

UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



PUNO – PERÚ

2014

UNIVERSIDAD NACIONAL DEL ALTIPLANO

FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

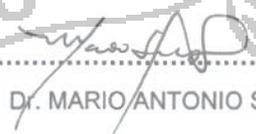
“DATAMART DE CONTRATACIONES PÚBLICAS A PARTIR DEL SEACE, Y SU APLICACIÓN EN LA TOMA DE DECISIONES DE LAS MICRO Y PEQUEÑAS EMPRESAS DE LA CIUDAD DE PUNO - 2013”

TESIS PRESENTADA POR:

RONAN FROYLAN ACERO CALIZAYA

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE SISTEMAS

APROBADA POR EL JURADO REVISOR CONFORMADO POR:

PRESIDENTE	:	 Dr. ANGEL MANUEL OLAZABAL GUERRA
PRIMER MIEMBRO	:	 Mg. ELMER COYLA IDME
SEGUNDO MIEMBRO	:	 Dr. YALMAR TEMISTOCLES PONCE ATENCIO
DIRECTOR DE TESIS	:	 Dr. MARIO ANTONIO SUÁREZ LÓPEZ

PUNO – PERÚ

2014

ÁREA: Informática

TEMA: Sistemas de información tradicionales y expertos

Agradecimiento

A la universidad Nacional del Altiplano

A los docentes de la Facultad de Ingeniería Mecánica Eléctrica, Electrónica y Sistemas, Escuela Profesional de Ingeniería de Sistemas.

Al Ing. Angel Manuel Olazabal Guerra, por la atenta lectura de este trabajo, al Ing. Elmer Coyla Idme por sus consejos y por el ánimo que nos inyectó para terminar nuestro trabajo y al Ing. Yalmar Ponce Atencio por haber confiado en nosotros, por la paciencia y por el direccionamiento de este trabajo.



Dedicatoria

A mi familia.

A mis padres Pedro y Gregoria que hicieron todo en la vida para que pueda lograr mis sueños.



ÍNDICE

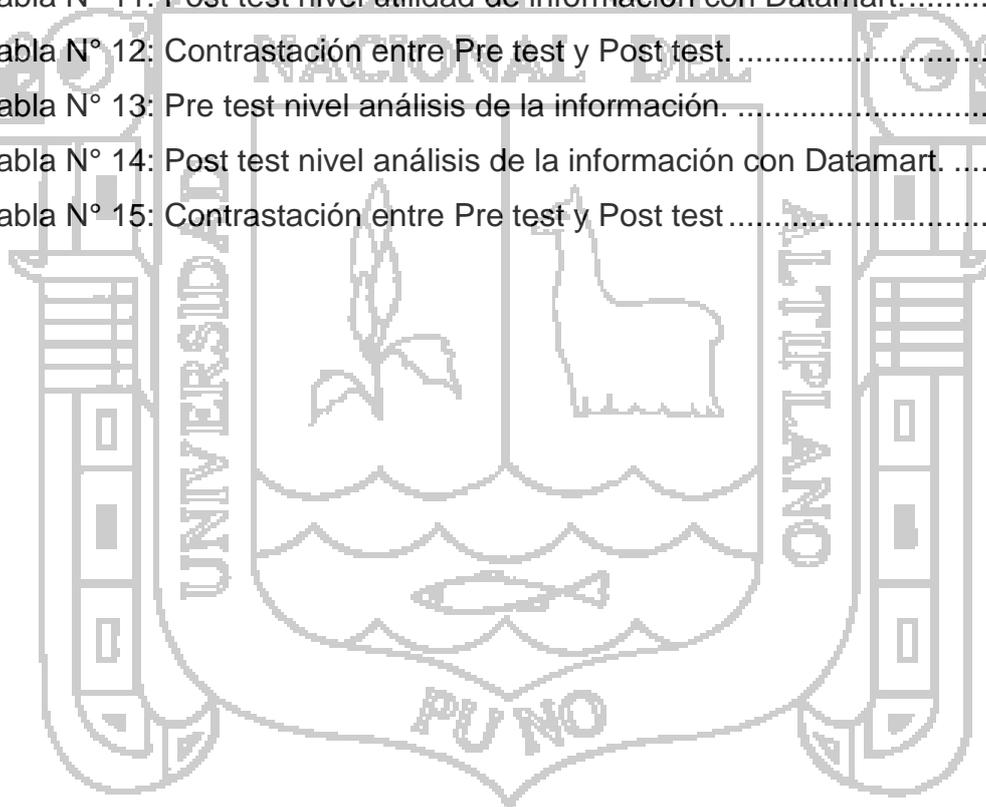
RESUMEN.....	10
INTRODUCCIÓN.....	12
CAPITULO I.....	14
PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN	
1.1. DESCRIPCIÓN DEL PROBLEMA.....	15
1.2. FORMULACIÓN DEL PROBLEMA.....	16
1.2.1. PROBLEMA GENERAL.....	17
1.2.2. PROBLEMAS ESPECÍFICOS.....	17
1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	17
1.4. OBJETIVOS DE LA INVESTIGACIÓN.....	19
1.4.1. OBJETIVO GENERAL.....	19
1.4.2. OBJETIVOS ESPECÍFICOS.....	19
1.5. DELIMITACIÓN DE LA INVESTIGACIÓN.....	20
1.6. LIMITACIÓN DE LA INVESTIGACIÓN.....	20
CAPITULO II.....	21
MARCO TEÓRICO	
2.1. ANTECEDENTES DE LA INVESTIGACIÓN.....	22
2.1.1. ANTECEDENTES NACIONALES.....	22
2.1.2. ANTECEDENTES INTERNACIONALES.....	24
2.2. BASE TEÓRICA.....	27
2.2.1. MYPES EN LA CIUDAD DE PUNO.....	27
2.2.1.1. CONCEPTO DE MICRO Y PEQUEÑA EMPRESA EN PERÚ.....	27
2.2.1.2. MYPES SEGÚN SU ORGANIZACIÓN EMPRESARIAL.....	27
2.2.2. CONTRATACIONES PÚBLICAS.....	28
2.2.2.1. PROCESOS DE SELECCIÓN.....	28
2.2.3. TOMA DE DECISIONES.....	29
2.2.3.1. TIPOS DE DECISIONES.....	31
2.2.3.2. EL PROCESO DE TOMA DE DECISIONES.....	32
2.2.4. DATAMART.....	34

2.2.4.1. DEFINICIÓN DE DATA WAREHOUSE.	34
2.2.4.2. DEFINICIÓN DE DATAMART.....	35
2.2.4.3. CARACTERÍSTICAS DEL DATAMART.....	36
2.2.4.4. DIFERENCIAS ENTRE UN DATAWAREHOUSE Y DATAMART.....	36
2.2.5. ARQUITECTURA DEL DATAMART.....	37
2.2.5.1. DATOS OPERACIONALES.....	37
2.2.5.2. PROCESO ETL (Extracción, Transformación y Carga).....	38
2.2.5.2. BASE DE DATOS MULTIDIMENSIONAL.....	42
2.2.6. MODELAMIENTO DE UN DATAMART.....	43
2.2.6.1. ESQUEMA ESTRELLA.....	43
2.2.7. METODOLOGÍA DE DESARROLLO DE DATAMART.....	44
2.2.8. METODOLOGÍA HEFESTO.....	45
2.2.8.1. CARACTERÍSTICAS.....	45
2.2.8.2. FASES DE LA METODOLOGÍA HEFESTO.....	47
2.2.9. BASE DE DATOS.....	54
2.2.9.1. MYSQL.....	54
2.2.10. PLATAFORMA PENTAHO.....	54
2.2.10.1. HERRAMIENTA ETL: TALEND OPEN STUDIO.....	55
2.2.10.2. HERRAMIENTA DE ANALISIS: PENTAHO ANALYSIS	
SERVICES.....	56
2.3. DEFINICIÓN DE TÉRMINOS.....	56
2.4. HIPÓTESIS Y VARIABLES.....	59
2.4.1. HIPÓTESIS GENERAL.....	59
2.4.2. HIPÓTESIS ESPECÍFICAS.....	59
2.5. VARIABLES.....	59
2.6. OPERACIONALIZACIÓN DE VARIABLES.....	60
CAPITULO III.....	61
DISEÑO METODOLÓGICO DE INVESTIGACIÓN	
3.1. TIPO Y DISEÑO DE LA INVESTIGACIÓN.....	62
3.2. POBLACIÓN.....	63
3.3. MUESTRA.....	63
3.4. TÉCNICA DE RECOLECCIÓN DE INFORMACIÓN.....	63

3.5. INSTRUMENTO.....	64
3.6. TÉCNICA PARA EL PROCESAMIENTO Y ANÁLISIS DE DATOS.	64
3.7. DISEÑO ESTADÍSTICO PARA LA PRUEBA DE HIPÓTESIS.....	64
3.8. MATERIAL EXPERIMENTAL.	67
3.8.1. METODOLOGÍA DE DESARROLLO DE DATAMART.	67
3.8.2. HARDWARE Y SOFTWARE DE DESARROLLO.....	68
3.8.3. ANÁLISIS DE REQUERIMIENTOS.	69
3.8.4. ANÁLISIS DEL SISTEMA TRANSACCIONAL.....	78
3.8.5. MODELO LÓGICO DEL DATAMART.....	83
3.8.6. PROCESO DE ETL.	87
3.8.7. CREACIÓN DE CUBOS MULTIDIMENSIONALES.....	89
3.8.8. DESPLIEGUE Y ANÁLISIS DE DATOS PARA EL APOYO DE TOMA DE DECISIONES.....	91
CAPITULO IV.....	94
ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	
4.1. CONTRASTE DE HIPÓTESIS.....	95
4.1.1. CONTRASTE DE HIPÓTESIS ESTADÍSTICA 1.....	95
4.1.2. CONTRASTE DE HIPÓTESIS ESTADÍSTICA 2.....	101
4.1.3. DECISIÓN FINAL.....	107
CONCLUSIONES.....	108
RECOMENDACIONES.....	109
BIBLIOGRAFÍA.....	110
ANEXOS.....	114

ÍNDICE DE TABLAS

Tabla N° 1: Diferencia entre Data Warehouse y Datamart.....	37
Tabla N° 2: Hardware y Software de desarrollo	68
Tabla N° 3: Análisis de Requerimientos.....	70
Tabla N° 4: Descripción de la tabla DimProducto	84
Tabla N° 5: Descripción de la Tabla DimProveedor	85
Tabla N° 6: Descripción de la Tabla DimEntidad	85
Tabla N° 7: Descripción de la tabla DimProceso	85
Tabla N° 8: Descripción de la tabla DimTiempo.....	85
Tabla N° 9: Descripción de la Tabla de FactContratación.....	86
Tabla N° 10: Pre test nivel utilidad de información.....	96
Tabla N° 11: Post test nivel utilidad de información con Datamart.....	97
Tabla N° 12: Contrastación entre Pre test y Post test.....	98
Tabla N° 13: Pre test nivel análisis de la información.....	102
Tabla N° 14: Post test nivel análisis de la información con Datamart.....	103
Tabla N° 15: Contrastación entre Pre test y Post test.....	105



ÍNDICE DE FIGURAS

Figura N° 1: Transformación de la información en acción.....	31
Figura N° 2: Tipos y características de decisión.	32
Figura N° 3: Etapas en la Toma de decisiones.	33
Figura N° 4: Data Warehouse.	34
Figura N° 5: Transformación: Codificación.....	40
Figura N° 6: Transformación: Medida de Atributos.	40
Figura N° 7: Cubo Multidimensional.....	43
Figura N° 8: Ejemplo de esquema estrella.....	44
Figura N° 9: Fases de la metodología HEFESTO.....	47
Figura N° 10: Modelo conceptual.....	49
Figura N° 11: Modelo conceptual ampliado.	51
Figura N° 12: Diseño de tablas de dimensiones	52
Figura N° 13: Tabla de hechos (Esquema estrella).....	53
Figura N° 14: Arquitectura funcional de Pentaho.....	55
Figura N° 15: Modelo conceptual. Perspectiva-Indicadores.....	75
Figura N° 16: Proceso de solución.....	76
Figura N° 17: Lista de Procesos de selección.....	77
Figura N° 18: Detalle de un Proceso de selección.....	77
Figura N° 19: Relación de perspectivas y Datos fuente.....	80
Figura N° 20: Modelo conceptual ampliado.	83
Figura N° 21: Modelo físico del Datamart	87
Figura N° 22: Configuración de proceso ETL para las Dimensiones.....	88
Figura N° 23: Resultado de la transformación para DimTiempo	89
Figura N° 24: Definición de Indicadores, Atributos y Jerarquías.....	90
Figura N° 25: Pantalla de Inicio de Pentaho.	91
Figura N° 26: Análisis de información, de general a específico.	92
Figura N° 27: Precio unitario adjudicado en el tiempo	93
Figura N° 28: Gráfico de Barras – Contrastación de Pre test y Post test.	99
Figura N° 29: Gráfico de Barras – Contrastación de Pre test y Post test.	105

RESUMEN

El presente trabajo investigación titulado "Datamart de Contrataciones públicas a partir del SEACE, y su aplicación en la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013", tuvo como objetivo general determinar el grado de influencia de la aplicación de la Datamart de Contrataciones públicas en la toma de decisiones de las MYPEs. Durante el desarrollo de la investigación se abordó los conceptos de análisis, exploración, información histórica, procesos selección, contratación de bienes y servicios. Se utilizó la metodología HEFESTO para el desarrollo del Datamart, desde los requerimientos hasta la difusión de los resultados, permitiendo que los usuarios finales (MYPEs) analicen y exploren la información de manera sencilla e intuitiva. El Datamart de Contrataciones públicas contiene el comportamiento de los precios, cantidades, montos adjudicados de contratación de bienes y servicios por entidad, montos adjudicados por proveedor (MYPE), la información es presentado en función al tiempo. Para el despliegue de la solución Datamart se utilizó la herramienta Pentaho con capacidad de personalizar reportes de acuerdo a las necesidades del negocio y MySQL como motor de base de datos. La investigación siguió el diseño de investigación pre test y post test, donde se demostró estadísticamente la hipótesis planteado y como resultado final se obtuvo que la aplicación del Datamart de Contrataciones públicas a partir del SEACE apoya significativamente la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

Palabras claves: Datamart, análisis, información, bienes, servicios y contratación.

ABSTRACT

The present work investigation titled "Datamart of public contractings from the SEACE, and his application in the capture of decisions of the Micro and small enterprises of Puno's city - 2013 ", it had as general aim determine the degree of influence of the application of the Datamart of public contractings in the capture of decisions of the MYPEs. During the development of the investigation there were approached the concepts of analysis, exploration, historical information, processes selection, contracting of goods and services. There was in use the methodology HEFESTO for the development of the Datamart, from the requirements up to the diffusion of the results, allowing that the final users (MYPEs) should analyze and explore the information of a simple and intuitive way. The Datamart of public contractings contains the behavior of the prices, quantities, amounts awarded of contracting goods and services for entity, amounts awarded by supplier (MYPE), the information is presented in function to the time. For the deployment of the solution Datamart there was in use the tool Pentaho with aptitude to personalize reports of agreement to the needs of the business and MySQL as engine of database. The investigation followed the design of investigation pretest and posttest, where the hypothesis was demonstrated statistically raised and as final result there was obtained that the application of the Datamart of public Contractings from the SEACE supports significantly the capture of decisions of the Micro and small enterprises of Puno's City - 2013.

Key words: Datamart, analysis, information, goods, services and contracting.

INTRODUCCIÓN

La capacidad para tomar decisiones con rapidez basadas en un adecuado conocimiento de la realidad de la Empresa así como los mercados y sus tendencias, ha pasado a convertirse en nueva fuente de ventaja competitiva. La aplicación del Datamart genera escenarios, pronósticos y reportes. Con el cual se da a conocer las tendencias del mercado y entender su propia posición de una Empresa con respecto a sus competidores y esta información permite que el proceso de toma de decisiones esté fundamentado.

El presente proyecto de investigación muestra la aplicación del Datamart en la toma de decisiones de las MYPEs para contratar con el estado. Para ello se realiza la recolección de los requerimientos, extracción de la información desde una fuente (SEACE), Integración y carga de datos en un formato utilizable para el análisis (Datamart), Análisis y producción (Pentaho) y finalmente se realiza la difusión que permite a los MYPEs explorar los datos de manera sencilla e intuitiva.

Este trabajo de investigación contiene cuatro capítulos, todos ellos debidamente concordantes, pretendiendo hacer un trabajo coherente y de fácil comprensión para los lectores, tal como se detalla a continuación.

Capítulo I: Planteamiento del Problema, Justificación de la Investigación, Objetivos de la Investigación, Limitaciones de la Investigación, estos puntos sitúan al lector dentro de la problemática de la Investigación.

Capítulo II: El Marco Teórico compuesto por antecedentes de la investigación, marco conceptual, definición de términos básico más utilizados a lo largo de la

Investigación, Base teórica, donde reside toda la teoría necesaria para el mejor entendimiento y finalmente la hipótesis del presente trabajo de investigación.

Capítulo III: Diseño metodológico de investigación está constituido por la definición del tipo y diseño de investigación, técnicas empleadas para la recolección de datos y su posterior análisis estadístico, material experimental donde se muestra el desarrollo del Datamart siguiendo la metodología HEFESTO.

Capítulo IV: Y este capítulo presenta el análisis y la interpretación de los resultados de la investigación.

Y por último se presenta las conclusiones a las que se ha llegado con la aplicación del Datamart en la toma de decisiones y las recomendaciones para las futuras investigaciones.





1.1. DESCRIPCIÓN DEL PROBLEMA.

El sector público en su conjunto es uno de los mayores compradores del mercado, donde el presupuesto para la compra es elevado y se incrementa año tras año, el acceso a este mercado es una oportunidad para mantener o aumentar el volumen de ventas para una Micro y Pequeña Empresa(MYPE).

Según el Organismo Supervisor de las Contrataciones del Estado (OSCE) en un análisis de las Contrataciones públicas en cada región del país, concluyó que durante el año 2012 el Estado Peruano adquirió bienes y servicios por 15 453.80 millones de nuevo soles, cifra 41,1% mayor que la registrada en el año 2011 (11 447.20 millones de nuevo soles). También se indica que el número promedio de participantes por proceso de selección es de 1.8 participantes por un proceso de selección. (Informe OSCE, 2013). Según este estudio se afirma que existe una gran capacidad de compra por parte del sector público y poca participación de la Empresas en los procesos de selección.

Para las MYPEs según las cifras descritas anteriormente, vender al estado es un objetivo que se toma en consideración, porque el sector público tiene alto poder adquisitivo. Esto significa que se tiene muchas oportunidades para ofertar, por lo tanto abrir nuevas expectativas del negocio.

La información de los procesos de selección realizadas por las Entidades Públicas para proveerse de bienes y servicios, están difundidas mediante el Sistema Electrónico de Contrataciones del Estado (SEACE), esta información muestra la capacidad de compra del sector público, información que no es aprovechada adecuadamente por las Micro y Pequeñas Empresas (MYPEs), para evaluar y participar de los procesos de selección para proveer bienes o

servicios a las entidades públicas. Esta información se convierte en una necesidad para promover el negocio y la función de ser una herramienta principal para entender las tendencias del mercado del sector público y entender su propia posición frente a la competencia. Por lo cual es fundamental la formulación de estrategias que permitan alcanzar los objetivos de una organización.

Ante la falta de herramientas de consolidación de información y análisis se pierden mucho tiempo generando reportes estáticos que muchas veces no contienen todas las variables que se quieren analizar o solo muestran datos actuales y no permiten enfocar datos históricos, ocasionando muchas veces, que estos reportes sean presentados fuera de tiempo, el cual no ayudan en la toma de decisiones de las MYPEs generando la pérdida de oportunidades.

En el entorno de los negocios de hoy, se requiere una aplicación más eficiente para tener siempre información disponible, transformada en medidas y formatos comunes, que la información esté indexada en el tiempo para poder analizar la información por periodos. Por consiguiente surge la necesidad de contar con herramientas que cumpla dichas expectativas, una solución como el Datamart para transformar los datos en conocimiento, y permitir que el proceso de toma de decisiones esté fundamentado sobre la información generada dentro y el entorno de una Empresa.

1.2. FORMULACIÓN DEL PROBLEMA.

Para el presente trabajo de investigación se formularon los siguientes interrogantes.

1.2.1. PROBLEMA GENERAL.

¿Cuál es el grado de influencia de la aplicación de Datamart de Contrataciones públicas a partir del SEACE en la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013?

1.2.2. PROBLEMAS ESPECÍFICOS.

PROBLEMA ESPECÍFICO 1.

¿En qué medida la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE mejora la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013?

PROBLEMA ESPECÍFICO 2.

¿En qué grado favorece la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013?

1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN.

Actualmente las Empresas están inmersas en entornos cada vez más competitivos, en los cuales es fundamental disponer de una información valiosa con la que se pueda generar un plan de acción, para situar el negocio por delante de sus competidores. Para este propósito, es indispensable disponer de

información que permita a los directivos de las Empresas tomar decisiones sobre como participar en un negocio.

La información es ahora la principal herramienta para entender las tendencias del mercado y entender su propia posición dentro del mismo con respecto a sus competidores. Con la aplicación del Datamart se aprovecha la información histórica para generar escenarios, pronósticos y reportes que apoyan al proceso de toma de decisiones.

La decisión de participar de un proceso de selección para proveer bienes o servicios a una entidad pública, es una tarea difícil para los MYPEs, y es cuando la información histórica comienza a tener valor para apoyar la decisión. El Datamart de Contrataciones públicas muestra la información sobre comportamientos de los precios en el tiempo, cantidades, los productos más adquiridos, proveedores(MYPEs) que venden más al estado, la información es presentada en forma numérica así como también mediante gráficas, el cual permite analizar y explorar la información en diferentes periodos, para realizar comparaciones de datos en diferentes periodos de tiempo que identifican los comportamientos y evoluciones, con el cual se confirma o se descubre tendencias. Y esta información mejora el proceso de toma de decisiones de la Micro y pequeñas Empresas (MYPEs) que participan o los que están interesados de las Contrataciones públicas.

El Datamart proporciona información exacta y a tiempo, para identificar y corregir situaciones, antes de que se conviertan en problemas y en potenciales pérdidas de oportunidades, una decisión fundamentada se traduce en un incremento de utilidades para una MYPE.

1.4. OBJETIVOS DE LA INVESTIGACIÓN.

Para el presente trabajo de investigación se plantearon los siguientes objetivos.

1.4.1. OBJETIVO GENERAL.

Determinar el grado de influencia de la aplicación de Datamart de Contrataciones públicas a partir del SEACE en la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno-2013.

1.4.2. OBJETIVOS ESPECÍFICOS.

OBJETIVO ESPECÍFICO 1.

Probar que la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE mejoran la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013.

OBJETIVO ESPECÍFICO 2.

Determinar que la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE favorecen el nivel de análisis de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

1.5. DELIMITACIÓN DE LA INVESTIGACIÓN.

La investigación se llevó específicamente con las Micro y pequeñas Empresas (MYPEs) que tienen participación e interesados en los procesos selección para proveer bienes y servicios a las entidades públicas, y la fuente del Datamart está compuesta por la información de contratación de bienes y servicios realizadas mediante procesos de selección en el Gobierno Regional de Puno sede central (REGIÓN PUNO), Municipalidad Provincial de Puno (MPP) y la Universidad Nacional del Altiplano de Puno (UNAP), durante el año 2013.

1.6. LIMITACIÓN DE LA INVESTIGACIÓN.

Para la puesta en funcionamiento de la aplicación del Datamart de contrataciones públicas fue necesario contar con MYPEs que hayan participado de las Contrataciones públicas, lo que implica que la muestra se reduce a MYPEs donde sus Gerentes o Administradores conozcan el procedimiento de las Contrataciones públicas en el Perú.



2.1. ANTECEDENTES DE LA INVESTIGACIÓN.

Para el presente trabajo de investigación se consideraron los siguientes antecedentes.

2.1.1. ANTECEDENTES NACIONALES.

Hermoza (2013) en su trabajo de investigación presenta y describe un modelo general OLAP y un prototipo de un Sistema Data Warehouse para una Empresa pública: Editora Perú S.A. La investigación inicia con la descripción del sistema de procesamiento de transacciones en línea (OLTP) implementado en la Empresa, y señala que la generación de reportes detallados, resumes y comparaciones se realizan mediante procedimientos tediosos y la información que se presenta es poco confiable.

El modelo propuesto tiene la finalidad de consolidar información para la toma de decisiones rápidas, mejorar el tiempo procesamiento y entrega de la información, diseñar un solo repositorio de información exclusivo para explotar la información histórica y proponer una herramienta de análisis que permite acceder a la información histórica.

Se concluye que el Datamart de ventas permite que el usuario pueda contar con una herramienta en línea totalmente automatizada, de fácil uso, y presenta la información necesaria en menor tiempo y muestra más opciones para su análisis de los datos.

Rodríguez y Mendoza (2011) en su trabajo de tesis describe los problemas que se presenta en el momento de la toma de decisiones por falta de información sólida en las áreas de ventas y compras de una Empresa comercializadora, y

plantean como solución la implementación de un Datamart de compras y un Datamart de ventas, herramientas de inteligencia de negocios.

Determinan que la Inteligencia de negocios, por medio del análisis de los datos, brinda información disponible y rápida, permitiendo detectar fallas en los procesos, descubriendo oportunidades de negocio y cuantificando relaciones con proveedores y clientes. Una solución que se basa en la extracción de datos de diversas fuentes, transformándolas y almacenándolas en un repositorio, desde el cual se genera información mediante reportes para los usuarios finales.

En el trabajo de tesis de **Mendoza (2011)**, se tiene como objetivo realizar el análisis del negocio en base a métricas o indicadores generados mediante los datos históricos que se mantienen a través de los sistemas transaccionales.

Señala que muchas Empresas manejan sus áreas de Planeamiento Estratégico, Análisis Táctico y Decisiones Operacionales de forma desintegrada, lo que dificulta la toma de decisiones del corte estratégico. Por consiguiente, muchas áreas como son Recursos Humanos, Contabilidad y Finanzas y Comercial no realizan una adecuada explotación de sus datos históricos de manera integrada que les permita obtener información estratégica, útil para la toma de decisiones.

Y concluye que la manera cómo las Empresas aprovechan la información histórica que se tiene en sus sistemas transaccionales determina una ventaja competitiva en el mercado, ya que permite realizar mejores decisiones en base a una información consistente y consolidada.

2.1.2. ANTECEDENTES INTERNACIONALES.

Gutiérrez (2012) en su tesis presenta como objetivo principal proporcionar a las PyMEs una metodología que les permita la integración de una solución de Inteligencia de Negocios que sea adecuada a su nivel de madurez, y que la implementación sea amigable con aceptación de los empleados de los nuevos cambios que se producen dentro de una Empresa. De esa forma poner alcance a las PyMEs este tipo de metodologías y por tanto este tipo de soluciones para que sean capaces de incrementar su productividad, lo que lleva a ser más competitivos.

En la investigación se desarrolla una metodología enfocada principalmente en conocer el negocio, para el cual se pretende desarrollar cierto conocimiento, para que permita guiar al dueño del negocio en establecer prioridades y realizar un desarrollo gradual que sea congruente con su nivel de madurez.

En el paper de **Yost (2004)** se presenta una solución frente a los problemas de interpretación del comportamiento de datos de una encuesta, donde la mejora de las técnicas estadísticas, y calidad de los datos se encuentran a menudo en un análisis cuidadoso de los datos históricos. En un tratamiento tradicional de los datos de un encuesta se tiene como resultado muchas preguntas que quedan sin respuesta, por falta de disponibilidad de datos, por acceso limitado a los datos, por que los metadatos no estaban bien definidas, o el rendimiento de las consultas era intolerablemente lentos.

En este paper describe las técnicas de modelado de bases de datos que permite a los usuarios finales tener un acceso rápido y fácil a las grandes cantidades de datos a nivel de detalle, datos sistematizados y en diferentes

marcos de tiempo . Además, también se analizan las técnicas para el seguimiento de los cambios de metadatos y estandarización. Se concluye que un modelo dimensional generalizada puede ser utilizada para cualquier censo o encuesta para poder realizar un seguimiento de la historia completa de la serie de datos.

El paper de Satyanarayana, **R., Rallabandi, S., Poorna, Ch., & Srikanth, R. (2010)**, muestran una visión general de Data Warehousing, Data Mining, OLAP, tecnologías OLTP, las aplicaciones y la arquitectura de Data Warehousing. Y señala que el Data Warehouse y el procesamiento analítico en línea (OLAP) son elementos esenciales para soporte de una decisión.

Se determinan que el Data Warehousing y OLAP han surgido como líderes en las tecnologías que facilitan el almacenamiento de datos dimensional, que ayuda a la organización con el soporte significativo de la toma de decisiones. Para ello los requisitos son muy diferentes sobre la tecnología de base de datos en comparación con las aplicaciones de procesamiento de transacciones en línea tradicionales.

Se concluye que un Data Warehouse puede decirse que es un almacén de datos semánticamente coherentes que sirve como una implementación física de un modelo de datos de soporte de decisiones que almacena información que necesita una Empresa para tomar decisiones estratégicas.

