



Carrera de INGENIERIA DE SISTEMAS COMPUTACIONALES.

“APLICACIÓN DE ANALITICA WEB PARA LA TOMA DE DECISIONES ACERCA DEL TRAFICO DE VISITAS DE LAS APLICACIONES WEB, EN EL INEI, LIMA 2019”

Tesis para optar el título profesional de:
Ingeniero de Sistemas Computacionales

Autor:

Michael John Macavilca Mejia

Asesor:

Ing. Manuel Ricardo Mondragón Vilela

Lima - Perú

2019

DEDICATORIA

A la memoria de mi padre, Aurelio Juan Macavilca Flores.

A mi madre Alicia, mi hermana Socorro, a Alexandra, la luz del camino de mi vida. Y a
Jannet por regalarme la dicha de ser padre.

AGRADECIMIENTO

Al INEI que me permite día a día, aprender, consolidar y aplicar los conocimientos del mundo de la Inteligencia de Negocios y la Analítica

INDICE

DEDICATORIA	2
AGRADECIMIENTO	3
ÍNDICE DE TABLAS	6
INDICE DE FIGURAS.....	8
ÍNDICE DE ECUACIONES	10
CAPÍTULO I. INTRODUCCIÓN	15
1.1 REALIDAD PROBLEMÁTICA	15
1.2. FORMULACIÓN DEL PROBLEMA.....	55
1.3. OBJETIVOS	56
1.4. HIPÓTESIS.....	56
CAPÍTULO II. METODOLOGÍA.....	58
2.1. TIPO DE INVESTIGACIÓN	58
2.2. POBLACIÓN Y MUESTRA (MATERIALES, INSTRUMENTOS Y MÉTODOS).....	59
2.3. PROCEDIMIENTO	61
2.3.1 CRONOGRAMA.....	62
2.3.2 ROLES Y RESPONSABILIDADES.....	63
2.3.3 SPRINT BACKLOG	64
2.3.4 DESARROLLO SPRINT 1	65
2.3.5 DESARROLLO SPRINT II.....	77
2.3.6 DESARROLLO SPRINT III	82
2.3.7 DESARROLLO SPRINT IV	83
CAPÍTULO III. RESULTADOS.....	111
3.1 CONTRASTACIÓN DE HIPÓTESIS.....	111
3.1.1 HIPÓTESIS GENERAL	111
3.1.2 HIPÓTESIS ESPECÍFICAS	113
CAPÍTULO IV. DISCUSIÓN Y CONCLUSIONES	121
4.1 DISCUSIÓN	121

4.2 CONCLUSIONES	124
4.3 RECOMENDACIONES	125
REFERENCIAS.....	126
ANEXO 1 – MATRIZ DE CONSISTENCIA	129
ANEXO 2 – ADMINISTRACIÓN DE CUENTAS DE GOOGLE ANALYTIC 1	132
ANEXO 3 – ADMINISTRACIÓN DE CUENTAS DE GOOGLE ANALYTIC 2	133
ANEXO 4 – ADMINISTRACIÓN DE SISTEMA DE ESTADÍSTICAS WEB 1	134
ANEXO 5 – ADMINISTRACIÓN DE SISTEMA DE ESTADÍSTICAS WEB 2	135
ANEXO 6 – FICHA DE EVALUACIÓN DE LAS DIMENSIONES DE LA CALIDAD DE LA INFORMACIÓN 1	136
ANEXO 7 – FICHA DE EVALUACIÓN DE LAS DIMENSIONES DE LA CALIDAD DE LA INFORMACIÓN 2	138
ANEXO 8 – FICHA DE EVALUACIÓN DE LAS DIMENSIONES DE TOMA DE DECISIÓN EN LA ANALÍTICA WEB.....	140
ANEXO 9 – FICHA DE OBSERVACIÓN PARA EL REGISTRO DEL INDICADOR: TIEMPO DE GENERACIÓN DE INFORMACIÓN PARA LA TOMA DE DECISIONES.....	141
ANEXO 10 – FICHA DE EVALUACIÓN DE LAS PREGUNTAS SOBRE LA SATISFACCIÓN DEL USUARIO CON LAS INTERFACES DE ANÁLISIS 1	143
ANEXO 10 – FICHA DE EVALUACIÓN DE LAS PREGUNTAS SOBRE LA SATISFACCIÓN DEL USUARIO CON LAS INTERFACES DE ANÁLISIS 2	145

