



UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

**Datamart para mejorar la productividad del Equipo de Evaluación de
Calidad de Aguas Residuales de SEDAPAL**

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

AUTOR:

Ccasani Condori, Christian (ORCID: 0000-0001-6684-7245)

ASESOR:

Dr. Chávez Pinillos, Frey Elmer (ORCID: 0000-0003-3785-5259)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

ATE - PERÚ

2021

DEDICATORIA

A mis padres, por el ejemplo de perseverancia y superación que incondicionalmente me han dado, por su apoyo, por los consejos, valores y principios que me han inculcado.

AGRADECIMIENTO

Quiero agradecer a mi familia, por haber sido mi apoyo durante todo este tiempo, también a mis profesores y en general a la universidad por permitirme un espacio en el cual pueda adquirir los conocimientos de un ingeniero.

Índice de Contenidos

Carátula	i
Dedicatoria	iii
Agradecimiento	iv
Índice de contenidos	vi
Índice de tablas	vii
Índice de gráficos y figuras	viii
Resumen	ix
Abstract	x
I. INTRODUCCIÓN	11
II. MARCO TEÓRICO	15
III. METODOLOGÍA	21
3.3 Población, muestra y muestreo	23
3.4 Técnicas e instrumentos de recolección de datos.....	24
3.5 Procedimientos.....	25
3.6 Método de análisis de datos	25
3.7 Aspectos éticos	26
IV. RESULTADOS	26
V. DISCUSIÓN	33
VI. CONCLUSIONES.....	36
V. RECOMENDACIONES.....	37
REFERENCIAS	38
ANEXOS.....	48

Índice de tablas

Tabla 1 Operacionalización de la variable dependiente	23
Tabla 2 Técnica e instrumento de recolección de datos	24
Tabla 3 Validez a través de juicio de expertos	25
Tabla 4: Estadísticos descriptivos del Índice de eficiencia en el Pre y Post de la implementación del Datamart.....	26
Tabla 5: Estadísticos descriptivos del Índice de eficacia en el Pre y Post de la implementación del Datamart.....	27
Tabla 6: Prueba de normalidad del Pre-Test y Post-Test del indicador Índice de Eficiencia.....	29
Tabla 7: Prueba de normalidad del Pre-Test del indicador Índice de eficacia	30

Índice de gráficos y figuras

Figura 1: índice de eficiencia pre y post de la implementación del datamart.....	27
Figura 2: índice de eficiencia pre y post de la implementación del datamart.....	28
Figura 3: Prueba de normalidad del Pre-Test Índice de Eficiencia.....	30
Figura 4: Prueba de normalidad del Post-Test Índice de Eficiencia	30
Figura 5: Prueba de normalidad del Pre-Test Índice de Eficacia.....	31
Figura 6: Prueba de normalidad del Post-Test Índice de Eficacia.....	31

Resumen

La investigación tuvo como objetivo determinar la influencia de la implementación de un datamart en la productividad en el EEC-AR de SEDAPAL.

Teniendo como problemática principal la falta del área de sistemas por ende soporte técnico y la importancia de la gestión de los recursos de tecnologías de información (TI) en la empresa, los cuales impiden dedicar servicios de calidad hacia los usuarios de la empresa, ya que no se encuentran alineados a las TI.

La investigación fue de tipo aplicada, de nivel explicativa, de diseño experimental (pre experimental), considerando una población de 13 registros. Se utilizó como técnica el fichaje y como instrumento la ficha de registro para los indicadores, el análisis estadístico se hizo utilizando el programa SPSS 21.0, lo cual arrojó resultados confiables en la investigación. Finalmente se logró determinar que un datamart en el EEC-AR mejora el índice de eficiencia mejora en 36.46% y mantiene el índice de eficacia al 100%, generando así una mejora en la productividad.

Palabras clave: Datamart, Inteligencia de Negocios, ETL.

ABSTRACT

The objective of the research was to determine the influence of the implementation of a datamart on productivity in SEDAPAL's EEC-AR.

Having as main problem the lack of the systems area, therefore technical support and the importance of the management of information technology (IT) resources in the company, which prevent dedicating quality services to the users of the company, since they are not aligned to IT.

The research was applied, explanatory level, experimental design (pre-experimental), considering a population of 13 records. The recording was used as a technique and as an instrument the registration form for the indicators, the statistical analysis was done using the SPSS 21.0 program, which yielded reliable results in the investigation. Finally, it was determined that a datamart in the EEC-AR improves the efficiency index by 36.46% and maintains the efficiency index at 100%, thus generating an improvement in productivity.

Keywords: Datamart, Business Intelligence, ETL.

I. INTRODUCCIÓN

En la actualidad, las distintas empresas están en una necesidad de almacenamiento de datos, ya sean provenientes de la relación con sus proveedores, clientes, procesos internos, etc.; siendo idónea, la instauración de un repositorio de datos para tener los datos agrupados y en un solo lugar (Budi 2019), es de esta manera que las bases de datos tienen un fin de acumulamiento y consulta, que al interactuar con el usuario estos datos pasan a ser información. El Perú no es la excepción, puesto que las empresas cuentan con la necesidad mencionada, en apoyo, Khalaf (2020) comenta que los datos estructurados, en consecuencia la información, se ha vuelto un elemento de gran valor para las organizaciones, pues es usada en sus operaciones del día a día, como a la vez puede ser usada para un tema estratégico, pudiendo contribuir a la inteligencia empresarial, es así que Jurado y Miguelena (2019) mencionan que la información puede resultar en ventajas competitivas si se analiza correctamente, de esta manera es que las empresas pasan de la información al conocimiento.

Siendo así, SEDAPAL es una empresa estatal peruana que brinda servicios de agua potable que va acompañado de un servicio de saneamiento; el agua que suministra a sus usuarios se trata en La Atarjea, con la cual se abastecen 46 de 49 distritos existentes entre Lima y Callao; dentro de la organización, el Equipo de Evaluación de Calidad de Aguas Residuales (EEC-AR), área perteneciente a SEDAPAL, tiene como función el control de las aguas residuales mediante la metodología y procesos de control determinadas por la normativa VMA, permitiendo el reúso de las mismas para la conservación y el cuidado del ambiente.

Es así que en el Equipo de Evaluación de Calidad de Aguas Residuales, se detectó que, de acuerdo a la entrevista (Anexo 7), el EEC-AR cuenta con una productividad de calidad deficiente puesto que no se realiza la elaboración de reportes en los tiempos asignados (2 horas promedio), excediendo estos en un 35% del tiempo asignado, es de este modo que la productividad se ve afectada, además se cuenta con una mala calidad de datos, haciendo que el EE-CAR tenga que retrasar sus otras labores por priorizar la realización de los reportes puesto que los datos no están organizados.

Es entonces, que la causa de la situación actual resulta de venir trabajando con una base de datos que no está apta para manipular los datos de manera óptima, pues el esquema relacional no lo permite, la cual al comienzo de las operaciones del EECAR ciertamente resultó útil teniendo como fin el almacenamiento y consulta de información como menciona Casas et al. (2016); sin embargo, detalla Castillo, Medina y Fariña (2018) esto genera que los datos se acumulen a tal punto que cuando se necesita analizar los datos estructurados para posibles estrategias de fortaleza o reducción de debilidades en base al análisis, estos datos se encuentran desorganizados por lo que operar con ellos sería riesgoso.

De seguir trabajando con la base datos actual es posible la toma de decisiones incorrecta, en apoyo Braa, Heywood y Sahay (2012), afirman que cuando los volúmenes de datos van en aumento, y estos no se analizan mediante indicadores sino en bruto, se cuenta con una mala calidad de datos, resultando en una toma de decisiones no acertada e ineficiente.

El problema existente en mención se puede tratar mediante 3 alternativas de solución: la primera solución podría ser una reestructuración del área en mención en donde se tengan bien definidos los roles y responsables de las distintas actividades, la segunda podría ser la creación de un sistema de registro automático para la entrada de datos y finalmente la implementación de un Datamart, por temas de viabilidad y por necesidad del área se eligió el último, es así que se propone el desarrollo de un Datamart para el EEC-AR, en el que, como detalla Tolee (2015), se podrá extraer la información de distintas fuentes y tenerlas en una sola base de datos en un menor tiempo en comparación del sistema actual; permitiendo al área realizar sus funciones con normalidad, ahorrando recursos, y dando un mayor tiempo a realizar sus otras actividades, permitiendo a la vez tener acceso a la información a los usuarios correspondientes, pues como detalla Wenda (2016) los componentes de data warehouse proporcionan información correspondiente respecto a un total, es decir se enfoca en un área, siendo así el datamart permitirá manipular la información, y mediante el análisis definir estrategias en base a la información visualizada. Finalmente en apoyo Díaz (2018) menciona que la mejora en la productividad resulta

en una mejora en la definición de estrategias en base a la información visualizada, toma de decisiones, puesto que los tiempos eficientes que se logran mediante la instauración de soluciones de inteligencia de negocio permiten un mayor margen al análisis de datos.

Sobre la base de realidad problemática presentada se planteó el problema general y los problemas específicos de la investigación. El problema general de la investigación es: ¿Cuál será el efecto de la implementación de un datamart en la productividad en el EEC-AR de SEDAPAL? Los problemas específicos de la investigación son los siguientes:

- PE1: ¿Cuál será el efecto de la implementación de un datamart en el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL?
- PE2: ¿Cuál será el efecto de la implementación de un datamart en el índice de eficacia en la productividad en el EEC-AR de SEDAPAL?

El objetivo general es determinar el efecto de la implementación de un datamart para la productividad en el EEC-AR de SEDAPAL. Los objetivos específicos son los siguientes:

- OE1: Determinar el efecto de la implementación de un datamart en el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL.
- OE2: Determinar el efecto de la implementación de un datamart en el índice de eficacia en la productividad en el EEC-AR de SEDAPAL.

La hipótesis general es que la implementación de un datamart mejora la productividad en el EEC-AR de SEDAPAL. Las hipótesis específicas son las siguientes:

- HE1: La implementación de un datamart mejora el índice de eficiencia de la productividad en el EEC-AR de SEDAPAL.
- HE2: La implementación de un datamart mejora el índice de eficacia de la productividad en el EEC-AR de SEDAPAL.

Se justifica de manera operativa, puesto que tras la implementación del Datamart en el EECAR, como detalla Zhu (2017), los usuarios podrán realizar consultas de manera rápida y eficiente pues la información estará agrupada en un solo repositorio, encontrándose disponible y siendo fiable, ahorrando y optimizando el tiempo en la realización de reportes y por consecuencia una mejor toma de decisiones. En apoyo a lo mencionado, Celis (2016) en su tesis tras implementar un datamart en el área de ventas, permitió agilizar los procesos que son realizados manualmente, puesto que genera las solicitudes de reportes con alta eficacia, finalmente Valle (2019) menciona que el datamart permite extraer y analizar datos de manera eficaz, permitiendo aprovechar estos datos en la toma de decisiones.

Se justifica de manera tecnológica, puesto que la iniciación de un proyecto de business intelligence y al mismo tiempo seguir una metodología para la creación del mismo para conseguir un data warehouse o datamart en un área de una organización resulta en una ventaja mayor a la que es la existencia de una base de datos, otorgando beneficios como el rápido acceso a los datos y mayor velocidad de consulta, que en consecuencia permitirá mejorar la definición de estrategias en base a la información visualizada de la organización. En apoyo a lo mencionado, Sousa, Yamanari y Guerrini (2020), detallan la importancia de la migración de una base de datos a un datamart, diciendo que las empresas que buscan posicionarse en el mercado, están en intensa competencia y al mismo tiempo con la necesidad de mejorar su desempeño en cuanto a su productividad, siempre con la intención de un proceso de mejora continua, es decir realizar cambios constantes, procesos apoyados en el datamart, en apoyo Wickramasinghe et al. (2020) menciona que estos cambios deben realizarse mediante el análisis de información y no por intuición, por consecuencia se recomienda tratar la información con herramientas de Business Intelligence (BI) para que los datos que se usan en el análisis se consideren fiables.

Se justifica de manera económica, puesto que no solo se aprovecha de mejor manera el tiempo en el acceso a la información que está asociado con el pago por tiempo a los trabajadores, sino que también permite a los tomadores de decisiones la posibilidad de descubrir tendencias que pueden resultar en nuevas oportunidades de negocio, lo

cual significaría un mayor ingreso económico. En apoyo a lo mencionado, Caballero (2019), menciona que al tener implementado un datamart, se reduce la saturación de los usuarios encargados del análisis de la información, puesto que antes descuidaban sus actividades para atender las solicitudes de información, lo cual se traduce en un ahorro ya que el trabajador resulta ser más productivo dotándolo de un mayor tiempo para la realización de otras labores determinadas por la empresa.

II. MARCO TEÓRICO

En este segundo capítulo se presentan los trabajos anteriormente realizados por autores que abordaron la misma temática, temática similar o que pueden contribuir a esta investigación, los cuales permitieron orientar la investigación y representan casos de éxito en la implementación de un datamart, lo que refuerza el éxito de culminar la investigación presente.

A continuación, se presentan los antecedentes revisados que abordan temas de Inteligencia de Negocios o Datamart; en dichas investigaciones se aprecia una mejora al implementar algún tipo de solución mencionada, permitiendo realizar un trabajo con la misma temática pues se evidencian casos de éxito y mejoras en el contexto donde se realizó cada una de las investigaciones.

Montes (2018) en su tesis, tuvo como objetivo instaurar una solución de inteligencia de negocios, específicamente un datamart, para la optimización de los resultados en la toma de decisiones. Esta investigación resultó ser de tipo experimental - pre experimental. La población y muestra fue de 10 personas. Finalmente, el autor determinó como conclusiones que el instaurar un sistema de BI mejora la toma de decisiones, pues presenta los datos en base a las necesidades del usuario. Asimismo, Montes (2018) recomienda el uso de los sistemas de BI a las autoridades de las organizaciones o en todo caso a quien corresponda tomar las decisiones para el manejo de los datos pues permite una visualización de lo que se requiere.