En el artículo de **Mohammed, M., Mohammed, A., Anwar, M. & Ahme, H., (2012)** presentan una aplicación E-Government Monitor (eGovMon) que garantiza los nuevos servicios en línea del Gobierno electrónico, que responden con eficacia las necesidades reales de los Ciudadanos, Empresas y Agencias

Gubernamentales. eGovMon utiliza muchas herramientas para ayudar a mejorar la calidad de los sitios web públicos.

Se muestra una propuesta para crear una arquitectura del Data Warehouse para eGovMon, con el que se busca mejorar los tratamientos de los datos mediante el uso de técnicas de Data Warehouse e inteligencia de negocios (BI), herramientas como ETL, DBMS, OLAP, DM y OLAM. Para hacer que la información sea más comprensible con buena calidad de los datos y mejorar el rendimiento del sistema de servicios del Gobierno electrónico.

Y se determinó que la creación de una arquitectura para eGovMon, con la herramienta y tecnología del Data Warehouse, muestran la descripción de los datos, también presenta altos detalles sobre sus datos. Y este almacén de información independiente a eGovMon es utilizado para minería de datos para el análisis información y predicción.

Gottlob, et al. (2010) en el paper señalan que el conocimiento de la evolución del mercado y las actividades de la competencia en el mercado es cada vez más un factor crítico de éxito para las Empresas. El Internet proporciona información de dominio público que puede ser recuperada por ejemplo de sitios web o tiendas en línea. La revisión o extracción de fuentes de información es sobre todo en forma manual, demanda mucho mucho tiempo. Este artículo describe cómo la información pública puede ser extraída de forma automática desde los sitios Web, transformado en formatos de datos estructurados, y se utiliza para el análisis de datos en los sistemas de Business Intelligence.

2.2. BASE TEÓRICA.

2.2.1. MYPES EN LA CIUDAD DE PUNO.

2.2.1.1. CONCEPTO DE MICRO Y PEQUEÑA EMPRESA.

Es una pequeña unidad de producción, comercio o prestación de servicios según Ley de Promoción y Formalización de la Micro y Pequeña Empresa, Ley N° 28015. Una MYPE es una unidad económica constituida por una persona natural o jurídica bajo cualquier forma de organización o gestión Empresarial, que tiene como objeto desarrollar actividades de extracción, transformación, producción, comercialización de bienes o prestación de servicios (PRODUCE, 2012).

2.2.1.2. MYPES SEGÚN SU ORGANIZACIÓN EMPRESARIAL.

Según Abad G. (2008) para el desarrollo de su actividad Empresarial, los MYPES adoptan una forma de organización, sea como persona natural o persona jurídica, según de la actividad económica, al capital, al número de personas que invierten en la Empresa, al mercado o público objetivo del nuevo negocio, etc. La Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT) recoge esta información y la clasifica por tipo de contribuyente, como sigue:

Persona Natural, es la denominación legal que el Código Civil otorga a los individuos capaces de adquirir derechos, deberes y obligaciones. Al constituir una Empresa como Persona Natural (con negocio), la persona asume a título personal todos los derechos y obligaciones de la Empresa, lo que implica que la persona asume la responsabilidad y garantiza con todo el patrimonio que posea

(los bienes que se encuentre a su nombre), las deudas u obligaciones que contrae la Empresa.

Persona **Jurídica**, tiene existencia legal pero no tiene existencia física y puede estar representada por una o más personas natural. Ley General de Sociedades, Ley N° 26887.

2.2.2. CONTRATACIONES PÚBLICAS.

Las Contrataciones públicas en el Perú se rigen por la Ley de Contrataciones del Estado, Ley N° 1017 que contiene las disposiciones y lineamientos que deben observar las Entidades del sector Público en los procesos de Contrataciones de bienes, servicios u obras y regula las obligaciones y derechos que se derivan de los mismos y su objeto es establecer las normas orientadas a maximizar el valor del dinero del contribuyente en las Contrataciones que realicen las Entidades del sector Público, de manera que éstas se efectúen en forma oportuna y bajo las mejores condiciones de precio y calidad, a través del cumplimiento de los principios señalados en su artículo N° 4.

2.2.2.1. PROCESOS DE SELECCIÓN.

Los procesos de selección son: Licitación pública, Concurso público, Adjudicación directa y Adjudicación de menor cuantía, los cuales se podrán realizar de manera corporativa o sujeto a las modalidades de selección de Subasta Inversa o Convenio Marco, de acuerdo a lo que defina el Reglamento. (Ley N° 1017, Art N° 15., 2008)

a) Licitación pública y concurso público.

La Licitación pública se convoca para la contratación de bienes, suministros y obras. El concurso público se convoca para la contratación de servicios de toda naturaleza. En ambos casos, se aplican los márgenes que establece la Ley de Presupuesto del Sector Público. (Ley N° 1017, Art N° 16., 2008)

b) Adjudicación directa.

La Adjudicación directa se aplica para las Contrataciones que realice la Entidad, dentro de los márgenes que establece la Ley de Presupuesto del Sector Público. La adjudicación directa puede ser pública o selectiva. El Reglamento señalará la forma, requisitos y procedimiento en cada caso. (Ley N° 1017, Art N° 17., 2008)

c) Adjudicación de menor cuantía.

La Adjudicación de menor cuantía se aplica a las Contrataciones que realice la Entidad, cuyo monto sea inferior a la décima parte del límite mínimo establecido por la Ley de Presupuesto del Sector Público para los casos de licitación pública y concurso público. (Ley N° 1017, Art N° 18., 2008)

2.2.3. TOMA DE DECISIONES.

Continuamente, las personas deben elegir entre varias opciones aquella que consideran más conveniente. Es decir, han de tomar gran cantidad de decisiones en su vida cotidiana, en mayor o menor grado importantes, a la vez que fáciles o difíciles de adoptar en función de las consecuencias o resultados derivados de cada una de ellas.

Es posible trasladar este planteamiento general al ámbito de la Empresa. La toma de decisiones abarca a las cuatro funciones administrativas, así los administradores cuando planean, organizan, conducen y controlan, se les denomina con frecuencia los que toman las decisiones. (Córdoba, 2004)

Ahora bien: ¿qué se entiende por decidir? Schackle (1966) define la decisión como un corte entre el pasado y el futuro. Otros autores definen la decisión como la elección entre varias alternativas posibles, teniendo en cuenta la limitación de recursos y con el ánimo de conseguir algún resultado deseado.

Como tomar una decisión supone escoger la mejor alternativa de entre las posibles, se necesita información sobre cada una de estas alternativas y sus consecuencias respecto a nuestro objetivo. La importancia de la información en la toma de decisiones queda patente en la definición de decisión propuesta por Forrester (1994), entendiendo por esta "el proceso de transformación de la información en acción". La información es la materia prima, el input de la decisión, y una vez tratada adecuadamente dentro del proceso de la toma de decisión se obtiene como output la acción a ejecutar. La realización de la acción elegida genera nueva información que se integrará a la información existente para servir de base a una nueva decisión origen de una nueva acción y así sucesivamente. Todo ello debido a una de las características de los sistemas cibernéticos que es la retroalimentación o Feed-back.

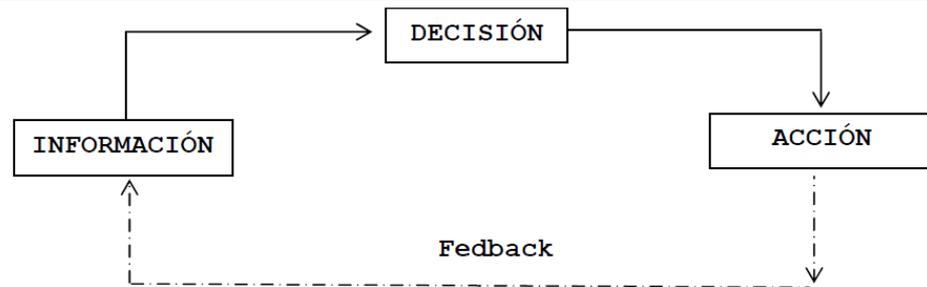


Figura N° 1: Transformación de la información en acción.
Fuente: Menguzzato & Renau (1991).

Le Moigne (1990) define el término decidir como identificar y resolver los problemas que se le presenta a toda organización. Por tanto, el desencadenante del proceso de toma de decisiones es la existencia de un problema, pero ¿cuándo existe un problema? Para Huber (1992) existirá un problema cuando hay diferencia entre la situación real y la situación deseada. La solución del problema puede consistir en modificar una u otra situación, por ello se puede definir como el proceso consciente de reducir la diferencia entre ambas situaciones.

2.2.3.1. TIPOS DE DECISIONES.

Las diferencias en la toma de decisiones se pueden clasificar según su nivel en la organización, que comprende a Gerencia de nivel superior, Gerencia de nivel medio y Gerencia operacional. (Laudon & Laudon, 2012). Dentro de cada uno de los niveles de toma de decisiones se clasifican en decisiones estructuradas, decisiones no estructuradas y semiestructuradas.

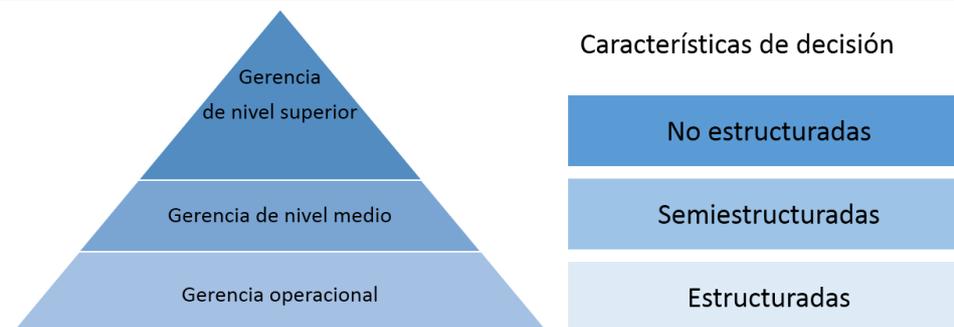


Figura N° 2: Tipos y características de decisión.
Fuente: Laudon & Laudon (2012).

- **Las decisiones no estructuradas:** Son aquellas en las que el encargado de tomarlas debe aportar juicio, evaluación y entendimiento de la definición del problema.
- **Las decisiones estructuradas:** Son repetitivas y rutinarias, e implican un procedimiento definido para manejarlas, así que no se tienen que tratar en cada ocasión como si fueran nuevas.
- **Las decisiones semiestructuradas:** Solo una parte del problema tiene una respuesta bien definida provista por un procedimiento aceptado.

2.2.3.2. EL PROCESO DE TOMA DE DECISIONES.

Tomar una decisión es un proceso que consta de varios pasos. Simon (1960) describió cuatro distintas etapas en la toma de decisiones: inteligencia, diseño, elección e implementación (vea la figura N° 3).

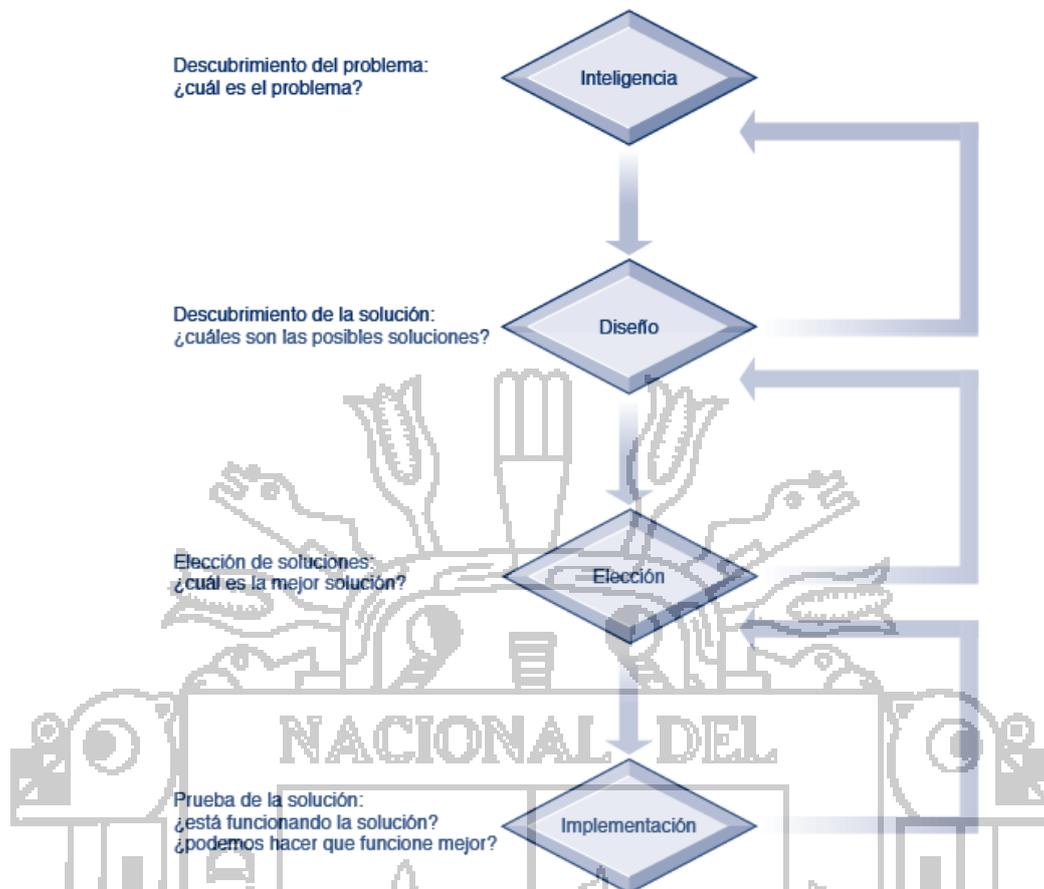


Figura N° 3: Etapas en la Toma de decisiones.
Fuente: Laudon & Laudon (2012).

- La inteligencia consiste en descubrir, identificar y comprender los problemas que ocurren en la organización: por qué existe un problema, en dónde y qué efectos tiene sobre la firma.
- El diseño implica identificar y explorar varias soluciones para el problema.
- La elección consiste en elegir una de varias alternativas de solución.
- La implementación implica hacer que funcione la alternativa elegida y continuar monitoreando qué tan bien funciona esa solución.

2.2.4. DATAMART.

Antes de definir un Datamart, conceptualizamos el Data Warehouse.

2.2.4.1. DEFINICIÓN DE DATA WAREHOUSE.

“Un Data Warehouse es un conjunto de datos orientado a temas, integrado, no volátil, estable y que se usa para el proceso de toma de decisiones”. Un DataMart es un subconjunto sectorial del DW a menudo perteneciente a un departamento concreto. (Gestión de Proyectos, 2013).

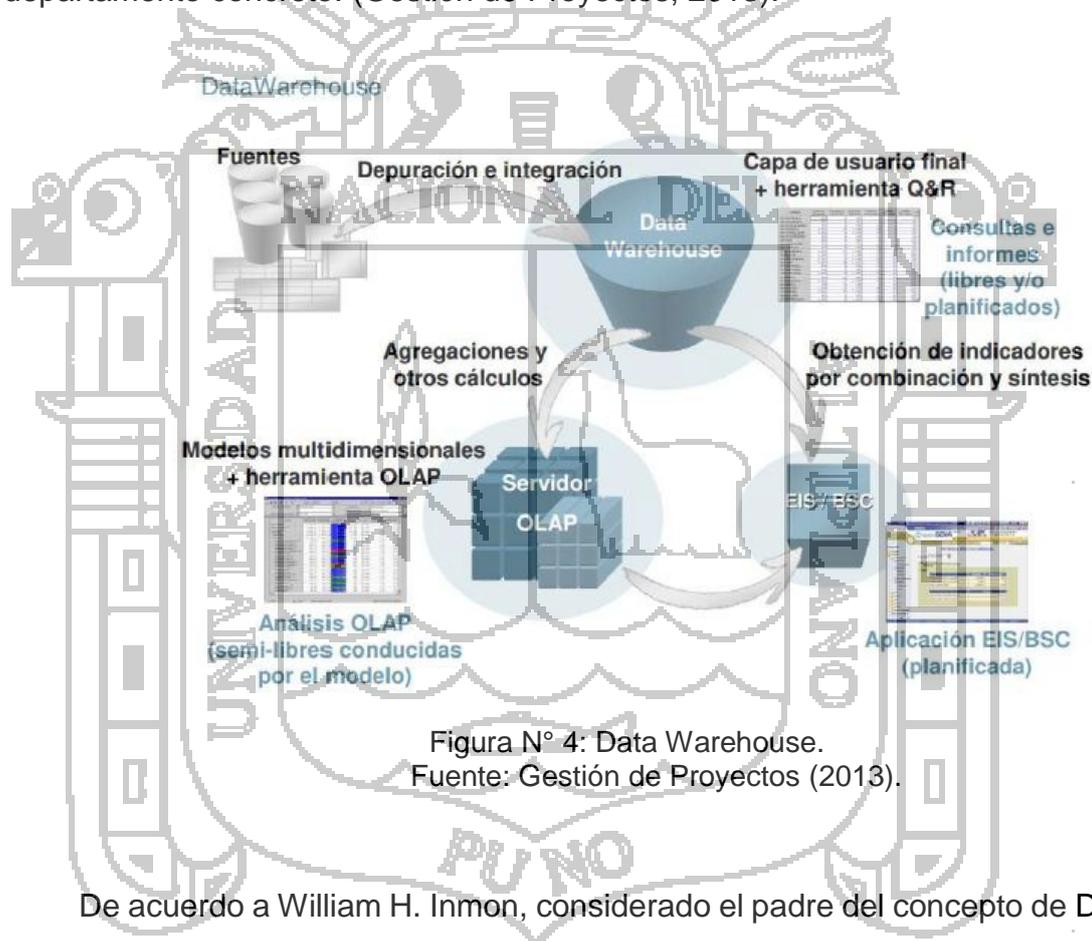


Figura N° 4: Data Warehouse.
Fuente: Gestión de Proyectos (2013).

De acuerdo a William H. Inmon, considerado el padre del concepto de Data Warehouse (almacén de datos) en los años ochenta, Data Warehouse es "un conjunto de datos integrados orientados a una materia, que varían con el tiempo y que no son transitorios, los cuales soportan el proceso de toma de decisiones de una administración". (Harjinder & Prakash, 1996)

El Data Warehouse es un almacén de datos relacionados a las actividades de una organización y grabadas en una base de datos diseñada específicamente con el propósito de hacer informes para después analizar estos informes y conseguir información estratégica. (Harjinder S., 1996)

Con las definiciones anteriores podemos mencionar que un Data Warehouse proporciona una visión global, común e integrada de los datos de la organización, independiente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios, con las propiedades siguientes: estable, coherente, fiable y con información histórica.

2.2.4.2. DEFINICIÓN DE DATAMART.

Un Datamart es un subconjunto de los datos del Data Warehouse cuyo propósito es el de responder a un determinado análisis, función o necesidad y con una población de usuarios específica. Al igual que en un Data Warehouse, los datos están estructurados en modelos de estrella o copo de nieve y un Datamart puede ser dependiente o independiente de un Data Warehouse. (Reinhardt, 2007).

Un Datamart según Lane (1999) es una forma más sencilla de un Data Warehouse que está enfocado a una sola área funcional tales como ventas, finanzas o mercadeo. Debido a que se centra únicamente en una sola área, los Datamart se constituyen de menor cantidad de fuentes de datos que los Data Warehouse, las cuales pueden ser sistemas operacionales internos o un Data Warehouse interno o externo.

Para Inmon (2002), lo más importante en la definición de un Datamart, constituye que el departamento de la organización propietario del mismo posea

el hardware, el software y datos que lo constituyen. Al poseer los derechos de propiedad de Datamart el departamento tiene el control y disciplina de los datos encontrados en el mismo.

Los Datamarts por las consideraciones mencionadas, vienen a ser una excelente herramienta de análisis de datos y soporte para la toma de decisiones para las pequeñas y medianas Empresas, por su versatilidad, desarrollo en corto tiempo y de bajo costo económico, así como la obtención de los resultados esperados a un corto plazo.

2.2.4.3. CARACTERÍSTICAS DEL DATAMART.

Inmon (2002) plantea que un Datamart debe presentar las siguientes características:

- Se centra únicamente en un área de la Empresa.
- El volumen de datos que maneja es inferior al de un Data Warehouse.
- Son empleados en su mayoría como soporte para la toma de decisiones.

2.2.4.4. DIFERENCIAS ENTRE UN DATAWAREHOUSE Y DATAMART

Las diferencias entre un Data Warehouse y un Datamart se pueden resumir en las siguientes:

	DATA WAREHOUSE	DATAMART
Alcance	Construido para satisfacer las necesidades de información de toda la organización	Construido para satisfacer las necesidades de un área de negocios específica.
Objetivos	Diseñado para optimizar la integración y la administración de los datos fuente	Diseñado para optimizar la entrega de información de soporte a decisiones
Características de los Datos	Administra grandes cantidades de datos históricos a nivel de detalle	Administra grandes cantidades de datos históricos a nivel atómico
Pertenencia	Pertenece a toda la organización	Pertenece al área de negocio al cual está orientado
Administración	Es administrado por la unidad de sistema de la organización	Es administrado por el personal de sistema de la unidad propietaria del Datamart

Tabla N° 1: Diferencia entre Data Warehouse y Datamart.
Fuente: Trak y Zambrano (2002).

2.2.5. ARQUITECTURA DEL DATAMART.

2.2.5.1. DATOS OPERACIONALES.

OLTP(On Line Transaction Processing), representa toda aquella información transaccional que genera una organización en su accionar diario, además, de las fuentes externas con las que puede llegar a disponer. Y este es el origen o fuente de datos para poblar el componente de almacenamiento físico Datamart.

Según Bernabeu (2007), entre los OLTP más habituales que pueden existir en cualquier organización se encuentran

- Archivos de textos.
- Hipertextos.
- Hojas de cálculos.
- Informes semanales, mensuales, anuales, etc.
- Bases de datos transaccionales.

2.2.5.2. PROCESO ETL (Extracción, Transformación y Carga).

Para poder extraer los datos desde los OLTP, para luego manipularlos, integrarlos y transformarlos, para posteriormente cargar los resultados obtenidos en el Datamart, es necesario contar con algún sistema que se encargue de ello. Precisamente, los ETL son los que cumplen con tal fin. (Bernabeu, 2007)

Los procesos ETL no se limitan a tomar datos de un lugar y colocarlos en otro; para Palomar (2001) son capaces de hacer otras tareas:

- Reunir datos de sistemas operacionales separados.
- Convertir los datos de un formato a otro.
- Modificar los datos para que sean más comprensibles o más completos.
- Cargar los datos modificados en el Data Warehouse o Datamart destino.

a) Extracción.

La extracción es la operación que tiene como meta obtener los datos desde las fuentes de información. Esta operación se puede llevar a cabo mediante la extracción de los datos desde un archivo que los contenga o directamente mediante una conexión de red desde el sistema que los contiene. Los datos

obtenidos de esta etapa serán posteriormente transformados y cargados en el almacén de datos, ya sea un Data Warehouse o un Datamart.

Según Palomar (2001), las fuentes de datos que normalmente suministran los insumos para esta operación son: Base de datos operacionales o Base de datos externas, Archivos en texto en plano, Hojas de cálculo o Datamart existente.

b) Transformación.

Esta función es la encargada de convertir aquellos datos inconsistentes en un conjunto de datos compatibles y congruentes, para que puedan ser cargados en el Datamart. Estas acciones se llevan a cabo, debido a que pueden existir diferentes fuentes de información, y es vital conciliar un formato y forma única, definiendo estándares, para que todos los datos que ingresarán al Datamart estén integrados. (Bernabeu, 2007).

Los casos más comunes que se realizan en la integración, son los siguientes:

- Codificación.
- Medida de atributos.
- Convenciones de nombramiento.
- Fuentes múltiples.
- Limpieza de datos.

1) Codificación.

Una inconsistencia muy típica que se encuentra al intentar integrar varias fuentes de datos, es la de contar con más de una forma de codificar un atributo

en común. Lo que se debe realizar en estos casos, es seleccionar o recodificar estos atributos, para que cuando la información llegue al Datamart, esté integrada de manera uniforme. (Bustos y Mosquera, 2013)

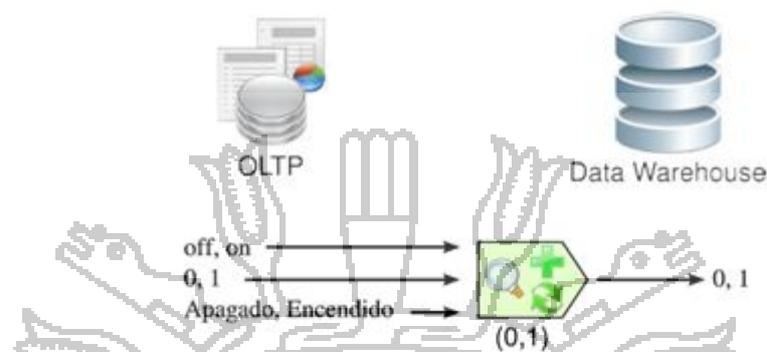


Figura N° 5: Transformación: Codificación.
Fuente: Bernabeu (2007).

2) Medida de Atributos.

Los tipos de unidades de medidas utilizados para representar los atributos de una entidad, varían considerablemente entre sí, a través de los diferentes OLTP. En esta ocasión, se deberán estandarizar las unidades de medidas de los atributos, para que todas las fuentes de datos expresen sus valores de igual manera. Los algoritmos que resuelven estas inconsistencias son generalmente los más complejos.

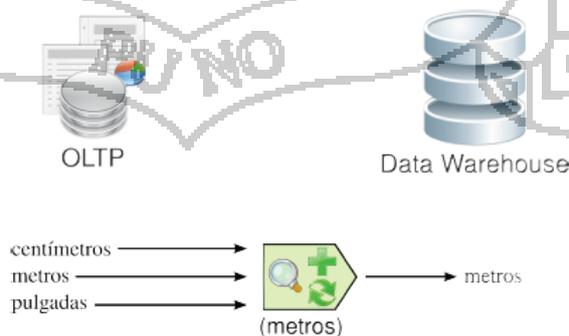


Figura N° 6: Transformación: Medida de Atributos.
Fuente: Bernabeu (2007).

3) Convenciones de nombramiento.

Usualmente, un mismo atributo es nombrado de diversas maneras en los diferentes OLTP. Por ejemplo, al referirse al nombre del proveedor, puede hacerse como “nombre”, “razón_social”, “proveedor”, etc. Aquí, se debe utilizar la convención de nombramiento que para los usuarios sea más comprensible.

4) Fuentes múltiples.

Un mismo elemento puede derivarse desde varias fuentes. En este caso, se debe elegir aquella fuente que se considere más fiable y apropiada.

5) Limpieza de datos.

Su objetivo principal es el de realizar distintos tipos de acciones contra el mayor número de datos erróneos, inconsistentes e irrelevantes.

c) Carga.

Este proceso es el responsable de cargar la estructura del Datamart con:

- Aquellos datos que han sido transformados y que residen en un almacenamiento intermedio.
- Aquellos datos de los OLTP que tienen correspondencia con el depósito de datos.

Se debe tener en cuenta, que los datos antes de moverse al Datamart, deben ser analizados con el propósito de asegurar su calidad, ya que es un factor clave, que debe dejarse de lado.

2.2.5.2. BASE DE DATOS MULTIDIMENSIONAL.

Una base de datos multidimensional es una base de datos en donde su información se almacena en forma multidimensional, es decir, a través de tablas de hechos y tablas de dimensiones.

Para Bernabeu (2010) las bases de datos multidimensionales implican tres variantes posibles de modelamiento, que permiten realizar consulta de soporte de decisión:

- Esquema en estrella: Consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a estas.
- Esquema copo de Nieve: Este esquema representa una extensión del modelo en estrella cuando las tablas de dimensiones se organizan en jerarquías de dimensiones.
- Esquema constelación: Este modelo está compuesto por una serie de esquemas en estrella.

a) Tabla de dimensiones.

Una dimensión es una estructura que clasifica los datos en el orden que los usuarios establecen mediante las consultas Empresariales. Algunas de las dimensiones más utilizadas son Cliente, Producto, Región Geográfica y Tiempo. Para identificar las dimensiones es necesario revisar las tablas, especialmente las llamadas tablas de hechos. El conjunto de dimensiones define lo que se denomina Cubo de Dimensiones el cual se puede apreciar en la figura:

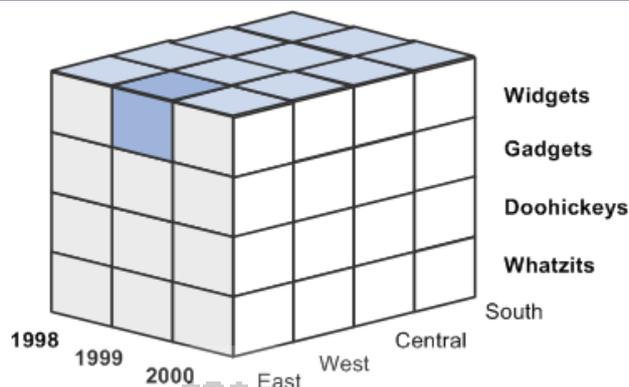


Figura N° 7: Cubo Multidimensional.
Fuente: Oracle Data Warehousing Guide.

b) Tabla de hechos.