ÍNDICE DE TABLAS

Tabla 1. Ejemplos de Análisis clave de información.....	19
Tabla 2. Características a evaluar según la fuente de información.....	25
Tabla 3. Comparando modelos de procesos.....	31
Tabla 4. Análisis Pre test – Criterios de evaluación de las 3 dimensiones para la toma de decisiones.....	42
Tabla 5. Pre Test - Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web.....	43
Tabla 6. Pre Test – Análisis de la Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web.....	43
Tabla 7. Datos del Pre test de Tiempo de proceso de generación de información para la toma de decisiones.....	44
Tabla 8. Análisis Pre test – Criterios de evaluación de las 8 dimensiones de la calidad de información.....	44
Tabla 9. Datos del Pre test de la evaluación de la calidad de información.....	46
Tabla 10. Pre Test - Preguntas de satisfacción del usuario con las interfaces de análisis.....	46
Tabla 11. Pre Test - Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis.....	47
Tabla 12. <i>Pre Test – Análisis de la Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis.....</i>	47
Tabla 13. Materiales utilizados.....	59
Tabla 14. Instrumentos de recolección de información.....	60
Tabla 15. Cronograma.....	61
Tabla 16. Roles.....	62
Tabla 17. Sprint Backlogs.....	63
Tabla 18. Entregables Sprint 1.....	64
Tabla 19. Control de Acceso a usuarios.....	64
Tabla 20. Visualización de total de información cargada.....	65
Tabla 21. Análisis de información de aplicaciones.....	65
Tabla 22. Visualización de características de información de Visitantes y Visitas.....	66
Tabla 23. Análisis de información según estado de navegación.....	66

Tabla 24. Análisis de información geográfica por País según número de visitantes.....	67
Tabla 25. Análisis de información geográfica por País según número de visitas.....	67
Tabla 26. Análisis de información geográfica por Ciudad según número de visitantes.....	68
Tabla 27. Análisis de información geográfica por Ciudad según número de visitas.....	68
Tabla 28. Análisis de listado de aplicaciones.....	68
Tabla 29. Análisis por Mes de diferentes indicadores.....	69
Tabla 30. Análisis indicadores por Departamentos del Perú.....	69
Tabla 21. Predicción por Número de Visitantes.....	70
Tabla 32. Predicción por Número de Visitas.....	70
Tabla 33. Requerimientos Funcionales.....	71
Tabla 34. Requerimientos No Funcionales.....	72
Tabla 35. Entregables Sprint II.....	76
Tabla 36. Tabla BI_LogAplicaciones.....	78
Tabla 37. Tabla Servidor_Aplicaciones.....	79
Tabla 38. Tabla ip2location_db3.....	79
Tabla 39. Entregables Sprint III.....	80
Tabla 40. Indicadores.....	80
Tabla 41. Entregables Sprint IV.....	81
Tabla 42. Entregables Sprint 5.....	95
Tabla 43. Análisis Pos test – Criterios de evaluación de las 3 dimensiones para la toma de decisiones.....	109
Tabla 44. Post Test - Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web.....	110
Tabla 45. Post Test – Análisis de la Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web.....	110
Tabla 46. Datos del Post test de Tiempo de proceso de generación de información para la toma de decisiones.....	112
Tabla 47. Pruebas de normalidad de Tiempo de proceso de generación de información para la toma de decisiones.....	112
Tabla 48. Análisis Pos test – Criterios de evaluación de las 8 dimensiones de la calidad de información.....	115
Tabla 49. Análisis Pos test – Criterios de evaluación de las 8 dimensiones de la calidad de información.....	117
Tabla 50. Post Test - Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis.....	118
Tabla 51. Pre Test y Post Test – Análisis de la Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis.....	118

