T-Researcher



*Memoria Descriptiva*

**T-RESEARCHER**

**Autores:**

1

2

3

**ÍNDICE**

**04**

**Memoria Descriptiva**

**05**

**Interfaz**

**07**

**Tecnologia**

**08**

**Conclusión**

**MEMORIA DESCRIPTIVA**

T-Researcher es una innovadora herramienta interactiva diseñada específicamente para simplificar y agilizar el proceso de realización de pruebas t de Student, una técnica estadística esencial para la comparación de medias en investigaciones científicas, estudios de mercado, análisis clínicos, y muchas otras áreas donde la comparación entre grupos es fundamental. En el análisis estadístico, la capacidad de determinar si existen diferencias significativas entre dos grupos o entre un grupo y un valor teórico es crucial para sacar conclusiones informadas y tomar decisiones basadas en evidencia. Sin embargo, el proceso de implementar correctamente estas pruebas y, más aún, interpretar sus resultados puede ser complejo, especialmente para aquellos que no tienen un trasfondo profundo en estadística avanzada. T-Researcher se ha desarrollado con el objetivo de romper estas barreras, proporcionando una plataforma accesible y potente que no solo permite realizar análisis estadísticos precisos, sino que también facilita la interpretación de los resultados a través de visualizaciones gráficas intuitivas y la generación de informes detallados. La aplicación está dirigida tanto a estudiantes, que buscan aprender y aplicar conceptos estadísticos en un entorno amigable, como a profesionales e investigadores que necesitan una herramienta confiable para realizar análisis de datos rápidos y eficientes. Desde el primer paso, donde los usuarios pueden cargar sus datos en formatos populares como CSV o Excel, T-Researcher guía al usuario a través de un flujo de trabajo ordenado y lógico. Una vez cargados los datos, la aplicación permite seleccionar las variables numéricas de interés y elegir entre diferentes tipos de pruebas t: comparación entre dos muestras independientes, comparación de una muestra con un valor teórico, o comparación de muestras emparejadas. Cada una de estas opciones está diseñada para cubrir una amplia gama de necesidades analíticas, permitiendo a los usuarios adaptar la prueba a su contexto específico. Además de realizar cálculos estadísticos utilizando la potente librería SciPy, T-Researcher se destaca por su capacidad de generar visualizaciones gráficas interactivas con Seaborn y Matplotlib, lo que facilita una comprensión más profunda de las diferencias entre los grupos comparados. Estas visualizaciones, que incluyen diagramas de caja, histogramas y gráficos de dispersión, permiten explorar los datos de manera visual, ofreciendo una capa adicional de interpretación que complementa los resultados numéricos. Finalmente, una de las características más valiosas de T-Researcher es su capacidad para generar informes en formato PDF, que no solo documentan los resultados de las pruebas t realizadas, sino que también incluyen los gráficos generados, proporcionando un recurso completo y profesional que puede ser fácilmente compartido o archivado para referencias futuras. Con su enfoque en la usabilidad, precisión y flexibilidad, T-Researcher se posiciona como una herramienta indispensable para cualquier persona que necesite realizar análisis de comparación de medias, facilitando la toma de decisiones informadas en una amplia gama de contextos académicos, científicos y profesionales.

**INTERFAZ**

La interfaz de la aplicación está cuidadosamente diseñada en Streamlit, un framework en Python que facilita la creación de aplicaciones web interactivas con una estructura sencilla y eficiente. El objetivo principal de la interfaz es guiar al usuario a través del proceso de análisis estadístico de manera intuitiva, asegurando que cada paso sea claro y fácil de seguir, independientemente del nivel de experiencia del usuario. La aplicación está organizada en varias secciones que desglosan el flujo de trabajo en etapas lógicas, permitiendo un manejo ordenado de los datos y la realización de pruebas estadísticas.

Pantalla de Carga de Archivo: En la primera sección, los usuarios tienen la opción de cargar sus archivos de datos en formato CSV o Excel, lo que les permite trabajar con conjuntos de datos comunes y fácilmente accesibles. Al cargar un archivo, la aplicación lee los datos y presenta los primeros registros en pantalla, permitiendo una rápida verificación visual de que los datos se han cargado correctamente. Esta funcionalidad es fundamental para asegurar que el análisis posterior se base en los datos correctos y que cualquier problema con el formato o contenido del archivo pueda ser identificado desde el principio.

