 'Ocasional', 'No consume'], N)
#La decima sera ¿Qué tipo de cerveza prefiere? Cerveza artesanal__ cerveza industrial__
tipo_cerveza = np.random.choice([0, 1], N)
#La onceava sera ¿Se considera usted como conocedora de la cerveza artesanal? Si__ No__
conocedora = np.random.choice([0, 1], N)
#La doceava es una pregunta con multiples respuestas cada respuesta debe ser una columna y debe ser si o no para cada valor
pregunta = ['Ligera', 'Pesada', 'Fuerte', 'Refrescante', 'Dulce', 'Amarga', 'Acida', 'Robusta', 'Oscura', 'Frutal', 'Otra']
respuestas = np.random.choice([0, 1], size=(N, 11))
df_respuestas = pd.DataFrame(respuestas, columns=pregunta)
#La treceava sera ¿Qué color de cerveza considera usted más agradable? Dorada__ ambar__ roja__ café__ negra__
color = np.random.choice(['Dorada', 'Ambar', 'Roja', 'Café', 'Negra'], N)
#convertimos a one hot encoding
color = pd.get_dummies(color)
#Para la catorceava la respuesta sera una lista respecto a esta pregunta: Ordene de mayor a menor (siendo 1 el mayor y 7 el menor)
factores = ['Precio', 'Origen', 'Sostenibilidad', 'Comunidad', 'Sentido de pertenencia', 'Representación femenina', 'Empoderamiento']
orden = np.random.choice(range(1, 8), size=(N, 7))
df_orden = pd.DataFrame(orden, columns=factores)
#La quinceava sera ¿Se siente usted cómoda y segura cuando bebe cerveza en lugares públicos? Si__ No__
seguridad = np.random.choice([0, 1], N)
#La dieciseisava sera Prefiere usted consumir cerveza en presentación en: lata__ botella__ barril/jarra/vaso__
presentacion = np.random.choice(['Lata', 'Botella', 'Barril/Jarra/Vaso'], N)
#hacemos one hot encoding
presentacion = pd.get_dummies(presentacion)
#La diecisieteava sera ¿Qué tipo de marca prefiere? Local__ internacional__
marca = np.random.choice([0, 1], N)
#La dieciochoava sera ¿Se ha sentido discriminada en un entorno cervecero? Si__ No__
discriminacion = np.random.choice([0, 1], N)
#Y la última sera ¿Qué estilo de cerveza prefiere? Ale__ IPA__ Lager__ stout__ trigo__ Pilsner__ otra__ no conoce o no sabe__
estilo = np.random.choice(['Ale', 'IPA', 'Lager', 'Stout', 'Trigo', 'Pilsner', 'Otra', 'No conoce o no sabe'], N)
#convertimos a one hot encoding
estilo = pd.get_dummies(estilo)
#Finalmente vamos a unir todos los datos en un solo data frame
df_encuesta = pd.DataFrame({'Mujer': Mujer, 'Edad': edad, 'Ciudad': ciudad, 'Hijos': hijos, 'Pareja': pareja, 'Fuma': fuma, 'Ingresos': ingresos, 'Educacion': educacion, 'Cerveza': cerveza, 'Tipo Cerveza': tipo_cerveza, 'Conocedora': conocedora, 'Respuestas': df_respuestas, 'Orden': df_orden, 'Seguridad': seguridad, 'Presentacion': presentacion, 'Marca': marca, 'Discriminacion': discriminacion, 'Estilo': estilo})
df_encuesta = pd.concat([df_encuesta, df_respuestas, df_orden, color, presentacion, estilo], axis=1)
```

Nota. Adaptado de Roncancio, O, 2024,

https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

Las variables numéricas como indica su nombre son números, generalmente enteros o decimales. Por otra parte, las variables categóricas suelen ser etiquetas generalmente en texto. Un ejemplo de variable categórica puede ser una variable animal que contiene valores como “león”, “perro” o “pájaro” y cada valor corresponde a una categoría distinta y estas variables categóricas se pueden dividir a su vez en 2 grupos específicos, las de tipo nominal y las de tipo ordinal (Ibid., 2020).

Las de tipo nominal corresponden a variables categóricas que no tienen ninguna relación entre sí cómo es el caso del ejemplo dado en el párrafo anterior, mientras que las ordinales como su nombre indica tienen un valor ordenado, por ejemplo, la frecuencia con la cual las personas se lavan las manos en el día siendo 1 el menor valor, luego 5 un valor intermedio y finalmente 10 como valor máximo. Debido a que gran parte de los modelos de ML no pueden usar variables categóricas, se crearon técnicas que permiten convertir este tipo de variables a números, logrando una mayor eficiencia en los modelos que ya podían usar variables categóricas y permitiendo la implementación de otros tipos de modelos que no podían ser usados previamente (Ibid., 2020). Durante este trabajo se usaron dos de estas técnicas.

La más común de estas técnicas se llama *OneHot Encoding* esta técnica consiste en tomar una variable categórica que solo tiene dos valores y reemplazarlos por 0 y 1 respectivamente (Ibid., 2020), cómo es el caso de la primera pregunta en la que se pide responder a la persona encuestada si se identifica o no como mujer. En segundo lugar, la variable “edad” se reemplaza mediante *Ordinal Encoding* en uno de cinco grupos, cada uno de estos grupos corresponde a un rango de edades, esto se observa en la pregunta número 2 de la encuesta propuesta. *Ordinal Encoding* es la segunda técnica usada para el desarrollo de la simulación. Este tipo de *Encoding* reemplaza cada uno de los rangos por un valor numérico, para este caso del 1 al 5, siendo 1 el rango de edades más bajo y 5 el rango más alto.

La ciudad de residencia de cada encuestado es la tercer variable, para la cual se usó el “df_ciudades” asignado previamente. Con ayuda de la función *random* y la variable “N” se escogieron aleatoriamente ciudades para los 10000 valores simulados, esta variable no se tomó en cuenta en el entrenamiento de los modelos de ML para reducir la carga generada, en caso de usarse en futuras investigaciones se recomienda realizar *OneHot Encoding* para esta variable categórica ya que es de tipo nominal. Las variables “hijos” y “pareja”

correspondientes a las preguntas 4 y 5, al igual que en la pregunta 1, se reemplazaron con *OneHot Encoding*, siendo (0) afirmativo y (1) negativo. Para las variables “fuma”, “ingresos” y “educación” que son las preguntas 6, 7 y 8 respectivamente se usó *Ordinal Encoding*, cómo en la pregunta número 2.

En la pregunta número 9, correspondiente a la variable “cerveza”, no se realizó ningún tipo de *Encoding*, ya que está fue escogida como variable objetivo. Es válido recordar que en el desarrollo de esta investigación se trabajó con modelos de aprendizaje supervisado, específicamente *Decision Trees* y *Random Forest*, por lo tanto, siempre se debe trabajar con una variable objetivo tal como se explicó en el marco teórico. Teniendo en cuenta que el propósito de la investigación es determinar preferencias y factores se debe tomar como variable objetivo la pregunta que corresponda al consumo de cerveza, permitiendo así evaluar la importancia de las características una vez entrenado el modelo de ML.