INDICE DE FIGURAS

Figura 1. Como se registran los archivos logs de los servidores web. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.....	21
Figura 2. Como se registran los datos con web beacons. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.....	22
Figura 3. Como se registra los datos con javascripts tags. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.....	23
Figura 4. Como trabaja la captura de datos con Packet Sniffing. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.	23
Figura 5. Categorías de la arquitectura técnica de BI. De “Business Intelligence Guidebook” por Rick Sherman, 2015. Copyright 2015 por Elsevier Inc.....	27
Figura 6. El modelo Trinity. De “Web Analytics One Hour One Day” por Avinash Kaushik, 2007. Copyright 2007 por Wiley Publishing.	30
Figura 7. 2019 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms. Recuperado de https://info.microsoft.com/ww-landing-gartner-mq-bi-analytics-2019.html?ls=Website	35
Figura 8. Evaluación de data Discovery. Recuperado de https://bi-survey.com/business-intelligence-software-comparison	36
Figura 9. Evaluación de self service reporting. Recuperado de https://bi-survey.com/business-intelligence-software-comparison	37
Figura 10. Evaluación de Embedded analytics. Recuperado de https://bi-survey.com/business-intelligence-software-comparison	37
Figura 11. Organigrama Funcional del INEI. Recuperado de http://www.inei.gob.pe	39
Figura 12. Diagrama de Ishikawa de la problemática actual	40
Figura 13. Proceso actual de recolección de información de la cuenta de Google Analytic	42
Figura 14. Como se visualiza un datawarehouse Web e2e. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.....	55
Figura 15. Diagrama UML de Requerimientos Funcionales	73
Figura 16. Arquitectura de solución BI.....	76
Figura 17. Proceso de carga de información.....	78
Figura 18. Modelo de Base de datos	79
Figura 19. Prototipo de Tablero de datos generales.....	84

Figura 20. Prototipo de Tablero de Análisis de información de aplicaciones.....	85
Figura 21. Prototipo de Tablero de Visualización de características de información de Visitantes y Visitas	86
Figura 22. Prototipo de Tablero de estado de Navegación	87
Figura 23. Prototipo de Tablero de Análisis de información geográfica por País según número de visitantes	88
Figura 24. Prototipo de Tablero de Análisis de información geográfica por País según número de visitas	89
Figura 25. Prototipo de Tablero de Análisis de información geográfica por Ciudad según número de visitantes	90
Figura 26. Prototipo de Tablero de Análisis de información geográfica por Ciudad según número de visitas	91
Figura 27. Prototipo de Tablero de Listado de Aplicaciones.....	92
Figura 28. Prototipo de Tablero de Indicadores por Mes y Año.....	93
Figura 29. Prototipo de Tablero de indicadores por Departamento.....	94
Figura 30. Prototipo de Tablero de Predicción de Número de Visitantes.....	95
Figura 31. Prototipo de Tablero de Predicción de Número de Visitas.....	96
Figura 32. Tablero de estado de la información.....	98
Figura 33. Tablero de Análisis por Aplicación	99
Figura 34. Tablero de Análisis por número de visitantes y visitas	100
Figura 35. Tablero de Análisis por estado de navegación	101
Figura 36. Tablero de Análisis geográfico por País según el número de visitantes.....	102
Figura 37. Tablero de Análisis geográfico por País según el número de visitas.....	103
Figura 38. Tablero de Análisis geográfico por Ciudad según el número de visitantes.....	104
Figura 39. Tablero de Análisis geográfico por Ciudad según el número de visitas.....	105
Figura 40. Tablero de Análisis por aplicación	106
Figura 41. Tablero de Análisis de aplicaciones por mes.....	107
Figura 42. Tablero de Análisis de indicadores por departamento.....	108
Figura 43. Tablero de Predicción de Número de Visitantes	109
Figura 44. Tablero de Predicción de Número de Visitas	110
Figura 45. Porcentaje de valoración de las dimensiones de toma de decisiones	113
Figura 46. Tiempo de proceso de generación de información para la toma de decisiones	115
Figura 47. Tiempo de proceso de generación de información para la toma de decisiones	116
Figura 48. Comparativo de los tiempos obtenidos en el pre test y el post test	116
Figura 49. Tiempo de proceso de generación de información para la toma de decisiones	117
Figura 50. Visualización de la administración de las cuentas en el servicio Google Analytic del INEI.	132
Figura 51. Visualización de la información por cada sitio web configurado en el servicio Google Analytic del INEI.....	133
Figura 52. Pantalla de edición o registro de información en el sistema de Estadísticas Web del INEI	134
Figura 53. Visualización de información en el sistema de Estadísticas Web del INEI.	135

ÍNDICE DE ECUACIONES

RESUMEN

La presente tesis se enmarcó dentro de la línea de investigación de la analítica web y se enfocó como objetivo principal en determinar su influencia en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web del INEI. Por otro lado, a nivel específico; el resultado de su influencia en la calidad de la información, el tiempo de proceso de generación de la información y la satisfacción del usuario con las interfaces de análisis.

