Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Económicas
Escuela Profesional de Economía

Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos: Caso Banco de la Nación periodo 2022

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de Economista

AUTOR

Angel JORGE SALAZAR

ASESOR

Mg. Ivonne Yanete VARGAS SALAZAR

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Jorge, A. (2022). *Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos: Caso Banco de la Nación periodo 2022*. [Trabajo de Suficiencia Profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Económicas, Escuela Profesional de Economía]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Angel Jorge Salazar
Tipo de documento de identidad	DNI
Número de documento de identidad	71552326
URL de ORCID	https://orcid.org/0009-0002-3041-9016
Datos de asesor	
Nombres y apellidos	Ivonne Yanete Vargas Salazar
Tipo de documento de identidad	DNI
Número de documento de identidad	16655197
URL de ORCID	https://orcid.org/0000-0002-0836-5309
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Madalyne Motta Flores
Tipo de documento	DNI
Número de documento de identidad	23998380
Miembro del jurado 1	
Nombres y apellidos	Encarnación Margarita Palomino Ladrón De Guevara
Tipo de documento	DNI
Número de documento de identidad	25583244
Datos de investigación	

Línea de investigación	No aplica.
Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	Edificio: Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Lima Centro poblado: No aplica Urbanización: No aplica Manzana y lote: No aplica Calle: Av. Germán Amézaga N.° 375 Latitud: -12.05659 Longitud: -77.08598
Año o rango de años en que se realizó la investigación	Marzo 2022 - Setiembre 2022
URL de disciplinas OCDE	Ciencias de la computación, inteligencia artificial https://purl.org/pe-repo/ocde/ford#1.02.01 Ciencias de la computación, sistemas de información https://purl.org/pe-repo/ocde/ford#1.02.02 Ingeniería de sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.01

ACTA DE SUSTENTACIÓN(VIRTUAL) DEL TRABAJO DE SUFICIENCIA PROFESIONAL PARA OPTAR EL TITULO PROFESIONAL DEECONOMISTA

Reunidos en la Plataforma Virtual Google Meet el día sábado 10 de setiembre de 2022 a las **4:00** pm, ante el Jurado presidido por el catedrático **Mg. MADALYNE MOTTA FLORES** e integrado por los catedráticos: **Econ. ENCARNACIÓN MARGARITA PALOMINO LADRÓN DE GUEVARA y Mg. IVONNE YANETE VARGAS SALAZAR.**

Se presenta el Bachiller **ANGEL JORGE SALAZAR** para realizar la defensa de su Trabajo de Suficiencia Profesional **"PROYECTO DE LOCALIZACIÓN DE ZONAS GEOGRÁFICAS PARA LA IMPLEMENTACIÓN DE CAJEROS AUTOMÁTICOS: CASO BANCO DE LA NACIÓN PERIODO 2022 "** presentada para la obtención del Título Profesional de Economista.

De acuerdo a las normas vigentes el Bachiller después de finalizada la exposición absolvió las preguntas planteadas por los miembros del jurado. Concluida esta fase se suspendió el Acto Virtual, a fin de que los señores catedráticos en votación secreta procedieran a la calificación correspondiente

Acto seguido el presidente del Jurado informó que el graduando había sido **APROBADO** por **UNANIMIDAD** con el calificativo de **18 (dieciocho)**

Para dar fe del acto, suscriben la presente acta el presidente, los Miembros del Jurado y el Vicedecano Académico de la Facultad de Ciencias Económicas, a los 10 días del mes de Setiembre de 2022.

Mg. MADALYNE MOTTA FLORES
Presidente

Econ. ENCARNACIÓN MARGARITA PALOMINO LADRÓN DE GUEVARA
Miembro

Mg. IVONNE YANETE VARGAS SALAZAR
Asesor

Mg. MIGUEL HILDEBRANDO PISFIL CAPUÑAY
Vicedecano Académico

UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

Anexo 1

CERTIFICADO DE SIMILITUD

Yo, Ivonne Yanete Vargas Salazar en mi condición de asesor acreditado con el Dictamen EP N° 000002-2026-EPE-FCE/UNMSM del Trabajo de Suficiencia Profesional, cuyo título es PROYECTO DE LOCALIZACIÓN DE ZONAS GEOGRÁFICAS PARA LA IMPLEMENTACIÓN DE CAJEROS AUTOMÁTICOS: CASO BANCO DE LA NACIÓN PERIODO 2022, presentado por el bachiller ANGEL JORGE SALAZAR, para optar el Título Profesional de Economista.

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 16% de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título profesional de Economista.

Lima, 20 de febrero de 2026

Firma del Asesor _____

Nombres y Apellidos del asesor: Ivonne Yanete Vargas Salazar

DNI: 16655197

Tabla de contenido

RESUMEN iii

INTRODUCCIÓN 1

I. INSTITUCIÓN Y FUNCIONES DESEMPEÑADAS 2

1.1. Banco de la Nación 2

1.2. Cargo y funciones desempeñadas 3

II. ACTIVIDADES 4

2.1. Actividad 1: Desarrollo del proyecto de localización de zonas geográficas para la implementación de cajeros automáticos. 4

2.1.1. Antecedentes metodológicos y teóricos. 4

2.1.2. Desarrollo del proyecto 5

2.1.2.1. Funciones vinculadas al proyecto 6

2.1.2.2. Finalidad y objetivo del proyecto 6

2.1.2.3. Problemática vinculada al proyecto 7

2.1.2.4. Metodología y procedimientos del proyecto 7

a. Comprensión del negocio 7

b. Comprensión de los datos 8

c. Preparación de los datos 9

d. Modelamiento 10

e. Evaluación 13

f. Implementación 13

2.1.2.5. Resultados del proyecto 14

2.2. Conocimientos técnicos de la profesión utilizados 15

2.3. Bases teóricas de la profesión utilizadas 16

III. ANÁLISIS Y DIAGNÓSTICO DEL PROBLEMA 17

3.1. Análisis de actividades desempeñadas 17

3.1.1. Planteamiento del problema acerca de las funciones desempeñadas 17

3.1.2. Diagnóstico del problema acerca de las funciones desempeñadas 18

IV. RECOMENDACIONES 19
4.1. Recomendaciones de mejora de actividades 19
4.1.1. Fundamentación 19
4.1.2. Medidas específicas 21
V. PERFORMANCE FUTURO DEL DESEMPEÑO PROFESIONAL 23
REFERENCIAS BIBLIOGRÁFICAS 26
WEBGRAFÍA 28
ANEXOS 29
GLOSARIO 37

LISTA DE FIGURAS

Figura 1: Estructura funcional 3
Figura 2: Zona rural 11
Figura 3: Zona metrópolis 12
Figura 4: Automatización del análisis financiero-económico 13
Figura 5: Dashboard de resultados 14
Figura 6: Diagrama de Ishikawa de data lake 18
Figura 7: Propuesta de implementación de data lake 23
Figura 8: Línea de carrera 25

LISTA DE ANEXOS

Anexo 1: Organigrama general de Banco de la Nación 29
Anexo 2: Organigrama específica del Banco de la Nación 30
Anexo 3: Contrato de locador de servicios 31
Anexo 4: Gantt del proyecto 30
Anexo 5: Variables utilizadas en la zona rural 31
Anexo 6: Variables utilizadas en la zona metrópolis 31
Anexo 7: Variables utilizadas para análisis financiero-económico 32
Anexo 8: Ejemplo de buffer 32
Anexo 9: Reporte de zonas geográficas para la implementación de cajeros automáticos 33
Anexo 10: Impacto en cobertura 34
Anexo 11: Ficha de iniciación del proyecto 35
Anexo 12: Ficha de finalización del proyecto 36

RESUMEN

El presente informe titulado Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos: Caso Banco de la Nación periodo 2022, se detallan las funciones que desempeño como científico de datos en el Banco de la Nación y el desarrollo del proyecto de localización de zonas para la implementación de cajeros automáticos y las distintas metodologías utilizadas en el desarrollo del proyecto. Además, en este informe se detallan los conocimientos técnicos de mi profesión utilizados dentro de mis funciones y para el desarrollo del proyecto. Finalmente se mencionan las recomendaciones para mi profesión como economista desde el punto de vista de mis funciones como científico de datos.

El proyecto de localización de zonas geográficos para la implementación de cajeros automáticos, tiene como objetivo para la actividad 1, realizar un estudio para la determinación de la cantidad de zonas geográficas para la implementación de cajeros automáticos. Y además tiene como objetivo general detallarnos las funciones que se desempeñan en la elaboración de la actividad 1 y generar recomendaciones.

El presente informe se realiza mediate un análisis descriptivo, analítico, reflexivo y propositivo de los criterios que se tomó en cuenta para la realizar el estudio de localización de zonas geográficas para la implementación de cajeros automáticos.

Palabras claves: formación en horas de trabajo, Instituciones financieras, Análisis costes-beneficio, máquina, ATM.

ABSTRACT

This report entitled Project for the location of geographical areas for the implementation of ATMs: Case of Banco de la Nación period 2022, details the functions that I perform as a data scientist at Banco de la Nación and the development of the project for locating areas for the implementation of ATMs and the different methodologies used in the development of the project. In addition, this report details the technical knowledge of my profession used within my functions and for the development of the project. Finally, the recommendations for my profession as an economist are mentioned from the point of view of my functions as a data scientist.

The project for the location of geographic areas for the implementation of ATMs, has as its objective for activity 1, to carry out a study to determine the number of geographic areas for the implementation of ATMs. And it also has the general objective of detailing the functions that are performed in the preparation of activity 1 and generating recommendations.

This report is carried out through a descriptive, analytical, reflective and purposeful analysis of the criteria that were taken into account to carry out the study of the location of geographical areas for the implementation of ATMs.

Keywords: Inservice training, Financial institutions, Cost benefit analysis, Machine, ATM.

INTRODUCCIÓN

El informe a continuación presenta el análisis del proyecto de la localización de zonas geográficas para la implementación de cajeros automáticos y la importancia de los cajeros automáticos en el Perú para disposición de efectivo y otras operaciones como alternativa al uso de canales físicas tales como las agencias y agentes y para el 2020 la cifra era de 124.3 cajeros automáticos por cada 100000 adultos peruanos (Banco Mundial, 2020).

Se optó en dividir el informe en cinco capítulos. En el capítulo I se describe a la empresa Banco de la Nación, el área y funciones desempeñadas dentro de ella. En el capítulo II se plantea el caso específico: desarrollo del proyecto de localización de zonas geográficas para la implementación de cajeros automáticos, Además se detalla la metodología y bases teóricas utilizadas para la elaboración del proyecto y análisis detallado del estudio y conocimientos técnicos utilizados de la profesión y las bases teóricas utilizadas de la profesión. En el capítulo III se analiza las actividades desempeñadas en la institución. En el capítulo IV se plantea las recomendaciones para la carrera de economía en nuestra casa de estudios, cuando he desempeñado el cargo de científico de datos y las conclusiones del proyecto y las funciones, en el capítulo V se detalla el performance futuro del desempeño profesional.

En el informe se presenta la institución en la cual laboro, mis funciones como científico de datos, el desarrollo del proyecto de localización de zonas geográficos para implementación de cajeros automáticos, las bases teóricas de mi carrera que me sirve para el desarrollo de mis funciones, recomendaciones para mejorar los problemas observados en la institución de labor y de estudio y finalmente la línea de carrera que se tiene como científico de datos y cómo se relaciona con mi carrera de economía.

I. INSTITUCIÓN Y FUNCIONES DESEMPEÑADAS

1.1. Banco de la Nación

El Banco de la Nación es una empresa de derecho público, integrante del Sector Economía y Finanzas, que opera con autonomía económica, financiera y administrativa. Actúa como agente financiero del Estado, atiende la deuda pública externa y las operaciones de comercio exterior, cuando el Ministerio de Economía y Finanzas lo requiera. Asimismo, recauda tributos y efectúa pagos, sin que esto sea exclusivo.

