



UNIVERSIDAD MAYOR DE SAN SIMÓN
FACULTAD DE CIENCIAS Y TECNOLOGÍA

DIRECCIÓN DE POSGRADO



**DIPLOMADO EN ESTADÍSTICA APLICADA A
LA TOMA DE DECISIONES
PRIMERA VERSIÓN**

**ANÁLISIS DE ENCUESTAS REALIZADAS EN
BOLIVIA REFERIDAS A LA INCLUSIÓN DE TICs
EN EL SISTEMA EDUCATIVO**

**PROYECTO PRESENTADO PARA OBTENER EL GRADO DE LICENCIATURA EN
INGENIERÍA DE SISTEMAS
MODALIDAD DOBLE TITULACIÓN**

POSTULANTE : Joel Abelardo Villca Gallardo
TUTOR : M. Sc. Rolando Amir Mirabal Hinojosa

**Cochabamba – Bolivia
2023**

ANÁLISIS DE ENCUESTAS REALIZADAS EN BOLIVIA REFERIDAS A LA INCLUSIÓN DE TICs EN EL SISTEMA EDUCATIVO

Por

Joel Abelardo Villca Gallardo

El presente documento, Trabajo de Grado es presentado a la Dirección de Posgrado de la Facultad de Ciencias y Tecnología en cumplimiento parcial de los requisitos para la obtención del grado académico de Licenciatura (o solo diplomado) en Ingeniería..., modalidad Doble Titulación, habiendo cursado el Diplomado en Estadística Aplicada a la Toma de Decisiones propuesta por el Centro de Estadística Aplicada (CESA) en su primera versión.

ASESOR/TUTOR

M.Sc. Rolando Amir Mirabal Hinojosa

COMITÉ DE EVALUACIÓN

Lic., M.Sc. (Presidente)

Ing., M.Sc. (Asesor 1)

Ing., M. Sc. (Docente FCyT o del Diplomado)

Lic., M.Sc. (Tribunal externo)



Aclaración

**DIRECCIÓN DE POSGRADO, FACULTAD DE CIENCIAS Y
TECNOLOGÍA**

Carretera de la Universidad, 28040 Madrid, España

Este documento describe el trabajo realizado como parte del programa de estudios de Diplomado en Estadística Aplicada a la Toma de Decisiones en el Centro de Estadística Aplicada CESA y la Dirección de Posgrado de la Facultad de Ciencias y Tecnología. Todos los puntos de vista y opiniones expresadas en el mismo son responsabilidad exclusiva del autor y no representan necesariamente las de la institución.

Resumen

Texto Garamond tamaño 11.5, interlineado múltiple 1.2 con espaciado anterior y posterior de 6 puntos.

Máximo una plana. Resume la idea completa del trabajo

Palabras clave

Insertar palabras clave en un máximo de 6 recomendable, recogen los temas principales, puede incluir palabras combinadas. Las palabras deben conducir a los temas principales de la investigación. Las palabras deben de ser entendibles para cualquier persona en el mundo (no usar palabras específicamente usadas en alguna región o zona)

No usar palabras muy genéricas, ejemplo: Energía NO, mejor Energía Hidroeléctrica.

*Dedicatoria en Texto Garamond en cursiva
tamaño 11.5, interlineado múltiple 1.2 con
espaciado anterior y posterior de 6 puntos.
Se sugiere un máximo de 2 párrafos breves.*

Agradecimientos

Agradecimientos con Texto Garamond en cursiva tamaño 11.5, interlineado múltiple 1.2 con espaciado anterior y posterior de 6 puntos.

Se sugiere un máximo de 1/2 página.

Ejemplo:

Al(la) Ing. MSc. por su valiosa y desinteresada colaboración.

A todo el equipo de profesionales del..... por la transmisión de sus conocimientos y por brindarme la información generada en utilizada en el presente proyecto.

A los técnicos de por los datos brindados.

Tabla de contenidos

1. Introducción	1
1.1. Antecedentes	1
1.2. Justificación	2
1.3. Planteamiento del problema	3
1.4. Objetivo general	3
1.4.1. Objetivos específicos	3
2. Marco teórico	4
2.1. Ingeniería de Software	4
2.2. TICs	5
2.2.1. Tecnología	5
2.2.2. Información	5
2.2.3. Comunicación	5
2.3. Tecnologías	6
2.3.1. Lenguaje de programación	6
2.3.2. Python	6
2.3.3. Pandas	6
2.3.1. Jupyter	7
2.4. Sistema educativo	7
3. Marco metodológico	9
3.1. Área de estudio	9
3.2. Flujograma metodológico	10
3.2.1. Planteamiento del problema	10
3.2.2. Recopilación de datos	10
3.2.3. EDA (Exploratory Data Analysis)	10
3.2.4. Limpieza y ordenamiento de datos	10
3.2.5. Procesamiento de datos	11
3.2.6. Determinar indicadores y pesos	11
3.2.7. Dataset enriquecido	11
3.2.8. Dashboards descriptivos	11
3.3. Fuentes de información	11
3.3.1. Fuente secundaria	11
3.3.2. Fuente primaria	12
3.4. Factores.....	12
3.5. Clasificación de(si corresponde)	12

3.6. Subtitulo	12
3.7. Subtítulo, ejemplo: Recopilación de datos de.....	13
3.8. Subtítulo.....	14
3.9. Subtítulo/ Diseño/ Modelacion	14
4. Resultados y Discusión	16
4.1. Resultados de	16
4.2. Discusión de resultados	16
5. Conclusiones	18
6. Recomendaciones	19
Referencias bibliográficas	20
Anexos	22
Anexo 1. Resultados de encuesta sobre	22

Lista de figuras

Figura 3-1: Poner un mapa que ayude a ubicar el área de estudio (incluir la fuente, año)	3
Figura 3-2: Flujograma metodológico	3
Figura/tabla 3-3: Factores para	3
Figura 3-4: Parámetros de	4
Figura 3-6: Planilla relevamiento de datos de conocimiento local en la encuesta	5
Figura 3-7: Puntos de relevamiento de datos en la encuesta estructurada	5
Figura 3-8: Fotografías de puntos de relevamiento de datos en la encuesta estructurada	6
Figura 3-10: Planilla relevamiento de datos de conocimiento experto en la encuesta	6
Figura 4-1: Título de la figura	8
Figura 4-10: figura de otra investigación, de otro autor (incluir fuente)	8
Figura 4-11: Figura comparativa con nuestros resultados	8

Lista de tablas

Tabla 3-1: Rango de valores para (Gracia, y otros, 2010)	4
Tabla 3-2: Factor de	4

1. Introducción

Análisis de datos públicos respecto al uso de TICs con el fin de ver la aceptación o rechazo de la inclusión de esta tecnología en el sector educativo y dónde se aprecian o se rechazan más este tipo de innovaciones. Este estudio va guiado por la siguiente premisa: probar que los sectores rurales rechazan en su mayoría estas nuevas tecnologías porque o bien no se tiene el conocimiento adecuado de las mismas o porque genera en muchos casos costes extras a la población donde se intenta aplicarlas.