Zegarra (2017) en su tesis, tuvo como objetivo llevar a cabo la instauración de un datamart para la mejora de los procesos necesarios para llevar a cabo la toma de decisiones, esta solución de inteligencia de negocios estaría realizada en base a la metodología Kimball, una de las tantas metodologías de BI. Esta investigación resultó ser de tipo experimental - pre experimental. La población fue igual a la muestra siendo 8 personas. Finalmente, el autor determinó como conclusiones que el instaurar un sistema de BI mejora la toma de decisiones, pues los datos resultan confiables puesto que se realizaron los procesos necesarios para garantizar que se está trabajando con datos que no son incorrectos, repetidos o duplicados, es decir se garantiza una toma de decisiones de calidad. Asimismo (Zegarra), recomienda la implementación de un datamart no solo en el área de logística o área que estén asociadas a esta área, sino en cualquier área que requiera analizar información, pudiendo muy a futuro converger todas estas soluciones en un data warehouse.

Duran (2017) en su tesis, tuvo como objetivo la instauración de una solución de business intelligence (datamart) en un área educativa y el mejoramiento en base a esa implementación de la toma de decisiones. Dicha investigación fue de tipo aplicada, el tipo de investigación fue explicativa y el diseño fue experimental. La población fue de 7 personas y la muestra fue de 2. Se concluyó que la instauración del datamart resultó favorable a el proceso de toma de decisiones en el área educativa pues según los resultados un 66.7% y un 33,3% de los encargados del análisis de la información y usuarios del datamart calificaron en el nivel muy bueno y bueno respectivamente. Asimismo, Duran (2017) recomendó motivar el uso de la inteligencia de negocios en nuevos indicadores como el seguimiento académico para garantizar la educación de calidad en base decisiones y monitorear los resultados de los estudiantes.

Rodríguez (2016) en su tesis, tuvo como objetivo averiguar que tanto es el impacto de un datamart en un área contable como lo es la Gerencia de Ventas de la empresa PERÚ PIMA S.A. Dicha investigación fue de tipo aplicada y el diseño fue experimental de tipo pre-experimental. La población y muestra fue de 9 reportes, se considera de esta forma pues los elementos a estudiar son pequeños. Se concluyó que la

implementación de un datamart en el área de ventas mejoró el proceso de toma de decisiones en la empresa mejorando el índice de eficacia de 66.29% a 90,50%, resultando en base a este porcentaje en un mayor logro de objetivos que fueron gracias a la implementación del datamart. Asimismo, Rodríguez (2016) recomendó en un mediano plazo implementar un data warehouse en toda la organización con el fin de que cada área pueda estar abastecida de la información necesaria para la toma de decisiones.

Caballero (2019) en su tesis, tuvo como objetivo determinar la influencia de un Datamart en la mejora de toma de decisiones en la gestión de proyecto. Dicha investigación fue descriptiva explicativa y el diseño fue pre-experimental. La población y la muestra fue de 30 personas. Finalmente, el autor determinó como conclusiones que el instaurar un sistema de BI mejora la toma de decisiones em cualquier área como lo es un área encargada del seguimiento de proyectos en una entidad estatal, pues los datamart están contruidos en base a requerimientos específicos de los propios futuros usuarios logrando aumentar el puntaje de 34,43 hasta 91,73 en la toma de decisiones. Asimismo, Caballero (2019) recomendó promover el uso de las herramientas BI en no solo a nivel escritorio como se realizó en esta investigación sino también en dispositivos móviles como los celulares, pues estos son más prácticos.

A continuación, con el propósito de tener un conocimiento más amplio sobre el presente trabajo se presentan los temas relacionados a la temática de la investigación:

Celis (2016) define el datamart como un repositorio de datos enfocado en un área del negocio en particular, teniendo como objetivo proporcionar indicadores o KPI para el fácil manipulamiento y análisis de la información, comentan también Knrmeta, Jovanic y Marjanovic (2016), que el datamart es componente de lo que se conoce como data warehouse, el cual cumple las cualidades de un datamart, pero este abarca toda la organización; por lo que la diferencia entre estos términos se reduce en alcance. En adición, Caballero (2019) menciona que la implementación de un datamart otorga beneficios en cuanto a consulta de la información, en el área en donde se implemente

puesto que se reduce el volumen de información a recorrer ya que los datos se encuentran estructurados.

De acuerdo con Celis (2016), la toma de decisiones debe ser un proceso diligente en todo tipo de organización, puesto que es el proceso en el cual se identifican problemas u oportunidades, para en base a la información analizada elegir un curso de acción; Montero et al. (2015), indica que este curso puede favorecer a la organización y su desarrollo; siendo así, la toma de decisiones resulta en la elección del camino a tomar para el aprovechamiento de las oportunidades o la mitigación de los problemas detectados. En adición, Fernández et al. (2015) menciona que el proceso de toma de decisiones requerirá distintos elementos para actuar de forma lógica, siendo uno de estos los datos. Sobre los datos, Caballero (2019) menciona que los datos usados en la toma de decisiones deben cumplir las características fiabilidad y disponibilidad para realizarse de manera correcta.

Sobre la fiabilidad de los datos respecto de la toma de decisiones, Caballero (2019) detalla que es el aspecto por el cual se comprueba si los datos cumplen las especificaciones para el posterior análisis en alguna herramienta de Business

Intelligence. Respecto a lo mencionado se puede decir que los datos son fiables cuando estos cumplen su función, al conocer que estos datos serán la entrada para el análisis de la información y posteriormente ser usados en la toma de decisiones, se recomienda que estos no pueden estar desactualizados, duplicados o incompletos, pues la toma de decisiones con datos que cumplan con las características mencionadas resultará en un análisis poco certero o erróneo.

Sobre la disponibilidad de los datos respecto de la toma de decisiones, Caballero (2019) menciona que es la capacidad de garantizar que se podrán acceder y visualizar los datos cuando lo disponga el usuario, a la vez Cubillo (1997) apoya la idea mencionando que tener los datos a tiempo y actualizados sirve para apoyar la toma de decisiones. Se entiende entonces que la correcta disponibilidad de los datos resulta en no truncar los procesos de los usuarios, permitiendo el acceso de los datos cuando este lo requiera, evitando que esto pueda impactar negativamente en la productividad

del personal, en adición Hermawan (2018) menciona que esta característica resulta importante en la era actual puesto que el proceso de toma de decisiones rápido y eficiente resultará en una ventaja competitiva o, según Bermeo et al. (2020), permitir equilibrar la balanza competitiva.

Es de este modo que la toma de decisiones, resulta ser un proceso en el que se analiza, en base a la información, lo que le conviene a la organización, existiendo factores que dotan de objetividad a la decisión tomada y cuan acertada puede llegar a ser, siendo estos: la información fiable, la cual ayudará a evitar errores en el análisis y la información disponible, que presentará los datos cuando el usuario necesite contar con ellos. En apoyo, Castillo, Medina y Fariña (2018) menciona que mediante un datamart la toma de decisiones resulta ser un proceso diligente pues se trabajan con datos objetivos y no subjetivos, datos que al estar preparados con anterioridad tienen un impacto positivo en la productividad ya que se cumplen las metas y se optimiza el tiempo, a la vez (Fontalvo, De La Hoz y Morelos, 2018, p.12) mencionan que la productividad resulta muy importante al gestionar una organización puesto que mediante su análisis se conocen los recursos consumidos para llegar a cumplir las metas empresariales.

Siendo así, (Gutiérrez y de la Vara, 2013, p.7) mencionan que la productividad se puede entender como una relación donde se considera lo producido y los recursos que se emplearon, y de que querer mejorar la productividad se deben maximizar los resultados y reducir los tiempos en la elaboración de esos mismos resultados, dimensionando a la vez la productividad en dos elementos los cuales son: la eficiencia, existiendo un enfoque en recursos y la eficacia, con un enfoque en alcanzar las metas.

Con respecto a la eficiencia, manifiesta (Gutiérrez y de la Vara, 2013, p.7) que es la concordancia de los resultados que se logran y los recursos que se deben utilizar, y de querer mejorar la eficiencia debe existir un enfoque en la reducción de tiempos, añade a la vez que los procesos productivos no solo se deben enfocar en el 100% del

cumplimientos de estos, sino también se debe considerar la calidad y/o los recursos que se utilizan, lo cual justamente mide este indicador.

De esta manera, se presenta la fórmula de eficiencia.

$$\text{Eficiencia} = \frac{\text{Reportes Generados} / \text{Horas Trabajadas}}{\text{Reportes Objetivos} / \text{Horas Previstas}}$$

Con respecto a la eficacia, manifiesta (Gutiérrez y de la Vara, 2013, p.7) la eficacia es el nivel en donde los requerimientos establecidos por el área o la empresa realizados son acorde a los resultados estimados, existiendo una prioridad en la realización de metas u objetivos, de querer mejorar el indicador eficacia se debe mejorar los resultados y en general los procesos.

En apoyo, Rodríguez (2016) menciona que el nivel de eficacia se evalúa la comparativa entre lo alcanzado y lo esperado, con el fin de aumentar la posibilidad de éxito y realizar la respectiva toma de decisiones.

De esta manera, se presenta la fórmula de eficacia.

$$\text{Eficacia} = \frac{\text{Reportes Generados}}{\text{Reportes Objetivos}}$$

Finalmente, sobre la elección de la metodología de desarrollo de un datamart, Garcés (2015) realizó un estudio de tipo descriptivo-comparativo en el que determina las mejores metodologías en proyectos de inteligencia de negocio, pudiendo ser este un datamart o data warehouse, se analizan en 10 fases, cada una con sus respectivas tareas, teniendo como resultado que las tres (3) mejores metodologías son: Hefesto, Kimball e Inmon respectivamente en cuanto a metodologías tradicionales, mencionando a la vez que la metodología Hefesto resulta eficiente en proyectos medianos y pequeños, recordando que un proyecto mediano o pequeño de business intelligence es un datamart y no un data warehouse, puesto que la metodología no se enfoca en la documentación.

La metodología mencionada, considera 4 fases: análisis de requerimientos, esta fase resulta esencial pues como menciona Miglani, Kumar y Pahwa (2016) el datamart debe brindar datos en base a la necesidades del usuario y que sean de su interés; análisis de OLTP, modelo lógico del DW e integración de los datos, Díaz et al. (2015) resalta la importancia de esta última fase mencionando que el proceso ETL, tarea incluida en la cuarta fase de la metodología, es el más importante pues este módulo será el encargado de mantener el repositorio actualizado y garantizará que los datos estén y sean aptos para el análisis de la información que estará a cargo de los tomadores de decisiones o en todo caso quien es el usuario del datamart o data warehouse.

Es así que se opta por la metodología Hefesto para la creación del datamart, finalmente se dispondrá la información en dashboards, puesto que Valdés, Ramos y Gómez (2013) mencionan que las vistas del análisis realizado y reportes generados a partir del datamart ayudan al análisis en la toma de decisiones, apoyando la idea Florez et al. (2019) menciona que la aplicación de técnicas de inteligencia empresarial como cuadros de mando resulta útil para mejorar los procedimientos del negocio.

III. METODOLOGÍA

3.1 Tipo y diseño de investigación

El investigación resulta ser de tipo aplicada puesto que, detallan Hernández, Fernández y Baptista (2014), este tipo de investigación se enfoca en la resolución de problemáticas reales que agobian a la organización mediante la utilización del conocimiento y teorías que poseen los encargados de resolver dicha problemática, pasando de la teoría a la práctica; siendo así que con la aplicación del conocimiento en nuestra realidad es que la humanidad se viene desarrollando y evolucionando, a la vez detalla que el diseño es experimental – pre experimental puesto que no se manipulan todas las variables, siendo el grado de control mínimo, en esta investigación se mide el impacto de la variable dependiente (datamart) sobre la variable que en esta investigación no se manipula, variable independiente (productividad).

Sobre el tipo experimental que corresponde al diseño de esta investigación, resulta ser experimental pues tal como lo detallan Hernández, Fernández y Baptista (2014), es la

investigación en la que existe manipulación de variables independientes, para analizar los efectos en la variable dependiente al manipularla, existiendo así una relación entre estas dando pase a un experimento.

Sobre el tipo pre experimental, detallan Hernández, Fernández y Baptista (2014), que existe esta denominación puesto que no existe control en la variable independiente a comparación de los experimentos puros en donde se aplican niveles ya sea de ausencia o presencia. A continuación, se presenta el diseño preprueba/posprueba en un grupo único en donde se aplica un estímulo:

G O1 X O2

G: Grupo experimental: Compuesto por la población.

O1: Productividad en el Equipo de Evaluación de Calidad de Aguas Residuales (EEC-AR) de SEDAPAL – Pre Aplicación del Datamart

X: Aplicación del datamart

O2: Productividad en el Equipo de Evaluación de Calidad de Aguas Residuales (EEC-AR) de SEDAPAL – Post Aplicación del Datamart

Es entonces que la presente investigación sería de tipo aplicada de diseño experimental de tipo pre experimental.

3.2 Variables y operacionalización

Definición Conceptual:

VI: Celis (2016) define el Datamart como un repositorio de datos enfocado en un área del negocio en particular, teniendo como objetivo proporcionar indicadores o KPI para el fácil manipulamiento y análisis de la información.

VD: (Gutiérrez y de la Vara, 2013, p.7) define como la relación entre lo producido y los medios empleados a la productividad.

Definición Operacional:

VI: Datamart: Celis (2016) menciona que el Datamart permite almacenar los datos de la organización por departamento disponiendo una estructura eficiente para analizar la información apoyando de esta forma a la toma de decisiones.

VD: Productividad: En base a lo propuesto por Gutiérrez y de la Vara (2013), para medir la variable productividad se dimensiona en eficiencia teniendo como indicador el índice de eficiencia y en eficacia mediante el indicador índice de eficacia.

Tabla 1. Operacionalización de la variable dependiente

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIÓN	INDICADOR	FÓRMULA	ESCALA DE MEDICIÓN
PRODUCTIVIDAD	(Gutiérrez y de la Vara, 2013, p.7) define la Productividad como la relación entre lo producido y los medios empleados.	En base a lo propuesto por Gutiérrez y de la Vara (2013), para medir la variable productividad se dimensiona en eficiencia teniendo como indicador el índice de eficiencia y en eficacia mediante el indicador índice de eficacia y se medirá usando una ficha de registro.	Eficiencia	Índice de Eficiencia	$\text{Eficiencia} = \frac{\text{Reportes Generados/Horas Trabajadas}}{\text{Reportes Objetivos/Horas Previstas}}$ <p>Gutiérrez y de la Vara (2013)</p>	De razón
			Eficacia	Índice de Eficacia	$\text{Eficacia} = \frac{\text{Reportes Generados}}{\text{Reportes Objetivos}}$ <p>Gutiérrez y de la Vara (2013)</p>	De razón

Fuente: elaboración propia

3.3 Población, muestra y muestreo

Población: Hernández, Fernández y Baptista (2014) mencionan que la población son los elementos que cumplen distintos requisitos determinados por el investigador para satisfacer correctamente los datos que se vayan a recolectar de este grupo, es así que estos son los elementos o personas a los cuales se les estudiará. En el caso de esta investigación la población está dada por los 13 reportes, estos son reportes necesarios

para la toma de decisiones en el área y los solicitados por la jefatura en el tiempo determinado por el estudio, el tiempo considerado es un mes para el pre test, de la misma forma el post test es considerada un mes.

