

UNIVERSIDAD DEL VALLE DE GUATEMALA

CC3084 Data Science

Sección 10

Msc. Lynette García Pérez



Laboratorio 9 - Visualización Interactiva

Consumo y precios de gasolina

José Pablo Orellana 21970

Diego Alberto Leiva 21752

GUATEMALA, 17 de octubre del 2024

Conjunto de Datos Seleccionado

Descripción del Conjunto de Datos

Este conjunto de datos provee información detallada sobre el consumo y precios de combustibles en Guatemala, recopilada desde el año 2021 hasta mayo de 2024. Utilizado previamente en los laboratorios del curso, esta versión limpia y preprocesada es ideal para un análisis exhaustivo de tendencias y patrones en el mercado de combustibles. Incluye información sobre productos clave como gasolina súper, gasolina regular y diésel, con un enfoque en el consumo en la Ciudad Capital. Los datos permiten analizar el impacto de eventos históricos, en el comportamiento de consumo e importación de combustibles.

Desglose de los Conjuntos de Datos

Consumo de Combustibles

- **Periodo:** Enero 2021 a mayo 2024 (mensual).
- **Variables:** Fecha, consumo de combustibles en barriles de 42 galones para gasolina regular, gasolina superior, diésel y gas propano.
- **Estadísticas:** Consumo promedio de gasolina regular y superior y diésel con valores de referencia para analizar variaciones y detectar patrones.

Precios Promedios al Consumidor Final en Ciudad Capital

- **Periodo:** Enero 2021 a julio 2024 (diarios).
- **Variables:** Fecha, precios de combustible (en quetzales por galón) para gasolina regular, gasolina superior y diésel, precio de gas propano (en quetzales por cilindro de 25 lbs).
- **Estadísticas:** Muestra la volatilidad y el precio promedio, especialmente en años recientes, donde se observan aumentos en el precio promedio y una alta volatilidad en 2022.

Razón de la Selección

Estos datos fueron seleccionados por su relevancia en el análisis de series temporales y su aplicabilidad en el desarrollo de modelos predictivos. Además, proporcionan una base sólida para mostrar la efectividad de distintos métodos de modelado, como LSTM.

Representaciones Visuales Relacionadas con el Tema de Gasolina

Para que dashboard sea visualmente atractivo y esté en sintonía con el tema, se utilizarán representaciones gráficas relacionadas con el combustible, como:

- Bidones de Gasolina
- Bombas de Gasolina
- Surtidores de Combustible

Estas imágenes ayudarán a contextualizar el contenido y captar el interés de una audiencia general además de que se pueden representar datos con estas mismas imágenes, lo que lo hará mucho más fácil de entender para el público que no posee conocimiento técnico al respecto o que no sabe interpretar gráficas de manera sencilla.

Paleta de Colores

Colores Utilizados en la Infografía

La paleta de colores ha sido seleccionada para mantener una estética profesional y facilitar la interpretación de los gráficos y textos. Los colores seleccionados son:

- **Fondo:** Un color que se encuentre en la escala de beige, que permite que los elementos gráficos destaquen sin distracciones.
- **Texto:** Colores azul (en todas sus tonalidades) o bien un color que contraste para garantizar legibilidad.

Elementos visuales: Utilizaremos gráficos relacionados con la gasolina en colores naturales como rojo, gris, negro, amarillo y verde.

Colores en los Gráficos

Se ha seleccionado una paleta específica para los gráficos que permita una diferenciación clara entre las categorías y las predicciones:

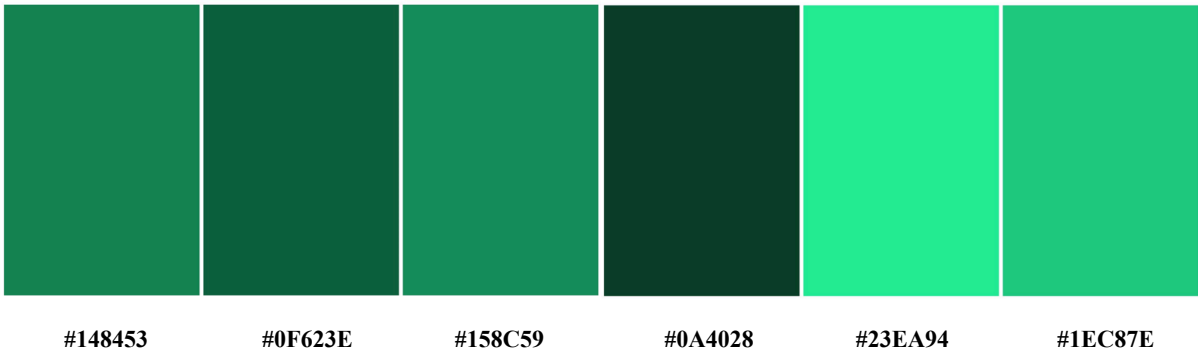
- **Azul y naranja:** Utilizados en los gráficos de líneas para diferenciar entre los datos históricos y las predicciones.
- **Verde, Rojo y Amarillo:** Asignados a los diferentes tipos de combustibles en los diagramas de barras apiladas, lo que hace evidente la segmentación de los datos.