Se define el presente trabajo, según Zorrilla; en el tipo de Investigación Aplicada y Mixta. Y analizando la clasificación de Carlos Sabino, lo definimos como tipo de Investigación Aplicada y Explicativa. Para la presente investigación se consideró como población, el total de 13 tableros de la solución de inteligencia de negocios implementada y que visualizan la información del datawarehouse integrado.

Para el desarrollo de esta solución, se aplicó la metodología scrum, con el fin de agilizar las coordinaciones y entregas de avance. Se integró en un datawarehouse, la información de los archivos de registro de los servidores web, la información de ubicación geográfica de los IPs a nivel mundial y la información del inventario de aplicaciones del INEI. Todo esto con el fin de poder analizar la información bajo distintas dimensiones agregadas, como país, local del INEI, tipo de usuario, aplicación, tipo de aplicación y otros, a los que se complementó el cálculo de 6 indicadores bajo el concepto de la analítica web; de tal manera que este análisis contribuya óptimamente a la toma de decisiones por parte de los usuarios responsables.

Se plantea finalmente que la solución de inteligencia de negocios desarrollada, sea accesible para las diferentes unidades de la Oficina Técnica de Informática, así como de la Dirección Técnica de Difusión para los fines respectivos. Además, debe sumar en su fase de integración, la información que proviene de la cuenta de Google Analytics y la de los servidores web que no son IIS de Microsoft. Por otro lado, se debe ampliar el número de tableros de análisis en la herramienta Qlik Sense, del tipo de análisis histórico y predictivo para así optimizar más aun, no solo la satisfacción del usuario, sino también la capacidad de poder tomar decisiones con un mejor entorno de análisis.

Palabras clave: analítica web, inteligencia de negocios, datawarehouse, Qlik Sense, análisis, Procesamiento de información, log file, archivos de registro web, internet information server, transformación de datos, integración de datos, calidad de la información, toma de decisiones, predicción.

ABSTRACT

This thesis was framed within the research line of web analytics and focused as the main objective in determining its influence in the decision making about the traffic of visits of the INEI web applications. On the other hand, at a specific level; the result of its influence on the quality of the information, the process generation time of the information and the satisfaction of the user with the analysis interfaces.

The present work is defined, according to Zorrilla; in the type of Applied and Mixed Research. And analyzing the classification of Carlos Sabino, we define it as a type of Applied and Explanatory Research. For the present investigation, the total of 13 boards of the business

intelligence solution implemented and displaying the information of the integrated datawarehouse was considered as a population.

For the development of this solution, the scrum methodology was applied, in order to speed up coordination and advance deliveries. It was integrated into a datawarehouse, the information from the log files of the web servers, the geographic location information of the IPs worldwide and the information from the INEI application inventory. All this in order to be able to analyze the information under different aggregate dimensions, such as country, INEI location, type of user, application, type of application and others, to which the calculation of 6 indicators was complemented under the concept of analytics Web; in such a way that this analysis contributes optimally to the decision making by the responsible users.

Finally, it is proposed that the developed business intelligence solution be accessible to the different units of the Technical Information Office, as well as the Technical Diffusion Directorate for the respective purposes. In addition, you must add in its integration phase, the information that comes from the Google Analytics account and that of non-Microsoft IIS web servers. On the other hand, the number of analysis boards in the Qlik Sense tool, of the type of historical and predictive analysis, must be expanded to optimize even more, not only user satisfaction, but also the ability to make decisions with better Analysis environment.

Keywords: web analytics, business intelligence, datawarehouse, Qlik Sense, analysis, Information processing, log file, web log files, internet information server, data transformation, data integration, information quality, decision making, prediction.

CAPÍTULO I. INTRODUCCIÓN

1.1 Realidad problemática

Davenport y Harris (2007) nos describen que la analítica se refiere al uso extensivo de datos, los análisis estadísticos y cuantitativos, los modelos explicativos y predictivos, y la gestión basada en hechos para impulsar decisiones y acciones. La analítica puede ser el input para decisiones humanas o puede conducir a decisiones totalmente automatizadas. La analítica

es un subconjunto de lo que se ha dado a llamar como inteligencia de negocios: un conjunto de tecnologías y procesos que utilizan datos para comprender y analizar el rendimiento del negocio. Entonces, concluimos que la inteligencia de negocios engloba el acceso a los datos, los informes y la analítica.