Para las preguntas 10, 11 y 13 correspondientes a las variables “tipo_cerveza”, “conocedora” y el df “color” se implementó *OneHot Encoding*. Las preguntas 10 y 11 manejan la misma estructura de la pregunta 1 por lo que ya se sabe con qué valores se realiza el *Encoding*. No obstante, la pregunta 13 es un poco diferente debido a que contiene varias respuestas. Para implementar *OneHot Encoding* en esta variable se usó la función `get_dummies` de la librería de pandas, internamente, esta función separa cada una de las posibles respuestas a esta pregunta y crea un df que contiene una columna por cada respuesta, a su vez cada una de estas columnas pueden tener un valor de (0) en caso de respuesta afirmativa o (1) en caso de respuesta negativa. Es importante aclarar que al ser una pregunta con una única respuesta solo una de las columnas tiene su valor en 0 manteniendo el concepto explicado anteriormente.

La pregunta número 12 también es un caso especial, al ser una pregunta de selección múltiple requiere un tratamiento similar a la pregunta 13 aunque con una variación, la diferencia es que en este caso podemos tener varias respuestas, mientras que en la

número 13 solo podíamos tener 1 respuesta, por lo que se realiza *OneHot Encoding* de la misma forma, pero permitiendo la posibilidad de que varios valores sean afirmativos, es decir “0” para una misma encuesta dando cómo resultante un df llamado “df_respuestas”.

La pregunta número 14 correspondiente al df “df_orden” ya tiene *Ordinal Encoding* desde su formulación ya que se pide ordenar por relevancia del 1 al 7, por lo que se genera un df donde cada una de las respuestas es una columna y tiene asignado un valor único del 1 al 7 por encuesta.

Finalmente, para las preguntas 15, 16, 17, 18 y 19 que son las variables “Seguridad”, “Presentación”, “Marca”, “Discriminación” y “estilo” también se implementó *OneHot Encoding* donde las preguntas 15, 16 y 18 corresponden al caso de la pregunta 1 mientras que la pregunta 17 y 19 corresponden al caso de la pregunta 13.

Después de realizar el *encoding* y la generación de valores aleatorios se creó un df resultante llamado “df_encuesta”. Este df se creó agrupando todas las variables y dataframes definidos y explicados en los párrafos anteriores. La visualización del “df_encuestas” se muestra en la figura 5.

Figura 5. Ejemplo de visualización del df.

	Mujer	Edad	Hijos	Pareja	Fuma	Ingresos	Educacion	Tipo_cerveza	Conocedora	Seguridad	...	Botella	Lata	Ale	IPA	L
0	0	4	1	1	4	4	0	0	0	1	...	False	True	False	False	
1	1	2	1	0	1	4	2	0	1	1	...	True	False	False	False	
2	1	4	1	1	4	2	2	1	0	1	...	True	False	True	False	
3	0	1	1	1	4	3	0	1	0	0	...	False	False	False	True	
4	1	3	0	1	1	3	2	1	1	1	...	False	False	False	False	
...
9995	0	4	0	0	4	3	2	1	1	1	...	False	False	False	False	
9996	1	3	1	0	4	1	3	0	1	1	...	False	True	False	False	
9997	0	4	0	1	2	1	2	0	0	0	...	False	True	False	False	
9998	1	1	1	1	2	4	1	0	0	1	...	True	False	False	False	
9999	0	1	0	1	3	1	0	1	1	0	...	False	True	False	False	

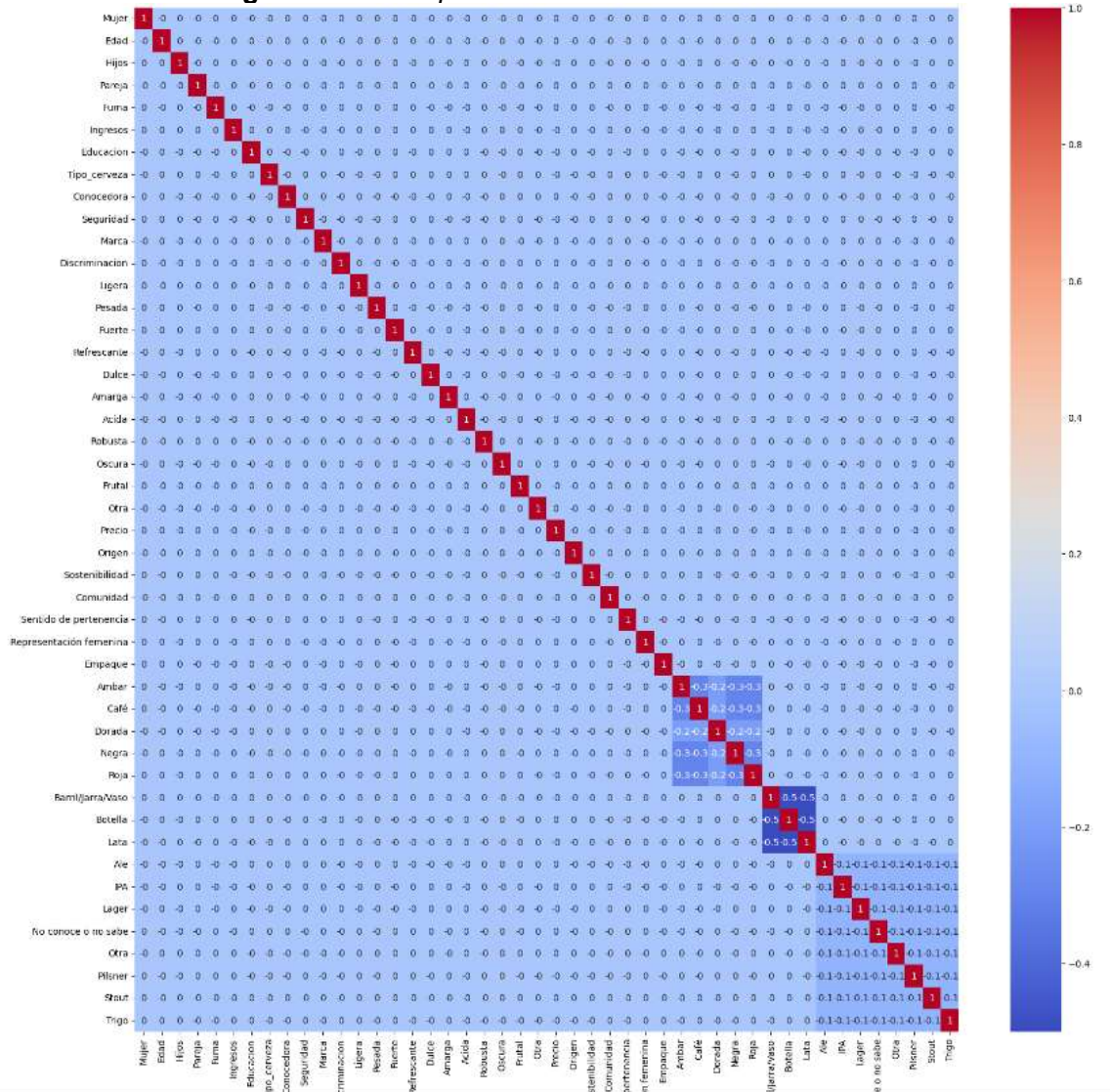
Nota. Adaptado de Roncancio, O, 2024, https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

En la figura 5 se ven algunos valores como *True* y *False*, a esto también se le puede denominar como una variable de tipo *boolean*, estas variables son conocidas porque solo pueden tomar 2 valores, para el presente caso no existe ninguna afectación, esto es un efecto netamente visual ya que cuando *Python* tiene solo valores de 0 y 1 dentro de una columna del *df* interpreta los 0 como *False* y el 1 como *True*, es simplemente otra forma en la que se puede representar la información (Llerena Izquierdo, 2020).