Justificación de la Paleta de Colores

Los colores fueron escogidos para ofrecer suficiente contraste y asegurar que los gráficos sean fáciles de leer e interpretar. Además, se emplean colores que son naturalmente asociados con el tema del combustible, lo que refuerza la coherencia visual de la infografía.

Escala de colores a utilizar (Bidones y elementos gráficos)

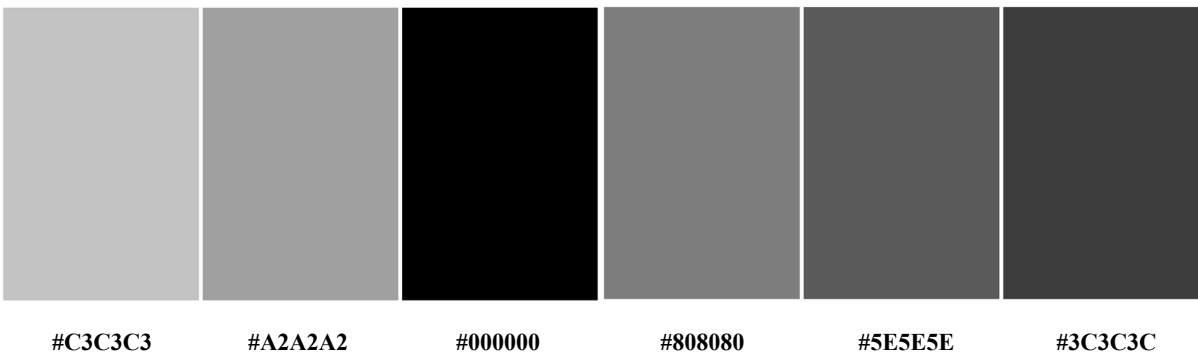
Gasolina Super (Verde)



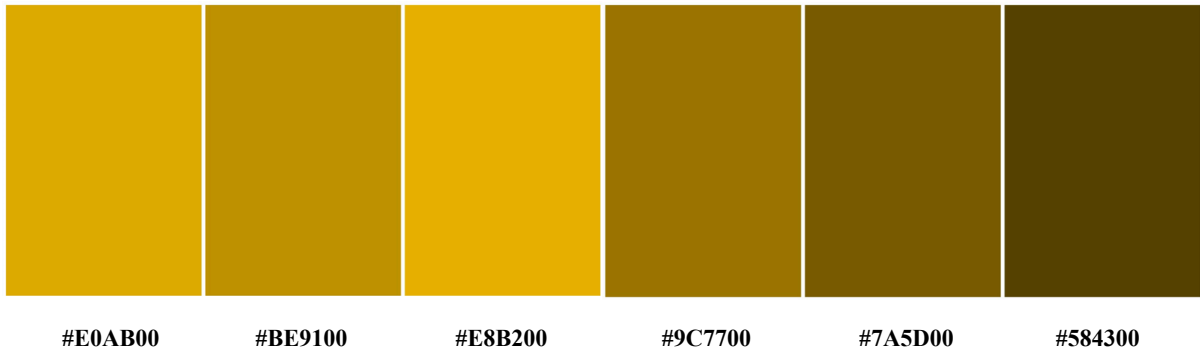
Gasolina Regular (Rojo)



Diesel (Negro)