La tabla de hechos, también llamada de detalles, es el núcleo central en un esquema de modelado de estrella y de copo de nieve y se caracteriza por poseer una clave compuesta, la cual se compone de las tablas de dimensiones. Un ejemplo de la tabla de hechos son los detalles de las ventas y los detalles de los consumos de un producto. (Rodríguez. y Mendoza, 2011)

2.2.6. MODELAMIENTO DE UN DATAMART.

2.2.6.1. ESQUEMA ESTRELLA.

Se tiene una sola tabla de hechos (tabla fact), la cual tiene datos para realizar el análisis rodeada de tablas dimensiones, esta tabla se envuelve de tablas más pequeñas, se asemeja a una estrella, es importante destacar que la tabla de dimensiones tiene una clave primaria simple a diferencia de la tabla de hechos, en donde la clave principal está compuesta por claves primarias de las tablas dimensionales (Bustos y Mosquera, 2013).

La principal ventaja de este esquema es su velocidad y simplicidad para realizar análisis multidimensionales, además para el usuario final las consultas no son complicadas. El esquema estrella por su rendimiento permite indexar dimensiones de manera individual sin que afecte al rendimiento de la base de datos.

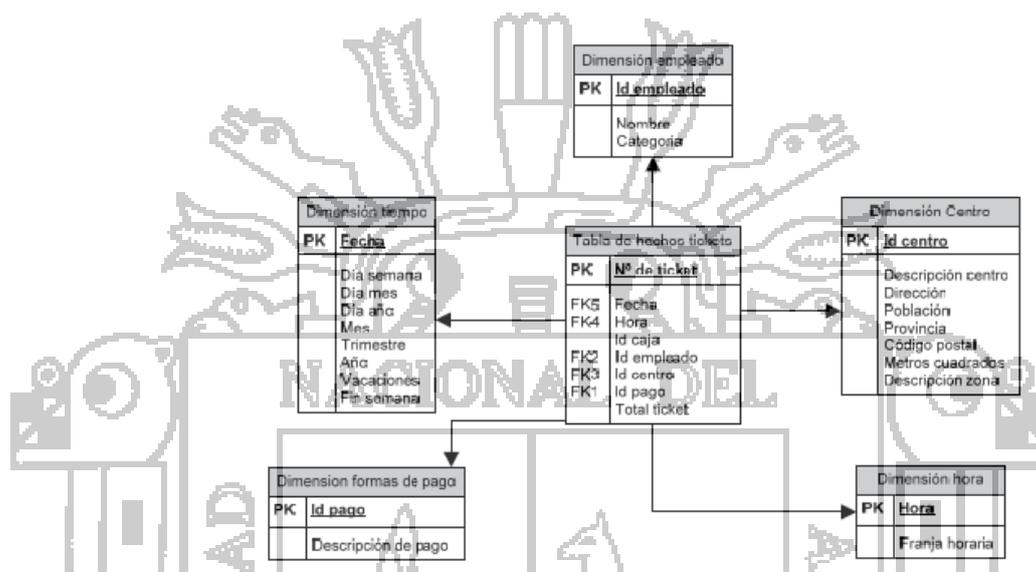


Figura N° 8: Ejemplo de esquema estrella.
Fuente: Lluís (2007).

2.2.7. METODOLOGÍA DE DESARROLLO DE DATAMART.

Existen varias tecnologías, utilizadas para el desarrollo e implementación de Soluciones de Datamart, entre las que encontramos:

Metodología desarrollada por Bill Inmon en el año 1992 en su libro "Building the Data Warehouse". Para esta metodología un Data Warehouse es considerado como un conjunto de datos orientados por temas, integrados, variantes en el tiempo y no volátiles, cuyo objetivo es ser apoyo importante para la toma de decisiones estratégicas. El Data Warehouse es una parte de un sistema de BI y los Data Marts obtienen su información a partir de este Data

Warehouse y para ser almacenada la información es necesario normalizarla antes de almacenarse.

En la Metodología Kimball el Data Warehouse se forma del conjunto de todos los Datamarts generados existentes en una Empresa. La información siempre se almacena en un modelo dimensional. Otra forma de denominar ésta aproximación es como "Bottom-up".

Metodología HEFESTO es una metodología propia para la construcción e implementación de un DW y se puede adaptar a cualquier ciclo de vida de desarrollo de software. Esta metodología permite realizar una primera implementación que satisface una parte de las necesidades, para demostrar sus ventajas y motivar a los usuarios. (Bernabeu, 2010)

2.2.8. METODOLOGÍA HEFESTO.

Se ha seleccionado la Metodología HEFESTO, debido a que las fases que esta comprende, se acopla al desarrollo y está enfocada al análisis de requerimientos, identificación de identificadores y perspectivas, que permitirán realizar el análisis de datos respectivos y obtener resultados que apoyen a la toma de definiciones de la organización.

2.2.8.1. CARACTERÍSTICAS.

- Los objetivos y resultados finales de cada fase se diferencian fácilmente y son sencillos de comprender.
- El usuario se encuentra involucrado estrechamente al final en cada etapa, para que tome decisiones respecto al comportamiento y funciones del Datamart.

- Utiliza modelos conceptuales y lógicos, sencillos de interpretar y analizar.
- Es independiente de las herramientas que se utilicen para su implementación.
- Es independiente de las estructuras físicas que contengan el Datmart y de su respectiva distribución.
- Cuando se finaliza con una fase o paso, los resultados obtenidos al terminar, sirven como punto de partida para llevar a cabo el siguiente paso.
- Se puede aplicar tanto para la implementación de Datamart como para Data Warehouse.
- Está basada en los requerimientos del usuario, lo que permite adaptarse con facilidad y rapidez ante los cambios en el negocio.
- La metodología HEFESTO permite la realización de Data Warehouse o un Datamart al mismo tiempo.

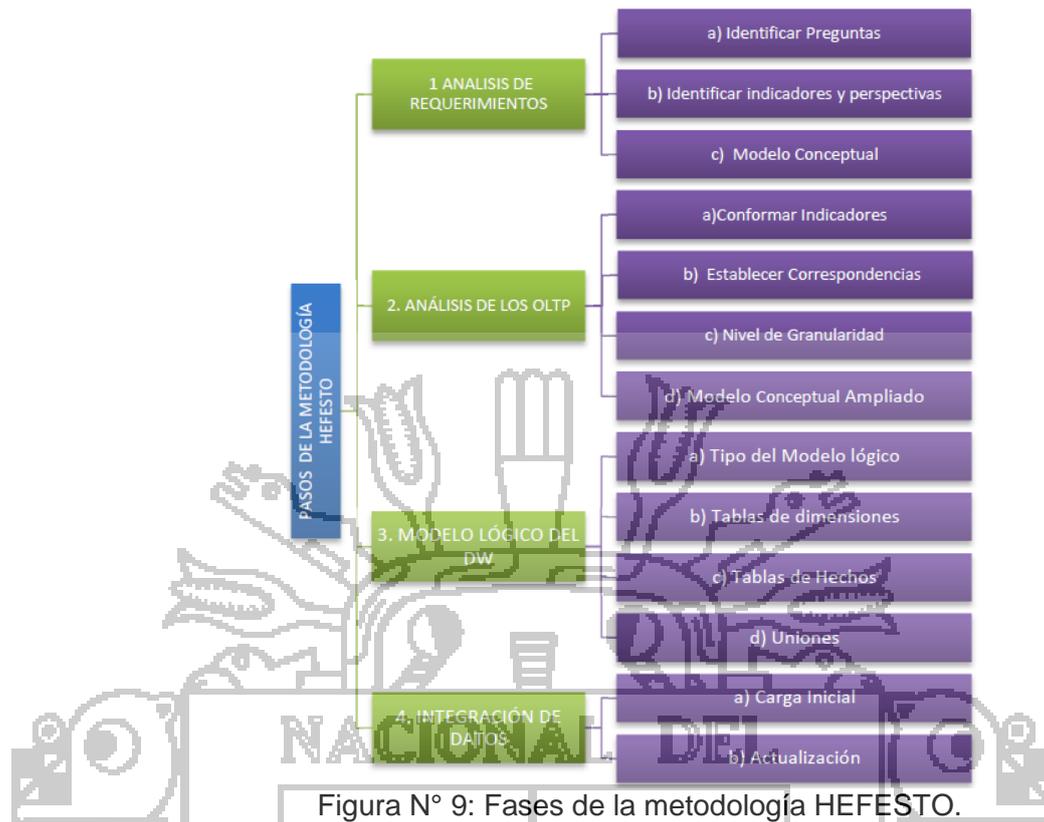


Figura N° 9: Fases de la metodología HEFESTO.

Fuente: Bernabeu (2010).

2.2.8.2. FASES DE LA METODOLOGÍA HEFESTO.

1) ANÁLISIS DE REQUERIMIENTOS.

Inicialmente se debe identificar los requerimientos de los usuarios, esta información se obtendrá mediante la elaboración de preguntas, que permitan conocer los objetivos de la organización. Para después de un análisis, poder determinar cuáles serán los indicadores y las perspectivas que se utilizarán que para la construcción del Data Warehouse o Datamart. (Bernabeu, 2010)

Al final se realizará un modelo conceptual que consolide la información de los indicadores, perspectivas, información que permitirá tener una visualización global del análisis inicial.

a) Identificar Preguntas.

Por la necesidad de adquirir información sobre los requerimientos de los usuarios, es posible aplicar diferentes técnicas, tales como entrevistas, cuestionarios, observaciones, entre otros.

Como punto de partida, se realiza el análisis de los requerimientos de los diferentes usuarios, ya que de ello depende, que el desarrollo muestre lo que se espera del depósito de datos, en relación a sus funciones y cualidades.

Principalmente se busca obtener e identificar las necesidades de información clave de alto nivel que proveerá el soporte para desarrollar las siguientes etapas, para cumplir con las metas y estrategias de la Empresa lo que apoyará en los procesos de toma de decisiones.

Un análisis adecuado, permitirá conocer claramente los objetivos estratégicos planteados por la Empresa, y sus principales procesos.

Para realizar el estudio y análisis de la información, se debe formular preguntas complejas sobre el negocio, que contemplen variables de análisis que se consideren relevantes, y permitan obtener fuentes de información notables e importantes.

b) Identificar indicadores y perspectivas de análisis.

Al referirse a Indicadores, se engloba todos aquellos valores numéricos que representan lo que se desea analizar concretamente (saldos, promedios, cantidades, sumatorias, fórmulas) mientras que las Perspectivas son aquellos objetos de los que se analizará los indicadores (clientes, productos, proveedores).

c) Modelo Conceptual Ampliado.

En esta etapa, se construye un modelo conceptual a partir de los indicadores y perspectivas obtenidas del análisis anterior para poder determinar los alcances del proyecto sobre los cuales se va a trabajar y que puede ser presentado a los usuarios. El modelo conceptual muestra una vista global sobre los resultados finales, las variables que se utilizarán para analizarlos y cuál es la relación que existe entre ellos.

Representación gráfica del modelo conceptual:

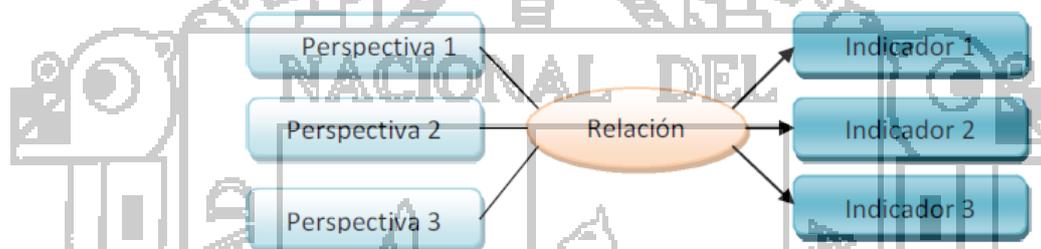


Figura N° 10: Modelo conceptual.
Fuente: Bernabeu (2010).

2) ANÁLISIS DE LOS OLTP.

En esta fase se efectuara un análisis de las fuentes OLTP para determinar cómo serán calculados los indicadores y así establecer las respectivas correspondencias que existen entre el modelo conceptual creado y las fuentes de datos que se tiene.

Después se definirá los campos que deben estar incluidos en las perspectivas, esta información estará contenida en un modelo conceptual ampliado.

a) Conformar Indicadores.

Se explica cuál será la forma de cálculo para obtener los indicadores definiendo de cada uno:

- Hecho/s que lo componen, con su respectiva fórmula de cálculo.
- Función de sumarización que se utilizará para su agregación.

b) Establecer correspondencias con los requerimientos.

En esta etapa, se examinan los OLTP con los que se va a trabajar, que contienen características e información necesaria, para definir las correspondencias existentes entre el modelo conceptual y las fuentes de datos.

De los indicadores, se deberá definir cuál será la forma de cálculo, sea mediante fórmulas u operaciones complejas.

c) Nivel de Granularidad.

Se seleccionan los campos que integrarán las perspectivas encontradas. Cuando se ha determinado las relaciones con los OLTP se procederá a definir los campos que conformarán cada perspectiva, por los que se manipularán y filtrarán a los indicadores.

Con las correspondencias establecidas, se mostrará al usuario los datos de análisis para las perspectivas. Se debe tener claro el significado de los campos o valores de los OLTP, para lo cual se debe programar reuniones con el personal encargado del sistema, para el análisis de los datos.

Se debe presentar al o los usuarios, los datos existentes y el significado de los mismos, con sus posibles valores y características para que los usuarios

puedan determinar qué datos tienen mayor importancia para obtener indicadores.

d) Modelo Conceptual Ampliado.

Al modelo Conceptual definido en pasos anteriores, se lo ampliará añadiendo los campos y atributos en las perspectivas, y se colocará bajo cada indicador, su fórmula de cálculo respectiva.

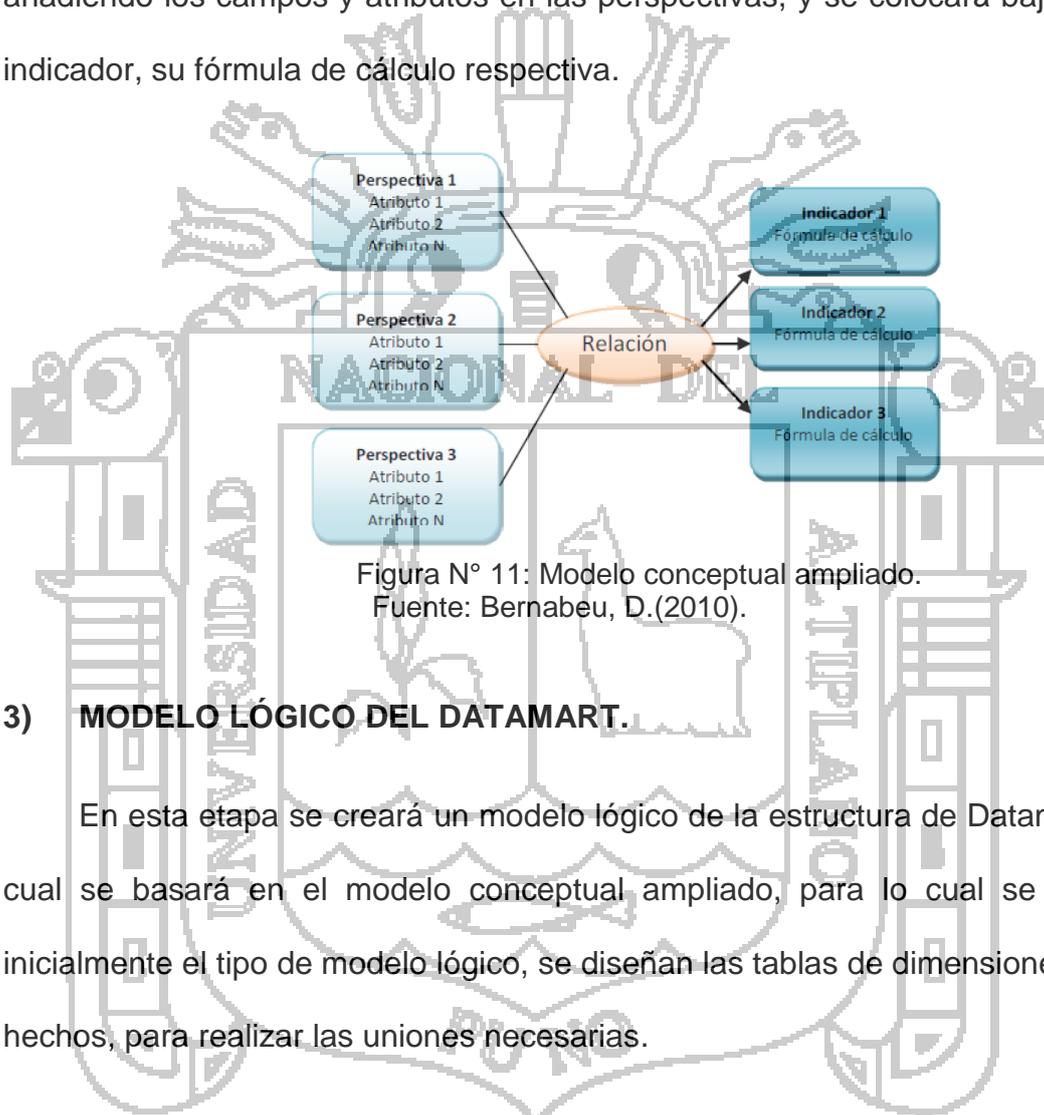


Figura N° 11: Modelo conceptual ampliado.
Fuente: Bernabeu, D.(2010).

3) MODELO LÓGICO DEL DATAMART.

En esta etapa se creará un modelo lógico de la estructura de Datamart, el cual se basará en el modelo conceptual ampliado, para lo cual se define inicialmente el tipo de modelo lógico, se diseñan las tablas de dimensiones y de hechos, para realizar las uniones necesarias.

a) Tipo del modelo lógico.

Se debe definir el esquema (estrella, constelación o copo de nieve) que se ajuste estrechamente a las necesidades y requerimientos de los usuarios.

b) Tablas de dimensiones.

Se debe crear las dimensiones del esquema lógico, tomando cada perspectiva con sus atributos correspondientes para:

- Elegir un nombre que identifique la dimensión.
- Añadir un campo que represente su clave principal.
- Redefinir nombres de los atributos si no son explicativos.

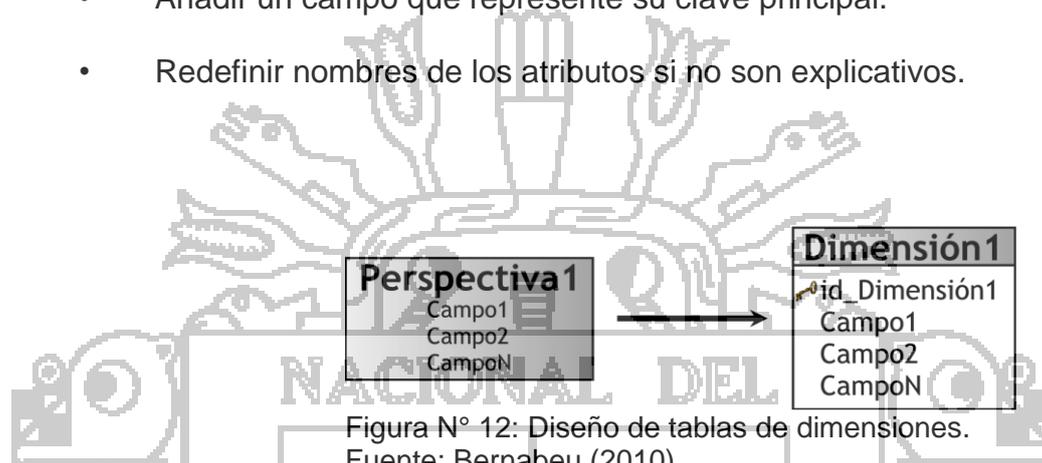


Figura N° 12: Diseño de tablas de dimensiones.
Fuente: Bernabeu (2010).

c) Tabla de hechos.

Las tablas de hechos contienen, los hechos que serán utilizados por los analistas de negocio para apoyar el proceso de toma de decisiones. Incluyen datos cuantitativos de los cuales se construirán los indicadores de estudio.

Esquemas en Estrella:

- Asignar un nombre a la tabla de hechos que indique la información analizada, área de estudio, negocio enfocado.
- Definir una clave primaria, compuesta por la combinación de las claves primarias de cada tabla de dimensión relacionada.
- Crear campos de hechos e indicadores, de acuerdo a los definidos en el modelo conceptual.

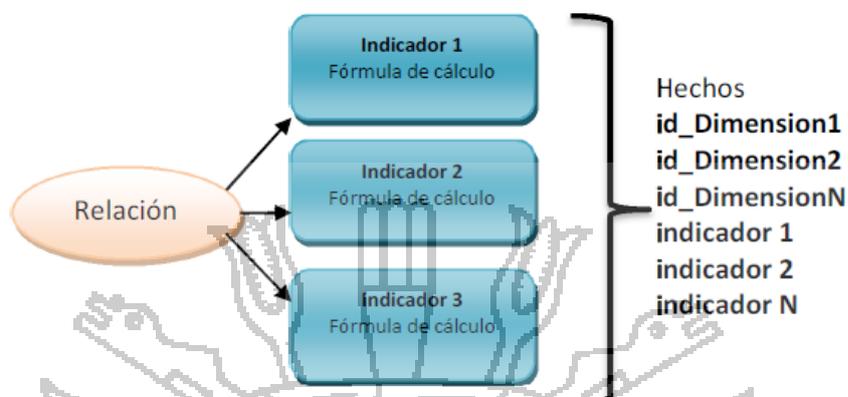
Representación Gráfica:

Figura N° 13: Tabla de hechos (Esquema estrella).
Fuente: Bustos y Mosquera (2013).

d) Uniones.

Las uniones correspondientes entre las tablas de dimensiones y tablas de hechos, se realizan para los tres tipos de esquemas.

4) INTEGRACIÓN DE DATOS.

El modelo lógico que se ha construido debe ser puesto a prueba con datos, utilizando técnicas calidad de datos como ETL, para definir reglas y normas para actualización.

a) Carga Inicial.

Se realiza una carga Inicial del Datamart, alimentando el modelo creado. Se debe llevar a cabo tareas de limpieza de datos, calidad de datos, procesos ETL, que tienen una lógica compleja.

En el Datamart debe contener datos válidos, quitando valores faltantes o anómalos, para así utilizar únicamente los datos útiles.

b) Actualización.

La actualización se la realizada teniendo la carga de datos en el Datmart:

- Establecimiento de políticas y estrategias ejecutar actualizaciones o refresco de datos.
- Especificar tareas de limpieza de datos, calidad de datos, procesos ETL, para actualizar los datos del Datamart.
- Especificar de forma general y detallada las acciones que deberá realizar cada software.

2.2.9. BASE DE DATOS.**2.2.9.1. MYSQL.**

Es un sistema de administración de base de datos relacional (RDBMS) se trata de un programa capaz de almacenar una enorme cantidad de datos de gran variedad y de distribución para cubrir la necesidad de cualquier tipo de organización, desde pequeños establecimientos comerciales a grandes Empresas y organismos administrativos, MySQL compite con sistemas RDBMS propietarios conocidos como Oracle, SQL Server y DB2. (Sabana, 2006)

2.2.10. PLATAFORMA PENTAHO.

Pentaho es una herramienta de Business Intelligence (BI) desarrollada bajo la filosofía del software libre para la gestión y toma de decisiones Empresariales. Es una plataforma compuesta de diferentes programas que satisfacen los requisitos de BI. Ofreciendo soluciones para la gestión y análisis de la

información, incluyendo el análisis multidimensional OLAP, presentación de informes, minería de datos y creación de cuadros de mando para el usuario (Gravitar, 2014).

En la siguiente ilustración se muestra la arquitectura funcional de Pentaho:

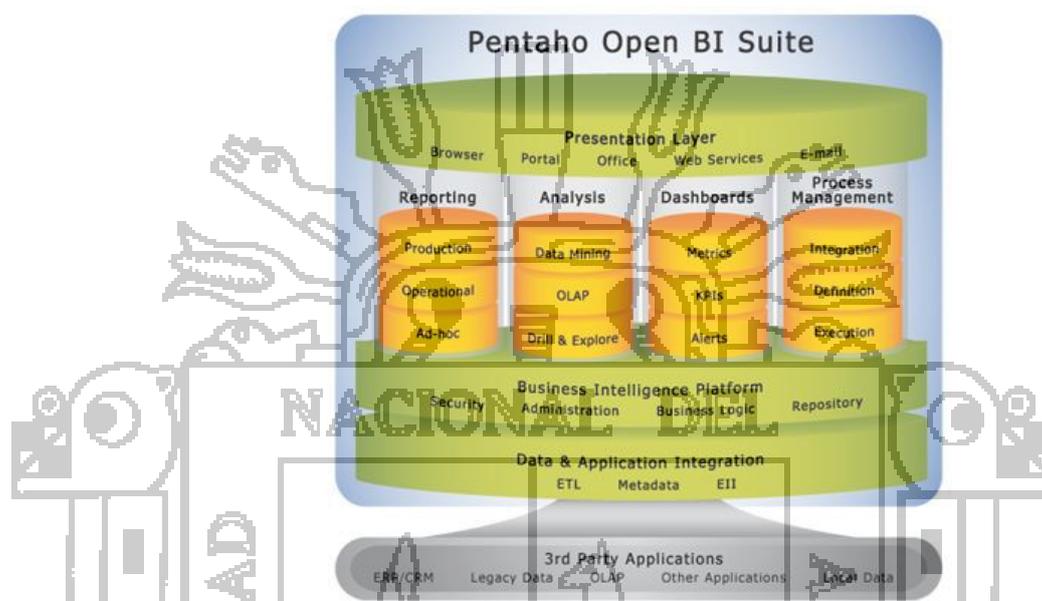


Figura N° 14: Arquitectura funcional de Pentaho.
Fuente: Pentaho, Documentation (2014).

En resumen la Plataforma BI integra componentes Open Source mostrando una combinación de flujos de trabajo y administración de procesos.

Para la integración de datos existe dos herramientas muy conocidas en el mundo de open source los cuales son Pentaho Data Integration (Kettle) y Talend Open Studio, para el presente investigación se decidió utilizar Talend Open Studio por la curva de aprendizaje menor y por la documentación existente.

2.2.10.1. HERRAMIENTA ETL: TALEND OPEN STUDIO.

Talend Open Studio es una herramienta Open Source que permite modelar transformaciones de datos de forma visual, básicamente es una herramienta ETL (Extract Transform and Load – Extracción Transformación y Carga). Entre los

usos que se le puede dar a esta herramienta es muy útil para realizar migraciones de datos, crear tablas de dimensiones para explorarlos con cubos OLAP, o sincronizar aplicaciones a nivel de base de datos. En esta serie de post vamos a ver cómo usarlo para migrar datos de una aplicación a otra. (Moya, 2010)

2.2.10.2. HERRAMIENTA DE ANALISIS: PENTAHO ANALYSIS SERVICES.

Ahora bautizado como “Pentaho Analysis Service” forma parte del motor OLAP integrado en el Suite BI de Pentaho. (Rodriguez,K. & Mendoza, A., 2011). Un ejemplo rápido sobre el flujo de datos que utiliza esta herramienta es:

- El cliente manda una solicitud de consulta bajo la interfaz web JPivot.
- Mondrian recibe la solicitud y bajo el esquema de metadatos que definen sus conceptos multidimensionales busca si ya tiene los datos en cache respondiendo rápidamente a la petición.
- Si los datos no se encontraron en cache ejecuta las sentencias SQL para generar los datos.
- Se almacenan los datos recibidos en cache para agilizar posteriores consultas.
- Y finalmente se devuelve el resultado al usuario cliente a través de la interfaz.

2.3. DEFINICIÓN DE TÉRMINOS.

2.3.1. DATAMART.

Es un subconjunto de un Data Warehouse en el que una parte de los datos de una organización se sitúan en bases de datos independientes para una población específica de usuarios. (Lane, 1999)

2.3.2. DATA WAREHOUSE.

Data Warehouse es una colección de datos orientado a temas, integrado, no volátiles y variantes en el tiempo, organizados para soportar necesidades Empresariales. (Harjinder S., 1996)

2.3.3. INFORMACIÓN.

Datos que ha sido procesado en forma intangible. La información se añade a una representación y dice al receptor algo que no le era conocido antes. Lo que es información para una persona puede no serlo para otra. (Mendoza S., 2006)

2.3.4. OSCE.

Es la entidad encargada de velar por el cumplimiento de las normas en las adquisiciones públicas del Estado Peruano. Tiene competencia en el ámbito nacional, y supervisa los procesos de contratación de bienes, servicios y obras que realizan las entidades públicas. (Ley N° 1017, 2008)

2.3.5. SEACE.

El SEACE es el sistema electrónico que permite el intercambio de información y difusión sobre las Contrataciones del Estado, así como la realización de transacciones electrónicas. (Ley N° 1017, 2008)

2.3.6. BIEN.

Son objetos que requiere una Entidad para el desarrollo de sus actividades y cumplimiento de sus fines. (OSCE, 2013)

2.3.7. CONTRATACIÓN.

Es la acción que deben realizar las Entidades para proveerse de bienes, servicios u obras, asumiendo el pago de precio o de la retribución correspondiente con fondos públicos, y demás obligaciones derivadas de la condición del contratante. (OSCE, 2013)

2.3.8. PROCESO DE SELECCIÓN.

Es un procedimiento administrativo especial conformado por un conjunto de actos administrativos, de administración o hechos administrativos, que tiene por objeto la selección de la persona natural o jurídica con la cual las Entidades del Estado van a celebrar un contrato para la contratación de bienes, servicios o la ejecución de una obra. (OSCE, 2013)

2.3.9. PROVEEDOR.

La persona natural o jurídica que vende o arrienda bienes, presta servicios generales o de consultoría o ejecuta obras. (OSCE, 2013)

2.3.10. MYPE.

Es una pequeña unidad de producción, comercio o prestación de servicios. (Ley N° 28015, 2003)

2.3.11. SERVICIO EN GENERAL.

La actividad o labor que realiza una persona natural o jurídica para atender una necesidad de la entidad, pudiendo estar sujeta a resultados para considerar terminadas sus prestaciones. (OSCE, 2013)

2.4. HIPÓTESIS Y VARIABLES.

Para el presente trabajo de investigación se plantearon las siguientes hipótesis.

