**🧠 1. Define tu pregunta de investigación**

Antes de modelar o visualizar, necesitas saber qué quieres responder. Algunas ideas potentes:

* ¿Cómo ha evolucionado el gasto por ministerio o programa en los últimos años?
* ¿Qué partidas muestran mayor rigidez o volatilidad presupuestaria?
* ¿Existe correlación entre el gasto en ciertos capítulos (como inversiones reales) y variables macroeconómicas?
* ¿Qué ministerios concentran el mayor peso en transferencias internas y cómo varía eso?

**🧰 2. Preprocesamiento y limpieza**

Ya que estás trabajando con series roja y verde, asegúrate de:

* Homogeneizar nombres de columnas y codificaciones.
* Convertir importes a euros reales si vas a comparar años (ajustar por inflación).
* Crear claves únicas por servicio-programa-económica para facilitar agrupaciones.

**📊 3. Análisis exploratorio (EDA)**

Aquí es donde empiezas a descubrir patrones:

* Histogramas de distribución por capítulo económico.
* Series temporales por servicio o programa.
* Heatmaps de correlación entre partidas.
* Boxplots para ver dispersión por tipo de gasto.

Usa librerías como pandas, matplotlib, seaborn, y si quieres algo más interactivo, plotly.

**🧮 4. Modelado estadístico o machine learning**

Dependerá de tu enfoque, pero aquí tienes opciones:

* **Clustering**: agrupa servicios o programas según similitud en estructura de gasto (KMeans, DBSCAN).
* **PCA**: reducción de dimensionalidad para entender qué variables explican más la variabilidad.
* **Regresión**: predicción de partidas futuras o análisis de impacto de una variable sobre otra.
* **Análisis de series temporales**: si tienes datos de varios años, puedes aplicar modelos como ARIMA, Prophet, etc.

**📈 5. Visualización avanzada**

Para el TFM, las visualizaciones son clave:

* **Treemaps** para ver distribución presupuestaria por capítulos.
* **Sunburst charts** para jerarquías (Ministerio → Servicio → Programa → Económica).
* **Dashboards** con Streamlit o Dash si quieres algo interactivo.

**🧾 6. Narrativa y conclusiones**

No te limites a mostrar gráficos. Interpreta:

* ¿Qué implicaciones tienen tus hallazgos para la eficiencia del gasto público?
* ¿Qué recomendaciones podrías hacer como analista?
* ¿Qué anomalías o patrones inesperados has detectado?

**🧠 Bonus: Enfoques alternativos**

Si quieres darle una vuelta más creativa:

* **Presupuestos como red**: modela las transferencias entre servicios como grafos (networkx).
* **Análisis semántico**: si incluyes textos de los informes, puedes aplicar NLP para ver cómo cambia el lenguaje presupuestario.