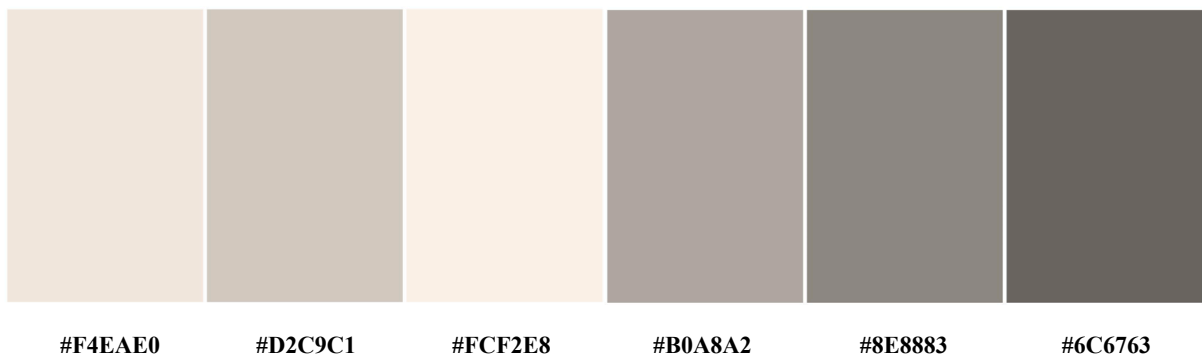


Color Auxiliar para Gasolina (Amarillo)

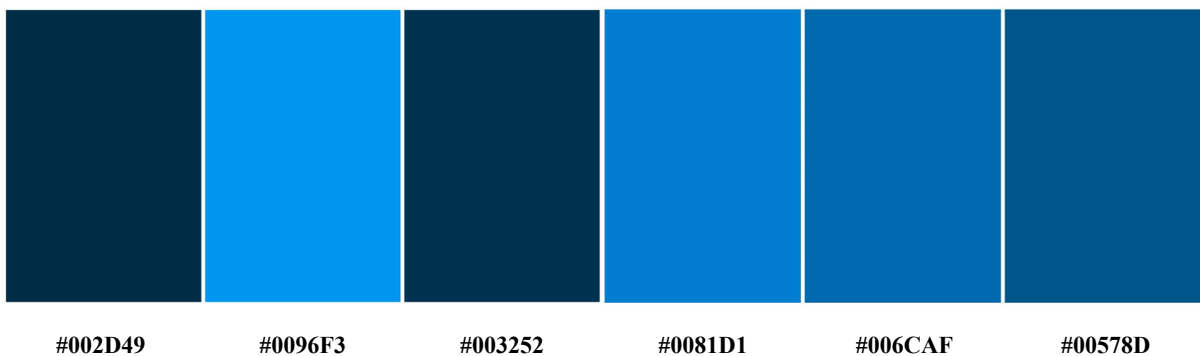


Escala de colores a utilizar (Elementos principales)

Paleta de Colores Para Fondo (Beige)

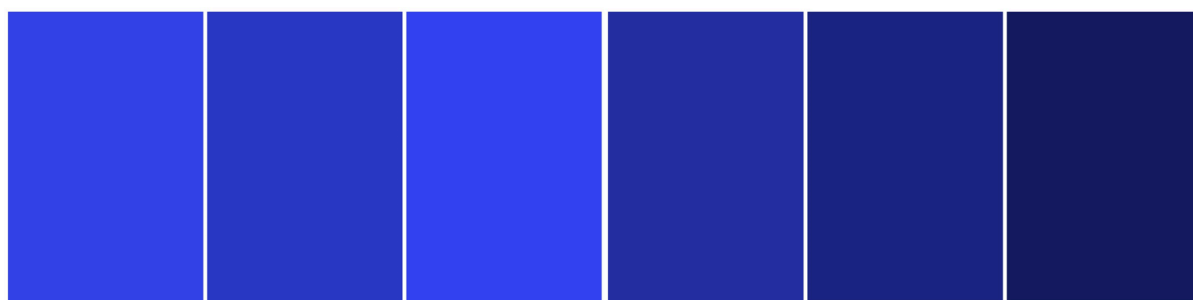


Paleta de Colores Para Texto (Azul)



Estas paletas de colores son para elementos básicos de la infografía, como lo pueden ser fondos, textos, etc.