Las funciones que se le asignaron al Banco de la Nación son varias, de las cuales resalto las funciones relacionadas con mi área de trabajo:

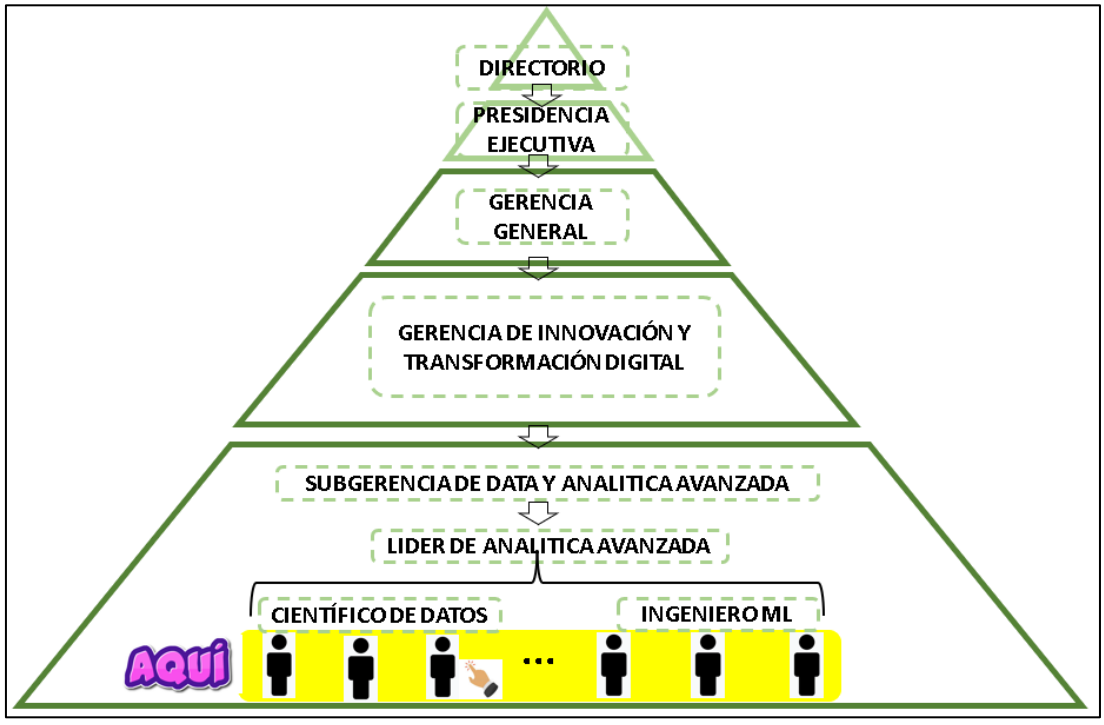
- Actuar como Agente Financiero del Estado.
- Actuar por cuenta de otros Bancos o Financieras, en la canalización de recursos.
- Brindar Servicios de Cuentas Corrientes a las Entidades del Sector Público Nacional y a Proveedores del Estado.
- Recibir depósitos de ahorros en lugares donde la banca privada no tiene oficinas.

La Misión del banco es “Brindar servicios de calidad a la ciudadanía y al Estado, ampliando nuestra cobertura de servicios y promoviendo la inclusión financiera, a través de una gestión moderna y auto-sostenible” y su visión es “Ser reconocido por brindar soluciones financieras, innovadoras, inclusivas y de calidad al Estado Peruano y la ciudadanía, dentro de un marco de gestión basado en la efectividad de sus procesos y prácticas de Buen Gobierno Corporativo”.

El Banco de la Nación está organizado de la siguiente forma, en el rango más alto está el directorio y la presidencia ejecutiva, luego viene la gerencia general, luego por debajo están las gerencias de distintas áreas, y finalmente las subgerencias de cada área y en mi caso soy parte

del equipo de científicos de datos del Banco de la Nación y tenemos un líder de data y analítica avanzada que estamos sujetos a la subgerencia de data y analítica avanzada y está sujeta a la gerencia de innovación y transformación digital y está por debajo de la gerencia general y presidencia ejecutiva, tal como se muestra en la Figura 1.

Figura 1: *Estructura funcional*



Nota: Elaboración del bachiller. Fuente: Banco de la Nación.

El organigrama del Banco de la Nación actualizado al año 2022 se encuentran en los Anexo 1 y Anexo 2.

1.2. **Cargo y funciones desempeñadas**

Me desempeño como científico de datos en el Banco de la Nación y mis funciones son los siguientes:

- Identificar y proponer de manera proactiva oportunidades de construcción de modelos de Machine Learning que contribuirán a la optimización de “gestión de Efectivo” en

ATM[1] y agencias.

- Obtener la data de los servidores del banco y realizar traslados de la información al entorno cluster de Advanced analytics para el desarrollo de los modelos de optimización de “Gestión de Efectivo” en ATM y agencias.
- Definir el objetivo, las variables involucradas y reportar los indicadores de calidad del modelo analítico definido los mismos que serán documentados de forma técnica (en relación al ítem anterior).
- Hacer seguimiento a los indicadores de los modelos generados para la optimización de “Gestión de Efectivo” en ATM y agencias, así como también definir la vigencia de los modelos y la necesidad de recalibración o update.

En el Anexo 3 se puede observar mi contrato de locador de servicios como científico de datos.

II. ACTIVIDADES

2.1. Actividad 1: Desarrollo del proyecto de localización de zonas geográficas para la implementación de cajeros automáticos.

2.1.1. Antecedentes metodológicos y teóricos.

En esta parte del informe se detallan algunos antecedentes y bases teóricas, que ayudan al desarrollo del proyecto de localización de zonas geográficas para la implementación de cajeros automáticos.

La importancia de contar con un análisis de ubicación geográfica se centra en la necesidad de utilizar herramientas que permitan optimizar la localización de los cajeros automáticos. Seguido

[1] Cajero automático (Automated treller machine)

en este sentido Muñoz Diaz (2013), presenta el análisis de las herramientas de información geográfica como apoyo en la ubicación de cajeros automáticos y oficinas bancarias, para mejorar la cobertura de los servicios bancarios para el caso de la ciudad de Bogotá.

La importancia del análisis de la ubicación de cajeros automáticos se centra en la necesidad de utilizar una metodología que permitan optimizar la localización de los cajeros automáticos. Seguido en este sentido Rivas Martínez (1999). Presenta un estudio que determina el número óptimo de Cajeros Automáticos en la ciudad de Guayaquil, para lo cual se aplicó un modelo de "Ubicación de Cajeros".

2.1.2. Desarrollo del proyecto

Dentro de este capítulo se detalla el flujo de trabajo, siguiendo la metodología CRISP-DM[2]. En el caso del Banco de la Nación para el desarrollo de proyecto se utiliza las herramientas tales como: Python (Es el lenguaje de programación popular hasta la fecha para implementación de proyectos de ciencia de datos), QGIS (Es una herramienta para manipular base de datos a nivel geográfico), Power BI (Herramienta de visualización), Excel (Herramienta de visualización y manipulación de datos) y VBA (Es una herramienta parte de Excel y lenguaje de programación que nos apoya para automatizar procesos).

Además, en el área de data y analítica avanzada se utiliza el Gantt, para hacer seguimiento y poner plazos al proyecto de localización de zonas geográficas que tiene una fecha de inicio de 02 de diciembre del 2021 y una fecha de fin de proyecto de 18 de mayo del 2022. El Gantt del proyecto se puede observar en el Anexo 4.

[2] CRISP-DM, son las siglas de Cross-Industry Standard Process for Data Mining, es un método probado para desarrollar proyectos de minería de datos.

2.1.2.1. Funciones vinculadas al proyecto

El desarrollo del proyecto de implementación de ATMs se trabajó de manera colaborativa entre 2 colaboradores científico de datos, en este caso particular trabajé en equipo con una compañera[3] y ambos desarrollamos las mismas funciones siendo responsables del proyecto al 50% cada uno. Mis funciones asignadas son las siguientes:

i) Comprensión del negocio, Identificación, propuesta y construcción de modelos de Machine Learning para la implementación de ATMs.

ii) Comprensión de los datos, definición del objetivo del proyecto y la definición de variables involucradas en el modelo analítico, descriptivo y predictivo.

iii) Preparación de los datos, obtención de la data de los servidores de banco y de acuerdo a lo que se requiere para el proyecto y realizar traslados de la información al entorno clúster de advanced analytics, para el desarrollo del análisis y modelos de optimización en ATMs,

iv) Modelamiento, la construcción de modelos de Machine Learning.

v) Evaluación, la validación de los resultados con el área usuaria de los resultados obtenidos.

vi) Implementación, hacer el seguimiento a los indicadores de los modelos y metodologías elaborados para ATMs,

2.1.2.2. Finalidad y objetivo del proyecto

La finalidad y objetivo del proyecto es la determinación de la cantidad de zonas geográficas para la implementación de ATMs (referencial, de acuerdo a la posibilidad de implementación y presupuesto del área usuaria), el análisis de rentabilidad financiero-económico y la ubicación

[3] Eliza Lastra Medina: Bachiller en economía por la UNMSM.

geográfica para cada punto detallado.

2.1.2.3. Problemática vinculada al proyecto

El área de planeamiento del Banco de la Nación como una parte de sus roles tiene la necesidad de implementar nuevos ATMs, ya que ellos reciben pedidos de congresistas o presidencia y se desconoce si el lugar es el adecuado para su implementación y además si un ATM está saturado donde está ocurriendo y se requiere adicionar un nuevo ATM y además para su ubicación se depende de un proveedor que lo ubica donde el proveedor tiene convenios pero no en el lugar idóneo o referencial dictado por el banco.

2.1.2.4. Metodología y procedimientos del proyecto

Antes de hacer el paso a paso es bueno aclarar que el proyecto de localización de zonas geográficas para la implementación de ATMs se divide en 2 formas de localización y la automatización del análisis financiero-económico para cada punto geográfico encontrado con las 2 formas anteriores mencionadas.
A continuación, detallamos el desarrollo del proyecto siguiendo la metodología CRISP-DM:

a. Comprensión del negocio

En esta etapa del proyecto, he participado de forma frecuente en reuniones con diversas áreas del banco, pero en especial con el área de planeamiento (área usuaria del proyecto), para entender la problemática acerca de la localización de nuevos puntos geográficos para la implementación de los nuevos cajeros automáticos ya que para el área de planeamiento es parte de su rol la entrega de puntos geográficos para la implementación de nuevos cajeros automáticos y una vez entendido el problema de la gerente y subgerente del área de planeamiento hice la

propuesta de localizar puntos geográficos con algoritmos de Machine Learning o con otras metodologías y todo depende de qué tipo de información se puede tener y obtener tanto información interna y externa. Además, he tenido reuniones con la gerente y subgerente del área de finanzas para entender los criterios e indicadores de la elaboración del análisis financiero y económico para la implementación del cajero automático.

b. Comprensión de los datos

Una vez que se ha comprendido las lógica de cada área usuaria acerca de la implementación, análisis financiero-económico de los cajeros automáticos, me dediqué a obtener los datos necesarios tanto internos como externos, las cuales me ayuden a fortalecer la propuesta inicial de localización de zonas geográficas con técnicas de Machine Learning y además obtener datos de forma separada tanto para la zona rural con cierta estructura que detallo más adelante, para la zona metrópolis con otra estructura y finalmente datos para la automatización del proceso de análisis financiero-económico. La recolección de datos separada y con otra estructura se da, porque para la zona rural se trabaja una metodología distinta que la zona metrópolis y en el análisis financiero ocurre lo mismo.