Este estudio pretende llegar a su objetivo mediante una serie de técnicas y estudios pero entre ellos prima el análisis de sentimientos de las encuestas de los datos recabados. De esta manera poder lograr llegar a una conclusión lo más confiable posible para poder rechazar o aceptar la hipótesis que se quiere demostrar.

Estos datos son obtenidos de manera libre de *Datos Abiertos Bolivia*. (s. f.). De donde se obtiene la información que se tratará a lo largo del desarrollo del estudio, tomando en cuenta los detalles de las encuestas realizadas se pretende poder hacer un proceso de ETL de los datos y poder verificar en los distintos lugares de donde se emite la opinión respecto al tema y así verificar si está a favor o en contra y por qué.

1.1. Antecedentes

A escala mundial, en la última década el uso de Internet ha crecido de manera acelerada, tanto en el número de usuarios como en la velocidad de acceso; al mismo ritmo, el precio del servicio ha disminuido. Puede decirse con toda propiedad que el uso y acceso a las TIC y a Internet son una necesidad primaria e importante para cualquier país, y deberían ser considerados fundamentales en sus planes de desarrollo. Adicionalmente, en Bolivia, la Constitución Política del Estado establece en su artículo 20 que todas las personas tienen derecho al acceso universal y equitativo a las telecomunicaciones, y que es responsabilidad del estado la provisión de este servicio (Mayra Palacios, 2019)

Bolivia al ser un país en vías de desarrollo está guiado por la necesidad de ponerse a la par tanto a nivel industrial, comercial y en este caso, tecnológico. Pero esta innovación no muchas veces viene acompañado de un buen plan de estudio sobre el área y menos de aplicación; es por eso que en muchos casos este entusiasmo por querer generar progreso e innovar se ve truncada por la falta de personal capacitado tanto para la implementación de esta tecnología como los que harán uso de ella.

Según las últimas cifras del regulador ATT, Bolivia tenía cerca de 11,6 millones de líneas móviles en servicio a fines de septiembre de 2019, un poco más que los 11,4 millones del año anterior. Los clientes móviles de prepago representaron el 89,1% del total de las líneas. Las conexiones al Servicio de Distribución de Señales (la televisión por cable y satelital) continúan creciendo mientras que los números para el servicio de telefonía local y de acceso público siguen en picada.

Los ingresos en el sector de telecomunicaciones alcanzaron los 6.298mn de bolivianos (US\$912mn) en el 1T19, la última cifra disponible, por debajo de los 6.471mn de bolivianos del mismo periodo de 2018. (La Voz, 2020).

Se tiene encuestas de datos abiertos de Bolivia donde se recopila opiniones sobre las TICs en el área educativa. Lo que nos brinda la posibilidad de tener opiniones reales tanto de profesores como de estudiantes sobre este conjunto de tecnologías que poco a poco va tomando más presencia en varios sectores. Este estudio se pretende focalizar únicamente en el tratado a los datos recopilados de: *Datos Abiertos Bolivia*. (s. f.). Pero esto no limitará la comparación con otros para lograr insights que nos sean de valor para este estudio.

1.2. Justificación

La inclusión de las Tecnologías de la Información y la Comunicación (TICs) en Bolivia es esencial para mejorar la competitividad económica del país, así como para mejorar la calidad de vida de la población. Las TICs pueden ayudar a mejorar la educación, el acceso a la información, la salud, el empleo, la seguridad y la eficiencia en los servicios públicos. Además, la inclusión de TICs puede ayudar a reducir las desigualdades sociales y económicas, al permitir el acceso a oportunidades y recursos que de otra manera estarían fuera del alcance de muchas personas. (Barreto & Diazgranados, 2017)

Incrementar el nivel de confianza para la toma de decisiones respecto a proyectos relacionados a las TICs y el sector educativo es el foco principal de este estudio que toma como referencia opiniones para canalizar esfuerzos de aplicación y de estudio en distintos lugares y no caer en la generalización de personas y lugares a la hora de la toma de decisiones, el poder tener una mayor seguridad de dónde serán más valiosos este tipo de proyectos y dónde tomar acciones previas a la aplicación o incursión de estas tecnologías.

En resumen, la inclusión de TICs en Bolivia es un paso importante para el desarrollo sostenible, inversiones inteligentes para un incremento en la productividad económica y la mejora de la calidad de vida de la población. Es por eso que se debe tomar decisiones inteligentes a la hora de presentarlas e incursionar en las mismas.

1.3. Planteamiento del problema

La falta de acceso a tecnologías y recursos educativos en línea en las escuelas de áreas rurales y de bajos recursos. Muchas escuelas en estas áreas no cuentan con equipos informáticos ni acceso a internet de alta velocidad, lo que dificulta el uso de herramientas educativas en línea y la implementación de metodologías educativas innovadoras. Además, la falta de capacitación para el uso de las TICs en la educación también puede ser un problema, ya que los maestros y estudiantes pueden no estar preparados para aprovechar al máximo las herramientas y recursos disponibles. Esto puede conducir a una brecha digital en la educación y a un desempeño académico inferior en las escuelas rurales y de bajos recursos.

De igual forma se tiene proyectos relacionados a TICs que ofrecen grandes beneficios al sector educativo pero están vistos de una manera muy generalizada o viendo casos y escenarios idóneos donde la mayoría de ellos no cuadran en muchos lugares/instituciones de nuestro departamento ya que las condiciones con las que fueron pensados no se cubren como deberían o no son bien aceptados por la población.

¿De qué depende el rechazo o aceptación de proyectos relacionados con TICs en el sistema educativo y cómo podemos identificar dónde aplicarlos?

1.4. Objetivo general

Analizar encuestas obtenidas generando indicadores que reflejan la opinión de las personas y ayuden a la toma de decisiones para la aplicación de proyectos relacionados con TICs en el sector educativo.

1.4.1. Objetivos específicos

- Evaluar las encuestas recopiladas obteniendo información de valor para nuestro caso de estudio.
- Ordenar y limpiar datos recopilados de las entrevistas para su posterior análisis y procesamiento.
- Procesar datos en búsqueda de insights.
- Determinar indicadores y sus pesos que nos ayuden en la modelación de un dataset.
- Generar un dataset enriquecido.
- Visualizar insights encontrados en los procesos previos mediante dashboards.