Escala de Colores a Utilizar en Elementos Gráficos (Modelos)



#3342E7

#2B38C5

#3544F0

#242FA3

#1C2581

#151B5F



#E78A10

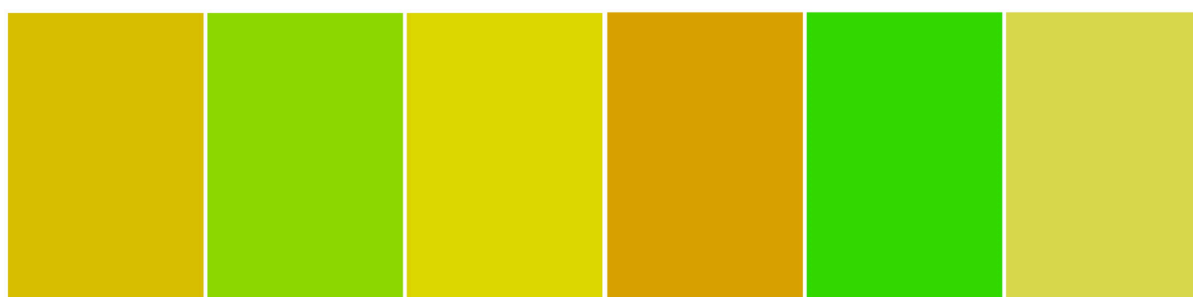
#C5760E

#F08F11

#A3610B

#814D09

#5F3907



#DBBD02

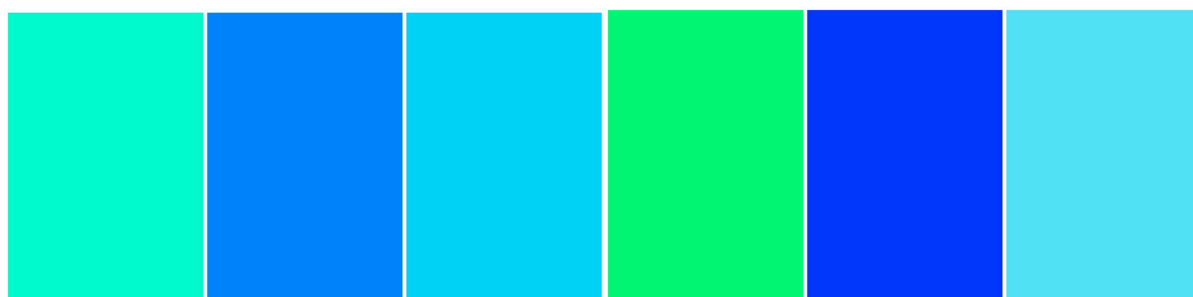
#8DDB02

#DCD703

#DBA302

#32DB02

#DBD94B



#00FACD

#0085FA

#00D3FA

#00FA78

#0037FA

#53E1FA

Como se mencionó anteriormente en el documento, se estarán utilizando llamativos para representar gráficas predictivas y que el contraste de colores muestre de mejor manera cómo se comportan dichas gráficas.

Planificación de las tareas para el grupo

La planificación de las tareas se enfoca en el análisis de datos y la creación de modelos de aprendizaje automático para entender mejor las tendencias de consumo y los precios de la gasolina. Aquí tenemos el proceso a detalle.

Preparación y limpieza de datos:

- Se utilizará un conjunto de datos que incluye información detallada sobre el consumo de gasolina y los precios a lo largo de un período específico. Esta etapa inicial incluye la limpieza y transformación de los datos.
- Se revisará la presencia de datos nulos o inconsistencias que puedan afectar el rendimiento de los modelos. Se tomará la limpieza del laboratorio 2.

Implementación de modelos de aprendizaje automático:

- Se desarrollarán tres modelos de aprendizaje automático para analizar y predecir el comportamiento del consumo y la variación de precios de la gasolina. Hasta el momento, ya se ha implementado un modelo (LSTM), por lo que el siguiente paso es completar los dos modelos restantes. Los modelos a implementar podrían incluir técnicas como modelos ARIMA o SARIMA (Modelos tentativos).

Implementación de visualizaciones en dashboard:

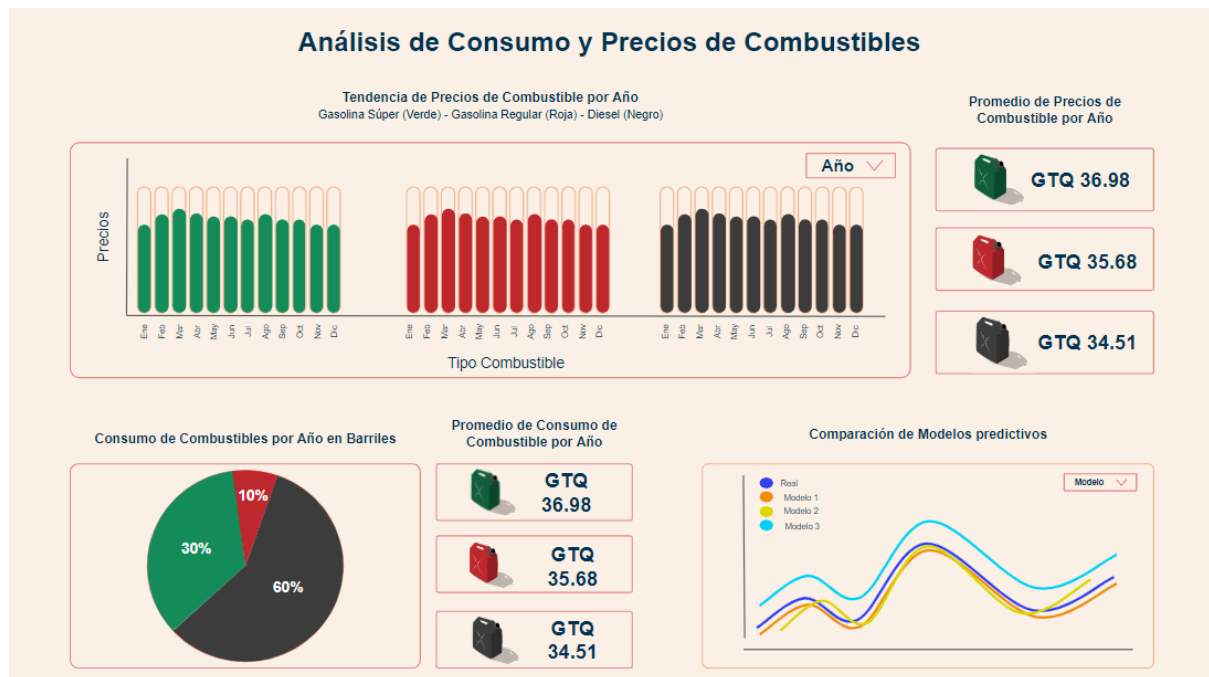
- Se diseñarán visualizaciones interactivas para presentar los resultados del análisis de datos y las predicciones de los modelos. Estas visualizaciones serán integradas en un dashboard que permitirá a los usuarios explorar y entender la evolución del consumo de gasolina y sus precios de forma dinámica.
- El diseño del dashboard seguirá los lineamientos planteados a lo largo de este informe, asegurando que sea intuitivo, fácil de usar y que presente la información de forma clara y concisa.

Pruebas y ajustes finales:

Antes de la entrega final, se realizarán pruebas exhaustivas del dashboard y los modelos para garantizar que todas las funcionalidades estén operativas y los resultados sean precisos.

Se realizarán ajustes en las visualizaciones y en los parámetros de los modelos si se detectan áreas de mejora, asegurando que el producto final cumpla con los objetivos planteados y sea capaz de aportar valor al análisis de datos sobre el consumo y los precios de la gasolina.

Bosquejo de diseño



Enlace para visualizar prototipado

<https://docs.google.com/presentation/d/1Nh239HlHIJup9e02dLorDelc2PIxpOHQJJbUcEw9mN0/edit?usp=sharing>

Selección de la herramienta

Durante el desarrollo de nuestro dashboard, evaluamos diversas opciones para crear visualizaciones interactivas que pudieran integrarse de manera efectiva en una plataforma accesible para los usuarios. Específicamente, realizamos pruebas con dos de las herramientas proporcionadas para este laboratorio: Python (Dash) y R (Shiny).

Aunque ambas herramientas ofrecían funcionalidades interesantes, finalmente nos decidimos por Dash, desarrollado en Python, debido a varios motivos. En primer lugar, nuestra familiaridad con Python fue un factor clave en la elección. Python es un lenguaje que utilizamos de manera habitual en proyectos de análisis de datos por su flexibilidad y el amplio ecosistema de bibliotecas que simplifican el procesamiento y la visualización de datos. Esto nos permitió acelerar el proceso de desarrollo y garantizar un manejo eficiente de las distintas partes del dashboard.

Además, tras realizar las pruebas, observamos que Dash nos ofrecía mayores posibilidades para construir visualizaciones más completas y estéticamente agradables. Es por ello que nos inclinamos por la utilización de dicha herramienta.

Enlace GitHub: <https://github.com/JPOrellana/Lab9-DataScience>