En el proyecto se obtiene la información para la zona rural, a nivel distrital de la cantidad de transacciones mensuales de los cajeros automáticos operativos actuales y del INEI[4] se obtiene información de la población, del MTC[5] se obtiene información de peajes y coberturas 3G, del MTPE[6] se obtiene información de trabajadores públicos y privados por distrito, del SBS[7], se obtiene la información de la situación financiera a nivel distrital. En el Anexo 5 se muestra la tabla. Para la zona Metrópolis se obtienen datos externos, realizando un levantamiento masivo de ubicaciones de comercios de Google Maps (Web scraping con EViews y la librería Selenium) y

[4] Instituto Nacional de Estadística e Informática
[5] El Ministerio de Transportes y Comunicaciones
[6] Ministerio del Trabajo y Promoción del Empleo
[7] Superintendencia de Banca, Seguros y AFP

se obtienen también datos de transacciones mensuales de cajeros automáticos, agentes y agencias. En el Anexo 6 se muestra la tabla.

Para el análisis financiero-económico se obtiene la información de transacciones mensuales de los cajeros automáticos operativos y base de datos brindados por el Área de finanzas y planeamiento del Banco de la Nación. Además, mencionar que para esta parte del desarrollo se requiere la información de puntos geográficos obtenidos tanto en la zona rural y metrópolis. Se omite la descripción de algunas variables por la sensibilidad de la información y se agrupa dentro de otras variables de costos y otras variables. En el Anexo 7 se muestra la tabla.

c. Preparación de los datos

Después de haber obtenido y comprendido los datos, continuamos con la fase de la preparación de los datos, tanto la preparación de datos para las zonas rurales, metrópolis y análisis económico financiero, a continuación, detallamos la preparación de datos para cada uno de ellos:

El procesamiento (limpieza de los datos, creación de nuevas variables, transformación de variables, estandarización de variables, tratamiento de valores nulos, etc.) de la información para la zona rural se hizo con la herramienta Python y finalmente se obtuvo una tabla en formato csv con los datos agrupados a nivel distrital y con los campos utilizados para la zona rural.

El procesamiento para las zonas metrópolis sigue el mismo flujo, que la zona rural, obteniendo un archivo final con la información a nivel de coordenadas geográficas de comercios y transacciones de cajeros automáticos, agentes y agencias.

El procesamiento para conseguir información para automatizar el análisis financiero-económico sigue el mismo flujo que las anteriores, obteniendo un archivo Excel a nivel de latitud y longitud, esto se da después de procesar los resultados de las zonas geográficas obtenidas de zona rural y metrópolis y combinarlos con la información de las plantillas compartidas del área de finanzas.

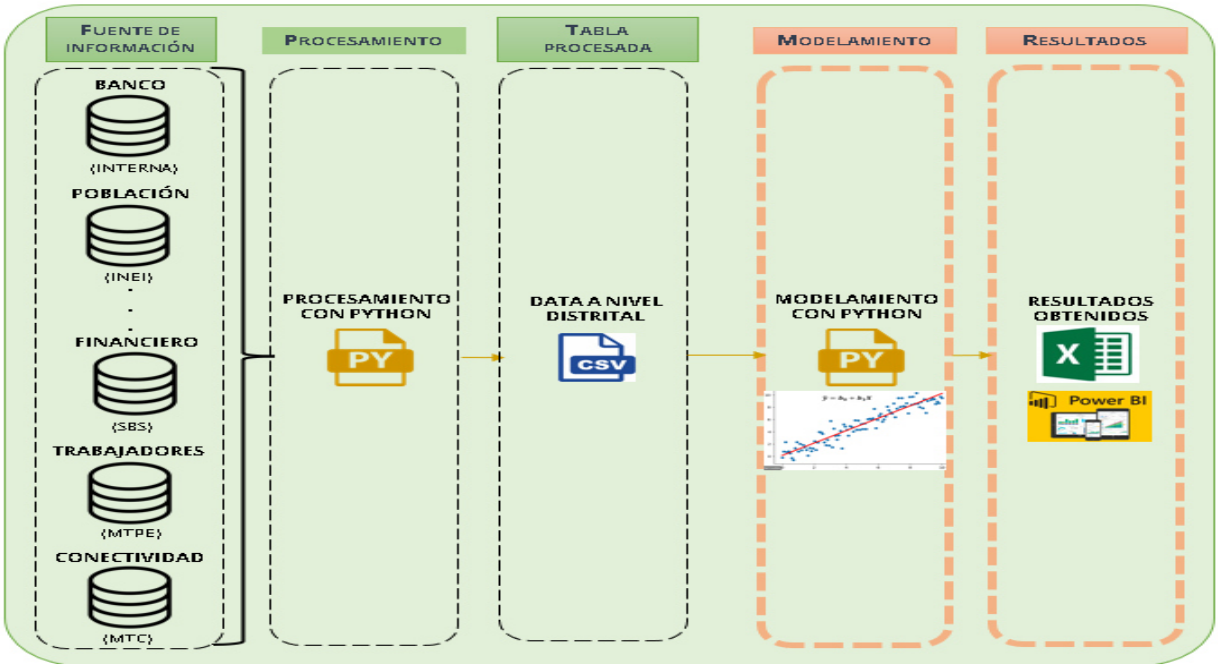
d. Modelamiento

Después de haber preparado los datos para zona rural y metrópolis, se procede con la parte de aplicar las distintas metodologías tanto para la zona rural y metrópolis, como se detalla a continuación,(1) para la zona rural, se emplea una regresión lineal múltiple, para proyectar transacciones en distritos que aún no cuentan con un ATM y (2) para la zona metrópolis se realiza una agrupación de comercios para, encontrar las zonas comerciales y estas zonas comerciales filtrarlas con los distintos canales(Agente, Agencia y ATMs) presentes del banco en dicho lugar y (3) Finalmente la parte de la automatización del análisis financiero-económico se da una vez se haya obtenido los resultados tanto para las zonas rurales y metrópolis, esta parte consiste en la evaluación financiera-económica (rentabilidad futura, si se implementa el cajero en dicho lugar) de cada posible punto geográfico para la implementación de un cajero automático. A continuación, se detalla las distintas metodologías aplicadas en los 3 casos mencionados.

En la zona rural, para obtener la cantidad de transacciones potenciales en distritos que no tienen oferta de cajeros automáticos, se desarrolló el modelo de regresión lineal, que nos apoya con aprender de la cantidad de transacciones existentes en dichos distritos que cuentan con al menos un cajero automático y para luego proyectar las transacciones potenciales en distritos que aún no cuentan con ningún cajero automático. Para nuestro proyecto la variable dependiente (y) es la suma de las transacciones de cajeros automáticos, agentes y agencias de operaciones realizables en un ATM a nivel distrital y las variables dependientes significativas son los siguientes: PEA, PENSIÓN 65, CRÉDITOS OTORGADOS POR EL BN y PRESENCIA BN y BP, etc. Una vez proyectado las transacciones en distritos que no cuentan con ningún cajero automático se resta las transacciones existentes de canales de agencia y agentes en el caso de que existe algunos de estos canales en dicho distrito y finalmente obtenemos las transacciones netas y aplicarles el siguiente criterio, siguiendo las directivas del banco, si se cumple que las

transacciones es mayor o igual a 4500 y PEA mayor o igual a 1500 personas, se concluye que es un distrito potencial, para la implementación de un cajero automático. Finalmente, los distritos potenciales, cuentan con algún centro poblado importante poblacionalmente, así que para ser más específico de elegir el lugar del potencial punto de implementación de un cajero automático, se elige el centro poblado más importante de los distritos potenciales obtenidos (Ver Figura 2).

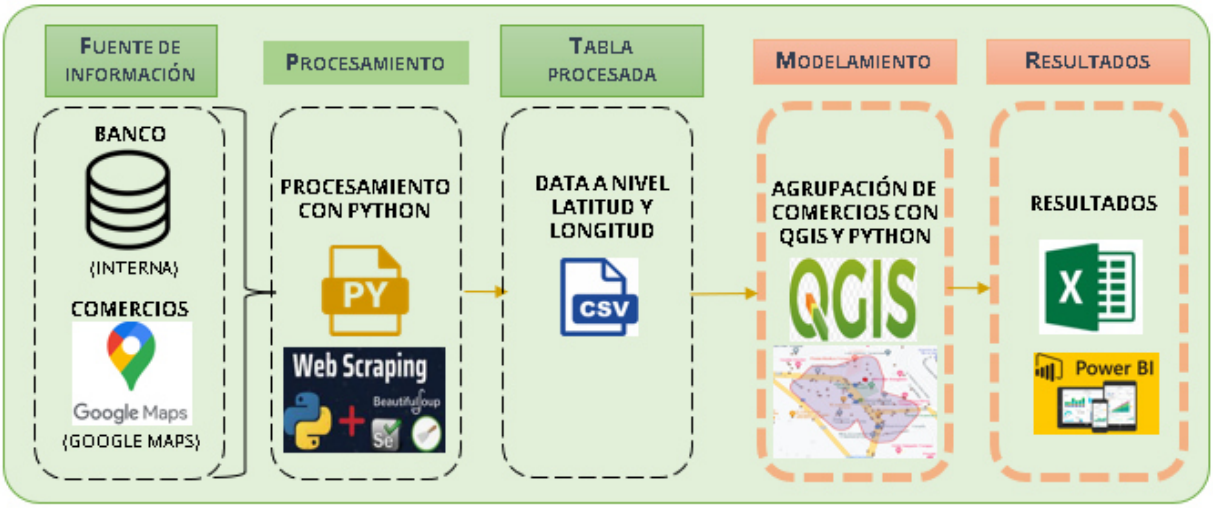
Figura 2: *Zona rural*



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

En la zona metrópolis, se trabaja con la agrupación de comercios con una cercanía entre cada comercio de aproximadamente 10 metros de distancia para encontrar zonas comerciales o polígonos (grupos) comerciales que son agrupaciones de varios comercios cercanos, para obtener los puntos comerciales, solo se consideran polígonos con un mínimo de 25 comercios dentro del polígono comercial. En la Figura 3 se muestra el flujo de trabajo.

Figura 3: *Zona metrópolis*



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Además, a la lista de los polígonos comerciales posibles para implementación de cajeros automáticos se le agrega la oferta de canales físicos existentes que encontramos dentro de cada polígono comercial, como filtro adicional se considera los que cumplen los siguientes criterios:

-1: polígonos comerciales sin ningún canal del banco

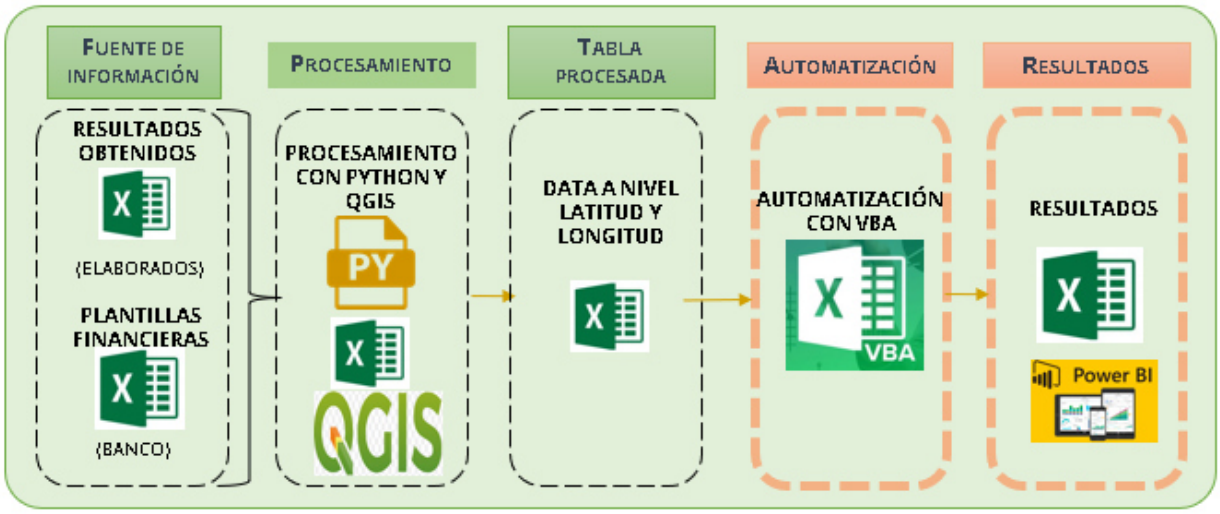
-2: polígonos comerciales que tienen al menos un cajero automático con alta demanda

-3: polígonos comerciales con al menos algún canal físico (agente y/o agencia). En el Anexo 8 se muestra el gráfico de un ejemplo de la zona comercial.