2.4.1. HIPÓTESIS GENERAL.

La aplicación del Datamart de Contrataciones públicas a partir del SEACE apoya la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013.

2.4.2. HIPÓTESIS ESPECÍFICAS.

HIPÓTESIS ESPECÍFICAS 1.

La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE mejoran la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

HIPÓTESIS ESPECÍFICAS 2.

La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE favorecen el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

2.5. VARIABLES.

VARIABLE INDEPENDIENTE: Datamart de Contrataciones públicas a partir del SEACE.

VARIABLE DEPENDIENTE: Toma de decisiones.

2.6. OPERACIONALIZACIÓN DE VARIABLES.

VARIABLES	DIMENSIONES	INDICADORES	ESCALA
VARIABLE INDEPENDIENTE: Datamart de Contrataciones públicas a partir del SEACE	Usabilidad	Nivel de usabilidad	- Muy buena - Buena - Regular - Mala - Muy mala
	Dimensiones	Cantidad de Escenarios	- Muy buena - Buena - Regular - Mala - Muy mala
	Tiempo	Cantidad de tiempo	- Menos de 15 min - Entre 15 y 30 min - Entre 30 y 45 min - Entre 45 min y 60 min - Mas de 1h
VARIABLE DEPENDIENTE: Toma de decisiones	Utilidad de la Información	Disponibilidad de la Información histórica.	- Muy buena - Buena - Regular - Mala - Muy mala
		Interpretación de la información.	- Muy buena - Buena - Regular - Mala - Muy mala
	Nivel de análisis de la información.	Participación de las Contrataciones	- Muy buena - Buena - Regular - Mala - Muy mala
		Oportunidades, Alternativas de Negocio.	- Regular - Mala - Muy mala
		Conocimiento del negocio, nivel de demanda de productos o servicios.	- Muy buena - Buena - Regular
		Información de la competencia	- Mala - Muy mala

Fuente: Elaboración propia.



3.1. TIPO Y DISEÑO DE LA INVESTIGACIÓN.

El presente trabajo de investigación, de acuerdo con las características de la hipótesis, los objetivos y la pregunta de investigación, se enmarca dentro del enfoque de investigación explicativa, es decir, una investigación que analizar la relación entre las variables: Datamart de Contrataciones públicas a partir de SEACE y Toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno.

El área y la línea de investigación es Administración de sistemas y Sistemas empresariales respectivamente, en la investigación se muestra la relación que existe entre los Datamarts como apoyo en la gestión de una Empresa.

El diseño de la investigación es pre experimental, Pre test y Post test con un solo grupo. Donde 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 tratamiento, para identificar una base de comparación o línea de referencia.

El diseño consiste en:

G: O1 X O2

Donde:

- G : Grupo conformado por los MYPEs de la Ciudad de Puno.
- X : Tratamiento con la aplicación del Datamart de Contrataciones públicas a partir de SEACE.
- O1 : Resultados Pre test.
- O2 : Resultados Post test.

Este diseño permite la comparación de resultados Pre test y Pos test, con un alto nivel de probabilidad, que el Datamart de Contrataciones públicas a partir del SEACE (variable independiente) ha sido factor determinante en el nivel de análisis de información y uso de la información para el apoyo en la Toma de decisiones (variable dependiente) de las Micro y pequeñas Empresas de la Ciudad de Puno en el año 2013.

3.2. POBLACIÓN.

Para el desarrollo de la investigación, la población está conformada por los MYPEs de la Ciudad de Puno. Según REMYPE, 2013, en la Ciudad de Puno existen 160 Micro y pequeñas Empresas debidamente registrados y acreditados.

3.3. MUESTRA.

Para el presente trabajo de investigación se empleó un muestreo no probabilístico de tipo intencional, es decir, un muestreo en el que el investigador escoge en función de los propósitos de la investigación (Tamayo, 2002). Ya que fue necesario hallar MYPEs que participen regularmente de las Contrataciones del Estado y que sus representantes, Gerentes o Administradores conozcan el marco normativo de las Contrataciones públicas.

Por lo tanto, la muestra para este trabajo de investigación lo constituyeron 12 MYPEs.

3.4. TÉCNICA DE RECOLECCIÓN DE INFORMACIÓN.

Se aplicó la técnica de la encuesta para el llevar a cabo la obtención de la información para el desarrollo de la investigación. El cual estuvo dirigida al Gerente o Administrador de la MYPE de la ciudad de Puno.

3.5. INSTRUMENTO.

El instrumento que se utilizó en la investigación es el cuestionario, un documento que contiene una serie de preguntas dirigidas a obtener información específica, el diseño del cuestionario está estructurado con preguntas cerradas. (Ver Anexo N° 2 y N° 3).

3.6. TÉCNICA PARA EL PROCESAMIENTO Y ANÁLISIS DE DATOS.

Para el tratamiento de datos se utilizó el siguiente esquema.

- a) Recopilación y tabulación de datos.
- c) Tabla de resultados
- d) Interpretación de los resultados

Para el análisis de los resultados y la validación de hipótesis se realizó mediante la prueba de hipótesis.

3.7. DISEÑO ESTADÍSTICO PARA LA PRUEBA DE HIPÓTESIS.

Se plantearon hipótesis estadística para cada una de las hipótesis específicas del presente trabajo de investigación:

a) HIPÓTESIS ESTADÍSTICA.

Hipótesis Estadística 1.

H1o: La usabilidad, los variables de análisis y el tiempo de acceso a la información de Datamart de Contrataciones públicas a partir del SEACE no mejora la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

H1i: La usabilidad, los variables de análisis y el tiempo de acceso a la información de Datamart de Contrataciones públicas a partir del SEACE mejora la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

Hipótesis Estadística 2.

H2o: La usabilidad, los variables de análisis y el tiempo de acceso a la información de Datamart de Contrataciones públicas a partir del SEACE no favorece el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

H2i: La usabilidad, los variables de análisis y el tiempo de acceso a la información de Datamart de Contrataciones públicas a partir del SEACE favorece el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

b) NIVEL DE SIGNIFICANCIA.

El nivel de significancia (α) escogido para la prueba de la hipótesis será del 5%. Siendo $\alpha = 0.05$ (Nivel de Significancia).

c) PRUEBA ESTADÍSTICA.

Se plantearon la aplicación de la prueba t para cada uno de los hipotesis estadísticos planteados anteriormente. La prueba t consiste en evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias. Y el valor “ t ” obtiene con la siguiente formula:

$$t = \frac{\bar{X}_d}{S_d / \sqrt{n}}$$

$$\bar{X}_d = \sum_1^n \frac{x_{i1} - x_{i2}}{n}$$

$$S_d = \sqrt{\frac{\sum_1^n (d_i - \bar{X}_d)^2}{n - 1}}$$

Donde:

\bar{X}_d : Media Aritmética de las diferencias (antes - después)

S_d : Desviación estándar de las diferencias

n : Numero de sujetos de la muestra, 12 MYPEs

c) REGLA DE DECISIÓN.

Para cada una de las hipótesis estadísticas planteadas se establece la siguiente regla de decisión con nivel de significancia $\alpha = 5$.

- Si $t < t_t$ se rechaza H_0 (Hipótesis nula)

Donde:

t = Valor de Calculado con la formula.

t_t = Valor de t en la Tabla t-student.

Por lo que se establece, que si las hipótesis específicas son aceptadas entonces validamos la hipótesis general: La aplicación del Datamart de Contrataciones públicas a partir del SEACE apoya la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013.

3.8. MATERIAL EXPERIMENTAL.

3.8.1. METODOLOGÍA DE DESARROLLO DE DATAMART.

Para el desarrollo del Datamart se seleccionó la metodología HEFESTO porque permite el desarrollo de Datamart de forma sencilla, ordenada e intuitiva, está fundamentada en una amplia investigación, comparación de metodologías existentes.

La metodología comprende las siguientes fases:

1) Análisis de requerimientos.

Identificar preguntas para las que queremos tener respuesta y los objetivos que se quieren conseguir con el nuevo sistema. Analizar las preguntas para determinar las perspectivas de análisis y los indicadores de negocio.

2) Análisis de los sistemas transaccionales.

Determinación de indicadores: identificar el origen de los indicadores en los sistemas transaccionales y determinar la forma de su cálculo. Establecer correspondencias entre los elementos definidos en el modelo conceptual y las fuentes de datos existentes en ellos OLTP (sistemas transaccionales). Y definir el nivel de granularidad que representa el nivel de detalle de los datos a obtener para cada dimensión de análisis.

3) Modelo lógico del Datamart.

En esta fase se confecciona el modelo lógico de la estructura del Datamart, teniendo como base el modelo conceptual creado en la anterior fase. Se selecciona el tipo de esquema a utilizar, en este caso elegimos estrella. Y se definen:

- Tabla de dimensiones: representa las perspectivas de análisis considerado.
- Tablas de Hechos: contiene información a partir de los cuales construye los indicadores de análisis.
- Uniones: presenta las relaciones entre las tablas de dimensiones y las tablas de hechos.

4) Procesos ETL.

En esta fase se realiza el análisis, definición y desarrollo de todos aquellos procesos necesarios para la extracción, transformación y carga de datos desde los sistemas origen hasta el llenado de datos en el Datamart.

3.8.2. HARDWARE Y SOFTWARE DE DESARROLLO.

HARDWARE	SOFTWARE
- Laptop: Intel Core i5	- Windows 7.
- Disco Duro 500 GB.	- Xampp 1.7.3.
- Memoria RAM 4 GB.	- Pentaho Suite BI 5.1.
- Memoria USB 8 GB.	- Mondrian Workbench 3.7.
- Impresora Canon Pixma 190.	- Talend Open studio 5.3.
- Tarjeta Wi - Fi D-Link DWA 510 108 Mbps.	- MySQL workbench 5.2 CE.
	- Microsoft Excel 2013
	- Mozilla Firefox 32.0.3.
	- Google chrome 37.02.

Tabla N° 2: Hardware y Software de desarrollo.
Fuente: Elaboración propia.

3.8.3. ANÁLISIS DE REQUERIMIENTOS.

Se realizaron reuniones con los Gerentes o Administradores de las MYPEs, para determinar las necesidades y requerimientos para poder definir qué información necesitan obtener de los procesos de selección para contratar con las entidades de Estado.

Las entrevistas se realizaron con el fin de recabar la información sobre la participación de las MYPEs en los procesos de selección en la provisión bienes o servicios a las entidades y la información que requieren revisar constantemente, de tal manera que facilite sus actividades diarias y no lo realicen de manera manual u operativa.

En las Contrataciones públicas los proveedores son personas naturales o jurídicas que vende bienes y presta servicios. En esta investigación los se relaciona los proveedores con las MYPEs,

3.8.3.1. REQUERIMIENTOS.

Con las entrevistas realizadas se entendió el proceso de compras de las entidades públicas. Y se definió los requerimientos de acuerdo a lo realizado, de la siguiente manera:

Tabla N° 3: Análisis de Requerimientos.

Fuente: Elaboración propia.

Nombre de Reporte	Título de Reporte	Descripción	Filtros				
Oportunidad de negocio	Demanda de Bienes/Servicios	Permite obtener Monto total de Contrataciones realizado de manera: - Anual - Mensual Agrupado por tipo de objeto de selección: - Mondo adjudicado	Año	Mes	Monto Adjudicado		
		Permite obtener Monto total de Contrataciones por entidad pública realizado de manera: - Entidad - Anual - Mensual Agrupado por tipo de objeto de selección: - Monto Adjudicado	Año	Mes	Producto	Entidad	
		Permite obtener Monto total de Contrataciones por tipo de proceso de selección realizado de manera: - Tipo de proceso - Anual - Mensual Agrupado por tipo de objeto de selección - Monto Adjudicado	Año	Mes	Producto	Tipo Proceso	
Nombre de Reporte	Título de Reporte	Descripción	Filtros				

Contrataciones por Bienes/Servicios	Cantidad de productos contratados	Permite obtener cuantas unidades de cada productos fueron contratados, realizado de manera: - Anual - Mensual Agrupado por productos - Cantidad adjudicada	Año	Mes	Cantidad adjudicada	Productos
Contrataciones por Bienes/Servicios	Cantidad de productos contratados por entidad	Permite obtener cuantas unidades de cada productos fueron contratados por cada entidad, realizado de manera: - Entidad - Anual - Mensual Agrupado por productos - Cantidad adjudicada	Año	Mes	Cantidad adjudicada	Productos/Entidad
Contratación de Bienes/Servicios por precio unitario	Precios unitarios por producto	Permite obtener a qué precio fueron contratados cada producto, realizado de manera. - Anual - Mensual Agrupado por productos - Precio unitario	Año	Mes	Precio adjudicado	Productos
Contratación de Bienes/Servicios por precio unitario	Precios unitarios por producto y entidad contratante	Permite obtener a qué precio fueron contratados cada producto, realizado de manera. - Entidad - Anual - Mensual Agrupado por entidad - Precio unitario	Año	Mes	Precio adjudicado	Productos/Entidad
Nombre de Reporte	Título de Reporte	Descripción	Filtros			



<p>Contratación de Bienes/Servicios por precio unitario</p>	<p>Precios unitarios por producto y modalidad de selección</p>	<p>Permite obtener a qué precio fueron contratados cada producto, realizado de manera. - Anual - Mensual Agrupado por productos - Tipo de modalidad de selección - Precio unitario</p>	<p>Año</p>	<p>Mes</p>	<p>Precio adjudicado</p>	<p>Productos/Modalidad de selección</p>
<p>Participación de proveedores(MYPE) en las Contrataciones publicas</p>	<p>Monto total adjudicado por proveedor</p>	<p>Permite obtener el monto adjudicado a favor de cada proveedor, realizado de manera. - Anual - Semestral - Mensual Agrupado por Proveedor - Monto adjudicado</p>	<p>Año</p>	<p>Mes</p>	<p>Monto adjudicado</p>	<p>Proveedor</p>
	<p>Productos proveídos</p>	<p>Permite obtener los productos proveídos a las entidades contratantes, realizado de manera. - Anual - Mensual Agrupado por Proveedor - Monto adjudicado - Producto</p>	<p>Año</p>	<p>Mes</p>	<p>Monto adjudicado</p>	<p>Proveedor/Producto</p>
	<p>Productos proveídos por entidad</p>	<p>Permite obtener los productos proveídos a las entidades contratantes, realizado de manera. - Anual - Mensual Agrupado por Proveedor - Monto adjudicado - Producto - Entidad</p>	<p>Año</p>	<p>Mes</p>	<p>Monto adjudicado</p>	<p>Proveedor/Producto/Entidad</p>
	<p>Productos proveídos por precio unitario</p>	<p>Permite obtener los productos proveídos por precios unitarios, realizado de manera. - Anual - Mensual - Producto – Precio unitario</p>	<p>Año</p>	<p>Mes</p>	<p>Monto adjudicado</p>	<p>Proveedor/Producto</p>

3.8.3.2. IDENTIFICADORES Y PERSPECTIVAS DE ANÁLISIS.

Con la ayuda de la definición de los requerimientos expuestos anteriormente, se determinan los datos necesarios para cumplir los requerimientos analíticos de los usuarios finales, Gerentes o Administradores de los MYPE.

Indicadores

- Monto total adjudicado en Bienes/Servicios
- Monto total adjudicado por Entidad
- Monto total adjudicado por Tipo de proceso de selección
- Cantidad total de productos contratados
- Cantidad total de productos contratados por Entidad
- Precio unitario de cada producto contratado.
- Precio unitario de cada producto contratado por Entidad
- Precio unitario de cada producto por modalidad de proceso de selección
- Monto total adjudicado a cada Proveedor
- Monto total adjudicado a cada Proveedores y por Entidad
- Monto total adjudicado en Productos por cada Proveedor
- Precio unitario de los productos por proveedor

Perspectivas

- Año
- Trimestre

- Mes
- Producto
 - Bienes
 - Servicios
- Entidades
- Tipo de proceso de selección
 - Adjudicación menor cuantía
 - Adjudicación directa selectiva
 - Adjudicación directa publica
 - Concurso público
- Modalidad de selección
 - Procedimiento clásico
 - Subasta inversa electrónica
 - Subasta inversa presencial
- Proveedores
 - Razón social

3.8.3.3. MODELO CONCEPTUAL.

En esta parte se realizó el diseño conceptual haciendo un análisis minucioso de los indicadores y perspectivas para establecer relaciones de los datos que se almacena en las tablas y campos.

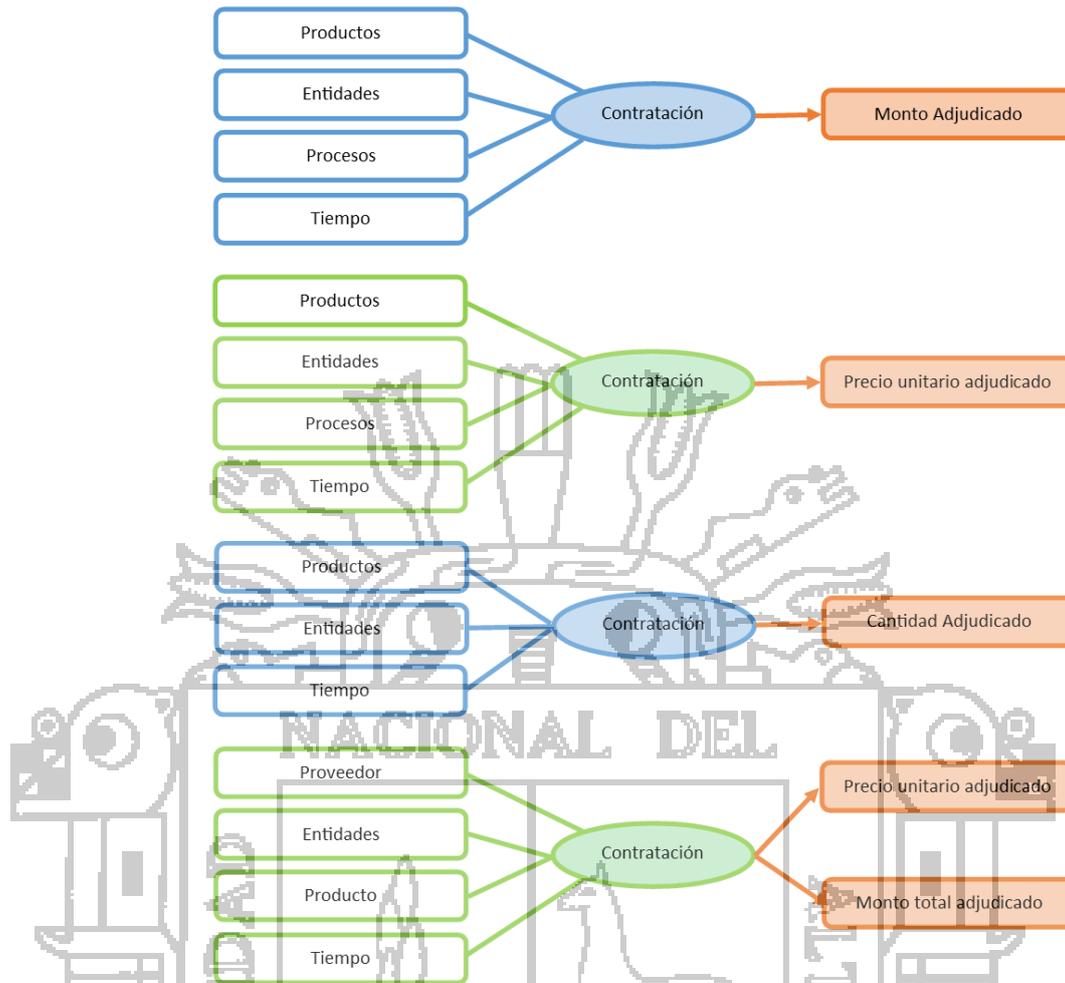


Figura N° 15: Modelo conceptual. Perspectiva-Indicadores.
Fuente: Elaboración propia.

3.8.3.4. PROCESO DE SOLUCIÓN.

Con los datos en la web SEACE se realizó los procesos ETL, para poblar el Datamart, a partir de los cuales se obtiene los datos necesarios para generar los indicadores requeridos, para realizar respectivo análisis y para la generación de reportes.



Figura N° 16: Proceso de solución.

Fuente: Elaboración propia.

3.8.3.5. FUENTE DE DATOS.

Se utilizó como fuente de datos, las publicaciones de los procesos de selección publicados en la página web del Sistema electrónico de Contrataciones del Estado (SEACE). Se consideraron los procesos de selección llevados durante el año 2013, de las entidades contratantes como son: Municipalidad provincial de Puno, Gobierno Regional de Puno sede central y Universidad Nacional del Altiplano de Puno.

Los procesos de selección están publicadas según orden cronológico en las que se llevó las Contrataciones de bienes y servicios mediante procesos de selección, y esta información es accesible para el público en general, como parte de la difusión e intercambio de información sobre las Contrataciones del Estado. Según la Ley de Contrataciones del Estado.



Fecha Presenta Propuesta	Nombre, Dirección y Teléfono de la Entidad	Nomenclatura del Proceso	Descripción del Objeto de la Contratación	Valor Referencial
03/01/2014	GOBIERNO REGIONAL DE PUNO SEDE CENTRAL (REGION PUNO) JR DEUSTUA 356 (PUNO) (Teléf: 051 356639)	LP SUBASTA INVERSA PRESENCIAL 57-2013/GRP/ICE-SIP (convocatoria : 1) (BIENES) Aviso de convocatoria con POR MOTIVOS DE FUERZA MAYOR el día 27/12/2013 18:16 horas	contratación de bien: cemento asfáltico obra: mejoramiento de la carretera macusani abra susuya distrito de macusani, provincia carabaya puno	Soles 1,283,000.00
03/01/2014	GOBIERNO REGIONAL DE PUNO SEDE CENTRAL (REGION PUNO) JR DEUSTUA 356 (PUNO) (Teléf: 051 356639)	ADS SUBASTA INVERSA PRESENCIAL 278-2013/GRP/ICE-SIP (convocatoria : 1) (BIENES) Aviso de convocatoria con POR MOTIVOS DE FUERZA MAYOR el día 27/12/2013 18:15 horas	contratación de bien: cemento asfáltico y asfalto líquido obra: mejoramiento de la infraestructura deportiva del estadio municipal chirivano del distrito de huancane	Soles 186,400.00
29/11/2013	GOBIERNO REGIONAL DE PUNO SEDE CENTRAL (REGION PUNO) JR DEUSTUA 356 (PUNO) (Teléf: 051 356639)	AMC PROCEDIMIENTO CLASICO 274-2013/GRP/CEP (convocatoria : 1) (BIENES) Aviso de convocatoria con PARA MAYOR PLURALIDAD DE POSTORES PARA EL BIEN A ADQUIRIR el día 28/11/2013 16:35 horas	contratación de bien: sulfato de potasio proyecto mej. de capac. técnicas y competitivas en la cadena productiva de frutales en la selva de la región puno	Soles 21,600.00
25/11/2013	GOBIERNO REGIONAL DE PUNO SEDE CENTRAL (REGION PUNO) JR DEUSTUA 356 (PUNO) (Teléf: 051 356639)	AMC PROCEDIMIENTO CLASICO 254-2013/GRP/CEP (convocatoria : 1) (SERVICIOS)	contratación de servicio: impresión de afiche proy. desarrollo de capacidades humanas para la conser. y uso sostenible de los rec naturales reg puno	Soles 15,952.50

Figura N° 17: Lista de Procesos de selección

Fuente: SEACE.

Síntesis
 contratación de bien: cemento portland tipo i obra: construcción y mejoramiento de infraestructura via en av. 2 de selenbre distrito de putina san antonio de putina puno
Requisitos de habilitación
 No se indican
Objeto
 BIENES
Valor Referencial
 Soles: 192,960.00
Derecho de participar
 Gratuito
Costo de reproducción de las bases
 Soles 15.00
Lugar de registro de participantes
 JR DEUSTUA 356- CIUDAD DE PUNO
 Informado el día 14/11/2013 18:20 horas

ULTIMA ETAPA-ACTO PUBLICO DE PRESENTACION DE PROPUESTAS, PUJA Y OTORGAMIENTO DE BUENA PRO el 21/11/2013, 1 ítem en BUENA PRO CONSENTIDA,
 Otras Acciones de la Convocatoria

- 1 INSTRUMENTO QUE APRUEBA O AUTORIZA LA REALIZACION DEL PROCESO, RESOLUCION ADMINISTRATIVA REGIONAL N° 722-2013-ORA-GR PUNO, , informado el día 14/11/2013 16:48
- 2 RESUMEN EJECUTIVO DE POSIBILIDADES QUE OFRECE EL MERCADO, , , informado el día 14/11/2013 16:48
- 3 ACTA DE BUENA PRO, , , informado el día 21/11/2013 18:36
- 4 BUENA PRO CONSENTIDA, , CONSENTIMIENTO AUTOMATICO, , informado el día 21/11/2013 19:26
- 5 CONTRATO U ORDEN DE COMPRA O DE SERVICIO, 854927344radB7868.PDF, , , informado el día 09/12/2013 21:17
- 6 GARANTIA CONTRATO, 854927344radCF636.PDF, , , informado el día 09/12/2013 21:18

Bases	Resumen Ejecutivo	Absolución de Consultas y Observaciones	Absolución Consultas	Absolución Observaciones	Pronuncia. Entidad	Pronuncia. OSCE	Bases Integradas	Acta de Evaluación Técnica	Cuadro Comparativo y/o Actas	Acta Buena Pro	Notificación Suscripción de Acuerdo	Resolución de Recurso de Apelación de la Entidad	Resolución de Recurso de Apelación del Tribunal
	14/11/2013 16:48 Tamaño 10,758 Kb.	14/11/2013 16:48 Tamaño 783 Kb.						21/11/2013 18:36 Tamaño 67,072 Kb.	21/11/2013 18:36 Tamaño 1,461,004 Kb.				

Item	Descripción	Familia Catálogo	Cantidad	Unidad de Medida	Valor Referencial Unitario	Valor Referencial	Plazo(s) de entrega o Fecha(s) de entrega	Lugares de entrega	Formas de entrega	Decremento mínimo	Garantía de seriedad de oferta	Beneficio de la Ley 27833	MYPE	Cantidad Adjudicada	Monto Adjudicado
1	CEMENTO PORTLAND TIPO I	A301116	7,200.00	Bolsa	192,960.00	0/C	05/12/2013	EN OBRA	CRONOGRAMA	1930	BUENA PRO CONSENTIDA informado el 21/11/2013 por SEACE			7200	188,856.00

Totales de registros encontrados : 1 Páginas : 1

Figura N° 18: Detalle de un Proceso de selección.

Fuente: SEACE.

3.8.4. ANÁLISIS DEL SISTEMA TRANSACCIONAL.

El objetivo de realizar este paso fue indagar sobre la presentación de datos en el SEACE para poder identificar la correspondencia entre el modelo conceptual y la fuentes de datos. Luego de analizar el formato de presentación de Contrataciones públicas en el SEACE se determinó las correspondencias de los campos principales de interés de cada perspectiva descrita anteriormente, y definición de los indicadores. Los resultados se exponen en la determinación de indicadores y establecimiento de correspondencias.

3.8.4.1. DETERMINACIÓN DE INDICADORES.

En esta fase se determinó la forma en que serán calculados los indicadores. A partir de los indicadores se establecen las repuestas a las preguntas para tomar una decisión, evaluación de un negocio o seguimiento de una actividad (procesos de selección).

- Monto total adjudicado
 - Hechos: Monto total adjudicado
 - Funciones de sumarización: SUM

Aclaración: El indicador “Monto total adjudicado” representa la sumatoria del monto total que se ha contratado

- Cantidad de total adjudicado
 - Hechos: Cantidad total Adjudicada
 - Funciones de sumarización: SUM

Aclaración: El indicador “Cantidad total Adjudicado” representa la sumatoria de cantidad total de productos en particular.

- Precio unitario adjudicado
 - Hechos: $(\text{Monto total adjudicado}) / (\text{Cantidad total adjudicado})$
 - Funciones de sumarización: AVG

Aclaración: El indicador “Precio unitario adjudicado” representa el promedio de precio unitario de cada producto contratado durante un periodo de tiempo.

