



**UNIVERSIDAD CÉSAR VALLEJO**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA  
ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA DE  
SISTEMAS**

“Machine Learning para la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú”

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:**

Ingeniero de Sistemas

**AUTOR:**

Br. Anthony Jean Claude Simon Altamirano (ORCID: 0000-0001-9880-5795)

Br. Julinho Raúl Ostos Mera (ORCID: 0000-0001-6571-1610)

**ASESOR:**

Mg. Orleans Galvez Tapia (ORCID: 0000-0002-4352-9495)

**LÍNEA DE INVESTIGACIÓN:**

Sistema de Información y Comunicaciones

**LIMA – PERÚ**

**2019**

## **DEDICATORIA**

La presente investigación la dedicamos a todos nuestros familiares que se preocuparon por nosotros y apoyaron para poder culminarlo, principalmente a las personas que estuvieron desde el inicio y fin dado que la educación y valores que nos inculcaron son el cimiento para la formación de nuestra vida profesional además de forjar las bases de nuestro deseo de superación. Gracias a Dios por otorgarme la mejor familia.

## **AGRADECIMIENTO**

Gracias a Dios por permitirnos vivir cada día, a nuestras familias por creer en nosotros y brindarnos sus consejos además de su apoyo en cada decisión que tomamos. Este largo trayecto no ha sido sencillo hasta ahora, pero gracias a su apoyo, su amor y su inmensa bondad, motivaron a continuar lograr esta meta. Les agradezco, y hago presente mi gran afecto hacia ustedes, mi hermosa familia.

## PÁGINA DEL JURADO

### DECLARATORIA DE AUTENTICIDAD

Yo, Anthony Jean Claude Simon Altamirano, estudiante de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Cesar Vallejo, Sede Lima Norte; presento el trabajo académico titulado "MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS EN LA EMPRESA KONECTA BTO, S.L. SUCURSAL EN PERÚ" Para la obtención del título Profesional de Ingeniería de Sistemas.

Por lo tanto, declaro lo siguiente:

He mencionado todas las fuentes empleadas en el presente trabajo de investigación, identificando correctamente toda la cita textual o de paráfrasis proveniente de otras fuentes, de acuerdo con lo establecido por las normas de elaboración de trabajos académicos.

Este trabajo de investigación no ha sido previamente presentado completa ni parcialmente para la obtención de otro grado académico o título profesional.

Soy consciente de que mi trabajo puede ser revisado electrónicamente en búsquedas de plagios.

De encontrar uso de material intelectual ajeno sin el debido reconocimiento de su fuente o autor, me someto a las sanciones que determinan el procedimiento disciplinario.

Lima, 16 de Julio del 2019



Simon Altamirano Anthony Jean Claude

70405567

#### DECLARATORIA DE AUTENTICIDAD

Yo, Julinho Raul Ostos Mera, estudiante de la Escuela Profesional de Ingeniería de Sistemas de la Universidad Cesar Vallejo, Sede Lima Norte; presento el trabajo académico titulado "MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS EN LA EMPRESA KONECTA BTO, S.L. SUCURSAL EN PERÚ" Para la obtención del título Profesional de Ingeniería de Sistemas.

Por lo tanto, declaro lo siguiente:

He mencionado todas las fuentes empleadas en el presente trabajo de investigación, identificando correctamente toda la cita textual o de paráfrasis proveniente de otras fuentes, de acuerdo con lo establecido por las normas de elaboración de trabajos académicos.

Este trabajo de investigación no ha sido previamente presentado completa ni parcialmente para la obtención de otro grado académico o título profesional.

Soy consciente de que mi trabajo puede ser revisado electrónicamente en búsquedas de plagios.

De encontrar uso de material intelectual ajeno sin el debido reconocimiento de su fuente o autor, me someto a las sanciones que determinan el procedimiento disciplinario.

Lima, 16 de Julio del 2019



Julinho Raul Ostos Mera

76278255

## ÍNDICE

	Página
Dedicatoria	ii
Agradecimiento	iii
Página del jurado	iv
Declaratoria de autenticidad	v
Índice	vii
Resumen	xiii
Abstract	xiv
<b>I. INTRODUCCIÓN</b>	
1.1. Realidad Problemática	1
1.2. Trabajos Previos	4
1.3. Teorías relacionadas al Tema	11
Planificación	22
Machine Learning	14
Metodología de Machine Learning	14
SEMMA	27
KDD	
CRISP DM	28
1.4. Formulación del Problema	21
Problema General	21
Problema Específicos	21
1.5. Justificación del Estudio	21
Justificación Tecnológica	21
Justificación Económica	22
Justificación Institucional	22
Justificación Operativa	22

1.6. Hipótesis	23
Hipótesis General	23
Hipótesis Específicos	23
1.7. Objetivos	23
Objetivo General	24
Objetivo Específicos	24
<b>II. MÉTODO</b>	
2.1. Diseño de Investigación	25
Método de Investigación	25
Hipotético Deductivo	25
Tipo de Investigación	25
Explicativa	25
Experimental	25
Aplicada	26
Diseño de Investigación	26
Pre-Experimental	26
2.2. Variables y Operacionalización	27
Definición Conceptual	27
Definición Operacional	27
Operacionalización de Variables	28
2.3. Población y Muestra	30
2.4. Técnicas e Instrumentos de Recolección de datos, validez y confiabilidad	31
2.5. Métodos de Análisis de Datos	31
2.6. Aspectos éticos	32

III. RESULTADOS	36
IV. DISCUSIÓN	49
V. CONCLUSIONES	51
VI. RECOMENDACIONES	53
VII. REFERENCIAS	55

## ANEXOS

Anexo 1: Matriz de consistencia

Anexo 2: Diagrama de proceso

Anexo 4: Árboles de problemas

Anexo 5: Tabla de Evaluación de Experto

Anexo 6: Tabla de Evaluación de Indicadores

Anexo 7: Ficha de Registro – Indicador Eficiencia Pres test

Anexo 8: Ficha de Registro – Indicador Productividad Pres test

Anexo 9: Ficha de Registro – Indicador Eficiencia Post test

Anexo 10: Ficha de Registro – Indicador Productividad Post test

Anexo 11: Cronograma del plan

Anexo 12 : Modelo Lógico

Anexo 13. Entrevista a Gerente General

## ÍNDICE DE TABLAS

Tabla 1: Selección de metodología de software

Tabla 2: Resultado de expertos

Tabla 3. Operacionalización de variables

Tabla 4 Indicadores de la planificación

Tabla 5. Validación de indicador Productividad

Tabla 6. Validación de indicador Eficiencia

Tabla 7: Medidas descriptivas de la Eficiencia en el proceso antes y después de implementar Machine Learning

Tabla 8: Porcentaje de la Eficiencia en el proceso antes y después de implementar Machine Learning

Tabla 9: Los resultados descriptivos de Productividad

Tabla 10: porcentaje de productividad

Tabla 11: Prueba normalidad eficiencia

Tabla 12: Eficiencia pretest

Tabla 13: Eficiencia post test

Tabla 14: Prueba de normalidad Productividad

Tabla 15: Productividad pretest

Tabla 16: Productividad pos test

Tabla 17: T student eficiencia

Tabla 18: T student Productividad

Tabla 19: Descripciones de atributos

Tabla 20: Cantidad de registro de muestra

## ÍNDICE DE FIGURAS

Figura 1: Productividad

Figura 2: Eficiencia

Figura 3: Planificación

Figura 4: Fases de la metodología SEMMA

Figura 5: Procedimientos de la metodología KDD

Figura 6: Fases de la metodología Crisp DM

Figura 7: Metodología Crisp DM

Figura 8: Niveles de confiabilidad

Figura 9: Indicador eficiencia Pretest

Figura 10: Indicador productividad pretest

Figura 11: Web scraping

Figura 15: Visualización de DATAFRAME

Figura 16: Análisis de datos

Figura 17: Reducción de dimensiones aplicando PCA

Figura 18 Esquema de una Neurona Artificial

Figura 19: Arquitectura de red Neuronal

Figura 20: Regresión Lineal entorno Jupyter

Figura 21: Ejecución del Score

Figura 22: Red neuronales en el entorno Spyder

Figura 23: Script Spyder

Figura 24: Inserción de datos

Figura 25: Procedimiento Almacenado

Figura 26: Inserción de nuevo JOB

## RESUMEN

La presente tesis detalla el desarrollo de Machine Learning para la planificación en el servicio de atención de llamadas en la empresa Konecta Bto, Sucursal en Perú, debido a que la situación empresarial previa a la aplicación del sistema presentaba deficiencias en cuanto al proceso de planificación. El objetivo de esta investigación fue determinar la influencia del Machine Learning en la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

Por ello, se describe previamente aspectos teóricos de lo que es el proceso de planificación, así como las metodologías que se utilizaron para el desarrollo del Machine Learning. Para el desarrollo, se empleó la metodología CRISP DM, por ser la que más se acomodaba a las necesidades y etapas del proyecto, además, se utilizó el lenguaje de programación Python y el gestor de base de datos SQL Server 2017.

El tipo de investigación es, explicativa, aplicada y experimental, el diseño de la investigación es pre-experimental y el enfoque es cuantitativo. La técnica de recolección de datos fue el fichaje y el instrumento fue la ficha de registro, los cuales fueron validados por expertos.

Se concluye que el Machine Learning influyo positivamente en la planificación en el servicio de atención de llamadas en la empresa Konecta Bto, SL Sucursal en Perú.

**Palabras clave:** Aprendizaje Automático, CRISP DM, Python

## ABSTRACT

This thesis details the development of Machine Learning for planning in the call answering service in the company Konecta Bto, Branch in Peru, due to the fact that the business situation prior to the application of the system presented deficiencies in the planning process. The objective of this research was to determine the influence of Machine Learning in Planning in the call answering service in the company Konecta BTO, S.L. Branch in Peru. Therefore, theoretical aspects of what the planning process is, as well as the methodologies that were used for the development of Machine Learning, are previously described. For the development, the CRISP DM methodology was used, as it was the one that best suited the needs and stages of the project, in addition, the Python programming language and the SQL Server 2017 database manager were used. The type of research is explanatory, applied and experimental, the research design is pre-experimental and the approach is quantitative. The data collection technique was the registration and the instrument was the registration form, which were validated by experts. It is concluded that Machine Learning positively influenced the planning of the call answering service in the company Konecta Bto, SL Branch in Peru.

**Keywords:** Machine Learning, CRISP DM, python

# I. INTRODUCCIÓN

## 1.1 Realidad problemática

En el plano internacional Según Galarza López (2016), en la revista el proceso de planificación estratégica manifiesta que: “Hoy en día la planificación en el mundo no se le está dedicando un tiempo determinado que permita analizar un conjunto de aspectos necesarios que poseen gran importancia para el crecimiento organizacional como la misión, visión, metas. Suele acelerar el procedimiento debido a la falta de tiempo, se aceleran algunas etapas y se termina con un estado anormal de la costumbre en pequeñas ocasiones y sin utilizar alguna tecnología básica. Esto hace que haya una dimensión segura de participación en las recomendaciones subyacentes de la planificación, afectando negativamente la ejecución y resultados que se obtienen posteriormente” (p.89).

En el plano nacional, según José María (2015), en la revista gestión nos indica que: “Perú está muy por debajo del estándar de tiempo que los gerentes de las organizaciones dedican a los temas de planificación para generar eficiencia en las áreas respectivas. Un gerente debería dedicarle un tiempo adecuado a la planificación en general. Si bien las organizaciones peruanas más conocidas si procuran que sus gerentes le den mayor tiempo, se dice que, a través del coaching, las tareas diarias les demanda mayor tiempo” (pag.1).

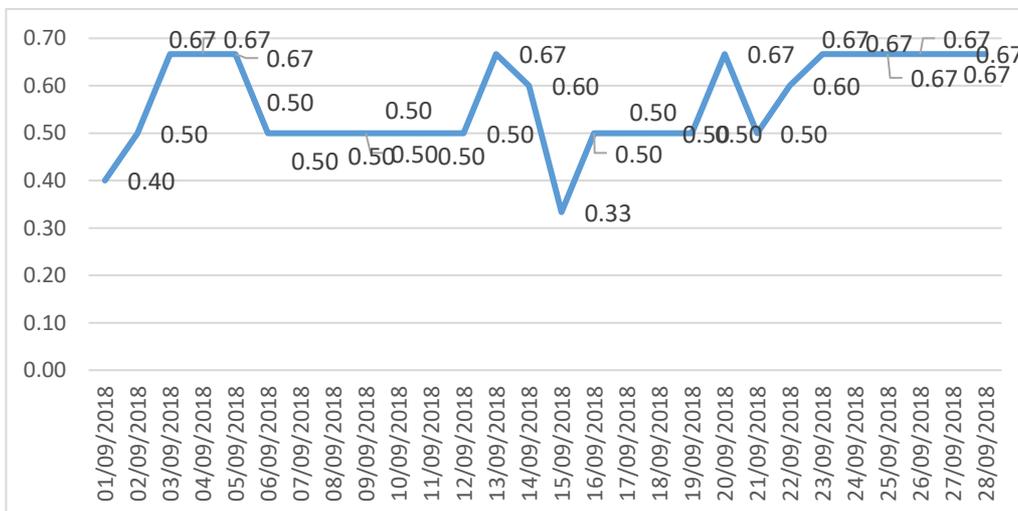
Por otro lado, en el ámbito nacional según Prieto Pulido (2017), en la revista Espacios nos indica que: “la planificación es un proceso que orienta, que proporciona las reglas para caracterizar los destinos de la asociación, los activos que se utilizarán y la cantidad de individuos que administrarán la organización de dichos activos. De la misma manera, las técnicas para conectarse con el personal afectarán su conducta dentro de la asociación y de esta manera en su intensidad de trabajo. Por lo tanto, en las organizaciones, los pasos deben hacerse con menos recursos y mejores resultados, lo que incita al personal a ser trabajadores ubicados en el cliente, con el objetivo final de cumplir objetivos jerárquicos, en este sentido, los directores deben estar interesados en las disposiciones de sus representantes, tales estados de ánimo

advierten sobre posibilidades y conductas de impacto, infiriendo dimensiones de eficiencia autoritaria” (pag.4).

Cabe mencionar que el Machine Learning hoy en día es una solución eficaz debido a las deficiencias que se vienen presentando en las organizaciones a la hora planificar actividades o recursos, ya que toda organización genera datos a través de sus movimientos diarios o sistemas transaccionales o ya sean datos almacenados manualmente en texto plano o en hojas de cálculo en Excel que son para las pequeñas empresas que no pueden costear herramientas que les permita un almacenamiento de los datos.

De igual manera estos problemas se presentan en todo tipo de empresas como es el caso de la empresa Konecta Perú. Visualizamos en la Figura 1 la productividad real se mide en base a los datos planificados, cuya planificación actualmente no viene siendo productiva cuando se compara con la realidad, en la elaboración de estos mismos.

**Figura 1: Productividad**

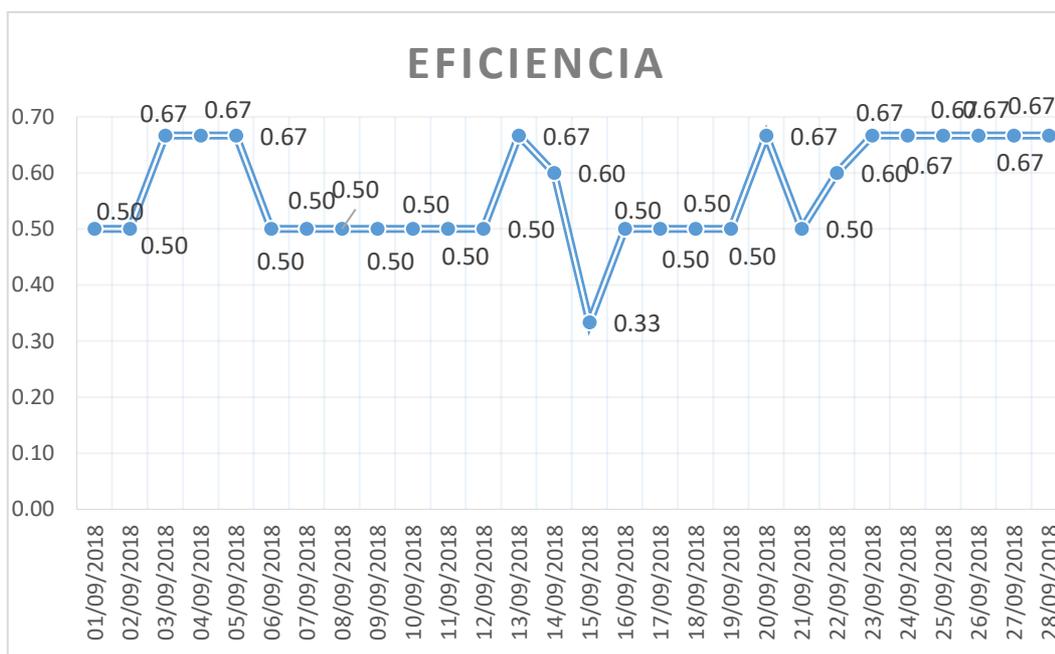


**Fuente: Elaboración propia**

La medición pretest del indicador productividad arrojó un resultado del 0.58 en un lapso de tiempo del 1 al 28 de septiembre del 2018 en la figura N°1 se muestra la productividad alcanzada.

De igual manera se observa en la Figura 2 la eficiencia real se mide en base a los datos planificados, cuya eficiencia actualmente no viene siendo la que se espera cuando se compara con la realidad, en la elaboración de esto mismos.

**Figura 2: Eficiencia**



**Fuente: Elaboración propia**

La medición test del indicador eficiencia dio un resultado de 0.58 en un lapso de tiempo del 1 al 28 de septiembre del 2018 En la figura N°2 se muestra la eficiencia alcanzada.

Actualmente en la empresa Konecta Perú del sector telecomunicaciones se realiza la planificación de indicadores como por ejemplo cuantas llamadas deben de ingresar, cuantas llamadas serán abandonadas, cuantos agentes deben estar conectados en cierto intervalo de hora, cuantos agentes deben estar disponibles para contestar las llamadas que van a ingresar, cuanto será el tiempo medio operativo (TMO), todos estos por intervalos de medias horas, día y mes, este proceso lo vienen realizando manualmente en hojas de cálculo Excel la cual demanda más horas de trabajo e incomodidad por parte del

personal teniendo mayor probabilidad de errores en la planificación, una vez realizada la planificación son enviados a operaciones para que en base a ello puedan posicionar a los agentes y puedan gestionar, actualmente la planificación no viene siendo eficaz al momento de compararlo con la realidad por ende conlleva a una mala gestión en operaciones por parte de los agentes, perjudicando a los indicadores que tienen como objetivos, por lo tanto dada estas planificaciones no eficaces se pueden dar penalizaciones monetarias por parte del cliente a quienes se les brinda el servicio como ejemplo Diners Club, Calidda, Movistar Perú, Movistar Chile, Latam, Entel Perú, entre otras empresas, por otro lado estas planificaciones tienden a ser enviados fuera del plazo establecido la cual hace que se genere una mala toma de decisiones a nivel gerencial y directivo, tampoco se tiene la información actualizada debido a ello la gerencia y directiva estarían siendo penalizadas por parte de los clientes ya mencionados.

## **1.2 Trabajos Previos**

En el año 2016, Medrano Parado, Sandra Zoraida realizo la investigación titulada "Modelo de minería de datos usando Machine Learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016" (Tesis para optar grado de Ingeniero de sistemas) en la Universidad Cesar Vallejo situada en Trujillo; muestra que este problema afecta a todo incluido y lo que se ha buscado durante mucho tiempo es tener una cultura que permita cambiar los estilos de trabajo, elevar los acuerdos y los sistemas para disminuir el efecto de los errores, de esta manera. Eso garantiza la seguridad del paciente como en los Estados Unidos: National Patient Safety Foundation, en México El Programa Nacional de la Cruzada por la Calidad de los Servicios de Salud y en diferentes naciones, a pesar de que la Organización Mundial de la Salud enmarcará la "Unión Mundial para la Seguridad del Paciente" A pesar de esto, se ha ido incrementando gradualmente y el avance procederá así si no se toman medidas para mejorarlo. La forma de vida del bienestar del paciente. Donde se propuso mejorar la conclusión de los

pacientes en la ciudad de Trujillo con el avance de un modelo de minería de información con el reconocimiento de ejemplos de efectos secundarios e infecciones respiratorias. El tipo de investigación según lo que está conectado, la confirmación del procedimiento de prueba es Pre Experimental, la población es la ciudad de Trujillo, los instrumentos utilizados en el examen son la encuesta y la investigación narrativa. El sistema que se ejecutó fue, en general, se trata de la mejora de los hallazgos terapéuticos de un modelo programado que utiliza el cálculo del árbol de opciones.

Por consiguiente, esta investigación nos da el aporte del uso de las redes neuronales como herramientas a base de su implementación mediante patrones generando un valor agregado.

En el 2013 Beltrán Castañón Norman, Realizo la siguiente investigación titulada “Modelo para proyección de balances energéticos nacionales del Perú según la estructura insumo-producto; Aplicando redes neuronales”. (Tesis para optar grado de Ingeniero Economista) de la Universidad Nacional del Altiplano. El objetivo en General es realizar y desglosar la proyección de la solicitud de vitalidad, por ejemplo, gas inflamable, petróleo, gas, etc. Según lo indicado por las partes monetarias en el Perú. El tipo de investigación es aplicada y el aspecto de la actividad, gráfico, ya que alude a la introducción del considerable número de efectos secundarios del examen; La investigación sobre la actividad debido al tratamiento numérico sugirió que los permisos para aprobar cuantitativamente los resultados mediante la utilización de modelos de proyección, esta técnica está relacionada con la razón para establecer los fines y resumir las consecuencias del examen. En el examen, el tamaño correspondiente de la inspección se evaluó en las áreas residenciales y se encontró que la vitalidad se ajusta, con una normalidad marginal de 2.4% en comparación con 2013, cifra menor que el desarrollo de 2011 a 2013 que superó el 3%. El Diesel es el interés más notable con el 53%, el 18% de GLP, el 17% de gasolina, el 9% de gasolina de aviación y el 3% restante. Pero para el 2025, se tendría un crecimiento de 15% para el Diesel en promedio por año, 12% para la gasolina y 13% para gasolina de aviación. Finalmente, los residuales tendrían un decrecimiento de 1.8% para el año 2025. Se concluye

en la necesidad de impulsar el modelo de proyección de Balances Nacionales y Regionales de Energía, según estructura de Leontief, utilizando un modelo de proyección, tomando como caso la información que se cuenta del Balance Nacional de Energía de Perú a teniendo competitividad de costos, diferenciación y flexibilidad.

El presente antecedente se ha utilizado como referencia para elaboración del tipo de estudio que es aplicada, también para entender algunos conceptos sobre los tipos de estudio y conocer el marco teórico.

Según Carrillo Ucuñay, Melissa y Vites Ángeles Celeste con la tesis "Aplicación de una red neuronal para el control domótica de encendido de luces en una vivienda de la urbanización Santa Ángela- Chiclayo 2015. Universidad Nacional Pedro Ruiz Gallo. Actualmente en una vivienda en la urbanización de Santa Ángela de la ciudad de Chiclayo, los métodos de interruptores, esto permite apagar las luces o controlar una y otra vez su poder, ya sea desde una habitación o básicamente desde una habitación. Asociado con un tomacorriente, el método para encender y apagar es insignificante en cualquier punto en que estos individuos puedan completarlo sin problemas o problemas, y colaborar de manera viable con los objetos de su vivienda de una manera característica y directa, ya que el medio en el cual se observa que fue estructurado y trabajado para ser utilizado a lo largo de estas líneas. El tipo de investigación es aplicada. El uso de una red neuronal para el control demótico está preparado para robotizar idealmente la iluminación de las luces, mejorando la vida de las personas con una versatilidad limitada en una casa de la Urbanización Santa Ángela. Terminando con el estudio, podría reconocer que con un marco de control de entrada que aplica la innovación biométrica fisiológica, al igual que el reconocimiento de la voz humana, se puede lograr una prueba reconocible y una validación de sí misma. Algunas razones nos llevan a suponer que un enfoque de red neuronal es una respuesta adecuada para el problema del reconocimiento de discurso, desde el principio, a la luz del hecho de que la voz tiene ejemplos delimitados por los fonemas inscritos; En segundo lugar, dado que los sistemas neuronales son versátiles y entrenables, con las pruebas de voz grabadas en documentos, pueden tener las cualidades esenciales para permitir una salida del sistema perfecto. Se razonó que la

mejora del rango se debe al examen de recurrencia de las wavelets, con lo cual podemos conocer aproximadamente el compromiso de las frecuencias por nivel de tiempo en las señales del discurso, ya que investiga con pequeños segmentos de wavelets de alta recurrencia en la bandera y con wavelets segmentos más grandes de baja recurrencia presentes en la bandera. Existen algunas técnicas utilizadas para la extracción de calidades en la preparación avanzada de banderas para el Reconocimiento de Voz Automático, destacándose entre las más vigorosas y utilizadas del MFCC, que presenta diversos tipos de ejecución. La suficiencia del reconocimiento de discurso 146 de MFCC es que es un modelo de representación decente de la creación e impresión del discurso, que se adquiere debido a la recopilación de diferentes estrategias, por ejemplo, el septum, la escala de Mel, la transformada de Fourier. Una certeza similar de la recopilación de algunas técnicas para obtener una viabilidad más prominente en el reconocimiento, aumenta el tiempo de ejecución del cálculo, lo que brinda la oportunidad de obtener un tiempo complejo, por segmento de tamaño  $n$ , y un tiempo de ejecución más largo. Como aporte a la investigación nos muestra los pasos a seguir de una red neuronal, es por ello que se tiene un análisis previo y asimismo elevar el tiempo de ejecución del algoritmo dando como resultado predicciones que puedan ser entrenables.

En el 2014, Alarcón Jaque, Mauricio. Realizo su investigación de “Redes neuronales aplicando en la predicción del signo de los fondos de AFP CUPRUM”. Universidad de Chile Economía y Negocios; Demuestra que las redes neuronales se convierten en una forma completamente razonable de lidiar con los problemas abiertos de la organización. En cualquier caso, el volumen expansivo de datos requerido cada día para preparar estos datos, son requisito exponencial de tiempo, requerido para absorberlos, arriesga increíblemente su ejecución en situaciones reales. El objetivo general de este estudio es decidir si la estrategia de redes neuronales tiene un límite de nivel de pronóstico de señal (PPS) precoz en la estimación de la cantidad de los multifondos de discapacidad del gobierno de A.F.P. Cuprum S.A. El tipo de investigación es deductivo con el objetivo de que su estudio proponga ciertas

especulaciones de que a través de la investigación se planifique anticipar los fines adquiridos, proponiendo modelos clave en las reservas de afirmación. Este estudio tuvo un tamaño de población de 84.4% en su calidad de ejemplo adicional para el Fondo de Aseguramiento, que se resolvió con una certeza intermedia del 95%, una centralidad medible del 5% y una estimación básica de 1.96, el mejor retorno obtenido al utilizar Redes Neuronales fue del 135,92% encontrado en uno de los modelos del Fondo A, superando en un 12,62% el método inactivo. Además, la Prueba DA estableció que el mejor multifondo cuando se dice todo en términos efectivos fue el Fondo D, ya que el 40% obtuvo ganancias más altas utilizando Redes Neuronales para Compra y Retención, y de esta manera los califica con capacidad presciente para descartar la especulación no válida. Inferimos que el uso de Neural Network Ward nos permite anticipar constantemente la diferencia en la indicación de la estimación de oferta de los multifondos AFP Cuprum S.A. Esto depende de la forma en que el 60% de los modelos de cada uno de los diversos multifondos superó el PPS objetivo del 64%, un 9% más que la estimación base del 55% para calificar un modelo con una gran naturaleza de límite de presciencia. Esta exploración permite que la utilización de las Redes neuronales como un aparato esencial en la ejecución de una metodología de especulación funcional para un especialista financiero en beneficios, produzca estima, ya que los resultados obtenidos con respecto a la productividad durante el marco de tiempo de la investigación, generaron rentabilidad sobre el procedimiento separado en la mayoría de los Fondos de AFP Cuprum SA.

El aporte de esta investigación ayudo a tener en cuenta sobre las diferentes técnicas de machine Learning que existen hoy en día y poder aprovecharlas para nuestra investigación.

En el 2018, Ibarra, José y Kimberly Paredes. Realizaron su investigación de “Redes neuronales artificiales para el control de acceso basado en reconocimiento facial”. En la Universidad Pontificia Universidad Católica del Ecuador; en el siguiente proyecto se ve el desarrollo de un software basado en redes neuronales artificiales para el reconocimiento rápido de rostros y su utilización en el control de acceso. El reconocimiento facial es un tema de gran interés por su impacto y posibles aplicaciones en temas de carácter laboral,

control de ingreso a espacios, seguridad ciudadana, activación y funcionamiento de equipos, entre otros. El control de acceso es una de las medidas que puede contribuir a garantizar la seguridad del personal dentro de una organización o en entornos domésticos, por lo cual es importante el uso de herramientas tecnológicas que ayuden al apropiado reconocimiento facial y automatizar los procesos de control de acceso. La inteligencia artificial ha demostrado su utilidad y aplicación en diversas áreas del quehacer cotidiano y particularmente las redes neuronales artificiales por su capacidad de aprendizaje y generalización es una técnica novedosa y con potencialidades para su uso en el reconocimiento facial. En el sistema desarrollado se utilizan las redes neuronales artificiales para el reconocimiento facial de forma rápida a través de la extracción de características tomadas de la imagen del rostro. La implementación del sistema incluye el desarrollo usando herramientas de hardware libre para el control de acceso. El sistema desarrollado fue implantado en una empresa en la ciudad de Ibarra, Ecuador, obteniéndose resultados favorables y que permiten el registro y control de acceso a los empleados y visitantes. Como conclusión con el desarrollo del control de acceso basado en reconocimiento facial mediante el uso de redes neuronales artificiales, se tiene un mayor control en el acceso a la empresa debido a que accede al reconocimiento facial una vez que se ingrese correctamente la contraseña del usuario, la misma que es generada por el sistema de forma automática, esto hace que el sistema sea más confiable.

El aporte de la investigación es mediante el uso de distintos modelos de que están dentro de la inteligencia artificial dando una proyección de resultados favorables referente a la capacidad de aprendizaje y generalización que tiene como técnica novedosa que se vienen usando hoy en día.

En el año 2014 Garcia Gazabon, Gisela. Realizo su investigación para optar su grado de Maestría titulada “Modelo de Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos” en la universidad Tecnológica de Bolívar – Cartagena 2014. Indica que en Medellín, Bogotá, Cuando todo está

dicho, hay un tema importante de consideración en las oficinas de crisis de los centros distintivos de la ciudad y se ve por el aplazamiento en la consideración y la decepción de los clientes de la administración de bienestar; Es probable que la caída de esta región sea causada por numerosas variables convincentes, por ejemplo, la ausencia de fundamentos físicos y de recursos humanos, la mala asignación de activos, la ausencia de personal adecuadamente preparado, el trabajo de desmotivación y, sobre todo, los pacientes que utilizan la administración de crisis de una instalación. Buscando consideración efectiva contrastada con consideración. Dado en reuniones externas. Esto provoca un encuentro negativo con el cliente. El objetivo esencial que este método aborda es demostrar cómo la probabilidad del evento de una ocasión, típicamente dicotómica, afecta la proximidad o la no asistencia de diferentes componentes y su estima o nivel. También se puede utilizar para medir la probabilidad de evento de cada uno de los resultados concebibles de una ocasión con más de dos clases (politómicas). Al incorporar los procedimientos de Aprendizaje automático en la reorganización del triaje con la zona de consideración de base, estamos contribuyendo con la seguridad de los modelos que pueden ayudar a encontrar esos ejemplos ocultos en la información que depende de la experiencia de los médicos y las escalas sugeridas por la escritura. En esta exploración nos centramos en el tipo de aprendizaje regulado. Finalmente, el orden de las crisis en las instalaciones de la ciudad es un problema que se está expandiendo, sin embargo, como se ha presentado en medio del avance de esta exploración desde puntos de vista alternativos, podemos hacer arreglos para vigilar la mejora constante de la exploración. Procedimiento de agrupación para la confirmación de pacientes en un Servicio de Urgencias.

Como aporte la técnica de clasificación de machine Learning permite clasificar distintos tipos de grupos que se tenga en una base de datos esto nos pueda ayudar para encontrar patrones en nuestra base de datos histórica que tenemos para la elaboración de la planificación.

En el 2016, Ján Vojt. Realizo su investigación de "Deep neural networks and their implementation". En la Universidad Charles University in Prague. Indica que; las redes neuronales profundas representan un modelo eficaz y universal

capaz de resolver una amplia variedad de tareas. Esta tesis se centra en tres diferentes tipos de redes neuronales profundas:

The multilayer perceptron, the convolutional neural network, and the deep belief network. Todos los modelos de red discutidos son implementados en hardware paralelo, y probado exhaustivamente para varias opciones de la arquitectura de la red y sus parámetros. El sistema implementado va acompañado.

Por una documentación detallada de las decisiones arquitectónicas y propuestas optimizaciones La eficiencia del marco de implementación es confirmada por el Resultados de las pruebas realizadas. Una parte significativa de esta tesis representa adicional Pruebas de otros marcos existentes que soportan redes neuronales profundas. Esta comparación indica un rendimiento superior a los marcos rivales probados de perceptrones multicapa y redes neuronales convolucionales. La creencia profunda. La implementación de la red funciona ligeramente mejor para capas RBM con hasta 1000 neuronas ocultas, pero tiene un rendimiento notablemente inferior para una RBM más robusta En comparación con el marco rival probado.

### **1.3 Teorías relacionadas al tema**

#### **1.3.1 Planificación**

Para definir el concepto de planificación Friedmann (2012) deduce que al respecto: “La planificación se considera un procedimiento dentro de la administración, a través del cual se hacen más racionales las decisiones referentes a los fines y métodos de las grandes organizaciones, se dice que es un proceso objetivo siendo definida una meta de una organización tomando en cuenta programas y parámetros específicos de acción” (p 22). Por otro lado, Navajo (2017) dice que “la planificación necesita que las organizaciones aporten recursos, y que el equipo dedique más horas a analizar y elaborar el plan, por lo tanto, para realizar este plan, las empresas deberían de tener una situación estable. No es un momento bueno abordar esta tarea en situaciones donde la empresa tenga problemas económicos” (p 40).

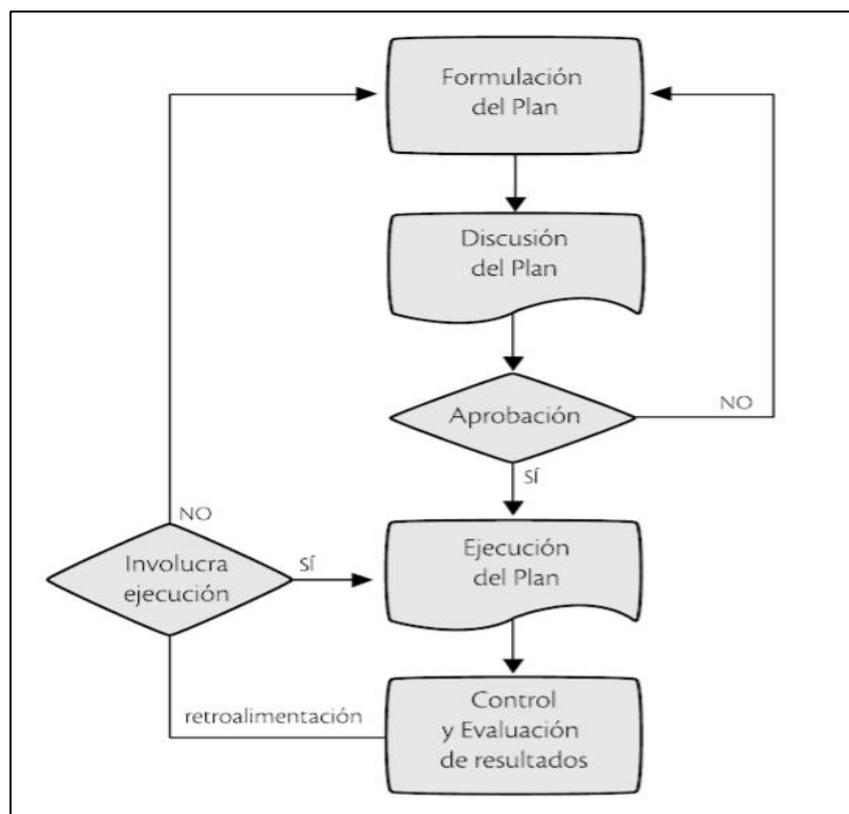
Al respecto existen diferentes conceptos que sustentan la realización del proyecto de investigación. Gonzalo Martner (2011) sostiene que “[...] La

planificación es el proceso de fijar los objetivos inmediatos ya sea a corto o largo plazo, consiste en identificar, ordenar y transformar equitativamente los recursos disponibles, para alcanzar objetivos de bienes y servicios en un plazo y localización determinados.” (p 74).

Con base a lo anterior, entendemos que la planificación se efectúa de la siguiente manera:

- Formulación del Plan.
- Discusión y aprobación del plan.
- Ejecución del Plan.
- Control y evaluación.

**Figura 3. Fases de la Planificación.**



**Fuente: Gonzalo Martner (2011)**

### **Formulación del Plan.**

Gonzalo Martner (2011, p 74) menciona que: “La formulación del plan, una guía, un proyecto o un programa que es reflejado mediante un documento final, abordando de decisiones ordenadas, integradas y coherentes, las cuales han sido evaluadas para la gestión que abordan estos objetivos, desarrollando alternativas potenciales para la organización, alcanzando las metas a través de ciertos medios con determinados recursos”

### **Discusión y Aprobación del Plan.**

Gonzalo Martner (2011, p 75) menciona que: “La discusión del plan se realiza con base al proyecto del mismo, es aprobado por el gerente del proyecto, el patrocinador del proyecto o los gerentes funcionales que proporcionan los recursos para el proyecto, identificando con éxito a todas las partes interesadas y sus requisitos y objetivos”.

### **Ejecución del Plan.**

Gonzalo Martner (2011, p 76) menciona que: “Cuando el plan ha sido aprobado y divulgado ampliamente se establece los medios para ejecutar, monitorear y controlar proyectos, la ejecución del plan sirve como el principal vehículo de comunicación para garantizar que todos conozcan los objetivos del proyecto y cómo se lograrán y así, alcanzar las metas propuestas”.

### **Control y Evaluación.**

Gonzalo Martner (2011, p 77) menciona que: La evaluación y el control se centran en los métodos de evaluación del desempeño de la empresa con respecto a los objetivos, permitiendo la satisfacción y el rendimiento del proyecto en relación con los objetivos y estándares. Estas medidas de evaluación y control permiten nuevas políticas, capacitación y procesos mejorados”.

### **Dimensión: Control y Evaluación.**

Gonzalo Martner (2011, p 77) menciona que: “Las funciones del director del plan, es la de controlar la realización de las actividades de los funcionarios. Esto permite asegurar el cumplimiento de los resultados esperados. Un posterior análisis sobre el estado de las actividades que se controlan, para

proponer o introducir las medidas correctivas de rigor, nos ubica en la función de evaluación de los resultados”

## **Indicadores.**

### **Eficiencia.**

Manuel Fernández (2011, p 67) indica que: “La Eficiencia es algo que hace corresponder a hacer las cosas correctas, por lo que las cosas deben hacerse o ejecutarse con el objetivo de que los activos estén conectados de la manera más razonable posible”.

$$\text{Eficiencia} = \frac{\text{Llamadas Reales}}{\text{Llamadas Planificadas}}$$

### **Productividad.**

Zacarías Torres (2014, p 302) menciona que “La productividad es el aprovechamiento eficiente de recursos y se mide por la producción obtenida y la cantidad de insumos utilizados en esa producción, sin duda un medio que facilita y hasta cierto grado garantiza tanto la rentabilidad como la competitividad. El insumo es lo que utilizo para poder producir como las horas trabajadas, herramientas, etc.”.

$$\text{Productividad} = \frac{\text{Llamadas Reales}}{\text{Insumo}}$$

### **1.3.2 Machine Learning**

Para definir el concepto de Machine Learning Miranda (2015, p 65) sostiene que “El Machine Learning, llamado en español como Aprendizaje Automático, es una disciplina científica habitualmente encuadra dentro del área de la

inteligencia Artificial, es una técnica donde el ordenador aprende con ejemplos y el aprendizaje se aplica para resolver nuevos casos”.