Finalmente, como parte del proyecto requiere del análisis financiero-económico (rentabilidad futura) para cada punto geográfico obtenido tanto de zonas rurales y metrópolis. Se ha analizado la plantilla de trabajo del área de finanzas sobre el análisis financiero-económico por cada punto donde se van a implementar los cajeros automáticos y se automatiza el proceso de evaluación financiera-económica de forma masiva con Macros y Visual Basic. Se obtiene el resultado de cuatro indicadores financieros-económicos (VAN financiero, VAN económico, Costo-beneficio,

punto de equilibrio y brecha de punto equilibrio-demanda) de los puntos posibles geográficos para la implementación de cajeros automáticos. En la Figura 4 se muestra la secuencia de trabajo.

Figura 4: *Automatización del análisis financiero-económico*



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

e. Evaluación

Se realiza la validación de algunos puntos geográficos obtenidos con área usuaria (área de planeamiento).

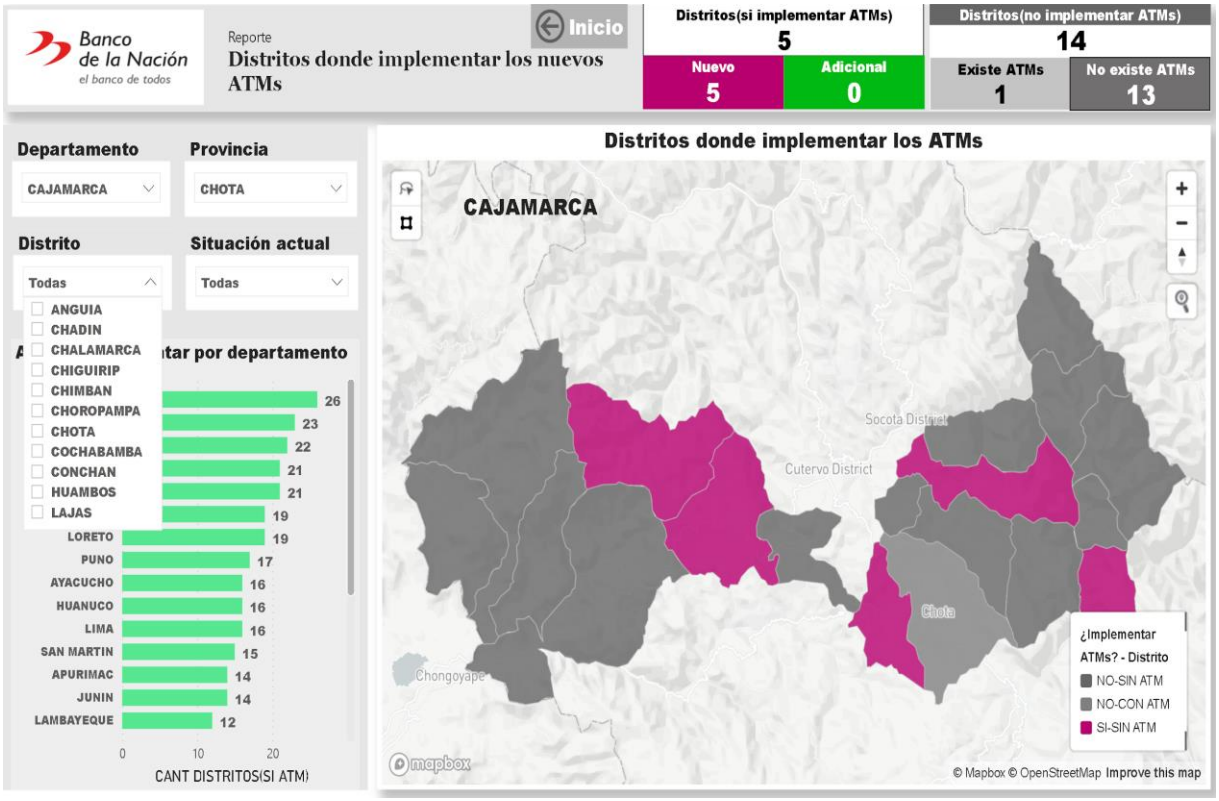
f. Implementación

La implementación para el proyecto que he desarrollado como automatización del uso de la información y con qué frecuencia. Se actualiza la información de posibles puntos geográficos para la implementación de cajeros automáticos, teniendo en cuenta la definición de implementación para nuestra área. Por la frecuencia de uso de la información por parte del área usuaria la automatización no fue posible, ya que requiere una frecuencia anual de actualización.

2.1.2.5. Resultados del proyecto

Los resultados del proyecto es un reporte en Excel con la información de 1399 puntos o coordenadas geográficas que nos indican que son los posibles puntos geográficos donde se pueden implementar ATMs y además en este reporte hay una columna adicional que nos indica si ese punto geográfico es rentable o no rentable con los indicadores financieros y finalmente esta misma información también lo tenemos en Power BI para realizar su respectivo monitoreo. En la Figura 5, se muestra la vista del Power BI y en el Anexo 9 y Anexo 10 podemos observar el reporte en Excel y el impacto.

Figura 5: *Dashboard de resultados*



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

2.2. Conocimientos técnicos de la profesión utilizados

- Econometría, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado análisis de correlación de variables, para la determinación de variables significativas y se utiliza para el desarrollo del modelo de regresión lineal múltiple.
- Estadística, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado el análisis de datos internos y externos del banco, por ejemplo: determinación de la cantidad de transacciones en un distrito sin oferta de cajero automático con un nivel de probabilidad.
- Matemática financiera, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado el análisis financiero y económico, por ejemplo, hallar el valor actual neto de las transacciones futuras de un posible punto para la implementación de un cajero automático.
- Análisis financiero, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado el análisis financiero y económico, por ejemplo, hallar el indicador de rentabilidad de las transacciones futuras de un posible punto para la implementación de un cajero automático.
- Evaluación de proyecto privados y públicos, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado la evaluación del proyecto al inicio y al final del proyecto, por ejemplo, hacer una evaluación y dar el tiempo de desarrollo del proyecto.
- Informática para economistas, es una herramienta necesaria que he utilizado como parte del proyecto, cuando he realizado el entrenamiento del modelo, la manipulación y el análisis de los datos, por ejemplo, Ejecutar el modelo con Python e interpretar el resultado y en la automatización del proceso de análisis financieros-económicos con Macros y VBA.

Las evidencias del desarrollo del proyecto son las fichas de iniciación y finalización del proyecto con las áreas implicadas y podemos observar en el Anexo 11 y Anexo 12.

2.3. Bases teóricas de la profesión utilizadas

(1) Predicción de transacciones

- La **econometría,** interviene como herramienta que permite cuantificar las relaciones entre las variables y ayuda a predecir las transacciones futuras.
- **La modelación econométrica,** trata de definir una simplificación de la realidad a partir de un modelo matemático que la explique y considerando el efecto del error aleatorio implicado en el proceso (Gujarati y Porter, 2010).
- **Modelos de regresión múltiple** pueden aceptar diversas variables explicativas que tal vez estén correlacionadas, puede esperarse inferir causalidad en casos en los que el análisis de regresión simple podría no dar buenos resultados (Wooldridge, 2010).
- **Coeficiente de determinación (R2),** el cual, según Wooldridge (2009), es un número que nos permite saber hasta qué punto la recta de regresión estimada a través de la estimación por Mínimos Cuadrados Ordinarios (MCO) se ajusta a los datos muestrales.
- **Los mínimos cuadrados ordinarios** (**MCO**) es un tipo de método de mínimos cuadrados lineales para estimar los parámetros desconocidos en un modelo de regresión lineal (Wooldridge, 2009).

(2) Probabilidades y promedios

- **La Teoría de la Probabilidad** es la encargada de estudiar los fenómenos aleatorios y mediante ciertos axiomas**.** Blaise Pascal (1623-1662).
- **Las medidas de tendencia central (MTC),** diversos autores señalan su importancia en la construcción y comprensión de otros conceptos estadísticos (e.g., Batanero, 2001; Cobo, 2003; Escobedo y Mayén, 2018; Mayén, 2009).

- **Estadística**, la palabra estadística procede el vocablo “establo”, pues era función principal de los gobiernos establecer los registros de la población, nacimientos, defunciones, cosechas e impuestos y los fundamentos de la estadística actual y muchos de los métodos de inferencia son debidos a R. A. Fisher y las aplicaciones de la estadística en la economía tienen el nombre propio de econometría.

(3) Indicadores financieros

- Welsch et al. (2005) definen al **Valor Presente Neto** como un indicador que: “compara el valor actual de los flujos netos de entradas de efectivo con el valor del costo inicial, de un proyecto de desembolso de capital”.

(4) Software estadístico

- El análisis estadístico tiene muchas formas diferentes y cada tipo es apropiado o beneficioso para diferentes casos de uso. Aunque los métodos son muchos, el objetivo final del análisis se divide en tres tipos: descriptivo, prescriptivo y predictivo (Reyes et.al., 2018)

III. ANÁLISIS Y DIAGNÓSTICO DEL PROBLEMA

3.1. Análisis de actividades desempeñadas

3.1.1. Planteamiento del problema acerca de las funciones desempeñadas

Durante el desarrollo de mis funciones dentro del Banco de la Nación me encuentro con los siguientes problemas principales, que limitan la agilidad del proyecto y estas son:

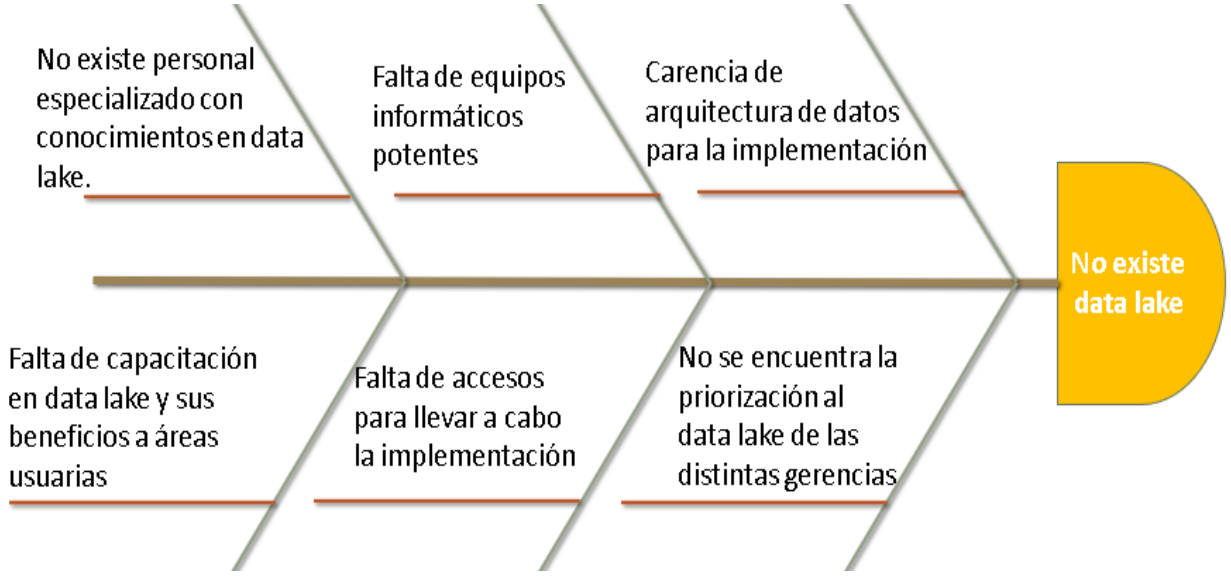
- No existe el **data lake**, es decir no tienen un repositorio general para almacenar los datos de distintas áreas y se considera como un problema, porque va ser causa del alargamiento de proyectos de ciencia de datos, por la carencia de información y no saber dónde

podemos conseguir la información, ya que para llevar a cabo los proyectos de ciencia de datos, necesitamos de datos históricos y datos maestros de nuestros clientes, trabajadores y cualquier tipo de información que nos aporte para las distintas tomas de decisiones. La solución a este problema es la implementación de la data lake.