Seguido a este proceso se separa el *df* en 2 dejando un *df* llamado "Y" y otro llamado "X". El *df* "Y" corresponde a la columna de la variable objetivo, para este caso llamada "cerveza" mientras que "X" contiene todas las demás columnas. Con el *df* separado se realizó una gráfica de *heatmap* con las librerías de *matplotlib* y *seaborn*, este tipo de gráfico permite observar gráficamente la correlación que hay entre cada una de las variables presentes en el *df* "X" como se puede observar en la Figura 6.

El *heatmap* o mapa de calor es un gráfico que permite conocer de una manera muy sencilla que tan correlacionadas están múltiples variables entre sí. Normalmente maneja una escala de color donde uno de los extremos corresponde a 1 y el otro a -1 respectivamente. Si el valor entre 2 variables se encuentra más cerca de 1 quiere decir que tienen una correlación positiva, mientras que en caso opuesto tendrían una correlación negativa. Una forma más sencilla de explicar esto es que si hay correlación positiva quiere decir que entre más aumente una de las variables la otra también lo hará, mientras que si tienen correlación negativa pasara el efecto contrario, es decir, si una de las variables aumenta entonces por consiguiente la otra disminuirá, o viceversa. Al realizar este análisis visual se puede empezar a identificar realmente cuál de las variables puede tener influencia sobre las demás. Es importante resaltar que la diagonal de este gráfico siempre será 1 ya que representa la correlación de cada variable contra sí misma (Narula, 2021).

Figura 6. Heatmap de las variables de la encuesta.



Nota. Adaptado de Roncancio, O, 2024,

https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

Con las variables “X” y “Y” definidas se procedió a importar las librerías necesarias para el modelo de *Decision Trees*. La librería de *Sklearn* es una de las librerías usadas para realizar proceso de ML en *Python*, contiene múltiples herramientas que ayudan al programador en el preprocesamiento de datos y también en la generación de modelos predictivos (Chandra, 2023).

De la librería de *Sklearn* se importó inicialmente el módulo para separar el df para su entrenamiento, así como los módulos correspondientes a las estadísticas para evaluar el

modelo y también los módulos de *Decision Trees* y *Random Forest*. El siguiente paso fue la división de los datos, por defecto *Sklearn* divide las variables “X” y “Y”. La variable “X” se divide en las variables “X_train” y “X_test”, de igual forma sucede con “Y”. Cuando no se especifica directamente en que proporciones se divide cada variable, *sklearn* asigna por defecto 0.3 para las variables “X_train” y 0.7 para las variables “X_test”, esto equivale al 30% y 70% respectivamente (Sklearn, 2024).

Cada uno de los modelos de ML existentes manejan parámetros que pueden ser modificados para dar una mayor precisión al modelo, estos parámetros suelen ser asignados de manera arbitraria. En este caso se hizo uso de otro de los módulos ofrecidos por *Sklearn* llamado *GridSearchCV*. Con ayuda de este módulo se definen varios de los posibles parámetros que puede tener el modelo, permitiendo así que éste sea evaluado con múltiples combinaciones entre los parámetros, como se puede apreciar en la figura 7.

Se declaró una variable llamada “param_grid” que contiene los posibles parámetros del modelo de *Decision Trees*. Posteriormente se inicializó el modelo de ML y se asignó a la variable “dt”, similar al parámetro *Seed* que se usó antes para reproducir la aleatoriedad de la simulación. En este punto se puede usar el parámetro “random_state” que realiza la misma función, pero para entrenamientos de modelos de ML. Posteriormente a este proceso se inicializó la variable “grid_search” y se le brindaron los parámetros estimados del modelo inicializado, es decir “dt”. La variable con los posibles parámetros “param_grid”, “cv” que es la cantidad de grupos, “n_jobs” con un valor de -1, lo que quiere decir que usará todos los procesadores disponibles en el dispositivo, y *verbose* con un valor de 2, que controla la cantidad de mensajes emitidos por el módulo (Sklearn, 2024). Como resultado, “*GridSearch*” (usando la función “best_params”) nos brindó como resultado la combinación que mejor resultado tuvo dentro de su análisis.

Figura 7. Uso de GridSearch.

```
#Gridsearch para encontrar Los mejores hiperparámetros para el modelo de decision tree
from sklearn.model_selection import GridSearchCV
from sklearn.tree import DecisionTreeClassifier

# Definir Los hiperparámetros
param_grid = {
    'criterion': ['gini', 'entropy'],
    'max_depth': np.arange(1,30),
    'min_samples_split': [2, 5, 10, 15, 20],
    'min_samples_leaf': [1, 2, 5, 10, 15]
}

# Inicializar el modelo
dt = DecisionTreeClassifier(random_state=0)

# Inicializar GridSearch
grid_search = GridSearchCV(estimator=dt, param_grid=param_grid, cv=5, n_jobs=-1, verbose=2)

# Entrenar el modelo
grid_search.fit(X_train, y_train)

# Mostrar Los mejores hiperparámetros
grid_search.best_params_
```

Fitting 5 folds for each of 1450 candidates, totalling 7250 fits

```
{'criterion': 'gini',
 'max_depth': 20,
 'min_samples_leaf': 2,
 'min_samples_split': 2}
```

Nota. Adaptado de Roncancio, O, 2024,

https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

Una vez encontrados los mejores parámetros para el modelo se creó una instancia haciendo uso de los parámetros previamente hallados y se asignó el resultado a la variable llamada modelo. Luego se usó la función *fit*, usando las variables “X_train” y “Y_train” para entrenar el modelo con la porción separada con anterioridad y se crea la variable “Y_pred”. Esta variable se usa para predecir los valores separados en la variable “X_test”. Finalmente, se imprimió la exactitud del modelo también llamado *accuracy*, para esto se le dan las variables “Y_test” y “Y_pred” para que compare las predicciones realizadas en la variable anterior con las clasificaciones verdaderas dando como resultado un valor decimal que al multiplicar por 100 queda en términos porcentuales.