Por otro lado, Müller (2016, p.1) nos dice que “el Machine Learning consiste en extraer el conocimiento de los datos. Es un campo de investigación en la intersección de las estadísticas, la inteligencia artificial y la informática y también es conocido como analítica predictiva o aprendizaje estadístico”.

También, Chio et al. (2018, p.53) nos dice que “el Machine Learning se refiere a un aspecto de este objetivo, específicamente, a los algoritmos y procesos que aprenden en el sentido de poder generalizar datos y experiencias pasadas para predecir resultados futuros. En su núcleo, el aprendizaje automático es un conjunto de técnicas, implementadas en sistemas informáticos, que permiten un proceso de minería de la información, descubrimiento de patrones y deducciones a partir de datos.”

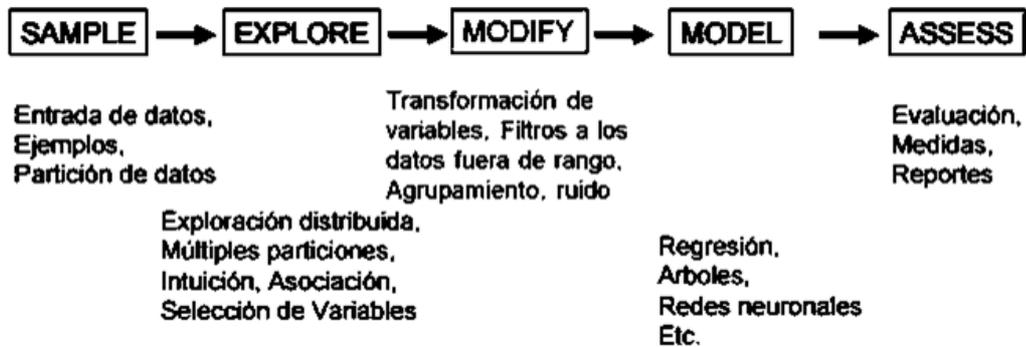
### **1.3.3 Metodologías de Machine Learning**

#### **1.3.3.1 SEMMA**

La metodología Semma en el mercado actual es también una herramienta comercial de análisis de datos, ya que tiene un repertorio básico de funciones estadísticamente para la preparación de datos.

David (2010, p.140) nos dice que “SEMMA (Sample, Explore, Modify, Model y assess) una metodología para el análisis de datos empleada por ENTERPRISE MINER SAS, Aporte técnicas específicas para asociaciones, patrones secuenciales, arboles de decisión, redes neuronales, regresión logística”

FIGURA 4: FASES DE LA METODOLOGÍA SEMMA

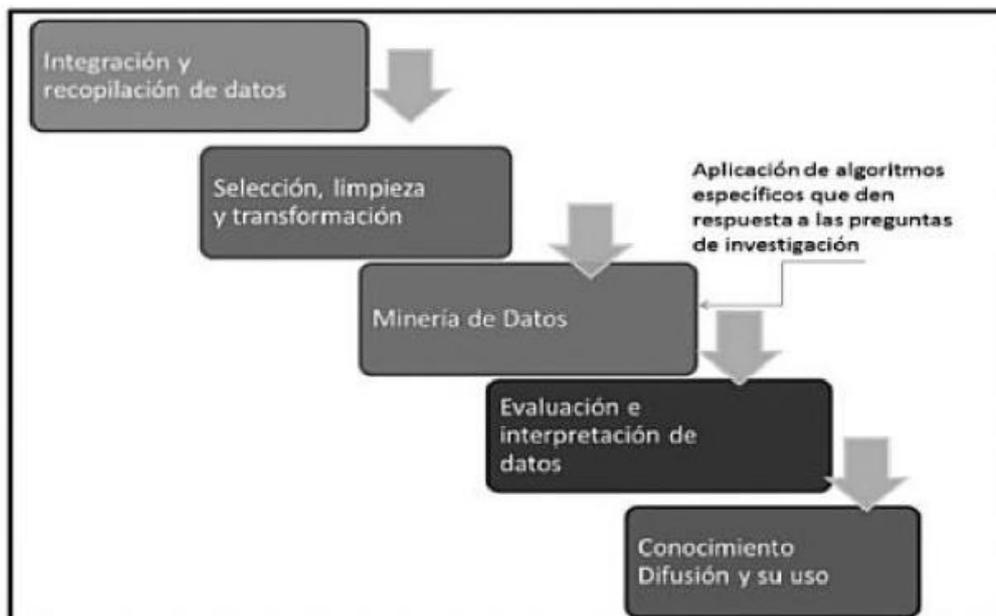


FUENTE: DAVID 2010

### 1.3.3.2 KDD

Díaz (2018, p 120) nos comenta “La metodología KDD es también llamado un paradigma centrándose en la exploración informatizada de grandes cantidades de almacenes de datos, descubriendo relevantes patrones”

FIGURA 5: PROCEDIMIENTO DE LA METODOLOGIA KDD



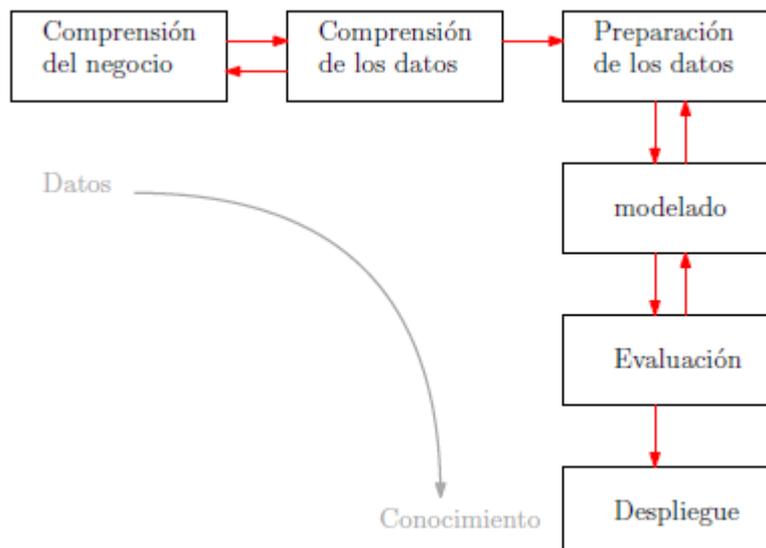
FUENTE: DIAZ 2018

### 1.3.3.3 CRISP DM

La metodología CRISP DM de IBM es el lenguaje de las estadísticas proponiendo un ciclo de 6 etapas visualizadas en la figura 6. Esta metodología provee algunos pasos sé que requieren para un proyecto de minería de procesos, como son el entendimiento de los objetivos del negocio, recopilar, describir, analizar y limpiar datos. Las tareas generales se componen a su vez de actividades específicas, y de un conjunto de resultados concretos. La metodología CRISP-DM constituye, por lo tanto, un mapa de ruta que permite determinar qué actividades desarrollar en qué etapa de manera de alcanzar los objetivos finales del proyecto.

Según Girones Roig (2017, p 28) menciona que: “La metodología Crisp DM fue unas de las pioneras en la aplicación de técnicas de minera de datos”.

FIGURA 6: FASES DE LA METODOLOGIA CRISP DM



FUENTE DE: GIRONES 2017

Tabla 1: Selección de la metodología

	CRITERIO	CRISP DM	KDD	SEMMA
1	El modelo es adecuado para el proyecto.	X		
2	Basado en los requerimientos del usuario para la implementación propuesta.	X		
3	El modelo es flexible a los cambios del proyecto de investigación.	X		
4	Involucra al cliente durante etapas del proyecto.	X		
5	Riesgos y costos son controlados a medida que surgen los requisitos y las limitaciones.	X		
6	Involucra al usuario como parte del equipo de desarrollo.	X		
7	La documentación es menor para el proyecto de investigación	X		

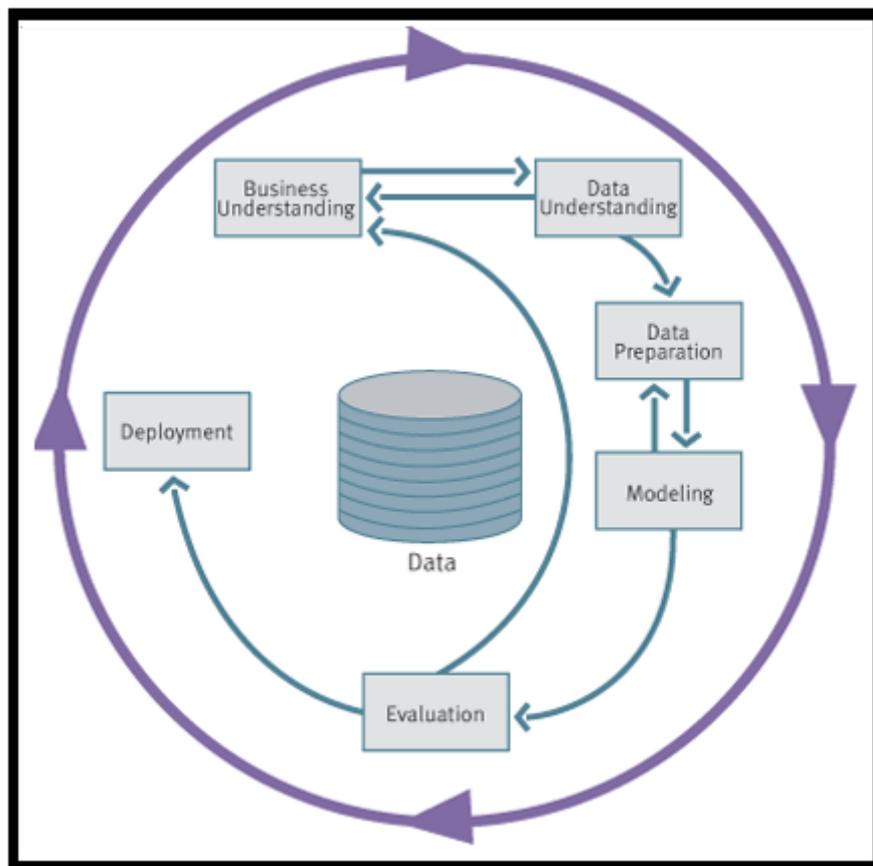
**Tabla 2: Evaluación de expertos**

EXPERTO	CRISP DM	KDD	SEMMA
Gálvez Tapia, Orleans Moisés	35	29	21
Huarote Zegarra Raúl	35	28	33
Chumpe Agosto Juan	31	31	28
<b>TOTAL</b>	<b>101</b>	<b>88</b>	<b>82</b>

La metodología Crisp DM tiene mayor aceptación, por lo que se ha seleccionado dicha metodología para el desarrollo del Machine Learning.

Según Braga (2012, p. 11) Nos dice que “La metodología Crisp DM es la más actual o la mejor, pero muy útil para la comprensión de datos o extraer ideas para realizar nuevos diseños y métodos de trabajo”.

**Figura 7: Metodología Crisp DM**



**Fuente de Braga (2012)**

## **Fases de la metodología Crisp DM**

Según Braga define las siguientes fases de la metodología de Crisp DM

**Comprensión del negocio** (Objetivos y requerimientos desde una perspectiva no técnica)

- Establecimiento de los objetivos del negocio (Contexto inicial, objetivos, criterios de éxito)
- Evaluación de la situación (Inventario de recursos, requerimientos, supuestos, terminologías propias del negocio,...)
- Establecimiento de los objetivos de la minería de datos (objetivos y criterios de éxito)
- Generación del plan del proyecto (plan, herramientas, equipo y técnicas)

**Comprensión de los datos** (Familiarizarse con los datos teniendo presente los objetivos del negocio)

- Recopilación inicial de datos
- Descripción de los datos
- Exploración de los datos
- Verificación de calidad de datos

**Preparación de los datos** (Obtener la vista minable o dataset)

- Selección de los datos
- Limpieza de datos
- Construcción de datos
- Integración de datos
- Formateo de datos

**Modelado** (Aplicar las técnicas de minería de datos a los dataset)

- Selección de la técnica de modelado
- Diseño de la evaluación
- Construcción del modelo
- Evaluación del modelo

**Evaluación** (De los modelos de la fase anteriores para determinar si son útiles a las necesidades del negocio)

- Evaluación de resultados
- Revisar el proceso
- Establecimiento de los siguientes pasos o acciones

**Despliegue** (Explotar utilidad de los modelos, integrándolos en las tareas de toma de decisiones de la organización)

- Planificación de despliegue
- Planificación de la monitorización y del mantenimiento
- Generación de informe final
- Revisión del proyecto

#### **1.4 Formulación del problema**

PG: ¿En qué medida el Machine Learning influye en la planificación en el servicio de atención de llamadas en la Empresa Konecta BTO, S.L. Sucursal en Perú?

##### **1.4.1 Problemas específicos**

P1: ¿En qué medida el Machine Learning influye en la eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú?

P2: ¿En qué medida el Machine Learning influye en la productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú?

#### **1.5 Justificación de estudio**

##### **1.5.1 Justificación tecnológica**

Philip Osborne (2013, p.56) menciona que: “El Machine Learning se puede usar para una variedad de problemas de planificación, incluyendo planes de viaje, planificación de presupuesto y estrategia comercial. Las ventajas de usar Machine Learning es que tiene en cuenta la probabilidad de resultados y nos permite controlar partes del entorno”.

Una solución es que un sistema adquiera automáticamente el conocimiento específico del dominio a través del aprendizaje. La idea de un sistema de planificación que pueda mejorar su rendimiento con experiencia es muy atractiva. Además, los avances en el aprendizaje automático han proporcionado una comprensión más profunda de los mecanismos de

aprendizaje relevantes para adquirir dicho conocimiento. Por esta razón, hay un gran interés en esta área de inteligencia artificial.

### **1.5.2 Justificación económica**

El Machine Learning es de naturaleza táctica, el uso de Machine Learning para automatizar u optimizar algún proceso. Algunos ejemplos de esta segunda categoría de aplicaciones Machine Learning es la planificación, detección de fraudes, la optimización de campañas de marketing, la recomendación de productos, la traducción de documentos, la conducción de un automóvil. Lo que estos ejemplos tienen en común es que tienden a involucrar la toma de decisiones automatizada. La automatización le permite hacer la misma cantidad de trabajo con menos personas, o liberar a las personas y otros activos costosos para hacer proyectos más valiosos. Un simpatizante puede responder el mismo problema una y otra vez, pero es incluso mejor si se les libera la atención para que tengan la capacidad de notar patrones de consulta emergentes y compartan las ideas de los clientes con los gerentes de productos para guiar la innovación y mejorar la experiencia del cliente. En primer lugar, como ocurre con otras tecnologías, la automatización de procesos conduce en última instancia a una reducción de los costes de personal, ya que actualmente en el área de Control de Gestión hay 15 analistas cuyo sueldo es de 1500 soles mensuales, anualmente es un gasto de 270,000 soles aproximados en recursos humanos, la cual aplicando Machine Learning estos costes en recursos humanos podrían disminuir notablemente.

### **1.5.3 Justificación Institucional**

Según Chase Charles (2018, p.47) menciona que: “Las técnicas de Machine Learning se pueden aplicar a todo tipo de actividades en su organización para reducir el trabajo repetitivo cotidiano y descubrir información clave para mejorar la efectividad de sus procesos, así como su fuerza laboral, se ha convertido en una métrica estándar para que la mayoría de las compañías mejoren la precisión del pronóstico mejorando así la eficiencia del proceso”.

### **1.5.4 Justificación Operativa**

Según Alexandra Gomez (2018, p.44) menciona que: “Actualmente, el Machine Learning está revolucionando el campo de los negocios a través de su aplicación. Ante un volumen de datos prácticamente incalculable, los sistemas

de Machine Learning son capaces de identificar y procesar patrones complejos y así lograr predecir el comportamiento. Esto es especialmente útil para desarrollar un modelo propio que permita descubrir tendencias futuras y todo ello, además, sin requerir ayuda externa. De esta forma, conseguimos automatizar trabajos sin necesidad de programar explícitamente, lo que nos permite maximizar la eficiencia”.

Cuando detectan un error, los sistemas de Machine Learning son capaces de corregirlo y volver a repetir el análisis en menos de un segundo, para realizar este tipo de procesos es necesaria tener una mayor capacidad de almacenamiento.

## **1.6 Hipótesis**

### **1.6.1 Hipótesis general**

HG: Si Machine Learning es confiable, entonces mejora la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

### **1.6.2 Hipótesis específicas:**

H1: Si Machine Learning es confiable, entonces aumenta la eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

H2: Si Machine Learning es confiable, entonces aumenta la productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

Objetivos

### **1.7.1 Objetivo general**

OG: Determinar la influencia del Machine Learning en la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

### **1.7.2 Objetivos específicos**

O1: Determinar la influencia del Machine Learning en la eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

O2: Determinar la influencia del Machine Learning en la productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

## **CAPÍTULO II**

### **MÉTODO**

## **II. MÉTODO**

Método de investigación: Hipotético – deductivo

Augusto Bernal (2010, p.60), manifiesta que el “Hipotético – deductivo es en un proceso que inicia de una afirmación en calidad de hipótesis y busca refutar o falsear tales hipótesis, deduciendo de ellas conclusiones que deben confrontarse con los hechos. El método hipotético-deductivo tiene varios pasos esenciales: observación del fenómeno a estudiar, creación de una hipótesis para explicar dicho fenómeno, deducción de consecuencias o proposiciones más elementales que la propia hipótesis, y verificación o comprobación de la verdad de los enunciados deducidos comparándolos con la experiencia. Este método obliga al científico a combinar la reflexión racional o momento racional (la formación de hipótesis y la deducción) con la observación de la realidad o momento empírico (la observación y la verificación)”.

El método de la presente investigación es hipotético deductivo, ya que, en base a los problemas definidos, se debe plantear la hipótesis y verificar los datos.

### **2.1.2 Tipo de Investigación**

#### **2.1.2.1 Explicativo**

Fidias Arias (2012, p. 26) menciona que “la investigación explicativa se encarga de indagar el porqué de los hechos mediante las relaciones de causa-efecto, pueden determinar las causas (investigación post facto), como de los efectos (investigación experimental), mediante la prueba de hipótesis, las conclusiones y resultados constituyen el nivel más profundo de conocimientos”.

#### **2.1.2.2 Experimental**

Rodríguez Miguel (2005, p. 80) menciona que: “se presenta mediante la manipulación de una variable experimental no comprobada, en condiciones rigurosamente controladas, con el fin de describir de qué modo o por que causa se produce una situación o acontecimiento particular”.

#### **2.1.2.3 Aplicada**

Carrasco Díaz (2013, p.30) manifiesta que “una investigación aplicada se diferencia por tener propósitos prácticos, inmediatos, bien definidos, es decir

se investiga para actuar, transformar, modificar o producir cambios en un determinado sector de la realidad. Para realizar investigaciones aplicadas es muy importante contar con el aporte de las teorías científicas que son producidas por la investigación básica y sustantiva”

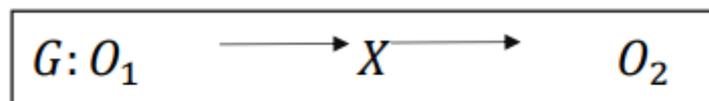
El tipo de investigación es aplicada, porque se implementará el Machine Learning para el servicio de atención de llamadas en la empresa Konecta BTO SL sucursal en Perú, el cual permitirá solucionar la problemática que presenta la empresa.

### 2.1.3 Diseño de investigación: Pre - experimental

El diseño de la presente investigación es pre experimental puesto que se evaluará la planificación en el servicio de atención de llamadas antes de implementar el Machine Learning y después para comparar resultados.

Hernández, Fernández y Baptista (2004, p.120) indican: “en el diseño de estudio pre experimental existe una subclase llamada diseño de pre prueba/pos prueba con un solo grupo, que consiste en que a un grupo se le aplica una prueba previa al estímulo o tratamiento experimental, después se le administra el tratamiento y finalmente se le aplica una prueba posterior al estímulo.” El esquema utilizado es siguiente:

Esquema del diseño:



**Diseño pre experimental de pre-prueba/post-prueba**

Dónde:

G: Grupo experimental: Es el grupo de los tipos de servicios de atención de llamadas (muestra) al cual se aplicará la medición para evaluar la eficiencia de planificación de un plazo determinado y productividad de la planificación.

X: Experimento (Machine Learning): Es la aplicación del Machine Learning en la planificación en el servicio de atención de llamadas en la empresa Konecta Bto, S.L Sucursal Perú, mediante dos evaluaciones (Pre – Test y Post – Test) para medir los cambios en la planificación del servicio de atención de llamadas.

O1: Pre – Test: Medición del grupo experimental antes de la implementación del Machine Learning en la planificación en el servicio de atención de llamadas. Esta medición se compara con la medición del Post – Test.

O2: Post – Test: Medición del grupo experimental después de la implementación del Machine Learning en la planificación en el servicio de atención de llamadas. Se comparan ambas mediciones para determinar la eficiencia en la planificación en el servicio de atención de llamadas en un plazo determinado y productividad en la planificación en el servicio de atención de llamadas antes y después de la implementación del Machine Learning.

## **Variables, Operacionalización**

### **Definición conceptual**

#### **2.2.1 Variable Independiente: Machine Learning**

Según, Antonio Miranda (2015, p.65) menciona que: “es una disciplina científica habitualmente encuadra dentro del área de la inteligencia Artificial, es una técnica donde el ordenador aprende con ejemplos y el aprendizaje se aplica para resolver nuevos casos”.

#### **2.2.2 Variable Dependiente: Planificación**

Gonzalo Martner (2015, p.77) indica que “la planificación consiste en identificar, ordenar y transformar racionalmente los recursos disponibles, para alcanzar objetivos de bienes y servicios en un plazo y localización determinados; Así mismo debe mantener unidad en el tiempo, es decir, que el plan que finalmente se obtenga como producto de esa planificación esté vinculado y sirva de referencia a otros planes de distintos plazos dados en ese tiempo.”

Las Variables que se han determinado son las siguientes:

#### **Variable Independiente (VI): Machine Learning**

Según, Antonio Miranda (2015, p.65) menciona que: “es una disciplina científica habitualmente encuadra dentro del área de la inteligencia Artificial, es una técnica donde el ordenador aprende con ejemplos y el aprendizaje se aplica para resolver nuevos casos”.

#### **Variable Dependiente (VD): Planificación**

Gonzalo Martner (2015, p.77) indica que “la planificación consiste en identificar, ordenar y transformar racionalmente los recursos disponibles, para alcanzar objetivos de bienes y servicios en un plazo y localización determinados; Así mismo debe mantener unidad en el tiempo, es decir, que el plan que finalmente se obtenga como producto de esa planificación esté vinculado y sirva de referencia a otros planes de distintos plazos dados en ese tiempo.”

A continuación, veremos la Operacionalización de variables en la Tabla 2.

**Tabla 3. Operacionalización de variables**

Operacionalización de variables					
Variable	Definición Conceptual	Definición Operacional	Dimensión	Indicadores	Escala de medición
<b>Independiente</b>	Según, Antonio Miranda (2015, p.65) menciona que: “es una disciplina científica habitualmente encuadra dentro del área de la inteligencia Artificial, es una técnica donde el ordenador aprende con ejemplos y el aprendizaje se aplica para resolver nuevos casos”.	El Machine Learning es un paso previo a la Inteligencia Artificial (IA) para que ésta se desarrolle en su totalidad. Concretamente, es un aprendizaje de máquinas que utilizan muchos datos con el objetivo de ser cada vez más inteligentes a la hora de hacer cosas. El ML se alimenta de algoritmos e información, es algo así como llevar la inteligencia al dato”.			
<b>M L A C H I N E L E A R I N G</b>					
<b>Dependiente</b>	Gonzalo Martner (2011, p.77) indica que “la planificación consiste en identificar, ordenar y transformar racionalmente los recursos disponibles, para alcanzar objetivos de bienes y servicios en un plazo y localización determinados; Así mismo debe mantener unidad en el tiempo, es decir, que el plan que finalmente se obtenga como producto de esa planificación esté vinculado y sirva de referencia a otros planes de distintos plazos dados en ese tiempo.”	La planificación consiste en formular planes a largo y corto plazo para diversas parte de la organización que ayudan con el cumplimiento de objetivos.	Control y Evaluación	Eficiencia	Razón
<b>P L A N I F I C A C I O N</b>				Productividad	

Fuente: Elaboración propia

**Tabla 4 Indicadores de la planificación**

Indicador	Descripción	Técnica	Instrumento	Unidad de medida	Formula
Eficiencia	Evaluará la eficiencia en la planificación en el servicio de atención de llamadas.	Fichaje	Ficha de registro	Unidad	Llamadas Reales / Llamadas Planificadas
Productividad	Evaluará la productividad en la planificación en el servicio de atención de llamadas.	Fichaje	Ficha de registro	Unidad	Llamadas Reales / Horas Trabajadas

Fuente: Elaboración propia

## 2.1 Población y muestra

### 2.3.1 Población

Hernández, Fernandez y Baptista (2014), nos señala “la población es un grupo de los casos que están en relación con una serie de especificaciones, con la finalidad de ver cuál será el parámetro muestral”. (p. 174).

En la investigación “Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú”, se considera una población de los 475 tipos de servicio de atención de llamadas.

### 2.3.2 Muestra

Según Hernández, Fernandez y Baptista (2014), “es un subgrupo de la población. Se dice que es un subconjunto de individuos que corresponden a ese grupo determinado en sus características al que llamamos población.” (p. 175)

Cálculo del tamaño de la muestra:

$$n = \frac{N * Z^2 * p * q}{d^2 * (N - 1) + Z^2 * p * q}$$

Dónde:

N: Población

Z: Nivel de confianza al 95% (1.96)

P: Proporción esperada (5%=0.05)

I: Precisión (0.05)

q: 1-p (0.95)

Empleando Formula:

Muestra indicadores: La eficiencia en el servicio de atención de llamadas en un plazo acordado y la productividad en el servicio de atención de llamadas.

$$n = \frac{475 * 1.96^2 * 0.05 * 0.95}{0.05^2 * (475 - 1) + 1.96^2 * 0.05 * 0.95} = 213$$

El resultado de la fórmula empleada dio una muestra de 213 llamadas de los tipos de servicio, por lo tanto, la muestra es conformada por 28 días estratificada de lunes a domingo.

### 2.3.3 Muestreo

Muestreo probabilístico:

Según Hernández, Fernandez et al (2014), En las muestras probabilísticas, consiste en dividir toda la población objeto de estudio en diferentes subgrupos o estratos disjuntos, de manera que un individuo sólo puede pertenecer a un estrato. Una vez definidos los estratos, para crear la muestra se seleccionan individuos empleando una técnica de muestreo cualquiera a cada uno de los estratos por separado. (p. 176).

### 2.3.4 Muestreo Estratificado:

Según (Hernández en el 2014 p. 140) se caracteriza por la división de la población en subgrupos o estratos debido a que las variables que deben someterse a estudio en la población presentan cierta variabilidad o distribución conocida que es necesario tomar en cuenta para extraer la muestra..

Técnicas e instrumentos de recolección de datos, validez y confiabilidad

Según, Hernández, Fernandez y Baptista (2014) indica que “la validez alude al grado donde un instrumento mide la variable que pretende medir asimismo la confiabilidad de un instrumento de medición, se refiere al grado en que la aplicación repetida tiene al mismo individuo u objeto produciendo resultados iguales.” (p.200).

En la investigación “Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú”, se utilizará como técnica el fichaje, será usado como muestra y para el instrumento se harán formatos (fichas de Registro), la cual es materia de investigación.

Figura 8. Niveles de confiabilidad

Escala	Nivel
$0.00 < \text{sig.} < 0.20$	Muy bajo
$0.20 \leq \text{sig.} < 0.40$	Bajo
$0.40 \leq \text{sig.} < 0.60$	Regular
$0.60 \leq \text{sig.} < 0.80$	Aceptable
$0.80 \leq \text{sig.} < 1.00$	Elevado

**Fuente: Hernández Fernández (2014)**

### 2.3.5 Validación

Según Hernández, Fernandez y Baptista (2014), demuestra que “un instrumento de estimación puede ser sólido, aunque no realmente legítimo. Por lo tanto, es una necesidad que el instrumento de estimación termine siendo confiable y legítimo.” (p. 204)

En la investigación “Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú”, se usará el juicio de expertos para la aprobación de los instrumentos de medición.

**Tabla 5. Validación de indicador Productividad**

	<b>Experto</b>	<b>Grado Académico</b>	<b>Puntaje</b>
1	Orleans Gálvez Tapia	Magister	80%
2	Huarote Zegarra Raúl	Magister	85%
3	Chumpe Agosto Juan	Magister	70.5%

**Fuente: Elaboración propia**

Se mostraron las fichas para que sean evaluadas por parte de los expertos para el indicador productividad obteniendo un nivel de aceptación y confianza a los instrumentos siendo los correctos para capturar los datos de dicho indicador.

**Tabla 6. Validación de indicador Eficiencia**

	<b>Experto</b>	<b>Grado Académico</b>	<b>Puntaje</b>
1	Orleans Gálvez Tapia	Magister	80%
2	Huarote Zegarra Raúl	Magister	90%
3	Chumpe Agosto Juan	Magister	70.5%

**Fuente: Elaboración propia**

Se mostraron las fichas para que sean evaluadas por parte de los expertos para el indicador eficiencia obteniendo como resultado un 80% dando un nivel de aceptación y confianza a los instrumentos siendo los correctos para capturar los datos de dicho indicador.

## Confiabilidad

Para la confiabilidad en la investigación “Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú”, se utilizará el SPSS versión 25.

**Figura 9. Indicador Eficiencia**

		Eficiencia_Test	Eficiencia_reTest
Eficiencia_Test	Correlación de Pearson	1	,737
	Sig. (bilateral)		,000
	N	28	28
Eficiencia_reTest	Correlación de Pearson	,737	1
	Sig. (bilateral)	,000	
	N	28	28

Debe notarse que la investigación de confiabilidad indicada por el SPSS produce un resultado similar, que muestra una dimensión aceptable de confiabilidad, es decir el instrumento es confiable.

**Figura 10. Indicador Productividad**

		Productividad_Test	Productividad_reTest
Productividad_Test	Correlación de Pearson	1	,708
	Sig. (bilateral)		,000
	N	28	28
Productividad_reTest	Correlación de Pearson	,708	1
	Sig. (bilateral)	,000	
	N	28	28

El SPSS produce un resultado similar, que demuestra una dimensión aceptable de confiabilidad, es decir el instrumento es confiable

### 2.3.6 Métodos de análisis de datos

En esta investigación se utilizará la prueba medible Student T, es cualquier prueba en la que la medición utilizada tenga una transferencia T de Student si la hipótesis nula es cierta. Se aplica cuando la población contemplada persigue una típica; sin embargo, la medida de ejemplo es demasiado pequeña para que la medición en la que se basa la inducción se transmita regularmente, utilizando un indicador de la desviación estándar en lugar del valor real.

#### Prueba t Student

Según Del Rio Sadornil (2013) nos indica que el propósito de la prueba T student permite comprobar las existencias o no de diferencias estadísticamente significativas entre dos grupos con respecto a un conjunto de variables dependientes consideradas asimismo conjuntamente. Esta prueba puede considerarse como una especulación para instancias específicas de variables independientes e incluso es el estadístico para terminaciones multivariante.

**Figura 11: Formula para el T student**

Fuente: Molina Quiñones, 2011	$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N}}}$	Donde: <ul style="list-style-type: none"><li>• <math>x</math> es la media muestral,</li><li>• <math>s</math> es la desviación estándar muestral</li><li>• <math>N</math> es el tamaño de la muestra.</li></ul>
-------------------------------	--	--

**Fórmula para hallar de t-Student**

### 2.3.7 Hipótesis de investigación

#### Hipótesis específico 1 (HE1)

Si Machine Learning es confiable, entonces aumenta la eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

#### Indicador Eficiencia:

Ia1: Eficiencia previamente al desarrollo de Machine Learning.

Ia2: Eficiencia posteriormente al desarrollo de Machine Learning.

#### Hipótesis nula

Se considera hipótesis de rechazo entre cada par de variables dependiendo donde no se encuentre ninguna diferencia significativa, eso quiere decir que llega a tomar estimación de parámetros constatando su hipótesis h.

### **Hipótesis estadística 1**

#### **Hipótesis Nula (H0):**

Si Machine Learning no es confiable, entonces no aumenta la eficiencia en la planificación en el servicio de atención de llamadas de la empresa Konecta BTO, S.L. Sucursal en Perú.

#### **Fórmula**

$$H_0 = \mu_1 - \mu_2$$

$\mu$  = Medida de la población

$\bar{X}$  = Medida de distribución de datos

n = Medida de la muestra

s = error estándar de la muestra

#### **Aspectos éticos**

De acuerdo con Bernal (2010), indica que la ciencia avanzada es una de las mejores secuelas de la humanidad y considerando que vivimos en un mundo representado por filosofías dependientes de la ciencia y la utilización de los instrumentos creados por esto, es importante saber que estos sistemas de creencias y esos instrumentos pueden ser utilizados de manera efectiva o inexacta, es decir, de manera moral o engañosa. En consecuencia, es importante pensar en una moral de la ciencia, es decir, una ciencia con voz tranquila, por así decirlo, una ciencia cuyos sistemas de creencias e instrumentos se utilizan con precisión.” (p.18)

De este modo, el investigador intenta considerar la veracidad de los resultados, la calidad inquebrantable de la información proporcionada por la organización y el carácter de las personas que participan en esta investigación.

## **CAPÍTULO III**

### **RESULTADOS**

### 3.1 Análisis Descriptivo

En el estudio se aplicó una técnica científica – Machine Learning para la planificación en el servicio de atención de llamadas; para ello se aplicó un Pre-Test que permita conocer las condiciones iniciales del indicador; posteriormente se implementó Machine Learning y nuevamente se registró los indicadores eficiencia y productividad. Los resultados descriptivos de estas medidas se observan en las tablas 7 y 8.

- **INDICADOR: Eficiencia**

Los resultados descriptivos de la eficiencia de estas medidas se observan en la Tabla

*Tabla 7: Medidas descriptivas de la Eficiencia en el proceso antes y después de implementar Machine Learning*

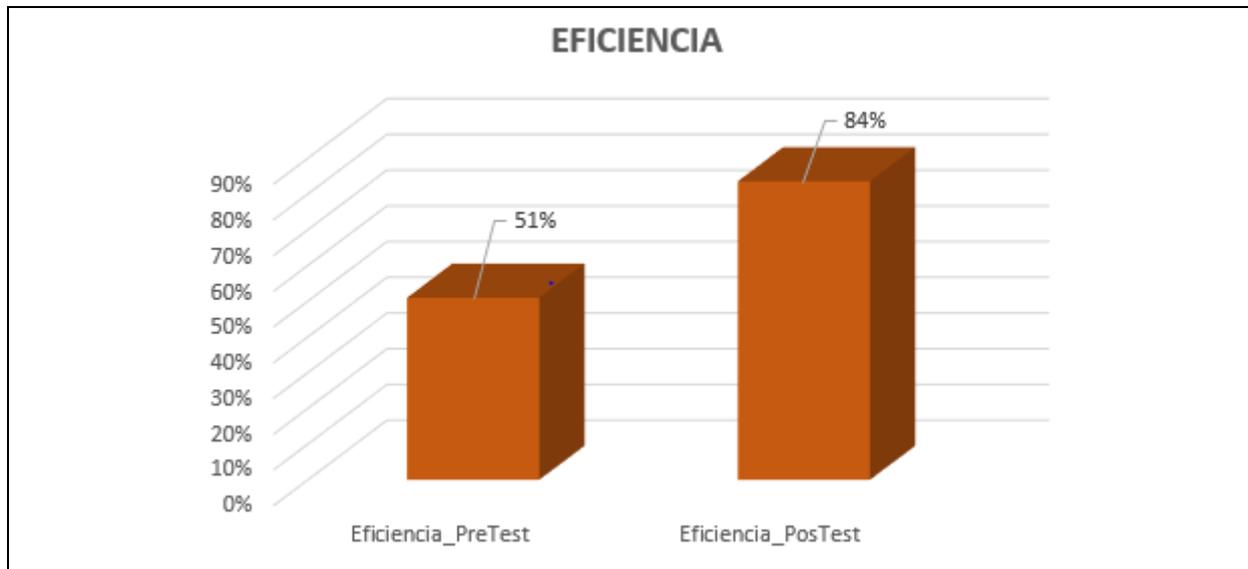
<b>Estadísticos descriptivos</b>					
	N	Mínimo	Máximo	Media	Desv. Desviación
Eficiencia_PreTest	28	,33	,75	,5096	,09946
Eficiencia_PosTest	28	,60	1,00	,8354	,09457
N válido (por lista)	28				

*Fuente: Elaboración Propia*

En el caso de Eficiencia en el proceso de la Planificación, en el pre-test se obtuvo un valor de 51%, mientras que en el post-test fue de 84% tal como se aprecia en la figura; esto indica una gran diferencia antes y después de la implementación de Machine Learning; así mismo, la Eficiencia mínima fue de 0,33 antes, y 0,60 (ver Tabla Descriptiva) después de la implementación del Machine Learning.

En cuanto a la dispersión, en el pre-test se tuvo una variabilidad de 0,99; sin embargo, en el post-test se tuvo un valor de 0,94.

**Tabla 8: Porcentaje de la Eficiencia en el proceso antes y después de implementar Machine Learning**



- INDICADOR: Productividad**

Los resultados descriptivos de Productividad de estas medidas se observan en la Tabla

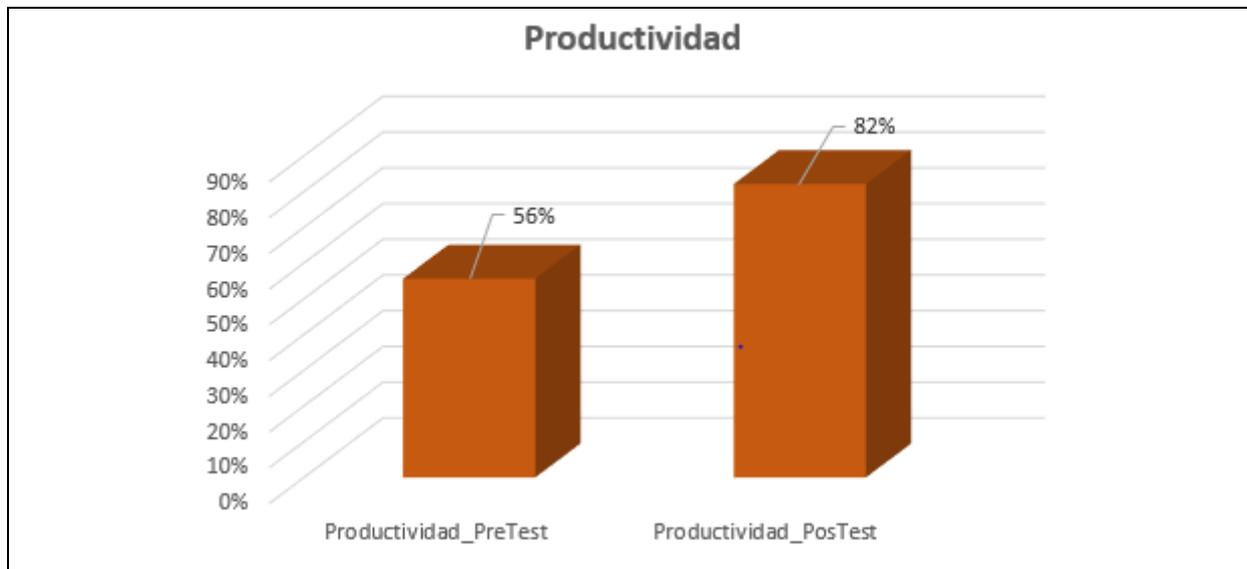
**Tabla 9: Los resultados descriptivos de Productividad**

Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desv. Desviación
Productividad_PreTest	28	,40	,75	,5561	,09585
Productividad_PosTest	28	,67	1,00	,8211	,07733
N válido (por lista)	28				

En el caso de productividad en el proceso de Planificación, en el pre-test se obtuvo un valor de 56%, mientras que en el post-test fue de 82% tal como se aprecia en la figura; esto indica una gran diferencia antes y después de la implementación de Machine Learning; así mismo, el porcentaje de Productividad mínima fue del 40 antes, y 67 (Ver tabla) después de la implementación de Machine Learning.