3.1.2. **Diagnóstico del problema acerca de las funciones desempeñadas**

Durante el desarrollo de mis funciones en el banco, he detectado el problema de no contar con el data lake. El Banco de la Nación es una de muchas de las instituciones peruanas que recién están migrando a trabajar con datos y viene contratando profesionales especializados en datos. Lo que se quiere es trabajar con los datos de forma masiva y tener alcance y soporte a todas las áreas de forma automatizada y rápida, además tener la posibilidad de trabajar con grandes volúmenes de datos. ¿Cómo llegamos a detectar estos problemas? Cuando he trabajado el proyecto de localización, no se encontraron los datos de forma histórica y cada área maneja su información sin coordinación con otras áreas. Además, los reportes de finanzas no estaban automatizados por lo cual cada análisis demoraba mucho tiempo (meses). En la Figura 6, se muestra el diagrama de Ishikawa para el problema sobre el data lake.

Figura 6: *Diagrama de Ishikawa de data lake*



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

A continuación, se detallan las causas del problema identificado de la falta de la data lake:

- En el Banco de la Nación actualmente no se cuenta con un arquitecto de datos (profesional con conocimientos en data lake), por falta de conocimiento de la existencia de estos perfiles que en el Perú son muy escasos y sus salarios son muy elevados.
- Existe carencia de equipos informáticos potentes, tales como los servidores, buenas computadoras con memorias RAM potentes.
- Al no tener un profesional con el perfil de arquitecto de datos no se tiene un diagrama de la arquitectura de datos.
- La falta de capacitación y conocimiento de los beneficios de la data lake de parte de las áreas usuarias de datos, para acelerar la implementación de la data lake.
- La falta de accesos y restricciones que existen para conseguirlos, resulta de visión distinta que tienen los directivos, sobre este tema. Dejando de lado la transformación digital, por lo tanto, se queda estancado la idea de implementar un data lake.
- No se prioriza la implementación de la data lake, a pesar de que dependemos mucho del uso de los datos, porque la implementación de la data lake toma entre 2 a 3 años de implementación y el costo en los 2 años es elevado.

IV. RECOMENDACIONES

4.1. Recomendaciones de mejora de actividades

4.1.1. Fundamentación

(1) Para el Banco de la Nación

- Se recomienda contratar personas con el perfil de arquitecto de datos y apoyarlos para armar la arquitectura de los datos y acelerar el proyecto.

- Realizar una revisión de los equipos y repotenciarlos o sino poder reemplazarlos por equipos nuevos potentes que nos ayuden con el uso de grandes volúmenes de datos.
- Realizar capacitaciones de los beneficios que nos genera tener el data lake con los gerentes, subgerentes y jefes de sección para tener el apoyo en la facilitar los accesos.
- Crear directivas y trabajar en conjunto con los equipos que brinden los accesos y acelerar el proyecto.
- Capacitar a los gerentes respecto de los beneficios que genera el data lake, porque se debe priorizar el proyecto.

(2) Para la facultad de ciencia económicas de la Universidad Nacional Mayor de San Marcos (UNMSM[8])

- Mi recomendación para la facultad de Ciencias Económicas en la formación de economistas es optar por enseñar algún lenguaje de programación, en especial Python o R y manipulación de datos dentro de la malla curricular, como un curso aparte tal como se enseña informática para economistas y econometría con R o Python y manipulación de bases de datos con SQL, porque los modelos econométricos y Machine Learning se trabaja bastante con las computadores y lenguajes de programación, además en la actualidad ya no se soluciona un modelo econométrico con lapicero y una hoja sino de forma masiva con algún lenguaje de programación o software, porque la malla curricular de la facultad a mí me ayudó bastante, en especial los cursos de matemática, estadística, econometría, finanzas, economía, informática para economistas y evaluación de proyectos.

  El talón de Aquiles de un economista de la facultad y la malla curricular del plan de estudios 2015 para empezar a trabajar como practicante o científico de datos junior podría decir que es el desconocimiento del lenguaje de programación y manipulación de base de

[8] Universidad Nacional Mayor de San Marcos

datos. Pero, permite ser principiante en ciencia de datos con conocimientos de matemática, estadística, negocios y modelos econométricos, que son bases para entender los modelos de Machine Learning. En mi caso ha sido necesario llevar cursos de Python, Excel, SQL en instituciones externas y una especialización en ciencia de datos.

4.1.2. Medidas específicas

Con la fundamentación de las recomendaciones en la sección anterior, propondré algunas medidas específicas para el Banco de la Nación como propuesta de solución al problema identificado con anterioridad.

(1) Banco de la Nación

- En primera instancia, es tener las reuniones con los gerentes del área y tener el respaldo para desarrollar el proyecto y brindarles la información de lo valioso que es el data lake, para tener el apoyo con los accesos y darle prioridad al proyecto.
- Luego contratar todo un equipo de arquitecto de datos para crear la arquitectura de datos, la cual se va implementar. Para crear o diseñar la arquitectura de datos, lo que se va hacer es observar y preguntar a cada área donde tienen almacenado sus datos, en que carpeta lo guardan para crear el flujo de automatización con la creación de Extraer, Transformar y Cargar (ETL[9]) y por sus siglas en español significa Extraer los datos de esas carpetas en las cuales las áreas lo tienen, para luego hacer la transformación de los datos y cargarlos al repositorio único y así sucesivamente automatizar la carga para todas las áreas con frecuencia diaria.

  Finalmente, una vez ya automatizado para todas las áreas y tener todo tipo de información almacenada darles el acceso a cada uno de las áreas usuarias para que realicen las

[9] Es un tipo de integración de datos que hace referencia a los tres pasos (extraer, transformar, cargar) que se utilizan para mezclar datos de múltiples fuentes.

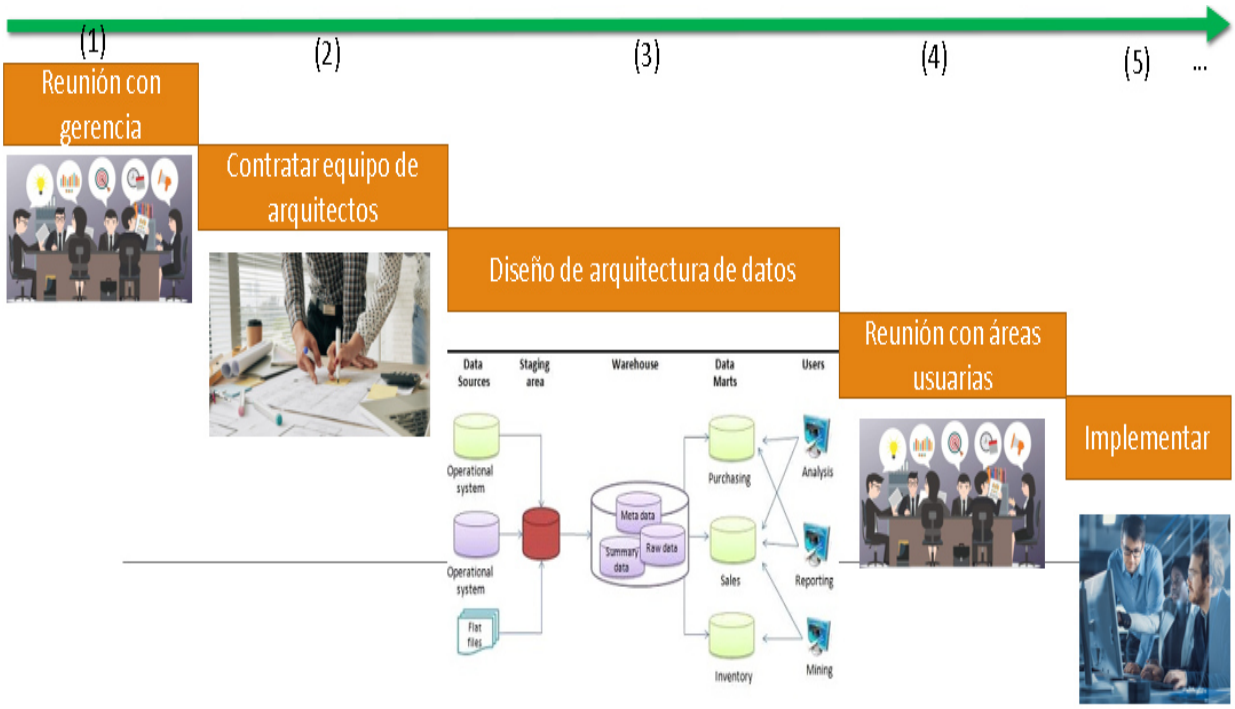
consultas necesarias de la información que requieran, tanto información histórica y no histórica. Para crear los ETLs se requiere saber distintos lenguajes de programación tales como SQL, Python u otro lenguaje de programación y visual studio y el trabajo duro y complicado se encuentra en esa parte y mi propuesta en esta parte del trabajo en primer lugar es crear job[10] en el SQL que ejecute los scripts creados en Python para obtener la información de alguna de las áreas, luego transformarlas y cargarlos al SQL en las madrugadas ya que es el horario en la cual los servidores no están en uso.

- Comprar o repotenciar equipos informáticos potentes.
- Capacitar a las áreas usuarias de datos para tener la facilidad y me puedan brindar los datos y permisos para almacenar la base datos de forma automática con los ETLs.

En la Figura 7 se puede ver mi propuesta.

[10] Paquete en SQL que nos ayuda a programar un horario para ejecutar algún código de SQL.

Figura 7: *Propuesta de implementación de data lake*



Nota: Elaboración del bachiller.

V. PERFORMANCE FUTURO DEL DESEMPEÑO PROFESIONAL

Para trabajar en ciencia de datos, se requiere tener dominio y conocimiento en estadística, matemática, lenguaje de programación y negocios. Partiendo de los temas que debe dominar un científico de datos, a continuación, se detalla la línea de carrera de un científico de datos dentro del Banco de la Nación.

(1) Practicante de ciencia de datos: Sus roles de un practicante netamente es apoyar en el desarrollo de proyectos de ciencia de datos a los científicos de datos senior e ir aprendiendo ya que se encuentra en la etapa de aprendizaje de los softwares y el desarrollo de algunas etapas del proyecto.

(2) Científico de datos junior: Al igual que el practicante un junior se encuentra en la etapa de aprendizaje y apoya en algunas partes del proyecto de ciencia de datos a los científicos de

datos senior, aunque un junior ya tiene conocimiento de algunos softwares de desarrollo de proyecto por la misma experiencia de haber hecho prácticas en otros trabajos.

(3) Científico de datos semi senior: Un profesional con este perfil ya tiene experiencia en proyectos de ciencia de datos y es capaz de desarrollar proyectos de ciencia de datos solo o con algunas pequeñas indicaciones de un científico de datos senior y líder de analítica avanzada y en mi caso me encuentro desempeñando este rol dentro del Banco de la Nación.