Sobre la muestra, Hernández, Fernández y Baptista (2014) mencionan que si los elementos estudiados lo que se entiende por población no excede de (50) elementos, la población no debería variar por lo que la muestra debería tener el mismo valor de la población. La población al no exceder el mínimo detallado por el autor (50) para definir un subgrupo que sería la muestra es entonces que se toma el total de los elementos para la muestra, es entonces que la muestra son 13 reportes.

Sobre el muestreo, no se utiliza algún muestreo en esta investigación puesto que no se existe muestra ya que no se subdividió y se trabaja con la población entera o completa, se realiza así puesto que las características de la investigación lo requieren, a la vez los resultados de la población determinada es la finalidad de la investigación. Es entonces que en este caso el muestreo no aplica.

3.4 Técnicas e instrumentos de recolección de datos

Con respecto a cómo se recolectarán los datos, considerando la recolección desde el pre test hasta el procesamiento de estos datos recolectados incluidos el post test Hernández, Fernández y Baptista (2014) mencionan que para recolectar datos es necesario elaborar un plan detallado que nos conduzca a reunir los datos, y que estos datos sirvan para concluir la investigación correctamente y no se trabajen con datos que no tienen propósito en la investigación, a la vez mencionan que un instrumento de recolección de datos debe ser confiable y valido (validez de contenido) para que los resultados de la investigación puedan tomarse en serio. Es así, que para la recolección de los datos de la presente investigación se recurrió a la técnica de fichaje.

Tabla 2. *Técnica e instrumento de recolección de datos*

Técnica	Instrumento	Fuente	Participante(s)
Fichaje	Ficha de Registro	EEC-AR	Jefe del EEC-AR

Fuente: elaboración propia

Es así que el instrumento, ficha de registro, que se utiliza es el de Lozada y Morales (2020) el cual fue revisado y validado por expertos:

Tabla 3. Validez a través de juicio de expertos

N.º	Expertos	Grado Académico
1	Ávila López, Bernardo Patricio	Magister
2	Granda Cotrina, Judith	Magister
3	Pérez Rojas, Even Deyser	Magister

Fuente: elaboración propia

3.5 Procedimientos

Para validar las hipótesis de esta investigación se recolectarán los datos para el pretest y postest de las fichas de registro las cuales están dirigidas a la población respectiva, participantes del Equipo de Evaluación de Calidad de Aguas Residuales (EEC-AR), los datos que se planean recolectar serán rellenos durante el proceso de creación de reportes por los mismos usuarios que realizan dicha labor, una vez recolectado los datos se analizarán los datos en el programa SPSS 21.0, contrastando así la información con la hipótesis de los resultados en el pre-test y post-test para verificar si cumple o no con la hipótesis, previo a eso se determinará si los datos presentan una distribución normal o no normal y finalmente presentar los datos en manera gráfica o de tabla para la facilidad de visualización.

3.6 Método de análisis de datos

La manera de analizar los datos para una investigación detallan Hernández, Fernández y Baptista (2014) consiste en procesar los datos obtenidos de la población para poder explicar y describir el fenómeno del estudio, dicho proceso se lleva a cabo por computadora con programas como el SPSS, Minitab, STATS u otro equivalente y busca generar resultados, en base a la información resultante los cuales se analizarán en base a los objetivos e hipótesis de la investigación, llegando así a una conclusión. Siendo así que la información generada se pasará a evaluar en el programa SPSS

mediante estadística para determinar que un Datamart mejorará o no la toma de decisiones en el Equipo de Evaluación de Calidad de Aguas Residuales de SEDAPAL, determinando así si se acepta la hipótesis alterna o no.

3.7 Aspectos éticos

Esta investigación y el investigador se comprometen en guardar confidencialidad de información personal que maneja la empresa como la que los colaboradores también especifiquen, específicamente los usuarios de la empresa puesto que resulta sensible la exposición al público de esta información, de la misma manera la identidad de los encuestados se mantendrá en privado. En el proceso de creación del datamart y reuniones que se realizan periódicamente con el área también se comenta la información que se esta mostrando en esta investigación, con el fin de no afectar a la empresa, al área y a sus clientes.

IV. RESULTADOS

4.1 Análisis Descriptivo

Indicador: Índice de eficiencia

Tabla 4. Estadísticos descriptivos del Índice de eficiencia en el Pre y Post de la implementación del Datamart

Estadísticos Descriptivos					
	N	Min	Max	Media	Desvi
Índice_eficiencia_Pre	13	.3261	.7709	.4897	.1312
Índice_eficiencia_Post	13	.6383	.9677	.8543	.1003
N	13				

Fuente: elaboración propia

Se verificó que la media en el pre test (Índice_eficiencia_Pre) fue 48,97% y en el post test (Índice_eficiencia_Post) 85,43% esto señaló a un aumento de un 36,46%, después de haber implementado el DataMart. Además, el indicador de índice de eficiencia la valorización mínima en el pre test fue 32,61% y en el post test fue 62,83%

respectivamente, como valorización máxima el pre test fue 77,09% y en el post test fue 96,77%. Para la desviación típica para el pre test fue 13,12% y en el post test se calculó que fue 10,03%, de manera respectiva.

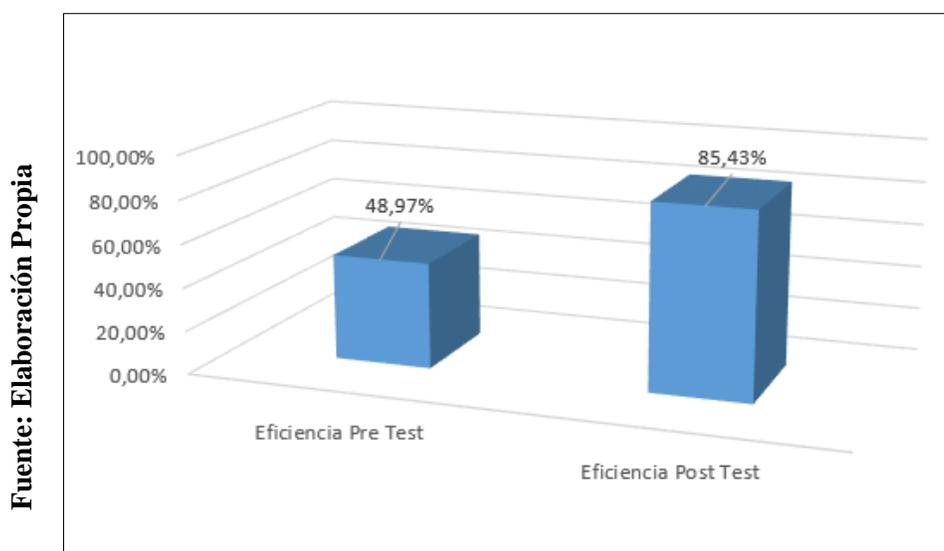


Figura 1: Índice de eficiencia pre y post de la implementación del datamart

Indicador: Índice de eficacia

Tabla 5: Estadísticos descriptivos del Índice de eficacia en el Pre y Post de la implementación del Datamart

Estadísticos Descriptivos					
	N	Min	Max	Media	Desvi
Índice_eficacia_Pre	13	Las variables son constantes y se ha desestimado.			
Índice_eficacia_Post	13				
N	13				

Fuente: Elaboración propia

Se verificó que el indicador índice de eficacia en el pre test (Índice_eficacia_Pre) y en el post test (Índice_eficacia_Post) son constantes por lo que el valor de la media, el valor mínimo, el valor máximo y el valor de la desviación típica se ha desestimado.

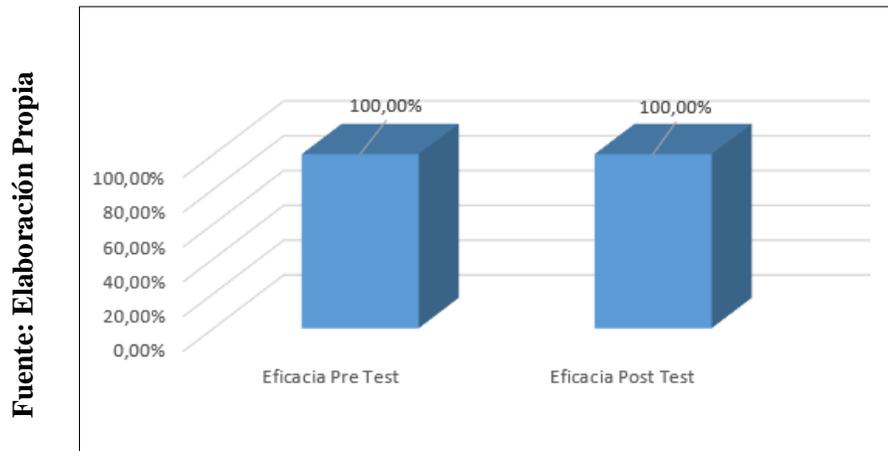


Figura 2: Índice de eficiencia pre y post de la implementación del datamart

4.2 Análisis inferencial

Se procedió con la comprobación de normalidad para el índice de eficiencia y índice de eficacia usando el método de Shapiro-Wilk ya que la población de los indicadores es menor a 50, alojando los datos en el software SPSS 21.0 con una confiabilidad de 95%.

Donde:

$\alpha > 0.05$ Normal -> En este caso los datos presentan una distribución normal o paramétrica

$\alpha < 0.05$ No Normal -> En este caso los datos presentan una distribución anormal o no paramétrica

Indicador: Índice de eficiencia

La finalidad del proceso es seleccionar la prueba de hipótesis y así corroborar si las informaciones de índice de eficiencia expresaban una distribución paramétrica en base al valor del Sig.

Tabla 6: Prueba de normalidad del Pre-Test y Post-Test del indicador Índice de Eficiencia

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Índice_eficiencia_Pre	,829	13	,015
Índice_eficiencia_Post	,904	13	,150

Fuente: Elaboración propia

Tanto se pudo verificar la validez de Sig. del Pre-Test del índice de eficiencia en la productividad fue 0,015, cuyo valor es mayor a 0,05. De la misma forma en el post test del índice de eficiencia en la productividad fue 0,150, esto indica que es mayor a 0,05. Estos resultados dan la confirmación que en ambas pruebas tanto del pre y post test para el indicador índice de eficiencia tienen una distribución normal o también llamado Paramétrico tal como se puede estimar en las figuras 3 y 4.

Fuente: Elaboración Propia

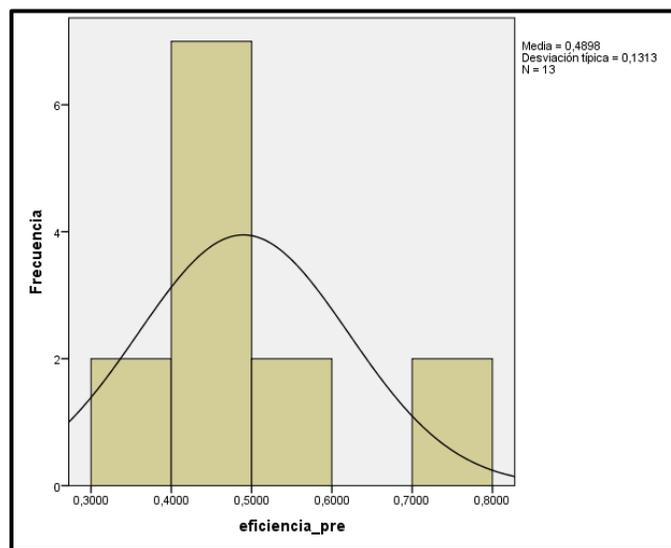


Figura 3: Prueba de normalidad del Pre-Test Índice de Eficiencia

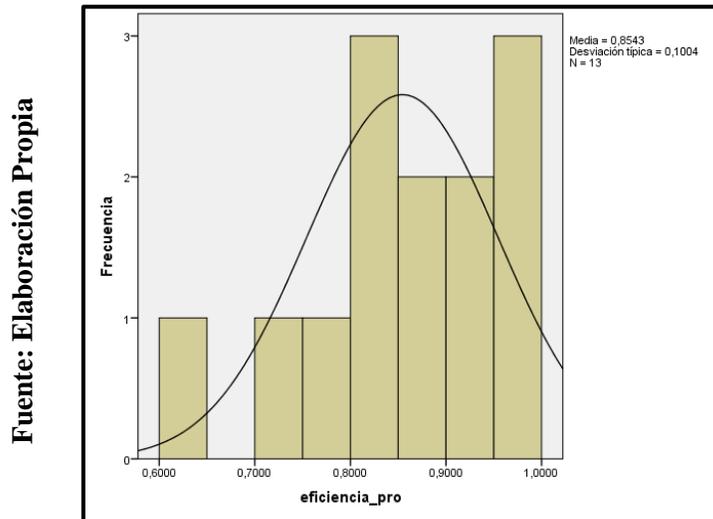


Figura 4: Prueba de normalidad del Post-Test Índice de Eficiencia

Indicador: Índice de Eficacia

La finalidad del proceso es seleccionar la prueba de hipótesis y así corroborar si las informaciones de índice de eficacia expresaban una distribución de datos paramétrica.

Tabla 7: Prueba de normalidad del Pre-Test del indicador Índice de eficacia

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Eficacia_Pre	Las variables son constantes y se han desestimado.		
Eficacia_Post			

Fuente: Elaboración propia

Tanto se pudo verificar la validez de Sig. del Pre-Test y Post-Test del índice de eficacia en la productividad se desestimó. Estos resultados dan la confirmación que estas variable son constantes tal como se puede visualizar en las figuras 5 y 6.