2. Marco teórico

2.1. Ingeniería de Software

Se considera a la Ingeniería del Software como la disciplina que se ocupa de las actividades relacionadas con los sistemas informáticos o sistemas de información en los que el software desempeña un papel relevante. Una definición alternativa podría ser: “La aplicación de métodos y conocimiento científico para crear soluciones prácticas y rentables para el diseño, construcción, operación y mantenimiento del software y los productos asociados, al servicio de las personas.” (Mary Shaw y David Garlan, 1996).

“En el contexto de la Ingeniería del Software, un proceso no es una prescripción rígida de cómo se construye un sistema software. Debe dar una aproximación adaptable que permita al equipo de ingenieros hacer su trabajo. Como hay muchos tipos diferentes de software, no existe un proceso software universal. No obstante, cualquier proceso debe incluir, de alguna manera, las cuatro actividades principales de la Ingeniería del Software.” (Pressman Roger, 2016, p. 12).

- Especificación del software: Donde se definen la funcionalidad del software y sus restricciones.
- Desarrollo del software: Se produce el software que cumple con las especificaciones.
- Validación del software. Se debe asegurar que el software cumple con lo que el cliente espera.
- Evolución del software. El software debe evolucionar para cumplir con las necesidades cambiantes del cliente.

“Las herramientas de desarrollo del software (llamadas en ocasiones herramientas de Ingeniería de Software Asistido por Computadora o CASE, por las siglas de Computer-Aided Software Engineering) son programas usados para apoyar las actividades del proceso de la ingeniería de software. En consecuencia, estas herramientas incluyen editores de diseño, diccionarios de datos, compiladores, depuradores, herramientas de construcción de sistema, etcétera. ” (Somerville Ian, 2011, p. 37)

Las herramientas de software ofrecen apoyo de proceso al automatizar algunas actividades del proceso y brindar información sobre el software que se desarrolla. Los ejemplos de actividades susceptibles de automatizar son:

- Desarrollo de modelos de sistemas gráficos, como parte de la especificación de requerimientos o del diseño del software.
- Generación de código a partir de dichos modelos de sistemas gráficos.
- Producción de interfaces de usuario a partir de una descripción de interfaz gráfica, creada por el usuario de manera interactiva.
- Depuración del programa mediante el suministro de información sobre un programa que se ejecuta.
- Traducción automatizada de programas escritos, usando una versión anterior de un lenguaje de programación para tener una versión más reciente.

Las herramientas pueden combinarse en un marco llamado ambiente de desarrollo interactivo o IDE (por las siglas de Interactive Development Environment). Esto ofrece un conjunto común de facilidades, que usan las herramientas para comunicarse y operar con mayor destreza en una forma integrada.

2.2. TICs

Las Tecnologías de la Información y la Comunicación (TICs) son un conjunto de herramientas y recursos que se utilizan para tratar, almacenar y transmitir información. Estas herramientas incluyen dispositivos y equipos electrónicos, software, redes de comunicación y servicios en línea. Las TICs se utilizan en una amplia variedad de ámbitos, incluyendo la educación, el negocio, la comunicación y el entretenimiento. (*Unión Internacional De Telecomunicaciones (UIT)*, n.d.)

2.2.1. Tecnología

Refiere al conjunto de conocimientos, herramientas, técnicas y procesos utilizados para crear, producir y mejorar bienes y servicios. Puede incluir desde herramientas y máquinas simples hasta sistemas complejos y avanzados de información y comunicaciones. La tecnología está en constante evolución y cambio, y juega un papel importante en el desarrollo económico y en la mejora de la calidad de vida de las personas.

2.2.2. Información

Es un conjunto de datos que tienen significado y son útiles para un propósito específico. Puede ser cualquier cosa, desde una simple fecha hasta una descripción detallada de un evento o proceso. La información es esencial para la toma de decisiones y el aprendizaje.

2.2.3. Comunicación

La comunicación en la tecnología permite a las personas conectarse entre sí en tiempo real a través de diferentes canales, como el correo electrónico, el chat, las redes sociales y la videoconferencia.

Además, las TICs han transformado la forma en que las organizaciones se comunican entre sí y con sus clientes, permitiendo una mayor eficiencia y efectividad en la comunicación.

La comunicación en la tecnología también ha cambiado la forma en que las personas acceden y comparten información, ya sea a través de la búsqueda en internet, la lectura de noticias en línea o el acceso a bibliotecas digitales. A medida que las TICs continúan evolucionando, la comunicación se vuelve cada vez más rápida, fácil y accesible para un número cada vez mayor de personas.

2.3. Tecnologías

2.3.1. Lenguaje de programación

Un lenguaje de programación es un conjunto de reglas y convenciones utilizadas para escribir programas informáticos. Estos programas pueden ser utilizados para controlar computadoras y automatizar tareas. Los lenguajes de programación se dividen en dos categorías: los lenguajes de bajo nivel y los lenguajes de alto nivel. Los lenguajes de bajo nivel se corresponden con las instrucciones que entiende la computadora, mientras que los lenguajes de alto nivel se corresponden con un lenguaje más cercano al lenguaje humano.

Los lenguajes de programación son el medio fundamental para la comunicación entre los seres humanos y las computadoras, y son una herramienta esencial para la programación de software. (Ghezzi, Jazayeri, & Mandrioli, 2017).

2.3.2. Python

Python es un lenguaje de programación de alto nivel, interpretado, dinámico y orientado a objetos. Fue creado por Guido van Rossum en 1991. Es ampliamente utilizado en la programación científica, la inteligencia artificial, el análisis de datos, el desarrollo web y la automatización de tareas. Se caracteriza por su sintaxis simple y legible, lo que lo hace fácil de aprender y usar para principiantes y programadores experimentados.

Python es un lenguaje de programación interpretado cuyo objetivo principal es proporcionar una forma fácil de trabajar con sistemas operativos, redes, sistemas de bases de datos y aplicaciones web. (Van Rossum & Drake, Jr., 1995).

Con Python y sus librerías incorporadas se pretende realizar desde el análisis exploratorio de datos hasta el modelamiento del dataset resultante.

2.3.3. Pandas

Pandas es una librería de Python que proporciona estructuras de datos y herramientas para el análisis y manipulación de datos. La librería está construida sobre las estructuras de datos de NumPy y se utiliza ampliamente en el análisis de datos y la ciencia de datos. Los dos principales tipos de estructuras de datos en Pandas son los DataFrame y los Series. Los DataFrame son similares a las tablas en una base de datos relacional o las hojas de cálculo, mientras que las Series son similares a las columnas en un DataFrame.