En cuanto a la dispersión del porcentaje de Productividad, en el pre-test se tuvo una variabilidad de 0,95; sin embargo, en el post-test se tuvo un valor de 0,77.

**Tabla 10: Porcentaje de Productividad**



### 3.2 Análisis Inferencial

#### Prueba de Normalidad

Se procedió a realizar las pruebas de normalidad para los indicadores de Eficiencia y Productividad a través del método Shapiro-Wilk, debido a que el tamaño de nuestra muestra estratificada está conformado por 28 días de planificación y es menor a 50, tal como lo indica Hernández, Fernández y Baptista (2014, p. 376). Dicha prueba se realizó introduciendo los datos de cada indicador en el software estadístico SPSS 24.0, para un nivel de confiabilidad del 95%, bajo las siguientes condiciones:

Si:

Sig. < 0.05 adopta una distribución no normal.

Sig.  $\geq$  0.05 adopta una distribución normal.

Dónde:

Sig. : P-valor o nivel crítico del contraste.

Los resultados fueron los siguientes:

- **INDICADOR: Eficiencia**

Con el objetivo de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos de la eficiencia contaban con distribución normal.

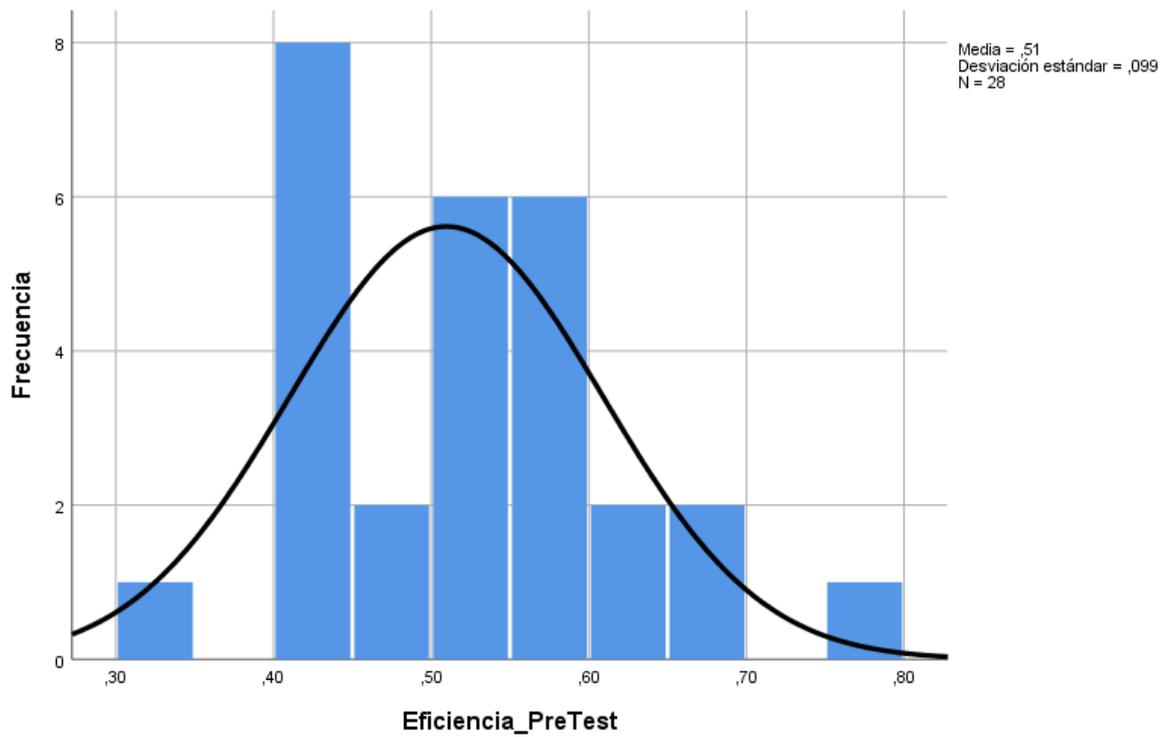
**Tabla 11: Prueba normalidad eficiencia**

Pruebas de normalidad			
		Shapiro-Wilk	
	Estadístico	gl	Sig.
Eficiencia_PosTest	,950	28	,202
Eficiencia_PreTest	,959	28	,332

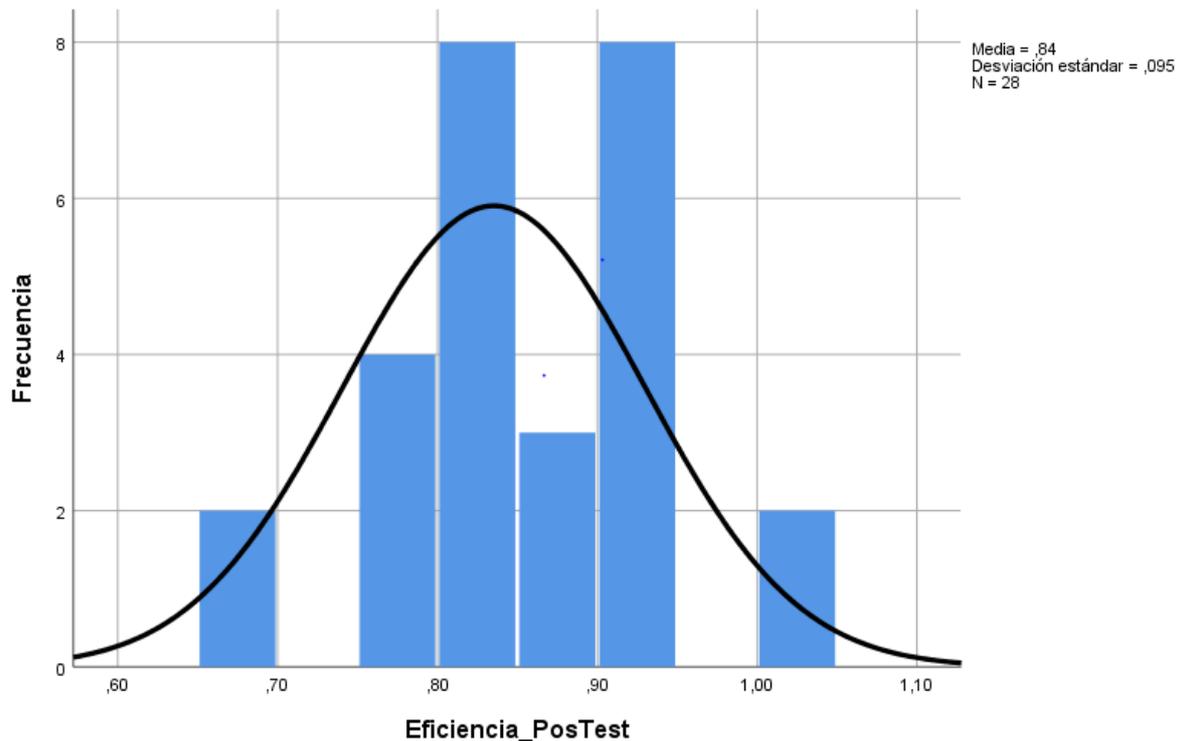
\*. Esto es un límite inferior de la significación verdadera.  
a. Corrección de significación de Lilliefors

Como se muestra en la Tabla, los resultados de la prueba indican que el Sig. De Eficiencia generados en el Pre-Test fue de 0.332, cuyo valor es mayor que 0.05. Por lo tanto, la eficiencia se distribuye normalmente. Los resultados de la prueba del Post-Test indican que el Sig. De la eficiencia fue de 0.202, cuyo valor es mayor que 0.05, por lo que indica que Eficiencia se distribuye normalmente. Lo que confirma la distribución normal de ambos datos de la muestra, se puede apreciar en las Figuras 11 y 12.

**Tabla 12: Eficiencia pre test**



**Tabla 13: Eficiencia post test**



- **INDICADOR: Productividad**

Con el objetivo de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos de Productividad con distribución normal.

**Tabla 14: Prueba de normalidad Productividad**

Pruebas de normalidad			
		Shapiro-Wilk	
	Estadístico	gl	Sig.
Productividad_PreTest	,938	28	,101
Productividad_PosTest	,966	28	,471

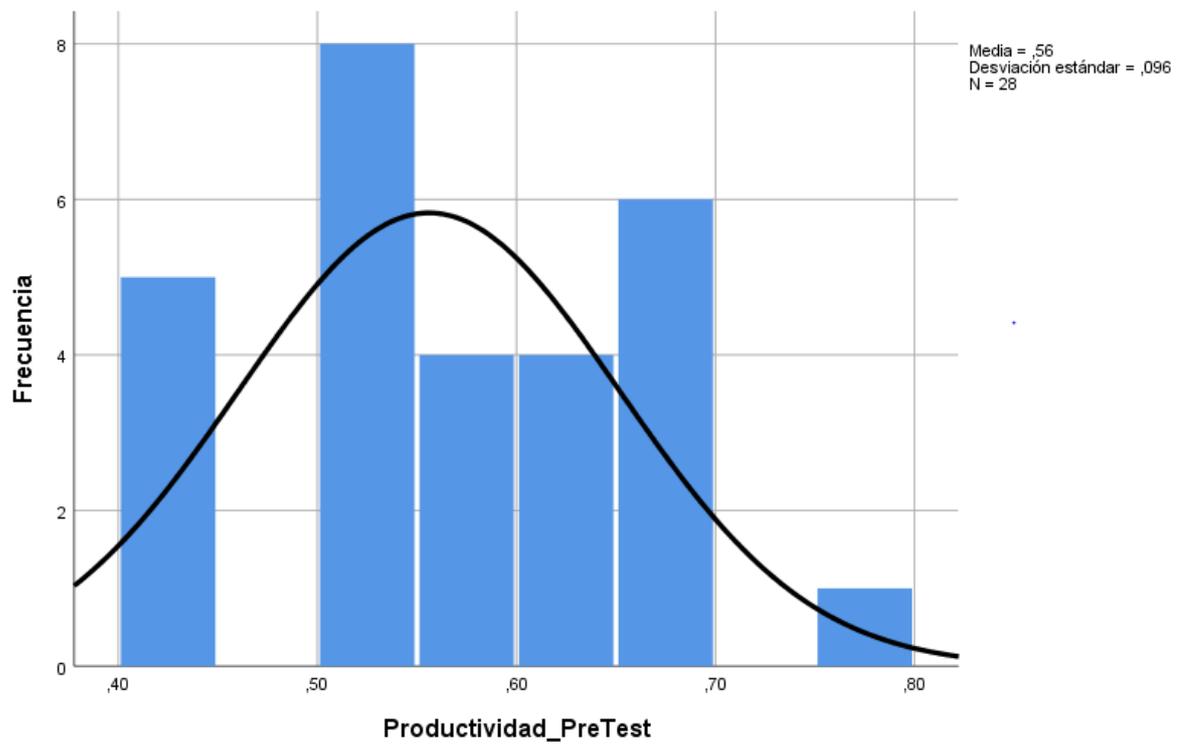
\*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

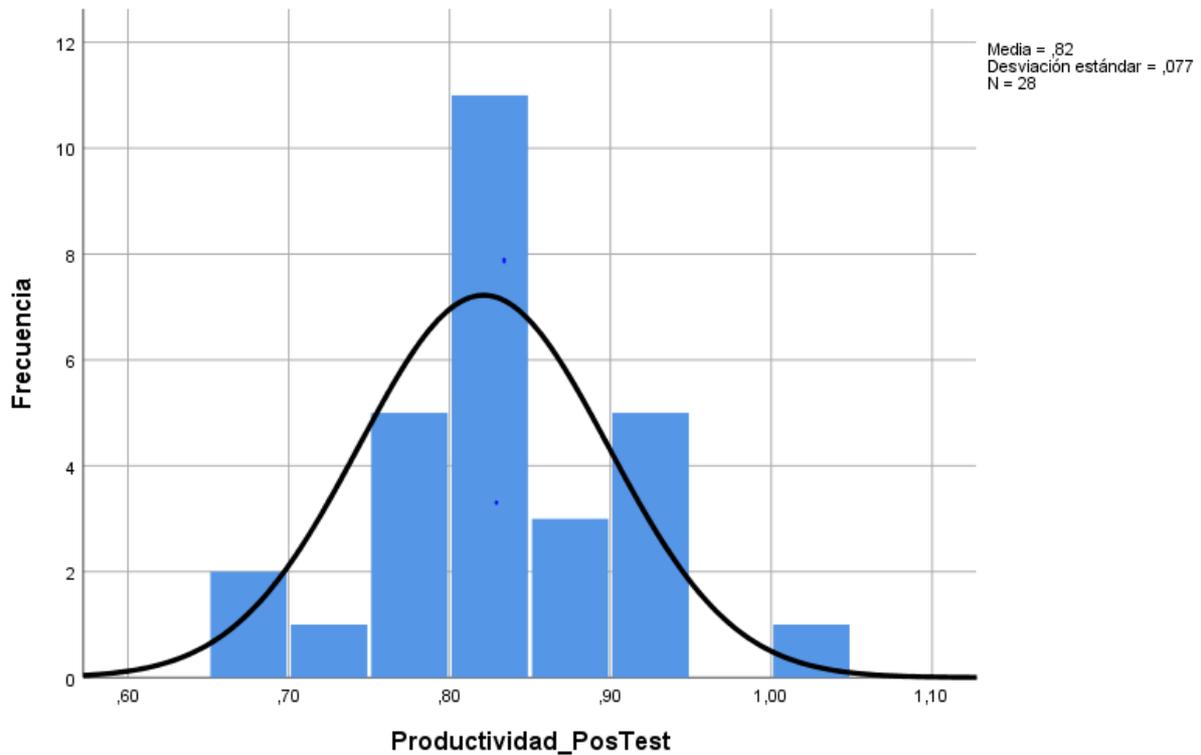
Como se muestra en la Tabla, los resultados de la prueba indican que el Sig. De los Productividad en el proceso de planificación en el Pre-Test fue de 0.101, cuyo valor es mayor que 0.05, por lo que indica que la productividad se distribuye normalmente. Los resultados de la prueba del Post-Test indican que el Sig. De la Productividad fue de 0.471, cuyo valor es mayor que 0.05, por lo que indica que la

Productividad se distribuye normalmente. Lo que confirma la distribución normal de ambos datos de la muestra, lo cual se puede apreciar en las Figuras.

**Tabla 15: Productividad pretest**



**Tabla 16: Productividad pos test**



### 3.3 Prueba de Hipótesis

**H1:** Si Machine Learning es confiable, entonces aumenta la eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

- **Indicador:** Eficiencia

#### Hipótesis Estadísticas

##### Definiciones de Variables:

- **EFa:** Eficiencia antes de utilizar Machine Learning.
- **EFd:** Eficiencia después de utilizar Machine Learning.

**H0:** Machine Learning no aumenta la Eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

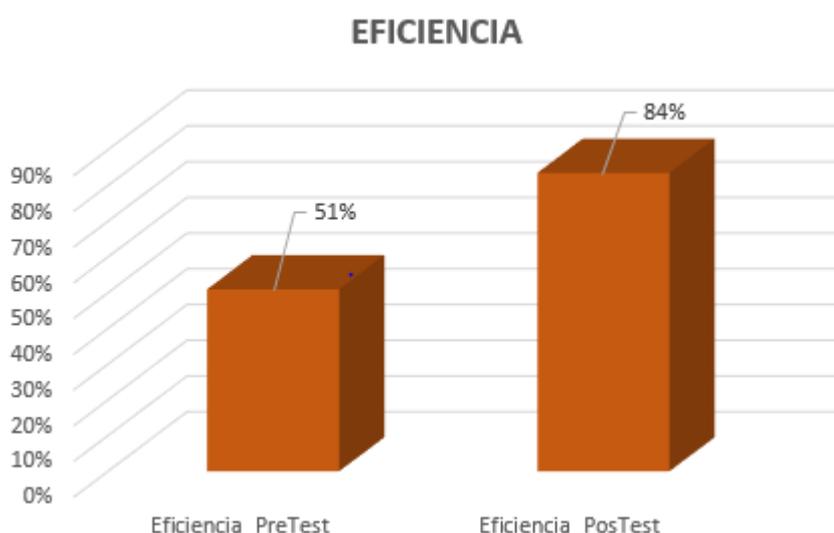
$$H0: EFa \geq EFd$$

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

**HA:** Machine Learning aumenta la Eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

Ha: EFa < EFd

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.



Se concluye de la Figura 21 existe un incremento en la eficiencia, el cual se puede verificar al comparar las medias respectivas, que asciende de 51% al valor de 84%.

En cuanto al resultado del contraste de hipótesis se aplicó la Prueba T-Student, debido a que los datos obtenidos durante la investigación (Pre-Test y Post-Test) se distribuyen normalmente. El valor de T contraste es de -10,403, el cual es claramente menor que 1.7033.

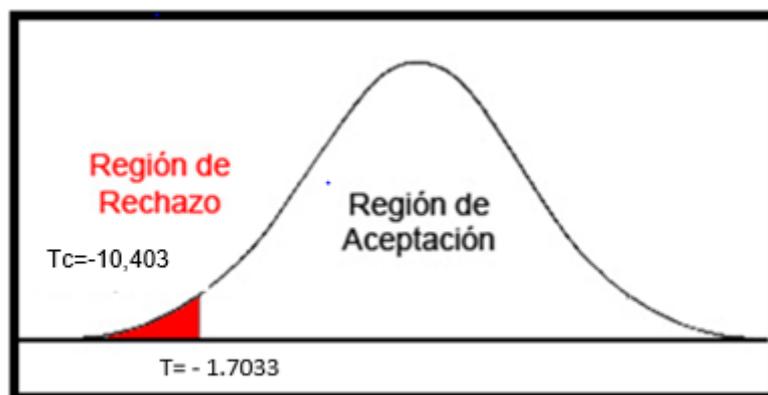
**Tabla 17: T student Eficiencia**

	Media	Prueba de T-Student		
		T	gl	Sig. (bilateral)
Eficiencia_pretest	,5096	-10,403	27	0.000
Eficiencia_postest	,8354			

Aplicando la formula T student.

$$T_c = \frac{x - u}{S / \sqrt{n}}$$
$$T_c = \frac{0,5096 - 0,8354}{0,16567 \sqrt{28}}$$
$$T_c = \frac{-0.3254}{0,16567 / 5.2915}$$
$$T_c = \frac{-0.3254}{0.0313}$$
$$T_c = -10,403$$

Entonces, se rechaza la hipótesis nula, aceptando la hipótesis alterna con un 95% de confianza. Además, el valor T obtenido, como se muestra en la Figura, se ubica en la zona de rechazo. Por lo tanto, Machine Learning aumenta la Eficiencia de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.



## Hipótesis de Investigación 2:

**H2:** Si Machine Learning es confiable, entonces aumenta la Productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

- **Indicador:** Productividad

### Hipótesis Estadísticas

#### Definiciones de Variables:

- **PRa:** Productividad antes de utilizar Machine Learning.
- **PRd:** Productividad después de utilizar Machine Learning.
- **H0:** Machine Learning no aumenta la Productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

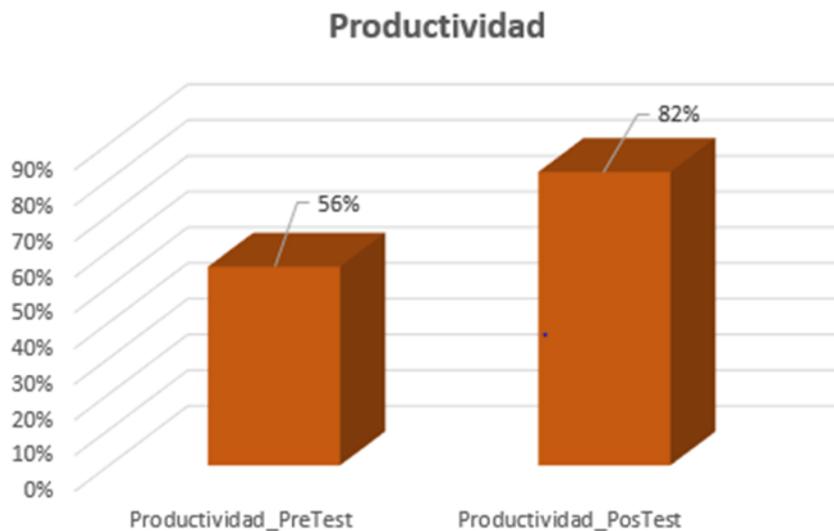
$$H_0: PRa \geq PRd$$

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

**HA:** Machine Learning aumenta la Productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

$$H_a: PRa < PRd$$

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.



Se concluye de la Figura 23 existe un incremento en la productividad, el cual se puede verificar al comparar las medias respectivas, que asciende de 56% al valor de 82%

En cuanto al resultado del contraste de hipótesis se aplicó la Prueba T-Student, debido a que los datos obtenidos durante la investigación (Pre-Test y Post-Test) se distribuyen normalmente. El valor de T contraste es de -11,147, el cual es claramente menor que 1.7033.

**Tabla 18: T student Productividad**

	Media	Prueba de T-Student		
		T	gl	Sig. (bilateral)
Productividad_pretest	,5561			
Productividad_posttest	,8211	-11,147	27	0.000

Aplicando la formula

$$T_c = \frac{x - u}{S / \sqrt{n}}$$

$$T_c = \frac{0,5561 - 0,8211}{0,12580 \sqrt{28}}$$

$$T_c = \frac{-0,265}{0,12580 / 5.2915}$$

$$T_c = \frac{-0,265}{0,02377}$$

$$T_c = -11,147$$

Entonces, se rechaza la hipótesis nula, aceptando la hipótesis alterna con un 95% de confianza. Además el valor T obtenido, como se muestra en la Figura, se ubica en la zona de rechazo. Por lo tanto, Machine Learning aumenta la Productividad de la planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.



## **CAPÍTULO IV**

### **DISCUSIÓN**

En la presente investigación, se tuvo un resultado que el Machine Learning aumenta la eficiencia de un 47% a un 79% y en la productividad de un 53% a un 79% en un plazo establecido por el proceso de planificación.

De la misma manera Sandra Zoraida, en su investigación Modelo de minería de datos usando machine learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016, llegó a concluir que el Machine Learning influye formidablemente en el aumento de diagnósticos en paciente de la ciudad de Trujillo con un 20.05% a 58.6%

Los resultados obtenidos en la presente investigación comprueba que la utilización del Machine Learning es una herramienta que dio resultados positivos en las predicciones cumpliendo con el objetivo de mejorar el servicio de las llamadas, abriendo puertas para ampliar la investigación y no solo centrarse en una área sino también en áreas específicas de alta gerencia.

## **CAPÍTULO V**

### **CONCLUSIONES**

Las conclusiones obtenidas en la presente tesis son las siguientes:

**PRIMERO:** Se concluye que el Machine Learning aumento el porcentaje de eficiencia a un 79% Teniendo inicialmente un 47,05% Por lo tanto se afirma que el Machine Learning aumenta la eficiencia en el servicio de atención de llamadas en la empresa Konecta BTO, SL, sucursal en Perú.

**SEGUNDO:** Se concluye que el Machine Learning aumento el porcentaje de Productividad a un 79% Teniendo inicialmente un 53,05% Por lo tanto se afirma que el Machine Learning aumenta la Productividad en el servicio de atención de llamadas en la empresa Konecta BTO, SL, sucursal en Perú.

**TERCERO:** Se concluye que el Machine Learning Si Machine Learning es confiable, para mejorar la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú.

## VI. RECOMENDACIONES

En base a las conclusiones y objetivos nos permitimos recomendar lo siguiente:

**Primero:**

Se debe fomentar la utilización de Machine Learning, para ello se pueden diseñar estrategias que nos permita dar el acceso a la población a tecnologías y herramientas para el aprendizaje autodidacta.

**Segunda:**

Promover la capacitación a empleados de herramientas tecnológicas para el aprendizaje y se puedan desarrollar como profesional.

**Tercera:**

Establecer alianzas estratégicas o convenios con entidades de cualquier sector que puedan brindar la posibilidad de fomentar el desarrollo de Machine Learning.

**Cuarta:**

Promover e influenciar en la evolución constante de las aptitudes y capacidades de cualquier persona que este involucrado en temas de Machine Learning, Ciencia de datos, Big Data, etc.

## IV REFERENCIAS

Almuiñas Rivero, José Luis, Galarza López, Judith, EL PROCESO DE PLANIFICACIÓN ESTRATÉGICA EN LAS UNIVERSIDADES: DESENCUENTROS Y RETOS PARA EL MEJORAMIENTO DE SU CALIDAD. Revista Gestão Universitária América Latina - GUAL [en línea] 2012, (agosto): [Consultado en octubre de 2018] Disponible en: <http://www.redalyc.org/articulo.oa?id=319327515006>

Almuiñas Rivero, José Luis, Galarza López, Judith, PROCESO DE PLANIFICACIÓN: ETAPAS EJECUTADAS EN PEQUEÑAS Y MEDIANAS EMPRESAS PARA OPTIMIZAR LA COMPETITIVIDAD. Revista Espacios [en línea] 2017, (agosto): [Consultado en octubre de 2018] Disponible en: <http://www.revistaespacios.com/a17v38n52/a17v38n52p16.pdf>

Medrano Parado, Sandra Zoraida (2016). Modelo de minería de datos usando Machine Learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016 [en línea] [Consulta: 01 de Octubre del 2018] Disponible en: <http://repositorio.ucv.edu.pe/handle/UCV/9852>

Beltrán Castañón, Norman (2013). Modelo para proyección de balances energéticos nacionales del Perú según la estructura insumo-producto; Aplicando redes neuronales [en línea] [Consulta: 01 de octubre del 2018] Disponible en: <http://repositorio.unap.edu.pe/handle/UNAP/3512>

Carrillo Ucuñay, Melissa y Vites Ángeles (2015). Aplicación de una red neuronal para el control domótica de encendido de luces en una vivienda de la urbanización Santa Angela- Chiclayo 2015 [en línea] [Consulta: 01 de octubre del 2018] Disponible en: <http://repositorio.unprg.edu.pe/handle/UNPRG/477>

Luna Domínguez, Ernesto (2015). Redes neuronales aplicando en la predicción del signo de los fondos de AFP CUPRUM [en línea] [Consulta: 01 de octubre del 2018] Disponible en: <http://repositorio.unprg.edu.pe/handle/UNPRG/477>

Ibarra, José y Kimberly Paredes (2018). Redes neuronales artificiales para el control de acceso basado en reconocimiento facial [en línea] [Consulta: 13 de octubre del 2018] Disponible en: [www.revistapuce.edu.ec/index.php/revpuce/article/download/140/242](http://www.revistapuce.edu.ec/index.php/revpuce/article/download/140/242)

García Gazabon, Gisela (2014) Modelo de Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos – Cartagena (2014) [en línea] [Consulta: Noviembre del 2014]. <https://biblioteca.utb.edu.co/notas/tesis/0068210.pdf>

Ján Vojt (2016). Deep neural networks and their implementation [en línea] [Consulta: 10 de octubre del 2018] Disponible en: <https://is.cuni.cz/webapps/zzp/download/120229251>

GIRONÉS Jordi, CASAS Jordi, MINGUILLÓN Julià, RAMON Alfonso. Minería de datos: modelos y algoritmos. Barcelona España: Editorial UOC, 2017 [Consulta: 10 de octubre de 2018]. ISBN: 978-84-9116-904-8

Almuiñas Rivero, José Luis, Galarza López, Judith, EL PROCESO DE PLANIFICACIÓN ESTRATÉGICA EN LAS UNIVERSIDADES: DESENCUENTROS Y RETOS PARA EL MEJORAMIENTO DE SU CALIDAD. Revista

Gestão Universitária América Latina - GUAL [en línea] 2012, (agosto): [Consultado en octubre de 2018] Disponible en: <http://www.redalyc.org/articulo.oa?id=319327515006>

Dionisio, del Rio Sadornil. Diccionario – Glosario de la metodología de la investigación Social. Madrid España: Editorial UNED, 2013 [Consulta: 10 de octubre de 2018]. ISBN: 978-84-362-6803-4

José, Maria Pareja, Gerentes en Perú solo dedican hasta 15% de su tiempo al planeamiento en sus empresas. Revista Peruana Gestión [en línea] 2015, (junio): [Consultado en octubre de 2018] Disponible en: <https://gestion.pe/tendencias/management-empleo/haygroup-gerentes-peru-dedican-15-planeamiento-empresas-92816>

Almuiñas Rivero, José Luis, Galarza López, Judith, PROCESO DE PLANIFICACIÓN: ETAPAS EJECUTADAS EN PEQUEÑAS Y MEDIANAS EMPRESAS PARA OPTIMIZAR LA COMPETITIVIDAD. Revista Espacios [en línea] 2017, (agosto): [Consultado en octubre de 2018] Disponible en: <http://www.revistaespacios.com/a17v38n52/a17v38n52p16.pdf>

(2016). Modelo de minería de datos usando Machine Learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016 [en línea] [Consulta:: 01 de Octubre del 2018] Disponible en: <http://repositorio.ucv.edu.pe/handle/UCV/9852>

Beltrán Castañón, Norman (2013). Modelo para proyección de balances energéticos nacionales del Perú según la estructura insumo-producto; Aplicando redes neuronales [en línea] [Consulta:: 01 de octubre del 2018] Disponible en: <http://repositorio.unap.edu.pe/handle/UNAP/3512>

Carrillo Ucuñay, Melissa y Vites Ángeles (2015). Aplicación de una red neuronal para el control domótica de encendido de luces en una vivienda de la urbanización Santa Angela- Chiclayo 2015 [en línea] [Consulta:: 01 de octubre del 2018] Disponible en: <http://repositorio.unprg.edu.pe/handle/UNPRG/477>

Luna Domínguez, Ernesto (2015). Redes neuronales aplicando en la predicción del signo de los fondos de AFP CUPRUM [en línea] [Consulta:: 01 de octubre del 2018] Disponible en: <http://repositorio.unprg.edu.pe/handle/UNPRG/477>

Ibarra, José y Kimberly Paredes (2018). Redes neuronales artificiales para el control de acceso basado en reconocimiento facial [en línea] [Consulta:: 13 de octubre del 2018] Disponible en: [www.revistapuce.edu.ec/index.php/revpuce/article/download/140/242](http://www.revistapuce.edu.ec/index.php/revpuce/article/download/140/242)

García Gazabon, Gisela (2014) Modelo de Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos – Cartagena (2014) [en línea] [Consulta:: Noviembre del 2014]. <https://biblioteca.utb.edu.co/notas/tesis/0068210.pdf>

Ján Vojt (2016). Deep neural networks and their implementation [en línea] [Consulta:: 10 de octubre del 2018] Disponible en: <https://is.cuni.cz/webapps/zzp/download/120229251>

GIRONÉS Jordi, CASAS Jordi, MINGUILLÓN Julià, RAMON Alfonso. Minería de datos: modelos y algoritmos. Barcelona España: Editorial UOC, 2017 [Consulta:: 10 de octubre de 2018]. ISBN: 978-84-9116-904-8

Hernández, R. Fernández, C. y Baptista, P. (2010). Metodología de la Investigación (5.a ed.). México: McGraw-Hill/Interamericana Editores S.A (p.408) ISBN: 978-968-460-950-0

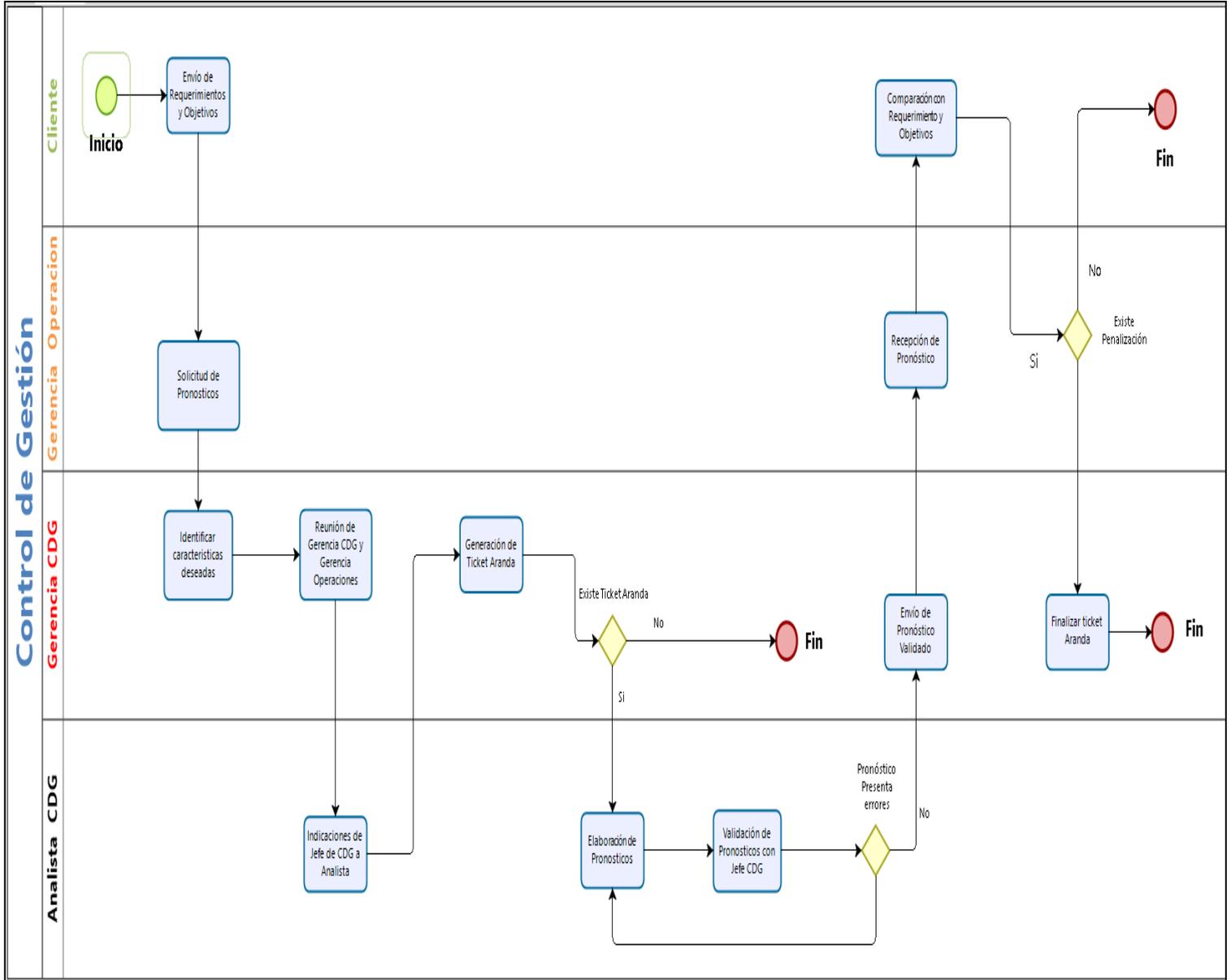
Dionisio, del Rio Sadornil. Diccionario – Glosario de la metodología de la investigación Social. Madrid España: Editorial UNED, 2013 [Consulta:: 10 de octubre de 2018]. ISBN: 978-84-362-6803-4

## **ANEXOS**



ANEXOS

Anexo 2. Diagrama de Proceso.



### Anexo 3. Árbol de Problemas



## Anexo 5. Tabla de Evaluación de Experto



UNIVERSIDAD CÉSAR VALLEJO

### TABLA DE EVALUACIÓN DE EXPERTOS

Apellidos y nombres del experto:

Champe Aguirre, Juan

Título y/o Grado

Doctor...()	Magister <input checked="" type="checkbox"/>	Ingeniero...()	Otros... Especifique
-------------	--	----------------	-------------------------

Universidad que labora: Universidad Cesar Vallejo sede Lima Norte

Fecha: 07/11/2018

### TÍTULO DE PROYECTO

MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS DE LA EMPRESA KONECTA, S.L. SUCURSAL EN PERÚ

Mediante la tabla de evaluación de expertos usted tiene la facultad de evaluar los 3 métodos propuestos para el desarrollo del sistema, mediante una serie de criterios con puntuaciones específicas al final de la tabla. Se debe colocar el puntaje correspondiente de cada criterio y seguidamente sumar puntaje.

Evaluar con la siguiente puntuación:

Muy bueno (5)      Bueno (4)      Regular (3)      Malo (2)      Muy (1)

ITEMS	CRITERIOS	CEISF DM	KDD	SEMMA
1	El modelo adecuado para el proyecto de corta duración	5	4	4
2	El modelo es adecuado para los requerimientos del usuario para la implementación propuesta	5	5	4
3	El modelo es flexible a los cambios de proyecto de investigación	5	5	4
4	El modelo involucra al cliente durante etapas del proyecto	4	5	4
5	Riesgos y costos son controlables e medibles que surge los requisitos y limitaciones	4	4	4
6	El modelo involucra al cliente como parte del equipo de desarrollo	4	4	4
7	El modelo que se adecua a la documentación (manera) al proyecto de investigación	4	4	4
TOTAL		31	31	31

SUGERENCIAS:



TABLA DE EVALUACION DE EXPERTOS

Apellidos y nombres del experto:

Ms. Huante Zojana Paul

Título y/o Grado

Doctor...()	Magister... <i>MA</i>	Ingeniero...()	Otros..... Especifique
-------------	-----------------------	----------------	---------------------------

Universidad que labora: Universidad Cesar Vallejo sede Lima Norte

Fecha *09/11/18*

TITULO DE PROYECTO

MACHINE LEARNING PARA LA PLANIFICACION EN EL SERVICIO DE ATENCION DE LLAMADAS DE LA EMPRESA KONNECTA S.O. S.L. SUCURSAL EN PERU

Mediante la tabla de evaluación de expertos usted tiene la facultad de evaluar los 3 métodos propuestas para el desarrollo del sistema, mediante una serie de criterios con puntuaciones específicas al final de la tabla. Se debe colocar el puntaje correspondiente de cada criterio y seguidamente sumar puntaje.

Evaluar con la siguiente puntuación:

Muy bueno (5)      Bueno (4)      Regular (3)      Malo (2)      Muy (1)

ITEMS	CRITERIOS	CRISP DM	KDD	SEMMA
1	El modelo adecuado para el proyecto de corta duración	<i>5</i>	<i>1</i>	<i>5</i>
2	El modelo es adecuado para los requerimientos del usuario para la implementación propuesta	<i>5</i>	<i>4</i>	<i>5</i>
3	El modelo es flexible a los cambios de proyecto de investigación	<i>5</i>	<i>4</i>	<i>4</i>
4	El modelo involucra al cliente durante etapas del proyecto	<i>5</i>	<i>4</i>	<i>4</i>
5	Riesgos y costos son controlados a medida que surge los requisitos y limitaciones	<i>5</i>	<i>4</i>	<i>5</i>
6	El modelo involucra al cliente como parte del equipo de desarrollo	<i>5</i>	<i>4</i>	<i>5</i>
7	El modelo que se adecua a la documentación (menor) al proyecto de investigación	<i>5</i>	<i>4</i>	<i>5</i>
<b>TOTAL</b>		<i>35</i>	<i>28</i>	<i>33</i>

SUGERENCIAS

*[Handwritten signature]*



TABLA DE EVALUACION DE EXPERTOS

Apellidos y nombres del experto:

Calle Tapia Oleas

Título y/o Grado

Doctor...()	Magister...()	Ingeniero...()	Otros..... Especifique
-------------	---------------	----------------	---------------------------

Universidad que labora: Universidad Cesar Vallejo ande Lima Norte

Fecha 9/11/2018

TITULO DE PROYECTO

MACHINE LEARNING PARA LA PLANIFICACION EN EL SERVICIO DE ATENCION DE LLAMADAS DE LA EMPRESA KONNECTA SDO. S.L. SURCUAL EN PERU

Mediante la tabla de evaluación de expertos usted tiene la facultad de evaluar los 3 métodos propuestos para el desarrollo del sistema, mediante una serie de criterios con puntuaciones específicas al final de la tabla. Se debe colocar el puntaje correspondiente de cada criterio y seguidamente sumar puntaje.