Según Beasley (2013) la analítica web es la forma de aprender, el cómo los usuarios interactúan con los sitios web y las aplicaciones móviles, conllevando al registro automático de los aspectos del comportamiento de los usuarios para luego, combinar y transformar el comportamiento registrado, en datos que puedan analizarse.

La navegación en internet desde un inicio, se conceptualiza de una manera simple. El digitar la dirección del sitio web en un navegador, nos permite acceder al mismo, permitiéndonos enlazar a otras direcciones internas o externas. Esta navegación e interacción con las páginas conlleva a descubrir que en ocasiones se producían errores. Entonces es cuando, alguien descubrió los archivos de registro de errores del servidor y los aprovechó para descubrir información relacionada a las acciones realizadas en los sitios web que estaban en estado de publicación en el servidor web. Así mismo, estos registros del servidor capturaban no solo el hecho de que alguien accedió al sitio web, sino también información adicional, como el nombre del archivo, la hora, la página que realiza la solicitud de enlace, el registro del identificador del navegador, el IP del usuario conectado, el sistema operativo, Etcétera.

A medida que los archivos de registro comenzaron a crecer, otra persona escribió el primer script que analizaría automáticamente los archivos de registro, y, en consecuencia, generó las métricas básicas. Es entonces cuando, la analítica web había nacido oficialmente. Las primeras generaciones de herramientas de análisis de archivos de registro, impulsaron la adopción del análisis web más allá del alcance del equipo de Tecnología de la Información (TI). Lo que generó que la gente de marketing, pueda entender lo que estaba sucediendo con la primera versión del análisis de esta información. Más tarde, a mediados de la década de los 90,

hace su ingreso al mundo del internet, un servicio llamado, “Contador”. Con él, los sitios web comenzaron a exponer sus estadísticas web al público, mostrando cuán popular era el sitio que exponía este contador.

La Analítica web comercial apareció en el escenario varios años después, con WebTrends convertido en su mejor representante. WebTrends tomó al básico analizador de archivos de registro y le agregó mejoras, como tablas y gráficos bonitos que finalmente arrastraron la analítica web a los equipos de dirección de negocios. Para el año 2000, con el crecimiento exponencial de la Web, se consolidaron empresas como Accrue, WebTrends, WebSideStory y Coremetrics, quienes proporcionaron soluciones cada vez más complejas y permitiendo a su vez, administrar grandes cantidades de datos. Al mismo tiempo, los proveedores y clientes de la analítica web basados en el uso de registros del servidor web, se encontraron con los siguientes desafíos:

- Almacenamiento de páginas en caché por los ISP: Cuando el ISP tuviera una copia de la página, todas las páginas subsiguientes serían atendidas por esta versión, perdiéndose información a adicionar en el registro del servidor web.
- Robots de búsqueda: Los robots de búsqueda de los “Search Engine” dirigen sus resultados a sitios ya almacenados en su base de datos.
- Visitantes únicos: El caso de usuarios asignados con direcciones IP dinámicas hizo difícil identificar visitantes únicos.

Es entonces cuando surgieron las etiquetas de JavaScript (líneas de código JavaScript), como un nuevo estándar para recopilar datos de los sitios web. Su fácil implementación en los mismos, fue uno de los factores positivos que contribuyó a su popularidad. Se agregan unas pocas líneas de código JavaScript a cada página y se activan cuando estas se cargan en el navegador, enviando datos a un servidor de recopilación de datos, que no es administrado por

el propietario del sitio web. Por consiguiente, se transfirió la responsabilidad de recopilar y procesar datos, de los departamentos de TI de la empresa hacia los proveedores de analítica web.

El siguiente paso evolutivo en la analítica web fue la introducción del “site overlay” (click density). Durante un período de tiempo elegido, la herramienta de analítica web muestra exactamente en qué sector de la página, hacen clic los visitantes. Lo que define otro tipo de análisis en el comportamiento de los usuarios. De ahí que, logran afianzarse cuatro grandes proveedores de analítica web: Coremetrics, Omniture, WebTrends y WebSideStory. Pero no es hasta en el 2005, cuando Google se hizo presente comprando a Urchin, y el 2006, lo lanzó como una herramienta de analítica web gratuita bajo el nombre de Google Analytics. Convirtiéndose de esta manera, en una de las herramientas de analítica web más populares a nivel mundial.