Pandas ofrece una gran cantidad de funciones y métodos para manipular y analizar datos, como la selección de datos, el filtrado, la agrupación, la unión, el agregado y la limpieza de datos. Además, también proporciona una interfaz para leer y escribir datos en diferentes formatos, como CSV, Excel, SQL y JSON.

La librería Pandas es una librería de Python que proporciona estructuras de datos flexibles y herramientas para el análisis y manipulación de datos, con el objetivo de facilitar el trabajo con datos en Python. (McKinney, 2011).

2.3.4. Jupyter

Jupyter es una herramienta de código abierto que permite crear y compartir documentos que contienen código, ecuaciones, visualizaciones y texto. Es ampliamente utilizada en la comunidad científica y de datos para realizar investigaciones, desarrollar software y crear tutoriales interactivos. Jupyter es compatible con varios lenguajes de programación, incluyendo Python, R, Julia y Scala.

Jupyter ofrece una interfaz de usuario amigable, conocida como Jupyter Notebook, que permite escribir, ejecutar y documentar el código de manera eficiente. Los Notebooks de Jupyter son archivos que contienen tanto código como texto explicativo, se pueden compartir fácilmente y son ideales para la creación de tutoriales y documentos de investigación. (Kluyver et al., 2016).

2.3.5. Tableau Prep

Tableau Prep es una herramienta de preparación de datos creada por Tableau Software que permite a los usuarios limpiar y estructurar sus datos antes de analizarlos con Tableau Desktop. Tableau Prep ofrece una interfaz de usuario visual para transformar, limpiar y combinar datos de diferentes fuentes y formatos.

Algunas de las características clave de Tableau Prep incluyen la capacidad de:

- Explorar y visualizar datos a medida que se transforman.
- Detectar y corregir automáticamente errores comunes de limpieza de datos.
- Unir, combinar y dividir conjuntos de datos.
- Aplicar transformaciones de datos a nivel de fila y columna.

En definitiva, Tableau Prep ayuda a simplificar el proceso de preparación de datos, lo que permite a los analistas de datos ahorrar tiempo y aumentar la eficiencia de su trabajo Tableau. (2021).

2.4. Sistema educativo

El sistema educativo en Bolivia ha enfrentado varios desafíos en los últimos años, incluyendo una baja tasa de alfabetización, una alta tasa de deserción escolar y una brecha de acceso a la educación entre zonas urbanas y rurales. Sin embargo, el gobierno boliviano ha adoptado medidas para mejorar el sistema educativo, incluyendo la inclusión de las Tecnologías de la Información y la Comunicación (TICs) en la educación.

En los últimos años, se ha implementado un plan nacional de informatización que busca mejorar el acceso y uso de las TICs en la educación. Este plan incluye la creación de espacios de innovación tecnológica en escuelas y la formación de docentes en el uso de las TICs. El Ministerio de Educación también ha desarrollado un sistema de educación en línea, conocido como "EduBolivia", que permite a los estudiantes acceder a recursos educativos en línea y participar en clases virtuales.

Con respecto al uso de las TIC y plataformas virtuales, este fenómeno no es desconocido para los jóvenes de la población, si bien las plataformas de formación virtual son consideradas por el profesorado como herramientas tecnológicas con fuertes potencialidades didácticas (Cabero y Ruiz-Palmero, 2018), las mismas han estado siendo estudiadas, de manera somera, en especial los jóvenes conocen que la educación en muchas partes del mundo ahora es virtual y de la cual a futuro ellos pueden ser parte.

Sin embargo, todavía hay desafíos en cuanto a la inclusión de las TICs en la educación en Bolivia. Uno de los principales desafíos es el acceso a las TICs, especialmente en zonas rurales y de bajos recursos, donde el acceso a internet y el equipamiento tecnológico sigue siendo limitado. Además, existe una brecha de habilidades digitales entre docentes y estudiantes, lo que dificulta la implementación efectiva de las TICs en el aula.

3. Marco metodológico

3.1. Área de estudio

Este estudio se centra de acuerdo a las encuestas obtenidas en el país de Bolivia tomando en cuenta todos sus departamentos.



Figura 3-1: Mapa de Bolivia. (2022). Google Earth.

3.2. Flujograma metodológico

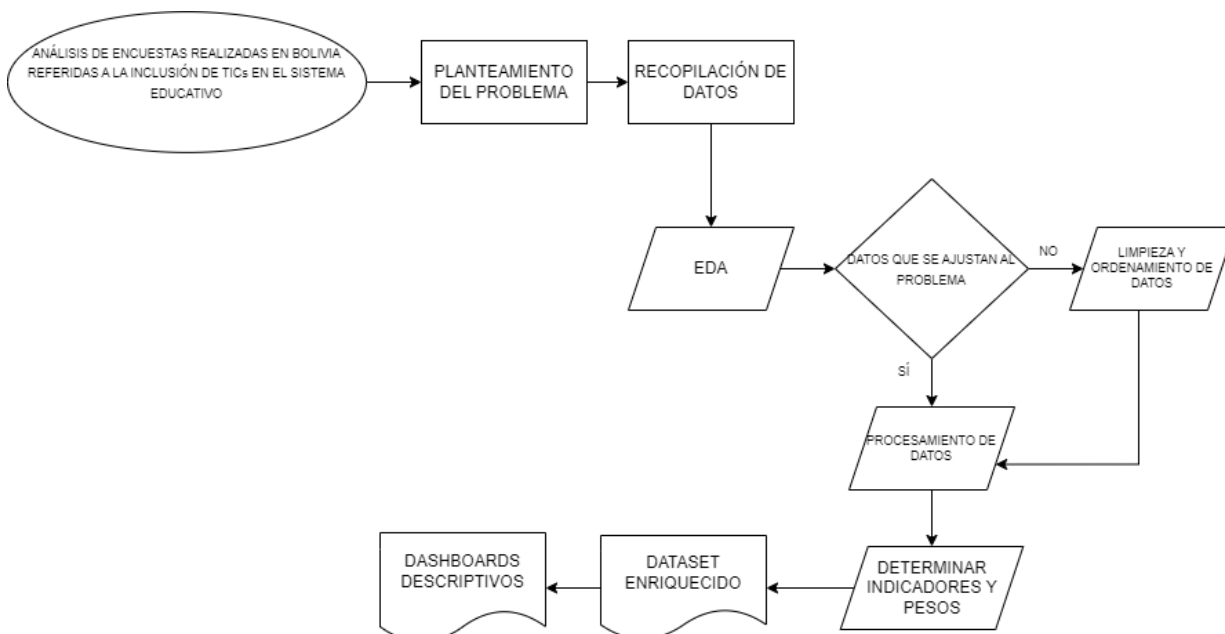


Figura 1. Flujograma metodológico (Elaboración Propia)

3.2.1. Planteamiento del problema

La problemática principal es la incertidumbre que existe sobre la aceptación o rechazo de las TICs en ciertos lugares de nuestro país y el por qué del mismo.