Existen múltiples maneras de revisar la efectividad de los modelos de ML. Con fines de guiar a futuros investigadores se plantearon 3 de estas técnicas durante la implementación de los modelos. La primera técnica es el *accuracy* mencionado en el anterior párrafo, esta medida representa la cantidad de valores que fueron correctamente catalogados entre la cantidad total de valores disponibles, hay que tener en cuenta que una sola técnica para medir la efectividad de un modelo no brinda demasiada confianza por si sola, por lo que se recomienda hacer uso de al menos las propuestas en esta guía. La segunda técnica implementada es llamada *Confusion matrix*, esta se compone de las siguientes medidas estadísticas: “precision”, “recall”, “f1-score”, “support”, “accuracy” y dos promedios (“macro avg” y “weighted avg”), esta matriz puede variar según la cantidad de valores que tenga la variable objetivo (Harikrishnan, 2019). Los investigadores que sigan esta guía tendrían una matriz mostrada en la figura 8.

Figura 8. Matriz de confusión.

Reporte del modelo:	precision	recall	f1-score	support
Frecuente	0.33	0.54	0.41	987
No consume	0.35	0.06	0.11	1060
Ocasional	0.32	0.39	0.35	953
accuracy			0.32	3000
macro avg	0.33	0.33	0.29	3000
weighted avg	0.33	0.32	0.28	3000

Nota. Adaptado de Roncancio, O, 2024, https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

Como se puede observar en la Figura 8, a la izquierda de la matriz se encuentran las 3 categorías disponibles en la variable objetivo, es decir, la variable cerveza. Posteriormente se encuentran la medida “precision” que indica qué tan efectivo es el modelo prediciendo un valor dado para cada etiqueta, seguida de la medida “recall” que indica qué tan efectivo es clasificando el modelo cada una de las etiquetas. Luego se encuentra “f1-score” que

mide la relación entre la medida “precision” y la medida “recall”. Finalmente, “support” es una medida que indica cómo se distribuyen en las categorías las muestras simuladas, el total de valores es de 3000 ya que como se indicó antes solo se deja 30% de los valores para el segmento de pruebas, y como se generaron 10000 encuestas simuladas 3000 corresponde al 30% separado (Harikrishnan, 2019).

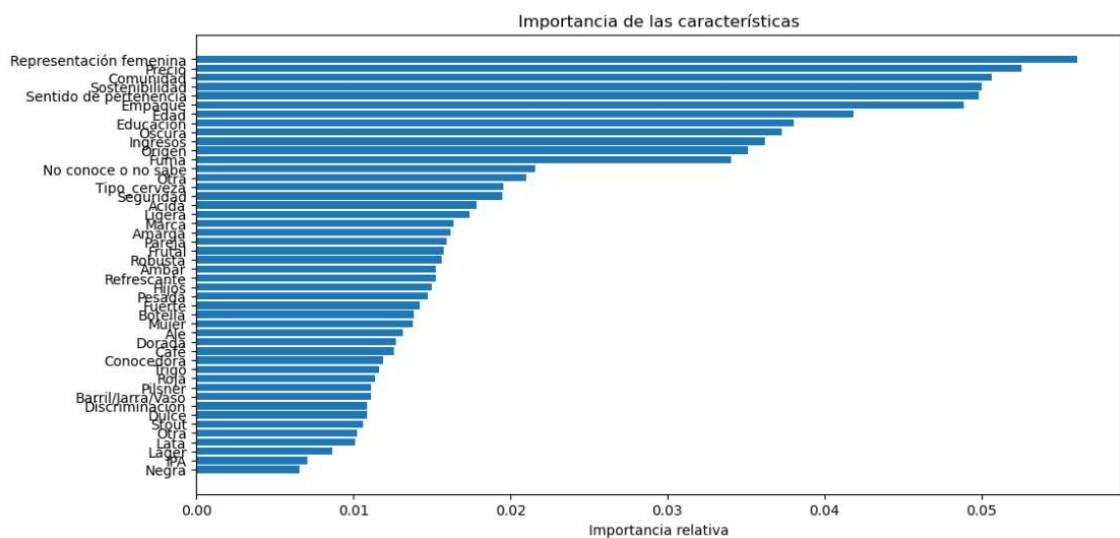
Hacia abajo se encuentra nuevamente la medida calculada “accuracy” y los promedios entre los posibles resultados para la variable objetivo. Sin embargo, al observar la Figura 8, se puede evidenciar que desde todas las perspectivas y con las 2 técnicas usadas este no es un modelo que aporte valor, esto se debe netamente a que los valores son totalmente aleatorios, lo que significa que el modelo no será capaz por sí solo de determinar patrones de comportamiento. Esto hace necesario usar datos reales para aplicar esta propuesta.

Adicional a las 2 técnicas exploradas, se usó una tercera para dar mayor soporte a la guía creada. La técnica de curva de aprendizaje o *learning curves* nos muestra de manera visual el comportamiento a través de la cantidad de muestras del “accuracy” mientras se entrena el modelo, así como la validación de este a través de las mismas muestras, entre más cerca estén ambas líneas quiere decir que los datos no están sesgados. Existe la posibilidad de que cuando se tomen las encuestas todas las respuestas se enfoquen únicamente a una de las categorías de la variable objetivo, en ese caso podríamos decir que el modelo puede estar sesgado. Idealmente se busca que la información este lo más uniforme posible para que el aprendizaje del modelo también logre ser óptimo para su uso. Con ayuda de estas curvas se puede evidenciar este tipo de alertas, además de indicar que tan usable puede llegar a ser el modelo (Brownlee, Machine Learning Mastery, 2019).

Finalmente se llega al apartado más importante en el que se determinan los factores y preferencias en el consumo de cerveza artesanal en mujeres. Para determinar esas preferencias y factores se hizo uso de otra función de los modelos de ML llamada “feature_importances_”, esta función usa el modelo generado con anterioridad y calcula

cuales son las características que más importan para determinar uno de los resultados de la variable objetivo. Se tomó como característica cada una de las columnas de nuestro df, estos resultados se dan a nivel porcentual y se almacenan en la una lista que se ordena posteriormente para ser visualizada en un gráfico de barras con ayuda de *matplotlib* y *seaborn*. En la figura 9 se muestra un ejemplo de cómo debería verse el resultado de este cálculo.

Figura 9. Feature importance



Nota. Adaptado de Roncancio, O, 2024,

https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipynb

Si las muestras que se tienen en esta guía no fueran aleatorias se podría concluir con la Figura 9 que la representación femenina, el precio y la comunidad son las 3 características más importantes a la hora de determinar el consumo de cerveza artesanal en la población de la encuesta, por lo que se podrían generar propuestas por parte de los productores para explotar este nicho de mercado.