Fuente: Elaboración Propia

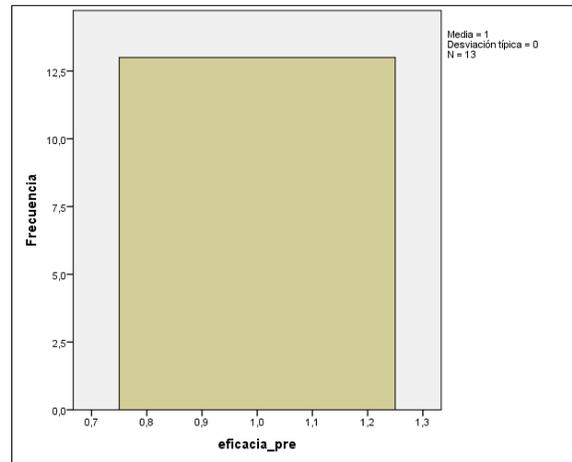


Figura 5: Prueba de normalidad del Pre-Test Índice de Eficacia

Fuente: Elaboración Propia

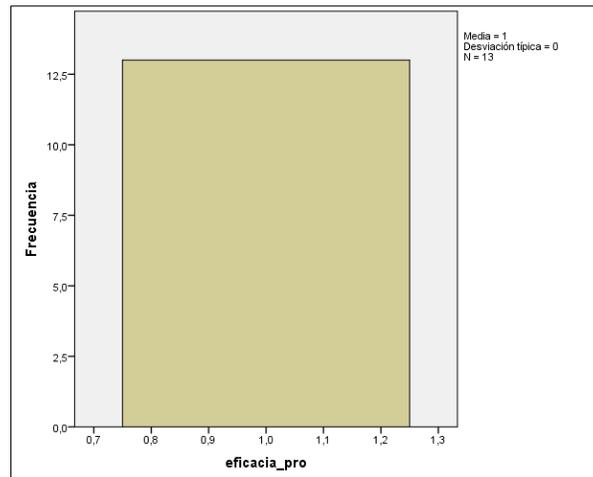


Figura 6: Prueba de normalidad del Post-Test Índice de Eficacia

Prueba de Hipótesis

Hipótesis de Investigación 1:

1. Planteamiento de hipótesis

H1: La implementación de un datamart mejora el índice de eficiencia de la productividad en el EEC-AR de SEDAPAL.

H0: La implementación de un datamart no mejora el índice de eficiencia de la productividad en el EEC-AR de SEDAPAL.

HA: La implementación de un datamart mejora el índice de eficiencia de la productividad en el EEC-AR de SEDAPAL.

2. Fijación de α

$\alpha > 0.05$ Normal -> Se reconoce la hipótesis nula (hipótesis de trabajo).

$\alpha < 0.05$ No Normal -> Se reconoce la hipótesis alterna (hipótesis del investigar).

3. Estadístico de prueba

Se aplicó la prueba de T Student porque la información obtenida durante la investigación del indicador Índice de eficiencia (Pre y Post Test) fue Paramétricos (tienen una distribución normal) y son muestras dependientes.

Tabla 8: Prueba de T Student para en índice de eficiencia de la productividad Pre y Post.

T Student	Eficacia_PRE
	Eficacia_POST
Sig. (bilateral)	,000

Fuente: Elaboración propia

4. Decisión estadística

Por otro lado, se observa en la prueba de hipótesis manifiestan que el Sig, del indicador Índice de eficiencia es 0,000, por consiguiente, se rechaza la hipótesis nula aceptando así la hipótesis alternativa con un 95% de confianza.

5. Conclusión

Se concluyó que, el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL mejoró a causa de la implementación del datamart en un 36.46%.

Hipótesis de Investigación 2:

1. Planteamiento de hipótesis

H2: La implementación de un datamart mejora el índice de eficacia de la productividad en el EEC-AR de SEDAPAL.

H0: La implementación de un datamart no mejora el índice de eficacia de la productividad en el EEC-AR de SEDAPAL.

HA: La implementación de un datamart mejora el índice de eficacia de la productividad en el EEC-AR de SEDAPAL.

1. Fijación de α

$\alpha > 0.05$ Normal -> Se reconoce la hipótesis nula (hipótesis de trabajo).

$\alpha < 0.05$ No Normal -> Se reconoce la hipótesis alterna (hipótesis del investigar).

2. Estadístico de prueba

No se puede calcular la correlación por trabajar con constantes, ni el valor T porque el error típico de la diferencia es 0 puesto que es una constante tanto en el pre como el post test.

3. Decisión estadística

No aplica porque los datos son constantes.

4. Conclusión

Se concluyó que, el índice de eficacia en la productividad en el EEC-AR de SEDAPAL se mantuvo al 100% por tanto el índice de eficiencia no mejoró a causa de la implementación del datamart.

V. DISCUSIÓN

En la recolección de datos del Pre Test del indicador índice de eficiencia realizado en el EEC-AR de SEDAPAL se alcanzó una media 48,97% y luego de la implementación del datamart se obtuvo una media de 85,43%. Los resultados que se obtuvieron indican un aumento de 36,46% para el índice de eficiencia, figurando ello como un resultado favorable. Al cotejar los resultados con los de Lozada y Morales en su tesis, posterior de la implementación del datamart, se halló una mejora en el índice de eficiencia del 6.92%, pasando de 88.12% a 95.04%, es así que se evidencia que ambos casos los resultados fueron favorables.

Los resultados mostrados sirven para afirmar que la implementación de un datamart mejora el índice de eficiencia en la productividad, si bien ambos resultados presentan variación entre ellos, ambos resultan en una mejora tras la implementación, demostrando así que al implementar un datamart mejora el índice de eficiencia el cual concuerda con lo planteado desde el inicio de la investigación que es que el datamart mejora el índice de eficiencia en el EEC-AR de SEDAPAL. Ello se sustenta en la apreciación de los autores Hurst, D. (et al) en la revista Journal of Personalized Medicine con la publicación detalla que el datamart resulta en que los datos sean más procesables y accesibles a comparación de los orígenes de datos, es así que permitirá acceder al usuario a los datos con eficiencia y transparencia, lo cual permitirá optimizar el tiempo en el análisis de datos.

Por otro lado, en la recolección del Pre Test del indicador índice de eficacia en el EEC-AR de SEDAPAL y luego de la implementación del datamart se obtuvo una media constante. Los resultados que se obtuvieron indican el cumplimiento de las acciones planificadas al 100%, para el índice de eficacia, figurando ello como un resultado favorable puesto que se cumplen las metas sin embargo este resultado no representa una mejora en el índice de eficacia. Al comparar los resultados con los de Rodríguez Briones en su tesis determinó que posterior de la implementación del datamart, se halló una mejora en el indicador índice de eficacia de 24.21% llegando a 90.50%, de similar manera se cotejaron los resultados con los de Lozada Huarcaya y Morales Panta en su tesis en la cual se obtuvo una mejora de 9.32% del indicador índice de eficacia tras implementar el datamart, obteniendo así ambas investigaciones resultados favorables en su investigación.

El resultado de la presente investigación si bien cumple el 100% en el indicador índice de eficacia no permite visualizar una mejora tras la implementación puesto que las variables son constantes en el pre y post test, esto se debe a la organización del área puesto que se enfocan en el logro de sus objetivos mas no se contemplan los recursos utilizados; en contraste con las otras investigaciones en donde sí se puede afirmar que

al implementar un datamart mejora el índice de eficacia en este caso el indicador solamente se mantiene, lo cual no concuerda con lo planteado “La implementación de un datamart mejora el índice de eficacia en la productividad en el EEC-AR de SEDAPAL”. De existir problemas en la eficacia el datamart debería poder incrementar el indicador mencionado puesto que Marazzi, Fabio (et al) sustenta en la revista *Journal of Personalized Medicine* en su publicación *Breast DataMart The Novel Breast Cancer Data Discovery System for Research and Monitoring: Preliminary Results and Future Perspectives* indica que la implementación de un datamart permitirá a los usuarios organizar la información de una manera eficaz ya que la solución de inteligencia de negocios mencionada centraliza la información relevante de los orígenes de datos y distintos archivos dispersados en un solo repositorio que está configurado para el cumplimiento de metas, puesto que el datamart se crea en base a los requerimientos del usuario o como menciona Belal (2017) también se basa en requisitos comerciales que se abarcan según el tamaño de la solución de inteligencia de negocios que se planea implementar, en cualquier caso las soluciones de inteligencia de negocios satisfacen necesidades particulares. A la vez Kangiama, Richar en su publicación, recalca que, este mini almacén de datos permite un la existencia de un repositorio en donde se recolectan los datos, este mini almacén es centralizado y al mismo tiempo universal del área donde se implemente, lo cual apoya a lograr las metas puesto que es una solución especializada para el departamento, área o función donde se encuentra el usuario.

Es así que en los resultados muestran que la implementación mejoró el índice de eficiencia pasando de una media menor al 50% a 85,43% lo cual significará un ahorro de recursos, en este caso tiempo, en la realización y actualización de reportes, permitiendo al usuario tener más tiempo en sus otras labores o incluso permitir el análisis de datos, estos beneficios nacen del método de construcción del datamart pues como menciona Guadaña en la *Revista de Investigación, Tecnología y Desarrollo* con su publicación concluyó que las herramientas business intelligence como el datamart permiten conocer los datos para transformarlos en un estándar el cual pueda manipularse y finalmente cargar estos convertidos en reportes o dashboards, todo este

proceso desde la fuente de datos, existiendo así un proceso automático que reduce los tiempos en la elaboración de reportes que impacta positivamente en el índice de eficiencia; sin embargo, el índice de eficacia se mantuvo al 100%, esta constante en el indicador mencionado sucede según la entrevista (Anexo 7) porque todas las tareas necesarias para la elaboración y actualización de reportes se vuelven prioritarias hasta concluir dicho elemento, siendo así el EEC-AR tiene conocimiento de su situación y por eso tiene días específicos para elaborar reportes específicos, este cronograma fue creado por el mismo área para lograr los objetivos en los días establecidos dando así una razón por la cual el índice de eficacia es constante y siempre se logra el 100%, este cronograma rige en cada uno de los meses y para poder cumplir las metas.

Finalmente, en base a la fórmula de productividad se evidencia que el índice de eficacia al ser constante, la productividad dependerá únicamente del índice de eficiencia el cual mejoró en 36.46% se entiende entonces que el valor de productividad se incrementó en la misma cantidad, esto concuerda con Delgado y Vichot en la revista Serie Científica de la Universidad de las Ciencias Informáticas con su publicación cuando detalló que la inteligencia de negocios para un área en específico con un datamart o en general en la empresa con un data warehouse permite a la organización que los usuarios puedan visualizar la información en base a indicadores, resultando esto en una posibilidad de mejorar el entorno mediante las estrategias que puedan concebirse mediante el análisis de la información, todo gracias al BI.

VI. CONCLUSIONES

Primera: Se concluyó que, el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL mejoró a causa de la implementación del datamart en un 36.46%. Por lo tanto, se confirma que la implementación de un datamart mejora el índice de eficiencia en la productividad.

Segunda: Se concluyó que, el índice de eficacia en la productividad en el EEC-AR de SEDAPAL se mantuvo constante a un nivel del 100%. Por lo tanto, no se confirma que la implementación de un datamart mejora el índice de eficacia en la productividad.

Tercera: Se concluyó que, la implementación de un datamart mejoró la productividad en el EEC-AR de SEDAPAL, pues permitió la mejora del índice de eficiencia y mantuvo el índice de eficacia, lo cual permitió alcanzar satisfactoriamente los objetivos planteados para esta investigación.

V. RECOMENDACIONES

Primera: Se recomienda, manejar un estándar en las fuentes de información con el fin de que todas las áreas manejen el mismo tipo de dato, y puedan compartir información sin inconvenientes.

Segunda: Se recomienda, extender las soluciones de inteligencia de negocios a otras áreas, en especial a las que realizan procesos repetitivos, puesto que se podría automatizar.

Tercera: Se recomienda, capacitar a los empleados en la herramienta Power BI a un nivel básico, para que estos puedan conocer las funcionalidades de mejor forma y explotar los datos.

REFERENCIAS

GENERATOR Breast DataMart The Novel Breast Cancer Data Discovery System for Research and Monitoring: Preliminary Results and Future Perspectives por MARAZZI, Fabio [et al]. Journal of Personalized Medicine [en línea] Enero. 2021. [fecha de consulta: 25 noviembre 2021].

Disponible en: <https://doaj.org/article/6253095dca904e1bb8657cd57de6f69b>

ISSN: 2075-4426

Improve HR Decision-Making Based On Data Mart and OLAP por Alaa Khalaf Hamoud [et al]. Journal of Physics: Conference Series [en línea]. 2020. [fecha de consulta: 26 de Setiembre de 2020]

Disponible en: <https://iopscience.iop.org/article/10.1088/1742-6596/1530/1/012058>

Enabling Value-Based Health Care with Business Analytics and Intelligence por Wickramasinghe, Nilmini [et al]. [en línea]. 2020. [fecha de consulta: 28 noviembre 2020]

Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=edssmt&AN=springer.978.3.030.17347.0.Chapter.17&lang=es&site=eds-live>

ISSN: 2191-5946 2191-5954

SOUSA, Thales, YAMANARI, Juliana y GUERRINI, Fabio. An enterprise model on Sensing, Smart, and Sustainable. SciELO [en línea]. Setiembre 2020. [fecha de consulta: 03 Octubre 2020].

Disponible en: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2020000300311&lang=en

ISSN: 1806-9649

Development of an electronic health records datamart to support clinical and population health research. HURST, Duke [et al]. Journal of Personalized Medicine [en línea] Junio. 2020. [fecha de consulta: 25 noviembre 2021].

Disponible en: <https://doi.org/10.1017/cts.2020.499>

KANGIAMA, Richard. Construction of a Datamart and knowledge extraction using the decision tree application to medical data. J. of Advancement in Medical and Life Sciences [en línea] Enero. 2020. [fecha de consulta: 25 noviembre 2021].

Disponible en: <https://doi.org/10.5281/zenodo.3780416>

Implementación de Data Mart; en Power BI; para el análisis de ventas a clientes; en los Ecomercios “Gransol” por Moyano Bermeo [et al]. [en línea]. 2020. [fecha de consulta: 29 de Setiembre de 2020]

Disponible en:

<http://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.944CB9DC&lang=es&site=eds-live>

CABALLERO Cabello, Raúl. Diseño e implementación de un DATAMART para la mejora de toma de decisiones en la gestión de proyectos de PROMPERÚ. Tesis (Licenciado en Ingeniería de Sistemas) Lima: Universidad Peruana de las Américas, 2019.