Evaluar con la siguiente puntuación:

Muy bueno (5) Buono (4) Regular (3) Malo (2) Muy (1)

ITEMS	CRITERIOS	CRISP DM	KDD	SEMMA
1	El modelo adecuado para el proyecto de corta duración	5	4	3
2	El modelo es adecuado para los requerimientos del usuario para la implementación propuesta	5	5	3
3	El modelo es flexible a los cambios de proyecto de investigación	5	4	3
4	El modelo involucra al cliente durante etapas del proyecto	5	4	3
5	Riesgos y costos son controlados o medida que surge los requisitos y limitaciones	5	4	3
6	El modelo involucra al cliente como parte del equipo de desarrollo	5	4	3
7	El modelo que se adecua a la documentación (planes) al proyecto de investigación	5	4	3
TOTAL				

SUGERENCIAS

*[Handwritten signature]*

## Anexo 6. Tabla de Evaluación de Indicadores



UNIVERSIDAD CÉSAR VALLEJO

Evaluación de Experto de Indicadores

Apellidos y Nombres del Experto: Huante Laguna Raúl

Fecha: 07/11/18

Título de Proyecto

MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS DE LA EMPRESA KONECTA BTO, S.L. SUCURSAL EN PERÚ

Mediante la tabla de evaluación de expertos, usted tiene la facultad para medir el indicador, mediante una serie de preguntas con puntuaciones especificada al inicio de la tabla. Asimismo, le exhortamos en la corrección de ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de preguntas.

0 – 20 %	21 – 50 %	51 – 70 %	71 – 80 %	81 – 100 %
Deficiente	Regular	Buena	Muy buena	Excelente

ITEMS	PREGUNTAS	VALORACIÓN				
		0 – 20 %	21 – 50 %	51 – 70 %	71 – 80 %	81 – 100 %
1	¿Cumple con el diseño adecuado?					82
2	¿Tiene relación con el título de la investigación?					83
3	¿Se mencionan las variables de la investigación?					85
4	¿Facilitará el logro de los objetivos de investigación?					82
5	¿Se relaciona con la variable de estudio?					90
6	¿Facilitará el análisis y procesamiento de datos?					89
7	¿Son entendible sus alternativas de respuestas?					85
8	¿Será accesible a la población sujeto de estudio?					88
9	¿Es claro, preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					90
10	¿Los indicadores son los adecuados para la medición?					85

Firma



# UNIVERSIDAD CÉSAR VALLEJO

## Evaluación de Experto de Indicadores

Apellidos y Nombres del Experto: Galvez Tapia Orleaus

Fecha:

Título de Proyecto

**MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS DE LA EMPRESA KONECTA BTO. S.L. SUCURSAL EN PERÚ**

Indicador: Productividad = Producción Real / Horas Trabajadas

Mediante la tabla de evaluación de expertos, usted tiene la facultad para medir el indicador, mediante una serie de preguntas con puntuaciones especificada al inicio de la tabla. Asimismo, le exhortamos en la corrección de items indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de preguntas.

0 – 20 % Deficiente	21 – 50 % Regular	51 – 70 % Bueno	71 – 80 % Muy bueno	81 – 100 % Excelente
------------------------	----------------------	--------------------	------------------------	-------------------------

ITEMS	PREGUNTAS	VALORACIÓN				
		0 – 20 %	21 – 50 %	51 – 70 %	71 – 80 %	81 – 100 %
1	¿Cumple con el diseño adecuado?				80	
2	¿Tiene relación con el título de la investigación?				80	
3	¿Se mencionan las variables de la investigación?				80	
4	¿Facilitará el logro de los objetivos de investigación?				80	
5	¿Se relaciona con la variable de estudio?				80	
6	¿Facilitará el análisis y procesamiento de datos?				80	
7	¿Son entendible sus alternativas de respuestas?				80	
8	¿Será accesible a la población sujeto de estudio?				80	
9	¿Es claro, preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?				80	
10	¿Los indicadores son los adecuados para la medición?				80	

Firma



# UNIVERSIDAD CÉSAR VALLEJO

## Evaluación de Experto de Indicadores

Apellidos y Nombres del Experto: Champe Aguilera, J. G.

Fecha: 07/11/2018

### Título de Proyecto

**MACHINE LEARNING PARA LA PLANIFICACIÓN EN EL SERVICIO DE ATENCIÓN DE LLAMADAS DE LA EMPRESA KONECTA BTO, S.L. SUCURSAL EN PERÚ**

**Indicador: Eficiencia = Producción Real / Producción Planificada**

Mediante la tabla de evaluación de expertos, usted tiene la facultad para medir el indicador, mediante una serie de preguntas con puntuaciones especificada al inicio de la tabla. Asimismo, le exhortamos en la corrección de ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de preguntas.

0 – 20 %	21 – 50 %	51 – 70 %	71 – 80 %	81 – 100 %
Deficiente	Regular	Bueno	Muy bueno	Excelente

ITEMS	PREGUNTAS	VALORACIÓN				
		0 – 20 %	21 – 50 %	51 – 70 %	71 – 80 %	81 – 100 %
1	¿Cumple con el diseño adecuado?			70		
2	¿Tiene relación con el título de la investigación?			70		
3	¿Se mencionan las variables de la investigación?			70		
4	¿Facilitará el logro de los objetivos de investigación?			70		
5	¿Se relaciona con la variable de estudio?			70		
6	¿Facilitará el análisis y procesamiento de datos?			70		
7	¿Adaptado para estimar aspectos del sistema metodológico?			70		
8	¿Será accesible a la población sujeto de estudio?			70		
9	¿Entiende los aspectos de cantidad y claridad?			70		
10	¿Los indicadores son los adecuados para la medición?			70		

Firma:   
07/11/2018

Promedio 70.5

## Anexo 7. Ficha de Registro – Indicador Eficiencia Pres test

<b>Empresa</b>	Konecta BTO S.L. Sucursal en Perú
<b>Investigación</b>	Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú
<b>Indicador</b>	Eficiencia
<b>Dimensión</b>	Control y Evaluación

<u>Items</u>	Fecha	ID_Servicio	Servicios	Llamadas planificadas	Llamadas reales	Eficiencia
1	01/09/2018	1001	TECNICO NEOGCIO	3	7	0,43
2	02/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
3	03/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
4	04/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
5	05/09/2018	1001	TECNICO NEOGCIO	3	5	0,60
6	06/09/2018	1001	TECNICO NEOGCIO	2	4	0,50
7	07/09/2018	1001	TECNICO NEOGCIO	3	4	0,75
8	08/09/2018	1001	TECNICO NEOGCIO	6	12	0,50
9	09/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
10	10/09/2018	1001	TECNICO NEOGCIO	6	12	0,50
11	11/09/2018	1001	TECNICO NEOGCIO	6	13	0,46
12	12/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
13	13/09/2018	1001	TECNICO NEOGCIO	4	9	0,44
14	14/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
15	15/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
16	16/09/2018	1001	TECNICO NEOGCIO	2	3	0,67
17	17/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
18	18/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
19	19/09/2018	1001	TECNICO NEOGCIO	2	3	0,67
20	20/09/2018	1001	TECNICO NEOGCIO	2	6	0,33
21	21/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
22	22/09/2018	1001	TECNICO NEOGCIO	6	11	0,55
23	23/09/2018	1001	TECNICO NEOGCIO	4	10	0,40
24	24/09/2018	1001	TECNICO NEOGCIO	6	13	0,46
25	25/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
26	26/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
27	27/09/2018	1001	TECNICO NEOGCIO	6	14	0,43
28	28/09/2018	1001	TECNICO NEOGCIO	5	8	0,63

## Anexo 8. Ficha de Registro – Indicador Productividad Pres test

Empresa	Konecta BTO S.L. Sucursal en Perú
Investigación	Machine Learning para la Planificación en el servicio de atención de llamadas en la empresa Konecta BTO, S.L. Sucursal en Perú
Variable	Planificación
Dimensión	Control y Evaluación

Items	Fecha	ID_Servicio	Servicios	Horas Trabajadas	llamadas reales	Productividad
1	01/09/2018	1001	TECNICO NEOGCIO	3	7	0,43
2	02/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
3	03/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
4	04/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
5	05/09/2018	1001	TECNICO NEOGCIO	3	5	0,60
6	06/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
7	07/09/2018	1001	TECNICO NEOGCIO	3	4	0,75
8	08/09/2018	1001	TECNICO NEOGCIO	7	11	0,64
9	09/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
10	10/09/2018	1001	TECNICO NEOGCIO	3	7	0,43
11	11/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
12	12/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
13	13/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
14	14/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
15	15/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
16	16/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
17	17/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
18	18/09/2018	1001	TECNICO NEOGCIO	4	7	0,57
19	19/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
20	20/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
21	21/09/2018	1001	TECNICO NEOGCIO	4	8	0,50
22	22/09/2018	1001	TECNICO NEOGCIO	6	9	0,67
23	23/09/2018	1001	TECNICO NEOGCIO	3	5	0,60
24	24/09/2018	1001	TECNICO NEOGCIO	3	5	0,60
25	25/09/2018	1001	TECNICO NEOGCIO	2	5	0,40
26	26/09/2018	1001	TECNICO NEOGCIO	4	9	0,44
27	27/09/2018	1001	TECNICO NEOGCIO	6	11	0,55
28	28/09/2018	1001	TECNICO NEOGCIO	4	8	0,50

## Anexo 9. Ficha de Registro – Indicador Productividad Post test

Items	Fecha	ID_Servicio	Servicios	Horas Trabajadas	Llamadas reales	Productividad
1	01/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
2	02/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
3	03/05/2018	1001	TECNICO NEOGCIO	10	11	0,91
4	04/05/2018	1001	TECNICO NEOGCIO	10	11	0,91
5	05/05/2018	1001	TECNICO NEOGCIO	10	11	0,91
6	06/05/2018	1001	TECNICO NEOGCIO	10	11	0,91
7	07/05/2018	1001	TECNICO NEOGCIO	4	5	0,80
8	08/05/2018	1001	TECNICO NEOGCIO	6	8	0,75
9	09/05/2018	1001	TECNICO NEOGCIO	6	8	0,75
10	10/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
11	11/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
12	12/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
13	13/05/2018	1001	TECNICO NEOGCIO	4	5	0,80
14	14/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
15	15/05/2018	1001	TECNICO NEOGCIO	6	8	0,75
16	16/05/2018	1001	TECNICO NEOGCIO	6	9	0,67
17	17/05/2018	1001	TECNICO NEOGCIO	9	10	0,90
18	18/05/2018	1001	TECNICO NEOGCIO	7	8	0,88
19	19/05/2018	1001	TECNICO NEOGCIO	6	9	0,67
20	20/05/2018	1001	TECNICO NEOGCIO	4	4	1,00
21	21/05/2018	1001	TECNICO NEOGCIO	10	13	0,77
22	22/05/2018	1001	TECNICO NEOGCIO	5	6	0,83
23	23/05/2018	1001	TECNICO NEOGCIO	4	5	0,80
24	24/05/2018	1001	TECNICO NEOGCIO	7	8	0,88
25	25/05/2018	1001	TECNICO NEOGCIO	5	7	0,71
26	26/05/2018	1001	TECNICO NEOGCIO	4	5	0,80
27	27/05/2018	1001	TECNICO NEOGCIO	6	7	0,86
28	28/05/2018	1001	TECNICO NEOGCIO	6	8	0,75

## Anexo 10. Ficha de Registro – Indicador Eficiencia Post test

Items	Fecha	ID_Servicio	Servicios	Llamadas planificadas	Llamadas reales	Eficiencia
1	01/05/2018	1001	TECNICO NEOGCIO	6	5	0,83
2	02/05/2018	1001	TECNICO NEOGCIO	10	9	0,90
3	03/05/2018	1001	TECNICO NEOGCIO	11	10	0,91
4	04/05/2018	1001	TECNICO NEOGCIO	11	10	0,91
5	05/05/2018	1001	TECNICO NEOGCIO	11	10	0,91
6	06/05/2018	1001	TECNICO NEOGCIO	11	10	0,91
7	07/05/2018	1001	TECNICO NEOGCIO	10	8	0,80
8	08/05/2018	1001	TECNICO NEOGCIO	8	6	0,75
9	09/05/2018	1001	TECNICO NEOGCIO	8	6	0,75
10	10/05/2018	1001	TECNICO NEOGCIO	10	8	0,80
11	11/05/2018	1001	TECNICO NEOGCIO	13	11	0,85
12	12/05/2018	1001	TECNICO NEOGCIO	6	5	0,83
13	13/05/2018	1001	TECNICO NEOGCIO	5	4	0,80
14	14/05/2018	1001	TECNICO NEOGCIO	6	5	0,83
15	15/05/2018	1001	TECNICO NEOGCIO	12	9	0,75
16	16/05/2018	1001	TECNICO NEOGCIO	12	8	0,67
17	17/05/2018	1001	TECNICO NEOGCIO	10	9	0,90
18	18/05/2018	1001	TECNICO NEOGCIO	8	7	0,88
19	19/05/2018	1001	TECNICO NEOGCIO	9	6	0,67
20	20/05/2018	1001	TECNICO NEOGCIO	5	5	1,00
21	21/05/2018	1001	TECNICO NEOGCIO	13	10	0,77
22	22/05/2018	1001	TECNICO NEOGCIO	10	9	0,90
23	23/05/2018	1001	TECNICO NEOGCIO	12	11	0,92
24	24/05/2018	1001	TECNICO NEOGCIO	8	8	1,00
25	25/05/2018	1001	TECNICO NEOGCIO	6	5	0,83
26	26/05/2018	1001	TECNICO NEOGCIO	9	8	0,89
27	27/05/2018	1001	TECNICO NEOGCIO	6	5	0,83
28	28/05/2018	1001	TECNICO NEOGCIO	10	6	0,60

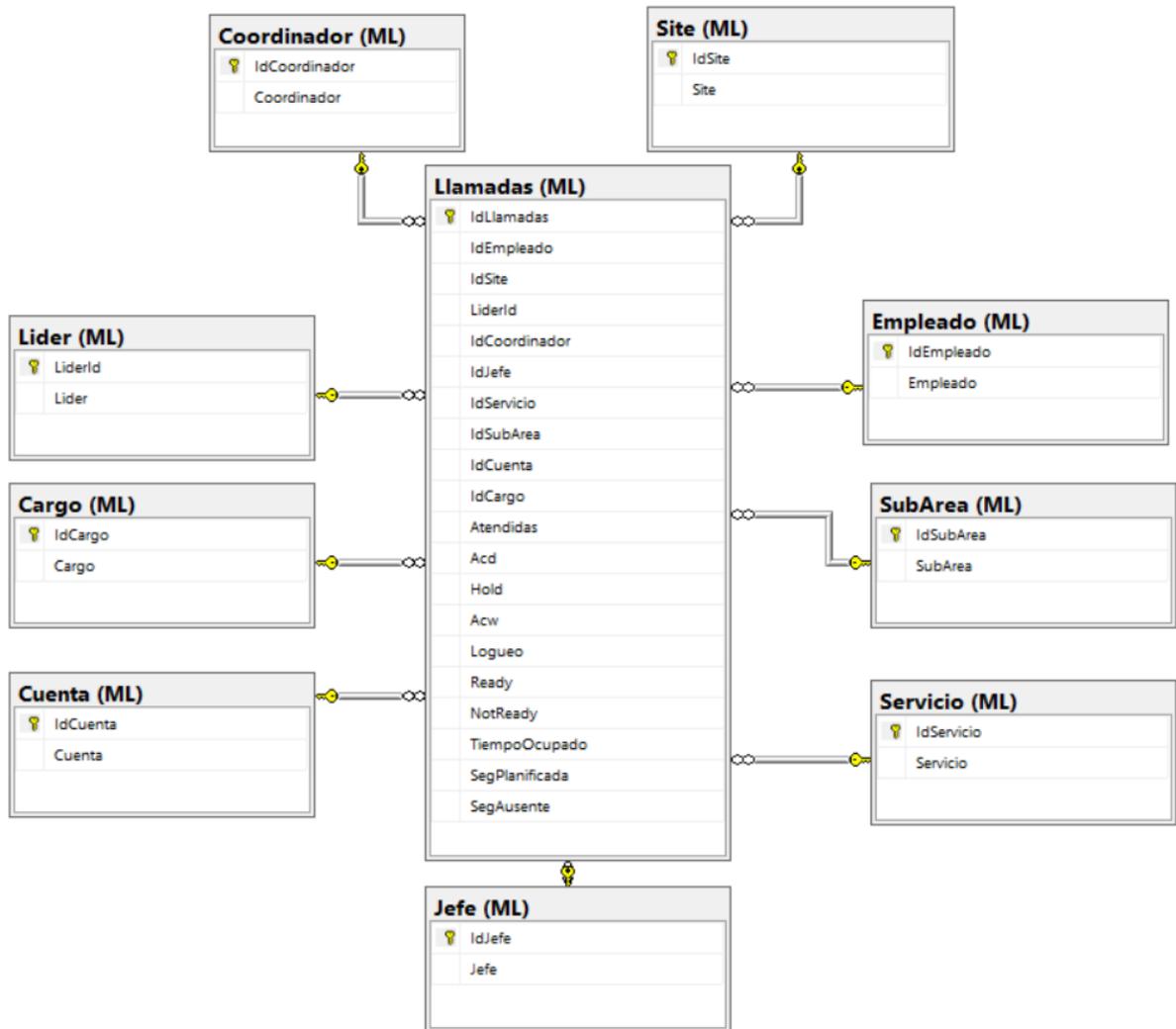
# Cronograma

Modo de tarea	Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
	<b>PROYECTO DE INVESTIGACIÓN</b>	<b>152 días?</b>	<b>mié 5/09/18</b>	<b>jue 4/04/19</b>	
	<b>Elección del tema de investigación</b>	<b>9 días</b>	<b>mié 5/09/18</b>	<b>lun 17/09/18</b>	
	Entrevista con el Gerente	1 día	mié 5/09/18	mié 5/09/18	
	Asignación de los temas de investigación	2 días	jue 6/09/18	vie 7/09/18	3
	Busqueda de la información	6 días	lun 10/09/18	lun 17/09/18	4
	<b>Planteamiento del problema</b>	<b>4 días</b>	<b>mar 18/09/18</b>	<b>vie 21/09/18</b>	
	Análisis y planteamiento de problema	4 días	mar 18/09/18	vie 21/09/18	5
	<b>Recopilación de información y redacción del informe</b>	<b>64 días?</b>	<b>lun 24/09/18</b>	<b>jue 20/12/18</b>	
	Busqueda de antecedentes	32 días	lun 24/09/18	mar 6/11/18	7
	Busqueda de bases teoricas	7 días	mié 7/11/18	jue 15/11/18	9
	Redacción de justificación	2 días	vie 16/11/18	lun 19/11/18	10
	Redacción hipótesis y objetivos	1 día	mar 20/11/18	mar 20/11/18	11
	Redacción de la metodología del estudio	2 días	mié 21/11/18	jue 22/11/18	12
	Redacción de variables y operacionalización	3 días	vie 23/11/18	mar 27/11/18	13
	Presentación del primer avance de informe	1 día	mié 28/11/18	mié 28/11/18	14
	Población y muestra	3 días	jue 29/11/18	lun 3/12/18	15
	Técnicas e instrumentos de recolección de datos	3 días	mar 4/12/18	jue 6/12/18	16
	Presentación del proyecto de investigación	1 día	vie 7/12/18	vie 7/12/18	17
	Levantamiento de observaciones	8 días	lun 10/12/18	mié 19/12/18	18
	Sustentación	1 día?	jue 20/12/18	jue 20/12/18	19
	<b>VACACIONES</b>	<b>75 días</b>	<b>vie 21/12/18</b>	<b>jue 4/04/19</b>	<b>20</b>
	Vacaciones	75 días	vie 21/12/18	jue 4/04/19	20
	<b>DESARROLLO DEL PROYECTO DE INVESTIGACIÓN</b>	<b>71 días</b>	<b>vie 5/04/19</b>	<b>vie 12/07/19</b>	
	<b>Investigación del Problema</b>	<b>4 días</b>	<b>vie 5/04/19</b>	<b>mié 10/04/19</b>	
	Análisis del problema	4 días	vie 5/04/19	mié 10/04/19	22
	<b>Recolectar Datos</b>	<b>10 días</b>	<b>jue 11/04/19</b>	<b>mié 24/04/19</b>	
	Investigación y obtención de datos	4 días	jue 11/04/19	mar 16/04/19	25
	Calidad y cantidad de datos	3 días	mié 17/04/19	vie 19/04/19	27
	Web scraping	3 días	lun 22/04/19	mié 24/04/19	28
	<b>Preparar los Datos</b>	<b>16 días</b>	<b>jue 25/04/19</b>	<b>jue 16/05/19</b>	
	Mezcla de datos	2 días	jue 25/04/19	vie 26/04/19	29
	Visualización de datos	2 días	lun 29/04/19	mar 30/04/19	31
	Revisar correlaciones de datos	2 días	mié 1/05/19	jue 2/05/19	32
	Selección de características	2 días	vie 3/05/19	lun 6/05/19	33
	Reducción de dimensiones aplicando PCA	2 días	mar 7/05/19	mié 8/05/19	34
	Balance de cantidad de datos	2 días	jue 9/05/19	vie 10/05/19	35
	Separación de datos (Entrenamiento y evaluación)	2 días	lun 13/05/19	mar 14/05/19	36
	Preproceso de datos - Normalización	2 días	mié 15/05/19	jue 16/05/19	37
	<b>Eligir Modelo</b>	<b>9 días</b>	<b>vie 17/05/19</b>	<b>mié 29/05/19</b>	
	Redes Neuronales	3 días	vie 17/05/19	mar 21/05/19	38
	Arboles de decisión	3 días	mié 22/05/19	vie 24/05/19	40
	Regresión Lineal	3 días	lun 27/05/19	mié 29/05/19	41
	<b>Entrenamiento de Maquina</b>	<b>10 días</b>	<b>jue 30/05/19</b>	<b>mié 12/06/19</b>	
	Set de datos de entrenamiento	2 días	jue 30/05/19	vie 31/05/19	42
	Inicializar los pesos	2 días	lun 3/06/19	mar 4/06/19	44
	Ajuste automático	2 días	mié 5/06/19	jue 6/06/19	45
	Revisar resultados obtenidos	2 días	vie 7/06/19	lun 10/06/19	46
	Corregir resultados obtenidos y volver a iterar	2 días	mar 11/06/19	mié 12/06/19	47
	<b>Evaluación</b>	<b>6 días</b>	<b>jue 13/06/19</b>	<b>jue 20/06/19</b>	
	Comprobar Maquina vs Set de datos de Evaluación	2 días	jue 13/06/19	vie 14/06/19	48
	Verificar precisión de modelo entrenado	2 días	lun 17/06/19	mar 18/06/19	50
	Validar exactitud de modelo	2 días	mié 19/06/19	jue 20/06/19	51
	<b>Parameter Tuning (Configuración de parámetros)</b>	<b>14 días</b>	<b>vie 21/06/19</b>	<b>mié 10/07/19</b>	
	Validación de precisión	2 días	vie 21/06/19	lun 24/06/19	52
	Validación de Learning Rate (taza de aprendizaje)	2 días	mar 25/06/19	mié 26/06/19	54
	Indicar máximo error permitido de modelo	2 días	jue 27/06/19	vie 28/06/19	55
	Ajuste de parametros	2 días	lun 1/07/19	mar 2/07/19	56
	Combinación de parametros	2 días	mié 3/07/19	jue 4/07/19	57
	Definición de cantidad de Hidden Layers (Capa Oculta)	2 días	vie 5/07/19	lun 8/07/19	58
	Definición de cantidad de Neuronas de Capa	2 días	mar 9/07/19	mié 10/07/19	59
	<b>Predicción o Inferencia</b>	<b>2 días</b>	<b>jue 11/07/19</b>	<b>vie 12/07/19</b>	
	Utilizar Modelo de Aprendizaje Automático	1 día	jue 11/07/19	jue 11/07/19	60
	Predecir o Inferir Resultados	1 día	vie 12/07/19	vie 12/07/19	62

### **Situación Actual:**

Actualmente en la empresa Konecta Perú del sector telecomunicaciones se realiza la planificación por llamadas que deben de ingresar, cuantas llamadas serán abandonadas, cuantos agentes deben estar conectados en cierto intervalo de hora, cuantos agentes deben estar disponibles para contestar las llamadas que van a ingresar, cuanto será el tiempo medio operativo (TMO), todos estos por intervalos de medias horas, día y mes, este proceso lo vienen realizando manualmente en hojas de cálculo Excel la cual demanda más horas de trabajo e incomodidad por parte del personal teniendo mayor probabilidad de errores en la planificación, una vez realizada la planificación son enviados a operaciones para que en base a ello puedan posicionar a los agentes y puedan gestionar, actualmente la planificación no viene siendo eficaz al momento de compararlo con la realidad por ende conlleva a una mala gestión en operaciones por parte de los agentes, perjudicando a los indicadores que tienen como objetivos, por lo tanto dada estas planificaciones no eficaces se pueden dar penalizaciones monetarias por parte del cliente a quienes se les brinda el servicio como ejemplo Diners Club, Calidda, Movistar Perú, Movistar Chile, Latam, Entel Perú, entre otras empresas, por otro lado estas planificaciones tienden a ser enviados fuera del plazo establecido la cual hace que se genere una mala toma de decisiones a nivel gerencial y directivo, tampoco se tiene la información actualizada debido a ello la gerencia y directiva estarían siendo penalizadas por parte de los clientes ya mencionados.

## Recolección de datos



### EMPLEADO:

- ID EMPLEADO
- EMPLEADO

### LLAMADA:

- IDLLAMADAS
- IDEMPLEADO
- IDSITE
- LIDERLD
- IDCOORDINADOR
- IDJEFE
- IDSERVICIO

- IDSUBAREA
- IDCUENTA
- IDCARGO
- ATENTIDAS
- ACD
- HOLD
- ACW
- LOGUEO
- READY
- NotREADY
- TiempoOcupado
- SegPLANIFICADA
- SegAUSENTE

**SERVICIO:**

- IDSERVICIO
- SERVICIO

**COORDINADOR:**

- IDCOORDINADOR
- COORDINADOR

**SITE:**

- IDSITE
- SITE

**SUBAREA:**

- IDSUBAREA
- SUBAREA

**SERVICIO:**

- IDSERVICIO
- SERVICIO

**JEFE:**

- IDJEFE
- JEFE

**CUENTA:**

- IDCUENTA
- CUENTA

**CARGO:**

- IDCARGO
- CARGO

**LIDER:**

- LIDERDL
- LIDER

**Descripción a las tablas a usar:**

TABLA	DESCRIPCION
EMPLEADO	Los datos generales del empleado
LLAMADAS	La atención de cumplimiento de llamadas
SERVICIO	Atención generadas por la persona
COORDINADOR	Los datos generales del coordinador
SITE	Posición de llamadas del empleado
SUBAREA	Área específica del servicio
JEFE	Los datos generales del jefe
LIDER	Datos generales del Responsable de Sub área
CARGO	Datos generales del cargo
CUENTA	Datos generales de la cuenta

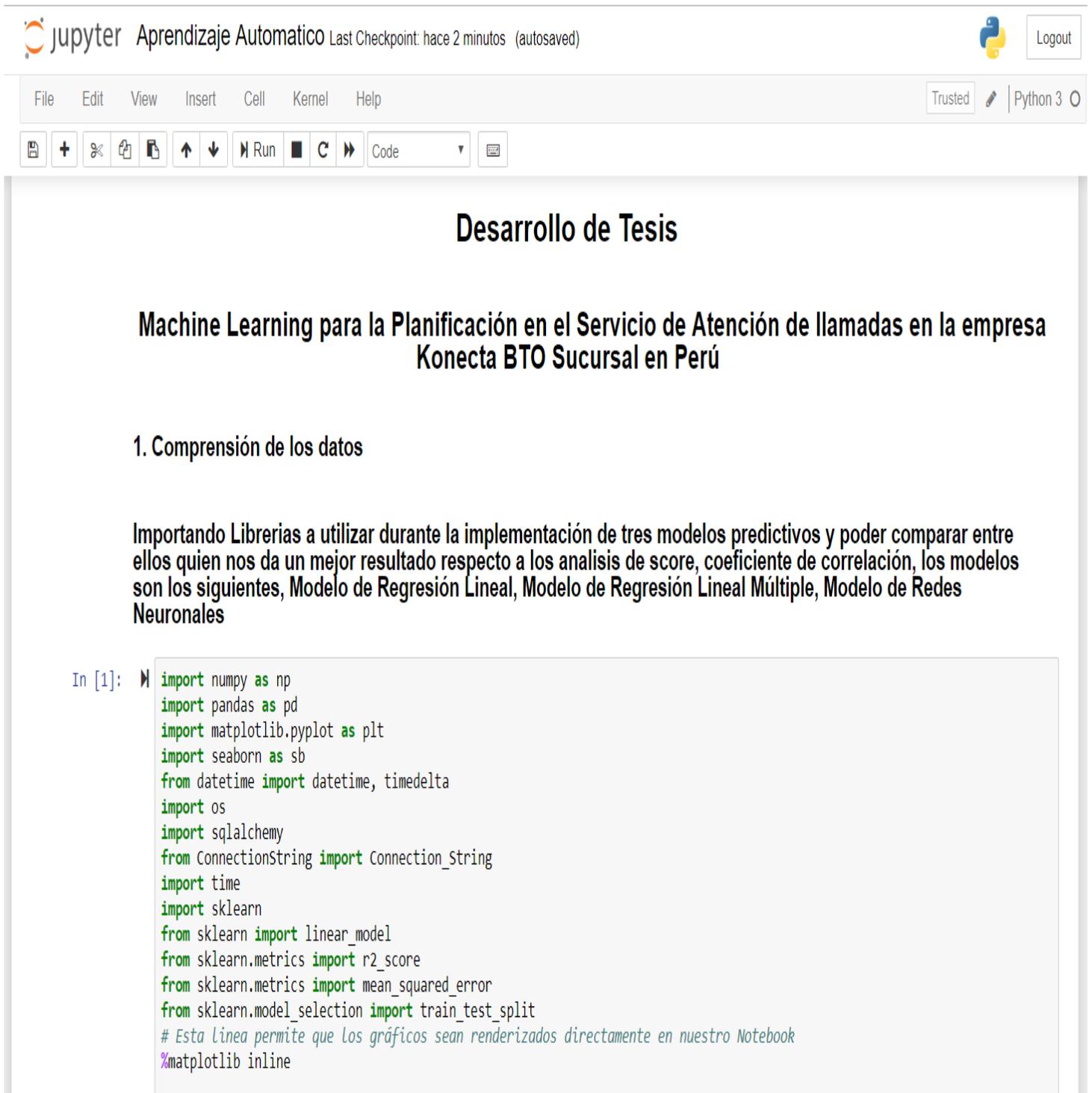
**Cantidad de registros de los datos muestras asignadas:**

TABLA	Nro. De registros
EMPLEADO	2141
LLAMADAS	550
SERVICIO	14
COORDINADOR	24
SITE	3
SUBAREA	2
JEFE	6
LIDER	142

CARGO	2
CUENTA	1

## Figura 12: preparación de los datos en el entorno jupyter

En la figura 12 Podemos observar la importación de las librerías a utilizar para la implementación de los modelos predictivos.



jupyter Aprendizaje Automatico Last Checkpoint: hace 2 minutos (autosaved)  Logout

File Edit View Insert Cell Kernel Help Trusted Python 3

Code

### Desarrollo de Tesis

#### Machine Learning para la Planificación en el Servicio de Atención de Llamadas en la empresa Konecta BTO Sucursal en Perú

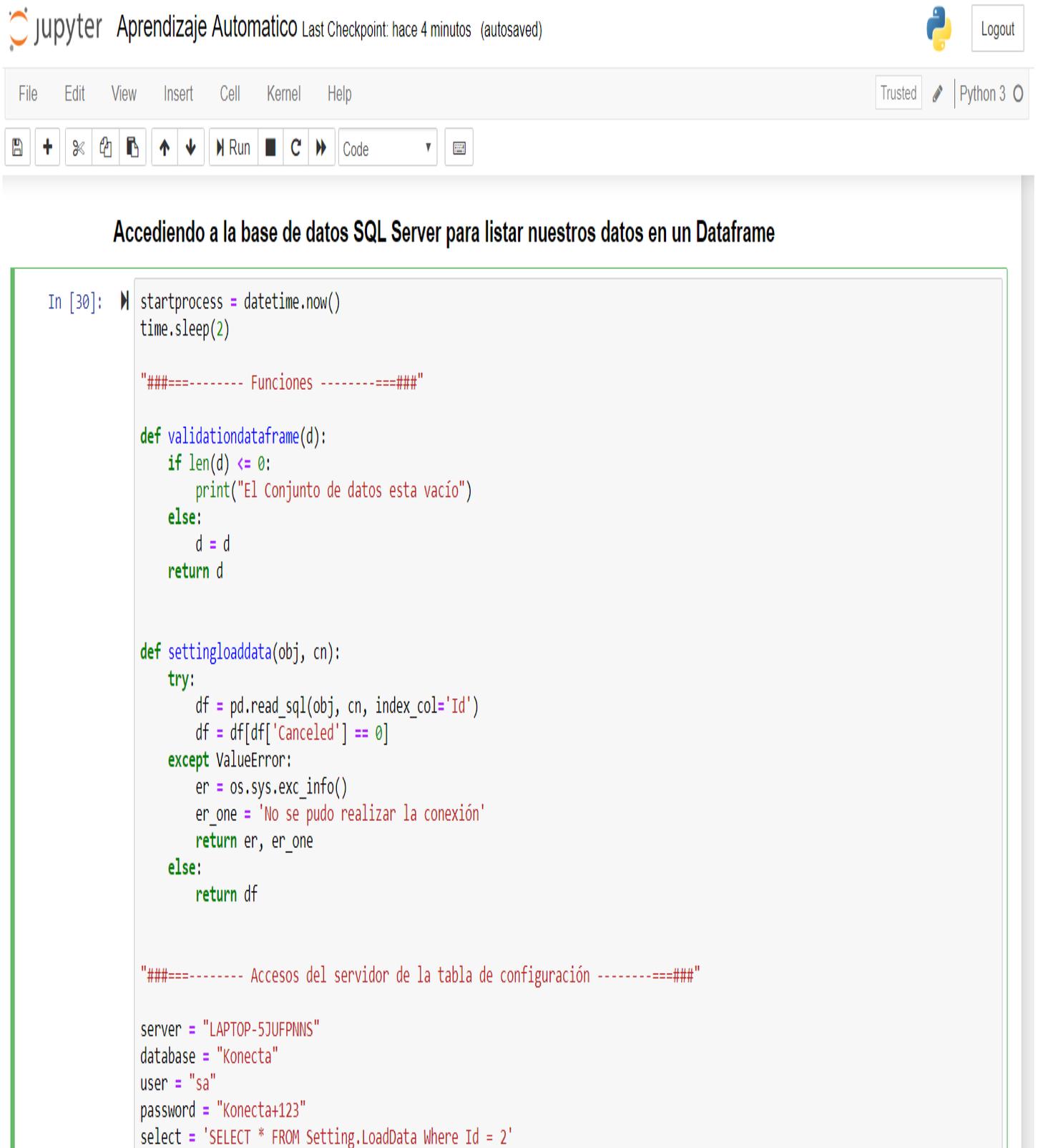
#### 1. Comprensión de los datos

Importando Librerías a utilizar durante la implementación de tres modelos predictivos y poder comparar entre ellos quien nos da un mejor resultado respecto a los analisis de score, coeficiente de correlación, los modelos son los siguientes, Modelo de Regresión Lineal, Modelo de Regresión Lineal Múltiple, Modelo de Redes Neuronales

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sb
from datetime import datetime, timedelta
import os
import sqlalchemy
from ConnectionString import Connection_String
import time
import sklearn
from sklearn import linear_model
from sklearn.metrics import r2_score
from sklearn.metrics import mean_squared_error
from sklearn.model_selection import train_test_split
# Esta línea permite que los gráficos sean renderizados directamente en nuestro Notebook
%matplotlib inline
```

## Figura 13

En la figura 13 podemos ver el acceso a la Base de Datos SQL Server para listar nuestros datos en un Dataframe (Librería Pandas).



The image shows a Jupyter Notebook interface with the following elements:

- Header:** "jupyter Aprendizaje Automatico Last Checkpoint: hace 4 minutos (autosaved)" on the left, and a Python logo and "Logout" button on the right.
- Menu:** "File Edit View Insert Cell Kernel Help" with a "Trusted" indicator and "Python 3" version.
- Toolbar:** Includes icons for save, add, delete, copy, paste, undo, redo, run, and code editor.
- Title:** "Accediendo a la base de datos SQL Server para listar nuestros datos en un Dataframe".
- Code Cell:** Contains the following Python code:

```
In [30]: ▶ startprocess = datetime.now()
time.sleep(2)

"###====----- Funciones -----###"

def validationdataframe(d):
    if len(d) <= 0:
        print("El Conjunto de datos esta vacío")
    else:
        d = d
    return d

def settingloaddata(obj, cn):
    try:
        df = pd.read_sql(obj, cn, index_col='Id')
        df = df[df['Canceled'] == 0]
    except ValueError:
        er = os.sys.exc_info()
        er_one = 'No se pudo realizar la conexión'
        return er, er_one
    else:
        return df

"###====----- Accesos del servidor de la tabla de configuración -----###"

server = "LAPTOP-5JUFPNNS"
database = "Konecta"
user = "sa"
password = "Konecta+123"
select = 'SELECT * FROM Setting.LoadData Where Id = 2'
```

*Elaboración propia*

```

user = "sa"
password = "Konecta+123"
select = 'SELECT * FROM Setting.LoadData Where Id = 2'
cnx = Connection_String(server, database, user, password)

##### Conexión a la Base de Datos #####

table_configuration = settingloaddata(select, cnx)
pd.set_option('display.expand_frame_repr', False)

try:
    for regconf in table_configuration.itertuples():

        ##### Obtener datos en variables de origen #####

        driver_source = regconf[1]
        server_source = regconf[2]
        user_source = regconf[3]
        password_source = regconf[4]
        database_source = regconf[5]
        scheme_source = regconf[6]
        objeto_source = regconf[7]

        ##### Abrir conexión Objeto Origen #####

        cnx_source = Connection_String(server_source, database_source, user_source, password_source)

        ##### Consultar objeto #####

        query = objeto_source

        try:
            print('Consultando: %s' % query, '...')
            query = pd.read_sql_query(query, cnx_source)
        except:
            print('Se produjo un error al consultar el objeto')
    