Se tiene como hipótesis principales las siguientes:

- El rechazo de las TICs es mayor en la población rural.
- El rechazo de las TICs se debe a la falta de información sobre las mismas.

3.2.2. Recopilación de datos

Recopilación de datos que a priori sean útiles para la propuesta de solución y relacionados con el problema principal. Estos datos serán analizados y recolectados de la fuente de Datos abiertos de Bolivia. Enfocándonos principalmente en la recopilación de encuestas sobre TICs que tengan relación con el sector educativo.

3.2.3. EDA (Exploratory Data Analysis)

EDA o Análisis Exploratorio de Datos es una técnica para comprender y describir los datos mediante la limpieza, transformación y visualización; es un proceso esencial en el ciclo de vida del análisis de datos para obtener un conocimiento profundo de los datos y prepararlos para el modelado y la toma de decisiones.

3.2.4. Limpieza y ordenamiento de datos

Si los datos lo requieren se procederá a realizar una limpieza para que el posterior tratado de los mismos sea mucho más llevadero. Así mismo, tener definido el diccionario de datos será muy vital para comprender los mismos y poder sacar el mayor provecho de ellos. Tanto la limpieza como el ordenamiento de datos se realizará de manera individual por cada dataset y diccionario de datos de los cuales hagamos uso ya que este proceso requiere un trato individual y muy particular dependiendo la calidad de datos con los que nos encontremos.

Para este proceso se pretende utilizar las herramientas de: Excel y Tableau Prep.

3.2.5. Procesamiento de datos

Se tiene la normalización y adaptación de los datos para que todos aquellos datasets que se lograron recopilar tengan cierta relación y si con los procesos previos no se descubrieron relaciones o similitudes en los mismos, en este punto se procesarán para relacionar las fuentes de datos de manera que formemos relaciones entre los mismos y así el esfuerzo para realizar las posteriores tareas no sean precisamente muy distintos para el tratado de un dataset u otro.

3.2.6. Determinar indicadores y pesos

Proceso en donde distinguimos las cualidades más importantes de los datos recopilados y procedemos a la selección de los más representativos para posteriormente, según el análisis de todos los datos recopilados, se pueda asignar un peso acorde a la relevancia descubierta en los análisis previos. Estos indicadores pertenecen principalmente a preguntas claves que descubramos en los análisis previos que no sirvan de manera directa a dirigir nuestros datos para el descubrimiento de aceptación o rechazo de las hipótesis principales que estamos siguiendo en este proceso.

3.2.7. Dataset enriquecido

De acuerdo a los procesos previos se generará un solo dataset compuesto de los indicadores elegidos que logre representar en esencia los análisis de las encuestas y las opiniones que más destacan según los indicadores que logramos seleccionar en el proceso previo.

3.2.8. Dashboards descriptivos

De acuerdo al dataset generado se procederá a contar la historia de los datos con ayuda de la elaboración de dashboards que den la información necesaria para ver de una mejor manera la inclinación de las personas sobre las TICs.

3.3. Fuentes de información

3.3.1. Fuente secundaria

Los datos obtenidos se encuentran distribuidos de manera libre en la página “Datos abiertos de Bolivia” donde encontramos los datos base utilizados para la realización de este proyecto en un formato .csv. Estos mismos se describen seguidamente:

- Encuesta Nacional de Opinión sobre TIC: El objetivo de la Encuesta Nacional de Opinión sobre TIC fue obtener información representativa a nivel nacional, urbano/ rural y departamental, sobre el acceso y usos de Tecnologías de Información y Comunicación (TIC), servicios de Gobierno Electrónico y equipamiento de la población internauta de 14 o más años de edad.
- Encuesta Final-Profesores de Inclusión Digital: Estudio con el propósito de integrar tecnologías de software libre a los métodos de enseñanza-aprendizaje en Unidades Educativas Públicas. El objetivo de este estudio es obtener información inicial de información relevante sobre Tecnologías de Información y Comunicación TIC de las las Unidades Educativas Públicas.

3.3.2. Fuente primaria

Para el caso del dataset: “Encuesta Nacional de Opinión sobre TIC” se procesó el diccionario de datos que obtuvimos en primera instancia dado que el mismo se encontraba en formato PDF que dificultaba bastante la interacción con el mismo y se obtuvo uno nuevo en formato JSON donde formamos un esquema que nos sería útil para tareas posteriores. A continuación las referencias al diccionario de datos obtenido y al notebook con el código que se usó para generar el mismo:

- Diccionario de datos de tipo JSON.
- Notebook con el código fuente utilizado.

3.4. Factores.....

Los factores para la generación de se resumen en la figura/tabla 3-3.

Figura/tabla 3-3: Factores para

La obtención de los factores es resultado de la simulación fue realizada con, aplicando el software....version.....; debido a

De la modelación se generan para

Figura 3-4: Parámetros de

3.5. Clasificación de(si corresponde)

Para la clasificación

.....según la tabla 3-1.

Tabla 3-1: Rango de valores para (Gracia, y otros, 2010)

3.6. Subtitulo

La detallados en la tabla 3-2.

Tabla 3-2: Factor de

Sigla	Descripción	Factor de
PR5	Período de retorno de 5 años	0.2

PONER FIGURA, FOTO O GRAFICA.

Figura 3 5: Mapas de

Fuente: Elaboración Propia

3.7. Subtítulo, ejemplo: Recopilación de datos de.....

Se realizó visitas a la zona de estudio, con la finalidad de obtención de información de la población a través de encuestas o entrevistas estructuradas

La hoja de encuesta (figura 3-6) recaba datos de

La identificación de la zona de estudio, para la recolección de información o entrevista, se la hizo mediante

..... recolección de datos con lo cual resultaron los 46 puntos de la recolección los que se observan en las figuras 3-7 y 3-8.