Los modelos de ML tienden a mantener una estructura genérica en su implementación, por lo que la propuesta del modelo de *Random Forest* tiene una gran similitud con la que se explicó para *Decision tree*. El único cambio evidenciable está en la búsqueda de parámetros con *GridSearch* donde aumenta la cantidad de parámetros en uno teniendo

ahora `n_estimators` que corresponde a la cantidad total de *Decision Trees* que se generaran para el modelo de *Random Forest*. Tanto la implementación de este modelo como todo el material explicado previamente fue recopilado de (Roncancio, 2024)

Discusión

Aún persiste el estereotipo de que existen cervezas artesanales masculinas y femeninas, y que beber cerveza es una actividad inherentemente masculina. Sin embargo, el número de mujeres en la comunidad de la cerveza artesanal está aumentando, no solo en cuanto a las mujeres que disfrutan de consumir cerveza artesanal, sino también en el número de mujeres que elaboran cerveza y poseen cervecerías (Horowiec, 2022).

El aumento en el consumo de cerveza artesanal por parte de las mujeres y el conocimiento de sus preferencias pueden ser beneficiosos para las cervecerías en sus esfuerzos de marketing y publicidad. Esto es relevante tanto para pequeñas cervecerías locales como para grandes cervecerías comerciales. Las mujeres representan un gran mercado objetivo para las cervecerías, y conocer sus preferencias puede ayudar en el desarrollo de estrategias de marketing efectivas. Comprender los estereotipos de género en la cultura de la cerveza también puede contribuir a que los espacios públicos de consumo sean más inclusivos y equitativos para ambos géneros (Horowiec, 2022).

Si bien la industria de la cerveza artesanal sigue siendo dominada por hombres, recientemente han aparecido estudios que proporcionan una corrección oportuna y necesaria a la escasez de investigaciones sobre las experiencias de las mujeres con las bebidas artesanales (Kuehn & Parker, 2018; Land et al., 2018; Rydzik & Ellis-Vowles, 2019). También es notable el aumento de individuos, grupos y organizaciones que representan a las consumidoras de cerveza artesanal, cerveceras y trabajadoras de la cerveza artesanal. Es de destacar la *Pink Boots Society*, una organización sin ánimo de lucro que tiene el objetivo de "asistir, inspirar y alentar a las mujeres profesionales de la industria cervecera a avanzar en sus carreras a través de la educación" (Pink Boots Society, 2024), la cual ha inspirado a otras organizaciones y ha realizado actividades para reunir a cerveceras en elaboraciones colaborativas regulares y otras actividades que

aumentan la visibilidad de las mujeres que trabajan en el sector, como el *Women's Craft Beer Collective* de San Diego, *FemAleBrewFest* en US y el *FemAle BrewsterFest* en UK (Thurnell-Read, 2022). Al resaltar y brindar una plataforma para que las mujeres en la industria de la cerveza artesanal compartan sus narrativas es posible generar diálogos sobre las complejidades y las luchas que enfrentan las mujeres, incentivando de esta forma el consumo de cervezas artesanales en las mujeres mediante la generación de sentido de pertenencia.

En Colombia ya se pueden encontrar espacios de inclusión que inspiran y buscan generar sentido de pertenencia mediante la cerveza artesanal. En 2020, nació un movimiento conocido como *Comunidad de mujeres cerveceras*, que busca empoderar a las mujeres productoras de cerveza y darles visibilidad. Este grupo de mujeres está compuesto por dueñas de locales, supervisoras de calidad, maestras cerveceras e influenciadoras (Alba, 2023). Algunas de las cervecerías creadas por las mujeres de este grupo incluyen a Green Hops, Yuma, Brewlinger, Madre Monte, Wica, Slow Beer y Non grata, las cuales elaboran una gran variedad de cervezas que van desde cervezas claras, poco amargas, dulces y afrutadas hasta cervezas negras con sabores fuertes, amargas y mucho aroma (Alba, 2023). Esta gran variedad de estilos y sabores de cervezas elaboradas por mujeres puede indicar que la percepción generalizada de la preferencia de las mujeres por cervezas dulces, afrutadas, poco amargas y ligeras es un constructo social que deja en evidencia el poco conocimiento que se tiene de las mujeres en el mercado de las cervezas en general y que refleja más bien los estereotipos de una población inexperta que más que un verdadero gusto del público femenino.

Adicionalmente, las redes sociales están creando nuevos espacios donde, principalmente, mujeres jóvenes pueden compartir su pasión y conocimiento sobre la cerveza artesanal. Ejemplos de esto son los grupos de Facebook como "Women in Beer" y "Crafty Beer Girls", y cuentas de Instagram como @girlswholikebeer, @thecraftbeergirl,

@msbeercraft y @comunidadmujerescerveceras. De Jong & Steadman (2021) destacan el impacto y empoderamiento que generan estos desarrollos en su estudio reciente sobre el evento *Independent Manchester Beer Convention* (Indy Man) en Manchester, Inglaterra. Durante este evento, se utilizaron hashtags de redes sociales como #womeninbeer y #beeryladies para coordinar una reunión de grupos de mujeres británicas aficionadas a la cerveza artesanal, apropiándose de una sección del lugar del evento como un proceso de territorialización tanto en línea como de manera presencial en el evento (de Jong & Steadman, 2021).

El desarrollo de la encuesta detallada en la Tabla 2 es una iniciativa clave para recolectar datos que permitan entender mejor este segmento de mercado con apoyo de los modelos de *Machine Learning* propuestos. Esta encuesta abarca una amplia gama de variables demográficas y comportamentales, lo que permite que la propuesta de un modelo predictivo se convierta en una herramienta robusta que, de ser aprovechada por el mercado de la cerveza artesanal, permitiría explotar un segmento del mercado poco explorado. La decisión de utilizar un enfoque de simulación de datos en *Python* responde a las limitaciones temporales del proyecto. Esta metodología permite generar un volumen considerable de datos sintéticos basados en la estructura de la encuesta propuesta, dando la posibilidad de proponer una guía tangible para futuros investigadores o incluso para las mismas productoras de cerveza artesanal que deseen ahondar en este nicho.

Conclusiones y Trabajo Futuro

Conclusiones

Las mujeres enfrentan barreras y desafíos significativos en el consumo de cerveza artesanal debido a estereotipos de género y percepciones culturales que históricamente han considerado la cerveza como una bebida masculina. A pesar de esto, el interés y consumo de cerveza artesanal por parte de las mujeres está en aumento, sugiriendo un cambio hacia un estado más equitativo en este mercado. Las mujeres tienden a preferir cervezas menos amargas y con sabores bien balanceados, mostrando una mayor apertura a explorar distintos estilos, incluyendo aquellos considerados masculinos como las IPAs. La promoción de cervezas que reflejen autenticidad, sostenibilidad y comunidad, junto con una mayor representación femenina en la producción y marketing de cerveza, pueden ayudar a superar estas barreras y atraer a más mujeres al mercado de la cerveza artesanal, contribuyendo a una cultura cervecera más inclusiva y diversa. Existe un gran potencial para captar nuevos consumidores de este segmento de la población, y descubrir qué atrae a las consumidoras femeninas podría crear enormes oportunidades para que los cerveceros puedan obtener una mayor cuota de mercado en el sector de las bebidas alcohólicas. En Colombia, hay pocos estudios sobre los gustos y preferencias de las mujeres en cervezas artesanales, por lo que emplear el cuestionario y el modelo predictivo desarrollado en este proyecto servirá de guía a las empresas productoras de cerveza artesanal en la creación de productos y espacios que se alineen con las preferencias y necesidades reales de la población femenina colombiana.