Selección de Variables: Una vez que los datos han sido cargados con éxito, la aplicación permite a los usuarios seleccionar las variables numéricas de interés para el análisis. Esta selección es crucial, ya que determina las variables que serán sometidas a las pruebas t de Student. Dependiendo del tipo de prueba que el usuario desee realizar, pueden seleccionarse una o dos variables. La interfaz está diseñada para ser intuitiva, presentando listas desplegables que filtran automáticamente las variables numéricas, lo que simplifica el proceso de selección y minimiza la posibilidad de errores.

Realización de Pruebas t: La aplicación ofrece tres tipos principales de pruebas t de Student, cada una diseñada para abordar diferentes necesidades analíticas:

Dos muestras independientes: Esta prueba se utiliza cuando se desea comparar las medias de dos grupos distintos, evaluando si existe una diferencia estadísticamente significativa entre ellos.

Una muestra: Esta prueba compara la media de un grupo con un valor teórico predefinido, permitiendo evaluar si la media del grupo difiere significativamente de ese valor.

Muestras emparejadas: Esta prueba es útil cuando se tienen dos conjuntos de datos relacionados (por ejemplo, mediciones antes y después de una intervención) y se desea comparar sus medias. Los resultados de cada prueba, incluyendo el estadístico t, el valor p, y los grados de libertad, se presentan de manera clara y accesible, permitiendo a los usuarios interpretar rápidamente los resultados sin necesidad de conocimientos avanzados en estadística.

Visualización Gráfica: Para mejorar la comprensión de los resultados estadísticos, la aplicación genera una variedad de gráficos interactivos que permiten explorar visualmente los datos. Estos gráficos incluyen diagramas de caja (boxplots), histogramas y gráficos de dispersión. Cada uno de estos gráficos está diseñado para resaltar diferentes aspectos de la distribución y la relación entre las variables seleccionadas, proporcionando una herramienta visual poderosa que complementa los resultados numéricos de las pruebas t. La interactividad de los gráficos permite a los usuarios acercarse, alejarse y explorar diferentes partes del conjunto de datos, lo que facilita una interpretación más profunda y detallada.

Generación de Informes: Finalmente, la aplicación ofrece la opción de generar un informe completo en formato PDF. Este informe incluye tanto los resultados detallados de las pruebas t como los gráficos generados durante el análisis. La posibilidad de descargar estos informes en un formato estándar como PDF es invaluable para documentar los resultados, compartir análisis con colegas o incluirlos en reportes más amplios. La interfaz para la generación de informes es simple y directa, permitiendo a los usuarios generar y descargar su informe con un solo clic, asegurando que la documentación de los resultados sea tan accesible como el proceso de análisis en sí.

**TECNOLOGIA**

La aplicación "T-Researcher" se ha construido utilizando una combinación de tecnologías modernas y robustas en el ecosistema Python, lo que permite un análisis estadístico preciso y una experiencia de usuario interactiva. A continuación, se describe el papel de cada librería utilizada:

Streamlit: es el núcleo de la aplicación, proporcionando el marco para crear interfaces web interactivas y reactivas con Python. Con Streamlit, se puede desarrollar rápidamente aplicaciones que responden a las interacciones del usuario sin necesidad de conocimientos avanzados en desarrollo web. La simplicidad de Streamlit permite que se enfoque en la lógica de análisis de datos, mientras que el framework maneja la presentación visual y la interacción con el usuario.

Pandas: Pandas es una de las librerías más utilizadas en Python para la manipulación y análisis de datos. En esta aplicación, Pandas facilita la carga, limpieza y manejo de los datos provenientes de archivos CSV o Excel. Ofrece estructuras de datos como DataFrames, que son esenciales para realizar operaciones estadísticas y para seleccionar y preparar las variables que serán utilizadas en las pruebas t de Student.

NumPy: proporciona soporte para grandes arreglos y matrices multidimensionales, junto con una colección de funciones matemáticas de alto rendimiento. En la aplicación, NumPy se utiliza para operaciones numéricas subyacentes que son fundamentales en el cálculo de estadísticas y en la manipulación de los datos antes de su análisis.

