|  |  |
| --- | --- |
| *Especialización en Finanzas Computacionales* | **Formato[[1]](#footnote-1)**  **Versión 005** |
| *Introducción a la ciencia de datos* |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Título del recurso** | OVA 1 Introducción al Análisis Exploratorio de Datos (EDA) | | | | | | | | | | | | | | |
| **Unidad temática** | UNIDAD 1 ANALISIS EXPLORATORIO DE DATOS | | | | | | | | | | | | | | |
| **Tipo de recurso** *(Marqué X)* | OVA > | X | Recurso interactivo > |  | Presentación > |  | Infografía > |  | Libro > |  | Juego > |  | OVI > |  |  |
| **Experto temático** | **Jorge Ivan Romero** | | | | | | | | | | | | | | |
| **Asesor pedagógico** | **Diana Gamar** | | | | | | | | | | | | | | |

**Por favor tener en cuenta las siguientes indicaciones:**

* Este formato se diligencia para recursos como: OVA, recurso interactivo, presentación, infografía, libro, juego, OVI.
* En el caso de imágenes que contengan, fórmulas matemáticas o similares, así como lenguajes de programación, estas deben ser generadas y suministradas por el autor en una carpeta con el nombre imágenes, el nombre de la imagen debe contener el nombre del recurso a que pertenece y la página o pantalla en donde se debe insertar. Existen herramientas en línea a través de las cuales se pueden generar este clase de imágenes, como por ejemplo:
  + <https://www.periodni.com/es/ecuaciones_matematicas_y_quimicas_en_la_web.html>
  + <https://editor.codecogs.com/>
  + <https://codepen.io/>

|  |
| --- |
| **Instrucciones generales** |
| ***Metáfora o Propuesta Temática:*** *La OVA para la Unidad de "Análisis Exploratorio de Datos" (EDA) representa un viaje científico a través del paisaje de los datos. Como una expedición de exploración, el estudiante viaja a través de diferentes etapas, descubriendo, transformando, modelando y validando datos. Cada etapa es una parada en este viaje, ofreciendo una oportunidad para aprender y aplicar habilidades específicas en ciencia de datos.*  ***Representación Gráfica:*** *La representación gráfica debe ser clara, coherente y atractiva, reflejando el mundo de la ciencia de datos. Debe incluir:*   * *Imágenes y símbolos representativos de cada etapa del EDA.* * *Animaciones educativas que expliquen procesos complejos de manera simple e intuitiva.* * *Tablas y gráficos legibles que muestren procesos y relaciones.*   ***Particularidades y Especificaciones para el Diseñador Gráfico:***  ***Estilo:*** *Utilizar un estilo gráfico moderno y profesional que sea coherente a lo largo de la OVA.*  ***Colores:*** *Utilizar una paleta de colores que refleje la ciencia y tecnología, como azules, grises y verdes.*  ***Tipografía:*** *Elegir fuentes legibles y atractivas que complementen la estética general.*  ***Animaciones:*** *Crear animaciones que sean informativas y atractivas, mostrando el flujo y las conexiones entre las etapas.*  *I****mágenes****: Utilizar imágenes de alta calidad y relevantes para el tema, como gráficos, diagramas y logotipos de herramientas (se agregan aquí imágenes creadas por midjourney y animaciones e imágenes creadas en Python).*  ***OVA 1: Introducción al Análisis Exploratorio de Datos***  ***Capítulo de Referencia: Capítulo 5: Grus***  ***Datos de Ejemplo: Datos de altura y peso de 5000 personas***  ***Librerías: matplotlib, numpy, statistics***  ***Objetivo: Calcular y graficar medidas de tendencia central, dispersión, forma y correlación*** |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **1** |
| **Instrucciones para el diseñador** | |
|  | **Diapositivas interactivas presentarían la importancia del análisis exploratorio de datos, con gráficos y texto explicativo.** | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **2** |
| **Instrucciones para el diseñador** | |
|  |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **3** |
| **Instrucciones para el diseñador** | |
| |  | | --- | | Parte 3: Conceptos básicos | | Diapositiva 1: Introducción a Conceptos Básicos  Texto: "Aquí aprenderemos conceptos estadísticos básicos que son fundamentales para el análisis de datos."  Interacción: Botón "Siguiente" para avanzar.  Diapositiva 2: Media, Mediana, Moda  Texto: "La media es el promedio de un conjunto de datos, la mediana es el valor medio y la moda es el valor más frecuente."  Interacción: Ejercicio en Google Colab para calcular estos valores en un conjunto de datos simple.  Diapositiva 3: Varianza y Desviación Estándar  Texto: "La varianza y la desviación estándar nos dan una idea de la dispersión de los datos."  Interacción: Ejercicio en Google Colab para calcular varianza y desviación estándar.  Diapositiva 4: Coeficiente de Correlación  Texto: "El coeficiente de correlación indica la relación lineal entre dos variables."  Interacción: Ejercicio en Google Colab para calcular la correlación entre dos variables.  Diapositiva 5: Resumen de Conceptos Básicos  Texto: "Estos conceptos básicos son la piedra angular para entender y analizar datos efectivamente."  Interacción: Cuestionario rápido para evaluar la comprensión del tema. | | **Diapositivas interactivas explicarían conceptos como media, mediana y rango intercuartil, con ejemplos y ejercicios en Google Colab.** | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **4** |
| **Instrucciones para el diseñador** | |
| **1** | **Se mostraría cómo cargar un conjunto de datos utilizando read\_csv en Python a través de un notebook en Google Colab.** | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **5** |
| **Instrucciones para el diseñador** | |
|  |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **6** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.2. Histogramas** | | |  |  | | financiero o la variabilidad en indicadores como el PIB o las tasas de interés. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **7** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.3. Diagramas de Caja (Boxplots)** | | |  |  | | **Descripción General:** Los diagramas de caja se utilizan para mostrar la variabilidad y la tendencia central de los datos. Los "bigotes" del gráfico pueden ayudar a identificar valores atípicos o volatilidad en los datos.  **En Finanzas:** Se utilizan para analizar la dispersión en series de datos como precios de acciones o tasasde interés, lo que puede ser útil para identificar volatilidad o riesgo. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **8** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.4. Gráficos de Dispersión (Scatter Plots)** | | |  |  | | **Descripción General**: Los gráficos de dispersión se utilizan para visualizar la relación entre dos variables numéricas. Pueden ayudar a identificar correlaciones o tendencias en los datos.  **En Finanzas:** Se utilizan para analizar la correlación entre diferentes activos financieros o entre un activo y un índice de mercado, lo que puede ser útil para la diversificación de carteras. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **9** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.5. Gráficos de Líneas** | | |  |  | | **Descripción General:** Los gráficos de líneas se utilizan para mostrar tendencias a lo largo del tiempo. Son útiles para identificar patrones y tendencias en series temporales de datos.  **En Finanzas:** Se utilizan para seguir la evolución de métricas como precios de acciones, tasas de interés o indicadores económicos a lo largo del tiempo. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **10** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.6. Mapas de Calor (Heatmaps)** | | |  |  | | **Descripción General:** Los mapas de calor se utilizan para visualizar matrices de correlación o tablas de contingencia. Los colores representan la intensidad de la correlación entre variables.  **En Finanzas:** Se utilizan para visualizar la correlación entre diferentes activos en una cartera o entre diferentes indicadores económicos. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **11** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.7. Diagramas de Violín** | | |  |  | | **Descripción General:** Los diagramas de violín combinan un diagrama de caja con un histograma para mostrar la distribución de los datos.  **En Finanzas:** Se utilizan para visualizar la distribución de métricas como rendimientos o precios de activos, lo que puede ser útil para entender tanto la tendencia central como la variabilidad. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **10** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.6. Mapas de Calor (Heatmaps)** | | |  |  | | **Descripción General:** Los mapas de calor se utilizan para visualizar matrices de correlación o tablas de contingencia. Los colores representan la intensidad de la correlación entre variables.  **En Finanzas:** Se utilizan para visualizar la correlación entre diferentes activos en una cartera o entre diferentes indicadores económicos. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **10** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.6. Mapas de Calor (Heatmaps)** | | |  |  | | **Descripción General:** Los mapas de calor se utilizan para visualizar matrices de correlación o tablas de contingencia. Los colores representan la intensidad de la correlación entre variables.  **En Finanzas:** Se utilizan para visualizar la correlación entre diferentes activos en una cartera o entre diferentes indicadores económicos. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **10** |
| **Instrucciones para el diseñador** | |
| |  |  | | --- | --- | | **1.5.6. Mapas de Calor (Heatmaps)** | | |  |  | | **Descripción General:** Los mapas de calor se utilizan para visualizar matrices de correlación o tablas de contingencia. Los colores representan la intensidad de la correlación entre variables.  **En Finanzas:** Se utilizan para visualizar la correlación entre diferentes activos en una cartera o entre diferentes indicadores económicos. | | |  | |