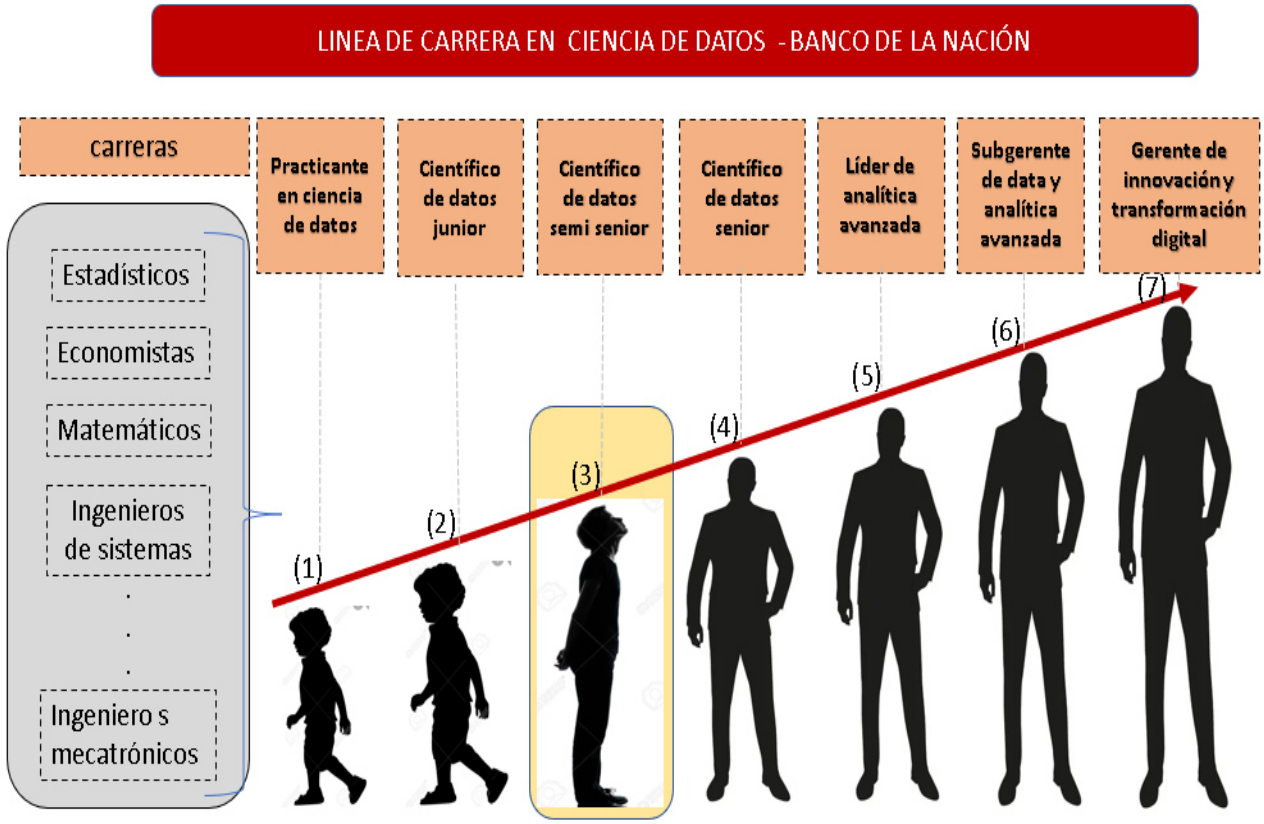
(4) Científico de datos senior: Es la persona con bastante conocimiento técnico y muchos proyectos de ciencia de datos desarrollados y tiene la capacidad de trabajar y liderar proyectos de ciencia de datos en equipo.

(5) Líder de analítica avanzada: Es la persona con bastante experiencia en liderar equipos de ciencia de datos y tener desarrollada las habilidades blandas, para un buen manejo de equipos de trabajo y llevar a cabo la culminación de los proyectos de ciencia de datos.

(6) Subgerente de data y analítica avanzada: Un profesional para asumir esta responsabilidad debe tener dominio técnico en ciencia de datos y desarrollado las habilidades blandas para liderar equipos y el rol es netamente gerencial.

(7) Gerencia de innovación y transformación digital: Un profesional para asumir esta responsabilidad, debe tener mucha experiencia liderando y dirigiendo equipos y algo de conocimiento técnico para entender a los científicos de datos cuando se realiza la presentación de los proyectos. Se enumera y detalla la línea de carrera dentro del Banco de la Nación y lo podemos ver en la Figura 8. Los requisitos para puestos de líder y las gerencias es ser bachiller y ser titulado de la maestría, en cambio para las posiciones de científicos de datos es ser bachiller de pregrado.

Figura 8: *Línea de carrera*



Nota: Elaboración del bachiller. Fuente: proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

REFERENCIAS BIBLIOGRÁFICAS

Galán Cortina, V. Castro Galán, E. (2015). Aplicación de la metodología CRISP-DM a un proyecto de minería de datos en el entorno universitario. proyecto fin de carrera. Universidad Carlos III de Madrid. España. Obtenido de: https://e-archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf

Montero Granados. R (2016). Modelos de regresión lineal múltiple. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España. Obtenido de: https://www.ugr.es/~montero/matematicas/regresion_lineal.pdf

Muñoz Diaz, S. (2013). Análisis geoespacial para ubicar cajeros automáticos y oficinas bancarias como herramienta para mejorar la cobertura de servicios de las entidades financieras en Bogotá. Universidad Militar Nueva Granada, Bogotá, Colombia. Obtenido de: https://repository.unimilitar.edu.co/bitstream/handle/10654/10914/Bibiana%20Mu%C3%B1oz%20Trabajo%20Final.pdf?sequence=1&isAllowed=y

Novales Cinca, A. (1993). Econometría (2da edición). Editora Isabel Capella. Obtenido de: https://econometriai.files.wordpress.com/2011/04/econometrc2a1a-2c2a6-ed-2000-alfonso-novales-mcgraw-hill.pdf

Portillo, F. (2006). Introducción a la econometría. Universidad de la Rioja. Obtenido de: https://www.unirioja.es/cu/faporti/ieTEMA01.pdf

Ramírez Carrillo, C. (2016). Análisis y comparación de proyectos de inversión mediante el método de valor actual neto. Universidad Técnica de Machala. Obtenido de:http://repositorio.utmachala.edu.ec/bitstream/48000/8816/1/ECUACE-2016-EC-CD00012.pdf

Ramirez C. García M. Pantoja C. Zambrano U. (2009). Fundamentos de matemáticas financieras. Universidad Libre. Colombia. Obtenido de: https://www.uv.mx/personal/cbustamante/files/2011/06/MATEMATICAS_FINANCIERAS.p

df

Reyes Robles, Gaona Montiel, Méndez Pérez, (2018). Uso de softwares estadísticos/econométricos, como herramientas en la investigación económica y administrativa. Universidad Autónoma Metropolitana. México.

Rivas Martínez, J. (1999). Estudio para la ubicación óptima de cajeros automáticos en la ciudad de Guayaquil (Tesis de grado). Escuela superior Politécnica del litoral, Guayaquil, Ecuador. Obtenido de: http://www.dspace.espol.edu.ec/bitstream/123456789/2091/1/4180.pdf

Rodríguez Sánchez, A. García García, C. The Coefficient of Determination in the Ridge Regression. Universidad de Granada. España. Obtenido de: file:///C:/Users/TOSHIBA/Downloads/Documat-TheCoefficientOfDeterminationInTheRidgeRegression-6641075%20(2).pdf

Salazar, C. Del Castillo, S. (2018). Fundamentos Básicos de estadística (1ra edición). Obtenido de: http://www.dspace.uce.edu.ec/bitstream/25000/13720/3/Fundamentos%20B%C3%A1sicos%20de%20Estad%C3%ADstica-Libro.pdf

Sulca G. Becerra E. Espinoza V. (2016). Análisis financiero. Universidad Central de Ecuador. Ecuador. Obtenido de : http://www.dspace.uce.edu.ec/bitstream/25000/21004/1/Analisis%20financiero.pdf

T. Sutton, O. Dassau, M. Sutton (2009). Documentación de QGIS 2.14. Obtenido de: https://docs.qgis.org/2.14/es/docs/gentle_gis_introduction/preamble.html

Villazán Olivares, F. (2010). Informática I. Universidad Michoacana de San Nicolás de Hidalgo. Obtenido de: https://www.upg.mx/wp-content/uploads/2015/10/LIBRO-31-Manual-de-Informatica.pdf

Vivallo P. Formulación y evaluación de proyectos. Obtenido de: https://economicas.unsa.edu.ar/afinan/informacion_general/book/ebooks/manual-evalua-proy.pdf

WEBGRAFÍA

https://datos.bancomundial.org/indicator/FB.ATM.TOTL.P5?locations=PE

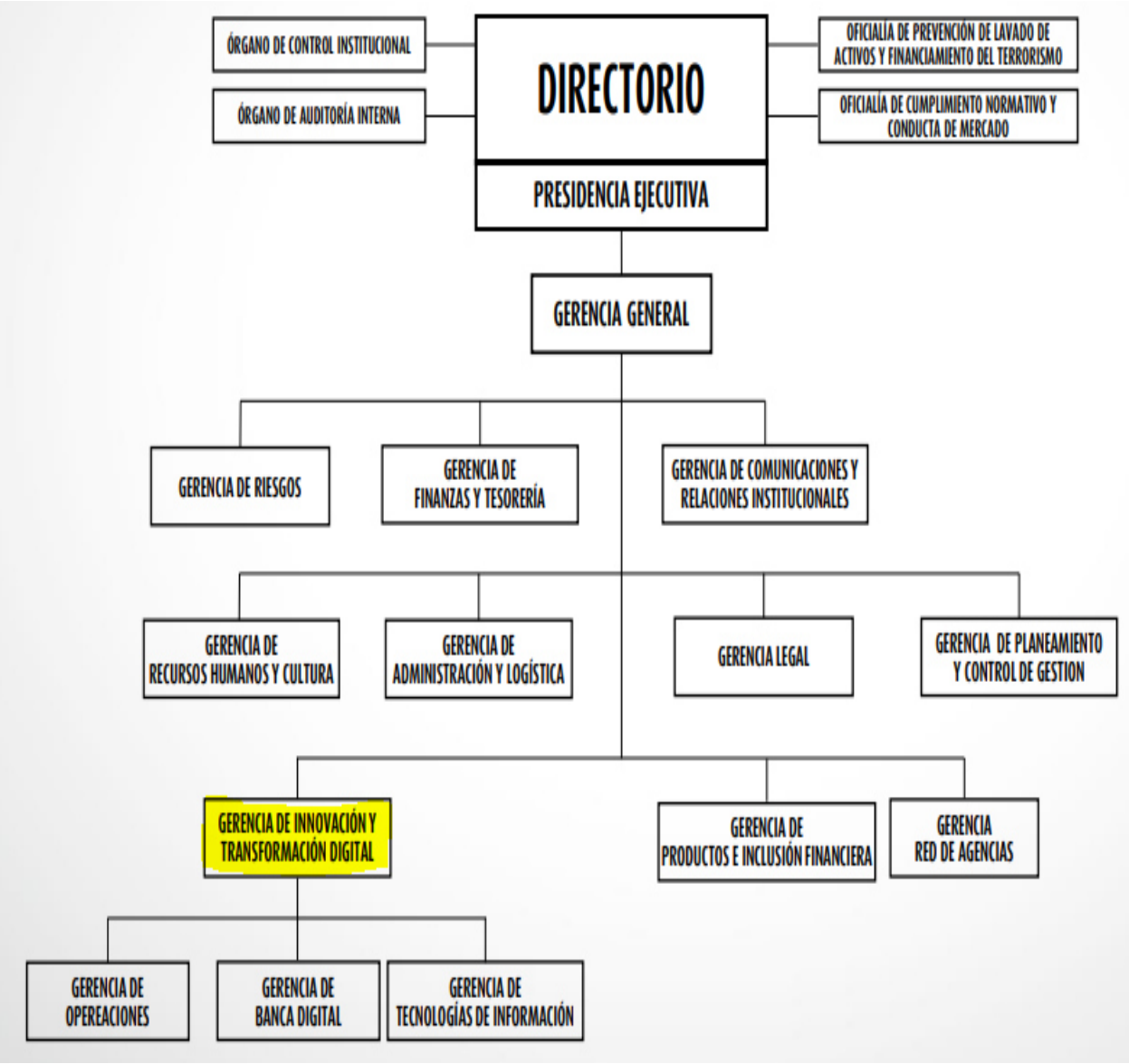
https://revistas.tec.ac.cr/index.php/eagronegocios/article/view/3687/3640