Development of an Advanced Analytics DataMart for Machine Learning por VALLE Carlos [et al]. Effectiveness Research, and Population Health Trends. 2019. [fecha de consulta: 04 octubre 2020].

Disponible en: <https://scholarlycommons.baptisthealth.net/cgi/viewcontent.cgi?article=4530&context=se-all-publications>

BUDI, Sudradjat. Analysis and Design of Data Mart Decision Support Systems at PT Marlindo Tirta Nusantara. [en línea]. 2019. [fecha de consulta: 26 Setiembre 2020].

Disponible en: <http://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.be398e01efa34ee3b883303c840a8949&lang=es&site=eds-live>

A Data Fusion System for Simulating Critical Scenarios and Decision-Making por Jimmy Florez Zuluaga [et al]. SciELO [en línea]. 2019. [fecha de consulta: 26 Setiembre 2020].

Disponible en: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0124-81702020000100089&lang=en

ISSN: 1909-7735

JURADO, Danitzel y Miguelena, Ramfis. Implementación de un Data Mart como una solución de Inteligencia de Negocios aplicando Metodología Ágil en base a la plataforma académica MOODLE de la Universidad Tecnológica de Panamá. [en línea]. 2019. [fecha de consulta: 29 Setiembre 2020].

Disponible en:

<http://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.38BE93F4&lang=es&site=eds-live>

CASTILLO, Wilson, MEDINA, Fernando y FARIÑA, Francisco. Data Mart para obtención de indicadores de productividad académica en una universidad. SciELO [en línea]. Noviembre 2018. [fecha de consulta: 26 Setiembre 2020].

Disponible en:

http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1646-98952018000100008&lng=es&tlng=en

ISSN: 07183305

CASTILLO, Wilson, MEDINA, Fernando y FARIÑA, Francisco. Una Metodología para Procesos Data WareHousing Basada en la Experiencia. SciELO [en línea]. Marzo 2018. [fecha de consulta: 26 Setiembre 2020].

Disponible en:

http://www.scielo.mec.pt/scielo.php?script=sci_arttext&pid=S1646-98952018000100008&lng=es&tlng=en

ISSN: 1646-9895

RODRÍGUEZ Briones, Eduardo. Datamart para la toma de decisiones en la gerencia de ventas de la empresa PERU PIMA S.A. Tesis (Licenciado en Ingeniería de Sistema) Lima, Perú: Universidad Cesar Vallejo, 2016.

Proceso de transformación de una base de datos relacional a modelo multidimensional haciendo uso de la metodología Hefesto para la elaboración de un Datamart por Edwin CASAS HUAMANTA [et al]. [en línea]. 2016. [fecha de consulta: 26 Setiembre 2020].

Disponible en: https://revistas.upeu.edu.pe/index.php/ri_ctd/article/view/632

MONTES Crespín, Cristina. Implementación y desarrollo de un Datamart para las consultas gerenciales del sistema terapéutico y grupales para el personal especializado del CDID. Tesis (Licenciado en Sistemas de Información) Guayaquil, Ecuador: Universidad de Guayaquil, 2018.

Modeling Data Mart Using ETL (Extract, Transform, Load) Webservice Concept On Feeder With A Dashboard por HERMAWAN, Aditiya [et al]. International Journal Of Engineering & Technology.2018. [fecha de consulta: 04 octubre 2020].

Disponible en:

https://www.researchgate.net/profile/Aditiya_Hermawan/publication/333746946_Modeling_data_mart_using_ETL_extract_transform_load_webservice_concept_on_feeder_with_a_dashboard/links/5d01ec98a6fdccd13096a7a1/Modeling-data-mart-using-ETL-extract-transform-load-webservice-concept-on-feeder-with-a-dashboard.pdf

FONTALVO, Tomás, De La Hoz, Efraín y MORELOS, José. La Productividad y sus factores: Incidencia en el Mejoramiento Organizacional. SciELO [en línea]. 2018. [fecha de consulta: 23 Octubre 2021].

Disponible en:

http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S1692-85632018000100047&lng=en&nrm=iso

ISSN: 1692-8563

DÍAZ, Luz. Relación entre toma de decisiones y productividad en las Fiscalías de Coronel Portillo, Ucayali 2018. Tesis (Master en Gestión Pública) Perú, Ucayali: Universidad César Vallejo, 2018

DURAN Colonia, Evelyn. Datamart como parte de una solución de inteligencia de negocios, para el soporte de la toma de decisiones de la gestión académica de la Facultad de Ciencias en la UNASAM. Tesis (Licenciado en Ingeniería de Sistemas). Huaraz, Perú: Universidad Nacional Santiago Antúnez de Mayolo, 2017.

ZEGARRA Mezarina, Alexander. Implementación de un datamart usando la metodología Ralph Kimball para el área de logística de la empresa comercializadora EASY WORK E.I.R.L. – Casma. Tesis (Licenciado en Ingeniería de Sistemas). Chimbote, Perú: Universidad Católica los Ángeles de Chimbote, 2017.

BELAL, Shbair. Design and Implementation of Data Warehouse using dynamic materialized views selection model: The Islamic University of Gaza as a case study. Tesis (Magister en ingeniería de computación). Gaza: The Islamic University of Gaza, 2017.

BERNABEU, Dario y GARCÍA, Mariano, Hefesto Data Warehousing [en línea]. 3. ra ed. Argentina: Free Software Foundation, 2017 [fecha de consulta: 10 de octubre de 2021].

Disponible en: <https://sourceforge.net/projects/bihefesto/files/Hefesto/>

Data mart for machine learning por ZHU, Qiang [et al]. Patent Application [en línea]. Julio 2017. [fecha de consulta: 04 octubre 2020].

Disponible en: <https://patents.google.com/patent/US20170193066A1/en>

GUADAÑA, Iván. Diseño de un Datamart para el apoyo en la toma de decisiones del Departamento de Admisión de la Universidad Peruana Unión, Filial Tarapoto. Revista de Investigación Ciencia, Tecnología y Desarrollo [en línea] Octubre. 2016. [fecha de consulta: 25 noviembre 2021].

Disponible en: <https://doi.org/10.17162/rictd.v2i2.624>

ISSN: 2313-7991

DELGADO, Frank y VICHOT, Andry. Mercado de Datos para la Dirección de Mercadotecnia Operativa de la División de Servicios Fijos de ETECSA. Serie Científica de la Universidad de las Ciencias Informáticas [en línea] Septiembre. 2016. [fecha de consulta: 25 noviembre 2021].

Disponible en: <https://publicaciones.uci.cu/index.php/serie/article/view/837>

ISSN: 2306-2495

CELIS Ballón, Javier. Datamart para la toma de decisiones en el área de ventas de la empresa Farmagro S.A. Tesis (Licenciado en Ingeniería de Sistemas) Lima, Perú: Universidad Cesar Vallejo, 2016.

KNRNETA, Dragoljub, JOVANOVIĆ, Vladan y MARJANOVIĆ, Zoran. An Approach to Data Mart Design from a Data Vault. Journal of Physics [en línea]. 2016. [fecha de consulta: 04 octubre 2020].

Disponible en: <https://infoteh.etf.ues.rs.ba/zbornik/2016/radovi/RSS-1/RSS-1-4.pdf>

WENDA, Jan. Systems and Methods for Data Mart Assembly. U.S. Patent Application. 2016. [fecha de consulta: 04 octubre 2020].

Disponible en: <https://patents.google.com/patent/US20160210315A1/en>

MIGLANI, Renu, KUMAR, Praveen y PAHWA, Payal. Swarm Intelligence Based Data Mart Maintenance. [en línea]2016. [fecha de consulta: 14 octubre 2020].

Disponible en:

<http://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.82F7DA&lang=es&site=eds-live>

Solución de inteligencia de negocio para métricas de gestión de proyectos por MONTERO, Anabel [et al]. SciELO [en línea]. 2015. [fecha de consulta: 04 octubre 2020].

Disponible en:

http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992015000500006&lang=es

CONCEPTUALIZACIÓN DE LA TOMA DE DECISIONES EN EL CUIDADO: ACERCAMIENTO DESDE LA VACUNACIÓN por Marta Fernández Batalla [et al]. ene: revista de enfermería [en línea]. 2015. [fecha de consulta: 03 de octubre de 2020]

Disponible en:

<https://www.ene-enfermeria.org/ojs/index.php/ENE/article/view/590>

GARCÉS, Magdalena. Estudio comparativo de metodologías e implementación de alternativas business intelligence opensource vs. propietarias en entornos tradicionales; caso prototipo en las pymes en el sector agroindustrial. Tesis (Master en Gerencia en Sistemas y Tecnologías de Información) Quito, Ecuador: Universidad de las Américas, 2015.

TOLEE, Alexandru. The Importance of Data Warehouses in the Development of Computerized Decision Support Solutions. A Comparison between Data Warehouses and Data Marts [en línea]. 2015. [fecha de consulta: 28 noviembre 2020].

Disponible

en:

<http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=115441819&lang=es&site=eds-live>

Técnicas para capturar cambios en los datos y mantener actualizado un almacén de datos por DÍAZ, Lisandra [et al.]. SciELO [en línea]. 2015. [fecha de consulta: 28 noviembre 2020].

Disponible en:

<http://search.ebscohost.com/login.aspx?direct=true&db=edssci&AN=edssci.S2227.18992015000400007&lang=es&site=eds-live>

ISSN: 2227-1899

HERNÁNDEZ, Roberto, FERNANDEZ, Carlos y BAPTISTA, Pilar. METODOLOGÍA DE LA INVESTIGACIÓN. 6° ed. México: Interamericana Editores, 2014.

ISBN: 9781456223960

VALDÉS, Yisel, RAMOS, Mayrin y GÓMEZ, Jandy. Mercado de datos para la dirección de cuadros de la Administración Provincial de Artemisa. SciELO [en línea]. 2013. [fecha de consulta: 04 octubre 2020].

Disponible en:

<http://search.ebscohost.com/login.aspx?direct=true&db=edssci&AN=edssci.S2227.18992013000100001&lang=es&site=eds-live>

BRAA, Jorn, HEYWOOD, Arthur y SAHAY, Sundeep. Improving quality and use of data through data-use workshops: Zanzibar, United Republic of Tanzania. SciELO [en línea]. Setiembre 2012. [fecha de consulta: 04 octubre 2020].

Disponible en: <https://www.scielosp.org/article/bwho/2012.v90n5/379-384/en/>

CUBILLO, Julio. La inteligencia empresarial en las pequeñas y medianas empresas competitivas de América Latina: algunas reflexiones. SciELO [en línea]. 1997. [fecha de consulta: 10 octubre 2020].

Disponible en:

https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000300005&lang=es

ISSN: 1518-8353

ANEXOS

Anexo 1: Matriz de Consistencia

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIÓN	INDICADOR	FÓRMULA	ESCALA DE MEDICIÓN
PRODUCTIVIDAD	(Gutiérrez y de la Vara, 2013, p.7) define la Productividad como la relación entre lo producido y los medios empleados.	En base a lo propuesto por Gutiérrez y de la Vara (2013), para medir la variable productividad se dimensiona en eficiencia teniendo como indicador el índice de eficiencia y en eficacia mediante el indicador índice de eficacia y se medirá usando una ficha de registro.	Eficiencia	Índice de Eficiencia	$\text{Eficiencia} = \frac{\text{Reportes Generados/Horas Trabajadas}}{\text{Reportes Objetivos/Horas Previstas}}$ <p>Gutiérrez y de la Vara (2013)</p>	De razón
			Eficacia	Índice de Eficacia	$\text{Eficacia} = \frac{\text{Reportes Generados}}{\text{Reportes Objetivos}}$ <p>Gutiérrez y de la Vara (2013)</p>	De razón

Anexo 2: Instrumento de Recolección de Datos

FICHA DE REGISTRO				
EMPRESA		SEDAPAL		
AREA		EEC-AR		
Indicador a medir		Eficacia		
$\text{Eficacia} = \frac{\text{Reportes Generados}}{\text{Reportes Objetivos}}$				
Item	Fecha	RG	RO	E
Día 1	01/08/2021			
Día 2	02/08/2021			
Día 3	03/08/2021			
Día 4	04/08/2021			
Día 5	05/08/2021			
Día 6	06/08/2021			
Día 7	07/08/2021			
Día 8	08/08/2021			
Día 9	09/08/2021			
Día 10	10/08/2021			
Día 11	11/08/2021			
Día 12	12/08/2021			
Día 13	13/08/2021			
Día 14	14/08/2021			
Día 15	15/08/2021			
Día 16	16/08/2021			
Día 17	17/08/2021			
Día 18	18/08/2021			
Día 19	19/08/2021			
Día 20	20/08/2021			
Día 21	21/08/2021			
Día 22	22/08/2021			
Día 23	23/08/2021			
Día 24	24/08/2021			
Día 25	25/08/2021			
Día 26	26/08/2021			
Día 27	27/08/2021			
Día 28	28/08/2021			
Día 29	29/08/2021			
Día 30	30/08/2021			
Día 31	31/08/2021			
PROMEDIO				

FICHA DE REGISTRO						
EMPRESA		SEDAPAL				
AREA		EEC-AR				
Indicador a medir		Eficiencia				
$\text{Eficiencia} = \frac{\text{Reportes Generados} / \text{Horas Trabajadas}}{\text{Reportes Objetivos} / \text{Horas Previstas}}$						
Horas -> Minutos						
Item	Fecha	RG	HT	RO	HP	EF
Día 1	01/08/2021					
Día 2	02/08/2021					
Día 3	03/08/2021					
Día 4	04/08/2021					
Día 5	05/08/2021					
Día 6	06/08/2021					
Día 7	07/08/2021					
Día 8	08/08/2021					
Día 9	09/08/2021					
Día 10	10/08/2021					
Día 11	11/08/2021					
Día 12	12/08/2021					
Día 13	13/08/2021					
Día 14	14/08/2021					
Día 15	15/08/2021					
Día 16	16/08/2021					
Día 17	17/08/2021					
Día 18	18/08/2021					
Día 19	19/08/2021					
Día 20	20/08/2021					
Día 21	21/08/2021					
Día 22	22/08/2021					
Día 23	23/08/2021					
Día 24	24/08/2021					
Día 25	25/08/2021					
Día 26	26/08/2021					
Día 27	27/08/2021					
Día 28	28/08/2021					
Día 29	29/08/2021					
Día 30	30/08/2021					
Día 31	31/08/2021					
PROMEDIO						