Hoy en día, las empresas de todo el orbe, invierten millones de dólares en analítica web. Direccionados en campañas para lograr el acceso de visitantes e implementaciones en sus sitios web, los mismos que a su vez tienen por objetivo primordial, generar ingresos en línea. El sitio web se ha convertido en un canal para generar ingresos, lo que conlleva ahora que exista una gran demanda para que este canal web rinda cuentas, tan igual como los demás canales de ventas de las empresas. Entonces, es ahora el momento, cuando se le pide que justifique el dinero invertido. Por lo tanto, es la analítica web, la herramienta que medirá el comportamiento de lo invertido y también el retorno esperado.

La analítica web es el proceso de analizar el comportamiento en línea para optimizar el uso del sitio web y aumentar el compromiso y las ventas.

Marr (2016) nos señala que hay dos tipos de analítica web: la que se aplica fuera del sitio web y en el sitio propiamente dicho. El análisis web fuera del sitio analiza lo que sucede en internet en su conjunto e incluye la medición de la audiencia potencial, la competencia y las

tendencias en línea de un producto o servicio. El análisis web en el sitio, es el análisis de su propio sitio web, esto incluye la recopilación de datos sobre cuántas personas visitaron el sitio, de dónde vinieron, cuánto tiempo se quedaron, cómo navegaron por el sitio y si la visita resultó en una venta o logrando el propósito del servicio brindado. El análisis web fuera del sitio es útil para evaluar el mercado y la oportunidad, mientras la que se aplica en el sitio, es útil para medir los resultados comerciales.

Al día de hoy, los usuarios tienen a su disposición, una variada gama de herramientas para poder lograr la implementación de la Analítica Web, a la vez, el acceso a los datos de su sitio web, lo pueden hacer de manera simple y de forma gratuita. Pero después de iniciar este proceso, es difícil determinar cuáles son las métricas de éxito, y, sobre todo, cómo hacer un análisis web correcto. Sobre esto, Beasley (2013) nos refiere que las métricas son medidas numéricas de varios aspectos del comportamiento de los usuarios, como cuánto tiempo pasan viendo una página o cuántas veces los usuarios vieron una página.

La base de la analítica web tradicional durante mucho tiempo, ha sido el análisis de los indicadores clave de rendimiento (KPI) como el número de visitas, número de visitantes únicos, páginas vistas, tiempo promedio de visita, % de rebote, % nuevas visitas, etc. En la actualidad para poder realizar un análisis más completo, permitiendo a este, adaptarse a las diferencias estratégicas en las operaciones comerciales de una empresa; es necesario adicionar a los KPI, los análisis de información clave (KIA).

Tabla 1.

Ejemplos de Análisis clave de información

Análisis clave de información	Descripción
(KIA)	
Análisis de densidad de clic	“site overlay”. Analizar en donde están haciendo clic los usuarios.

Propósito principal del visitante	Realizar una encuesta, hacer entrevistas telefónicas, busque clientes reales y pregúnteles por qué aparecen como visitantes de su sitio web.
Tasas de finalización de tareas	¿El hecho de que alguien haya visitado una página, realmente significa la finalización exitosa de una tarea? Entonces se debe centrar el estudio en aquellas páginas que urgen cuando alguna tarea ha finalizado.
Tendencias segmentadas de visitantes	Al día de hoy, pocas herramientas en el mercado tienen la capacidad real de segmentar los datos una vez que han sido capturados. Algunos de estos son: ClickTracks y Visual Sciences (ambos con costo).
Análisis de impacto multicanal	Tenemos que ver al sitio web como parte de un ecosistema en internet. Obtener información verdadera requiere medir el impacto de otros canales sobre el sitio web, y también de este sobre ellos.

En la última década, se ha establecido ya, Web Analytics 2.0. A diferencia de la 1.0 donde se centraba únicamente en el estudio del comportamiento del visitante en el sitio web (¿Qué es lo que hace?), la 2.0 se centra en estudiar el qué y, además, el motivo de las acciones del visitante. ¿Por qué?, y adicionalmente, de que es lo que espera encontrar o hacer. Para el logro de este fin, es de suma importancia que se creen los canales de retroalimentación de información disponibles y de fácil acceso en todas las aplicaciones web del servidor. Lo que nos permitirá obtener su percepción de lo esperado y luego esta información sea integrada al datawarehouse, para poder así, analizarla con la data cuantitativa que registro el comportamiento de los visitantes en el sitio web. En la actualidad algunas de las empresas más representativas en el mercado de herramientas de analítica web son: Omniture, Google

Analytics, IBM Unica's NetInsight, Webtrends, Yahoo! Web Analytics, Lyris HQ, Coremetrics.