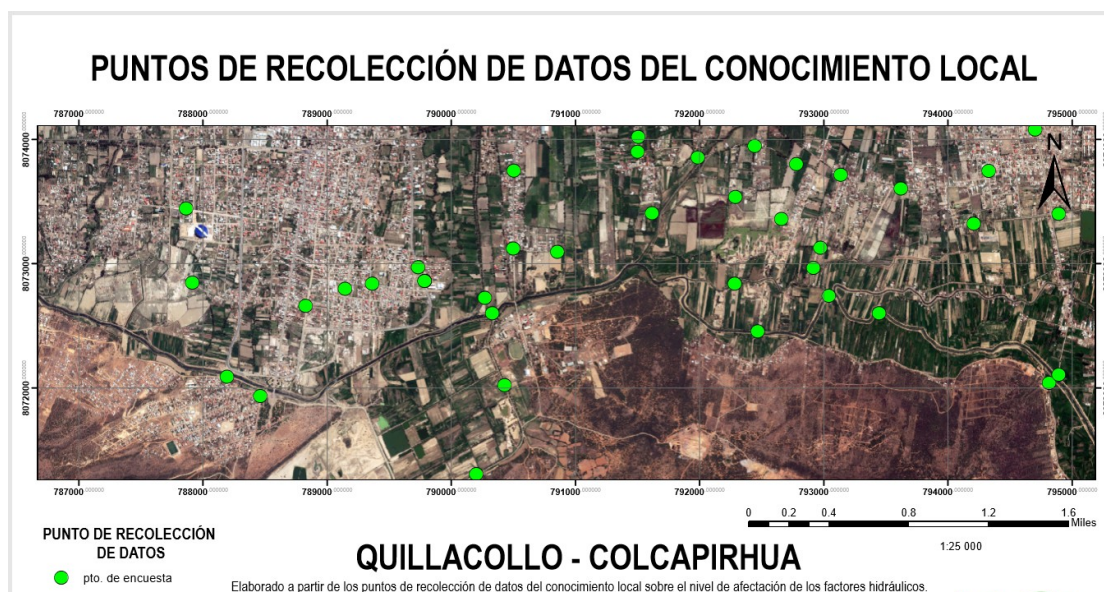


Figura 3-7: Puntos de relevamiento de datos en la encuesta estructurada

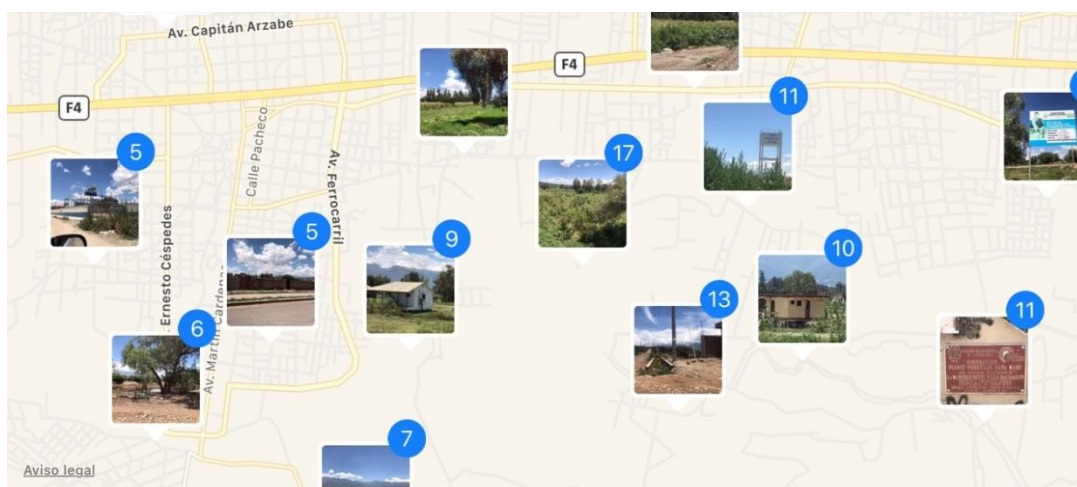


Figura 3-8: Fotografías de puntos de relevamiento de datos en la encuesta estructurada

El archivo digital que contiene las encuestas llenadas sobre....., puede ser consultado en el CD adjunto al presente trabajo (Anexo 2).

3.8. Subtítulo.....

La

3.9. Subtítulo/ Diseño/ Modelacion

La

Incluir tipos de procedimientos, diseño experimental, formulas, muestreo, técnicas de selección, variables y otros

Que tipos de equipos se utilizaron,

>No poner:

Encuestas completas

Lista de materiales

Tablas o figuras completas con datos obtenidos.

4. Resultados y Discusión

Evaluar las encuestas recopiladas obteniendo información de valor para nuestro caso de estudio.

Encuesta Final-Profesores de Inclusión Digital

a0	a1	a2	a3	a4	a40	b1.b11.
				Santa Cruz de la		
Sí	Masculino	32	Santa Cruz	Sierra	NACIONAL COTOCA	Sí
Sí	Masculino	38	La Paz	El Alto	PRIMAVERA B	Sí
					DOMINGO FAUSTINO	
Sí	Femenino	54	La Paz	La Paz	SARMIENTO	Sí
					ABRAHAM REYES FE Y	
Sí	Femenino	54	La Paz	La Paz	ALEGRIA	No
					ABRAHAM REYES FE Y	
Sí	Masculino	34	La Paz	La Paz	ALEGRIA	No

Table 1: Muestra del dataset: Encuesta Final-Profesores de Inclusión Digital

Este dataset nos brinda una serie de datos derivado de preguntas generalizadas sobre las TICs en archivo .csv y también cuenta con un diccionario de datos en un formato .ods en donde se encuentran las preguntas adjunto con la codificación de cada una de las que se realizaron en esta encuesta. Del tipo:

Código	Descripción
a0	¿Fue parte del proyecto Inclusión Digital?
a1	¿Cuál es su sexo?
a2	¿Cuántos años tiene?
a3	¿En qué departamento se encuentra la unidad educativa de la cual es profesor?
a4	¿En qué ciudad se encuentra la unidad educativa de la cual es profesor?
a40	¿Cuál es la Unidad Educativa donde imparte clases?
b1[b11]	¿Cuál es el nivel de instrucción adquirido? [Profesor Normalista]
b1[b12]	¿Cuál es el nivel de instrucción adquirido? [Técnico]
b1[b13]	¿Cuál es el nivel de instrucción adquirido? [Licenciatura]
b1[b14]	¿Cuál es el nivel de instrucción adquirido? [Especialidad]
b1[b15]	¿Cuál es el nivel de instrucción adquirido? [Maestría]
b1[b16]	¿Cuál es el nivel de instrucción adquirido? [Doctorado]
b1[other]	¿Cuál es el nivel de instrucción adquirido? [Otro]

Table 2: Muestra del diccionario de datos del dataset: Encuesta Final-Profesores de Inclusión Digital

Se tomará el enfoque de accesibilidad, uso y opinión que tienen los maestros sobre las TICs. Verificar si cuentan con acceso a tecnología como Smartphones, computadoras; si saben cómo usar las mismas y sacar provecho mediante el uso de internet y lo más importante, si es que les ayuda a comunicar esos nuevos descubrimientos o aprendizajes a sus estudiante. Esto mediante el análisis de preguntas pre-seleccionadas que apuntan a darnos más información sobre los casos mencionados.