3.8.4.2. FORMANDO CORRESPONDENCIAS.

Se establece las correspondencias con la fuente de datos origen y el modelo conceptual para determinar los campos o atributo para cada perspectiva.





Convocatoria													
AMC PROCEDIMIENTO CLASICO .32-2013/MPP (convocatoria : 1) MENOR CUANTIA en la modalidad de CONVOCATORIA													
Entidad Contratante MUNICIPALIDAD PROVINCIAL DE PUNO (MPP) Teléf : 051-301593													
Dirección JR. DE JUSTITIA N° 458-PUNO (PUNO)													
Síntesis gasohol 84 plus													
Objeto BIBNES													
Valor Referencial Soles : 13,45.75													
Derecho de participar Gratis													
Costo de reproducción de las bases Soles : 10.00													
Lugar de registro de participantes JIRON DE JUSTITIA N° 458-PLAZA DE ARMAS 2do Piso Informado el día 01/07/2013 18:23 horas													
Calendario													
Actividad	Fecha Inicio	Fecha Fin											
convocatoria	01/07/2013 00:00	01/07/2013 00:00											
registro de participantes	02/07/2013 07:45	02/07/2013 15:45											
presentacion de propuestas	03/07/2013 07:45	03/07/2013 12:00											
calificacion y evaluacion de propuestas	03/07/2013 00:00	03/07/2013 00:00											
otorgamiento de la buena pro	03/07/2013 00:00	03/07/2013 00:00											
ÚLTIMA ETAPA : OTORGAMIENTO DE LA BUENA PRO el 03/07/2013, 1 item en BUENA PRO CONSENTIDA,													
Otras Acciones de la Convocatoria													
1 INSTRUMENTO QUE APRUEBA O AUTORIZA LA REALIZACION DEL PROCESO, PROVEIDO N° 1255 GERENCIA MUNICIPAL, , , informado el día 01/07/2013 18:09													
2 ACTA DE BUENA PRO, , , informado el día 04/07/2013 10:05													
3 BUENA PRO CONSENTIDA, , CONSENTIMIENTO AUTOMATICO, , informado el día 05/07/2013 00:26													
4 CONTRATO U ORDEN DE COMPRA O DE SERVICIO, 870213981radEC80A.pdf, , informado el día 23/07/2013 12:03													
Bases	Resumen Ejecutivo	Absolución de Consultas y Observaciones	Absolución Consultas	Absolución Observaciones	Pronuncia. Entidad	Pronuncia. OSCE	Bases Integradas	Acta de Evaluación Técnica	Cuadro Comparativo y/o Actas	Acta Buena Pro	Notificación Suscripción de Acuerdo	Resolución de Recurso de Apelación de la Entidad	Resolución de Recurso de Apelación del Tribunal
01/07/2013 18:09 Tamaño 757 Kb.									04/07/2013 10:05 Tamaño 296,475 Kb.	04/07/2013 10:05 Tamaño 290,736 Kb.			
Item	Descripción	Familia Catálogo	Cantidad	Unidad de Medida	Valor Referencial	Plazo(s) de entrega	Garantía de seriedad de oferta	Beneficio de la ley 27633	MYPE	Cantidad Adjudicada	Monto Adjudicado		
1	gasohol 84 plus	A151015	1,055.00	Galón	13,345.75	0		BUENA PRO CONSENTIDA informado el 05/07/2013 por SEACE		1055	13,345.75		

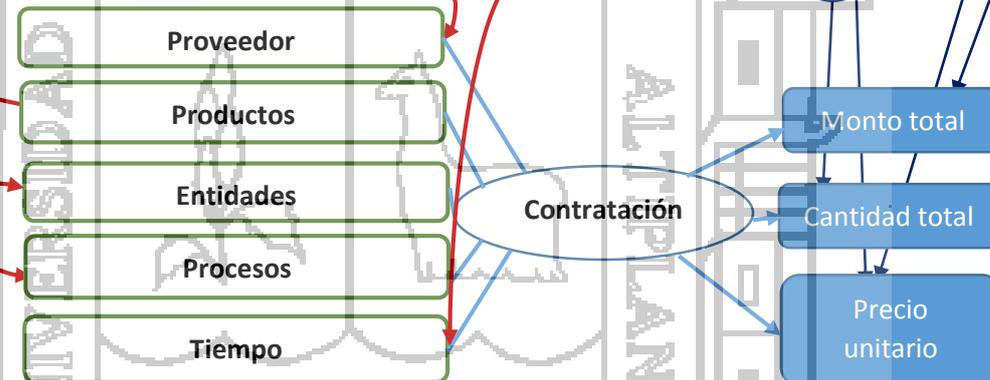


Figura N° 19: Relación de perspectivas y Datos fuente.
Fuente: Elaboración propia.

Las relaciones identificadas fueron los siguientes:

- El valor del campo “descripción” se relaciona con la perspectiva “Productos”.
- El valor del Campo “Buena pro” se relaciona con la perspectiva “Proveedor”.

- El valor del campo Entidad convocante se relaciona con la perspectiva “Entidades”.
- El valor del campo “convocatoria” se relaciona con la perspectiva “Procesos”.
- El valor “Buena pro consentida” se relaciona con la perspectiva “Tiempo”.
- El campo “cantidad adjudicada” se relaciona con el indicador “Cantidad total Adjudicada”.
- El campo “monto adjudicado” se relaciona con el indicador “Monto total adjudicado”.
- El campo “monto adjudicado” dividido con el campo “cantidad adjudicada”, se relaciona con el indicador “Precio unitario adjudicado”.

3.8.4.3. NIVEL DE GRANULARIDAD.

A partir de los requerimientos de información se define la granularidad, ¿Qué nivel de detalle debería estar disponibles en el modelo dimensional? Se describe para cada perspectiva, en el cual se determina los datos considerados de interés para analizar los indicadores ya expuestos. Teniendo el siguiente resultado:

- Perspectiva “Productos”
 - “nombre_prod”
 - “tipo”
 - “medida”

- Perspectiva “Proveedor”
 - “ruc”
 - “razón_social”
- Perspectiva “Entidad”
 - “nombre_entidad”
- Perspectiva “Proceso”
 - “tipo_proceso”
 - “modalidad”
 - “numero”
 - “convocatoria”
- La perspectiva “Tiempo”, es la que determina la granularidad del depósito de datos, los datos que se emplearon fueron:
 - Año
 - Semestre
 - Trimestre
 - Mes
 - Día

3.8.4.4. MODELO CONCEPTUAL AMPLIADO.

Después de establecer las correspondencias entre el modelo conceptual y el origen de datos, se forjan perspectivas más completas. Y las perspectivas se describen con el nivel de granularidad que determina el nivel de detalle en el análisis de la información.

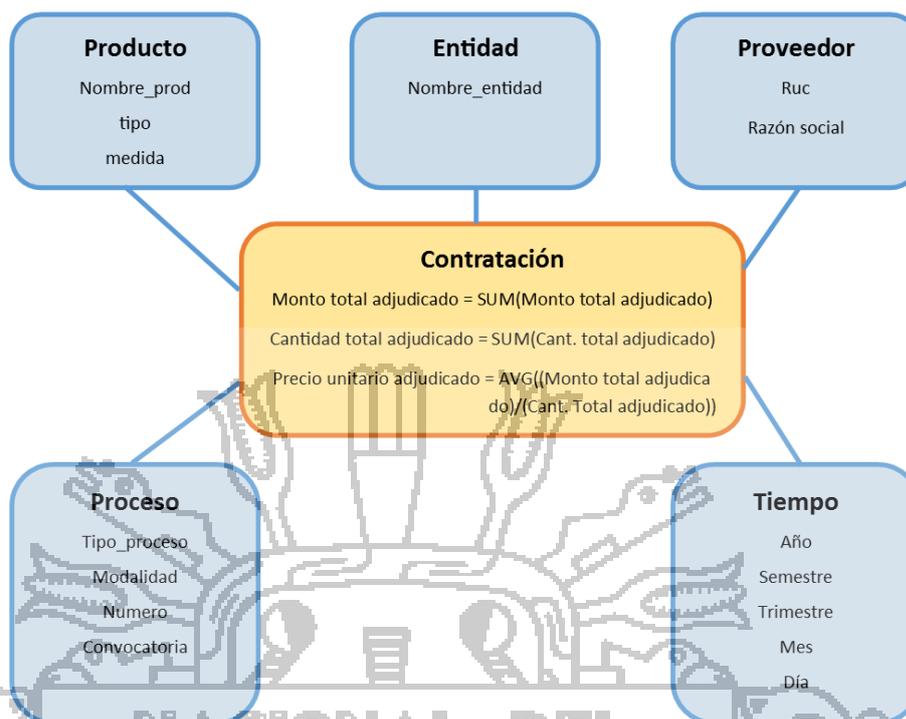


Figura N° 20: Modelo conceptual ampliado.
Fuente: Elaboración propia.

3.8.5. MODELO LÓGICO DEL DATAMART.

A continuación, se confeccionó el modelo lógico de la estructura del Datamart, teniendo como base el modelo conceptual. A partir de este último se desarrolló el Modelo físico de la base de datos o arquitectura de base de datos que se focalizó sobre la selección de las estructuras de almacenamiento (variables, tipo, longitud, entre otros).

3.8.5.1. TIPO DE MODELO LÓGICO DEL DATAMART.

El esquema que se utiliza es estrella, debido a sus características, ventajas y diferencias con los otros esquemas.

3.8.5.2. TABLA DE DIMENSIONES.

Cada perspectiva definida en el modelo conceptual constituye una tabla dimensión, para definir una dimensión se toma en cuenta lo siguiente:

- **Dim:** Representa un dimensión.
- **Nombre Tabla:** Nombre de la tabla a la que hace referencia la dimensión.
- **Formato:** Dim Nombre de la tabla a la que hace referencia.

a) Lista de dimensiones.

- DimProducto
- DimProveedor
- DimEntidad
- DimProceso
- DimTiempo

b) Descripción de las dimensiones.

Tabla N° 4: Descripción de la tabla DimProducto.

Tabla	DimProducto				
Descripción	Tabla de registro de Productos				
	Destino			Origen	
Columna	Desc.	Tipo	Tabla	Campo	Tipo
Idproducto	Llave primaria	INT		Proceso de transformación Java	
producto_desc	Nombre de producto	VARCHAR	web-SEACE	Descripcion	VARCHAR
medida_id	Id de medida	INT		Proceso de transformación Java	
medidad_desc	Nombre de unidad de medida	VARCHAR	web-SEACE	desc_medida	VARCHAR

Tabla N° 5: Descripción de la Tabla DimProveedor.

Tabla	DimProveedor				
Descripción	Tabla de registro de Proveedores				
Destino			Origen		
Columna	Desc.	Tipo	Tabla	Campo	Tipo
idproveedor	Llave primaria	CHAR	Web-SEACE	Buena pro	CHAR
Lugar	Lugar origen	VARCHAR	Web-SEACE	Buena pro	VARCHAR
tipoprov_id	Tipo de MyPE	INT	-	Proceso de transformación Java	-
tipoprov_desc	Nombre de tipo de MyPE	VARCHAR	-	Proceso de transformación Java	-

Tabla N° 6: Descripción de la Tabla DimEntidad.

Tabla	DimEntidad				
Descripción	Tabla de registro de Entidades				
Destino			Origen		
Columna	Desc.	Tipo	Tabla	Campo	Tipo
Identidad	Llave primaria	INT		Proceso de transformación Java	
nom_entidad	Nombre de entidad contratante	VARCHAR	web-SEACE	Entidad Contratante	VARCHAR

Tabla N° 7: Descripción de la tabla DimProceso.

Tabla	DimProceso				
Descripción	Tabla de registro de Procesos				
Destino			Origen		
Columna	Desc.	Tipo	Tabla	Campo	Tipo
Idproceso	Llave primaria	INT		Proceso de transformación Java	
tipoproceso_id	Tipo de proceso	TINYINT		Proceso de transformación Java	
tipoproceso_desc	Nombre de tipo de proceso	VARCHAR	web-SEACE	convocatoria	VARCHAR
modalidad_id	Modalidad	INT		Proceso de transformación Java	
modalidad_desc	Desc. de modalidad	VARCHAR	web-SEACE	convocatoria	VARCHAR
Numero	Numero de proceso	INT	web-SEACE	convocatoria	INT
Convocatoria	Numero de convocatoria	INT	web-SEACE	convocatoria	INT

Tabla N° 8: Descripción de la tabla DimTiempo.

Tabla	DimTiempo	
Descripción	Tabla de registro de Tiempos	
Destino		
Columna	Desc.	Tipo
ldtiempo	Llave primaria	INT
bp_fecha	Fecha	DATE
bp_fecha_desc	Fecha larga	VARCHAR
bp_anio	Año	VARCHAR
bp_semestre	Semestre	VARCHAR
bp_trimestre	Trimestre	VARCHAR
bp_numero_mes	Numero de mes	INT
bp_mes	Nombre de mes	VARCHAR
bp_dia_del_mes	Día del mes	INT

c) Tabla de hechos.

Tabla N° 9: Descripción de la Tabla de FactContratación.

Tabla	FactContratacion	
Descripción	Tabla de hechos contratación	
Destino		
Columna	Desc.	Tipo
monto_tot_adjudicado	Campo para monto de adjudicado	DOUBLE
cantidad_adjudicado	Campo para cantidad adjudicado	INT
monto_uni_adjudicado	Precio unitario adjudicado	DOUBLE
ldproceso	FK – tabla DimProceso	INT
lidentidad	FK – tabla DimEntidad	INT
ldproducto	FK – tabla DimProducto	INT
ldproveedor	FK – tabla DimProveedor	INT
ldtiempo_buenapro	FK – tabla DimTiempo	INT

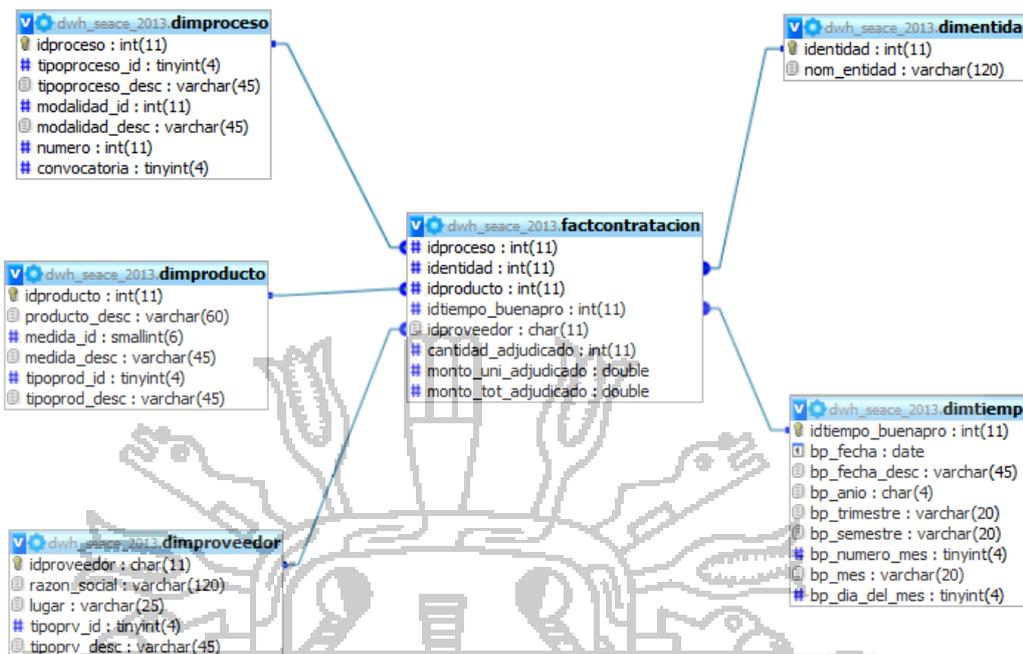


Figura N° 21: Modelo físico del Datamart.
Fuente: Elaboración propia.

3.8.6. PROCESO DE ETL.

Mediante el proceso de Extracción, Transformación y Carga (Extract, Transform, Load) se organizó el flujo de datos entre los datos publicados en el SEACE y el Datamart. De las publicaciones de los procesos de selección en el SEACE sólo se seleccionó información necesaria, para ello se utilizó herramientas que seleccionan datos determinados, con los cuales se realizó las transformaciones necesarias (limpieza y formateo) sobre los datos extraídos para luego mover esta fuente de información a una base de datos (Datamart). Y esta información resultante es la que se utilizó para el análisis sobre las Contrataciones públicas del estado.

Se utilizó la herramienta Import.io para leer los datos desde la web de SEACE y Talend open studio para realizar la Extracción, Transformación y Carga. Para mayor detalle ver Anexos N° 5 y 6.

3.8.6.1. PROCESO ETL DE DIMENSIONES.

Para todos los procesos ETL de las dimensiones DimProducto, DimProveedor, DimProceso, DimEntidad, se inicia con el archivo MS Excel generado por la herramienta Import.io, resultado de la lectura de datos desde la web SEACE el cual contiene la información necesaria para iniciar el poblamiento del Datamart.

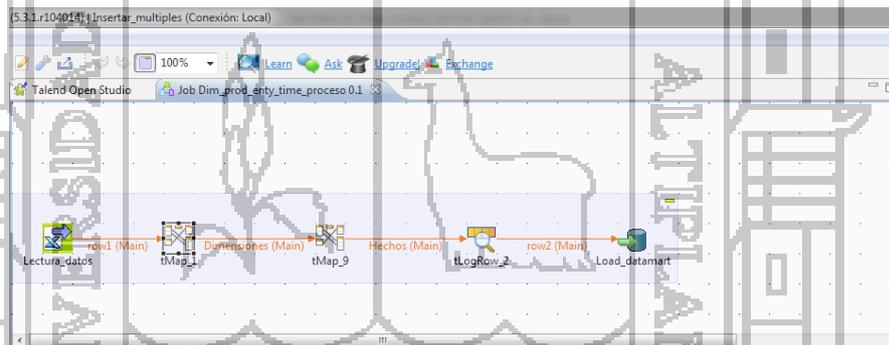


Figura N° 22: Configuración de proceso ETL para las Dimensiones.
Fuente: Elaboración propia.

El objetivo de la transformación es de formatear y realizar limpieza de los datos, y como resultado se alcanzaron datos estandarizados para ser almacenado en el Datamart. Ver Anexos.

3.8.6.2. PROCESO ETL DE DIMTIEMPO

Las fechas son registradas jerárquicamente empezando desde el año, mes y día.

idtiempo_buenapro	bp_fecha	bp_fecha_desc	bp_anio	bp_trimestre	bp_semestre	bp_numero_mes	bp_mes	bp_dia_del_mes
20130531	2013-05-31	viernes, 31 de mayo de 2013	2013	2do Trimestre	Primer Semestre	5	MAYO	31
20130601	2013-06-01	sábado, 1 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	1
20130602	2013-06-02	domingo, 2 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	2
20130603	2013-06-03	lunes, 3 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	3
20130604	2013-06-04	martes, 4 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	4
20130605	2013-06-05	miércoles, 5 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	5
20130606	2013-06-06	jueves, 6 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	6
20130607	2013-06-07	viernes, 7 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	7
20130608	2013-06-08	sábado, 8 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	8
20130609	2013-06-09	domingo, 9 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	9
20130610	2013-06-10	lunes, 10 de junio de 2013	2013	2do Trimestre	Primer Semestre	6	JUNIO	10

Figura N° 23: Resultado de la transformación para DimTiempo.
Fuente: Elaboración propia.

3.8.7. CREACIÓN DE CUBOS MULTIDIMENSIONALES.

Una vez poblado la base de datos multidimensional, se creó el cubo multidimensional o también conocido como cubo de decisión el cual nos permite tener una visión dinámica del análisis de datos, para ello se necesita distribuir las métricas (tabla de hechos) y dimensiones (tabla de dimensiones) en filas, columnas e indicadores. Para crear el cubo multidimensional definió lo siguiente:

- **Indicadores:** Está conformado por los tres indicadores, que conforman la tabla de hechos.
 - Monto total adjudicado
 - Cantidad total adjudicado
 - Precio unitario adjudicado
- **Atributos:** Esta conformados por los campos de las tablas:
 - DimProducto
 - DimProveedor
 - DimEntidad
 - DimProceso
 - DimTiempo
- **Jerarquías:** Las jerarquías representan al nivel en que puede ser evaluado un indicador.

La definición de dimensiones en atributos y presentados en niveles nos permite tener una mayor visión para el análisis de lo genérico a lo específico y viceversa. Cabe resaltar que para la creación de una jerarquía se necesita de un mínimo de dos niveles.

Para el presente trabajo se utilizó la herramienta Pentaho schema workbench, para diseñar el cubo, primer lugar se define el esquema general con la tabla hechos “factcontratacion” como tabla pivot y luego se establecen las jerarquías de cada una de las dimensiones que forman parte del base de datos multidimensional.

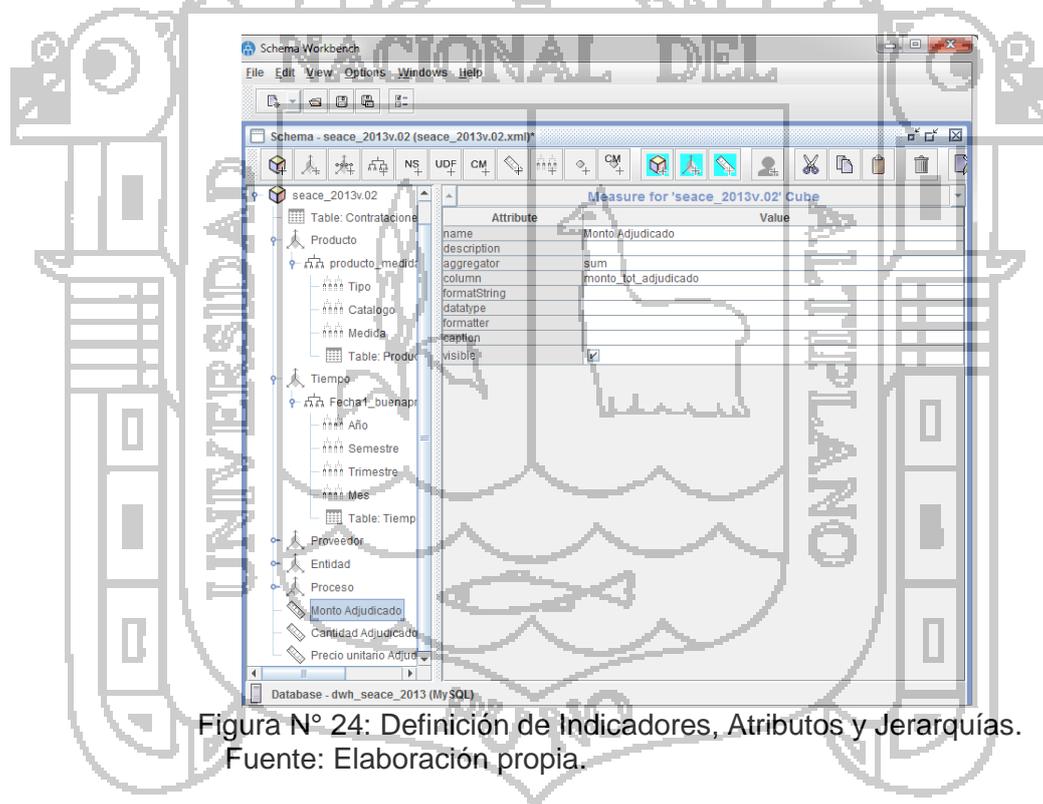


Figura N° 24: Definición de Indicadores, Atributos y Jerarquías.
Fuente: Elaboración propia.

El cubo diseñado con la aplicación Schema Worckbench es almacenado en un archivo XML para publicar en el servidor Mondrian, para el procesamiento analítico en línea. Ver Anexo N° 7.

3.8.8. DESPLIEGUE Y ANÁLISIS DE DATOS PARA EL APOYO DE TOMA DE DECISIONES.

En esta fase se pone en funcionamiento toda la plataforma dirigida a los usuarios finales.

Se ingresa a la plataforma a través de un explorador web, ya que esta plataforma es de acceso web en todo momento. Se muestra la página de inicio de plataforma Pentaho, donde se inicia sesión para acceso.

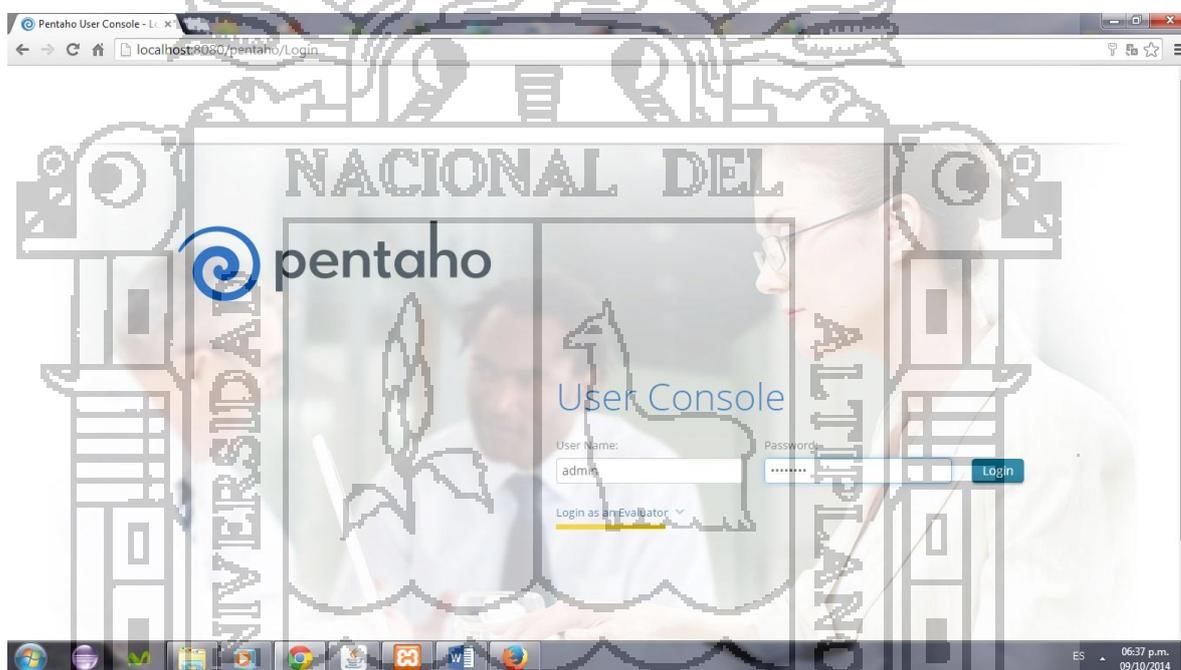


Figura N° 25: Pantalla de Inicio de Pentaho.
Fuente: Elaboración propia.

Una vez iniciado sesión se accede a la página principal de la plataforma de negocios, donde se tiene alcance a todas herramientas, los menús de administración y todas las funcionalidades cuenta con opciones configuraciones para la conexión con la base de datos dimensional, conexión con el servidor Mondrian para iniciar con la explotación de datos.

El resultado que se obtienen son reportes en sí, con opciones necesarias para obtener la información desde diferentes perspectivas, y recorrer las distintas dimensiones y sus cruces, visualizando para cada caso los valores resultantes de las medidas. Los resultados también son mostrados mediante gráficos.

Tipo	Año	Semestre	Trimestre	MUNICIPALIDAD PROVINCIAL DE PUNO	UNIVERSIDAD NACIONAL DEL ALTIPLANO	GOBIERNO REGIONAL PUNO SEDE CENTRAL
BIENES	2013	Primer Semestre	1er Trimestre	313.517,146	498.909,35	7.004.157,44
			2do Trimestre	1.764.764,870	6.019.876,98	12.116.338,48
		Segundo Semestre	3er Trimestre	3.973.834,6	3.051.767,78	9.531.787,61
			4to Trimestre	1.688.771,094	14.753.614,63	41.421.643,009
	2014	Primer Semestre	1er Trimestre - 2014	59.947,301	1.383.980,1	7.570.841,15
			2do Trimestre - 2014			87,620
		Segundo Semestre	1er Trimestre	19,000	21,999	3.332.949,1
			2do Trimestre	268.028,562		88.919.074,44
SERVICIOS	2013	Primer Semestre	3er Trimestre	231,624	103,920	2.341,819
			4to Trimestre	211,709	711.649,95	8.478.032,25
	2014	Primer Semestre	1er Trimestre - 2014		80.466,07	1.008,920

Figura N° 26: Análisis de información, de general a específico.
Fuente: elaboración propia.

En la figura anterior se muestra un análisis de los montos totales adjudicados mediante los procesos de selección para la contratación de bienes y servicios. Se inicia con una información general y bajando una jerarquía brinda la posibilidad de introducir un nuevo nivel o criterio de agregación en el análisis, separando grupos actuales que permiten apreciar los datos en un mayor detalle. Esta operación se denomina de ir de lo general a lo específico (Drill down). También se puede realizar un análisis de específico a lo general (Drill up).