https://www.ucm.es/data/cont/docs/518-2013-10-25-EctrGr-JAM-2.pdf

https://economia.unmsm.edu.pe/escuela/plan_Eco_2015_.pdf

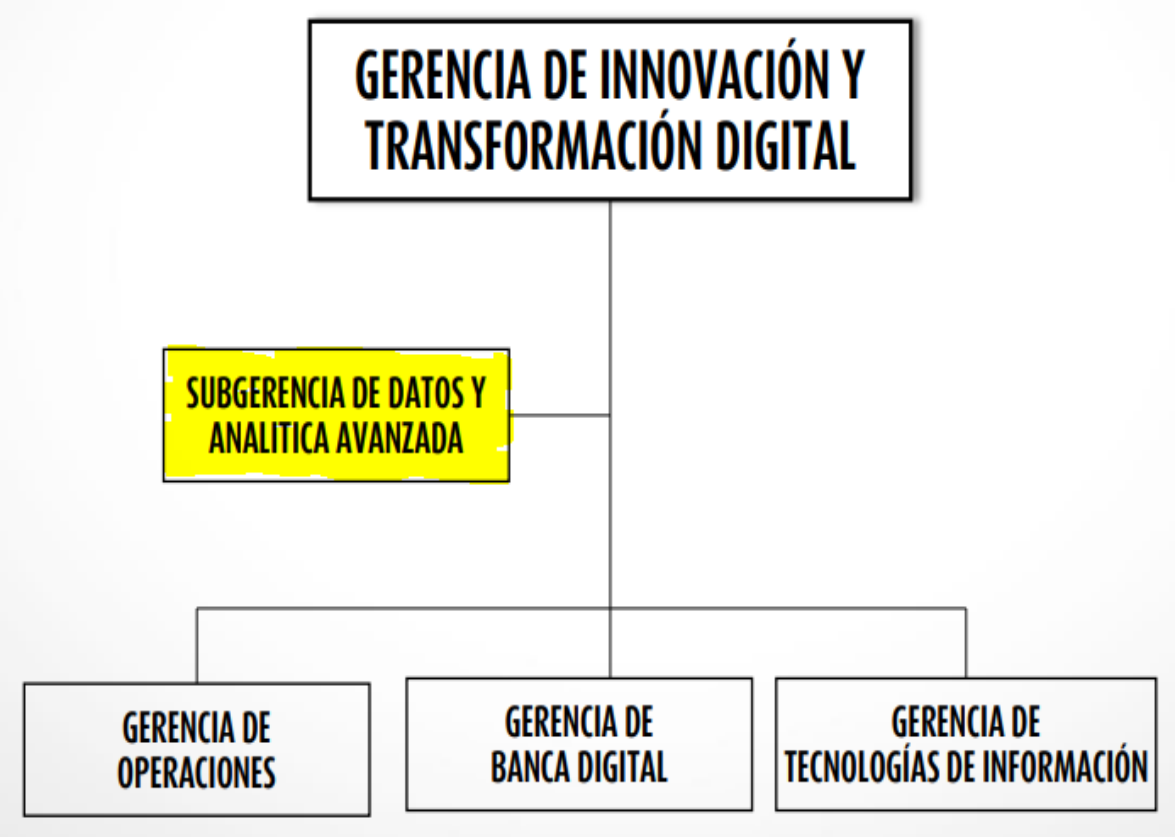
ANEXOS

Anexo 1: Organigrama general de Banco de la Nación



Fuente: Banco de la Nación 2022.

Anexo 2: Organigrama específica del Banco de la Nación



Fuente: Banco de la Nación 2022.

Anexo 3: Contrato de locador de servicios

Banco de la Nación

CO - 28139 - 21-BN

ANEXO N° 001 – CONDICIONES PARTICULARES
CONTRATO DE LOCACIÓN DE SERVICIOS

1. **IDENTIFICACIÓN DE EL LOCADOR**
   - **Nombres : ANGEL JORGE SALAZAR**
   - **DNI N° : 71552326**
   - **Domicilio : Mz O7 Lt21 Sector 1ro de Mayo - AH Pamplona Alta - San Juan de Miraflores**
   - **RUC : 10715523260**

2. **DETALLE DE EL SERVICIO:**
   - **Generación y desarrollo de modelos analíticos de optimización de "Gestiónde Efectivo" en ATM y agencias como Científico de Datos**

3. **PLAZO DE EL SERVICIO:**

   - 5 meses, desde el 2 de diciembre de 2021 y hasta el 1 de mayo de 2022 inclusive.

4. **ACTIVIDADES A REALIZAR**

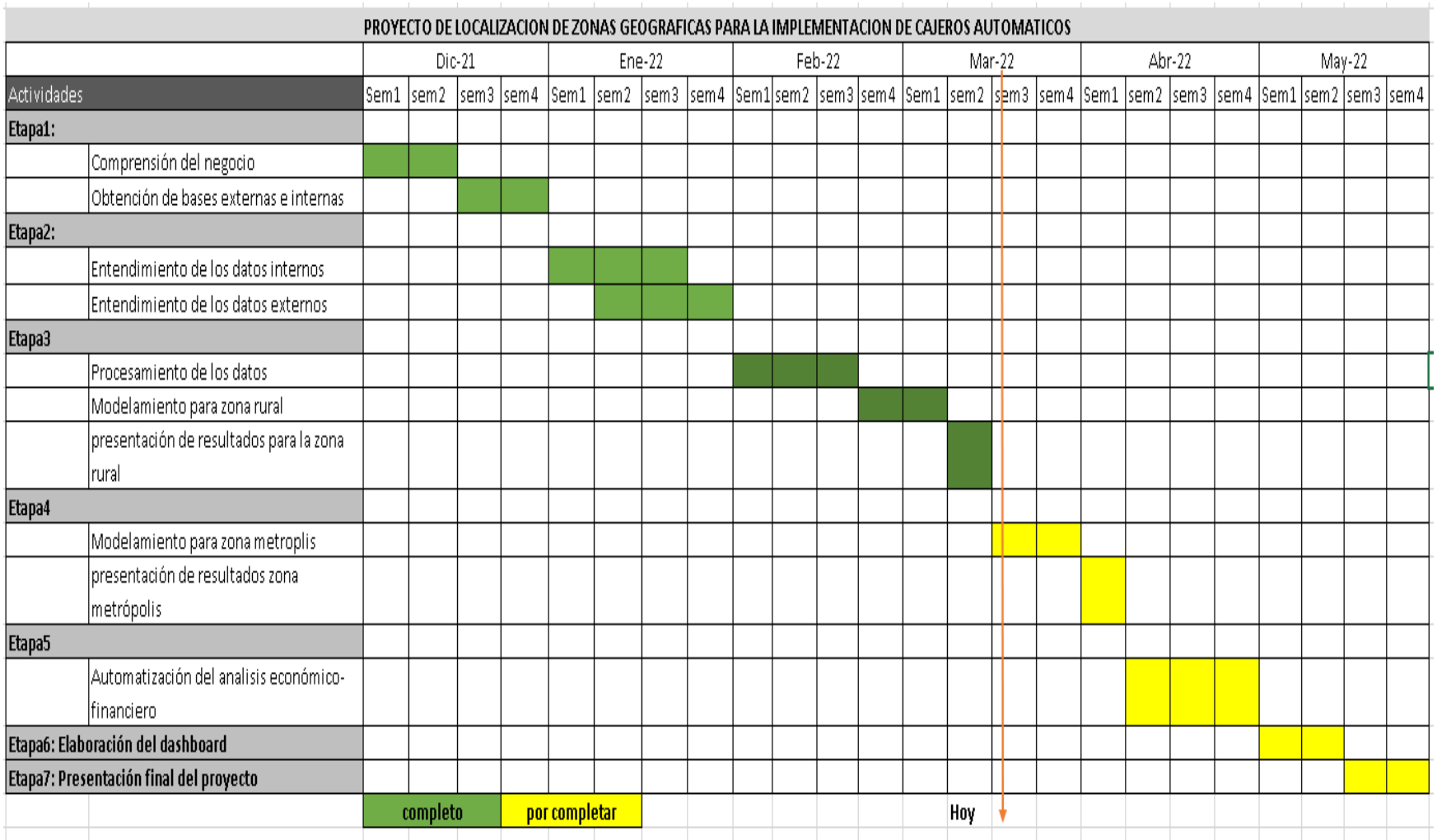
   1 Identificar y proponer de manera proactiva oportunidades de construcción de modelos de machine learning que contribuirán a la optimización de "Gestión de Efectivo" en ATM y agencias.

   2 Obtener la data de los servidores del banco y realizar traslados de información al entorno cluster de Advanced Analytics para el desarrollo de los modelos de optimización de "Gestión de Efectivo" en ATM y agencias

   3 Definir el objetivo, las variables involucradas y reportar los indicadores de calidad del modelo analítico definido los mismos que seran documentados de forma técnica (en relación al item anterior)

   4 Hacer seguimiento a los indicadores de los modelos generados para la optimización de "Gestión de Efectivo" en ATM y agencias, así como también definir la vigencia de los modelos y la necesidad de recalibración o update.

Anexo 4: Gantt del proyecto



Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 5: Variables utilizadas en la zona rural

VARIABLE	DESCRIPCIÓN
CTD_ATM	Cantidad de cajeros automáticos en el año 2019-2022
TRX_ATM_PROM	Suma del promedio mensual de transacciones en cajeros automáticos en el año 2019-2022
CTD_AGENTE	Cantidad de agentes al año 2022
CTD_AGENCIAS	Cantidad de agencias al año 2022
PEA	Población Económicamente Activa
COBERTURA	Distrito con cobertura de comunicaciones
BENEFICIARIO_BONO	N° de beneficiarios del bono Yanapay
PPSS	N° personas de programas sociales
PRESENCIA_BN y BP	1 si BN y BP, 0 en SOLO BN o SIN CANAL
POBLACION	Población

Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 6: Variables utilizadas en la zona metrópolis

Variable	Descripción
Comercio	Nombre de comercio
Tipo de comercio	Tipos de comercio
Latitud	Coordenada geográfica del comercio
Longitud	Coordenada geográfica del comercio

Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 7: Variables utilizadas para análisis financiero-económico

Variable	Descripción
# DE ATM	Id de cada ATM operativo
LATITUD	Coordenada geográfica
LONGITUD	Coordenada geográfica
CANTIDAD DE TRANSACCIONES DE RETIROS MENSUAL	Cantidad de transacciones de retiros mensual
OTRAS VARIABLES DE COSTOS	Costos que implica implementar ATMs
OTRAS VARIABLES	Otras variables implicadas en el análisis económico financiero

Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 8: Ejemplo de buffer



Ejemplo: *En el distrito de Yungay se encuentra un área donde se concentran 82 comercios, 1 agencia, 4 agentes, 1 cajero automático con alta demanda. La prioridad de la implementación es 2 por presentar un cajero con alta demanda.*