La encuesta detallada en la investigación se diseñó cuidadosamente para abarcar una amplia gama de variables demográficas y comportamentales. Estas incluyen aspectos como la edad, nivel educativo, ingresos, hábitos de consumo, preferencias de sabor, y motivaciones detrás del consumo de cerveza artesanal. La amplitud de la encuesta

asegura una comprensión integral de los factores que influyen en el consumo de cerveza artesanal en mujeres, permitiendo así generar una simulación relevante de datos aleatorios para la proposición, creación y entrenamiento de los modelos de ML, que han sido un componente central de esta investigación, brindando una guía práctica, replicable y explicada que no solo plantea y crea modelos genéricos, sino que además brinda recomendaciones y buenas prácticas a la hora de trabajar con valores reales permitiendo así en el futuro el crecimiento de productoras de cerveza que estén interesadas en trabajar en este sector y también a futuros investigadores que quieran recopilar la guía pactada para el presente proyecto.

Trabajo futuro

Para avanzar en la comprensión de las preferencias de sabor y estilo de cerveza artesanal entre las mujeres, es crucial realizar estudios empíricos específicos en contextos como Colombia, explorar estrategias de marketing que eviten estereotipos de género, y analizar el impacto de la representación femenina en la publicidad y producción de cerveza. Además, se debe investigar la influencia de factores socioculturales y demográficos en las preferencias de consumo y las experiencias de las mujeres en la industria cervecera artesanal. Esto permitirá identificar y abordar barreras, y promover una mayor inclusión y diversidad en este mercado creciente. Adicionalmente si se dispone de suficiente tiempo se puede recopilar una muestra relevante de encuestas que permita ejecutar la propuesta de ML y determinar cuáles son los factores que tienen más relevancia a la hora de revisar este sector.

Referencias

- 3cordilleras. (s.f.). 3 Cordilleras rosada: Dulce como siesta sin alarma. <https://3cordilleras.com/rosada-2/>
- Alba, J. (2023). Conozca a las mujeres cerveceras de Colombia que han revolucionado la industria. Revista Diners. https://revistadiners.com.co/estilo-de-vida/gastronomia/77963_mujeres-cerveceras-en-colombia/
- Ambar. (2019). *Desmontando mitos y leyendas sobre el consumo de cerveza por mujeres*. Recuperado de <https://ambar.com/noticias/mujeres-cerveceras/desmontando-mitos-y-leyendas/>
- Aprende Python. (s.f.). Aprende Python. Obtenido de Listas: <https://aprendepython.es/core/datastructures/lists/>
- Arias, E. R. (2020). Economipedia. Retrieved from Investigación mixta: <https://economipedia.com/definiciones/investigacion-mixta.html>
- Ayres, K. L., Harder, V. S., Rose, G. L., & Helzer, J. E. (2011). Drinking and stress: An examination of sex and stressor differences using IVR-based daily data. *Drug and Alcohol Dependence*, 115(1-2), 205-212. <https://doi.org/10.1016/j.drugalcdep.2010.10.022>
- Baiano, A. (2020). Craft beer: An overview. *Comprehensive Reviews in Food Science and Food Safety*, 20(2), 1829–1856. <https://doi.org/10.1111/1541-4337.12693>
- Bancolombia. (2024). Agenda tributaria 2024. Recuperado de <https://www.bancolombia.com/negocios/actualizate/legal-y-tributario/agenda-tributaria-2024>
- Banco de la República. (2024). Minutas de junio 2024. <https://www.banrep.gov.co/es/noticias/minutas-banrep-junio-2024>
- Banco Mundial. (2024). Colombia - Información general. Recuperado de <https://www.bancomundial.org/es/country/colombia/overview>.

BBC cervecería (2024). Cerveza rosada: Rosé.

<https://www.bbccerveceria.com/cerveza-rosada>

Becerra Elejalde, L. L. (2024). A qué ritmo deberían bajar las tasas de interés. Cambio Colombia. Recuperado de <https://cambiocolombia.com/economia/ritmo-deberian-bajar-tasas-de-interes>

Betancur, M. I., Motoki, K., Spence, C., & Velasco, C. (2020). Factors influencing the choice of beer: A Review. *Food Research International*, 137, 109367. <https://doi.org/10.1016/j.foodres.2020.109367>

Bobadilla, J. (2020). *Machine Learning y Deep Learning usando Python, Scikit y Keras* (Vol. 1). Bogotá, Colombia: Ra-ma Editorial y Ediciones de la U.

Brownlee, J. (2019). Machine Learning Mastery. Obtenido de How to use Learning Curves to Diagnose Machine Learning Model Performance: <https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/>

Brownlee, J. (2020). Machine Learning Mastery. Obtenido de Ordinal and One-Hot Encodings for Categorical Data: <https://machinelearningmastery.com/one-hot-encoding-for-categorical-data/>

Calle, A. M., & Giraldo, L. M. (2020). *Motivos de consumo de cerveza entre dos grupos etarios de consumidores en Santiago de Cali* [Tesis de maestría, Universidad Icesi]. Repositorio Institucional Universidad Icesi. https://repository.icesi.edu.co/biblioteca_digital/bitstream/10906/87453/1/TG02917.pdf

Colombia. (1991). Constitución Política de Colombia. Artículo 333.

Calvo-Porrá, C., Orosa-González, J., & Blazquez-Lozano, F. (2018). A clustered-based segmentation of beer consumers: From “Beer lovers” to “beer to fuddle.” *British Food Journal*, 120(6), 1280–1294. <https://doi.org/10.1108/bfj-11-2017-0628>

Challenger, I., Díaz, Y., & Becerra, R. (2014). Google Scholar. Obtenido de Redalyc: <https://www.redalyc.org/pdf/1815/181531232001.pdf>

Chandra, R. (2023). Ecoagi. Obtenido de ¿Qué es Scikit-Learn: La biblioteca de aprendizaje automático imprescindible: <https://ecoagi.ai/es/topics/Python/what-is-sklearn>

Chapman, N.G., Nanney, M., Lellock, J. S., & Mikles-Schluterman, J. (2018). Bottling gender: Accomplishing gender through craft beer consumption. *Food, Culture & Society*, 21(3), 296-313. <https://doi.org/10.1080/15528014.2018.1451038>

Cunha, A., Lopes, N., Barcia, M., Sautter, C., & Augusto Ballus, C. (2023). Production and characterization of craft beers with different additions of native fruits and agro-industrial residues: A Review. *Ciência Rural*, 53(9). <https://doi.org/10.1590/0103-8478cr20220194>

DANE. (2024). Estadísticas de Empleo y desempleo en Colombia. Departamento Administrativo Nacional de Estadística. <https://www.dane.gov.co/index.php/estadisticas-por-tema/mercado-laboral/empleo-y-desempleo>.