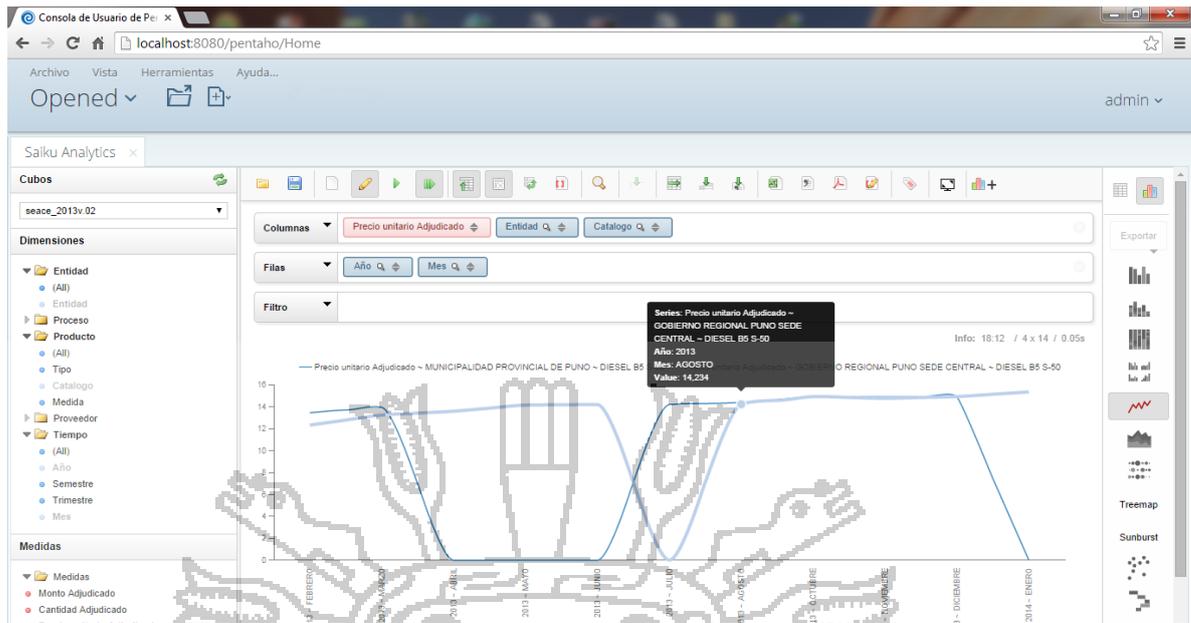


Figura N° 27: Precio unitario adjudicado en el tiempo
Fuente: Elaboración propia.

En la figura N° 27, se muestra un análisis de datos en el tiempo. Se aprecia los precios unitarios adjudicados de un bien específico durante el año 2013.





4.1. CONTRASTE DE HIPÓTESIS.

La contrastación de hipótesis se realizó con el método propuesto (Capítulo III: Materiales y métodos) Pre test – Post test, que nos permite aceptar o rechazar la hipótesis. Para esto se realizó una prueba por cada indicador.

4.1.1. CONTRASTE DE HIPÓTESIS ESTADÍSTICA 1.

H1o: La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE no mejoran la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

H1i: La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE mejoran la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

$$\mathbf{H1o: } P_A - P_B \geq 0$$

$$\mathbf{H1i: } P_A - P_B < 0$$

Donde:

P_A : Puntación promedio del sistema actual que representa el nivel de apoyo en la toma de decisiones.

P_B : Puntación promedio de Datamart propuesto que representa el nivel de apoyo en la toma de decisiones.

a) Nivel de Significancia.

El nivel de significancia (α) escogido para la prueba de la hipótesis es de 5%. Siendo $\alpha = 0.05$ (Nivel de Significancia) y $n - 1 = 4$ grados de libertad ($n =$ número de Indicadores), se tiene el valor crítico de T student (Anexo 4):

$$\text{Valor crítico: } t_{-0.05;4} = -2.1318$$

Región de rechazo, valores de t menores que $t_t = -2.1318$.

b) Situación actual.

Para definir la situación actual de los MYPEs con respecto a la utilización de la información para la toma de decisiones, con la usabilidad, los variables de análisis y el tiempo de acceso del SEACE para contratar con el estado, se realizó una encuesta (pre test) a los Gerentes o Administradores de los MYPEs de la Ciudad de Puno.

Tabla N° 10: Pre test nivel utilidad de información.

PREGUNTAS	CALIFICACIÓN					PUNTAJE TOTAL	PROMEDIO
	MB	B	R	M	MM		
	5	4	3	2	1		
1. ¿Cuál es el grado de usabilidad del SEACE para la toma de decisiones para contratar con el estado?	1	3	4	2	2	35	2.92
2. ¿Cómo considera los reportes del SEACE para el análisis desde diferentes ópticas para la toma de decisiones?	0	1	6	3	2	30	2.50
3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones?	0	2	3	4	3	28	2.33

4. ¿Cómo considera las Contrataciones públicas históricas para la toma de decisiones, disponibles en el SEACE?	0	1	3	5	3	26	2.17
5. ¿Cómo considera la presentación de la información estructurada del SEACE en contratación de bienes y servicios, para la toma de decisiones?	0	1	3	6	2	27	2.25

c) Datamart propuesto: Utilidad de la información

Después de poner el Datamart a disposición de los Gerentes o Administrativos para su utilización, se procedió con la segunda encuesta (pos test), para determinar el nivel de utilidad de información.

Tabla N° 11: Post test nivel utilidad de información con Datamart.

PREGUNTAS	CALIFICACIÓN					PUNTAJE TOTAL	PROMEDIO
	MB	B	R	M	MM		
	5	4	3	2	1		
1. ¿Cuál es el grado de usabilidad del DATMART para la toma de decisiones para contratar con el estado?	2	3	4	2	1	39	3.25
2. ¿Cómo considera los reportes del DATAMART para el análisis desde diferentes ópticas para la toma de decisiones?	2	4	3	2	1	40	3.33
3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones, con la aplicación DATAMART?	4	4	3	1	0	47	3.92



4. ¿Cómo considera las Contrataciones públicas históricas para la toma de decisiones, disponibles en el DATAMART?	1	4	5	2	0	40	3.33
5. ¿Cómo considera la presentación de la información estructurada del DATAMART en contratación de bienes y servicios, para la toma de decisiones?	1	5	4	2	0	41	3.42

d) Resultados de la hipótesis estadista 1.

Con los resultados obtenidos de la situación actual (Pre test) más los resultados obtenidos del Datamart propuesto (Post test), se procedió a realizar una comparación de ambas encuestas para poder observar los cambios que resultaron, para poder establecer si el Datamart propuesto mejora la utilidad de información para el apoyo de la toma de decisiones de las MYPES.

Tabla N° 12: Contrastación entre Pre test y Post test.

N°	Pregunta/Indicador	PRE TEST P_A	POST TEST P_B	d $= P_A - P_B$	d (%)	$d - \bar{X}_d$	$(d - \bar{X}_d)^2$
1	Nivel de Usabilidad	2.92	3.08	-0.16	11%	0.80	0.642
2	Cantidad de Escenarios	2.50	3.33	-0.83	33%	0.13	0.017
3	Cantidad de Tiempo	2.33	3.92	-1.59	68%	-0.63	0.396
4	Disponibilidad de información	2.17	3.33	-1.16	53%	-0.20	0.040
5	Interpretación de la información	2.25	3.42	-1.17	52%	-0.21	0.044
				$\sum_1^5 d$ $= -5.08$			$\sum_1^5 (d - \bar{X}_d)^2$ $= 0.88$

Fuente: Elaboración propia.

Representación gráfica:

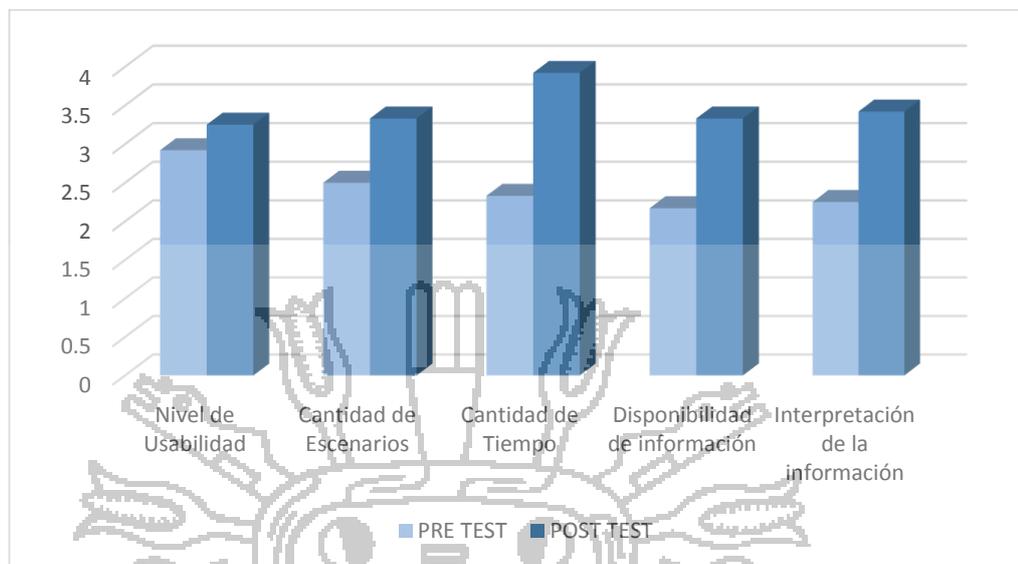


Figura N° 28: Gráfico de Barras – Contrastación de Pre test y Post test.

Fuente: Elaboración propia

Interpretación: Según los resultados se observó que cuando el nivel de usabilidad se incrementa en 11%, la cantidad de Escenarios se incrementan 33% y el tiempo de respuesta se mejora en 68%, la disponibilidad de información y la interpretación de la información sufrieron un incremento de 53% y 52% respectivamente.

Calculo de t para muestras relacionadas.

$$t = \frac{\bar{X}_d}{S_d / \sqrt{n}}$$

Donde:

\bar{X}_d : Media Aritmética de las diferencias

S_d : Desviación estándar de las diferencias

$$\bar{X}_d = \sum_{i=1}^n \frac{x_{i1} - x_{i2}}{n}$$

n : Numero de Indicadores

$$S_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{X}_d)^2}{n - 1}}$$

Remplazando datos:

Media aritmética

Desviación estándar

$$n = 5$$

$$n = 5$$

$$\sum_1^5 d = -5.08$$

$$\sum_1^5 (d - \bar{X}_d)^2 = 0.88$$

$$\bar{X}_d = \frac{-5.08}{5}$$

$$S_d = \sqrt{\frac{0.88}{5-1}}$$

$$\bar{X}_d = -1.016$$

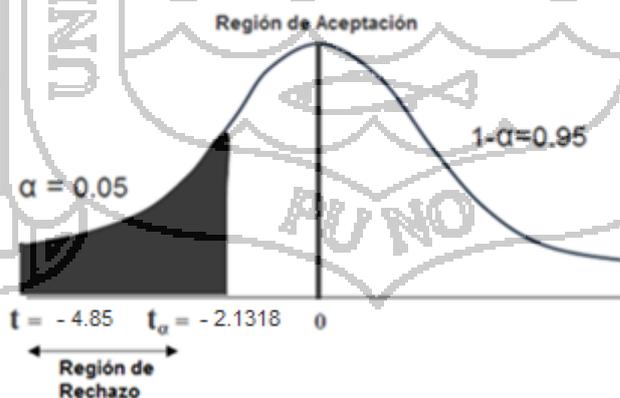
$$S_d = 0.47$$

Calculo de t:

$$t = \frac{-1.016}{0.47 / \sqrt{5}}$$

$$t = -4.85$$

e) Decisión: Utilidad de la información.



$$t = -4.85(\text{Calculado}) < t_{0.05;4} = -2.1318(\text{Tabular})$$

Al comparar los valores, se rechaza la hipótesis nula (H1o) con un margen de error de 5% y aceptamos la hipótesis alternativa (H1i) por cual se determina

que la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE mejoran la utilidad de la información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno – 2013.

4.1.2. CONTRASTE DE HIPÓTESIS ESTADÍSTICA 2.

H2o: La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE no favorecen el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

H2i: La usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE favorecen el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

$$\mathbf{H2o: } P_A - P_B \geq 0$$

$$\mathbf{H2i: } P_A - P_B < 0$$

Donde:

P_A : Puntación del sistema actual que representa el apoyo de la toma de decisiones con el nivel de información.

P_B : Puntación de Datamart propuesto que representa el apoyo en la toma de decisiones con el nivel de información.

a) Nivel de Significancia.

El nivel de significancia (α) escogido para la prueba de la hipótesis es de 5%. Siendo $\alpha = 0.05$ (Nivel de Significancia) y $n - 1 = 6$ grados de libertad ($n =$ número de Indicadores), se tiene el valor crítico de T student (Anexo 4):

$$\text{Valor crítico: } t_{-0.05;6} = -1.9432$$

Región de rechazo, valores de t menores que $t_t = -1.9432$.

b) Situación actual: Análisis de información.

Para definir la situación actual de los MYPEs con respecto al nivel de análisis de información para la toma de decisiones, con la usabilidad, los variables de análisis y el tiempo de acceso del SEACE para contratar con el estado, se realizó una encuesta (pre test) a los Gerentes o Administradores de los MYPEs de la Ciudad de Puno.

Tabla N° 13: Pre test nivel análisis de la información.

PREGUNTAS	CALIFICACIÓN					PUNTAJE TOTAL	PROMEDIO
	MB	B	R	M	MM		
1. ¿Cuál es el grado de usabilidad del SEACE para la toma de decisiones para contratar con el estado?	5	4	3	2	1	35	2.92
2. ¿Cómo considera los reportes del SEACE para el análisis desde diferentes ópticas para la toma de decisiones?	1	3	4	2	2		
3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones?	0	1	6	3	2		
	0	2	3	4	3	28	2.33



6. La información que se brinda sobre su participación de las Contrataciones públicas a través del SEACE es:	0	1	4	5	2	28	2.33
7. ¿Cómo considera la información de las Contrataciones publicadas sobre oportunidades, y alternativas de negocios en el SEACE?	1	3	4	2	2	35	2.92
8. El nivel de información que ofrece el SEACE sobre cantidad, precios y montos adjudicados de un bien o servicio es:	0	1	4	3	4	26	2.17
9. ¿Cómo considera la información que brinda el SEACE con respecto a montos totales, cantidades, y precios unitarios adjudicado a un proveedor en el tiempo?	0	1	2	4	5	23	1.92

c) Datamart propuesto: Análisis de la información.

Después de poner el Datamart a disposición de los Gerentes o Administrativos para su utilización, se procedió con la segunda encuesta (pos test), para determinar el nivel de Análisis de la Información.

Tabla N° 14: Post test nivel análisis de la información con Datamart.

PREGUNTAS	CALIFICACIÓN					PUNTAJE TOTAL	PROMEDIO
	MB	B	R	M	MM		
	5	4	3	2	1		
1. ¿Cuál es el grado de usabilidad del DATMART para la toma de decisiones para contratar con el estado?	2	3	4	2	1	37	3.08
2. ¿Cómo considera los reportes del DATAMART para el análisis desde diferentes ópticas para la toma de decisiones?	2	4	3	2	1	40	3.33

3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones, con la aplicación DATAMART?	4	4	3	1	0	47	3.92
6. La información que se brinda sobre su participación de las Contrataciones públicas a través del DATMART es:	2	4	5	1	0	43	3.58
7. ¿Cómo considera la información de las Contrataciones publicadas sobre oportunidades, y alternativas de negocios en el DATAMART?	1	2	6	2	1	36	3.00
8. El nivel de información que ofrece el DATAMART sobre cantidad, precios y montos adjudicados de un bien o servicio es:	2	4	4	2	0	42	3.50
9. ¿Cómo considera la información que brinda el DATAMART con respecto a montos totales, cantidades, y precios unitarios adjudicado a un proveedor en el tiempo?	1	3	4	3	1	36	3.00

d) Resultados de la hipótesis estadista 2.

Tabla N° 15: Contrastación entre Pre test y Post test

N°	Pregunta/Indicador	PRE TEST P_A	POST TEST P_B	d $= P_A - P_B$	d (%)	$d - \bar{X}_d$	$(d - \bar{X}_d)^2$
1	Nivel de Usabilidad	2.92	3.25	-0.33	11%	0.60	0.36
2	Cantidad de Escenarios	2.5	3.33	-0.83	33%	0.10	0.01
3	Cantidad de Tiempo	2.33	3.92	-1.59	68%	-0.66	0.44
6	participación de las Contrataciones	2.33	3.58	-1.25	54%	-0.32	0.10
7	Oportunidades, Alternativas de Negocio.	2.92	3	-0.08	3%	0.85	0.72
8	Conocimiento del negocio	2.17	3.5	-1.33	61%	-0.40	0.16
9	Información de la competencia	1.92	3	-1.08	56%	-0.15	0.02
				$\sum_1^7 d$ $= -6.49$			$\sum_1^7 (d - \bar{X}_d)^2$ $= 1.81$

Fuente: Elaboración propia.

Representación gráfica:

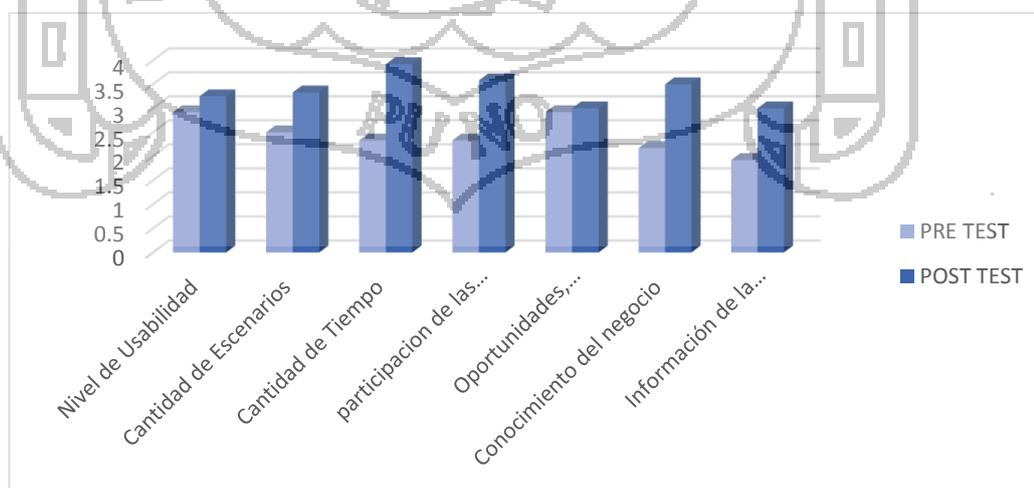


Figura N° 29: Gráfico de Barras – Contrastación de Pre test y Post test.
Fuente: Elaboración propia.

Interpretación: Según los resultados se observó que cuando el nivel de usabilidad se incrementa en 11%, la cantidad de Escenarios se incrementan 33% y el tiempo de respuesta se mejora en 68%, el análisis sobre la participación de Contrataciones públicas, oportunidades (alternativas) de negocio, conocimiento de negocio (demandas de bienes y servicios) y la información de la competencia sufrieron un incremento de 53%, 3%, 61% y 56% respectivamente.

Calculo de t para muestras relacionadas, reemplazamos los datos a las fórmulas que fueron descritas anteriormente:

Media aritmética

$$n = 7$$

$$\sum_{i=1}^7 d = -6.49$$

$$\bar{X}_d = \frac{-6.49}{7}$$

$$\bar{X}_d = -0.93$$

Desviación estándar

$$n = 7$$

$$\sum_{i=1}^7 (d - \bar{X}_d)^2 = 1.81$$

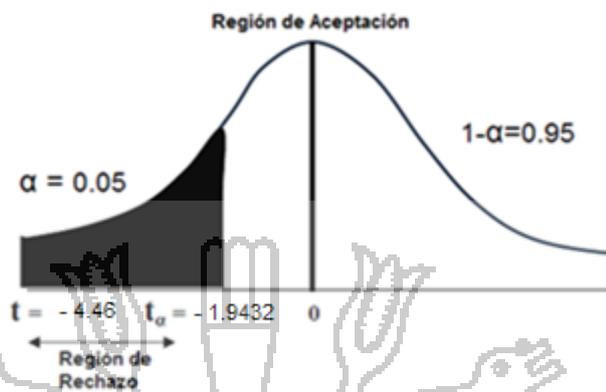
$$S_d = \sqrt{\frac{1.81}{7-1}} = 0.55$$

Calculo de t:

$$t = \frac{-0.93}{0.55/\sqrt{7}}$$

$$t = -4.46$$

e) Decisión: Nivel análisis de la información



$$t = -4.46(\text{Calculado}) < t_{0.05;6} = -1.9432(\text{Tabular})$$

Al comparar los valores, se rechaza la hipótesis nula (H_0) con un margen de error de 5% y aceptamos la hipótesis alternativa (H_1) por cual se determina que la usabilidad, las dimensiones y el tiempo de acceso al Datamart de Contrataciones públicas a partir del SEACE favorecen el nivel de análisis de información para la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno -2013.

4.1.3. DECISIÓN FINAL.

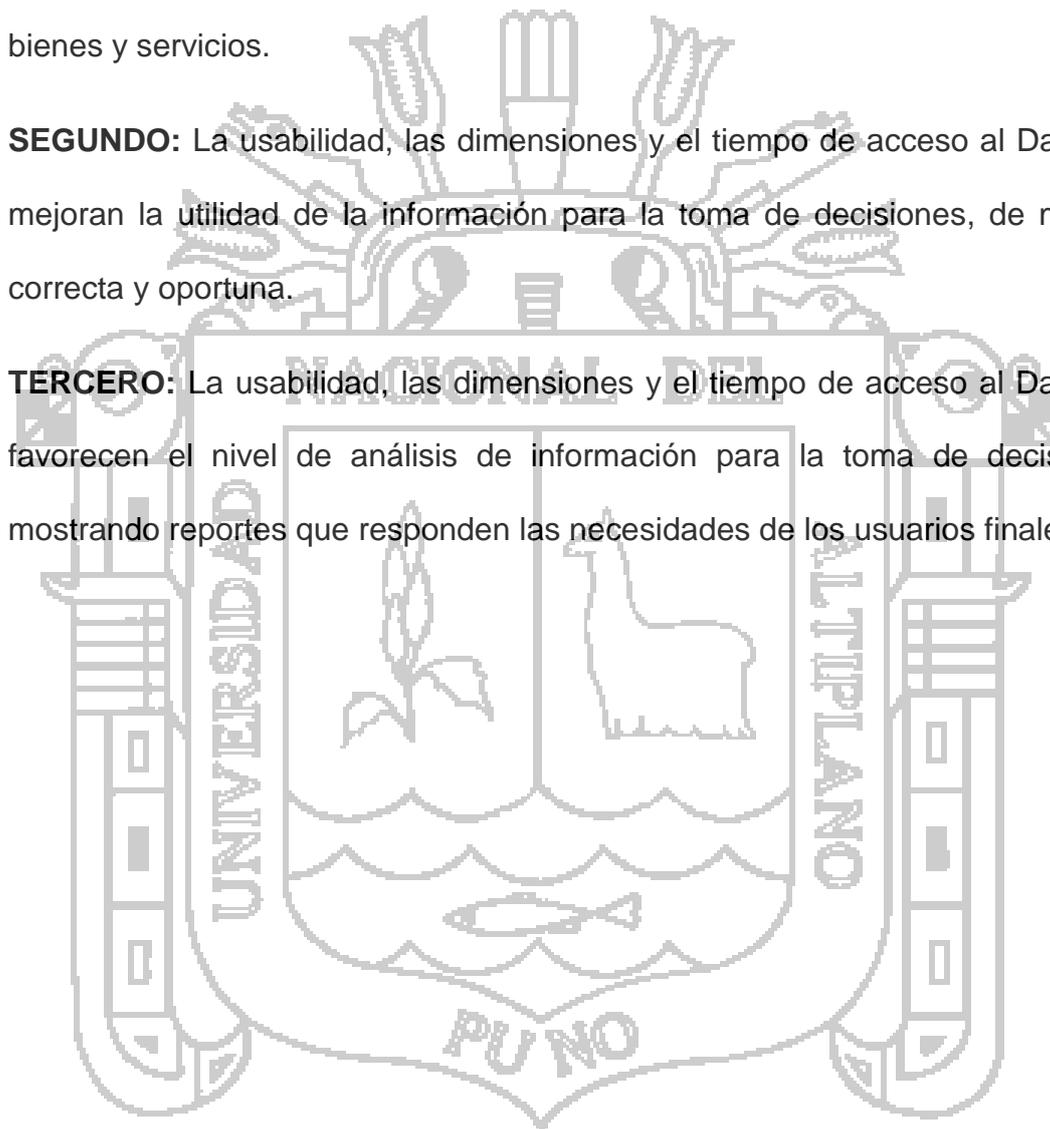
En los contrastes de las hipótesis estadísticas 1 y 2 se rechazó las hipótesis nulas, validando las hipótesis específicas del presente trabajo de investigación. Por lo tanto que aceptamos la hipótesis general de la investigación: La aplicación del Datamart de Contrataciones públicas a partir del SEACE apoya la toma de decisiones de las Micro y pequeñas Empresas de la Ciudad de Puno - 2013.

CONCLUSIONES

PRIMERO: Con la aplicación del Datamart Contrataciones públicas a partir del SEACE se logró demostrar que apoya la toma de decisiones de las MYPEs para contratar con el Estado, fundamentado en base de la información histórica de los procesos de selección realizados por las entidades públicas para proveerse de bienes y servicios.

SEGUNDO: La usabilidad, las dimensiones y el tiempo de acceso al Datamart mejoran la utilidad de la información para la toma de decisiones, de manera correcta y oportuna.

TERCERO: La usabilidad, las dimensiones y el tiempo de acceso al Datamart favorecen el nivel de análisis de información para la toma de decisiones, mostrando reportes que responden las necesidades de los usuarios finales.



RECOMENDACIONES

PRIMERO: El Datamart de Contrataciones públicas sólo representa la información externa o de entorno de una Micro y pequeña Empresa, se recomienda consolidar información interna para evaluar los objetivos y la capacidad de la Empresa para tomar mejores decisiones para contratar con el estado.

SEGUNDO: Para mejorar la utilidad de la información de las Contrataciones públicas se recomienda poblar el Datamart con información más reciente de los procesos de selección e incluir Plan Anual de Contrataciones (PAC) para conocer las futuras Contrataciones. Y establecer dimensiones con jerarquías y niveles necesarios para presentar información correcta y oportuna para la toma de decisiones.

TERCERO: Para alcanzar un buen nivel de análisis de información, se recomienda establecer indicadores y dimensiones según modelo de negocio y los requerimientos de los usuarios finales para reducir el tiempo de respuesta y mejorar la calidad de información.