```

Elaboración Propia



```

try:
    print('Consultando: %s' % query, ' ...')
    query = pd.read_sql_query(query, cnx_source)
except:
    print('---> Se produjo un error <----')
    print(os.sys.exc_info())
    cnx_source.close()
    print('Cierre de conexión: Source')
finally:
    cnx_source.close()
    print('Cierre de conexión: Source')
cnx.close()
print('Cierre de conexión: Local')

except:
    print('---> Se produjo un error <----')
    print(os.sys.exc_info())
    cnx.close()
    print('Cierre de conexión: Local')

endprocess = datetime.now()
timeexec = endprocess - startprocess

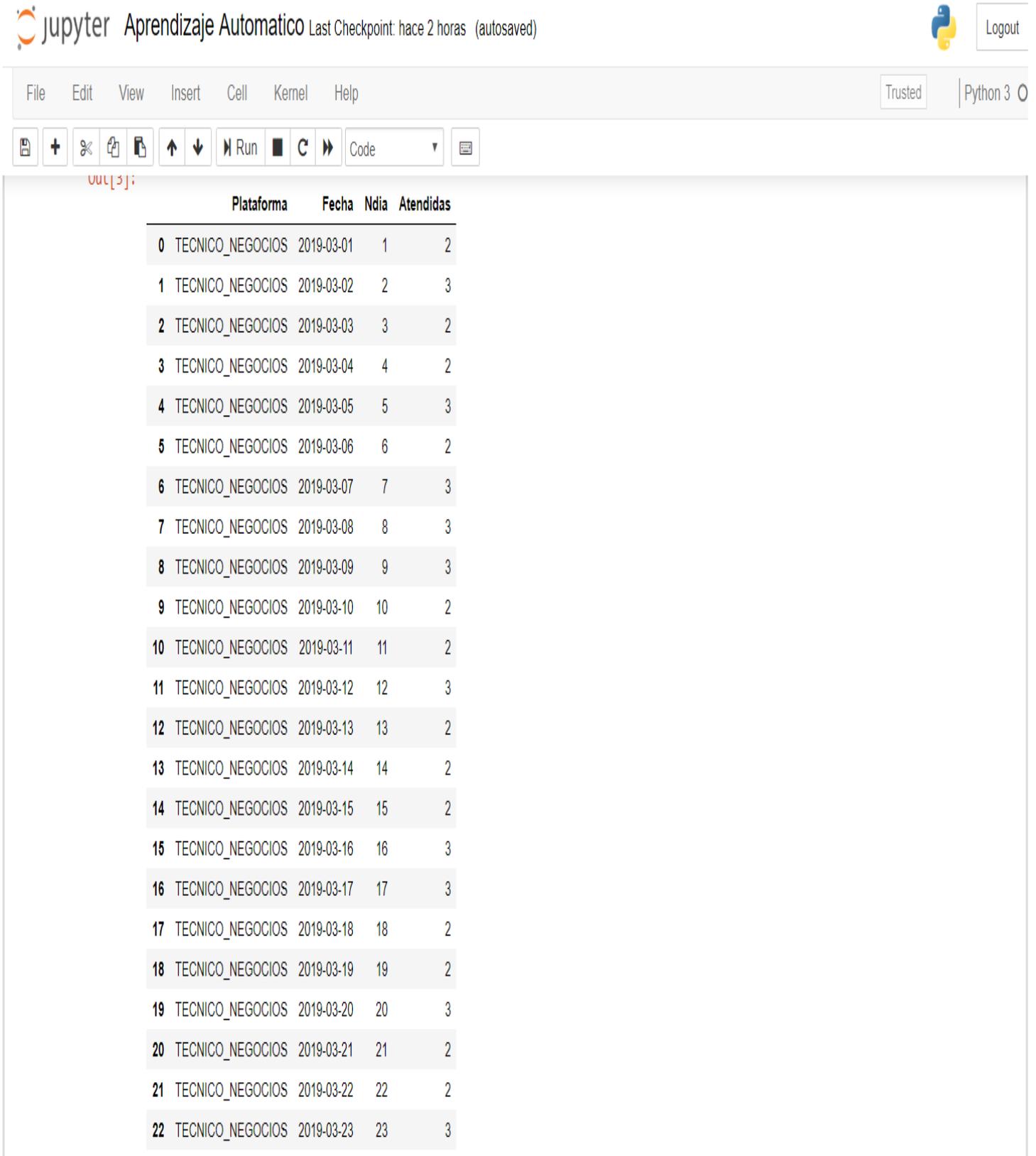
print('Inicio de Ejecución: %s' % startprocess.strftime('%Y-%m-%d %H:%M:%S'))
print('Fin de Ejecución: %s' % endprocess.strftime('%Y-%m-%d %H:%M:%S'))
print('Tiempo de ejecución: %s' % timeexec)
query_consolidado = query
query_consolidado
    
```

```

Consultando: Exec dbo.Lista_Consolidado ...
Cierre de conexión: Source
Cierre de conexión: Local
Inicio: 2019-06-08 14:06:12
Fin: 2019-06-08 14:06:14
Tiempo de ejecución: 0:00:02.197189
    
```

**Figura 14:**

En la figura 14 podemos observar los datos en un dataframe



*Elaboración Propia*

## Figura 15:

En la figura 15 visualizamos las columnas del dataframe.

### Visualizando columnas del Dataframe

```
In [6]: query_consolidado.columns
```

```
Out[6]: Index(['Plataforma', 'Fecha', 'Ndia', 'Atendidas'], dtype='object')
```

*Elaboración Propia*

## Figura 16:

En la figura 16 podemos visualizar el acceso a las columnas de nuestro dataframe

### Accediendo a las columnas del dataframe ¶

```
In [9]: columna1 = query_consolidado['Plataforma']  
columna1.head()
```

```
Out[9]: 0    TECNICO_NEGOCIOS  
1    TECNICO_NEGOCIOS  
2    TECNICO_NEGOCIOS  
3    TECNICO_NEGOCIOS  
4    TECNICO_NEGOCIOS  
Name: Plataforma, dtype: object
```

*Elaboración Propia*

## Figura 17:

En la figura 17 observamos el acceso a las líneas del dataframe

### Accediendo a líneas del dataframe

```
In [10]: linea = query_consolidado.loc[10,:]  
linea
```

```
Out[10]: Plataforma    TECNICO_NEGOCIOS  
Fecha                2019-03-11  
Ndia                  11  
Atendidas             2  
Name: 10, dtype: object
```

*Elaboración Propia*

## Figura 18:

En la figura “18” inspeccionamos los tipos de datos y el entendimiento de los mismos.

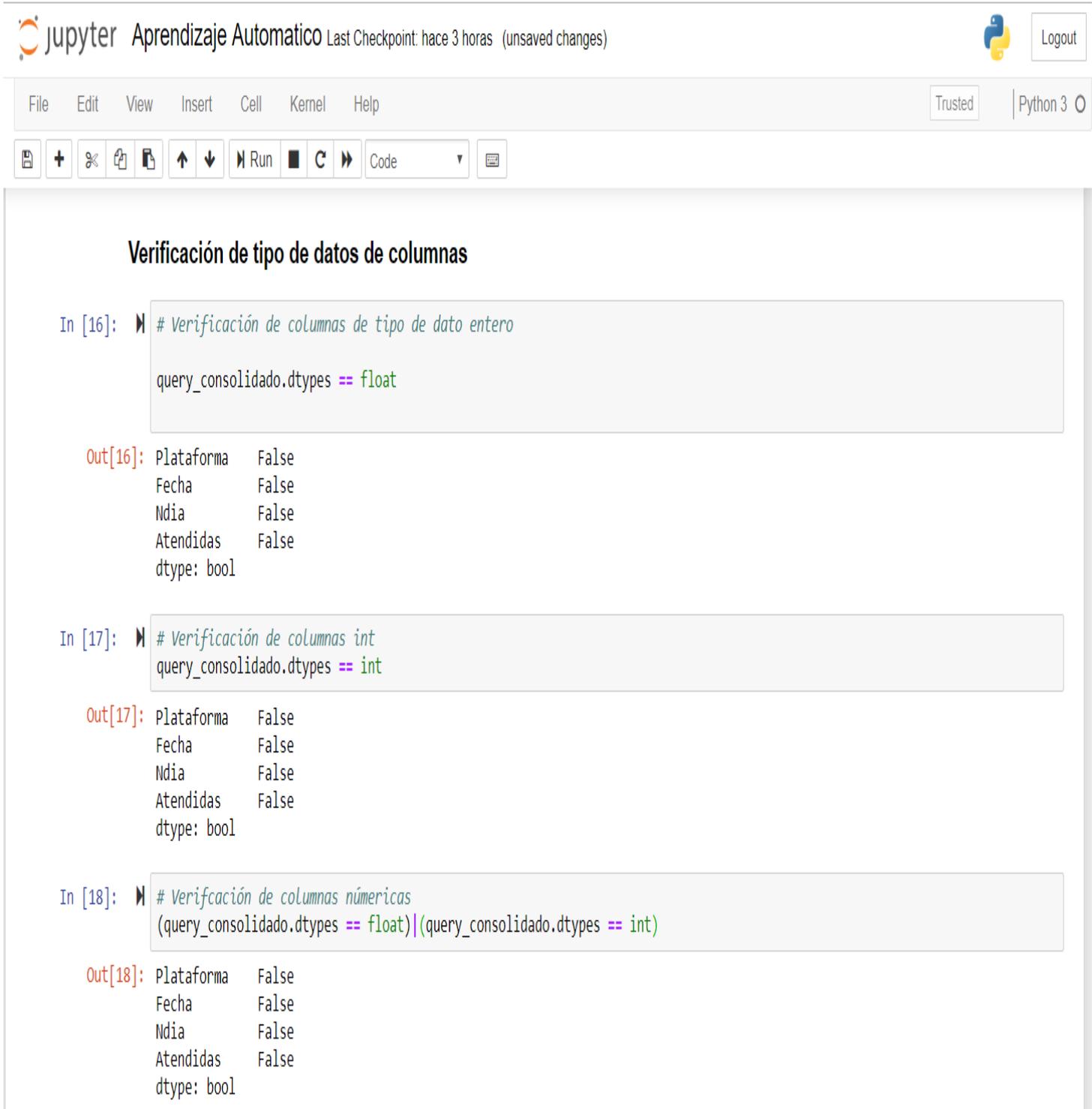
### Inspección de los tipos de datos y entendimiento

```
In [11]: query_consolidado.info()  
  
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 89 entries, 0 to 88  
Data columns (total 4 columns):  
Plataforma    89 non-null object  
Fecha         89 non-null object  
Ndia          89 non-null int64  
Atendidas     89 non-null int64  
dtypes: int64(2), object(2)  
memory usage: 2.9+ KB
```

*Elaboración Propia*

## Figura 19:

En la figura “19” verificamos los tipos de datos de las columnas del dataframe



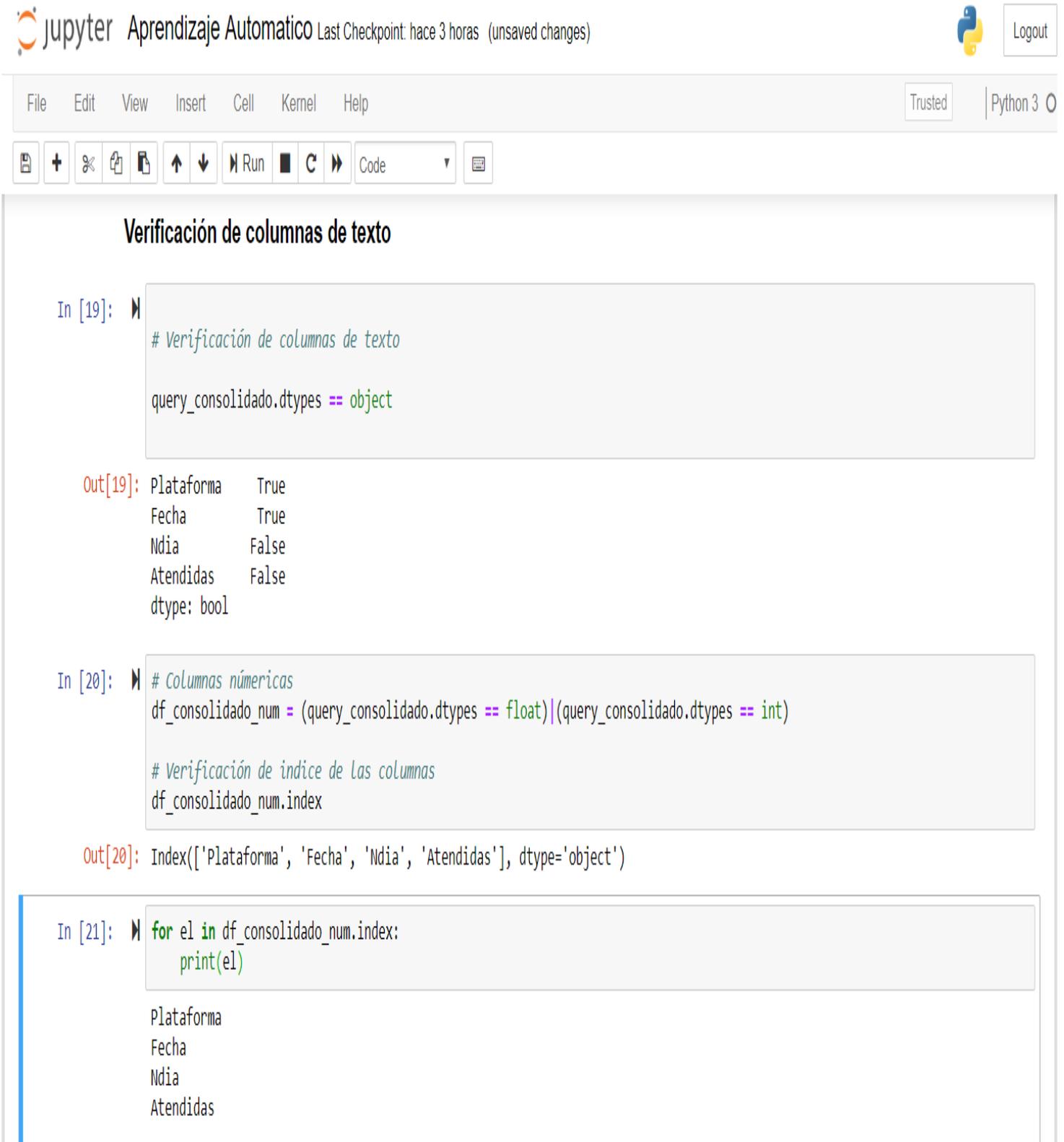
The screenshot shows a Jupyter Notebook interface with the following elements:

- Header:** "jupyter Aprendizaje Automatico Last Checkpoint: hace 3 horas (unsaved changes)" and a "Logout" button.
- Menu:** File, Edit, View, Insert, Cell, Kernel, Help.
- Trust Status:** Trusted
- Kernel:** Python 3
- Toolbar:** Includes icons for save, add, undo, redo, up/down arrows, run, clear, and code editor options.
- Section Title:** "Verificación de tipo de datos de columnas"
- Cell 16:** Code: `# Verificación de columnas de tipo de dato entero`  
`query_consolidado.dtypes == float`  
Output: `Plataforma False`  
`Fecha False`  
`Ndia False`  
`Atendidas False`  
`dtype: bool`
- Cell 17:** Code: `# Verificación de columnas int`  
`query_consolidado.dtypes == int`  
Output: `Plataforma False`  
`Fecha False`  
`Ndia False`  
`Atendidas False`  
`dtype: bool`
- Cell 18:** Code: `# Verificación de columnas numéricas`  
`(query_consolidado.dtypes == float)|(query_consolidado.dtypes == int)`  
Output: `Plataforma False`  
`Fecha False`  
`Ndia False`  
`Atendidas False`  
`dtype: bool`

Elaboración Propia

**Figura 20:**

En la figura 20 verificamos las columnas de tipo de datos texto o cadena



The image shows a Jupyter Notebook interface with the following elements:

- Header:** "jupyter Aprendizaje Automatico" with a logo, "Last Checkpoint: hace 3 horas (unsaved changes)", a Python logo, and a "Logout" button.
- Menu:** "File", "Edit", "View", "Insert", "Cell", "Kernel", "Help".
- Trust/Kernel:** "Trusted" and "Python 3".
- Toolbar:** Includes icons for save, add, copy, paste, undo, redo, run, and a dropdown menu set to "Code".
- Section Title:** "Verificación de columnas de texto".
- Code Cell 19:**

```
In [19]: # Verificación de columnas de texto
query_consolidado.dtypes == object
```
- Output 19:**

```
Out[19]: Plataforma    True
Fecha                True
Ndia                 False
Atendidas            False
dtype: bool
```
- Code Cell 20:**

```
In [20]: # Columnas numéricas
df_consolidado_num = (query_consolidado.dtypes == float)|(query_consolidado.dtypes == int)

# Verificación de índice de las columnas
df_consolidado_num.index
```
- Output 20:**

```
Out[20]: Index(['Plataforma', 'Fecha', 'Ndia', 'Atendidas'], dtype='object')
```
- Code Cell 21:**

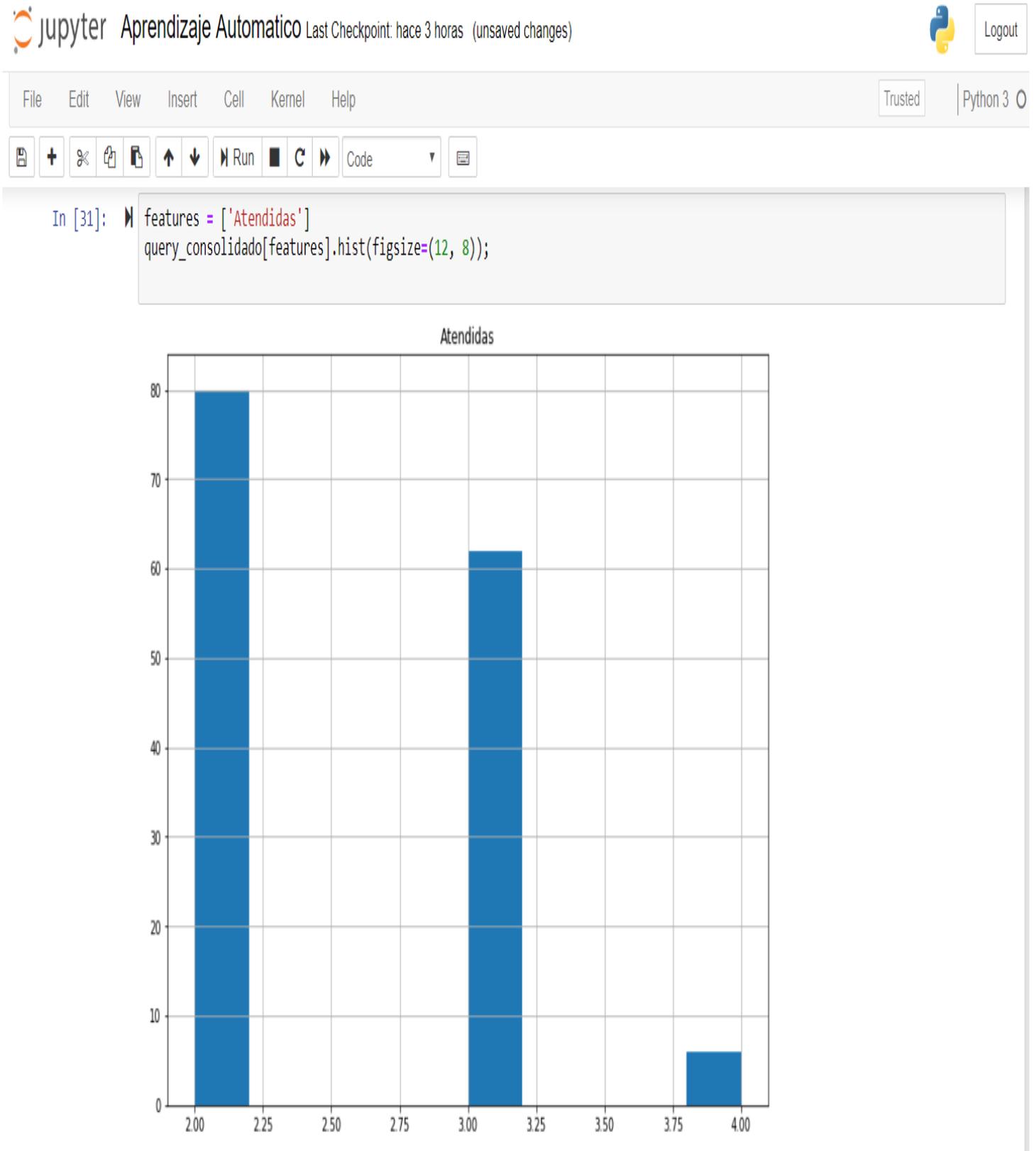
```
In [21]: for el in df_consolidado_num.index:
print(el)
```
- Output 21:**

```
Plataforma
Fecha
Ndia
Atendidas
```

*Elaboración Propia*

## Figura 21:

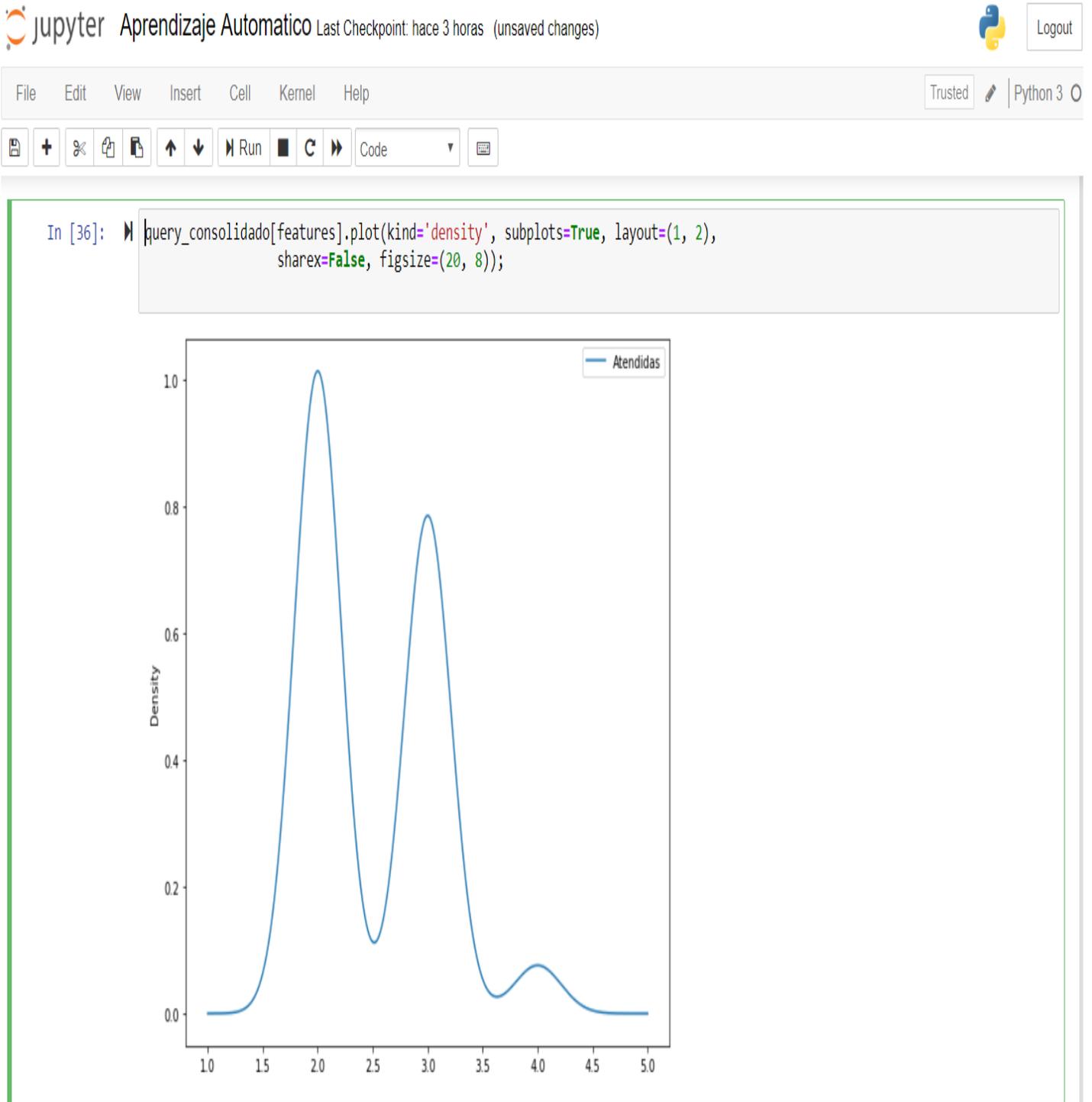
En la figura 21 analizamos los datos en un histograma.



*Elaboración Propia*

## Figura 22:

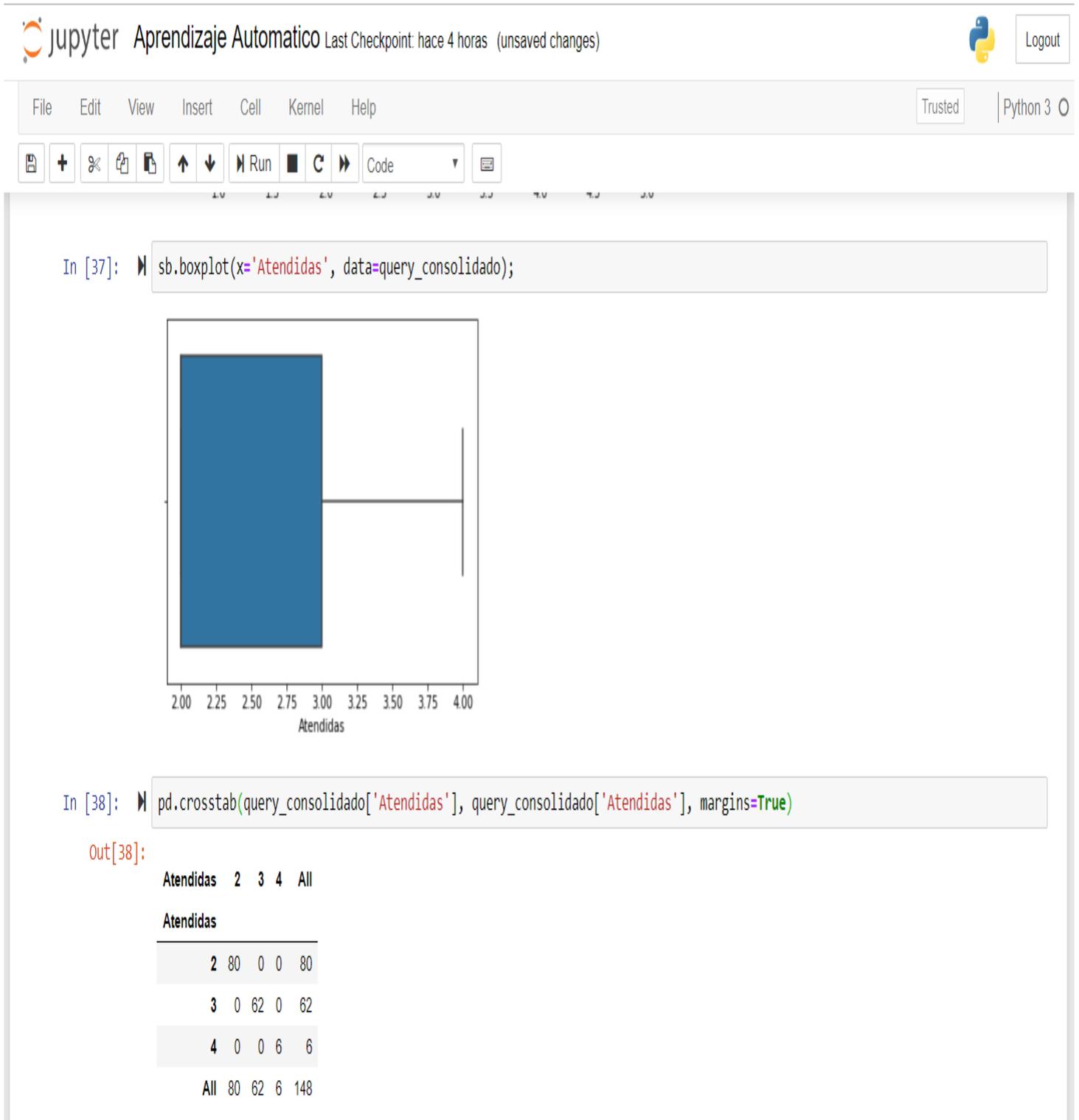
En la figura 22 analizamos los datos de las llamadas atendidas en un gráfico Plot.



*Elaboración Propia*

## Figura 23:

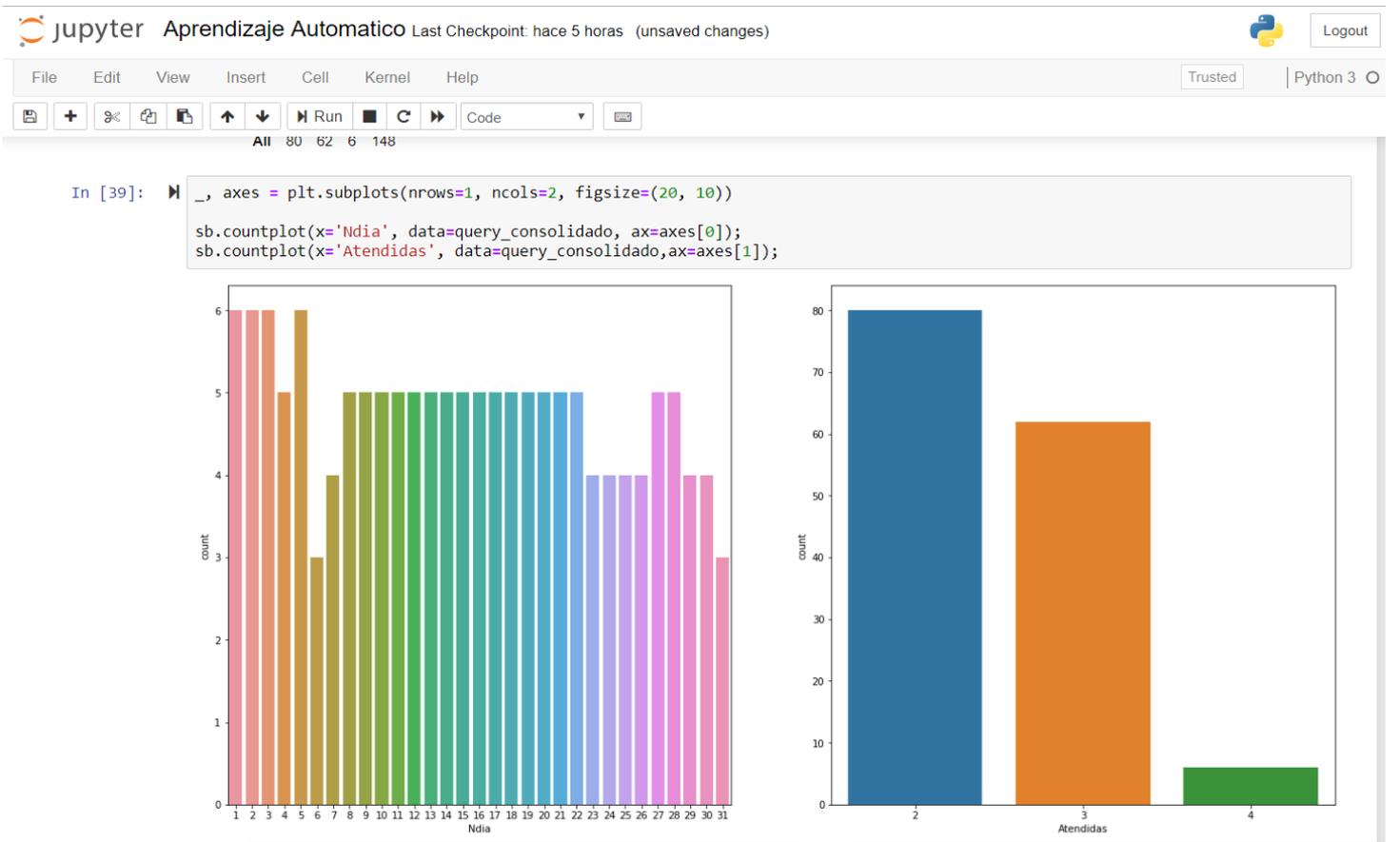
En la figura 23 observamos un gráfico boxplot de nuestros datos.



*Elaboración Propia*

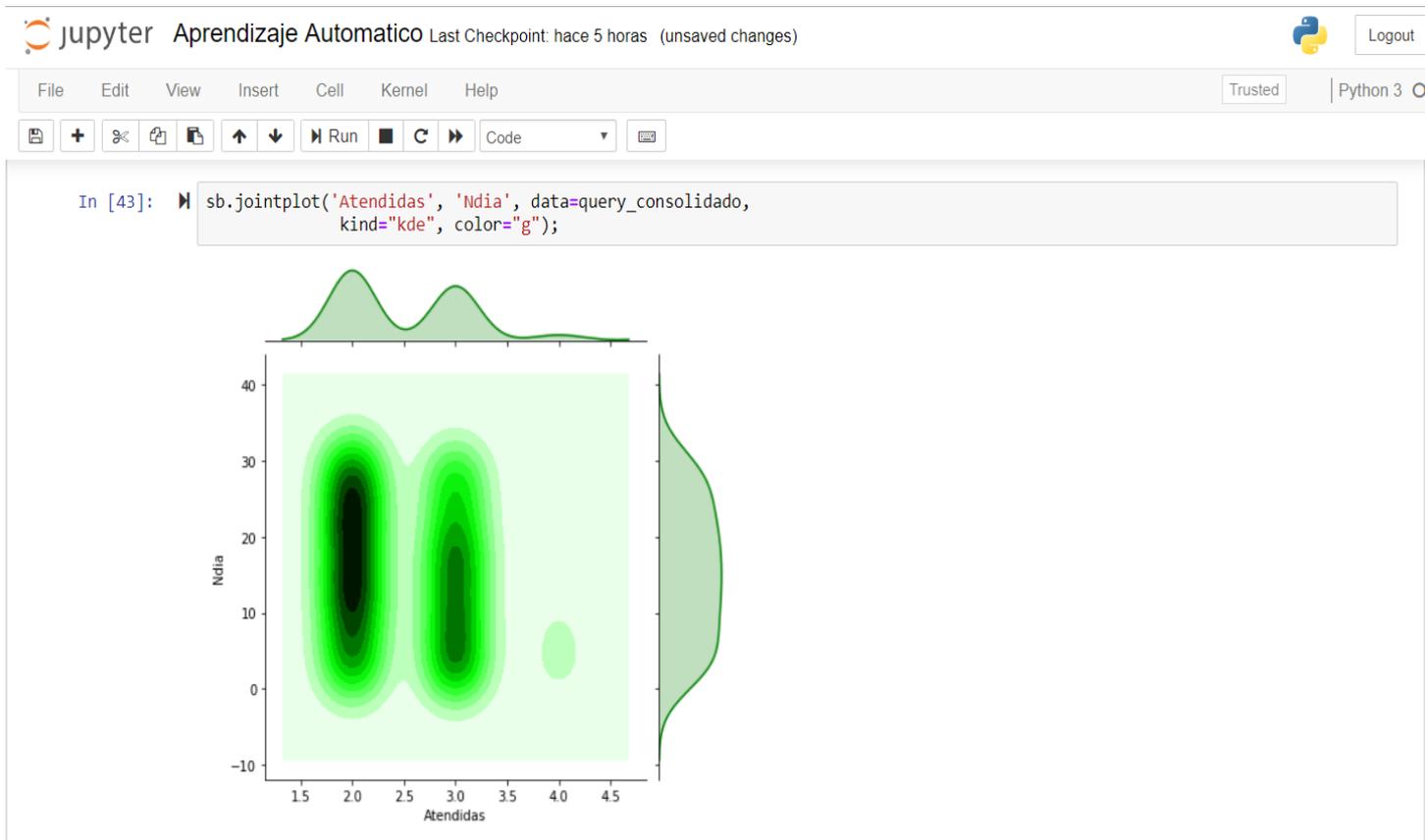
## Figura 24:

En la figura 24 se muestra un conteo gráfico de los datos.



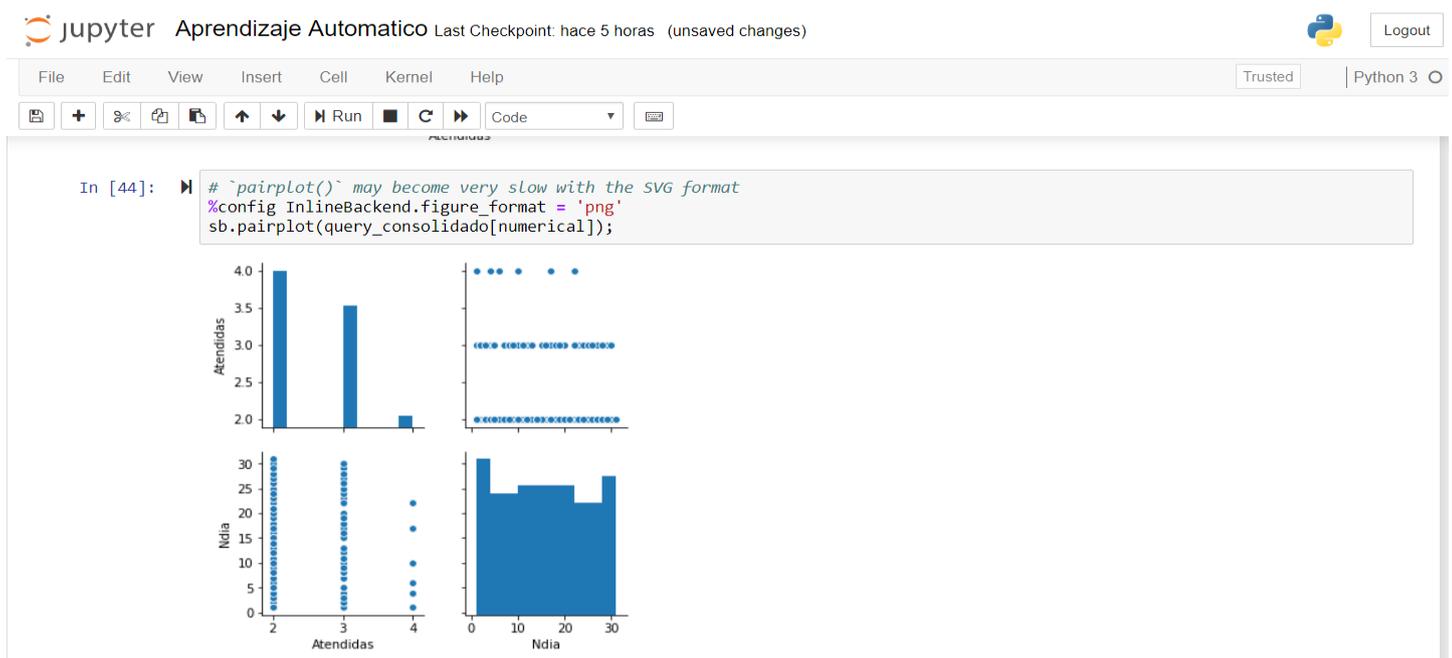
### Figura 25:

En la figura 25 se gráfica las variables Atendidas y Ndia

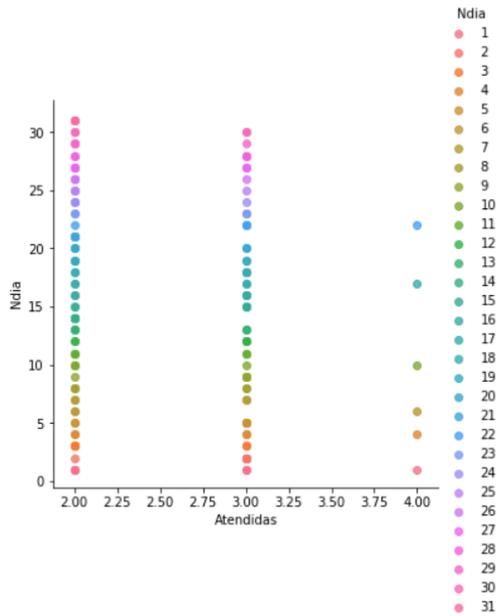


### Figura 26:

En la figura "26" hacemos un análisis de la variable Atendidas y Ndia

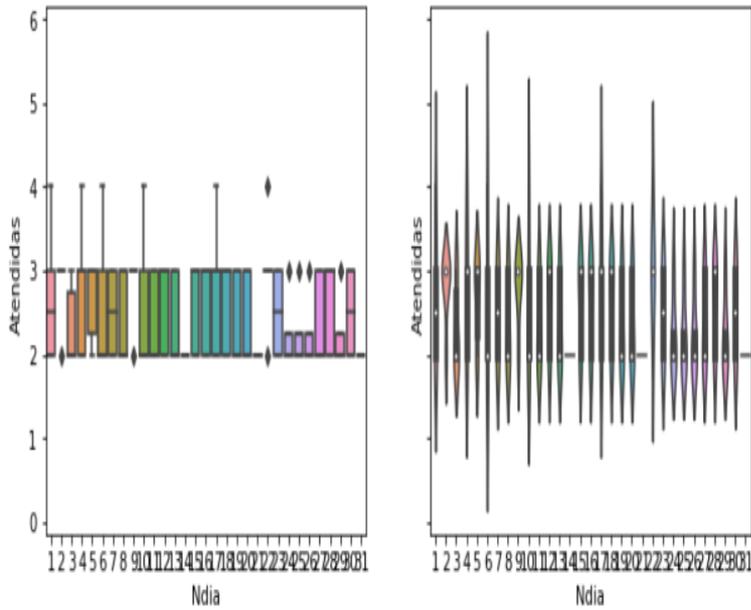