Encuesta Nacional de Opinión sobre TIC

Este dataset nos ofrece algo más generalizado, dirigido a internautas y no internautas. Estos datos se encuentran en un archivo .csv el mismo que cuenta con los registros para ambas encuestas, es decir; para los internautas y no internautas, también cuenta con un diccionario de datos en formato .pdf con todas las preguntas que se realizaron para la encuesta, la codificación de la pregunta, el tipo de respuesta obtenido y el tipo de respuesta que representa.

P1		
Valor		
Atributos	Posición	2
estándar	Etiqueta	1. ¿Me podría decir su edad por favor?
	Tipo	N Numérico
	Formato	F2
	Medición	Escala
P2		
Valor		
Atributos	Posición	3
estándar	Etiqueta	2. ¿Usted ha navegado por Internet o se ha conectado a alguna red social a través de Internet?
	Tipo	N Numérico
	Formato	F1
	Medición	N Nominal
Valores válidos	1	Si
	2	No
P3		
Valor		
Atributos	Posición	4
estándar	Etiqueta	3. ¿Cuándo fue la última vez que navegó en Internet o estuvo conectado a las redes sociales?
	Tipo	N Numérico
	Formato	F1
	Medición	N Nominal
Valores válidos	1	Hace 7 días o menos
	2	Entre 8 a 15 días
	3	Entre 16 a 30 días
	4	Hace más de 30 días

Figure 1: Muestra del diccionario de datos de "Encuesta Nacional de Opinión sobre TIC "

De forma similar, las dos encuestas recopilan información general sobre el acceso, uso y medio en el que se encuentran las personas que hacen uso de estas tecnologías. Al igual que el tratamiento del dataset anterior se procede a pre-seleccionar las preguntas que nos sean de más utilidad para obtener los datos que necesitamos para este análisis.

Ordenar y limpiar datos recopilados de las entrevistas para su posterior análisis y procesamiento.

Para el dataset “Encuesta Final-Profesores de Inclusión Digital” primeramente se procedió a pre-seleccionar las preguntas que nos puedan brindar más detalle según el objetivo de nuestro análisis por lo cual se seleccionaron las siguientes:

Cod. Pregunta	Pregunta
a0	¿Fue parte del proyecto Inclusión Digital?
a2	¿Cuántos años tiene?
a3	¿En qué departamento se encuentra la unidad educativa de la cual es profesor?
a4	¿En qué ciudad se encuentra la unidad educativa de la cual es profesor?
b1.b11.	¿Cuál es el nivel de instrucción adquirido? [Profesor Normalista]
b1.b12.	¿Cuál es el nivel de instrucción adquirido? [Técnico]
b1.b13.	¿Cuál es el nivel de instrucción adquirido? [Licenciatura]
b1.b14.	¿Cuál es el nivel de instrucción adquirido? [Especialidad]
b1.b15.	¿Cuál es el nivel de instrucción adquirido? [Maestría]
b1.b16.	¿Cuál es el nivel de instrucción adquirido? [Doctorado]
b1.other.	¿Cuál es el nivel de instrucción adquirido? [Otro]
b2	¿Cuántos años de experiencia tiene como Profesor en toda su vida laboral?
b6	¿Cuál es el área en la que imparte sus clases?

Table 3: Muestra de preguntas pre-seleccionadas del dataset: Encuesta Final-Profesores de Inclusión Digital

Esto en un archivo .txt donde se recopiló manualmente el código de las preguntas seleccionadas desde el diccionario de datos como se muestra en la *Tabla 4* en la columna: *Cod. Pregunta*

Posterior a ello se realizan los siguientes procesos de limpieza de datos:

- Limpieza la cabecera de datos del diccionario de datos para que concuerden los códigos de las preguntas del diccionario de datos como del dataset.
- Extracción de la lista de preguntas pre-seleccionadas según código de pregunta del archivo *main_qcodes.txt*.
- Filtración y exportación de un nuevo dataset que cuenta únicamente con el conjunto de preguntas pre-seleccionadas obtenidas previamente.

Para el dataset *Encuesta Nacional de Opinión sobre TIC* debido a que el diccionario de datos estaba en un formato poco amigable para su uso y extracción de información se realizó un Script para obtener los datos necesarios para el análisis posterior del dataset y de este nuevo diccionario de datos que estamos formando, para este caso se eligió como salida un archivo de tipo JSON con las siguientes características:

```
{
  "nq": "3.", // Número de pregunta. Ejm. P3
  "question": "¿Cuándo fue la última vez que navegó en Internet
o estuvo co nectado a las redes sociales?", // Pregunta
  "qtype": "Numérico", // Tipo de respuesta obtenido
  "rtype": "Nominal", // Tipo de respuesta que representa
  "options": { // Opciones de respuesta qtype: rtype
    "1": "Hace 7 días o menos",
    "2": "Entre 8 a 15 días",
    "3": "Entre 16 a 30 días",
    "4": "Hace más de 30 días"
  }
}
```

- nq: Representa la posición de la pregunta en el dataset.
- question: Representa la pregunta a la que hace referencia.
- qtype: Representa el tipo de dato de la respuesta obtenida.

- `rtype`: Representa el tipo de dato de la respuesta que representa.
- `options`: Representa la codificación de las respuestas posibles, de la forma: `qtype: rtype`. Algunas preguntas no cuentan con este atributo.

De esta manera tenemos un nuevo diccionario de datos que abarca un 93% de las preguntas totales del diccionario de datos original; esto debido a complicaciones con la extracción de datos desde el archivo PDF. Este nuevo diccionario de datos nos brinda la posibilidad de, al igual que en el caso anterior; pre-seleccionar las preguntas que nos puedan brindar más detalle según el objetivo de nuestro análisis por lo cual se seleccionaron de manera manual las siguientes:

Cod. Pregunta	Pregunta
P1	¿Me podría decir su edad por favor?
P2	¿Usted ha navegado por Internet o se ha conectado a alguna red social a través de Internet?
P1A	¿Cuáles son las razones por las cuales no se conecta a Internet regularmente?
P1B	¿Cuáles son las razones por las cuales no se conecta a Internet regularmente?
P1C	¿Cuáles son las razones por las cuales no se conecta a Internet regularmente?
P1D	¿Cuáles son las razones por las cuales no se conecta a Internet regularmente?
P6	¿Tiene computadora de escritorio, computadora portátil o tablet en su casa?
P11A	¿Qué usos le da usted a estos equipos?

Table 4: Muestra de preguntas pre-seleccionadas del dataset: "Encuesta Nacional de Opinión sobre TIC"

Esto en un archivo .txt donde se recopiló manualmente el código de las preguntas seleccionadas desde el diccionario de datos original.