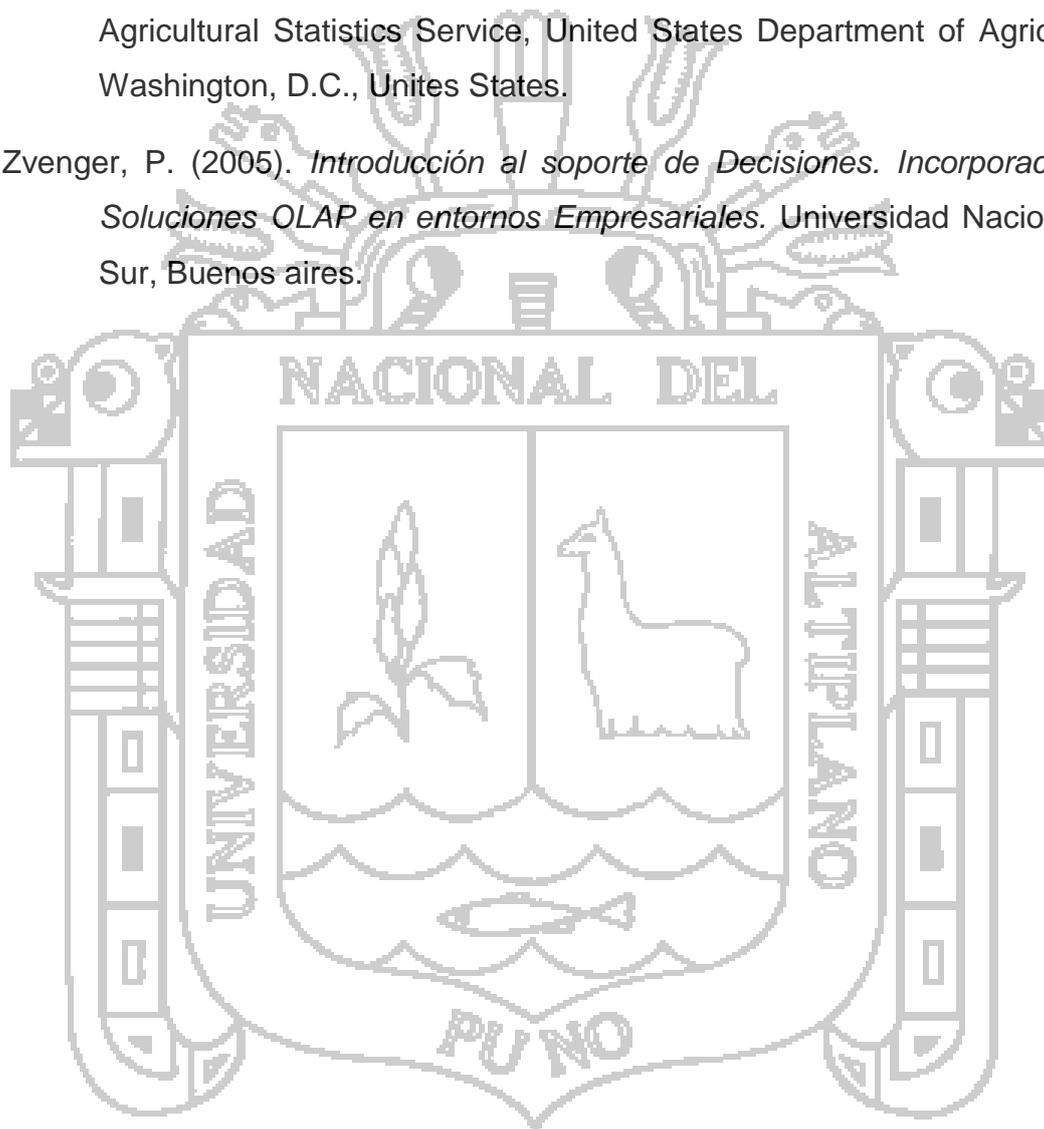
BIBLIOGRAFÍA

- Abad G., V. (2008). *Constitucion de MYPEs*. Lima: San Marcos.
- Bernabeu, D. (2007). *DATA WAREHOUSING: Investigacion y Sistematizacion de Conceptos*. Cordoba, Argentina.
- Bernabeu, D. (2010). *HEFESTO: Metodología para la Construcción de un Data Warehouse*. Cordoba, Argentina.
- Bustos, S. & Mosquera, V. (2013). *Análisis, Diseño e Implementación de una solución Business Intelligence para la generación de indicadores y control de desempeño, en la Empresa Otecel S.A., utilizando la metodología Hefesto v2.0*. Escuela Politécnica Del Ejército, Sangolquí.
- Córdoba, M. (2004). *Metodología para la toma de decisiones*. Madrid: Delta Publicaciones Universitarias.
- Forrester, J.W. (1994): "Policies, decisions and information sources for modeling". En John D. W. Morecroft and John D. Sterman (eds): *Modeling for learning organizations*. Portland, Oregon: Productivity Press.
- Gartner, G. (2012). Recuperado el 14 de Octubre de 2013, de <http://www.businessintelligence.info/mercado/gartner-businessintelligence-2012>
- Gestión de Proyectos. (2013). Recuperado el 12 de Setiembre de 2013, de http://www.larioja.org/upload/documents/458269_gestion_proyectos.pdf
- Gottlob, et al. (2010). *Web Data Extraction for Business Intelligence: the Lixto Approach*. Vienna, Austria.
- Gravitar. (2014). *Pentaho*. Recuperado el 10 de Setiembre de 2014, de <http://gravitar.biz/pentaho/>
- Gutiérrez, P. (2012). *Metodología de uso de Herramientas de Inteligencia de Negocios como Estrategia para aumentar la Productividad Y Competitividad de una PyME*. Instituto Politécnico Nacional, Mexico.
- Harjinder & Prakash. (1996). *Data Warehousing, la integración de información para la mejor toma de decisiones*. Mexico: Prentice Hall.

- Harjinder S., G. (1996). *Data Warehousing - La integración de información para la mejor toma de decisiones*. Mexico: Prentice Hall Hispanoamérica México.
- Hermoza, A. (2013). *Modelo de Datawarehouse para el sector Público (Caso Editora Perú S.A.)*. Lima, Peru.
- Huber, G. (1992). *Toma de decisiones en la gerencia*. Mexico D.F: Trillas S.A.
- Informe OSCE. (2013). *Contrataciones por Region*. Recuperado el 6 de Setiembre de 2013, de http://www.osce.gob.pe/documentos/regiones_2013.pdf
- Inmon, B. (2002). *Building the Data Warehouse*. (3ra Ed.). United States of America.
- Lane, P. (1999). *Oracle Data Warehousing Guide – Oracle Documentation*.
- Laudon, K. & Laudon, J. (2012). *Sistemas de Información Gerencial*(12da Ed.). Mexico: Pearson Educación.
- Le moigne, Jean-Louis. (1990). *La Modélisation des Systèmes Complexes*. Paris: Dunod.
- Ley de Promoción y Formalización de la Micro y Pequeña Empresa, Ley 28015. *Diario Oficial el Peruano*, N° 8411. Lima, jueves 3 de julio 2003.
- Ley de Contrataciones del Estado, Ley 1017. *Diario Oficial el Peruano*, N° 10252. Lima, miércoles 4 de junio del 2008.
- Ley General de Sociedades, Ley 26887. *Diario Oficial el Peruano*, N° 6374. Lima, martes 9 de diciembre de 1997.
- Mendoza, S.M. (2011). *Análisis, Diseño e Implementación de un Gerencial basado en un suite integrada de Datamarts para las áreas de Finanzas, Contabilidad, Recursos Humanos y Comercial*. Pontificia Universidad Católica del Perú, Lima.
- Menguzzato, M., Renau J. (1991). *La Dirección Estratégica de la Empresa: Un Enfoque Innovador del Management*. Barcelona: Ariel S.A.

- Mohammed, M., Mohammed, A., Anwar, M. & Ahme, H. (November de 2012). Meta-data and Data Mart solutions for better understanding for data and information in E-government Monitoring. *IJCSI International Journal of Computer Science Issues*, 9(3). Nasiriyah, Iraq.
- Moya, A. (2010). ETL con Talend Open Studio. Recuperado el 28 de Setiembre de 2014, de http://angelmoya.es/blog/otros_sistemas/etl-con-talend-open-studio/
- OSCE, (2013). Ley de Contrataciones del Estado y su Reglamento.
- Palomar, M. (2001). datos, Uso y diseño de Base de Datos Multidimensionales y almacenes de Datos. Recuperado el 2 de Octubre de 2013, de www.dlsi.ua.es/~jtrujillo/doctoradotema2bnpro.pdf
- Pentaho. Documentation. (2014). Recuperado el 12 de Setiembre de 2014, de <https://help.pentaho.com/Documentation/>
- Reinhardt, S. (2007). Reflexiones sobre la Tecnología de Información. Information Management. Recuperado el 07 de Diciembre de 2007, de <http://informationmanagement.wordpress.com/category/data-mart/>
- Rodriguez, K. y Mendoza, A. (2011). *Análisis Diseño e Implementación de una solución de Inteligencia de Negocios para el área de compras y ventas de una Empresa Comercializadora de Electrodomésticos*. Pontificia Universidad Católica del Perú, Lima.
- Sabana, M. (2006). Modelamiento e implementación de Base de Datos (1ra Ed.). Lima: Megabyte.
- Satyanarayana, R., Rallabandi, S., Poorna, Ch., Y Srikanth, R. (2010). Data Warehousing, Data Mining, Olap and OLTP Technologies are essential elements to support decision-Daking Process In Industries (Vol. 02). Hyderabad, College of Engg. & Technology, India.
- Shackle, G. (1966). Decisión, Orden y Tiempo. Madrid: Tecnos.
- Simon, H. A. (1960). The New Science of Management Decision. New York: Harper & Row. En Laudon, K. & Laudon, J. (2012). Sistemas de Información Gerencial (12da Ed.). Mexico: Pearson Educación.

- Tamayo, M. (2002). El proceso de la investigación científica. Mexico: Limusa S.A.
- Trak, A. y Zambrano, E. (2002). *Elaboración de un Data Mart de clientes PYME, a las unidades de mercadeo de CANTV Corporación*. Universidad Católica Andrés Bello, Caracas.
- Yost, M. (2004). Data Warehousing and Decision Support at the National Agricultural Statistics Service, United States Department of Agriculture. Washington, D.C., Unites States.
- Zvenger, P. (2005). *Introducción al soporte de Decisiones. Incorporación de Soluciones OLAP en entornos Empresariales*. Universidad Nacional del Sur, Buenos aires.



ANEXOS

ANEXO 01: CUESTIONARIO PARA REQUERIMIENTOS.

CUESTIONARIO PARA REQUERIMIENTO ENCUESTADO: _____ FECHA: <u>12/12/13</u>	
1.	¿Qué aplicaciones utiliza para buscar información acerca de las contrataciones del estado? <i>Portal SEACE.</i> <i>Portal Ouce mype.</i>
2.	¿Se encuentra en la capacidad de responder rápidamente las condiciones del mercado y asegurar una contratación con el estado? <i>No, porque la información está presentada en forma desordenada</i>
3.	¿Qué información clave se requiere para hacer o soportar las decisiones que usted toma en el proceso de alcanzar las metas y sobreponerse a los obstáculos? ¿Cómo obtiene esta información en este momento, con respecto a las contrataciones públicas? <i>Precio unitarios, Precio estimado por mes, Demanda de Bienes, Información de la competencia</i>
4.	¿Cuáles son los reportes que más le ayudan para participar en las contrataciones públicas? ¿Qué datos en el reporte son importantes? <i>Búsqueda de procesos en el portal SEACE, los datos importantes son la entidad convocante y el precio referencial</i>
5.	¿Qué información considera para visualizar oportunidades que no son aprovechadas en este momento, con respecto a las contrataciones públicas? <i>El comportamiento de los precios en el tiempo, la cantidad del producto solicitado.</i>
6.	¿Qué tan difícil es en este momento conseguir esa información? ¿Cómo debería presentarse, de qué forma? ¿Qué tan frecuente la necesita: Diario, Semanal, Mensual, Anual? <i>La información existe pero se hace difícil la revisión de estos procesos, la información es necesaria siempre.</i>
7.	¿Qué capacidades analíticas le gustaría tener, sobre la información acerca de las contrataciones públicas? <i>Reporte de licitaciones por producto, por meses por</i>

ANEXO 02: CUESTIONARIO PRE TEST.

CUESTIONARIO PARA PRE TEST ENCUESTADO: _____ FECHA: <u>18-12-13</u>	
SOBRE SEACE	
1. ¿Cuál es el grado de usabilidad del SEACE para la toma de decisiones para contratar con el estado?	a) Muy Buena b) Buena c) Regular <input checked="" type="checkbox"/> d) Mala e) Muy Mala
2. ¿Cómo considera los reportes del SEACE para el análisis desde diferentes ópticas para la toma de decisiones?	a) Muy Buena b) Buena <input checked="" type="checkbox"/> c) Regular d) Mala e) Muy Mala
3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones?	a) Menos de 15 min b) Entre 15 y 30 min c) Entre 30 y 45 min <input checked="" type="checkbox"/> d) Entre 45 y 60 min e) Más de 1 hora
SOBRE LA TOMA DE DECISIONES PARA CONTRATAR CON EL ESTADO	
4. ¿Cómo considera las contrataciones públicas históricas para la toma de decisiones, disponibles en el SEACE?	a) Muy Buena b) Buena c) Regular <input checked="" type="checkbox"/> d) Mala e) Muy Mala
5. ¿Cómo considera la presentación de la información estructurada del SEACE en contratación de bienes y servicios, para la toma de decisiones?	a) Muy Buena b) Buena c) Regular <input checked="" type="checkbox"/> d) Mala e) Muy Mala
6. La información que se brinda sobre su participación de las contrataciones públicas a través del SEACE es:	a) Muy Buena b) Buena c) Regular <input checked="" type="checkbox"/> d) Mala e) Muy Mala

7. ¿Cómo considera la información de las contrataciones publicadas sobre oportunidades, y alternativas de negocios en el SEACE?
- a) Muy Buena
b) Buena
 Regular
d) Mala
e) Muy Mala
8. El nivel de información que ofrece el SEACE sobre cantidad, precios y montos adjudicados de un bien o servicio es:
- a) Muy Buena
b) Buena
 Regular
d) Mala
e) Muy Mala
9. ¿Cómo considera la información que brinda el SEACE con respecto a montos totales, cantidades, y precios unitarios adjudicados a un proveedor en el tiempo?
- a) Muy Buena
b) Buena
c) Regular
 Mala
e) Muy Mala



ANEXO 03: CUESTIONARIO POST TEST.

CUESTIONARIO PARA POST TEST	
ENCUESTADO: _____	
FECHA: <u>21-07-17</u>	
SOBRE DATAMART.	
1. ¿Cuál es el grado de usabilidad del DATMART para la toma de decisiones para contratar con el estado?	
<input checked="" type="radio"/> a) Muy Buena <input type="radio"/> b) Buena <input type="radio"/> c) Regular <input type="radio"/> d) Mala <input type="radio"/> e) Muy Mala	
2. ¿Cómo considera los reportes del DATAMART para el análisis desde diferentes ópticas para la toma de decisiones?	
<input type="radio"/> a) Muy Buena <input checked="" type="radio"/> b) Buena <input type="radio"/> c) Regular <input type="radio"/> d) Mala <input type="radio"/> e) Muy Mala	
3. ¿Qué tiempo estima para obtener información de contratación de bienes y servicios de su interés para la toma de decisiones, con la aplicación DATAMART?	
<input type="radio"/> a) Menos de 15 min <input checked="" type="radio"/> b) Entre 15 y 30 min <input type="radio"/> c) Entre 30 y 45 min <input type="radio"/> d) Entre 45 y 60 min <input type="radio"/> e) Más de 1 hora	
SOBRE LA TOMA DE DECISIONES PARA CONTRATAR CON EL ESTADO	
4. ¿Cómo considera las contrataciones públicas históricas para la toma de decisiones, disponibles en el DATAMART?	
<input type="radio"/> a) Muy Buena <input type="radio"/> b) Buena <input checked="" type="radio"/> c) Regular <input type="radio"/> d) Mala <input type="radio"/> e) Muy Mala	
5. ¿Cómo considera la presentación de la información estructurada del DATAMART en contratación de bienes y servicios, para la toma de decisiones?	
<input type="radio"/> a) Muy Buena <input type="radio"/> b) Buena <input checked="" type="radio"/> c) Regular <input type="radio"/> d) Mala <input type="radio"/> e) Muy Mala	
6. La información que se brinda sobre su participación de las contrataciones públicas a través del DATMART es:	
<input checked="" type="radio"/> a) Muy Buena <input type="radio"/> b) Buena <input type="radio"/> c) Regular <input type="radio"/> d) Mala <input type="radio"/> e) Muy Mala	

7. ¿Cómo considera la información de las contrataciones publicadas sobre oportunidades, y alternativas de negocios en el DATAMART?

- a) Muy Buena
 b) Buena
c) Regular
- d) Mala
e) Muy Mala

8. El nivel de información que ofrece el DATAMART sobre cantidad, precios y montos adjudicados de un bien o servicio es:

- a) Muy Buena
 b) Buena
c) Regular
- d) Mala
e) Muy Mala

9. ¿Cómo considera la información que brinda el DATAMART con respecto a montos totales, cantidades, y precios unitarios adjudicado a un proveedor en el tiempo?

- a) Muy Buena
b) Buena
c) Regular
- d) Mala
e) Muy Mala





ANEXO 04: TABLA T-STUDENT.

<i>r</i>	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.995
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

ANEXO 05: EXTRACCIÓN DE DATOS DEL SEACE.

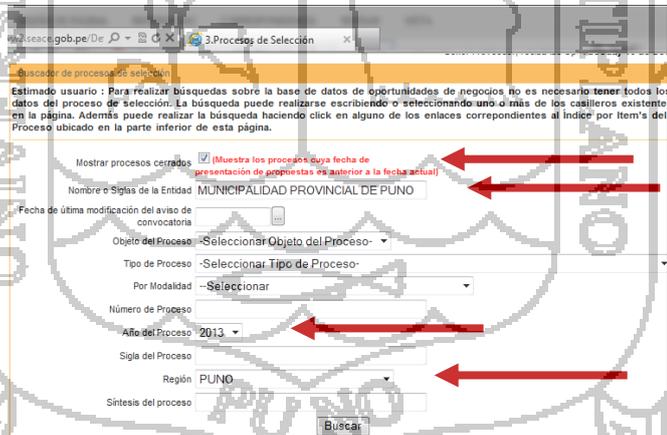
En este anexo se muestra la extracción de datos del SEACE sobre los procesos de selección llevados por las entidades públicas de la Ciudad de Puno durante el año 2013, para esta actividad se sigue los siguientes pasos:

Paso 01: Se accede a la página principal del SEACE: <http://www2.seace.gob.pe/> y se selecciona la opción **consultas SEACE v2.0**.



Una vez seleccionado esta opción, se muestra una lista de alternativas para la búsqueda de procesos de selección. Se elige la opción **Oportunidades de Negocio**

Paso 02: En esta sección se encuentra un formulario que sirve para buscar procesos de selección convocados por las entidades del Estado, se tiene completar los parámetros que requiere el sistema y luego presionar el botón **Buscar**.



- Mostrar procesos cerrados: Marcado.
- Nombre o Sigla de la Entidad: MUNICIPALIDAD PROVINCIAL DE PUNO, GOBIERNO REGIONAL DE PUNO SEDE CENTRAL Y UNIVERSIDAD NACIONAL DEL ALTIPLANO.
- Año de proceso: 2013.
- Región: PUNO.

La búsqueda de los procesos de selección se realiza por separado para cada entidad convocante.

El resultado es una relación de procesos de selección que cumplen con los parámetros ingresados anteriormente. En el detalle de los procesos de selección se presenta la ficha de proceso de selección que pertenece a cada una de los procesos de selección convocados por una entidad del Estado. Para este caso de estudio sólo se consideran procesos de selección con **Buena pro consentida**.

Proceso Electrónico
Convocatoria
 AMC PROCEDIMIENTO CLASICO 3-2013/MPP (convoc... MENOR CUANTIA
 en la modalidad de CONVOCATORIA

Entidad Contratante
 MUNICIPALIDAD PROVINCIAL DE PUNO (MPP)
 (Teléf : 051-301593)

Dirección
 JR. DEUSTUA N° 468-PUNO (PUNO)

Síntesis
 servicio de impresion

Objeto
 SERVICIOS

Valor Referencial
 Soles : 19,995,01

Derecho de participar
 Gratuito

Costo de reproducción de las bases
 Soles : 10,00

Lugar de registro de participantes

Informado el día 27/02/2013 20:05 horas

Inscribirse en proceso electrónico

Calendario		
Actividad	Fecha Inicio	Fecha Fin
convocatoria	27/02/2013 00:00	27/02/2013 00:00
registro de participantes	28/02/2013 00:01	05/03/2013 14:00
presentación de propuestas	28/02/2013 00:01	05/03/2013 14:00
calificación y evaluación de propuestas	05/03/2013 00:00	05/03/2013 00:00
otorgamiento de la buena pro	05/03/2013 00:00	05/03/2013 00:00

ULTIMA ETAPA : OTORGAMIENTO DE LA BUENA PRO el 05/03/2013, 1 item en BUENA PRO CONSENTIDA.

Otras Acciones de la Convocatoria

- 1** Fe de Errata, informado el día 27/02/2013 20:12, se precisa que el registro de participantes, presentación de propuestas y buena es hasta 01 de marz
- 2** Postergación, informado el día 01/03/2013 13:09, PARA DAR MAYOR COBERTURA A LOS POSTORES SEGUN LOS PRINCIPIOS DE CONTRATACIONES PUBLICAS
- 3** INSTRUMENTO QUE APRUEBA O AUTORIZA LA REALIZACION DEL PROCESO, PROVEIDO N° 331, , informado el día 27/02/2013 19:47
- 4** ACTA DE EVALUACION TECNICA, , , informado el día 06/03/2013 09:55
- 5** ACTA DE BUENA PRO, , , informado el día 06/03/2013 12:24
- 6** CONTRATO U ORDEN DE COMPRA O DE SERVICIO, 451616577radFDC92.pdf, , informado el día 02/04/2013 18:59

Bases	Resumen Ejecutivo	Absolución de Consultas y Observaciones	Absolución Consultas	Absolución Observaciones	Pronuncia. Entidad	Pronuncia. OSCE	Bases Integradas	Acta de Evaluación Técnica	Cuadro Comparativo y/o Actas	Acta Buena Pro	Notificación Suscripción de Acuerdo	Resolución de Recurso de Apelación de la Entidad	Resolución de Recurso de Apelación del Tribunal
27/02/2013 19:47 Tamaño 5,590 Kb.							08/03/2013 09:55 Tamaño 205 Kb.	08/03/2013 12:24 Tamaño 369,872 Kb.	08/03/2013 12:24 Tamaño 305,386 Kb.				

Item	Descripción	Familia Catálogo	Cantidad	Unidad de Medida	Valor Referencial	Plazo(s) de entrega	Garantía de seriedad de o	Beneficio de la ley 27633	MYPE	Cantidad Adjudicada	Monto Adjudicado
1	SERVICIO DE IMPRESION DE TICKET Y BOLETOS (10013262646) CHOQUE DE GONZALES LUCILA	A821215	1.00	Paquete	19,995,01	5 DIAS				1	19,000,00

A partir de la ficha de proceso selección, se selecciona los datos para el análisis posterior. Los cuales son:

1. Proceso de selección: Describe el tipo, modalidad y el número de proceso de selección.
2. Entidad contratante: Muestra la entidad del Estado que con convocó el procesos de selección.
3. Descripción: Muestra el bien o servicio que se contrata mediante un proceso de selección.

4. Unidad de medida: Presenta el tipo de unidad de medida del bien o servicio.
5. Proveedor: Empresa ganador del procesos de selección.
6. Cantidad adjudicado: Representa la cantidad del bien o del servicio contratado.
7. Monto adjudicado: Representa la monto total con que se contrata los bienes o servicios.
8. Buena pro consentida: Fecha en la que se publica el ganador del proceso de selección para la contratación de un bien o servicio.

Paso 03: Una vez identificado los datos a recuperar de la web, el objetivo es leer y almacenar datos automáticamente de cada procesos de selección llevados por las tres entidades públicas mencionadas anteriormente. Para ello se acuden a la herramienta **import.io** para que la lectura de datos desde la web sea automática.

En la aplicación **import.io** se configura la página de donde se extrae los datos. Para ello se ingresa la dirección de la página, en este caso la dirección de la ficha de un proceso de selección.



The screenshot shows the import.io interface with a browser window displaying the SE@CE website. The website header includes 'SE@CE Sistema Electrónico de Contrataciones del Estado' and 'OSCE Organismo Supervisor de las Contrataciones del Estado'. The main content area shows a 'Convocatoria' for 'AMC PROCEDIMIENTO CLASICO 3-2013/MPP (convocatoria : 1) MENOR CUANTIA en la modalidad de CONVOCATORIA' with details for the contracting entity 'MUNICIPALIDAD PROVINCIAL DE PUNO (MPP)'. A calendar table lists activities and their dates:

Actividad	Fecha Inicio	Fecha Fin
convocatoria	27/02/2013 00:00	27/02/2013 00:00
registro de participantes	28/02/2013 00:01	05/03/2013 14:00
presentación de propuestas	28/02/2013 00:01	05/03/2013 14:00
calificación y evaluación de propuestas	05/03/2013 00:00	05/03/2013 00:00
otorgamiento de la buena pro	05/03/2013 00:00	05/03/2013 00:00

Below the calendar, there is a red alert: 'ULTIMA ETAPA : OTORGAMIENTO DE LA BUENA PRO el 05/03/2013, 1 item en BUENA PRO CONSENTIDA.' A dialog box is overlaid on the page, asking 'OK, let's extract some data!' and 'Go to an example of the type of page you want to extract data from' with a pink 'I'm there!' button.

En la imagen se muestra la ficha de proceso de selección cargada en la aplicación **import.io**, y esta se utiliza como base para indicar las secciones de donde se extraen los datos, las fichas de los procesos de selección tiene un único formato por lo que esta acción se realiza una sola vez. Se presiona en la opción **I'm there!** para continuar la configuración, confirmamos que la detección de datos está completa presionando **Detect optimal settings**. Los datos de cada una de las fichas de los procesos de selección representan una sola fila del conjunto de procesos de selección realizadas por las entidades del Estado, para ello seleccionamos la opción **single**.

Síntesis
 servicio de impresion
Objeto
 SERVICIOS
Valor Referencial
 Soles : 19,995.01
Derecho de participar
 Gratuito
Costo de reproducción de las bases
 Soles : 10.00
Lugar de registro de participantes
 Informado el día 27/02/2013 20:05 horas

ULTIMA ETAPA : OTORGAMIENTO DE LA BUENA PRO el 05/03/2013, 1 ítem en BUENA PRO CONSENTIDA,
 Otras Acciones de la Convocatoria
1 Fe de Errata, informado el día 27/02/2013 20:12, se precisa que el registro de participantes, presentación de propuestas y buena es hasta 01 de marzo
2 Postergación, informado el día 01/03/2013 13:09, PARA DAR MAYOR COBERTURA A LOS POSTORES SEGUN LOS PRINCIPIOS DE CONTRATACIONES PUBLICAS
3 INSTRUMENTO QUE APRUEBA O AUTORIZA LA REALIZACION DEL PROCESO, PROVEIDO N° 331, , informado el día 27/02/2013 19:47

Rows per page
 Single: one thing to extract on each page, e.g. a product page
 Multiple: here lots of the same thing?

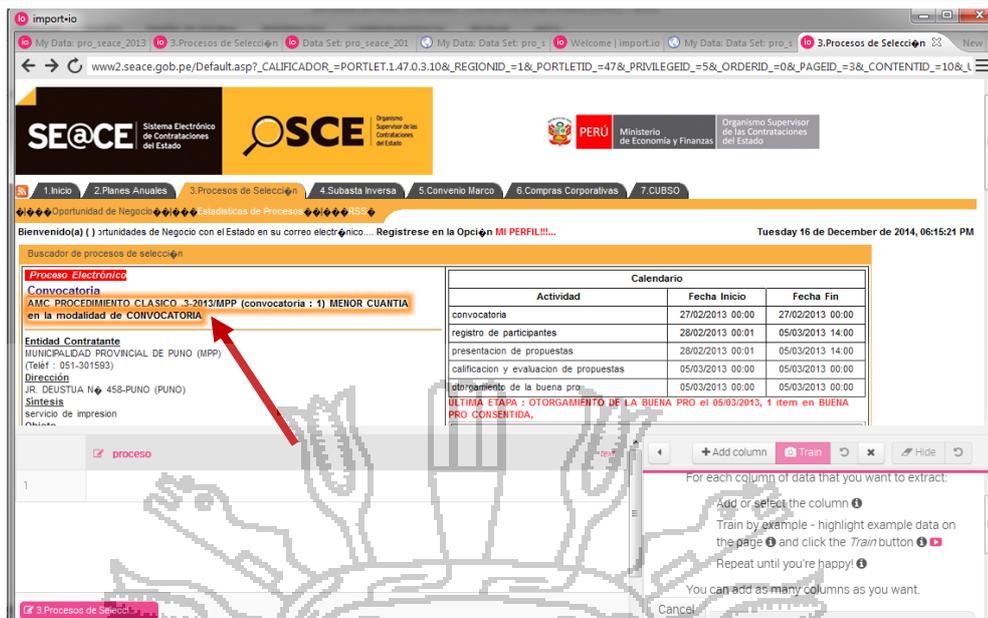
Con esta configuración se deja listo para seleccionar los campos que se desean recuperar desde la ficha. Se crea las columnas con la opción **add column**, se le asigna un nombre y se selecciona el tipo de dato.

Calendario		
Actividad	Fecha Inicio	Fecha Fin
convocatoria	27/02/2013 09:00	27/02/2013 00:00
registro de participantes	28/02/2013 09:01	05/03/2013 14:00
presentación de propuestas	28/02/2013 00:01	05/03/2013 14:00
calificación y evaluación de propuestas	05/03/2013 00:00	05/03/2013 00:00
otorgamiento de la buena pro	05/03/2013 00:00	05/03/2013 00:00

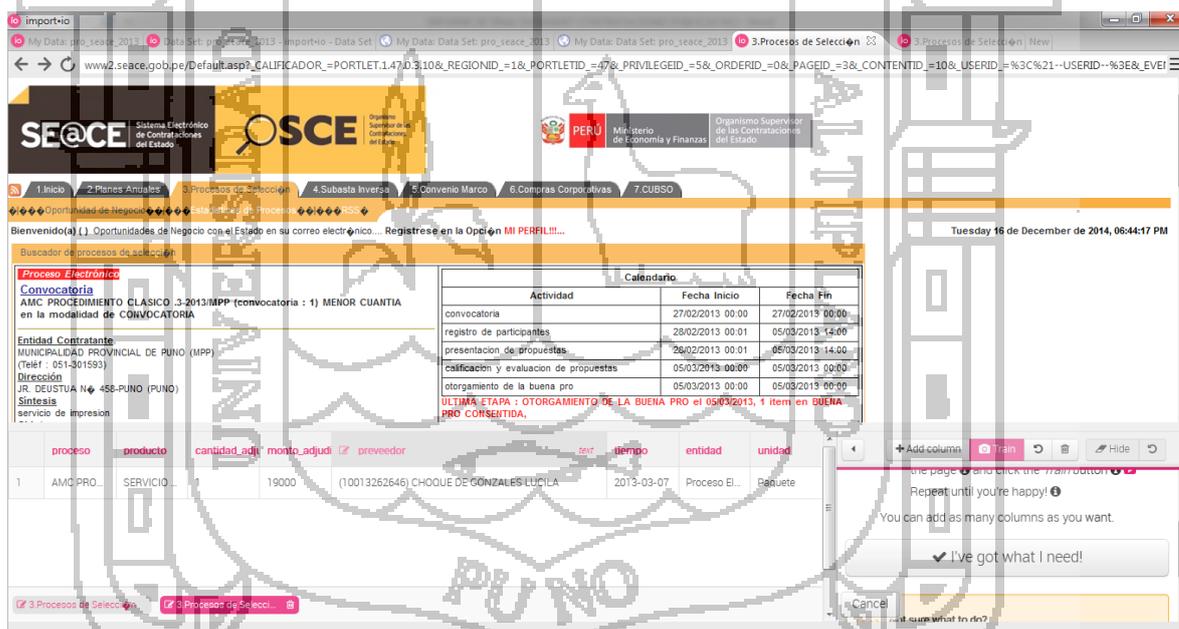
+ Add column
 For each column of data that you want to extract:
 Add or select the column
 Train by example - highlight example data on the page and click the Train button
 Repeat until you're happy!
 You can add as many columns as you want.
 Cancel

proceso
 Set default value (optional)
 Show advanced settings
 Text
 Date/Time
 Currency
 Whole number
 Number
 Language
 Country
 Boolean
 Image
 Link
 HTML
 Done

En este caso creamos la columna 'proceso' con tipo de dato 'Text' y terminamos el procedimiento presionando **Done**. A continuación se indica la parte de la ficha del proceso de selección que representa el 'proceso' para el caso de estudio.