```
In [45]: sb.lmplot('Atendidas', 'Ndia', data=query_consolidado, hue='Ndia', fit_reg=False);
```

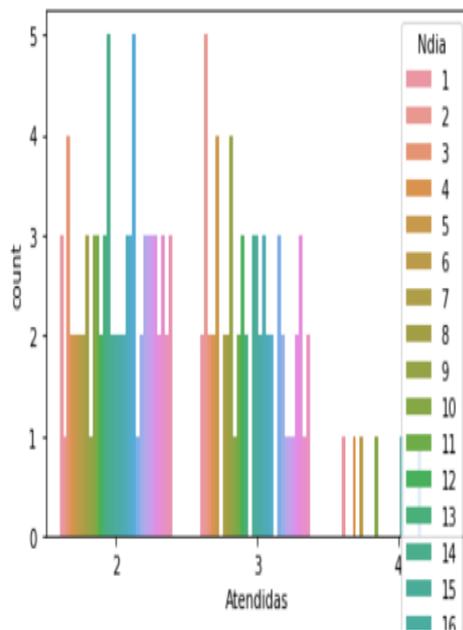




```
sb.violinplot(x= 'Ndia', y= 'Atendidas', data=query_consolidado, ax=axes[1]);
```

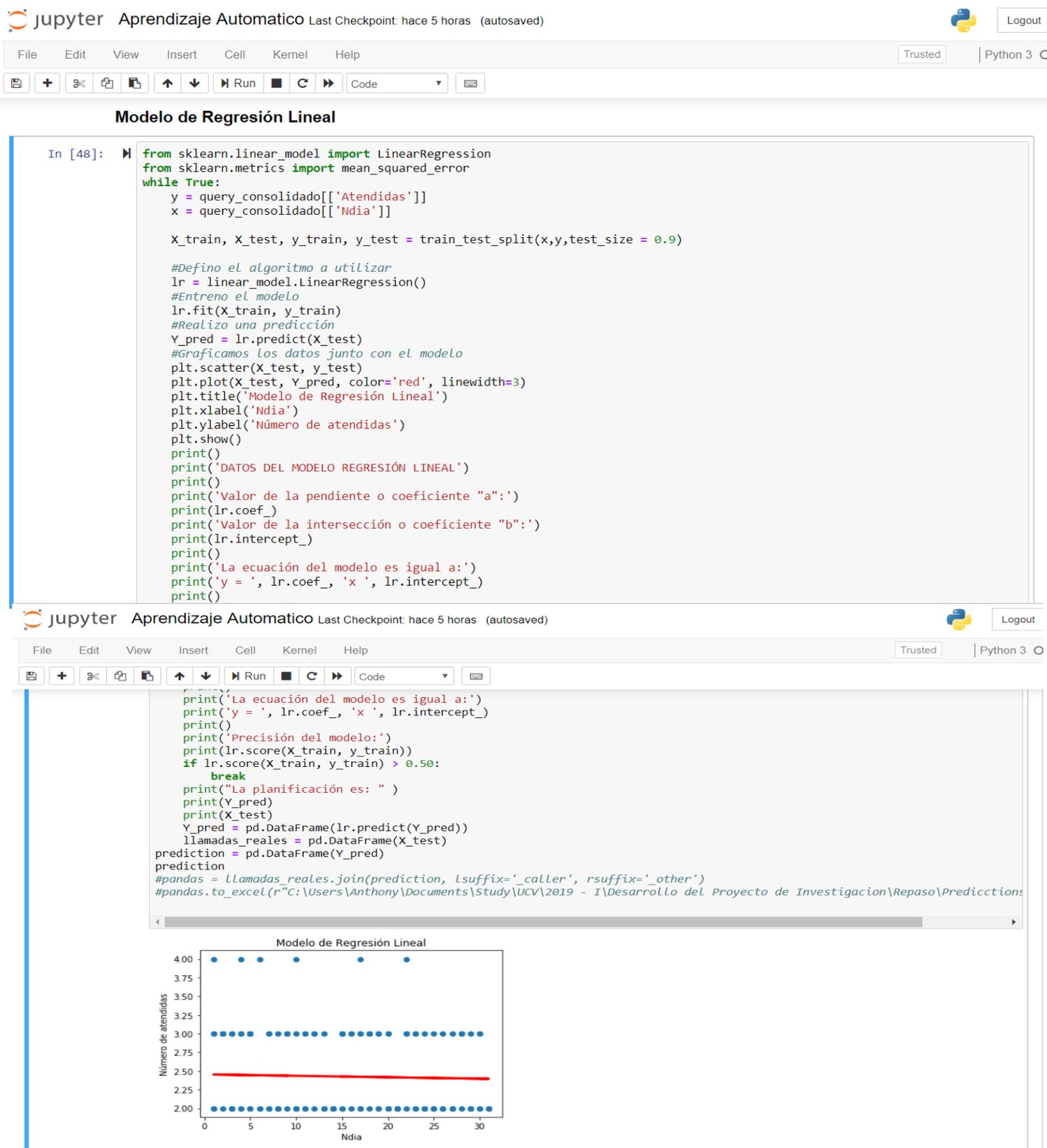


```
In [47]: sb.countplot(x='Atendidas', hue='Ndia', data=query_consolidado);
```



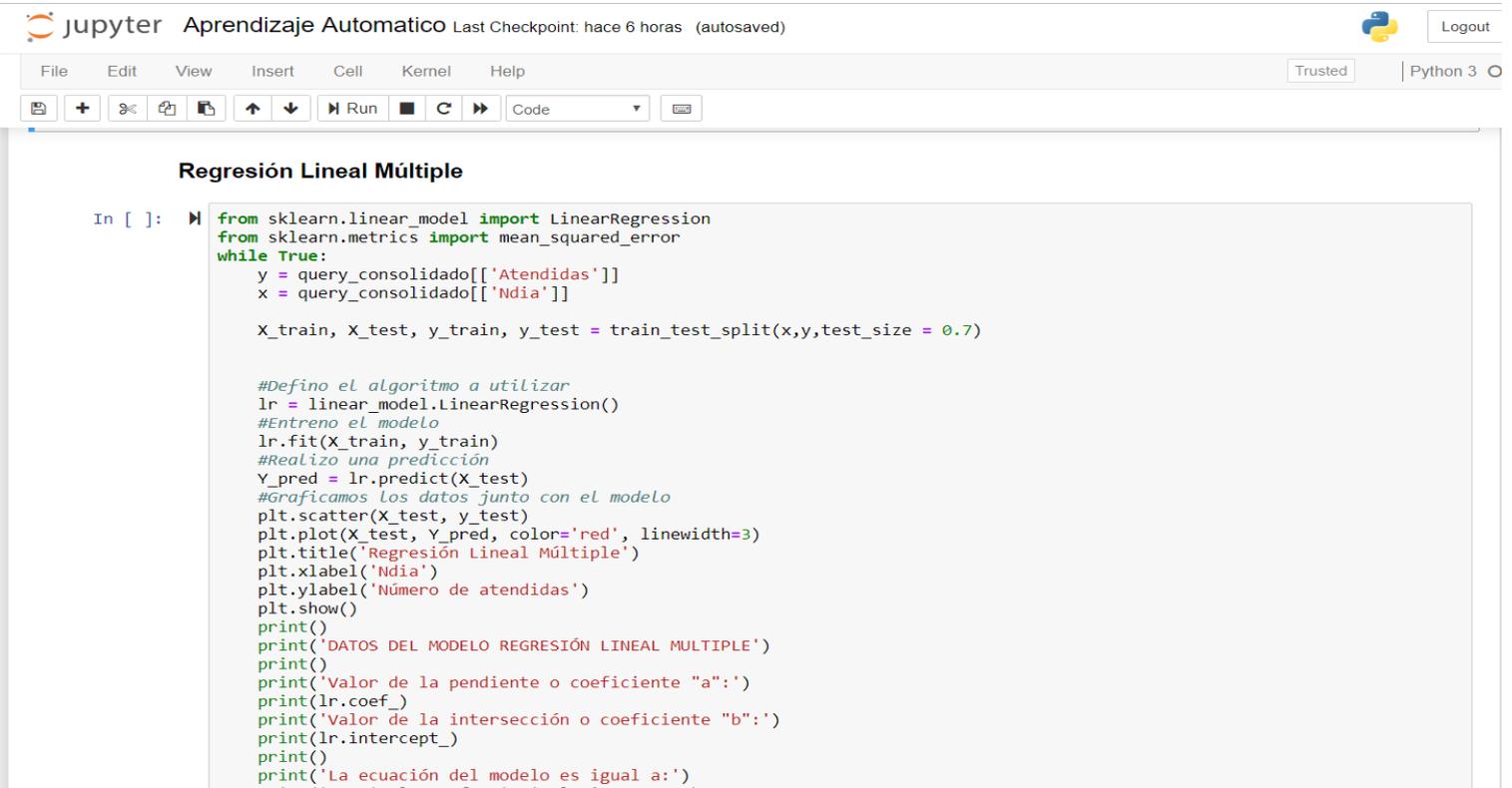
## Figura 27:

En la figura "27" podemos observar el modelo de regresión Lineal



## Figura 29:

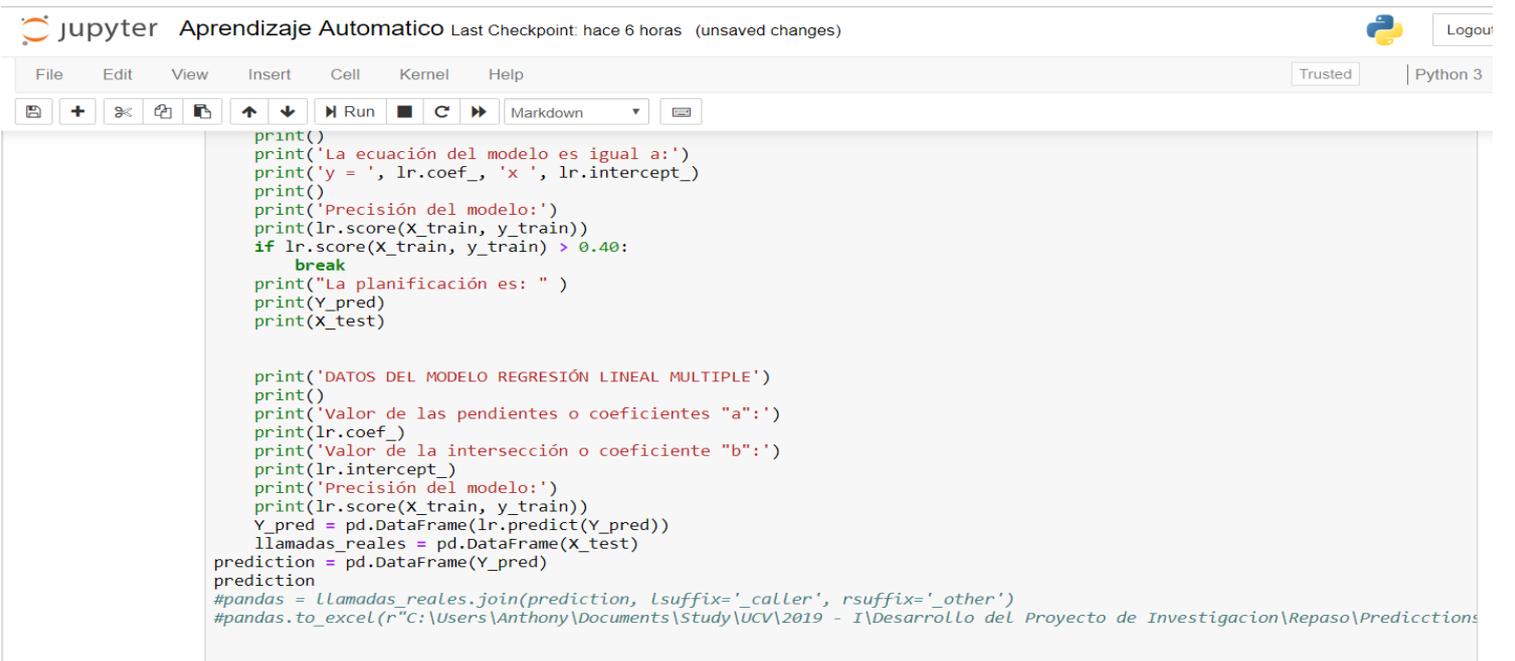
En la figura 29 podemos ver el modelo de regresión Lineal Múltiple.



```
In [ ]: from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
while True:
    y = query_consolidado[['Atendidas']]
    x = query_consolidado[['Ndia']]

    X_train, X_test, y_train, y_test = train_test_split(x,y,test_size = 0.7)

    #Defino el algoritmo a utilizar
    lr = linear_model.LinearRegression()
    #Entreno el modelo
    lr.fit(X_train, y_train)
    #Realizo una predicción
    Y_pred = lr.predict(X_test)
    #Graficamos los datos junto con el modelo
    plt.scatter(X_test, y_test)
    plt.plot(X_test, Y_pred, color='red', linewidth=3)
    plt.title('Regresión Lineal Múltiple')
    plt.xlabel('Ndia')
    plt.ylabel('Número de atendidas')
    plt.show()
    print()
    print('DATOS DEL MODELO REGRESIÓN LINEAL MULTIPLE')
    print()
    print('Valor de la pendiente o coeficiente "a":')
    print(lr.coef_)
    print('Valor de la intersección o coeficiente "b":')
    print(lr.intercept_)
    print()
    print('La ecuación del modelo es igual a:')
```



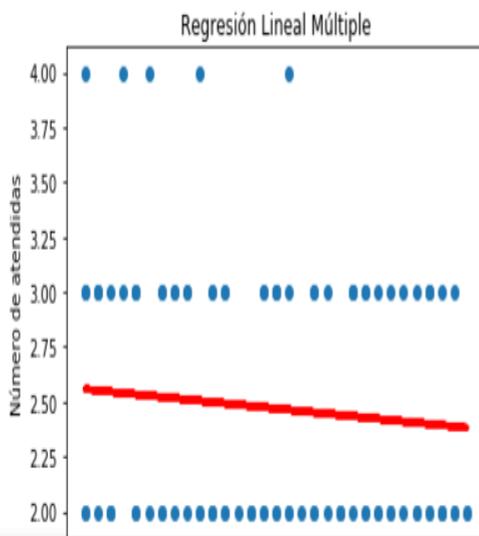
```
print()
print('La ecuación del modelo es igual a:')
print('y = ', lr.coef_, 'x ', lr.intercept_)
print()
print('Precisión del modelo:')
print(lr.score(X_train, y_train))
if lr.score(X_train, y_train) > 0.40:
    break
print("La planificación es: ")
print(Y_pred)
print(X_test)

print('DATOS DEL MODELO REGRESIÓN LINEAL MULTIPLE')
print()
print('Valor de las pendientes o coeficientes "a":')
print(lr.coef_)
print('Valor de la intersección o coeficiente "b":')
print(lr.intercept_)
print('Precisión del modelo:')
print(lr.score(X_train, y_train))
Y_pred = pd.DataFrame(lr.predict(Y_pred))
llamadas_reales = pd.DataFrame(X_test)
prediction = pd.DataFrame(Y_pred)
prediction
#pandas = llamadas_reales.join(prediction, lsuffix='_caller', rsuffix='_other')
#pandas.to_excel(r"C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigacion\Repaso\Prediccions
```



```
print(lr.intercept_)  
print('Precisión del modelo:')  
print(lr.score(X_train, y_train))  
Y_pred = pd.DataFrame(lr.predict(Y_pred))  
llamadas_reales = pd.DataFrame(X_test)  
prediction = pd.DataFrame(Y_pred)  
prediction  
#pandas = llamadas_reales.join(prediction, lsuffix='_caller', rsuffix='_other')  
#pandas.to_excel(r"C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigacion\Repaso\Prediccions
```

[[ -0.01875121]]  
Valor de la intersección o coeficiente "b":  
[ 2.779849]  
Precisión del modelo:  
0.06832768602539274



En la figura “30” podemos observar el modelo de Redes neuronales utilizando perceptrón Multicapa de la librería SkeaLearn.

The image shows the Spyder Python IDE interface. The main window is titled "Spyder (Python 3.7)" and contains a code editor on the left and a terminal window on the right. The code editor displays a Python script with the following content:

```
1 """
2 Created on Tue May 14 20:14:48 2019
3
4 @author: Anthony, Simon Altamirano
5         Julinho, Ostos Mera
6
7 Descripción: Redes Neuronales Perceptron Multicapa
8
9 """
10 #--- Importando conexión ---#
11 from ConnectionString import Connection_String
12
13 #--- Importación de Libreria Numpy ---#
14 import numpy as np
15
16 #--- Importación de Libreria Pandas ---#
17 import pandas as pd
18
19 #--- Importación de libreria Matplotlib ---#
20 import matplotlib.pyplot as plt
21
22 #--- Importación de libreria seaborn ---#
23 import seaborn as sb
24
25 #--- Importación de tiempos ---#
26 from datetime import datetime, timedelta
27
28 #--- Importación del sistema operativo (Archivos y directorios) ---#
29 import os
30
31 import sqlalchemy
32
33 import time
34
35
36 from sklearn.neural_network import MLPRegressor
37 from sklearn.model_selection import train_test_split
38
39
40 startprocess = datetime.now()
41 time.sleep(2)
42
43 """ Bloque de Funciones """
```

The terminal window on the right shows the output of the Python interpreter, including the version information and the start of an IPython session:

```
Python 3.7.3 (default, Mar 27 2019,
17:13:21) [MSC v.1915 64 bit (AMD64)]
Type "copyright", "credits" or "license" for
more information.

IPython 7.4.0 -- An enhanced Interactive
Python.

In [1]:
```

The status bar at the bottom of the IDE indicates the following information: Permisos: RW, Fin de línea: CRLF, Codificación: UTF-8, Línea: 16, Columna: 43, Memoria: 41 %.

# Figura 31:

The image shows the Spyder Python IDE interface. The main window is titled "Spyder (Python 3.7)" and contains a menu bar (Archivo, Editar, Buscar, Código fuente, Ejecutar, Depurar, Terminales, Proyectos, Herramientas, Ver, Ayuda) and a toolbar. The editor displays a Python script with the following code:

```
84
85 """ Consultar objeto """
86 query = objeto_source
87
88
89 try:
90     print('Consultando: %s' % query, '...')
91     query = pd.read_sql_query(query, cnx_source)
92     print(query)
93     y = np.array(query['Atendidas'])
94     x = np.array(query['Dia'])
95
96     X = x[:, np.newaxis]
97
98
99 while True:
100
101     X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = 0.8) #, test_size = 0.8
102
103     mlr = MLPRegressor(solver="lbfgs", alpha=1e-5, hidden_layer_sizes=(3,3), random_state=1)
104     mlr.fit(X_train, y_train)
105     print("El score es ", mlr.score(X_train, y_train))
106     if mlr.score(X_train, y_train) >= 0.80:
107         break
108
109     Y_pred = pd.DataFrame(mlr.predict(X_test))
110     llamadas_reales = pd.DataFrame(X_test)
111     prediction = pd.DataFrame(Y_pred)
112
113     pandas = llamadas_reales.join(prediction, lsuffix='Dia', rsuffix='PredicccionAtendida')
114
115     predicccionAtendida = pd.DataFrame(pandas['0PredicccionAtendida'].unique())
116     numeroDia = pd.DataFrame(pandas['0Dia'].unique())
117     pandasSQL = numeroDia.join(predicccionAtendida, lsuffix='Dia', rsuffix='PredicccionAtendida')
118
119     """ Antes de insertar eliminar datos de la tabla con la misma fecha. """
120
121     #print('Eliminando registros de: %s del %s al %s ...' % objeto_dest)
122     cursor = cnx_dest.cursor()
123     query_del = query_delete
124     with cursor.execute(query_delete):
125         print('Se eliminaron datos de: %s' % objeto_dest, " correctamente.")
```

On the right side, there is a help panel titled "Uso" with the following text:

En este panel es posible obtener la ayuda de cualquier objeto al oprimir **Ctrl+H** estando al frente del mismo, bien sea en el Editor o en la Terminal.

Esta ayuda también se puede mostrar automáticamente después de escribir un paréntesis junto a un objeto. Este comportamiento puede activarse en *Preferencias > Ayuda*.

Nuevo en Spyder? Lee nuestro [tutorial](#)

Below the help panel, there are tabs for "Explorador de variables", "Explorador de archivos", and "Ayuda". Below that is a "Terminal de Python" window showing the following output:

```
Python 3.7.3 (default, Mar 27 2019, 17:13:21) [MSC v.1915 64 bit (AMD64)]
Type "copyright", "credits" or "license()" for more information.

IPython 7.4.0 -- An enhanced Interactive Python.

In [1]:
```

At the bottom of the IDE, there is a status bar showing: "Permisos: RW", "Fin de línea: CRLF", "Codificación: UTF-8", "Línea: 1", "Columna: 1", "Memoria: 43 %".

```

10 -----
11 -- Author: <Anthony Simon Altamirano>
12 -- Create date: <15/04/2019/>
13 -- Description: <Store Procedure con datos de llamadas>
14 -----
15
16 ALTER PROCEDURE [dbo].[Lista_Consolidado]
17 AS
18
19 SET LANGUAGE SPANISH;
20
21 WITH CTE_Data
22 AS
23 (
24 SELECT [Anio]
25        ,[Mes]
26        ,[Fecha]
27        ,[DAY](Fecha)Ndia
28        ,CASE WHEN DATPART(MF, (MFKDAY (Fecha)) = 'Dominon' THEN 1 ELSE
29

```

Plataforma	Fecha	Ndia	Atendidas
TECNICO_NEGOCIOS	2019-03-01	1	2
TECNICO_NEGOCIOS	2019-03-02	2	3
TECNICO_NEGOCIOS	2019-03-03	3	2
TECNICO_NEGOCIOS	2019-03-04	4	2
TECNICO_NEGOCIOS	2019-03-05	5	3
TECNICO_NEGOCIOS	2019-03-06	6	1
TECNICO_NEGOCIOS	2019-03-07	7	3
TECNICO_NEGOCIOS	2019-03-08	8	8
TECNICO_NEGOCIOS	2019-03-09	9	5
TECNICO_NEGOCIOS	2019-03-10	10	2
TECNICO_NEGOCIOS	2019-03-11	11	2
TECNICO_NEGOCIOS	2019-03-12	12	3
TECNICO_NEGOCIOS	2019-03-13	13	2
TECNICO_NEGOCIOS	2019-03-14	14	2

Utilizamos la siguiente Query para obtener las columnas que tiene nuestro dataframe:

```

In [7]: print(query_consolidado.columns)

Index(['Plataforma', 'Fecha', 'Ndia', 'Atendidas'], dtype='object')

```

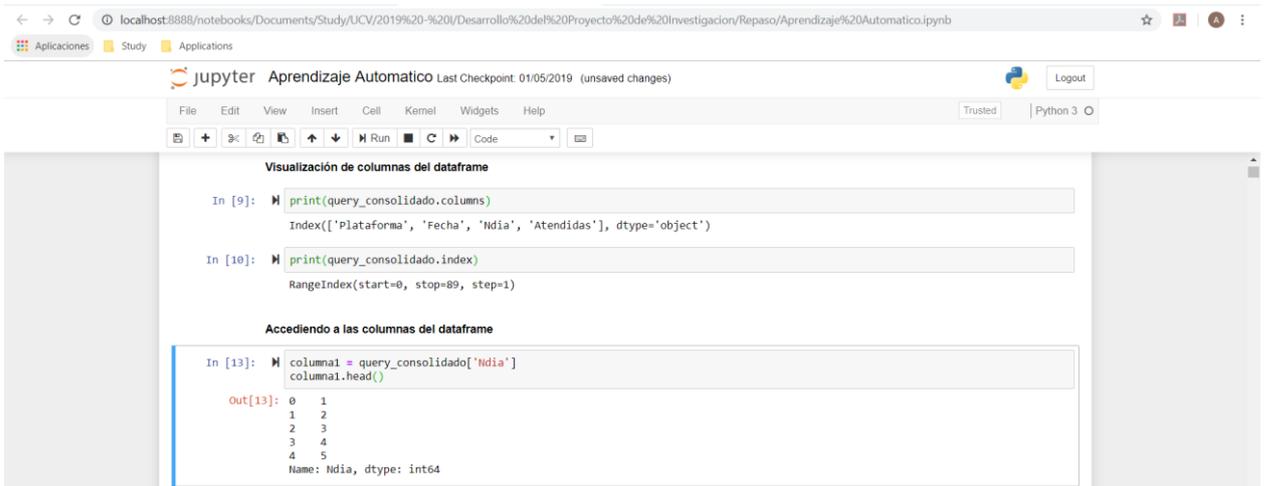
```

In [8]: # Cuantas lineas y columnas tiene nuestro dataframe
### Un dataframe es una estructura de datos que se compone de los elementos siguientes

query_consolidado.shape

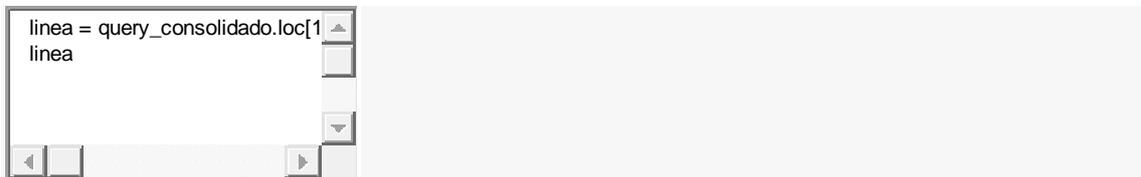
Out[8]: (89, 4)

```



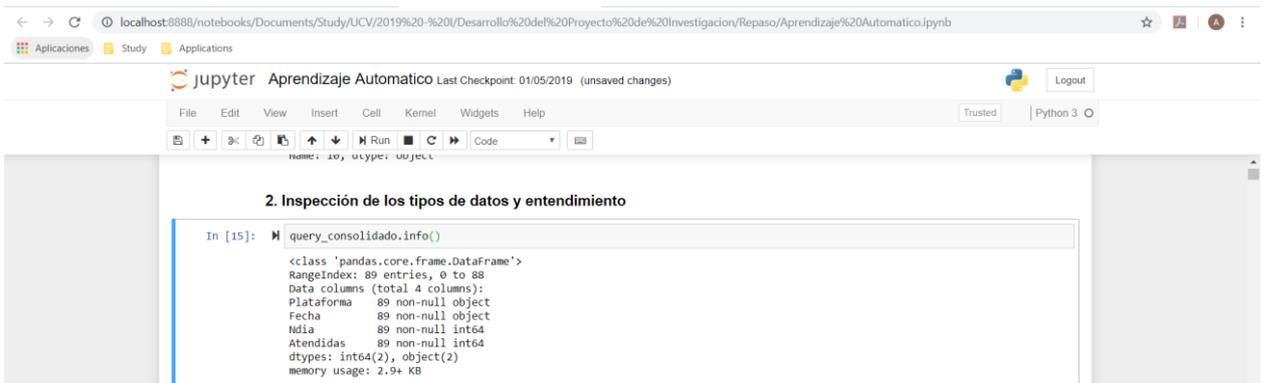
Accediendo a líneas del dataframe

In [14]:



## Figura 16: Análisis de datos

En la figura 16 se observa la Inspección de los tipos de datos y entendimiento



```
Aplicaciones Study Applications

jupyter Aprendizaje Automatico Last Checkpoint: 01/05/2019 (unsaved changes)
Python 3

File Edit View Insert Cell Kernel Widgets Help
+ Run Code

Columnas numéricas y columnas de texto

In [16]: query_consolidado.dtypes
Out[16]: Plataforma object
Fecha object
Ndia int64
Atendidas int64
dtype: object

Verificación de tipo de datos de columnas

In [17]: # Verificación de columnas float
query_consolidado.dtypes == float
Out[17]: Plataforma False
Fecha False
Ndia False
Atendidas False
dtype: bool

In [18]: # Verificación de columnas int
query_consolidado.dtypes == int
Out[18]: Plataforma False
Fecha False
Ndia False
Atendidas False
dtype: bool
```

```
localhost:8888/notebooks/Documents/Study/UCV/2019%20-%20/Desarrollo%20de%20Proyecto%20de%20Investigacion/Repaso/Aprendizaje%20Automatico.ipynb
Aplicaciones Study Applications

jupyter Aprendizaje Automatico Last Checkpoint: 01/05/2019 (unsaved changes)
Python 3

File Edit View Insert Cell Kernel Widgets Help
+ Run Code

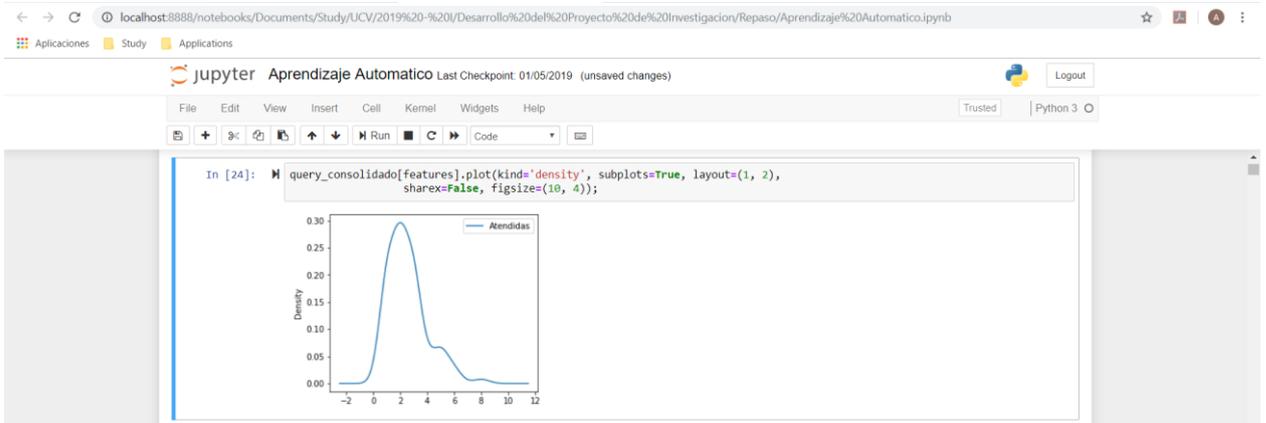
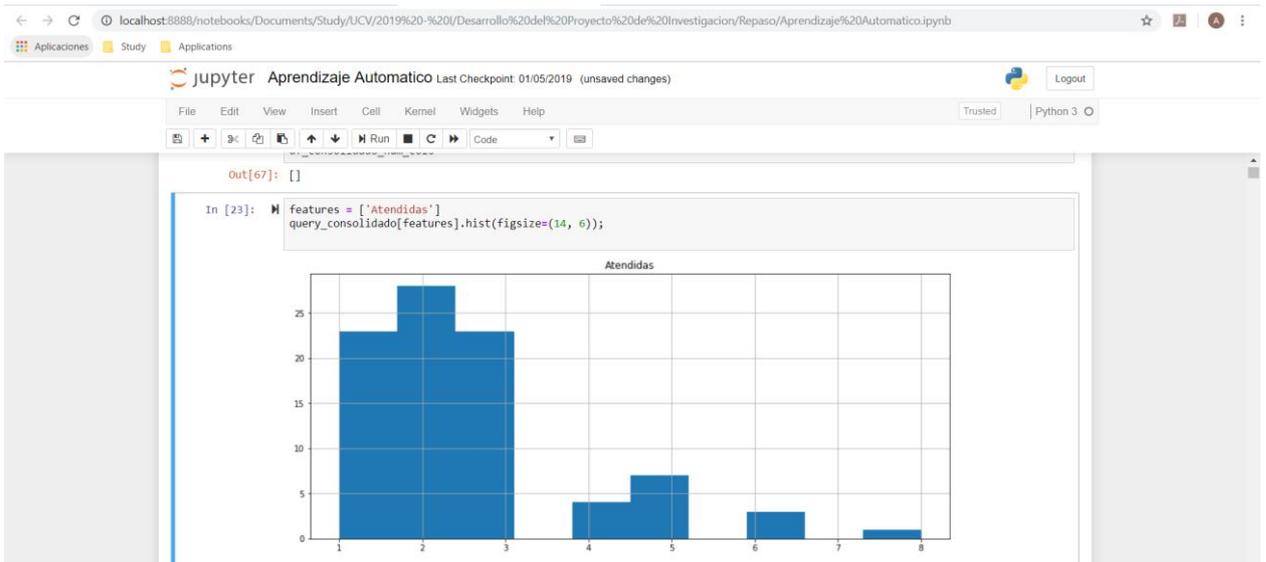
In [18]: # Verificación de columnas int
query_consolidado.dtypes == int
Out[18]: Plataforma False
Fecha False
Ndia False
Atendidas False
dtype: bool

In [19]: # Verificación de columnas numéricas
(query_consolidado.dtypes == float)|(query_consolidado.dtypes == int)
Out[19]: Plataforma False
Fecha False
Ndia False
Atendidas False
dtype: bool

In [20]: # Verificación de columnas de texto
query_consolidado.dtypes == object
Out[20]: Plataforma True
Fecha True
Ndia False
Atendidas False
dtype: bool

In [21]: # Columnas numéricas
df_consolidado_num = (query_consolidado.dtypes == float)|(query_consolidado.dtypes == int)
# Verificación de índice de las columnas
df_consolidado_num.index
Out[21]: Index(['Plataforma', 'Fecha', 'Ndia', 'Atendidas'], dtype='object')
```

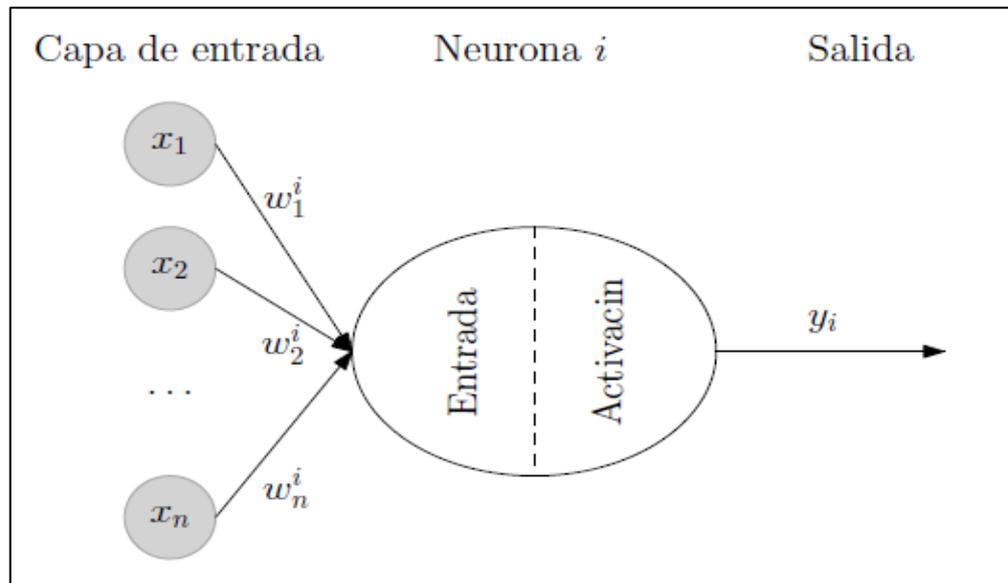
En la figura 16 se empezará a analizar la data mediante graficos que para ello se utilizará las librerías, Matplotlib y Seaborn.



## ELEGIR MODELOS

### Figura 18. Esquema de una Neurona Artificial

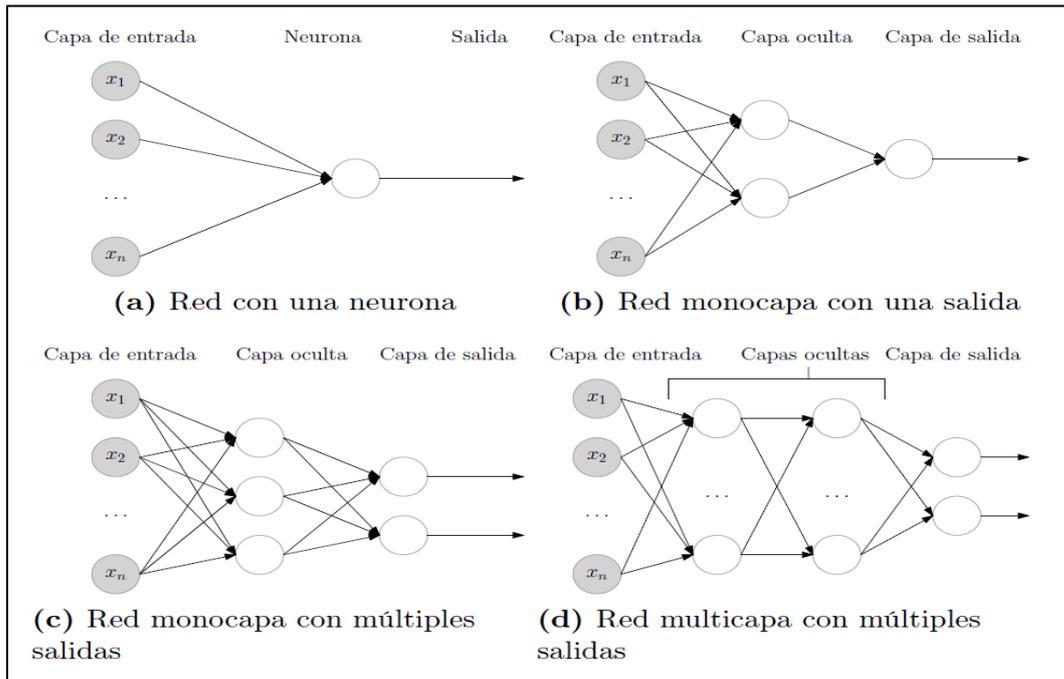
En la figura 18 se esquematiza la neurona artificial



Fuente: David Freeman (2018)

### Figura 19: Arquitectura de red Neuronal

En la figura 19 se observa los tipos de arquitectura de la red neuronal



**Figura 20: Regresión Lineal entorno Jupyter**

En la figura 20 se observa la codificación en el entorno Jupyter de la regresión Lineal

```

3. Regresión Lineal

In [ ]: from sklearn.linear_model import LinearRegression
        from sklearn.metrics import mean_squared_error
        while True:
            y = query_consolidado[['Atendidas']]
            x = query_consolidado[['Ndía']]

            X_train, X_test, y_train, y_test = train_test_split(x,y,test_size = 0.9)

            #Defino el algoritmo a utilizar
            lr = linear_model.LinearRegression()
            #Entreno el modelo
            lr.fit(X_train, y_train)
            #Realizo una predicción
            y_pred = lr.predict(X_test)
            #Graficamos los datos junto con el modelo
            plt.scatter(X_test, y_test)
            plt.plot(X_test, y_pred, color='red', linewidth=3)
            plt.title("Modelo de Regresión Lineal")
            plt.xlabel("Ndía")
            plt.ylabel("Número de atendidas")
            plt.show()
            print()
            print("DATOS DEL MODELO REGRESIÓN LINEAL")
            print()
            print("Valor de la pendiente o coeficiente "a":")
            print(lr.coef_)
            print("Valor de la intersección o coeficiente "b":")
            print(lr.intercept_)
            print()
            print("La ecuación del modelo es igual a:")
            print('y = ', lr.coef_, 'x', lr.intercept_)
            print()

```