Anexo 3: Matriz de Consistencia

PROBLEMA	HIPÓTESIS	OBJETIVOS	OPERACIONALIZACIÓN DE VARIABLES				METODOLOGÍA
			VARIABLE	DIMENSIONES	INDICADORES	INSTRUMENTO	
General	General	General	Independiente				
¿Cuál será el efecto de la implementación de un datamart en la productividad en el EEC-AR de SEDAPAL?	La implementación de un datamart mejora la productividad en el EEC-AR de SEDAPAL.	Determinar el efecto de la implementación de un datamart para la productividad en el EEC-AR de SEDAPAL.	Datamart				METODOLOGÍA APLICADA: HEFESTO
Específicos	Específicos	Específicos	Dependiente				
¿Cuál será el efecto de la implementación de un datamart en el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL?	La implementación de un datamart mejora el índice de eficiencia de la productividad en el EEC-AR de SEDAPAL.	Determinar el efecto de la implementación de un datamart en el índice de eficiencia en la productividad en el EEC-AR de SEDAPAL.	Productividad	Eficiencia	Índice de Eficiencia	$\text{Eficiencia} = \frac{\text{Reportes Generados/Horas Trabajadas}}{\text{Reportes Objetivos/Horas Previstas}}$ Gutiérrez y de la Vara (2013)	TIPO DE INVESTIGACIÓN: APLICADA
¿Cuál será el efecto de la implementación de un datamart en el índice de eficacia en la productividad en el EEC-AR de SEDAPAL?	La implementación de un datamart mejora el índice de eficacia de la productividad en el EEC-AR de SEDAPAL.	Determinar el efecto de la implementación de un datamart en el índice de eficacia en la productividad en el EEC-AR de SEDAPAL.		Eficacia	Índice de Eficacia	$\text{Eficacia} = \frac{\text{Reportes Generados}}{\text{Reportes Objetivos}}$ Gutiérrez y de la Vara (2013)	DISEÑO DE INVESTIGACIÓN: Pre-Experimental

Anexo 4: Base de datos en SPSS

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	idReporte	Númérico	8	0		Ninguna	Ninguna	8	Derecha	Escala	Entrada
2	eficiencia_pre	Coma	8	4		Ninguna	Ninguna	11	Derecha	Escala	Entrada
3	eficiencia_pro	Coma	8	4		Ninguna	Ninguna	10	Derecha	Escala	Entrada
4	eficacia_pre	Coma	8	0		Ninguna	Ninguna	8	Derecha	Escala	Entrada
5	eficacia_pro	Coma	8	0		Ninguna	Ninguna	8	Derecha	Escala	Entrada
6											

	idReporte	eficiencia_pre	eficiencia_pro	eficacia_pre	eficacia_pro	var
1	1	.5418	.9412	1	1	
2	2	.5128	.9486	1	1	
3	3	.4762	.9677	1	1	
4	4	.4380	.9524	1	1	
5	5	.4528	.6383	1	1	
6	6	.3261	.8333	1	1	
7	7	.4348	.8571	1	1	
8	8	.4412	.8333	1	1	
9	9	.4511	.9524	1	1	
10	10	.3704	.7317	1	1	
11	11	.4054	.7595	1	1	
12	12	.7453	.8633	1	1	
13	13	.7709	.8276	1	1	
14						

Anexo 5: Desarrollo de la metodología de desarrollo de datamart

1.- Análisis de Requerimiento

Lo primero que se hará será identificar los requerimientos de los usuarios a través de preguntas que expliciten los objetivos de su organización. Luego, se analizarán estas preguntas a fin de identificar cuáles serán los Indicadores y Perspectivas que serán tomadas en cuenta para la construcción del DW. Finalmente se confeccionará un Modelo Conceptual en donde se podrá visualizar el resultado obtenido en este primer paso.

Preguntas del Negocio

El objetivo principal de esta fase, es la de obtener e identificar las necesidades de información clave de alto nivel, que es esencial para lograr las metas y ejecutar las estrategias de la empresa, además facilitará que la toma de decisiones sea eficaz y eficiente.

Se desea conocer:

- Total de Usuarios Obturados por Centro de Servicio por Muestreo por Gerencia en un determinado tiempo.
- Total Usuarios Obturados Anexo 1 por Centro de Servicio por Muestreo por Gerencia en un determinado tiempo.
- Total Usuarios Obturados Anexo 2 por Centro de Servicio por Muestreo por Gerencia en un determinado tiempo.
- Total Usuarios Obturados Anexo Ambos por Centro de Servicio por Muestreo por Gerencia en un determinado tiempo.

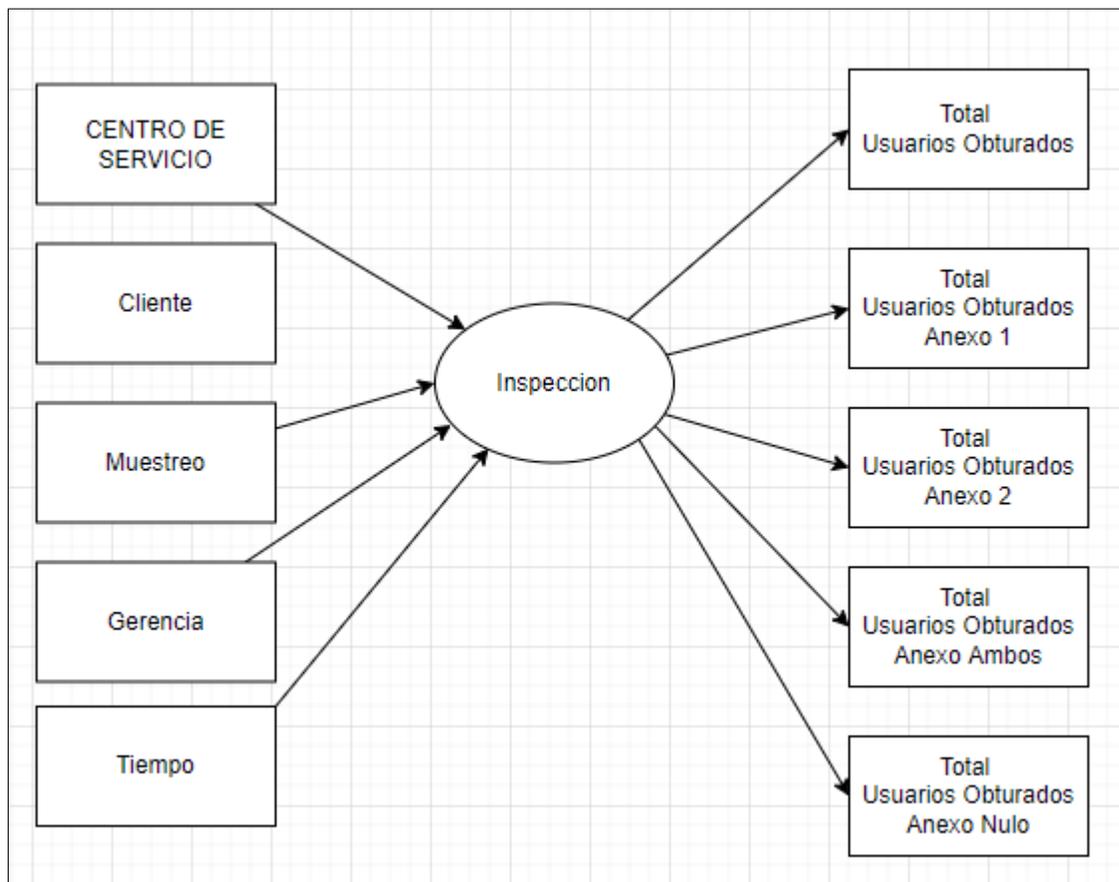
Identificar indicadores y perspectivas

Una vez que se han establecido las preguntas de negocio, se debe proceder a su descomposición para descubrir los Indicadores que se utilizarán y las Perspectivas de análisis que intervendrán.

INDICADORES	PERSPECTIVAS
- Total Usuarios Obturados	- Centro de Servicio
- Total Usuarios Obturados Anexo 1	- Cliente
- Total Usuarios Obturados Anexo 2	- Muestreo
- Total Usuarios Obturados Anexo Ambos	- Gerencia
- Total Usuarios Obturados Anexo Nulo	- Tiempo

Modelo Conceptual

El Modelo Conceptual permite de un solo vistazo y sin poseer demasiados conocimientos previos, comprender cuáles serán los resultados que se obtendrán, cuáles serán las variables que se utilizarán para analizarlos y cuál es la relación que existe entre ellos



2.- Análisis de DATA SOURCES

Hechos e Indicadores

En este paso se explicitan cómo se calcularán los Indicadores, definiendo los siguientes conceptos para cada uno de ellos:

Indicador: “Total de Usuarios Obturados”

Hechos: Total de Usuarios Obturados Función de sumarización: COUNT

Aclaración: el indicador “Total de Usuarios Obturados” representa el conteo de los usuarios de la BASE TOTAL.

Indicador: “Total Usuarios Obturados Anexo 1”

Hechos: Total Usuarios Obturados Anexo 1 Función de sumarización: COUNT

Aclaración: el indicador “Total Usuarios Obturados Anexo 1” representa el conteo de los usuarios de la BASE TOTAL por Anexo 1.

Indicador: “Total Usuarios Obturados Anexo 2”

Hechos: Total Usuarios Obturados Anexo 2 Función de sumarización: COUNT

Aclaración: el indicador “Total Usuarios Obturados Anexo 2” representa el conteo de los usuarios de la BASE TOTA por Anexo 2.

Indicador: “Total Usuarios Obturados Anexo Ambos”

Hechos: Total Usuarios Obturados Anexo Ambos Función de sumarización: COUNT

Aclaración: el indicador “Total Usuarios Obturados Anexo Ambos” representa el conteo de los usuarios de la BASE TOTAL por Anexo 1 y 2.

Indicador: “Total Usuarios Obturados Anexo Ambos”

Hechos: Total Usuarios Obturados Anexo Ambos Función de sumarización: COUNT

Aclaración: el indicador “Total Usuarios Obturados Anexo Ambos” representa el conteo de los usuarios de la BASE TOTAL por Anexo 1 y 2.

Indicador: "Total Usuarios Obturados Anexo Nulo"

Hechos: Total Usuarios Obturados Anexo Nulo Función de sumarización: COUNT

Aclaración: el indicador "Total Usuarios Obturados Anexo Nulo" representa el conteo de los usuarios de la BASE TOTAL sin Anexo.

Mapeo

No se muestra el diagrama relacional del OLTP pues contiene información privada, aun así, se especifican las correspondencias.

- Centro de Servicio: Son 6 posibles opciones que determina SEDAPAL
- Cliente: Usuarios Obturados
- Muestreo: El tipo de muestreo son 2 tipos.
- Gerencia: Son 3 posibles opciones que determina SEDAPAL.
- Tiempo: Se usa la fecha de muestreo.
- Total de Usuarios Obturados según Anexo: Anexo 1, anexo 2, ambos anexos y no excede (4 opciones)

Granularidad

-Perspectiva Clientes:

CLIENTE: Razón Social de la empresa.

DIRECCION: Ubicación de empresa o domicilio.

DISTRITO: Nombre del distrito de la dirección.

NIS: Código según SEDAPAL

NIA: Código 2 según SEDAPAL

-Perspectiva CCSS:

CCSS: Nombre del Centro de Servicio

-Perspectiva TIPOMUESTREO:

TPO DE MUESTREO: Nombre del tipo del muestreo

-Perspectiva Gerencia:

GERENCIA: Nombre del tipo de gerencia.

-Perspectiva Tiempo:

Año: Como 2018

MesCorto: Como Ene, Feb y Dic

NMes: Numero de mes 1

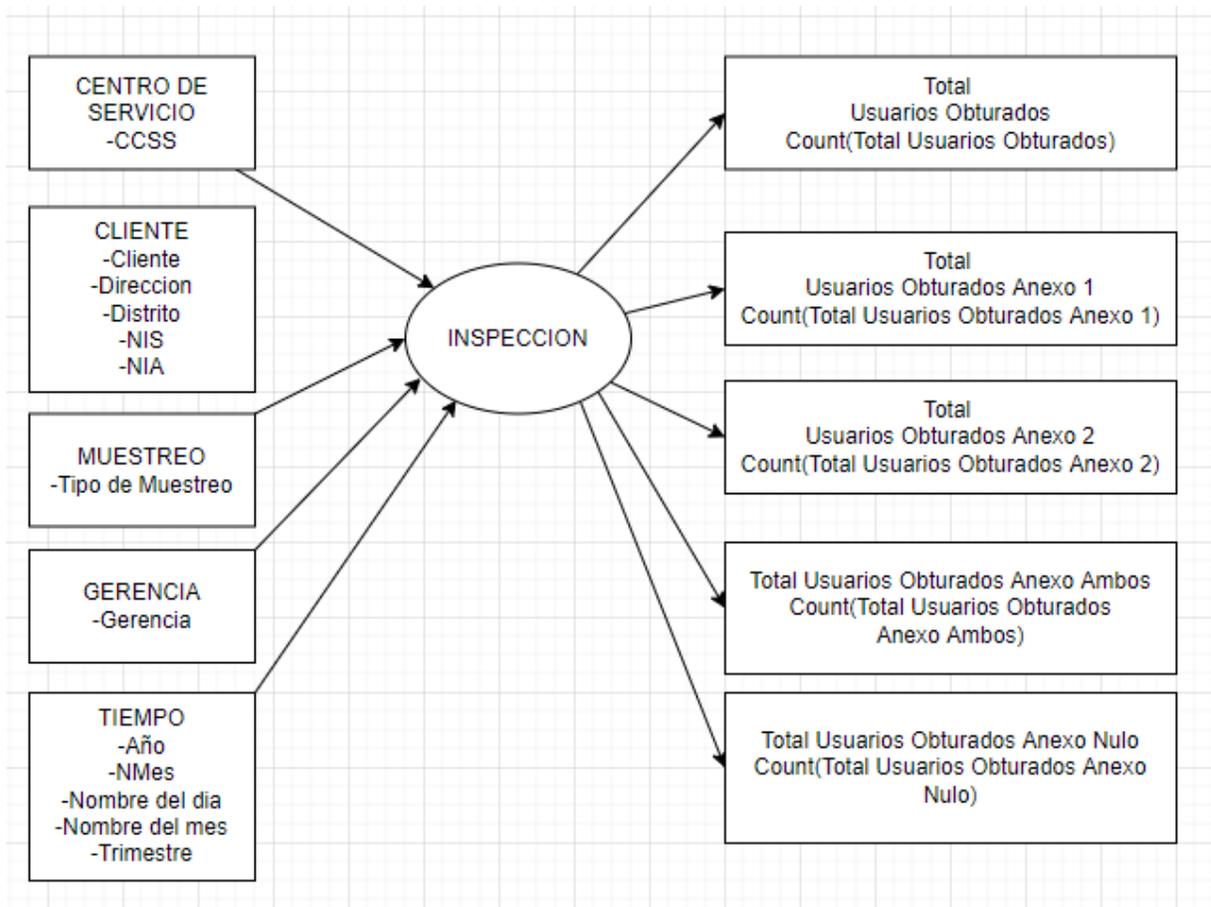
Nombre del día: Lunes, martes, ...