Se completa la selección presionando el botón **Train**. El procedimiento es similar para el resto de los datos seleccionados y descritos anteriormente, teniendo en cuenta el tipo de dato que representa y se muestra en la siguiente imagen el resultado.



Para que los datos seleccionados sean los esperados se realiza la validación de la selección de campos, para ello se agrega otras fichas de procesos de selección y se realiza la extracción presionando la opción **I've got what I need**. Y se termina el proceso con la publicación el extractor con la opción **Upload to import.io** a partir de ella se realiza la extracción de datos en paralelo con sólo alimentación de las direcciones de las fichas de procesos de selección.

	Proceso	Producto	Cantidad adjudica...	Monto adjudicado	Proveedor	Tiempo	Entidad
1	AMC PROCEDIMIENTO...	adquisicion de suminis...	1	29516	(20145496170) UNIVE...	2013-12-04 00:00	Convoc...
2	AMC PROCEDIMIENTO...	SERVICIO DE IMPRESI...	1	19000	(10013262646) CHOQ...	2013-03-07 00:00	Proces...

Waiting for www.woopra.com...

La aplicación import.io tiene la posibilidad de exportar los datos a MS Excel, este archivo se utiliza para la transformación y posterior carga al Datamart.



ANEXO 06: TRANSFORMACIÓN Y CARGA DE DATOS AL DATAMART.

En esta sección se detalla la transformación y limpieza de datos, esta operación se realiza a partir del archivo Ms Excel generado en el anterior sección, que contiene la lista de procesos de selección de las tres entidades públicas de la Ciudad de Puno, durante el año 2013. Para la carga de datos se debe tener creado el Datamart, presentado en el capítulo III del trabajo de investigación.

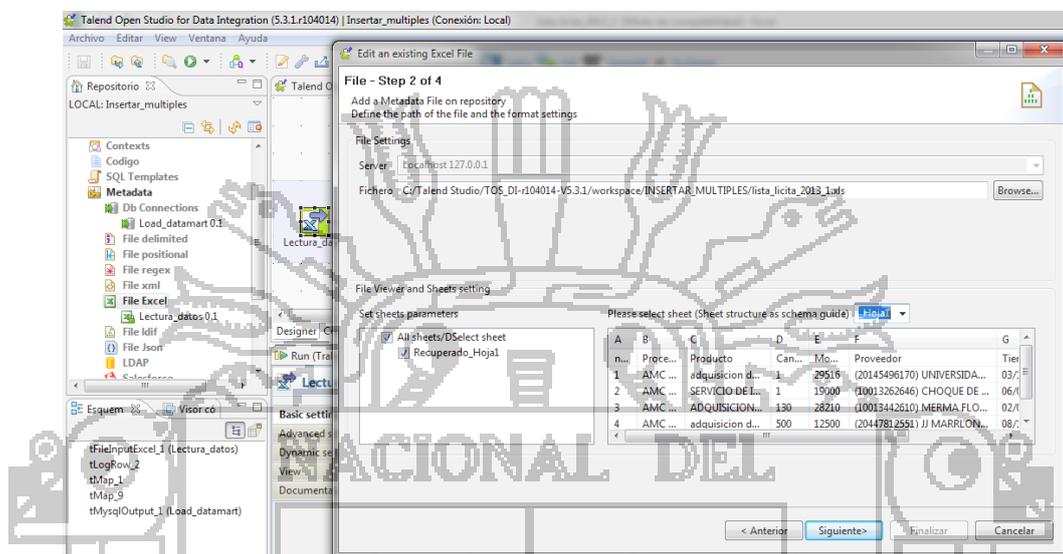
numero	Proceso	Producto	Cantidad adjudicada	Monto adjudicado	Proveedor	Tiempo	Entidad
1	AMC PROCEDIMIENTO CLASICO -89-2013/MPP (convocatoria: 1) MENOR CUANTIA en la modalidad de CONVOCATORIA	adquisición de suministro de diesel b5 s 50 y gasohol 84 octanos	1	29516	(20145436170) UNIVERSIDAD NACIONAL DEL ALTIPLANO	03/12/2013	Convocatoria AMC PROCEDIMIENTO CLASICO -89-2013/MPP (convocatoria: 1) MENOR CUANTIA en la modalidad de CONVOCATORIA Entidad Contratante MUNICIPALIDAD PROVINCIAL DE (MPP) (Teléf. -051-3014593) Dirección
2	AMC PROCEDIMIENTO CLASICO -3-2013/MPP (convocatoria: 1) MENOR CUANTIA en la modalidad de CONVOCATORIA	SERVICIO DE IMPRESION DE TICKET Y BOLETOS	1	19000	(10013262646) CHOQUE DE GONZALES LUCILA	06/03/2013	Proceso Electrónico Convocatoria 1 PROCEDIMIENTO CLASICO -3-20 (convocatoria: 1) MENOR CUANTIA modalidad de CONVOCATORIA Entidad Contratante MUNICIPALIDAD PROVINCIAL DE (MPP) (Teléf. -051-3014593) Dirección
3	AMC PROCEDIMIENTO CLASICO -18-2013/MPP (convocatoria: 3) MENOR CUANTIA en la modalidad de CONVOCATORIA	ADQUISICION DE PIEDRA LAJA PARA PISO	130	28210	(10013442610) MERMA FLORES OLENKA GAILY	02/06/2013	Convocatoria AMC PROCEDIMIENTO CLASICO -18-2013/MPP (convocatoria: 3) MENOR CUANTIA en la modalidad de CONVOCATORIA Entidad Contratante MUNICIPALIDAD PROVINCIAL DE (MPP) (Teléf. -051-3014593) Dirección
4	AMC PROCEDIMIENTO CLASICO -90-2013/MPP (convocatoria: 1) MENOR CUANTIA en la modalidad de CONVOCATORIA	adquisición de hormigon	500	12500	(20447812551) JJ MARLON EJECUTORES Y SERVICIOS MULTIPLES EMPRESA INDIVIDUAL DE RESPONSABILIDAD LIMITADA	08/12/2013	Convocatoria AMC PROCEDIMIENTO CLASICO -90-2013/MPP (convocatoria: 1) MENOR CUANTIA en la modalidad de CONVOCATORIA Entidad Contratante MUNICIPALIDAD PROVINCIAL DE (MPP) (Teléf. -051-3014593) Dirección

En la imagen se presenta los datos obtenidos desde el SEACE, cada fila de la representa un proceso de selección y cada columna representa una Dimensión del Datamart. Estos datos no tienen formato para ser almacenado directamente, es necesario separar los datos según los campos de las Dimensiones del Datamart:

- Dimproceso
 - Tipo de proceso
 - Modalidad de proceso
 - Numero
 - Convocatoria
- Dimproducto
 - Descripción
 - Unidad
- Dimproveedor
 - RUC
 - Razón social
- Dimentidad
 - Nombre

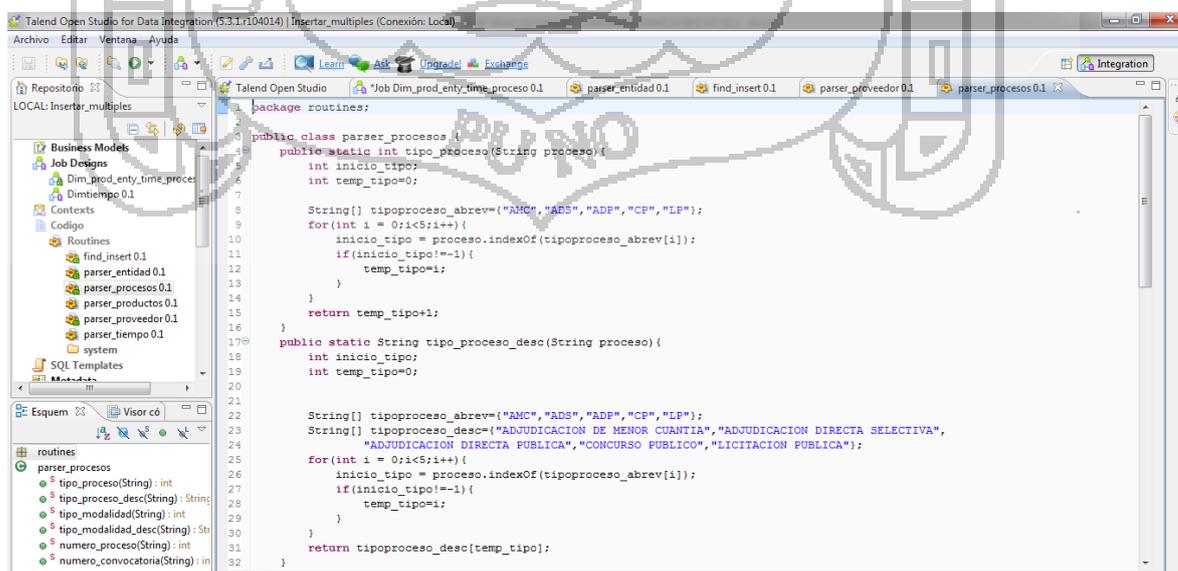
Para realizar este procedimiento se realiza los siguientes pasos.

Paso 01: Para realizar esta actividad es necesario contar con herramientas de ETL, en este caso se utiliza **Talend open studio 5.3**, se inicia con la configuración del componente para la entrada de flujo de datos, ingresando la ruta de la ubicación del archivo y asignando un nombre.

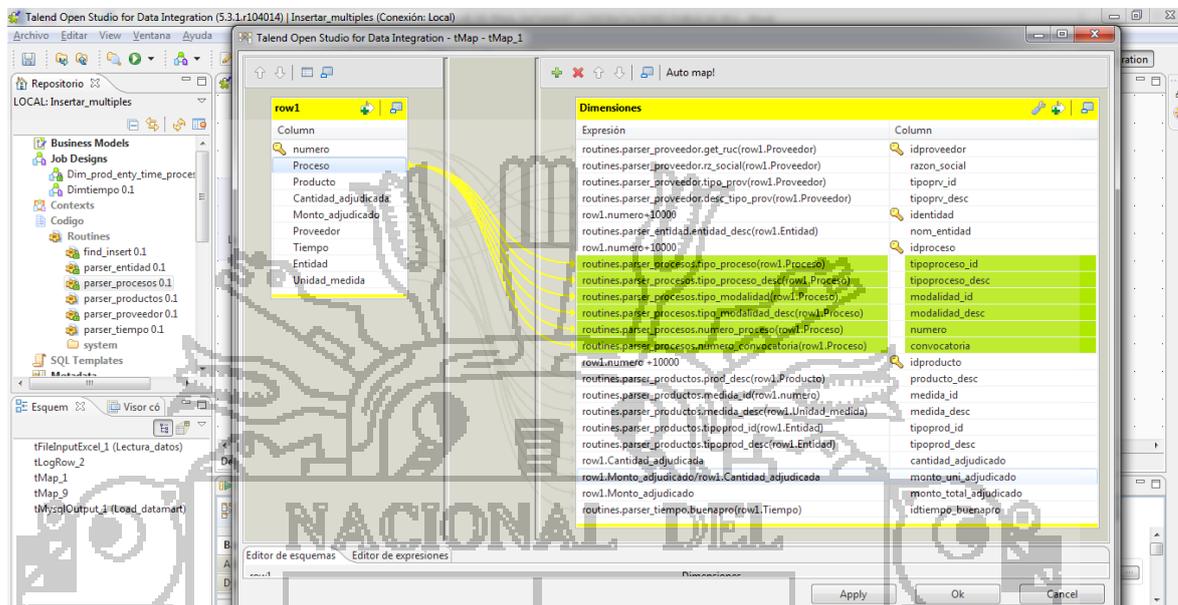


Una vez configurado la entrada, se organiza el flujo de datos para realizar la transformación.

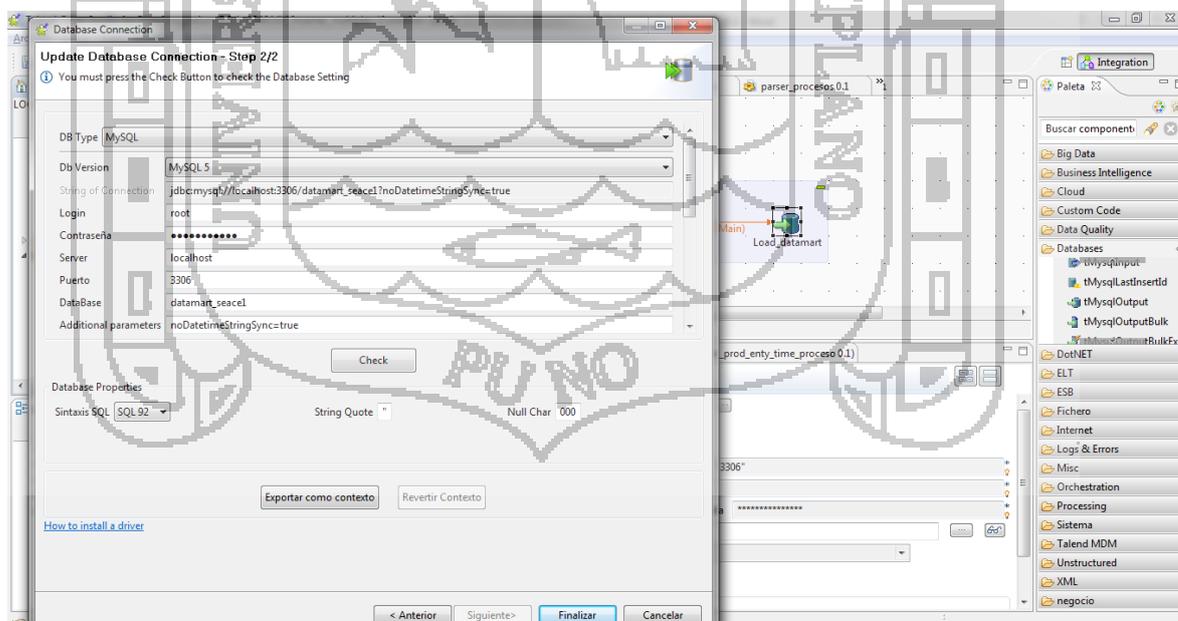
Paso 02. Para la transformación de datos se desarrolla funciones específicas para cada caso, para este propósito se utiliza lenguaje de programación java.



Las funciones devuelven fracciones de cadena que tiene formato para ser almacenados en las tablas del Datamart. Una vez desarrollado funciones de transformación y formateo se realiza la asociación de campos con la entrada y salida con el componente t_Map del herramienta Talend y se invoca a las funciones de transformación.



Paso 03. Una vez formateado y estandarizado los datos se almacenan los datos en el Datamart con el componente tMysqlOutput y termina el flujo de datos, para ello se proporcionan los parámetros de conexión al gestor de base de datos MySQL.



Y como resultado se muestra una parte del contenido de las tablas que representan las dimensiones del Datamart.

Dimproceso:



	idproceso	tipoproceso_id	tipoproceso_desc	modalidad_id	modalidad_desc	numero	convocatoria
<input type="checkbox"/>	10001	1	ADJUDICACION DE MENOR CUANTIA	1	PROCEDIMIENTO CLASICO	89	1
<input type="checkbox"/>	10002	1	ADJUDICACION DE MENOR CUANTIA	1	PROCEDIMIENTO CLASICO	3	1
<input type="checkbox"/>	10003	1	ADJUDICACION DE MENOR CUANTIA	1	PROCEDIMIENTO CLASICO	18	3
<input type="checkbox"/>	10004	1	ADJUDICACION DE MENOR CUANTIA	1	PROCEDIMIENTO CLASICO	90	1

Dimproducto:

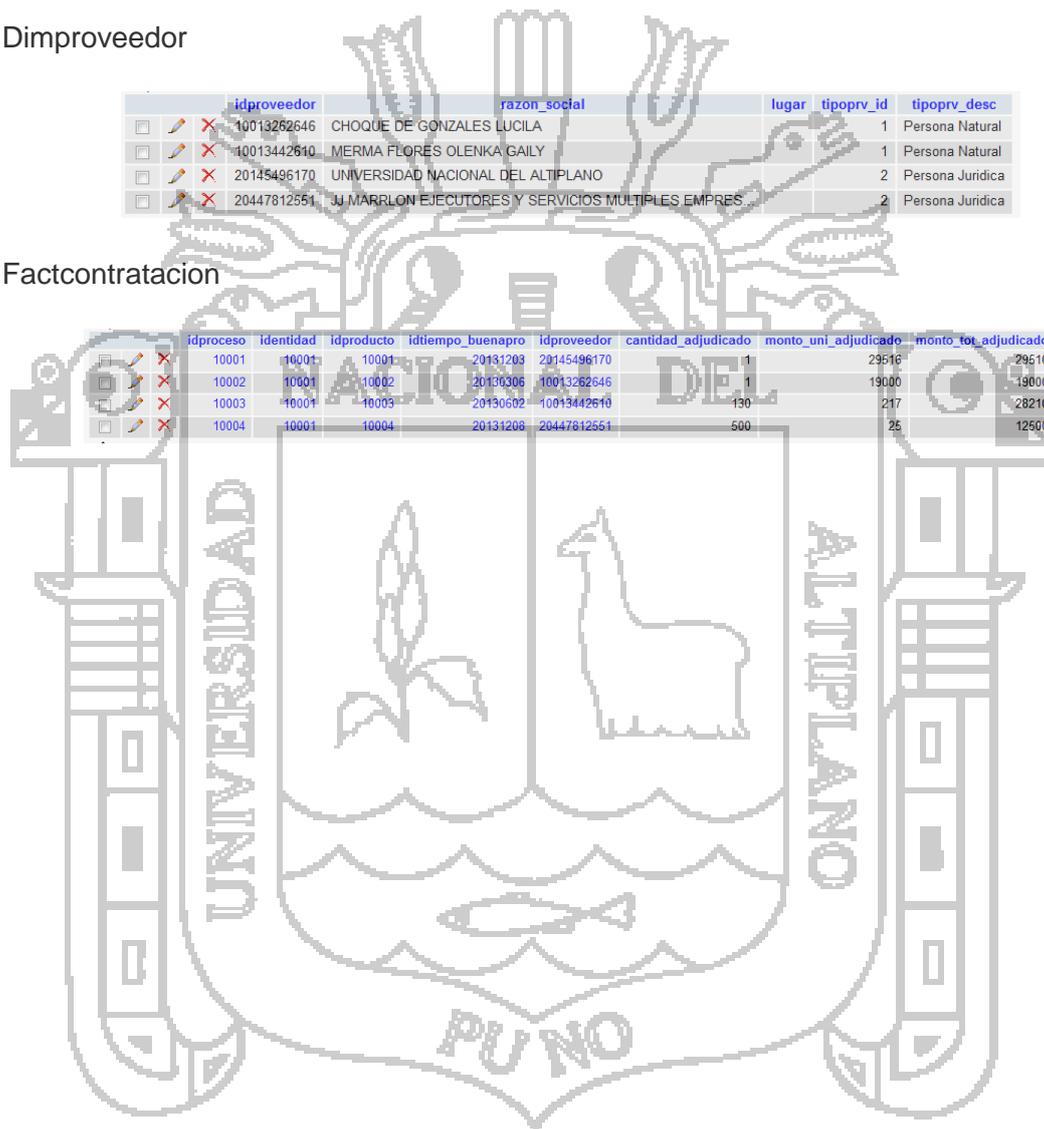
	idproducto	producto_desc	medida_id	medida_desc	tipoprod_id	tipoprod_desc
<input type="checkbox"/>	10001	DIESEL B5 S 50 Y GASOHOL 84 OCTANOS	101	PAQUETE	1	BIENES
<input type="checkbox"/>	10002	SERVICIO DE IMPRESION DE TICKET Y BOLETOS	102	PAQUETE	2	SERVICIOS
<input type="checkbox"/>	10003	PIEDRA LAJA	103	M2	1	BIENES
<input type="checkbox"/>	10004	HORMIGON	104	M3	1	BIENES

Dimproveedor

	idproveedor	razon_social	lugar	tipopriv_id	tipopriv_desc
<input type="checkbox"/>	10013262646	CHOQUE DE GONZALES LUCILA		1	Persona Natural
<input type="checkbox"/>	10013442610	MERMA FLORES OLENKA GAILY		1	Persona Natural
<input type="checkbox"/>	20145496170	UNIVERSIDAD NACIONAL DEL ALTIPLANO		2	Persona Juridica
<input type="checkbox"/>	20447812551	UJ MARRLON EJECUTORES Y SERVICIOS MULTIPLES EMPRES		2	Persona Juridica

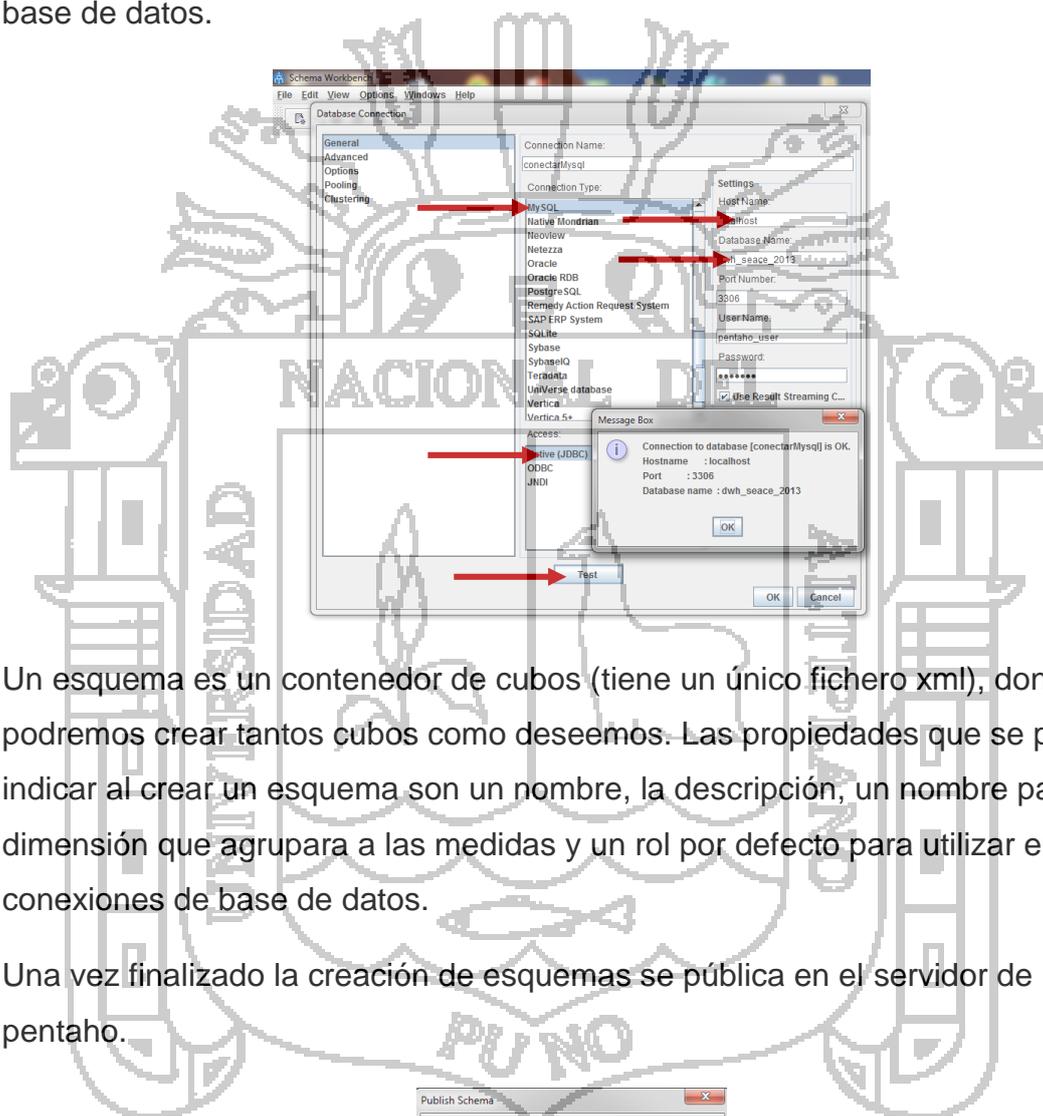
Factcontratacion

	idproceso	identidad	idproducto	ldiempo_buenapro	idproveedor	cantidad adjudicado	monto_uni adjudicado	monto_tot adjudicado
<input type="checkbox"/>	10001	10001	10001	20131203	20145496170	1	29516	29516
<input type="checkbox"/>	10002	10001	10002	20130306	10013262646	1	19000	19000
<input type="checkbox"/>	10003	10001	10003	20130602	10013442610	130	217	28210
<input type="checkbox"/>	10004	10001	10004	20131208	20447812551	500	25	12500



ANEXO 07: CONFIGURACIÓN DE PENTAHO SCHEMA WORKBENCH.

Pentaho Schema Workbench es la herramienta gráfica que permite la construcción de los esquemas de Mondrian. Los ficheros de esquema XML, describen las relaciones entre las dimensiones y medidas del cubo (modelo multidimensional) con las tablas y campos de la base de datos, a nivel relacional. Antes de crear el esquema Mondrian se configura la conexión de la base de datos.



Un esquema es un contenedor de cubos (tiene un único fichero xml), donde podremos crear tantos cubos como deseemos. Las propiedades que se pueden indicar al crear un esquema son un nombre, la descripción, un nombre para la dimensión que agrupara a las medidas y un rol por defecto para utilizar en las conexiones de base de datos.

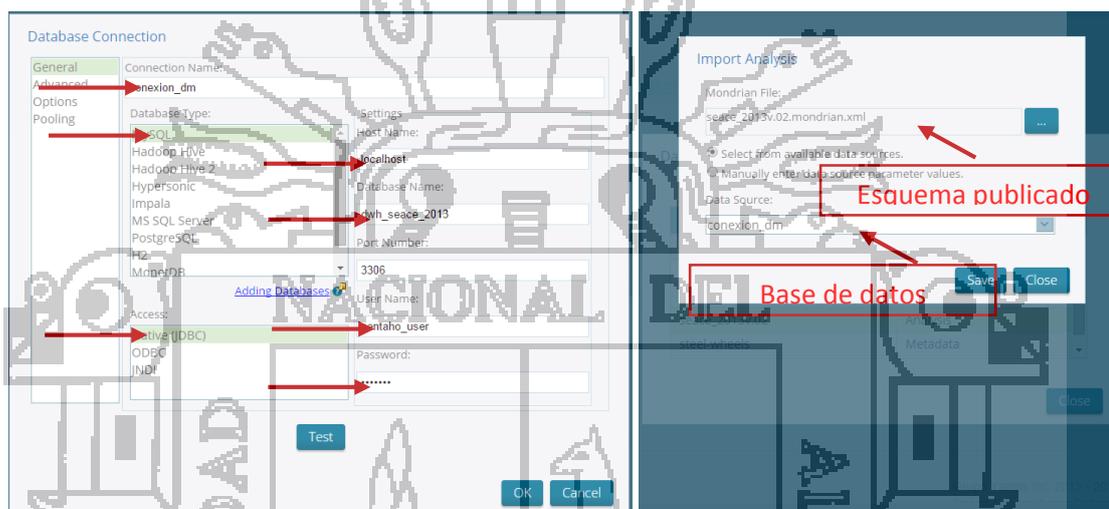
Una vez finalizado la creación de esquemas se publica en el servidor de pentaho.



ANEXO 08: CONFIGURACIÓN DE PENTAHO.

Las operaciones básicas son:

- Configurar la conexión a la base de datos.
- Relacionar el esquema publicado con la herramienta schema workbench y la conexión de la base de datos.
- Seleccionar el cubo que contiene el esquema.
- Y realizar el análisis interactivo de los datos.



En la configuración de la base de datos se proporciona los parámetros de conexión y se le asigna un nombre a la conexión, y se relaciona con el esquema.