```
#realizo una prediccion
Y_pred = lr.predict(X_test)
#Graficamos los datos junto con el modelo
plt.scatter(X_test, y_test)
plt.plot(X_test, Y_pred, color='red', linewidth=3)
plt.title('Modelo de Regresión Lineal')
plt.xlabel('Dia')
plt.ylabel('Número de atencidas')
plt.show()
print()
print('DATOS DEL MODELO REGRESIÓN LINEAL')
print()
print('Valor de la pendiente o coeficiente "a":')
print(lr.coef_)
print('Valor de la intersección o coeficiente "b":')
print(lr.intercept_)
print()
print('La ecuación del modelo es igual a:')
print('y = ', lr.coef_, 'x', lr.intercept_)
print()
print('Precisión del modelo:')
print(lr.score(X_train, y_train))
if lr.score(X_train, y_train) > 0.95:
    break
print("La planificación es: ")
print(Y_pred)
print(X_test)
Y_pred = pd.DataFrame(lr.predict(Y_pred))
llamadas_reales = pd.DataFrame(X_test)
prediction = pd.DataFrame(Y_pred)
pandas = llamadas_reales.join(prediction, lsuffix='_caller', rsuffix='_other')
pandas.to_excel(r"C:\Users\Anthony\Documents\Study\OCV\2019 - 1\Desarrollo del Proyecto de Investigacion\Repaso\Predicciones")
```

En la figura 20 se ejecutará un score donde la predicción sea aceptable para este caso, dentro de las líneas de la regresión lineal pusimos un WHILE y tener un score mayor a 0.95, mientras más cercano sea a 1 es mejor.

## Figura 21: Ejecución del Score

En la figura 21 se observa la ejecución teniendo como score 0.59 gráficamente en la regresión Lineal

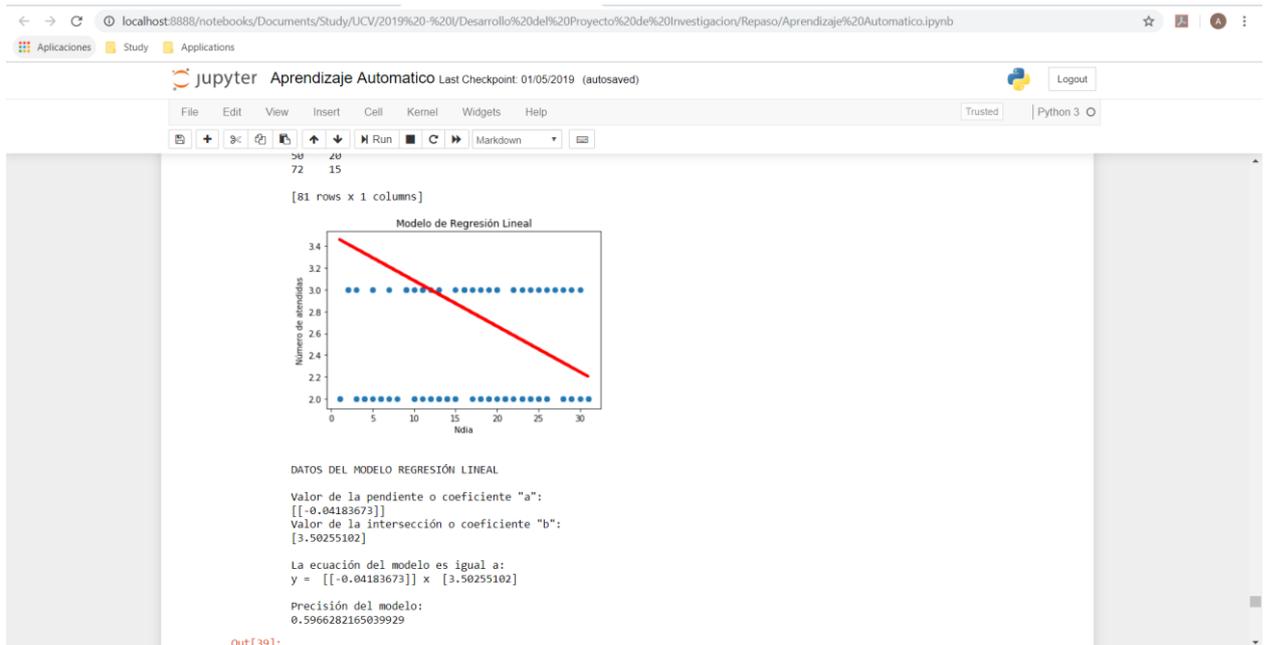


Figura 22: Red neuronal en el entorno Spyder

En la figura 22 se observa el entorno Spyder donde se utilizará para mejores resultados de la predicción.

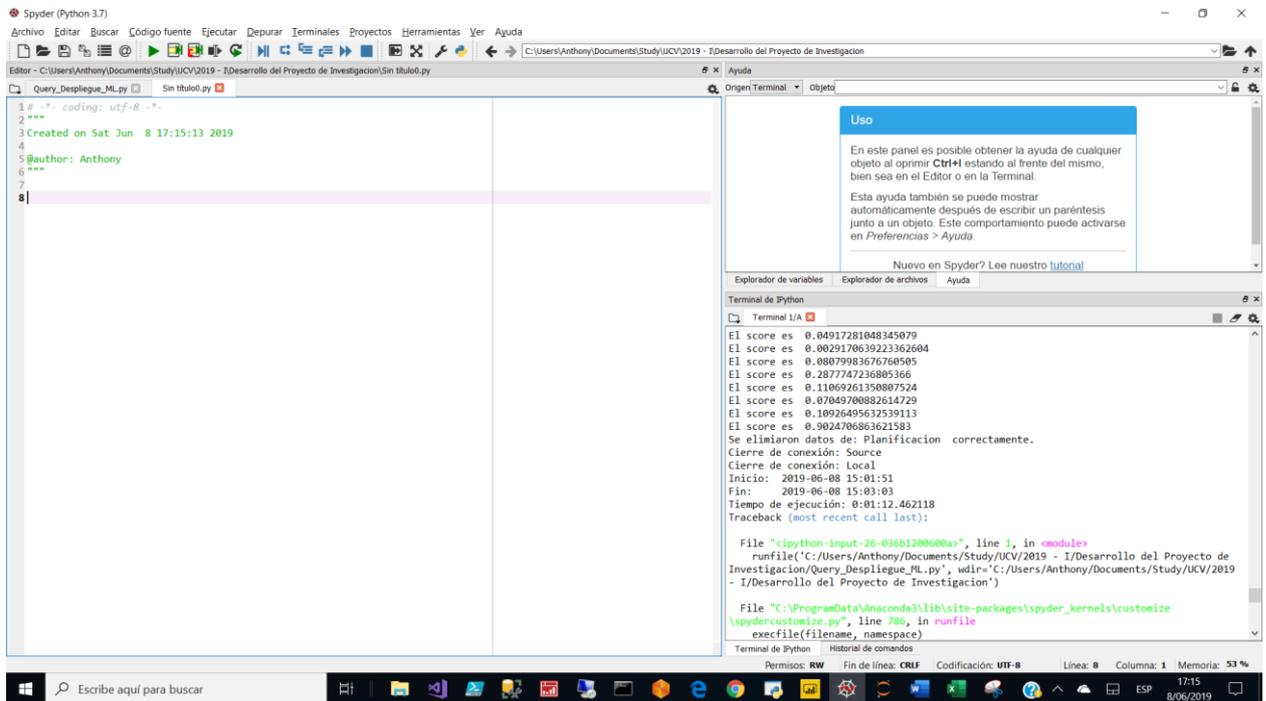


Figura 23: Script Spyder

```

Spyder (Python 3.7)
Archivo Editar Buscar Código fuente Ejecutar Depurar Terminales Proyectos Herramientas Ver Ayuda
Editor - C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigacion\Query_Despliegue_ML.py
Query_Despliegue_ML.py
1 #-*- coding: utf-8 -*-
2 """
3 Created on Tue May 14 20:14:48 2019
4
5 @author: Anthony Simon Altamirano
6 """
7
8 import numpy as np
9 import pandas as pd
10 import matplotlib.pyplot as plt
11 import seaborn as sb
12 from datetime import datetime, timedelta
13 import os
14 import sqlalchemy
15 import time
16 from ConnectionString import Connection_String
17 from sklearn.neural_network import MLPRegressor
18 from sklearn.model_selection import train_test_split
19
20
21 startprocess = datetime.now()
22 time.sleep(2)
23
24 """ Bloque de Funciones """
25
26 def validadiondataframe(d):
27     if len(d) <= 0:
28         print("El Conjunto de datos esta vacio")
29     else:
30         d = d
31     return d
32
33
34 def settingloaddata(obj, cn):
35     try:
36         df = pd.read_sql(obj, cn, index_col='Id')
37         df = df[df['canceled'] == 0]
38     except ValueError:
39         er = os.sys.exec_info()
40         er_one = 'No se pudo realizar la conexión'
41         return er, er_one
42     else:
43

```

**Uso**

En este panel es posible obtener la ayuda de cualquier objeto al oprimir **Ctrl+H** estando al frente del mismo, bien sea en el Editor o en la Terminal.

Esta ayuda también se puede mostrar automáticamente después de escribir un paréntesis junto a un objeto. Este comportamiento puede activarse en **Preferencias > Ayuda**.

Nuevo en Spyder? Lee nuestro [tutorial](#)

Terminal de Python

```

Terminal I/A
El score es 0.04917281048345079
El score es 0.0029170639223362604
El score es 0.08079983676760585
El score es 0.2877747236805366
El score es 0.11069261350807524
El score es 0.07049700882614729
El score es 0.10926495632539113
El score es 0.9024706863621583
Se elaiaron datos de: Planificacion correctamente.
Cierre de conexión: Source
Cierre de conexión: Local
Inicio: 2019-06-08 15:01:51
Fin: 2019-06-08 15:03:03
Tiempo de ejecución: 0:01:12.462118
Traceback (most recent call last):
File "C:\python-input-26-036b1200600a", line 1, in <module>
runfile('C:/Users/Anthony/Documents/Study/UCV/2019 - I/Desarrollo del Proyecto de Investigacion/Query_Despliegue_ML.py', wdir=C:/Users/Anthony/Documents/Study/UCV/2019 - I/Desarrollo del Proyecto de Investigacion')
File "C:\ProgramData\Anaconda3\lib\site-packages\spyder_kernels\customize\spydercustomize.py", line 786, in runfile
execfile(filename, namespace)

```

```

Spyder (Python 3.7)
Archivo Editar Buscar Código fuente Ejecutar Depurar Terminales Proyectos Herramientas Ver Ayuda
Editor - C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigacion\Query_Despliegue_ML.py
Query_Despliegue_ML.py
40 er_one = 'No se pudo realizar la conexión'
41 return er, er_one
42 else:
43     return df
44
45
46 """ Credenciales del servidor de origen de la tabla de configuración """
47 server = "localhost"
48 database = "Konecta"
49 user = "sa"
50 password = "Konecta123"
51 select = 'SELECT * FROM Setting_LoadData Where Id = 2'
52 cnx = Connection_String(server, database, user, password)
53
54
55 """ Conexión a base de datos """
56 table_configuration = settingloaddata(select, cnx)
57 pd.set_option('display.expand_frame_repr', False)
58
59 try:
60     for regconf in table_configuration.itertuples():
61         """ Obtener datos en variables de origen """
62         driver_source = regconf[1]
63         server_source = regconf[2]
64         user_source = regconf[3]
65         password_source = regconf[4]
66         database_source = regconf[5]
67         scheme_source = regconf[6]
68         objeto_source = regconf[7]
69         objeto_dest = regconf[15]
70         user_dest = regconf[11]
71         password_dest = regconf[12]
72         server_dest = regconf[10]
73         database_dest = regconf[13]
74         driver_dest = regconf[16]
75         scheme_dest = regconf[14]
76         query_delete = regconf[17]
77
78         """ Abrir conexión Objeto Origen """
79         cnx_source = Connection_String(server_source, database_source, user_source, password_source)
80

```

**Uso**

En este panel es posible obtener la ayuda de cualquier objeto al oprimir **Ctrl+H** estando al frente del mismo, bien sea en el Editor o en la Terminal.

Esta ayuda también se puede mostrar automáticamente después de escribir un paréntesis junto a un objeto. Este comportamiento puede activarse en **Preferencias > Ayuda**.

Nuevo en Spyder? Lee nuestro [tutorial](#)

Terminal de Python

```

Terminal I/A
El score es 0.04917281048345079
El score es 0.0029170639223362604
El score es 0.08079983676760585
El score es 0.2877747236805366
El score es 0.11069261350807524
El score es 0.07049700882614729
El score es 0.10926495632539113
El score es 0.9024706863621583
Se elaiaron datos de: Planificacion correctamente.
Cierre de conexión: Source
Cierre de conexión: Local
Inicio: 2019-06-08 15:01:51
Fin: 2019-06-08 15:03:03
Tiempo de ejecución: 0:01:12.462118
Traceback (most recent call last):
File "C:\python-input-26-036b1200600a", line 1, in <module>
runfile('C:/Users/Anthony/Documents/Study/UCV/2019 - I/Desarrollo del Proyecto de Investigacion/Query_Despliegue_ML.py', wdir=C:/Users/Anthony/Documents/Study/UCV/2019 - I/Desarrollo del Proyecto de Investigacion')
File "C:\ProgramData\Anaconda3\lib\site-packages\spyder_kernels\customize\spydercustomize.py", line 786, in runfile
execfile(filename, namespace)

```

Spyder (Python 3.7)

Archivo Editar Buscar Código fuente Ejecutar Depurar Terminales Proyectos Herramientas Ver Ayuda

Editor - C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigación\Query\_Despliegue\_ML.py

```

79 """ Abrir conexión Objeto Origen """
80 cnx_source = Connection_String(server_source, database_source, user_source, password_source)
81
82 """ Abrir conexión Objeto Destino """
83 cnx_dest = Connection_String(server_dest, database_dest, user_dest, password_dest)
84
85 """ Consultar objeto """
86 query = objeto_source
87
88
89
90 try:
91     print('Consultando: %s' % query, '....')
92     query = pd.read_sql_query(query, cnx_source)
93     print(query)
94     y = np.array(query['Atendidas'])
95     x = np.array(query['Dia'])
96
97     X = x[:, np.newaxis]
98
99
100
101 while True:
102     X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.8) #, test_size = 0.8
103
104     mlr = MLPRegressor(solver="lbfgs", alpha=1e-5, hidden_layer_sizes=(3,3), random_state=1)
105     mlr.fit(X_train, y_train)
106     print("El score es " ,mlr.score(X_train, y_train))
107     if mlr.score(X_train, y_train) >= 0.75:
108         break
109
110 Y_pred = pd.DataFrame(mlr.predict(X_test))
111 llamadas_reales = pd.DataFrame(X_test)
112 prediction = pd.DataFrame(Y_pred)
113
114 pandas = llamadas_reales.join(prediction, lsuffix='Dia', rsuffix='PrediccionAtendida')
115
116 predictionAtendida = pd.DataFrame(pandas['@PrediccionAtendida'].unique())
117 numeroDia = pd.DataFrame(pandas['@Dia'].unique())
118 pandasSQL = numeroDia.join(predictionAtendida, lsuffix='Dia', rsuffix='PrediccionAtendida')
119
120 """ Antes de insertar eliminar datos de la tabla con la misma fecha. """

```

Origen Terminal | Objeto

Uso

En este panel es posible obtener la ayuda de cualquier objeto al oprimir **Ctrl+H** estando al frente del mismo, bien sea en el Editor o en la Terminal.

Esta ayuda también se puede mostrar automáticamente después de escribir un paréntesis junto a un objeto. Este comportamiento puede activarse en **Preferencias > Ayuda**.

Nuevo en Spyder? Lee nuestro [tutorial](#)

Explorador de variables | Explorador de archivos | Ayuda

Terminal de Python

Terminal 1/A

```

El score es 0.04917281048345079
El score es 0.0029170639223362604
El score es 0.08079983676766585
El score es 0.2877747236885366
El score es 0.11069261358807524
El score es 0.0704970882614729
El score es 0.10926495632539113
El score es 0.9024706863621583
Tiempo de ejecución: 0:01:12.462118
Se eliminaron datos de: Planificación correctamente.
Cierre de conexión: Source
Cierre de conexión: Local
Inicio: 2019-06-08 15:01:51
Fin: 2019-06-08 15:03:03
Traceback (most recent call last):
  File "C:\ProgramData\Anaconda3\lib\site-packages\spyder_kernels\customize\spydercustomize.py", line 786, in runfile
    execfile(filename, namespace)

```

Permisos: RW Fin de línea: CRLF Codificación: UTF-8 Línea: 156 Columna: 48 Memoria: 53 %

Spyder (Python 3.7)

Archivo Editar Buscar Código fuente Ejecutar Depurar Terminales Proyectos Herramientas Ver Ayuda

Editor - C:\Users\Anthony\Documents\Study\UCV\2019 - I\Desarrollo del Proyecto de Investigación\Query\_Despliegue\_ML.py

```

119 """ Antes de insertar eliminar datos de la tabla con la misma fecha. """
120
121 #print('Eliminando registros de: %s del %s al %s ...' % objeto_dest)
122 cursor = cnx_dest.cursor()
123 query_del = query_delete
124 with cursor.execute(query_delete):
125     print('Se eliminaron datos de: %s' % objeto_dest, " correctamente.")
126
127 """ Insertar datos """
128 #print('Insertando datos en %s' % objeto_dest)
129 #engine = sqlalchemy.create_engine("mysql+pymysql://" + user_dest + ":" + password_dest + "@" + serve
130 #pandasSQL.to_sql(objeto_dest, engine, schema=schema_dest, if_exists='append', chunksize=50, index=
131 #print('Se insertaron datos en: %s' % objeto_dest, " correctamente.")
132
133 except:
134     print('---- Se produjo un error ----')
135     print(os.sys.exc_info())
136     cnx_source.close()
137     cnx_dest.close()
138     print('Cierre de conexión: Source')
139 finally:
140     cnx_source.close()
141     cnx_dest.close()
142     print('Cierre de conexión: Source')
143     cnx.close()
144     print('Cierre de conexión: Local')
145
146 except:
147     print('---- Se produjo un error ----')
148     print(os.sys.exc_info())
149     cnx.close()
150     print('Cierre de conexión: Local')
151
152 endprocess = datetime.now()
153 timeexec = endprocess - startprocess
154
155 print('Inicio: %s' % startprocess.strftime('%Y-%m-%d %H:%M:%S'))
156 print('Fin: %s' % endprocess.strftime('%Y-%m-%d %H:%M:%S'))
157 print('Tiempo de ejecución: %s' % timeexec)
158 #query consolidado = query
159 #print(query consolidado.info)

```

Origen Terminal | Objeto

Uso

En este panel es posible obtener la ayuda de cualquier objeto al oprimir **Ctrl+H** estando al frente del mismo, bien sea en el Editor o en la Terminal.

Esta ayuda también se puede mostrar automáticamente después de escribir un paréntesis junto a un objeto. Este comportamiento puede activarse en **Preferencias > Ayuda**.

Nuevo en Spyder? Lee nuestro [tutorial](#)

Explorador de variables | Explorador de archivos | Ayuda

Terminal de Python

Terminal 1/A