Nombre del mes: Enero, febrero, ...

Trimestre: Numero de Trimestre

Modelo Conceptual Ampliado

En este paso, y con el fin de graficar los resultados obtenidos en los pasos anteriores, se ampliará el Modelo Conceptual, colocando debajo de cada Perspectiva los campos seleccionados y debajo de cada Indicador su respectiva fórmula de cálculo.



3.- Modelo Lógico del Datamart

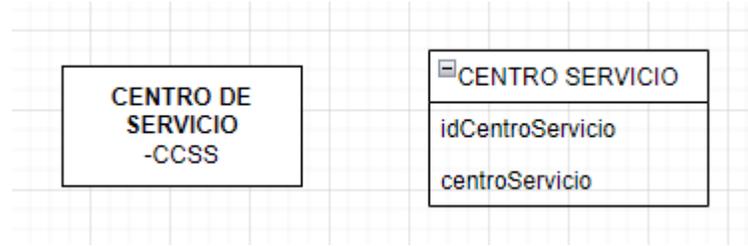
Tipología

Se ha seleccionado el Esquema en Estrella ya que cumple con los requerimientos planteados y es simple de implementar y comprender.

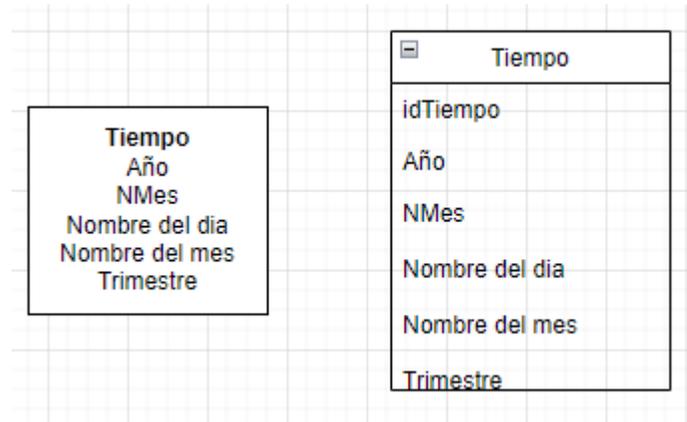
Tabla de dimensiones

En este paso diseñaremos las tablas de Dimensiones que formarán parte del DM.

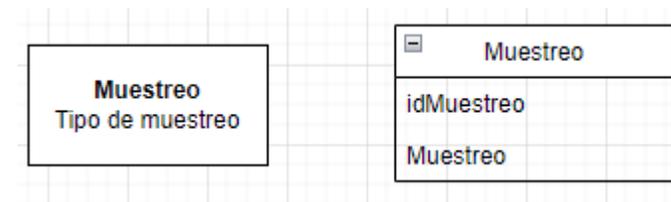
Dimensión Ccss



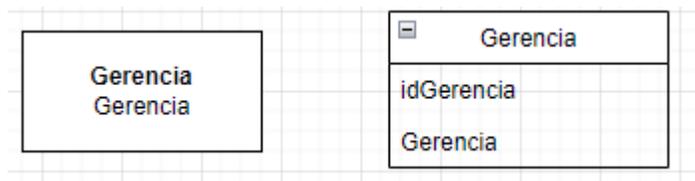
Dimensión Tiempo



Dimensión Muestreo



Dimensión Gerencia



Dimensión Cliente

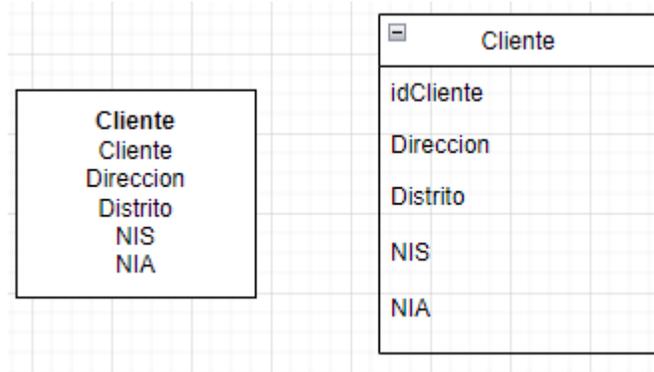
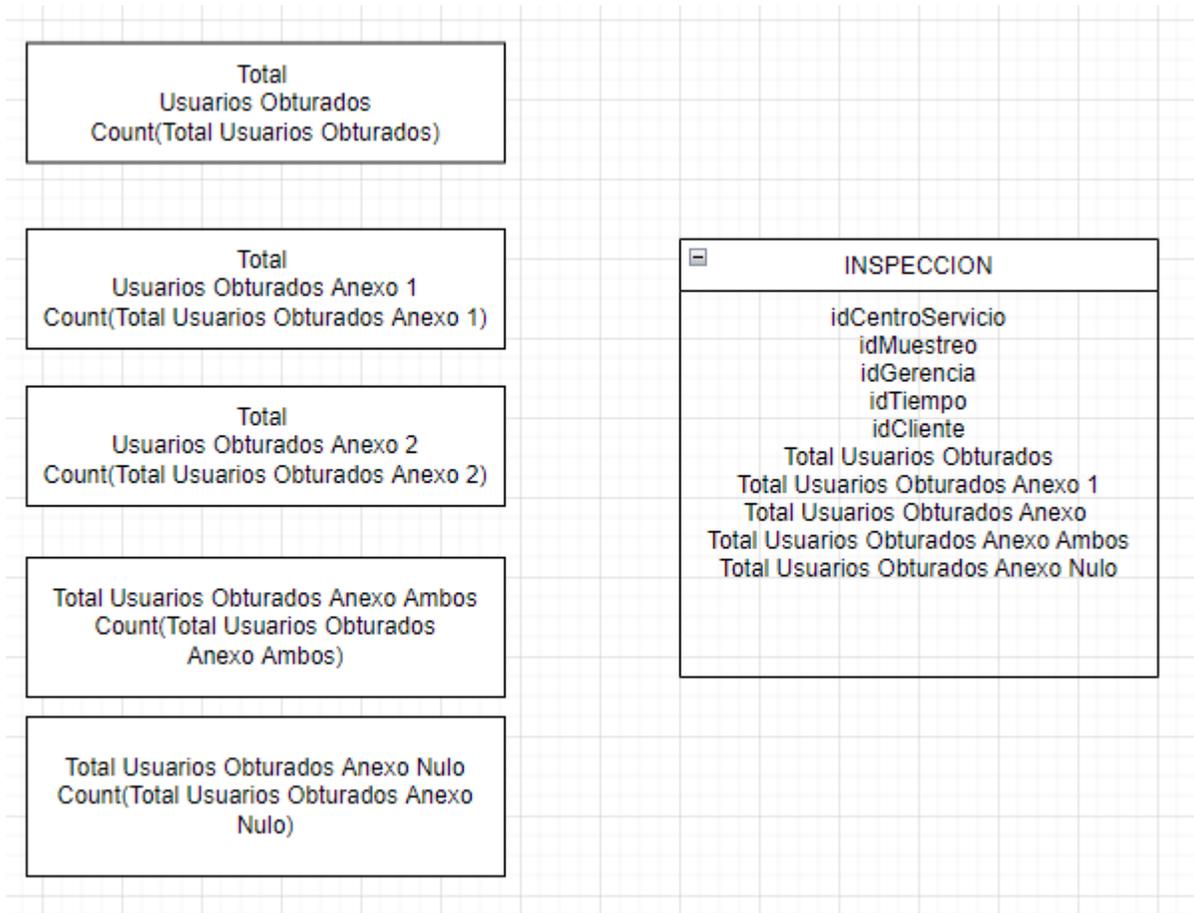


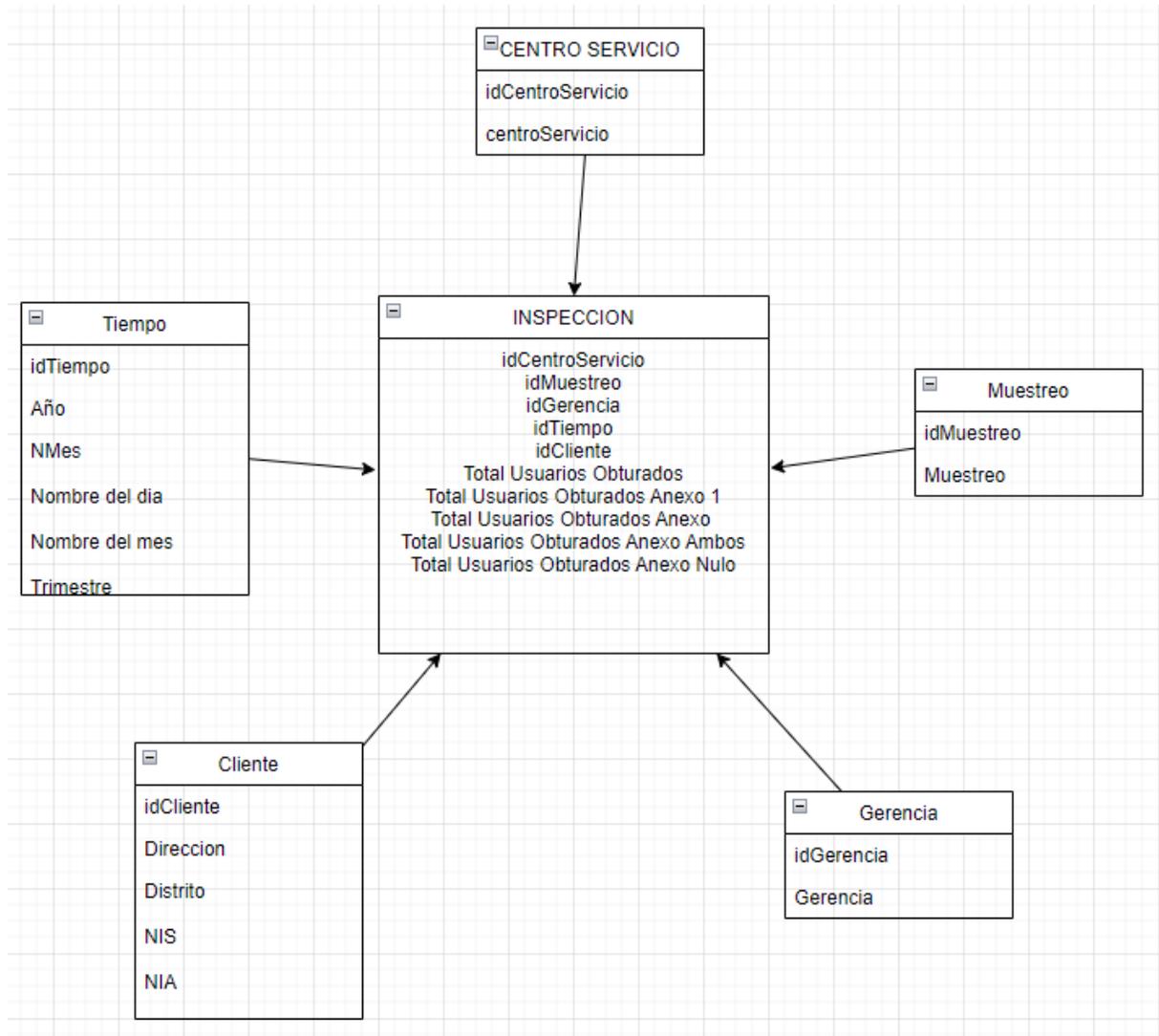
Tabla de Hechos

En este paso, se definirán las tablas de Hechos.



Uniones

Se realizarán las uniones correspondientes entre sus tablas de Dimensiones y sus tablas de Hechos.