SciPy: es una librería de Python que se basa en NumPy y proporciona rutinas de alta eficiencia para cálculos matemáticos, científicos e ingenieriles. Dentro de esta aplicación, el módulo stats de SciPy es crucial para realizar las pruebas t de Student. Este módulo contiene un conjunto de funciones estadísticas que permiten calcular valores p, estadísticos t, y otros indicadores relevantes para la comparación de medias en diferentes contextos.

Seaborn: es una librería de visualización de datos basada en Matplotlib, que proporciona una interfaz de alto nivel para dibujar gráficos estadísticos atractivos y útiles. En esta aplicación, Seaborn se utiliza para crear gráficos como diagramas de caja (boxplots), histogramas y gráficos de dispersión, que permiten a los usuarios visualizar la distribución y la relación entre las variables seleccionadas para las pruebas t. La capacidad de Seaborn para integrar fácilmente las variables categóricas y numéricas facilita la creación de visualizaciones complejas de manera sencilla.

Matplotlib: es la librería de visualización fundamental en Python, utilizada para crear gráficos 2D en una amplia variedad de formatos. En esta aplicación, Matplotlib trabaja en conjunto con Seaborn para generar los gráficos que se muestran en la interfaz de usuario. Además, Matplotlib es esencial para la creación de gráficos que se incluyen en los informes PDF, ya que permite un control detallado sobre el estilo y la presentación de las visualizaciones.

BytesIO: es una clase del módulo io en Python, que permite trabajar con archivos binarios en memoria como si fueran archivos normales. En esta aplicación, BytesIO se utiliza para manejar la generación de los informes PDF. Los gráficos generados por Matplotlib se almacenan temporalmente en la memoria mediante BytesIO, lo que permite incorporarlos directamente en el documento PDF sin necesidad de escribir en el disco, mejorando la eficiencia y la velocidad del proceso.

FPDF: es una librería de Python que facilita la generación de documentos PDF. En esta aplicación, FPDF se utiliza para crear informes detallados que incluyen texto y gráficos. La librería permite la personalización de los informes, añadiendo encabezados, pies de página, y múltiples páginas, todo ello controlado de manera programática. Esto es especialmente útil para generar informes estandarizados que pueden ser descargados y compartidos por los usuarios de la aplicación.

tempfile: se utiliza para crear archivos temporales en el sistema de archivos que pueden ser utilizados durante la ejecución de la aplicación. En esta aplicación, tempfile se emplea para almacenar temporalmente los gráficos que luego se insertan en los informes PDF generados por FPDF. Una vez que se ha completado el proceso de generación del informe, estos archivos temporales se eliminan automáticamente, lo que asegura que el almacenamiento en el disco no se vea comprometido.

**CONCLUSIÓN**

La aplicación T-Researcher es una herramienta integral que facilita la realización de análisis estadísticos detallados y precisos de comparación de medias, todo ello dentro de una interfaz accesible y amigable. Aprovechando la potencia de tecnologías como Streamlit, Pandas, SciPy, y Matplotlib, la aplicación permite a los usuarios cargar datos, seleccionar variables y realizar pruebas t con solo unos pocos clics. Además, ofrece una visualización gráfica clara y la posibilidad de generar informes completos en PDF, lo que convierte la aplicación en una solución eficaz para estudiantes, investigadores y profesionales que buscan realizar análisis rigurosos sin necesidad de profundos conocimientos técnicos en programación o estadística. La capacidad de combinar la facilidad de uso con un análisis estadístico robusto posiciona a esta aplicación como una herramienta valiosa en diversas disciplinas, desde la investigación académica hasta el análisis de datos en el ámbito empresarial.

Además de sus funcionalidades actuales, la aplicación tiene un gran potencial de expansión para incluir otras pruebas estadísticas y herramientas de limpieza y transformación de datos, lo que aumentaría aún más su utilidad y alcance. Esta capacidad de adaptación y crecimiento, junto con su enfoque en la accesibilidad y la eficiencia, subraya la relevancia de la aplicación en un entorno cada vez más orientado al análisis de datos. En definitiva, la aplicación de Prueba t de Student no solo simplifica el proceso de análisis estadístico, sino que también democratiza el acceso a herramientas avanzadas, haciendo que el análisis de datos sea más accesible y comprensible para una amplia gama de usuarios.