```

El score es 0.04917281048345079
El score es 0.0029170639223362604
El score es 0.08079983676766585
El score es 0.2877747236885366
El score es 0.11069261358807524
El score es 0.0704970882614729
El score es 0.10926495632539113
El score es 0.9024706863621583
Tiempo de ejecución: 0:01:12.462118
Se eliminaron datos de: Planificación correctamente.
Cierre de conexión: Source
Cierre de conexión: Local
Inicio: 2019-06-08 15:01:51
Fin: 2019-06-08 15:03:03
Traceback (most recent call last):
  File "C:\ProgramData\Anaconda3\lib\site-packages\spyder_kernels\customize\spydercustomize.py", line 786, in runfile
    execfile(filename, namespace)

```

Permisos: RW Fin de línea: CRLF Codificación: UTF-8 Línea: 156 Columna: 48 Memoria: 53 %

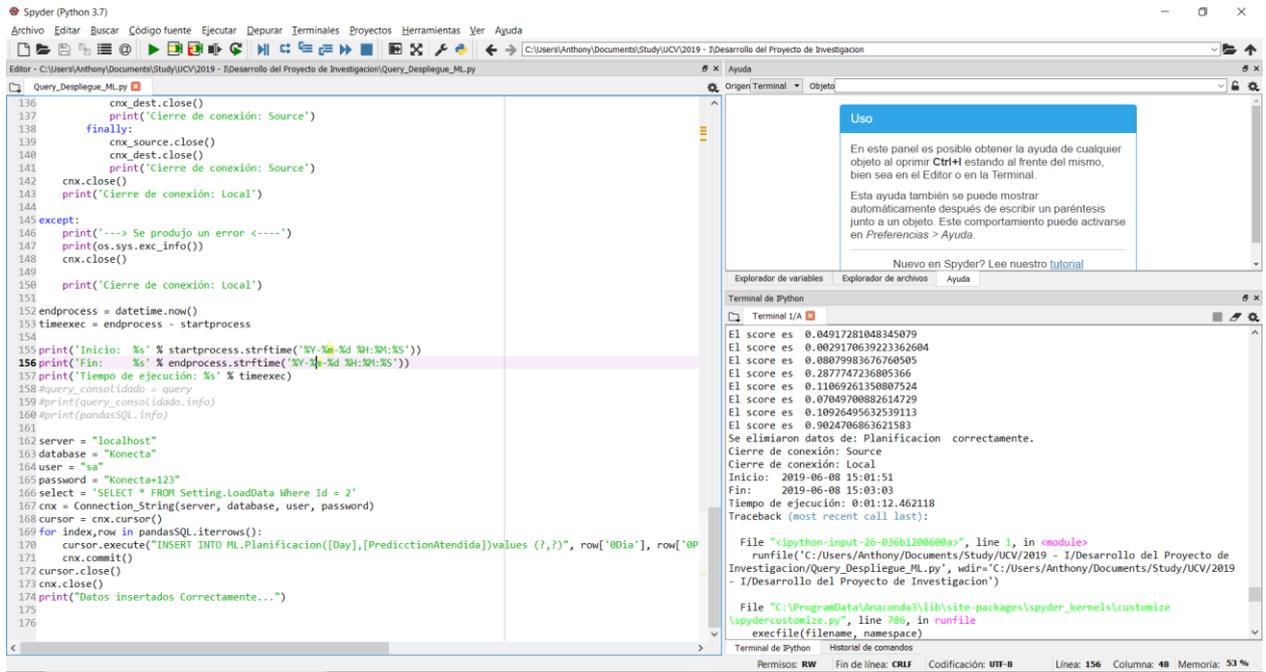


Figura 24: Inserción de datos

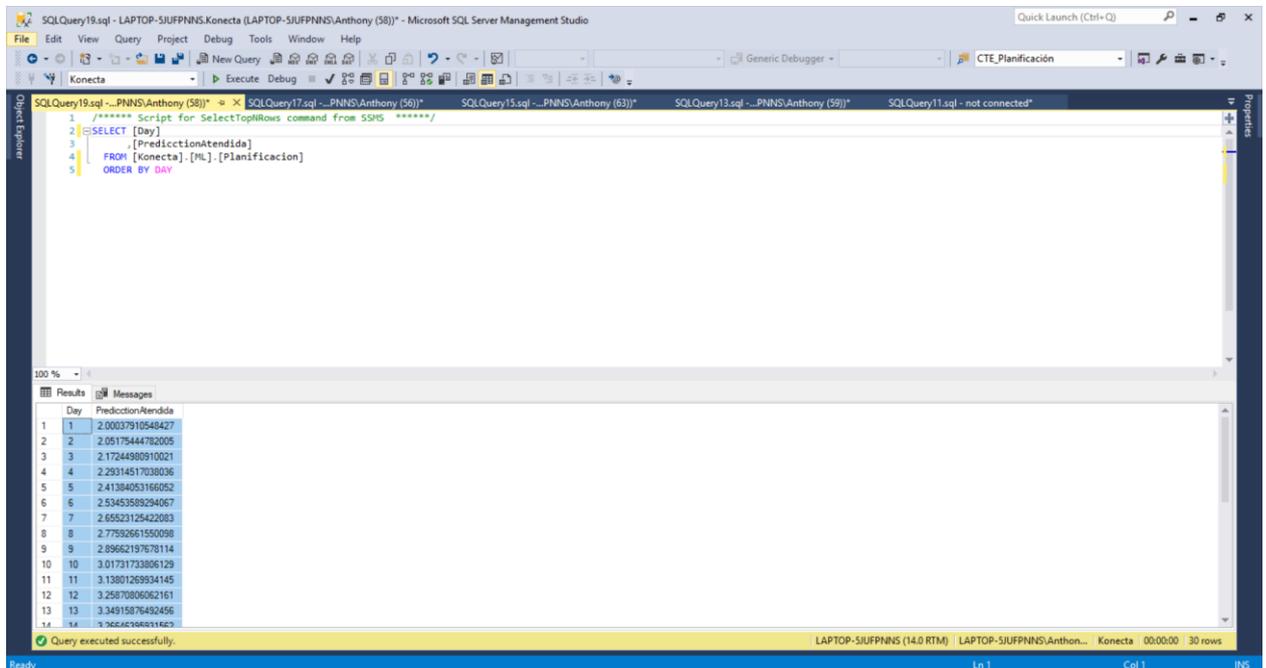
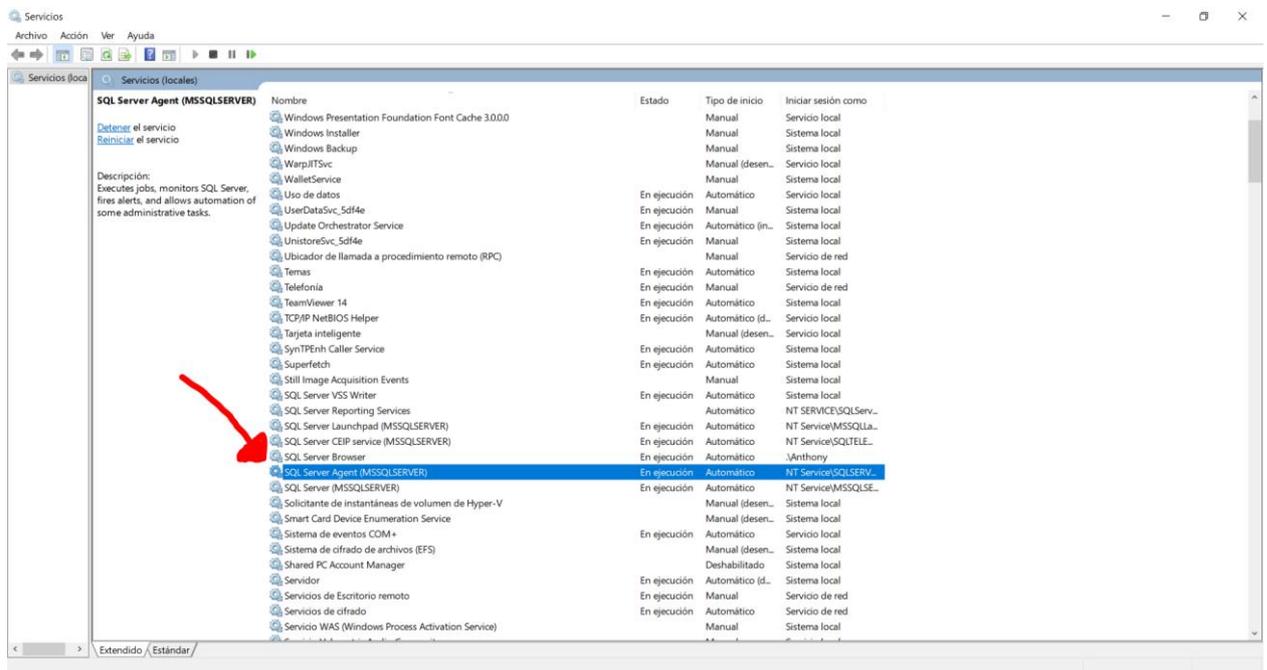


Figura 25: Procedimiento Almacenado

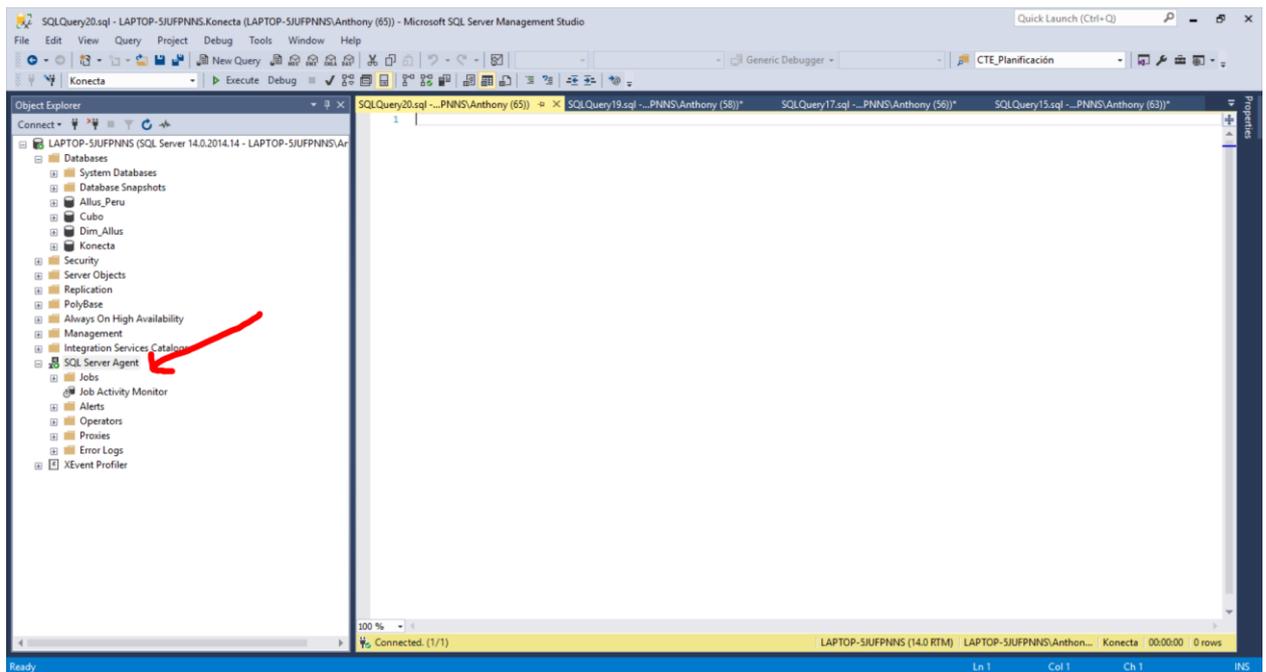
```
1 USE [Koneccta]
2 GO
3 /***** Object: StoredProcedure [dbo].[Consolidado_PhBI]    Script Date: 6/06/2019 21:00:54 *****/
4 SET ANSI_NULLS ON
5 GO
6 SET QUOTED_IDENTIFIER ON
7 GO
8
9 ALTER PROCEDURE [dbo].[Consolidado_PhBI]
10 AS
11
12 SET LANGUAGE SPANISH
13
14 ;
15 WITH
16 CTE_Planificación
17 AS
18 (
19 SELECT 2019 AS Anio, 5 AS Mes, 'Mayo' mMes, CAST('2019-' + '05-' + IIF(LEN(CAST([DAY] AS varchar(2))) < 2, '0' + CAST([DAY] AS varchar(2)), CAST([DAY] AS varchar(2))) AS DATE) Fecha, 'TECNICO_NEGOCIOS' AS Plata
20 FROM [HL].[Planificacion]
21 )
22 ,
23 CTE_Integrado
24 AS
25 (
26 SELECT Fecha, SRV_ID, SERVICIO, SUM(COALESCE(Logeo, 0)) / 3600.0 HorasLogeo
27 FROM [Koneccta].[dbo].[Integrado]
28 WHERE SERVICIO = 'TECNICO_NEGOCIOS' --AND Fecha >= '20190501'
29 GROUP BY Fecha, SERVICIO, SRV_ID
30 )
31 ,
32 CTE_PrePla
33 AS
34 (
35 SELECT COALESCE(c.Anio, p.Anio) Anio
36 , COALESCE(c.Mes, p.Mes) Mes
37 , COALESCE(c.NMes, P.NMES) NMES
38 , COALESCE(c.Fecha, P.FECHA) FECHA
39 , COALESCE(c.Plataforma, P.Plataforma) Servicio
```

Ahora para la implementación del despliegue utilizaremos el servicio de SQL Server Agent.

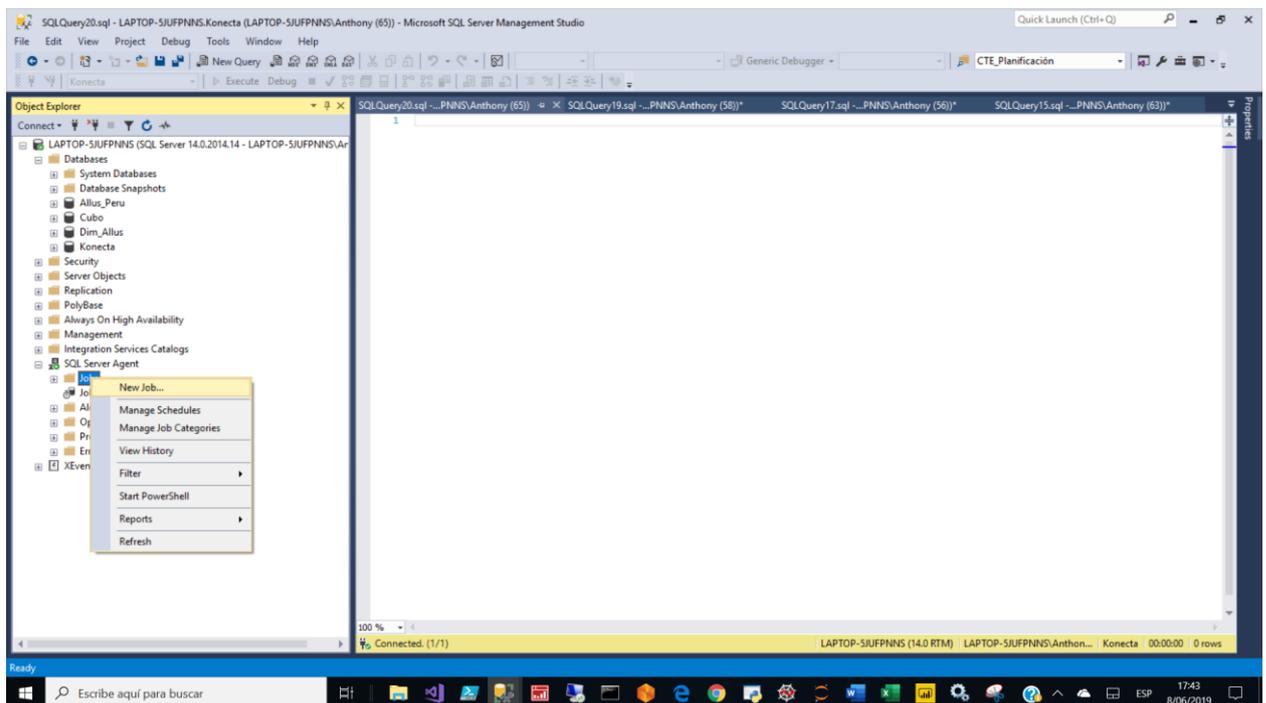
Verificamos el estado del servicio tiene que estar en ejecución:



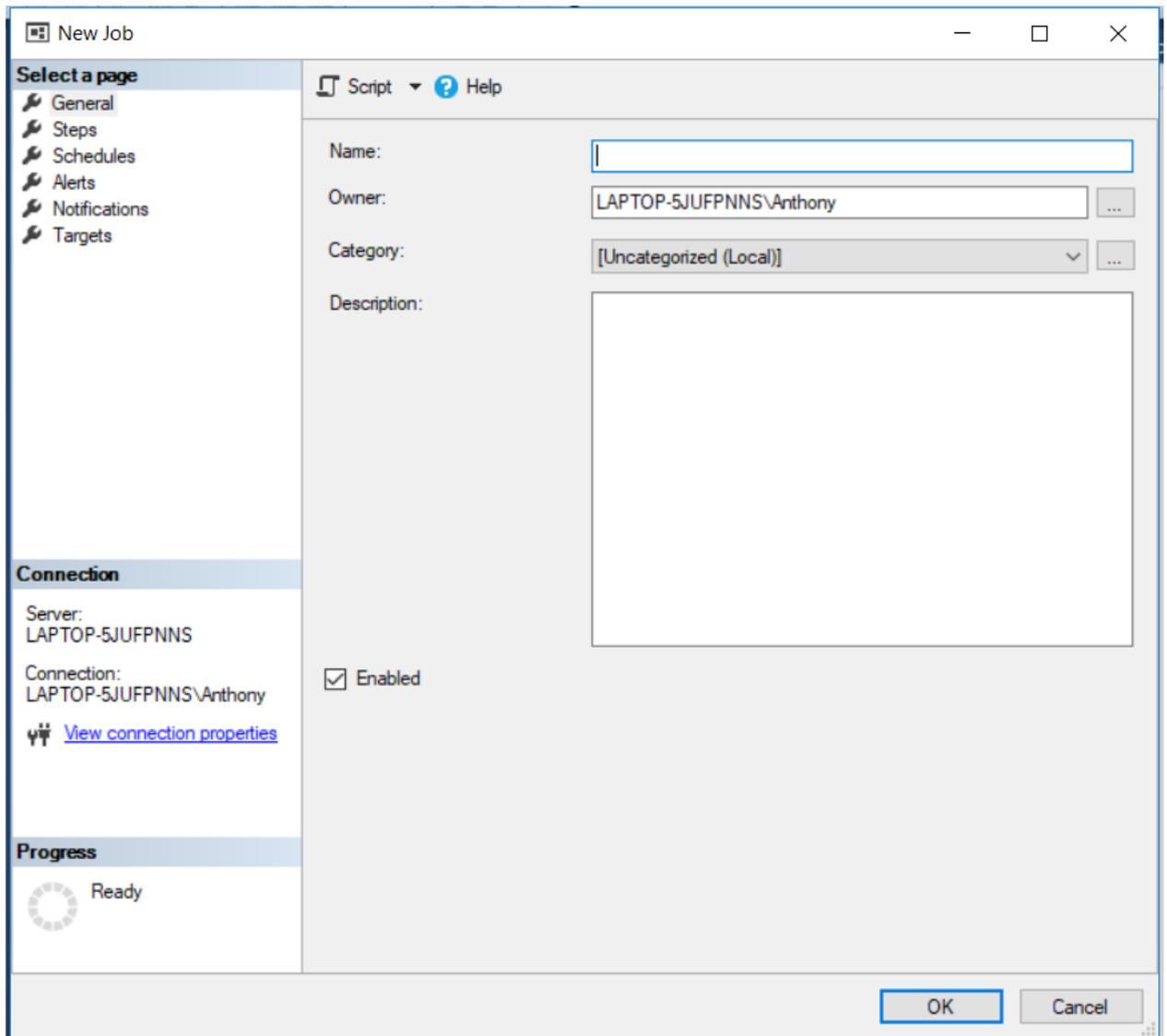
Ahora lo vemos en el SQL Server:



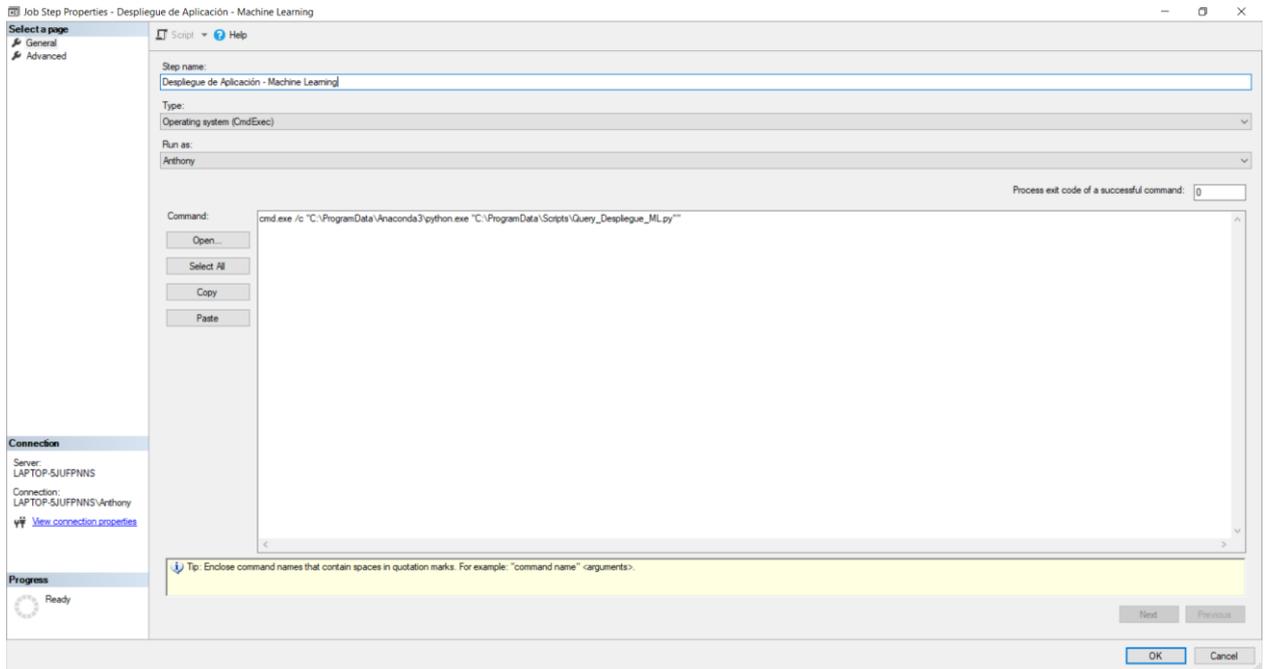
Ahora generamos un nuevo job de la siguiente manera:



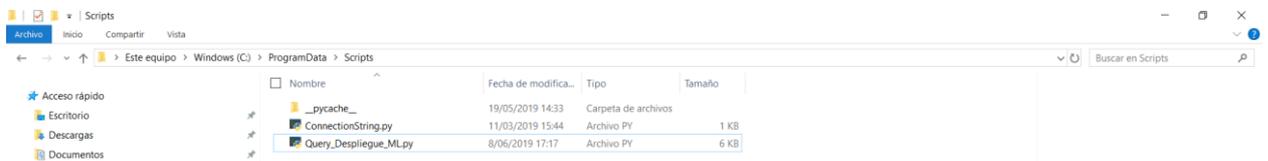
**Figura 26: Inserción de nuevo JOB**



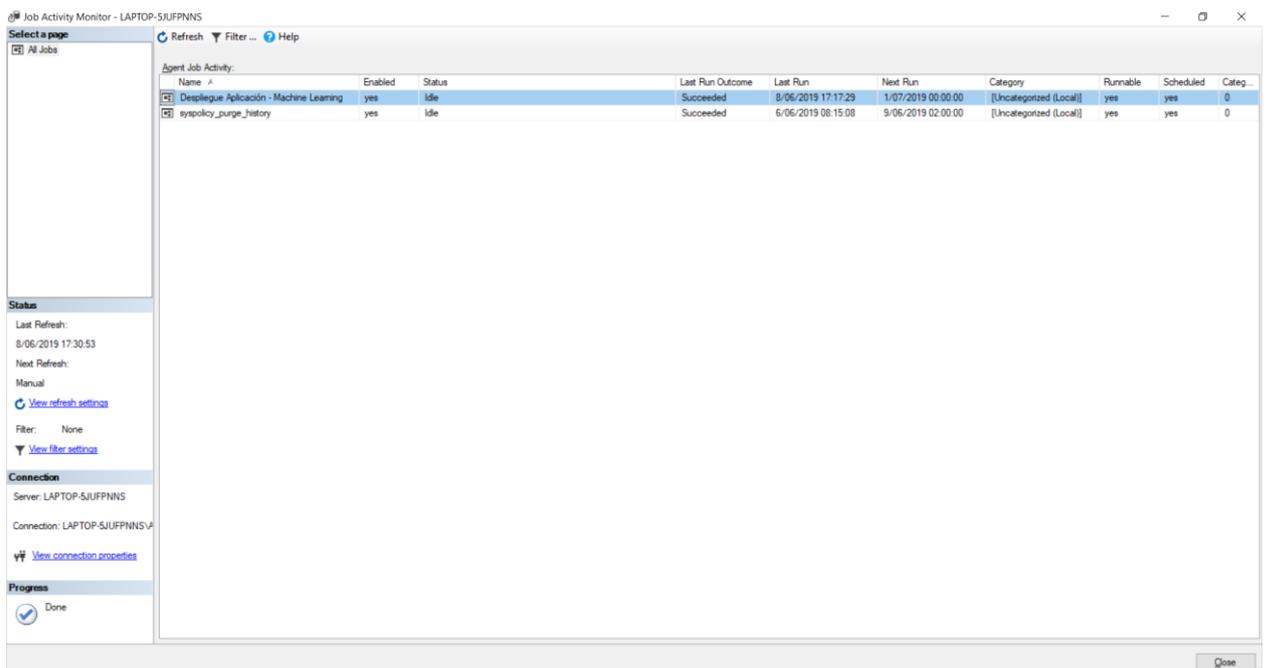
Y escogemos las siguientes opciones para ejecutar el código Python:

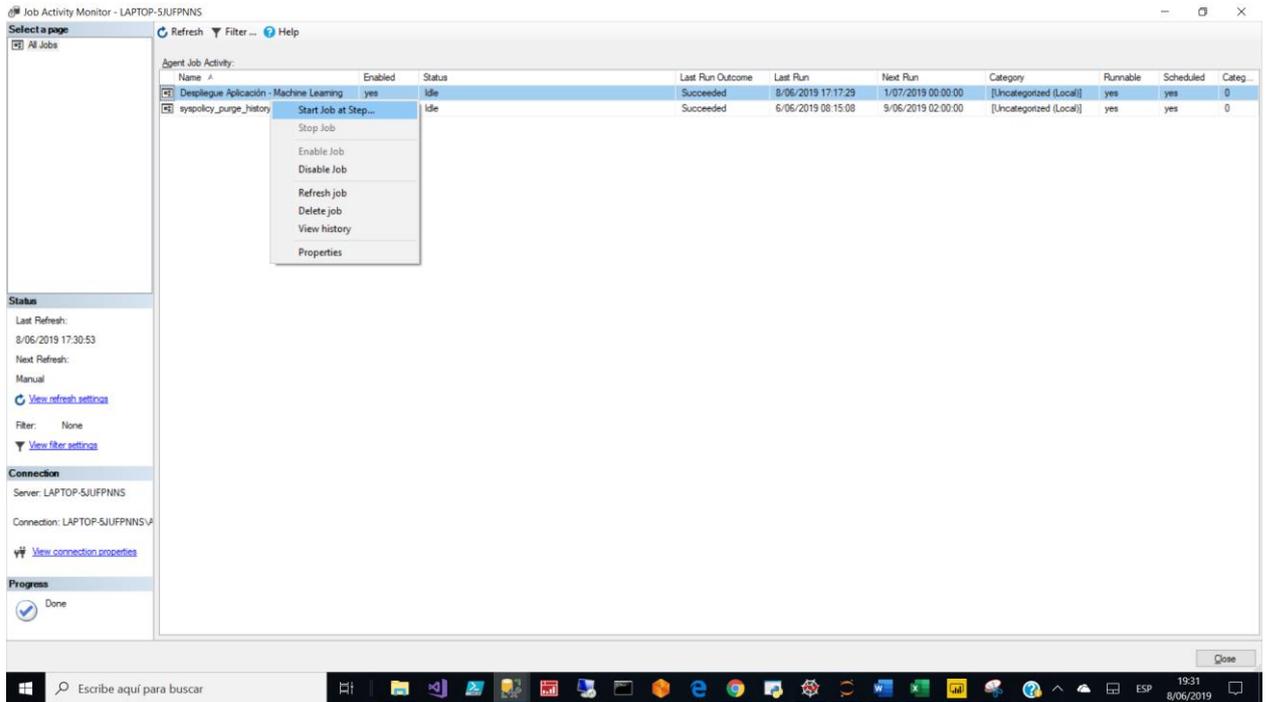


Previamente debemos guardar el script de Python en la siguiente ruta:

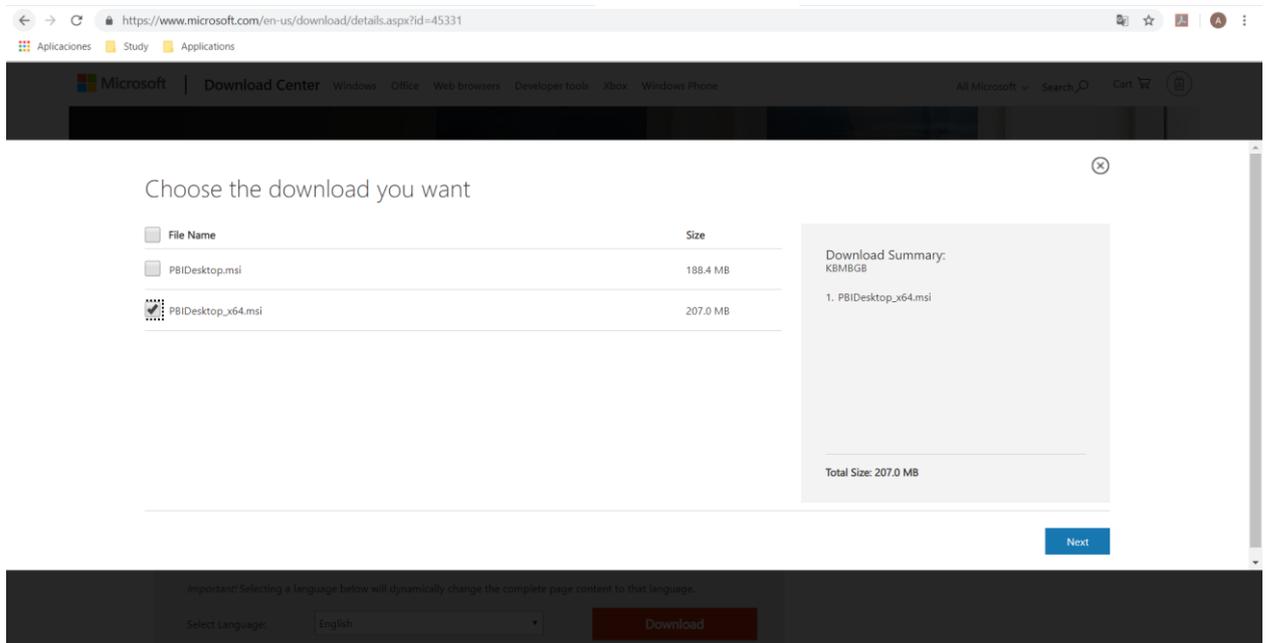


Posteriormente ya podemos observar el Job creado en el SQL Server Agent:





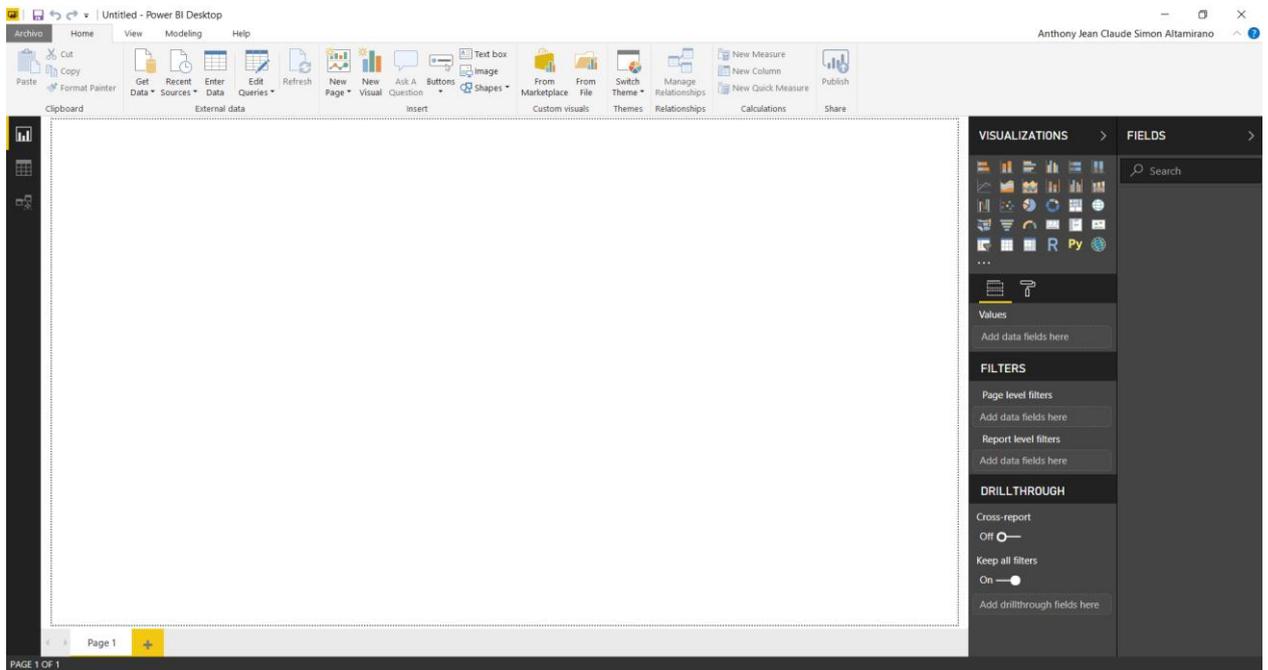
Una vez realizado el despliegue ya podemos empezar a utilizar Power BI como nuestra interfaz donde visualizaremos los datos planificados o las predicciones.



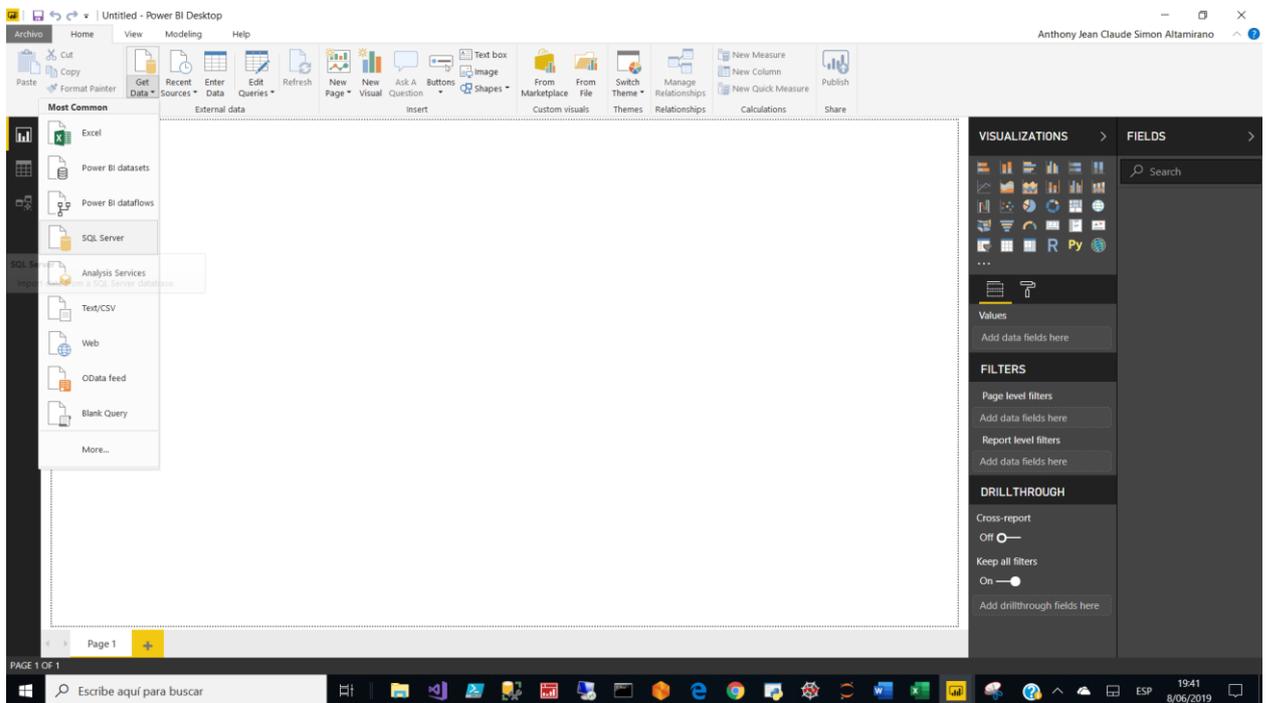
Lo descargamos y ejecutamos.

Ejecutamos el Power BI desktop y se abrirá la aplicación para empezar a realizar la interfaz.

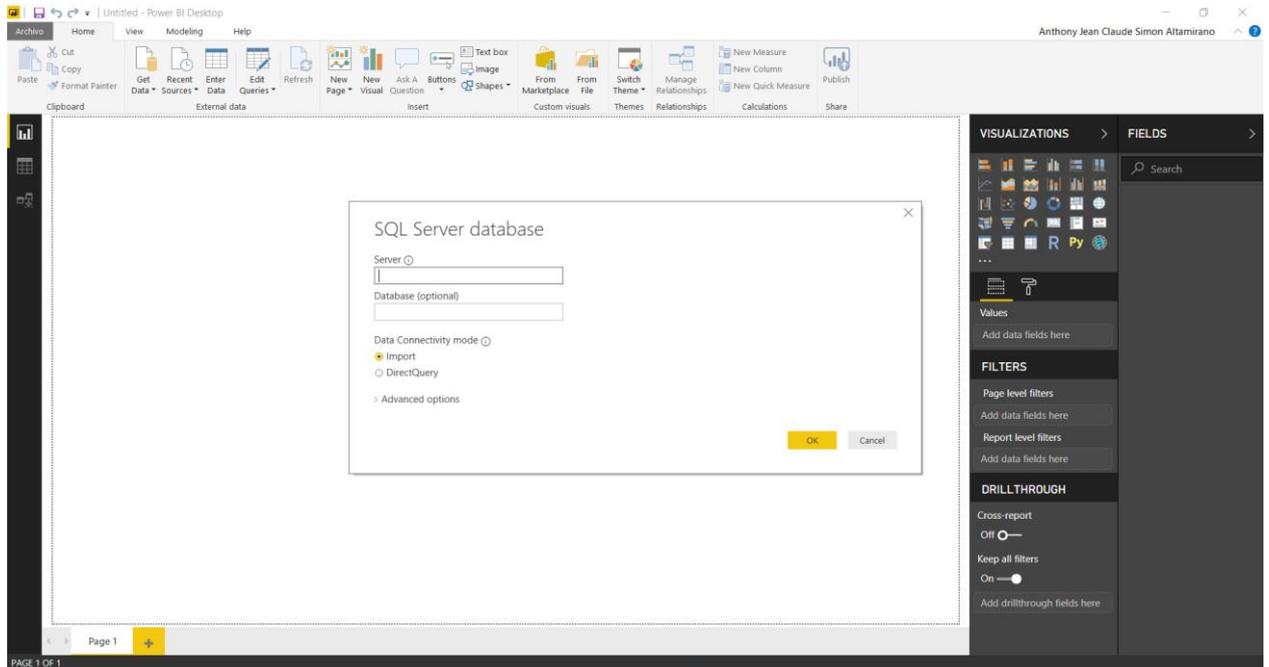
Figura 27: Interfaz Power BI



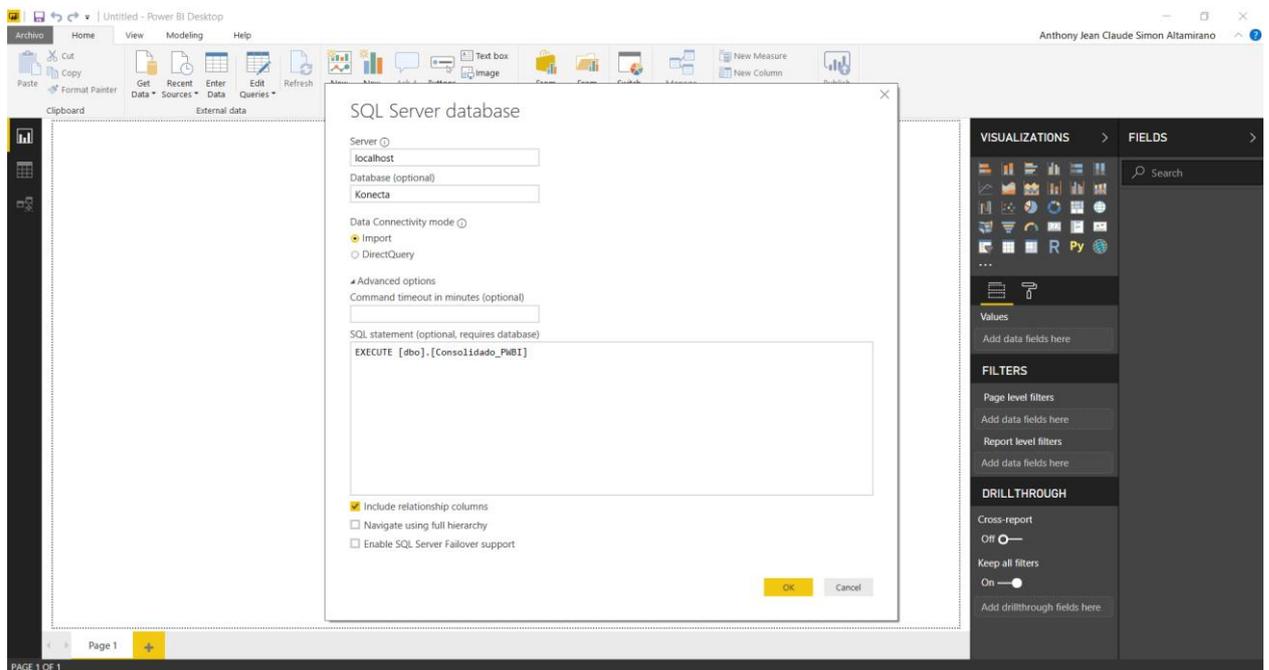
Y nos aparecerá las siguientes opciones:



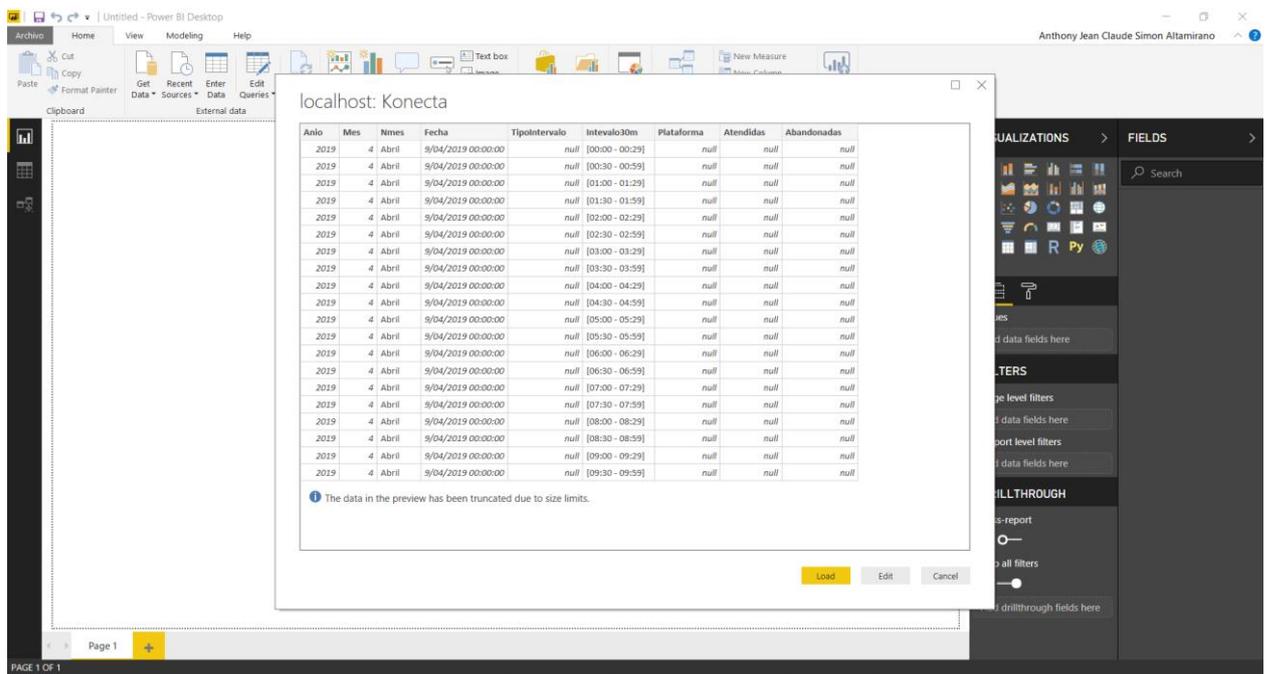
Presionamos en SQL Server y nos solicitara el origen de datos el servidor y las credenciales:



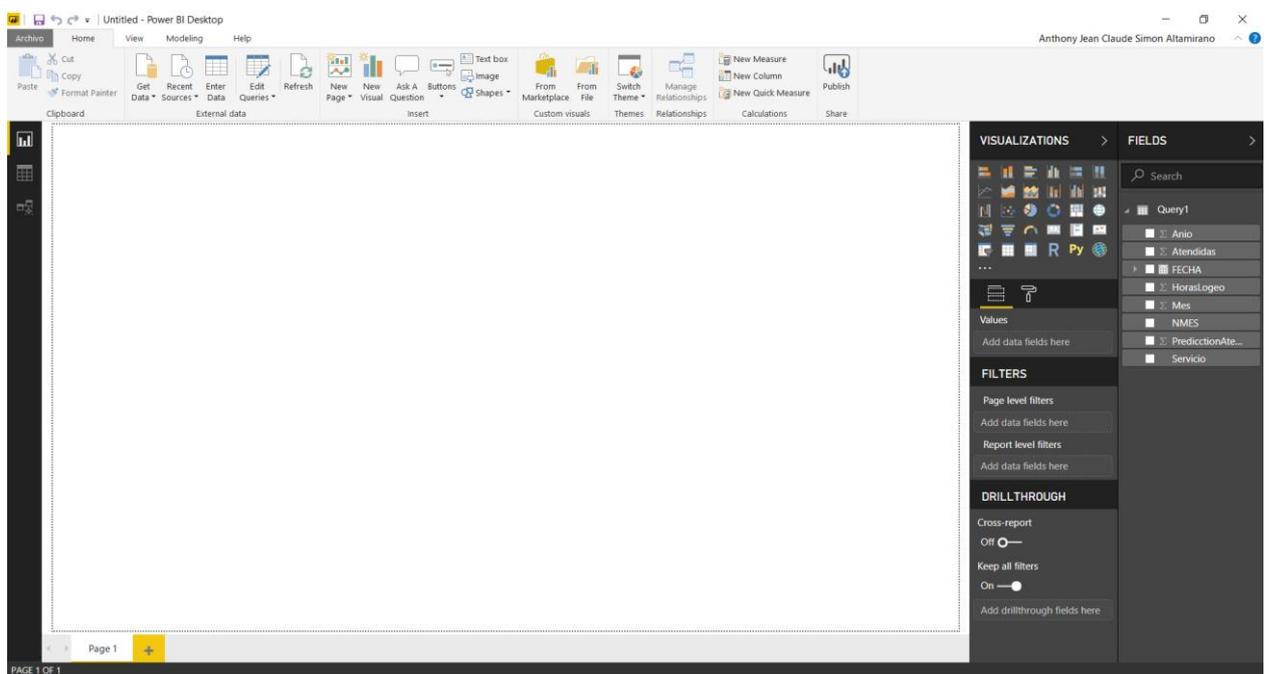
Ahora hacemos la conexión a la base de datos y ejecutaremos un procedimiento almacenado para acceder a los datos como se observa a continuación:



Empezará a cargar los datos y aparecerá la siguiente ventana:



Ahora podemos visualizar nuestros datos:



Y empezamos armar el Dashboard.

En el siguiente Dashboard podemos ver las predicciones comparando con los datos reales:

Y vemos nuestro indicador eficiencia.

Seleccione Año:

2019

Seleccione Mes:

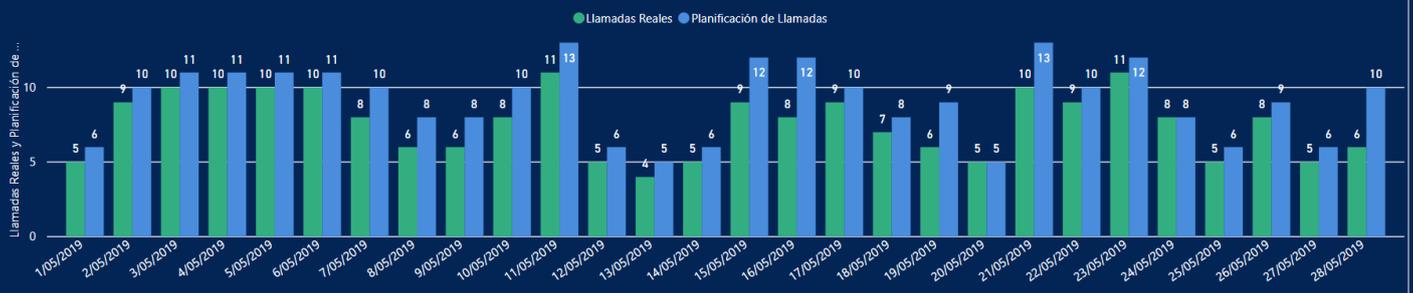
5

Seleccione Servicio:

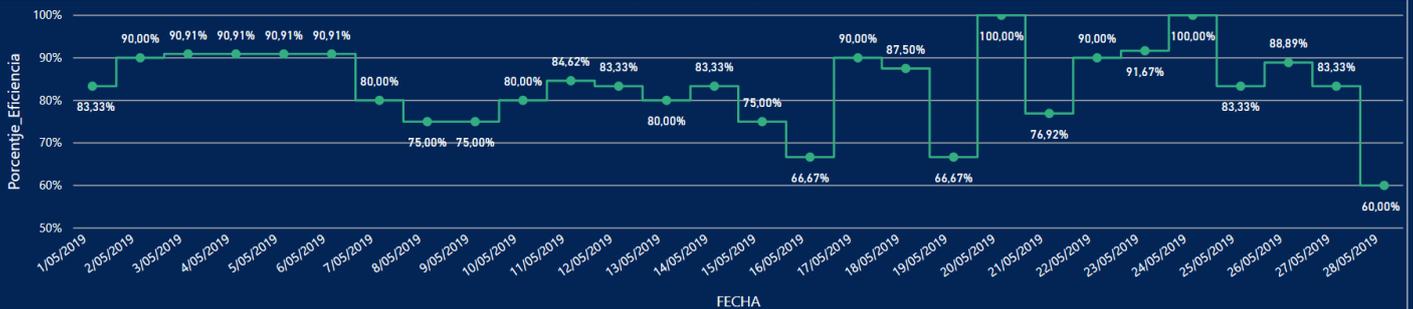
TECNICO\_NEGOCIOS

FILTROS

Llamadas Reales vs Planificación de Llamadas por Fecha



Porcentaje de Eficiencia por Fecha



Eficiencia

Productividad

Consolidado

Integrado

Seleccione Año:

2019

Seleccione Mes:

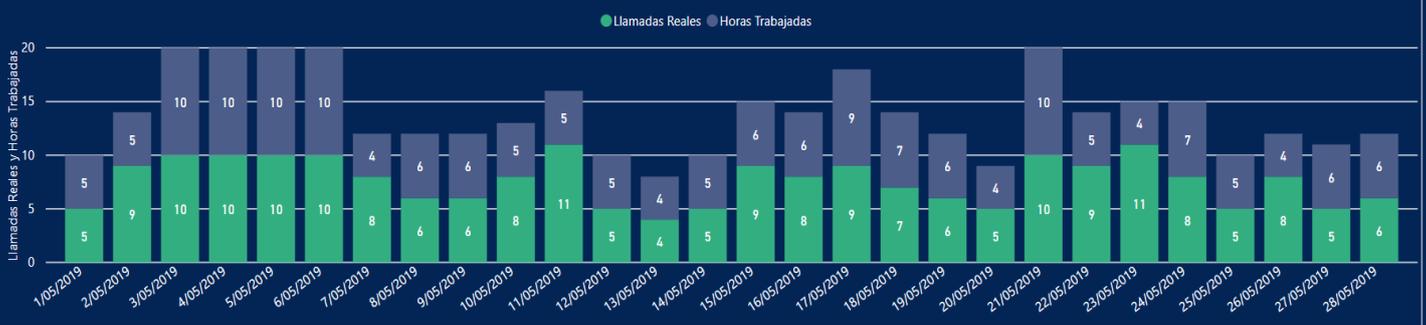
5

Seleccione Servicio:

TECNICO\_NEGOCIOS

FILTROS

Llamadas Reales vs Horas Trabajadas por Fecha



Porcentaje de Productividad por Fecha

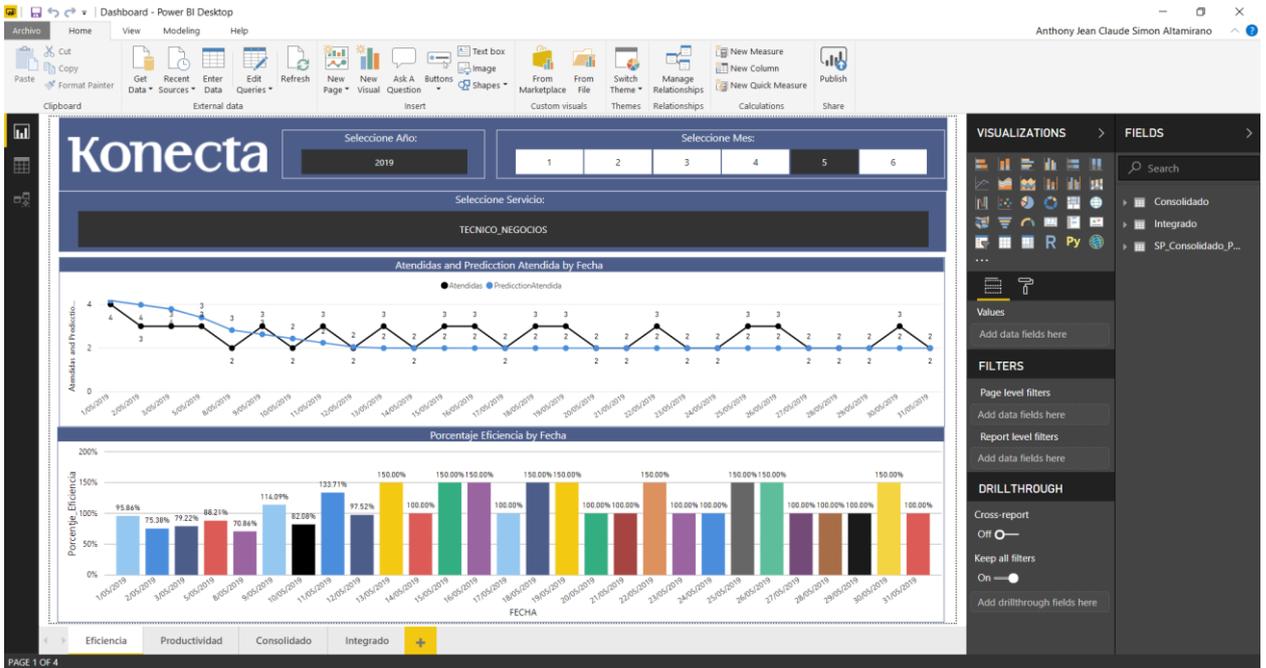


Eficiencia

Productividad

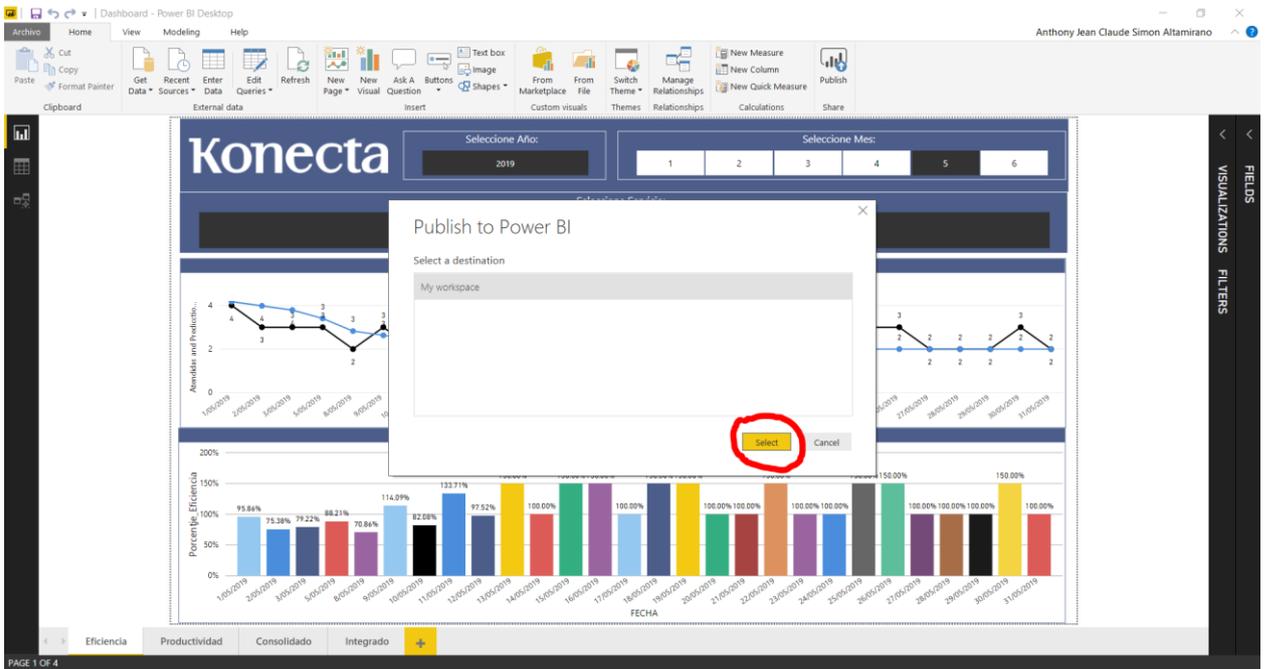
Consolidado

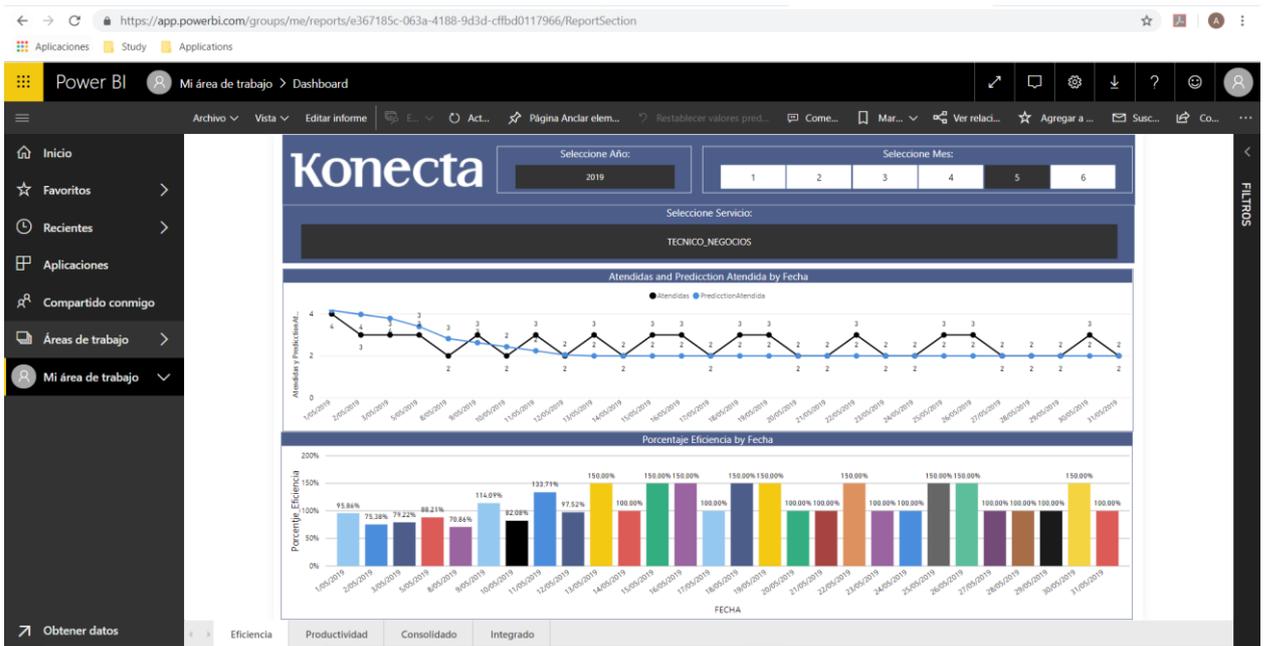
Integrado



Posteriormente lo publicamos en la web como se muestra a continuación, dando click al botón publish debemos haber iniciado sesión antes de dar publish:

Seleccionamos nuestro espacio de trabajo como se ve a continuación:





Aquí se muestra en pantalla completa:



Figura 27: Anaconda



Anaconda brinda a los científicos de datos la información que pueden implementar regularmente en proyectos de inteligencia artificial y aprendizaje automático en la producción a escala, y entregar rápidamente información a los tomadores de decisiones.

¿Cómo afectaría eso a su negocio?

Anaconda Enterprise es compatible con la organización sin importar el tamaño, escalando fácilmente de un solo usuario en una computadora portátil a miles de máquinas. Llegando así a toma decisiones rápidamente con herramientas de uso factible.