4.- Integración de Datos

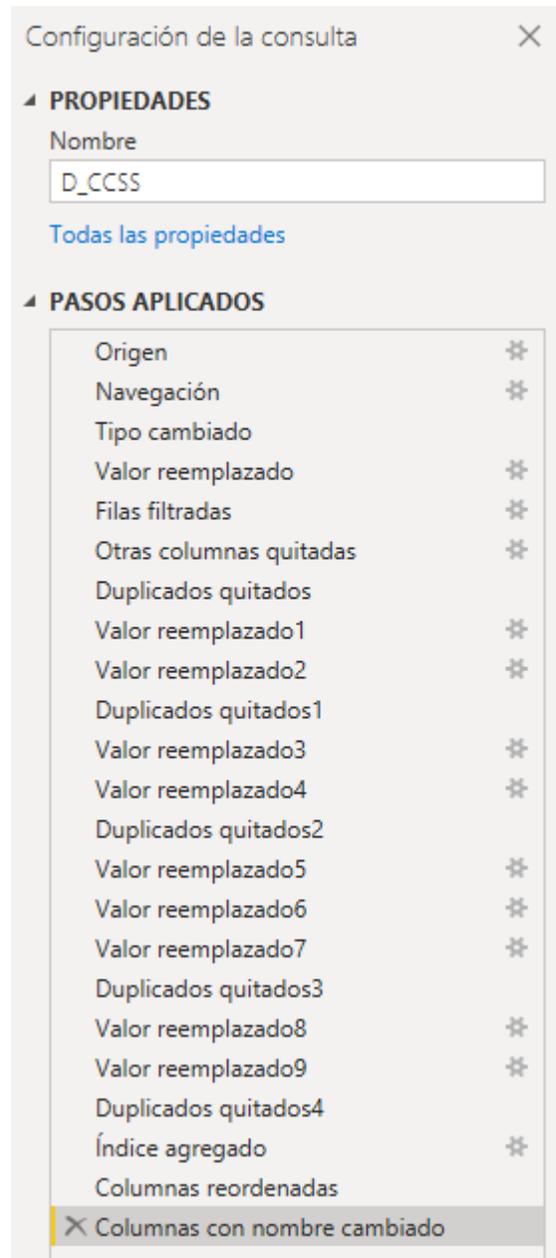
Carga inicial

Las tablas de dimensiones y hechos son alimentadas por la tabla propia que maneja la empresa, al crear cada tabla de dimensión se realizó los procesos ETL con PowerBI necesarios para que contenga valores únicos y puedan interactuar correctamente

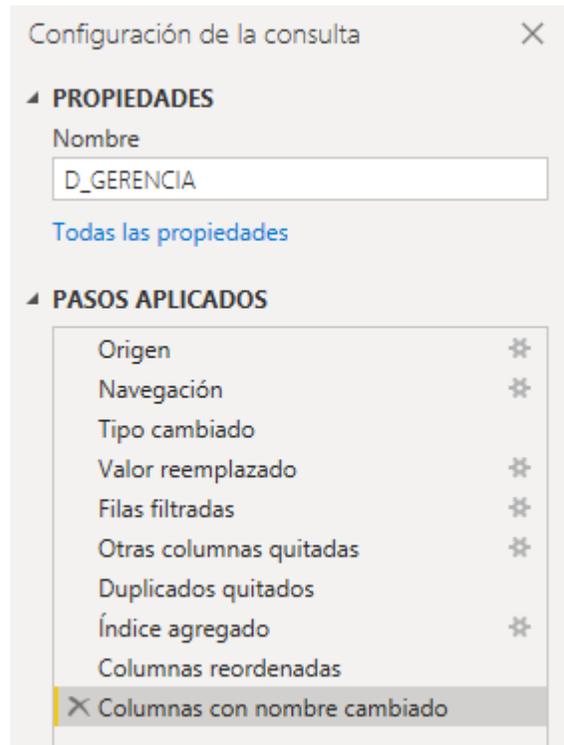
con la tabla de hechos al mismo tiempo que puedan ser analizadas desde las perspectivas planteadas.

Todos estos pasos están registrados y se puede ver el antes y después desde el Power Query o Editor de Power Query.

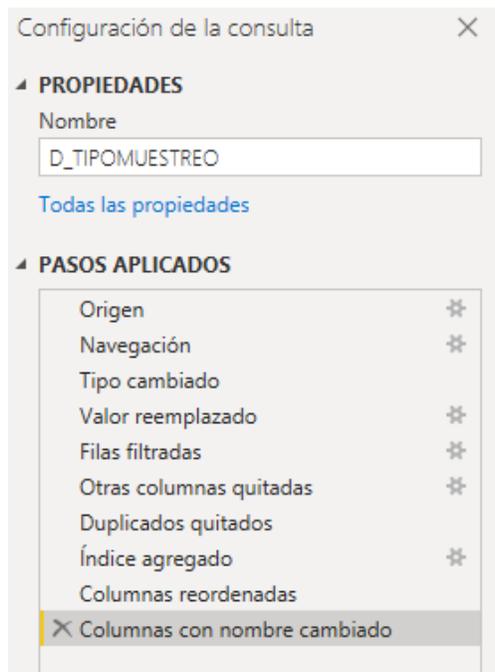
Dimensión CCSS



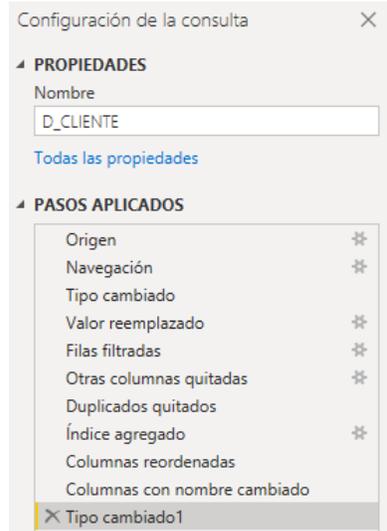
Dimensión Gerencia



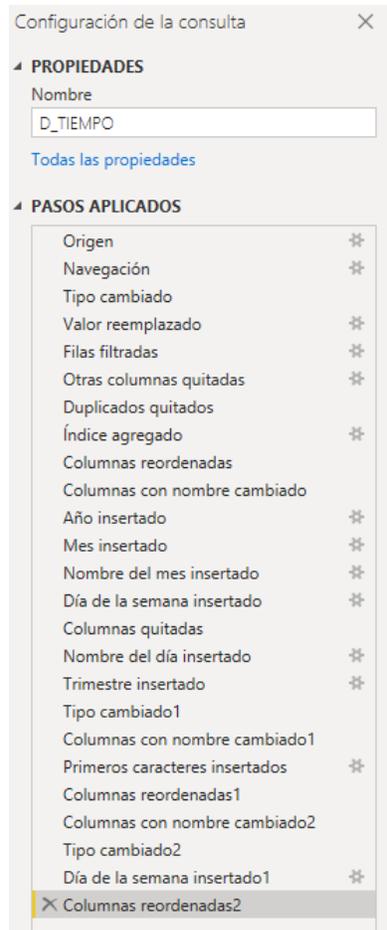
Dimensión TipoMuestreo



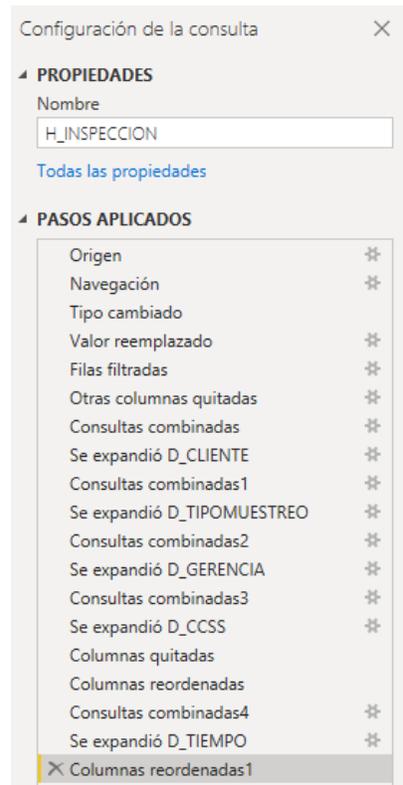
Dimensión Cliente



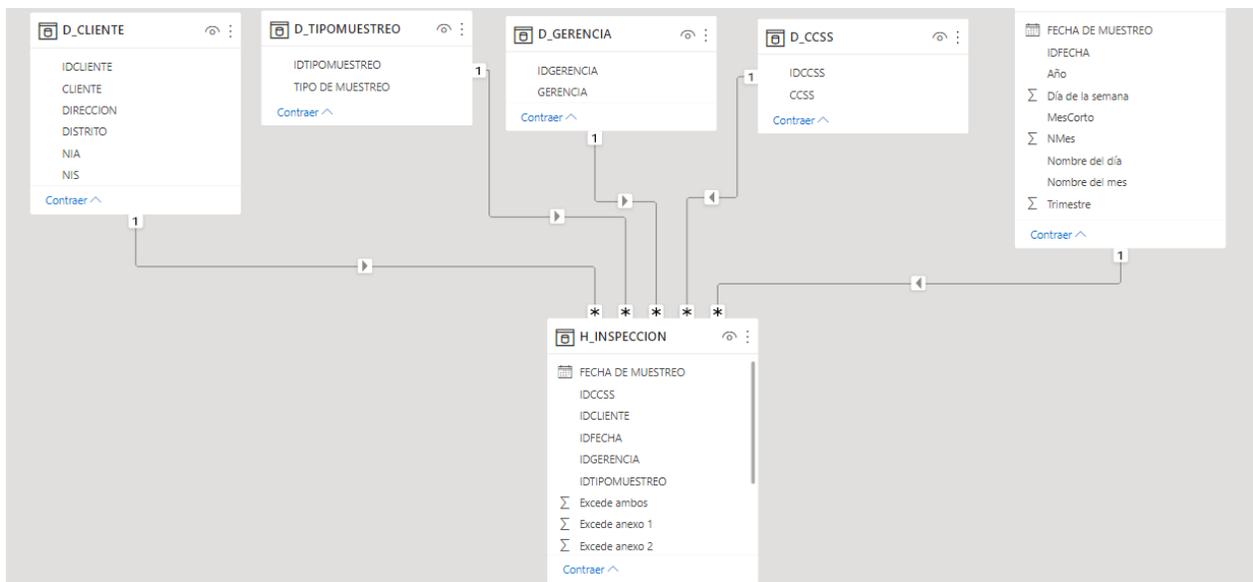
Dimensión Tiempo



Hecho Inspeccion

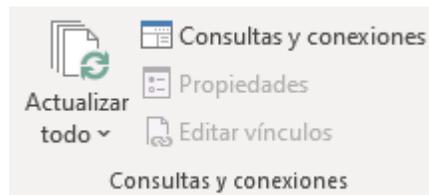


Dando el resultado de el resultado de cubo multidimensional, donde se cargan primero las tablas de dimensiones y finalmente los registros enlazados con la tabla de hechos.

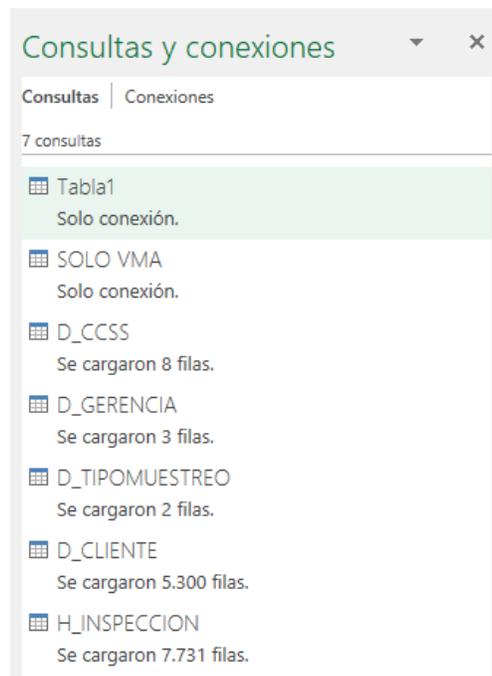


Actualización

Tras cualquier modificación o eliminación en los registros se podrán actualizar los datos desde el botón mostrado, puesto que el modelo de datos desarrollado en Power Pivot ya fue estructurado para que realice los procesos ETL necesarios desde el Power Query para filtrar la información independientemente de los datos que se utilicen de entrada.



Al actualizar, todas la conexiones y tablas se actualizarán dando inicio a la ejecución de procesos programados para que los datos se integren.



Anexo 6: Autorización de la empresa para realizar la investigación



SERVICIO DE AGUA POTABLE Y ALCANTARILLADO DE LIMA

Equipo Evaluación de Calidad de Aguas Residuales

"Decenio de la Igualdad de Oportunidades para mujeres y hombres"
"Año del Bicentenario del Perú: 200 años de Independencia"

Asunto : Constancia de trabajo de investigación

Que el señor Ccasani Condori Christian, está realizando el trabajo de investigación titulado: **"Datamart para mejorar la productividad del Equipo de Evaluación de Calidad de Aguas Residuales de SEDAPAL"**. Dicho trabajo se realiza con el personal que corresponde al Equipo de Evaluación de Calidad de Aguas Residuales (EEC-AR). El área está presta a colaborar con el estudiante brindando información general, mas no detallada que a nuestro criterio es confidencial.

Se expide la presente a solicitud del interesado para los fines que estime conveniente

Atentamente,

Carmen Inés Ccasani Condori
Jefa Equipo Evaluación de Calidad de Aguas Residuales (e)

Lima, 23 de octubre del 2021.

Anexo 7: Entrevista N°1

Dirigido a: EECAR – SEDAPAL

Entrevistador: Ccasani Condori, Christian

Fecha: 12/09/2020

Duración: 50 min

¿Cuál es el objetivo del EECAR?

Cumplir y hacer cumplir la normativa VMA, a los usuarios de SEDAPAL, la cual está dividida en Anexo 1 y Anexo 2, esta normativa tuvo una actualización en 2015, definiendo 23 parámetros, los cuales tienen el objetivo de mantener estándares de calidad en las aguas residuales con el fin social de preservar el medio ambiente.

¿Se cumplen los objetivos del área?

Se están cumpliendo, sí, pero también detectamos que el sistema con el que se trabaja y la organización del área no es la mejor. Se cumplen las metas, pero hay exceso de trabajo y falta de tiempo para otras tareas después de terminar los reportes.

¿Qué interfiere con las operaciones del EECAR?

La base de datos con la que se trabaja no viene siendo actualizada en los tiempos programados, por lo que siempre que hay que trabajar, hay momentos en donde debemos parar para priorizar la actualización de la base de datos; a pesar de que el equipo pone esfuerzos en que esto no ocurra o si no sería un desastre.

¿Ocurrió algún incidente en el área?

Sucedan a veces, cuando el seguimiento de un usuario falla en alguna parte de la metodología, este tiene el derecho de por más que haya sido un infractor en la normativa, sea eximido por un tiempo.

¿El proceso de toma de decisiones mediante reportes como lo nota?

Funciona, tiene que funcionar, pero es deficiente ya que los datos están desorganizados.

¿Cuál es el tiempo promedio en la elaboración de reportes?

En promedio 2 horas.

¿En qué medida cree que los datos no están disponibles y no son fiables?

En 4 de 10 casos generalmente no se cuenta con la información necesaria así que tenemos que priorizar la actualización y los datos son más fiables, considerando que si están incompletos o redundantes sería 3 de cada 10 casos.

¿Planean solucionar el problema?

Si y no, obviamente se planea la solución, pero nunca se llega en nada concreto y trabajamos con lo actual.

¿Hubo algún intento de mejora?

Se planeo hace un tiempo el automatizar los registros, pero nunca se completó, el proyecto que fue idea de un practicante nunca se pasó a la elaboración y entonces quedo en nada.

¿Existe perdida de datos?

Los datos están protegidos puesto que un área técnica asiste en esas funciones al EECAR y solo al EECAR.

¿Qué desearía el equipo para realizar sus funciones con normalidad?

Disponer de la información cuando se requiera y que esta esté libre de fallas, para no dejar de realizar nuestras otras labores por priorizar la actualización de la información.

¿Qué piensa sobre la implementación de un datamart en el área?

Si, eh escuchado de eso, y sé que se trabaja de manera más organizada; al área le vendría bien un sistema como ese.