 

**PIA – Entregable 4**

**Universidad Autónoma de Nuevo León**

**Facultad:** Facultad de Ciencias Físico Matemáticas

**Materia:** Programación básica

**Nombre del maestro:** Perla Marlene Viera González

**Grupo:** 073

**Fecha de entrega:** 18 de mayo de 2025

| **Matricula** | **Nombre** |
| --- | --- |
| 2136071 | Santiago Resendiz Badillo |
| 1909518 | Victor Adrian Rodriguez Ortiz |

A continuación, se presenta la documentación detallada para el último código, con especial énfasis en la selección y el razonamiento detrás de cada tipo de gráfica utilizada. Se explica cómo cada gráfica aporta a la comprensión del comportamiento de los asteroides durante 2024, se incluyen ejemplos de los resultados obtenidos y se incorporan algunas notas de código relevantes.

1. **Propósito General del Script**

Este script está diseñado para analizar los datos sobre asteroides cercanos obtenidos previamente en el archivo asteroides\_2024.txt, y transformarlos en información estadística y visual que permita:

* Evaluar y comparar características clave como velocidad, tamaño, distancia a la tierra y peligrosidad.
* Exportar los datos procesados en un archivo Excel.
* Generar múltiples tipos de gráficas que faciliten la interpretación de:
* Promedios mensuales.
* Distribución de tamaños y velocidades.
* Cantidades relativas de asteroides por mes y por peligrosidad.

1. **Lectura y Estructuración de Datos**

El script inicia leyendo el archivo de texto en el que cada asteroide está representado mediante bloques de líneas con formato clave-valor. Se realiza lo siguiente:

* **Lectura línea por línea:** Se extrae la información agrupando los datos de cada asteroide.
* **Procesamiento de valores:** Se identifican y separan las claves y sus respectivos valores, convirtiendo las cadenas numéricas a tipo float para realizar cálculos posteriores (tamaños, velocidades y distancias).
* **Almacenamiento:** Cada asteroide se guarda en un diccionario y se agrupa en una lista que luego se exporta a Excel, facilitando una revisión estructurada de la información.

**Nota:** Se manejan bloques de datos separados por líneas en blanco, lo que permite identificar cada grupo de datos de forma ordenada.

1. **Cálculos Estadísticos**

Utilizando la biblioteca statistics, el script calcula:

* **Media:** Para obtener el valor promedio de la velocidad, el tamaño y la distancia, se usan las funciones mean().
* **Mediana:** Se calcula con median() para determinar el valor central de las distribuciones.
* **Desviación estándar:** Con stdev() se evalúa la dispersión de los datos respecto a la media.

Estos cálculos permiten identificar características típicas y la variabilidad de los asteroides, lo que es fundamental para interpretar su comportamiento.

1. **Exportación a Excel**

El procesamiento de datos se complementa con la exportación a un archivo Excel (asteroides.xlsx) mediante la biblioteca openpyxl. Se realiza lo siguiente:

* **Encabezados y filas:** Se escriben los nombres de los campos en la primera fila y cada asteroide se almacena como una fila posterior.
* **Formato organizado:** Esto facilita la revisión y permite la integración de los datos en el archivo de Excel.

1. **Selección y Justificación de Gráficas**

La metodología para la selección de los tipos de gráficas se basa en la naturaleza de los datos y en el objetivo de resaltar las características relevantes para el análisis. A continuación, se describen cada uno de los tipos de gráficas implementadas:

1. **Gráfica de Barras – Promedio de Velocidad por Mes**

Metodología y Razonamiento:

* **Elección del gráfico:** La gráfica de barras se eligió porque permite comparar fácilmente los promedios de velocidad de cada mes.
* **Ventajas:** Visualmente es clara para detectar variaciones mensuales.

Permite identificar meses en los que los asteroides presentan mayor o menor velocidad.

1. **Gráficas de Líneas – Tamaño de Asteroides (Mínimo y Máximo) por Mes**

Metodología y Razonamiento:

* **Elección del gráfico:** Se usan gráficas de líneas para visualizar la tendencia a lo largo de los meses en cuanto a los extremos (mínimo y máximo) del tamaño de los asteroides.
* **Ventajas:** Permiten observar cómo varían los tamaños extremos (más pequeños y más grandes) a lo largo del año.

Ayudan a identificar patrones o anomalías en determinados meses.

1. **Diagrama de Dispersión – Tamaño versus Velocidad**

Metodología y Razonamiento:

* **Elección del gráfico:** El diagrama de dispersión es ideal para identificar correlaciones entre del tamaño y la velocidad.
* **Ventajas:** Permite determinar si existe relación entre el tamaño de un asteroide y su velocidad.

Facilita la detección de relaciones entre agrupaciones de datos.

1. **Gráficas de Pastel – Distribución de Asteroides por Mes y por Peligrosidad**

Metodología y Razonamiento:

* **Elección del gráfico:** Las gráficas de pastel fueron seleccionadas para representar porcentajes y proporciones, ofreciendo una visión inmediata de la distribución de asteroides tanto por mes como por su peligrosidad.
* **Ventajas:** Resultan útiles para mostrar de manera visual la contribución relativa de cada categoría.

Facilitan la comprensión de la información proporcionada por los asteroides en cada mes y de la proporción entre asteroides peligrosos y no peligrosos.

1. **Retos enfrentados en el análisis**

Durante el desarrollo de esta etapa del proyecto, centrada en el análisis gráfico y visual de los datos de asteroides, se presentaron varios desafíos, que se resolvieron progresivamente a través de pruebas:

1. **Selección adecuada del tipo de gráfica**

Uno de los principales retos fue elegir el tipo de gráfica más apropiado para representar cada tipo de dato. Al principio, no era evidente si usar barras, líneas, pastel o dispersión. Se requirió:

* Comparar ejemplos y documentación.
* Entender las diferencias entre variables como los meses y numéricas (como la velocidad o el tamaño).
* Probar distintas opciones hasta encontrar la más clara y útil para cada caso.

1. **Manipulación de datos agrupados por mes**

Para representar promedios o extremos por mes, fue necesario:

* Recorrer la lista de asteroides y agruparlos según el mes.
* Manejar correctamente los datos para calcular estadísticas individuales por grupo.
* Asegurarse de que los meses estuvieran en orden cronológico en las gráficas.

1. **Personalización de las gráficas con Matplotlib**

Inicialmente, las gráficas generadas eran muy básicas o difíciles de leer. Fue necesario aprender a:

* Ajustar títulos, ejes, etiquetas, márgenes y límites (plt.axis()).
* Añadir porcentajes en gráficas de pastel usando autopct='%1.1f%%'.

NOTA: Este parámetro permite mostrar los porcentajes con un decimal, mejorando la legibilidad.

* Elegir colores y marcadores adecuados para diferenciar bien los datos.

1. **Diagramas de dispersión y relación entre variables**

Comprender cómo usar plt.scatter() correctamente fue un reto, especialmente al decidir qué variables graficar y cómo interpretar la dispersión:

* Fue necesario probar combinaciones hasta lograr una visualización significativa (por ejemplo, tamaño vs. velocidad).
* También se exploró si existían agrupaciones o tendencias visuales.

1. **Interpretación de resultados visuales**

No bastaba con generar las gráficas, fue necesario aprender a leerlas para identificar:

* Patrones mensuales.
* Relación entre variables como el mes en el eje x y variables como el tamaño, velocidad, etc. en el eje y.