|  |  |  |
| --- | --- | --- |
| **Contenido** | **Número** | **10** |
| **Instrucciones para el diseñador** | |
| **Conclusión**  El Análisis Exploratorio de Datos (EDA) es una metodología poderosa y esencial en el análisis de datos en diversos campos, desde la investigación científica hasta el análisis de negocios y la toma de decisiones en Machine Learning. A través de técnicas visuales y gráficos efectivos, el EDA permite descubrir patrones, identificar relaciones y revelar información valiosa en los datos.  Usa la siguiente guía y practica tus visualizaciones:  **ACCESO A GOOGLE COLAB**   1. **Abrir Navegador:** Abre tu navegador web favorito y accede a tu cuenta de Google si no lo has hecho aún. 2. **Ir a Colab:** Escribe "Google Colab" en el buscador o entra directamente a colab.research.google.com 3. **Nuevo Notebook:** Haz clic en "Nuevo notebook" para crear un nuevo cuaderno de Jupyter, o puedes abrir uno existente desde Google Drive.   **CONFIGURACIÓN INICIAL**   1. **Renombrar Notebook:** Haz clic en el título para cambiar el nombre del notebook. 2. **Seleccionar Entorno:** En la esquina superior derecha, puedes seleccionar el tipo de entorno (CPU, GPU).   **IMPORTAR DATOS**   1. **Desde Google Drive:**   **Texto  Descripción generada automáticamente**   1. **Desde un archivo csv**     **EDA BÁSICO** |  | |

1. Este formato hace parte de los procesos de diseño y desarrollo de contenidos de la UNIDAD VIRTUAL de la UNIVERSIDAD JORGE TADEO LOZANO, se prohíbe su reproducción y modificación total o parcial” [↑](#footnote-ref-1)