Darwin, H. (2018). Omnivorous masculinity: Gender capital and cultural legitimacy in craft beer culture. *Social Currents*, 5(3), 301-316. <https://doi.org/10.1177/2329496517748336>

de Jong, A., & Steadman, C. (2021). (Re) crafting belonging: Cultural-led regeneration, territorialization and craft beer events (pp. 1–19). *Social & Cultural Geography*.

Deloitte. (2017). *Cerveza artesanal en México: Análisis de situación y perspectivas 2017*. Deloitte México. <https://www2.deloitte.com/content/dam/Deloitte/mx/Documents/consumer-business/2017/Cerveza-Artesanal-Mexico-2017.pdf>

Delwiche, J. (2004). The impact of perceptual interactions on perceived flavor. *Food Quality and Preference*, 15, 137–146.

Donadini, G., Fumi, M. D., Kordialik-Bogacka, E., Maggi, L., Lambri, M., & Sckokai, P. (2016). Consumer interest in specialty beers in three European markets. *Food Research International*, 85, 301–314. <https://doi.org/10.1016/j.foodres.2016.04.029>

El Tiempo (2007). Conquistar a las mujeres, el nuevo reto de Bavaria Redd's hace parte del portafolio de las nuevas marcas de la firma cervecera. <https://www.eltiempo.com/archivo/documento/MAM-2688605>

Escobar Fernández, J. (2023). Oxford Economics alertó que reformas del gobierno Petro pueden aumentar el riesgo económico en Colombia. Infobae. Recuperado de <https://www.infobae.com/colombia/2023/12/06/oxford-economics-alerto-que-reformas-del-gobierno-petro-pueden-aumentar-el-riesgo-economico-en-colombia/>

Euromonitor (2024). *Passport: Beer market in Colombia*. Euromonitor International. <https://www.euromonitor.com>

Eyre, I. (2024). Real Python. Obtained Visualizing Data in Python With Seaborn: <https://realpython.com/python-seaborn/#introducing-seaborns-contemporary-objects-interface>

Frankel, S., Benjamin, S., & Stephens, C. (2021). Crafty women: Exploring how southeastern female brewers navigate emotional labour within the craft beer industry. *Annals of Leisure Research*, 26(3), 372–395. <https://doi.org/10.1080/11745398.2021.1902356>

Forbes. (2023). Venta de cervezas artesanales crece 51% en Colombia. Forbes Colombia. <https://forbes.co/2023/10/04/actualidad/venta-de-cervezas-artesanales-crece-51-en-colombia>

Gabrielyan, G., McCluskey, J. J., Marsh, T. L., & Ross, C. F. (2014). Willingness to pay for sensory attributes in beer. *Agricultural and Resource Economics Review*, 43(1), 125–139. <https://doi.org/10.1017/s1068280500006948>

Gavilán, M., & Muniesa, J. (2023). Cinco tendencias de consumo de la cerveza artesana. *Gastronomía & Cía*. <https://www.gastronomiaycia.com/cinco-tendencias-de-consumo-de-la-cerveza-artesana/>

Gómez-Corona, C., Valentin, D., Escalona-Buendía, H. B., & Chollet, S. (2017). The role of gender and product consumption in the mental representation of Industrial and Craft Beers: An exploratory study with Mexican consumers. *Food Quality and Preference*, 60, 31–39. <https://doi.org/10.1016/j.foodqual.2017.03.008>

Harikrishnan (2019). Medium. Obtained Confusion Matrix, Accuracy, Precision, Recall, F1 Score: <https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd>

Hernández Naranjo, D. (2023). Reforma tributaria: se viene apretón en 2024 a los que ganen más de 13 millones. *Portafolio*. Recuperado de <https://www.portafolio.co/economia/reforma-tributaria/reforma-tributaria-se-viene-apreton-en-2024-a-los-que-ganen-mas-de-13-millones-594754>.

Horowiec, J. (2022). Beer Preferences of Women: Looking at Gender Stereotypes through the Consumption of Craft Beer. Johnson & Wales University, Hospitality Graduate Student Scholarship. https://scholarsarchive.jwu.edu/cgi/viewcontent.cgi?article=1028&context=hosp_graduate

IBM. (s.f.). IBM. Retrieved from ¿Qué es machine learning (ML)?: <https://www.ibm.com/mx-es/topics/machine-learning>

Impactotic. (2024). Internet en zonas rurales en Colombia. Impactotic. <https://impactotic.co/innovacion/internet-en-zonas-rurales-en-colombia/>

Kappele A (2015). A study of advertising: The role of gender representations on craft beer labels. MA Thesis, University of Missouri, Columbia, MO.

Kuehn, K. M., & Parker, S. (2018). One of the blokes: Brewsters, branding and gender (in)visibility in New Zealand's craft beer industry. *Journal of Consumer Culture*, 21(3), 519–538. <https://doi.org/10.1177/1469540518806956>

Land, C., Sutherland, N., & Taylor, S. (2018). 'Back to the brewster: Craft brewing, gender and the dialectical interplay of retraditionalisation and innovation' in bell. In E. , G. Mangia, S. Taylor, & M. L. Toraldo (Eds.), *The organization of craft work: Identities, meanings, and materialities* (pp. 150–168). Routledge.

Las Dos Orillas. (2024, febrero 27). La única universidad en Colombia que tiene un pregrado en inteligencia artificial. https://www.las2orillas.co/la-unica-universidad-en-colombia-que-tiene-un-pregrado-en-inteligencia-artificial/#google_vignette

Lazzari, A., Barbosa, H. D., Filho, E. R., Dada, A. P., Saraiva, B. R., & Matumoto-Pintro, P. T. (2021). Gender behavior and influence in acceptability of beers produced with Rubim and Mastruz. *Journal of Sensory Studies*, 37(2). <https://doi.org/10.1111/joss.12731>

Llerena Izquierdo, J. (2020). *Codifica en Python* (Vol. 1). Quito, Ecuador: Universidad Politécnica Salesiana.