Posterior a ello se realizan los siguientes procesos de limpieza de datos:

- Agregado de datos faltantes a nuestro nuevo diccionario de datos según las preguntas pre-seleccionadas.
- Sustitución de código de pregunta por índice de pregunta en el nuevo diccionario de datos generado.
- Extracción de la lista de preguntas pre-seleccionadas según código de pregunta del archivo *main_qa.txt*.
- Filtración y extracción de un nuevo dataset únicamente con el conjunto de preguntas pre-seleccionadas obtenido previamente.

Procesar datos en búsqueda de insights

5. Determinar indicadores y sus pesos que nos ayuden en la modelación de un dataset.
6. Generar un dataset enriquecido.
7. Visualizar insights encontrados en los procesos previos mediante dashboards.

Extensión sugerida del capítulo, entre 8 y 12 páginas. En esta sección los resultados son presentados y analizados. Se presentan porcentajes, números/cantidades/valores correspondientes a los resultados además del ANÁLISIS de los mismos.

El módulo V, correspondiente a la analítica y visualización de datos será utilizado en toda su extensión, presentando gráficos elaborados, llamativos, descriptivos, adecuados, que muestren de la mejor manera los resultados de interés. Tomar en cuenta todo lo aprendido en este módulo (V) sin dejar de lado el análisis de los mismos.

La discusión corresponde a realizar un contraste entre nuestros resultados y los de otro proyecto, comparando, detectando similitudes y diferencias aplicando criterios de análisis de donde surgirán recomendaciones.

Ante cualquier duda, en cualquier sección, es muy recomendable buscar ejemplos en proyectos de investigación de otras instituciones, otros países, otros ámbitos, etc. La revisión bibliográfica respecto al tema es esencial.

7.1. Resultados de

El (figura 4-1) muestra que el 16% del área tiene una menor a, ver figura 4-5.

Figura 4-1: Título de la figura

El

Se encontró -----breve y clara.

Se ha encontrado o se encontró

Se ha obtenido o se obtuvo (pasado simple)

Incluye tablas y figuras con datos relevantes (no datos secundarios ni tablas ampulosas, poner estas en anexos en el CD).

7.2. Discusión de resultados

Los resultados de ----otro autor, muestran -----(ver figura 4.x); por ende a efectos de poder efectuar la comparación _____ resultado de ello se aprecia en la figura 4-11.

Figura 4-10: figura de otra investigación, de otro autor (incluir fuente)

Figura 4-11: Figura comparativa con nuestros resultados

El resultado de la otra investigación..... Comparando visualmente ambos ____ se observa que continúa mostrando los mismos resultados en..... Similitudes en..... diferencias en.....

Estas diferencias o igualdades, se debe a que en el presente estudio lo cual muestra que....., puede ser analizada en función a.....

Poner de forma clara la decisión que toma el autor del presente proyecto, explicando el porqué.

8. Conclusiones

Poner afirmaciones basadas o respaldadas/fundamentadas en el presente proyecto..

Extensión sugerida entre $\frac{3}{4}$ y 1 página.

9. Recomendaciones

A partir

Se recomienda

Se debe tener presente que los resultados obtenidos son producto de, que tienen involucradas hipótesis y condiciones iniciales particulares por lo tanto, estos resultados deben irse calibrando con el transcurso del tiempo.

Extensión sugerida entre $\frac{3}{4}$ y 1 página.

EXTENSION TOTAL DE DOCUMENTO, ENTRE 30 - 60 PÁGINAS.

Referencias bibliográficas

Campero Bustillos, J. C. B. (2016). Infraestructura de telecomunicaciones y TIC en Bolivia. Internetbolivia.

<https://internetbolivia.org/wp-content/uploads/2017/05/Campero-merged.pdf>

La Voz. (2020, May 22). *Iniciativas TIC de Bolivia para manejar la crisis*. La Voz de Bolivia. Retrieved January 10, 2023,

<https://lavoizdebolivia.com/iniciativas-tic-de-bolivia-para-manejar-la-crisis/>

Palacios, M. (2019). Diseño de un plan de negocio para la producción y comercialización de fruta deshidratada en la ciudad.

<https://www.docsity.com/es/plan-de-negocio-fruta-deshidratada/5776838/>

M. Shaw y D. Garlan, Software architecture: Perspectives on an emerging discipline. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

I. Sommerville, Software Engineering, 10th ed. Essex, England: Pearson Education Limited, 2016.

PyPDF2 (3.0.1). (s. f.). [Librería]. readthedocs.

<https://pypdf2.readthedocs.io/en/3.0.0/index.html>

Barreto, C. R., & Diazgranados, F. I. (2017, 1 1). Las TIC en la educación superior: experiencias de innovación.

[https://books.google.com.bo/books?](https://books.google.com.bo/books?id=YLBJDwAAQBAJ&lpg=PP1&ots=XZGfvDZkhO&dq=tics%20en%20educaci%C3%B3n&lr&pg=PP1#v=onepage&q&f=false)

[id=YLBJDwAAQBAJ&lpg=PP1&ots=XZGfvDZkhO&dq=tics%20en](https://books.google.com.bo/books?id=YLBJDwAAQBAJ&lpg=PP1&ots=XZGfvDZkhO&dq=tics%20en%20educaci%C3%B3n&lr&pg=PP1#v=onepage&q&f=false)

[%20educaci%C3%B3n&lr&pg=PP1#v=onepage&q&f=false](https://books.google.com.bo/books?id=YLBJDwAAQBAJ&lpg=PP1&ots=XZGfvDZkhO&dq=tics%20en%20educaci%C3%B3n&lr&pg=PP1#v=onepage&q&f=false)

Datos Abiertos Bolivia. (n.d.). Datos Abiertos Bolivia. Retrieved Diciembre 20, 2022, from <https://datos.gob.bo/about>

Bnamericas. (2020, May 20). Un vistazo a las iniciativas de TIC de Bolivia para manejar la crisis. BNamericas. Retrieved January 8, 2023,

from <https://www.bnamericas.com/es/analisis/un-vistazo-a-las-iniciativas-tic-de-bolivia-para-manejar-la-crisis>

Van Rossum, G., & Drake Jr, F. L. (1995). Python tutorial. Technical report, Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

Ghezzi, C., Jazayeri, M., & Mandrioli, D. (2017). Fundamentos de ingeniería del software. Pearson.

McKinney, W. (2011). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. O'Reilly Media.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., ... IPython Development Team. (2016). Jupyter Notebook-A Publishing Format for Reproducible Computational Workflows. In Positioning and Power in Academic Publishing (pp. 87-90). International Society for Computers and their Applications.

Tableau. (2021). Tableau Prep Builder. <https://www.tableau.com/products/prep-builder>

Anexos

Anexo 1. Resultados de encuesta sobre