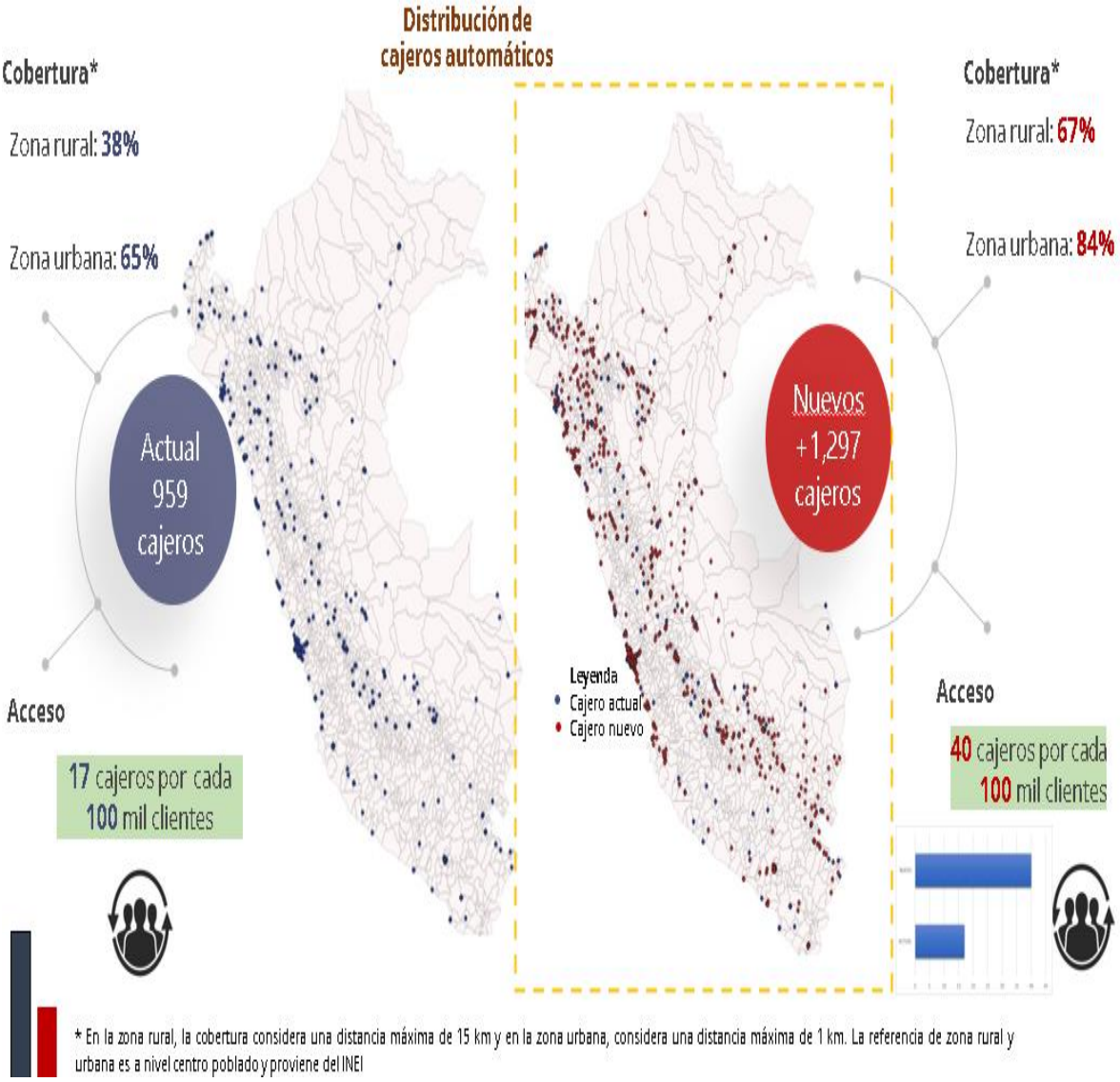
Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 9: Reporte de zonas geográficas para la implementación de cajeros automáticos

UBIGEO	DEPARTAMENTO	PROVINCIA	DISTRITO	GEOMETRY	LATITUD	LONGITUD
200703	PIURA	TALARA	LA BREA	Punto((-81.30587	-4.6546397	-81.305879
200702	PIURA	TALARA	EL ALTO	Punto((-81.22125	-4.268616	-81.2212978
200705	PIURA	TALARA	LOS ORGANOS	Punto((-81.12941	-4.17925108	-81.1294131
200504	PIURA	PAITA	COLAN	Punto((-81.05643	-4.90065268	-81.0564307
200505	PIURA	PAITA	LA HUACA	Punto((-80.96133	-4.91016196	-80.9613328
200505	PIURA	PAITA	LA HUACA	Punto((-80.91920	-4.92044443	-80.9192038
200603	PIURA	SULLANA	IGNACIO ESCUDE	Punto((-80.87304	-4.84604358	-80.8730481
200801	PIURA	SECHURA	SECHURA	Punto((-80.82223	-5.55723442	-80.8222356
200805	PIURA	SECHURA	VICE	Punto((-80.80878	-5.45296354	-80.8087831
200805	PIURA	SECHURA	VICE	Punto((-80.80693	-5.43970842	-80.8069334
200805	PIURA	SECHURA	VICE	Punto((-80.77625	-5.42219154	-80.776253
200605	PIURA	SULLANA	MARCAVELICA	Punto((-80.76185	-4.85826786	-80.7618575
200605	PIURA	SULLANA	MARCAVELICA	Punto((-80.71789	-4.86740315	-80.7178981
200109	PIURA	PIURA	LA ARENA	Punto((-80.70359	-5.34296403	-80.7035994
200107	PIURA	PIURA	CURA MORI	Punto((-80.66543	-5.32357184	-80.6654353
200607	PIURA	SULLANA	QUERECOTILLO	Punto((-80.64820	-4.83927391	-80.6482031
240103	TUMBES	TUMBES	LA CRUZ	Punto((-80.59010	-3.63718133	-80.5901042
200604	PIURA	SULLANA	LANCONES	Punto((-80.50789	-4.63201997	-80.5079
240102	TUMBES	TUMBES	CORRALES	Punto((-80.48058	-3.60151693	-80.4805847
240105	TUMBES	TUMBES	SAN JACINTO	Punto((-80.44519	-3.64096023	-80.445195
200404	PIURA	MORROPON	LA MATANZA	Punto((-80.09044	-5.21367037	-80.0904458
140306	LAMBAYEQUE	LAMBAYEQUE	MORROPE	Punto((-80.01564	-6.54025371	-80.0156426

Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 10: Impacto en cobertura



Nota: Con la lista de 1399 cajeros mejoramos en cobertura a nivel nacional. En zona rural pasamos de tener una cobertura de 38% a 65% y en la zona urbana pasamos de 67% a 84% de cobertura.

Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022.

Anexo 11: Ficha de iniciación del proyecto

FICHA TÉCNICA PROYECTOS DE ANALÍTICA AVANZADA BN			
Subgerencia(s) usuaria(s)	Gerenta de Planeamiento y Control de Gestión	Fecha	02/12/2021

A. PARTICIPANTES	
Product owner	Katerine Andrea Matos Vega
Área usuario o negocio	Gerencia de Planeamiento y control de gestión
Jefe del negocio	Helen Estefany Garrido Maldonado
Equipo de negocio	Sofía Mejía Carranza y Amós Canchari Soto
Equipo Data Scientist	Eliza Lastra Medina y Angel Jorge Salazar
Jefe Sección Analítica Avanzada	
Subgerente de Datos y Analitica Avanzada	

¿No vemos nada de crédito de última milla hasta cerrar agente comunidad?

B. IDENTIFICACIÓN	
Nombre del proyecto	**Estimación de nuevos cajeros automáticos**
Descripción del proyecto	El área de planeamiento del banco de la nación como una parte de sus roles tiene la necesidad de implementar nuevos ATMs por ello requieren saber cuánto es la cantidad óptima de ATMs que se requiere implementar para cerrar la brecha entre la oferta y la demanda y ver la rentabilidad financiera-económico de cada nuevo ATM a implementar, además requieren saber dónde ubicarlos.
Objetivo	Dado el problema, el objetivo es la determinación de la cantidad de cajeros automáticos que se requiere para reducir la brecha entre la oferta y la demanda, el análisis de rentabilidad financiero-económico y la ubicación geográfica.
Objetivos específicos	1. Identificar población objetivo / perfil
	2. Estimar cantidad de transacciones potenciales por distrito
	3. Determinar cantidad de cajeros automaticos por centro poblado y en ciudades metropolis y no metropolis

Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022, pág. 1

Anexo 12: Ficha de finalización del proyecto

Subgerencia Datos y Analítica Avanzada Sección: Sección Analítica y Avanzada	ACTA DE CONFORMIDAD DEL MODELO ANALITICO

1. INFORMACION :	
1.1. Proyecto	Estimación de nuevos cajeros automáticos
1.2 Objetivo	Es la **determinación de la cantidad de cajeros automáticos** que se requiere para reducir la brecha entre la oferta y la demanda, el análisis de **rentabilidad financiero-económico** y la **ubicación geográfica**
1.3 Fecha	18/05/2022
1.4 Propietario	Eliza Lastra Medina y Angel Jorge Salazar

2. USUARIOS				
Nombre y Apellido	**Entidad**	**Cargo**	**Área**	**Firma**
Katerine Andrea Matos Vega	BN	Gerenta	Planeamiento y Control de Gestión	
Helen Estefany Garrido Maldonado	BN	Subgerenta encargada	Planeamiento y Control de Gestión	
Sofía Mejía Carranza y Amós Canchari Soto	BN	Analista	Planeamiento y Control de Gestión	

3. DESARROLLO DEL MODELO ANALITICO

1. Metodología zona rural

Para obtener la cantidad de puntos geográficos, en primer lugar, se aprendió de distritos que tienen cajeros automáticos y proyectar las transacciones potenciales en distritos que no cuentan con cajeros automáticos, luego si el distrito supera los 5mil transacciones potenciales por mes y el nivel de PEA es mayor a 1500, se considera como un punto geográfico donde implementar un cajero automático, finalmente por importancia de centros poblados se determina la ubicación a nivel centro poblado

2. Metodología zona metrópolis y ciudades

Para esta metodología se trabaja con coordenadas geográficas de comercios obtenidos por web scraping de Google Maps. Esta metodología consiste en la agrupación de comercios con una cercanía entre cada comercio de aproximadamente 10 metros de distancia para encontrar zonas comerciales.

Nota: Elaboración del bachiller. Fuente: Proyecto de localización de zonas geográficas para la implementación de cajeros automáticos del Banco de la Nación 2022, pág. 1.

GLOSARIO

Machine Learning: es una forma de la IA que permite a un sistema aprender de los datos en lugar de aprender mediante la programación explícita.

Web scraping: Hacer scrapeo es el proceso de extraer datos de sitios web(freecodcamp.org).

CRISP-DM: que son las siglas de Cross-Industry Standard Process for Data Mining, es un método probado para orientar sus trabajos de minería de datos.

Minería de datos: La minería de datos es el proceso de hallar anomalías, patrones y correlaciones en grandes conjuntos de datos para predecir resultados. Empleando una amplia variedad de técnicas, puede utilizar esta información para incrementar sus ingresos, recortar costos, mejorar sus relaciones con clientes, reducir riesgos y más.

ATMs: Cajero automático (en inglés, Automated Teller Machine), máquina expendedora de dinero.

Data science: o ciencia de datos es un enfoque multidisciplinario para extraer insights accionables de los grandes y cada vez mayores cantidades de datos recopilados y creados por las organizaciones actuales.

Regresión lineal: Un modelo de regresión lineal múltiple es un modelo estadístico versátil para evaluar las relaciones entre un destino continuo y los predictores.

Diagrama de Gantt: es una herramienta útil para planificar proyectos. Al proporcionarte una vista

general de las tareas programadas, todas las partes implicadas sabrán qué tareas tienen que completarse y en qué fecha.

Análisis Espacial Vectorial (Buffers), una zona buffer es cualquier área que sirve al propósito de mantener las entidades del mundo real distantes entre ellas.

La gestión de efectivo, conocida también como cash management, es el conjunto de estrategias que busca optimizar la gestión de los recursos financieros de una empresa. Es decir, permite planificar y controlar la liquidez, administrar y gestionar los saldos y mejorar el flujo monetario de la organización.

Recalibración, ajustar, con la mayor exactitud posible, las indicaciones de un instrumento de medida con respecto a un patrón de referencia.

Update proviene de la lengua inglesa que significa “actualizar”, “modernizar.

Analítica avanzada (advanced analytics) emplea modelos predictivos, métodos estadísticos, Machine Learning y técnicas de automatización de procesos más allá de las capacidades de las herramientas tradicionales de inteligencia de negocio (BI) para analizar datos o información empresarial.