Lombo Delgado, J. S. (2024). *Reforma tributaria sería llevada al Congreso en 2025, anunció el presidente Petro*. *El Tiempo*. <https://www.eltiempo.com/politica/gobierno/reforma-tributaria-seria-llevada-al-congreso-en-2025-anuncio-el-presidente-petro-3333497>

Mendoza, D. (2024). *Reforma tributaria sería llevada al Congreso en 2025, anunció el presidente Petro*. *El Tiempo*. <https://www.eltiempo.com/politica/gobierno/reforma-tributaria-seria-llevada-al-congreso-en-2025-anuncio-el-presidente-petro-3333497>

Ministerio de Tecnologías de la Información y las Comunicaciones. (2024). *Datos Abiertos*. Obtenido de Departamentos y municipios de Colombia:

https://www.datos.gov.co/Mapas-Nacionales/Departamentos-y-municipios-de-Colombia/xdk5-pm3f/about_data

Muggah, E. & McSweeney, M. (2017a). Using preferred attribute elicitation to determine how males and females evaluate beer. *Journal of Food Science*, 82(8), 1916–1923. <https://doi.org/10.1111/1750-3841.13799>

Muggah, E. & McSweeney, M. (2017b). Females' attitude and preference for beer: A Conjoint Analysis Study. *International Journal of Food Science & Technology*, 52(3), 808–816. <https://doi.org/10.1111/ijfs.13340>

Murcia, J. D. (2024). *Los nueve puntos que propone Petro discutir en el proceso de una constituyente*. La República. <https://www.larepublica.co/economia/los-nueve-puntos-que-propone-petro-discutir-en-el-proceso-de-una-constituyente-3830656>

Nanney, M., Chapman, N.G., Lellock, J.S., & Mikles-Schluterman, J. (2020). Gendered expectations, gatekeeping, and consumption in craft beer spaces. *Humanity & Society*, 44(4), 449-468. <https://doi.org/10.1177/0160597620932888>

Narula, M. (2021). DelftStack. Obtenido de Mapa de calor de correlación en Seaborn: <https://www.delftstack.com/es/howto/seaborn/correlation-heatmap-seaborn-python/>

Navone, E. (2023). FreeCodeCamp. Obtenido de ¿Qué es programación? Manual para principiantes: <https://www.freecodecamp.org/espanol/news/que-es-programacion-manual-para-principiantes/>

Nave, E., Duarte, P., Rodrigues, R., Paço, A., Alves, H., & Oliveira, T. (2021). Craft beer – A systematic literature review and Research Agenda. *International Journal of Wine Business Research*, 34(2), 278–307. <https://doi.org/10.1108/ijwbr-05-2021-0029>

pandas. (2024). pandas. Obtenido de About pandas: <https://pandas.pydata.org/about/index.html>

pandas. (2024). pandas. Obtenido de pandas.DataFrame: <https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.html>

Pink Boots Society. (2024). *About us*. <https://www.pinkbootsociety.org/about-us/>

Poon, S.T.F. (2021). Objectification and sensibility: A critical look at sexism as subtext in postfeminist advertising. *Asian Social Science*, 17(2), 17-30.

Portafolio. (2024). Analistas bajan su expectativa de crecimiento del PIB para Colombia en 2024. Recuperado de <https://www.portafolio.co/economia/finanzas/analistas-bajan-su-expectativa-de-crecimiento-del-pib-para-colombia-en-2024-600923>

Rao, H. (2008). *Market Rebels: How Activists Make or Break Radical Innovations*. Princeton University Press. Retrieved from <https://doi.org/10.1515/9781400829743>

Rico Muñoz, A. (2023). Las nuevas claves legales que deben tener en cuenta las empresas y sectores en 2023. *Asuntos Legales*. Recuperado de <https://www.asuntoslegales.com.co/consumidor/las-nuevas-claves-legales-que-deben-tener-en-cuenta-las-empresas-y-sectores-en-2023-3531413>.

Rivaroli, S., Lindenmeier, J., & Spadoni, R. (2020). Is craft beer consumption genderless? exploratory evidence from Italy and Germany. *British Food Journal*, 122(3), 929–943. <https://doi.org/10.1108/bfj-06-2019-0429>

Roch, E. (s.f.). *lovtechnology*. Obtenido de ¿Qué es NumPy: cómo funciona y para qué sirve?: <https://lovtechnology.com/que-es-numpy-como-funciona-y-para-que-sirve/#:~:text=NumPy%20es%20una%20biblioteca%20de%20Python%20gratuita%20y,a s%C3%AD%20como%20una%20amplia%20gama%20de%20operaciones%20matem%C3%A1ticas>.

Roncancio, O. (2024). *Github*. Obtenido de FactoresConsumo: https://github.com/OscarRoncancio/FactoresConsumo/blob/main/Simulacion_Datos.ipyn

Ruiz, L. (2024). Informe de Moody's ve menor crecimiento e inflación en 2024 y 2025 para Colombia. *Valora Analitik*. Recuperado de <https://www.valoraanalitik.com/2024/02/05/informe-de-moodys-ve-menor-crecimiento-e-inflacion-en-2024-y-2025-para-colombia/>

Rydzik, A., & Ellis-Vowles, V. (2019). 'Don't use "the weak word"': Women brewers, identities and gendered territories of embodied work. *Work, Employment and Society*, 33(3), 483–499.

Saavedra, F. (2024). El Banco de la República de Colombia anticipó incremento en la inflación y la tasa de interés. Infobae. Recuperado de <https://www.infobae.com/colombia/2024/03/19/el-banco-de-la-republica-de-colombia-antipico-incremento-en-la-inflacion-y-la-tasa-de-interes/>

Scikit-Learn. (s.f.). Scikit-Learn. Retrieved from Decision Trees: <https://scikit-learn.org/stable/modules/tree.html>

Sklearn. (2024). Sklearn. Obtenido de train_test_split: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

Sklearn. (2024). Sklearn. Obtenido de GridSearchCV: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

Stewart, L. (S.f). Atlas.ti. Obtenido de ¿Qué es la investigación descriptiva y cómo se utiliza?: <https://atlasti.com/es/research-hub/investigacion-descriptiva>

Suárez Bernal, L. (2024). Internet en zonas rurales en Colombia. Impactotic. Recuperado de <https://impactotic.co/innovacion/internet-en-zonas-rurales-en-colombia/>

Spence, C. (2019a). Do men and women really live in different taste worlds? *Food Quality and Preference*, 73, 38–45. <https://doi.org/10.1016/j.foodqual.2018.12.002>

Spence, C. (2019b). Problems Associated with Marketing Food and Drink Specifically at Women. Preprints. <https://doi.org/10.20944/preprints201907.0034.v1>

Stewart, L. (S.f). Atlas.ti. Retrieved from ¿Qué es la investigación descriptiva y cómo se utiliza?: <https://atlasti.com/es/research-hub/investigacion-descriptiva>

Tak, E., Correll, S. J., & Soule, S. A. (2019). Gender inequality in product markets: When and how status beliefs transfer to products. *Social Forces*, 98(2), 548–577. <https://doi.org/10.1093/sf/soy125>

Thurnell-Read, T. (2022). Gender and craft drinks. *Sociology Compass*, 16(9).

<https://doi.org/10.1111/soc4.13018>

Unipyhton. (2017). Unipython. Obtenido de MATPLOTLIB: FUNCIONES PRINCIPALES: <https://unipython.com/matplotlib-funciones-principales/>

Wheelen & Hunger (2013). *Administración estratégica y política de negocios: Conceptos y casos*. Pearson. 12ª edición.

Withers, E. T. (2017). The impact and implications of Craft Beer Research: An interdisciplinary literature review. *Craft Beverages and Tourism*, Volume 1, 11–24.

https://doi.org/10.1007/978-3-319-49852-2_2