Existen cuatro formas principales de capturar datos acerca del flujo de clics realizados por los visitantes en un sitio web, estos son: los registros web, web beacons, etiquetas JavaScript y packet sniffing. A continuación, se detalla una breve descripción de cada uno.

- **Registros Web:** Los registros web han sido la fuente original de recopilación de datos desde los inicios de la web. Nacieron para poder capturar los errores generados por las aplicaciones en los servidores web y a través del tiempo han ido evolucionando para capturar más datos a medida que las necesidades analíticas cambiaron de técnicas a marketing.

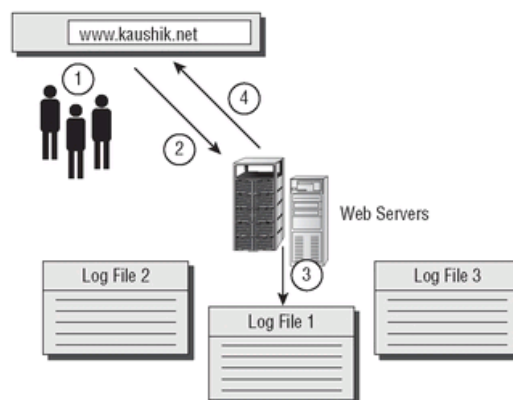


Figura 1. Como se registran los archivos logs de los servidores web. De "Web Analytics One Hour One Day" por Kaushik, 2007. Copyright 2007 por Wiley Publishing.

- **Web Beacons:** Los web beacons se desarrollaron durante una época en que los banners gobernaban la web como la principal forma de capturar información de los visitantes, siendo esta data registrada en sitios web externos. El flujo de su proceso consiste en que una empresa publica anuncios publicitarios en muchos sitios web, en consecuencia, surgió la necesidad de analizar diferentes indicadores, como: cuántas fueron las personas que accedieron a los anuncios publicitarios, cuantos hicieron clic en uno o los diferentes tipos de banners segmentados,

cuántos de estos, eran la misma persona, canales de origen (anuncios de banner, correo electrónico, texto de anuncios etc.), además del proveedor del servicio de la publicación de los banners.

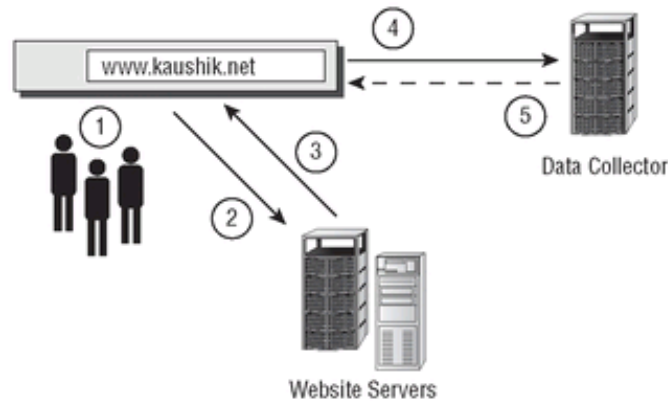


Figura 2. Como se registran los datos con web beacons. De "Web Analytics One Hour One Day" por Kaushik, 2007. Copyright 2007 por Wiley Publishing.

- JavaScript Tags: El etiquetado de JavaScript es el favorito de la industria en este momento. La mayoría de los proveedores y las soluciones de analítica web dependen del etiquetado de JavaScript para recopilar datos. El etiquetado de JavaScript permitió que se recopilaran más datos con mayor precisión. Redujo la dependencia de los departamentos de TI para la captura de datos, ya que la responsabilidad de esta captura de datos se trasladó hacia los proveedores externos de analítica web. Quienes a su vez procesan lo recolectado y publican una gama de informes que estarán disponibles en línea. Entonces las empresas ya no necesitan implementar la infraestructura para recopilar y procesar los datos, ni los sistemas para publicar los informes.

Su implementación es la más fácil. Consiste en agregar unas pocas líneas estándar de código JavaScript en cada una de las páginas del sitio web. Si la empresa renta el servicio de publicación de su sitio web y no tiene acceso a los registros del servidor web, esta opción es la ideal para hacer analítica web. Pero

también debemos tener en cuenta que no todos los visitantes del sitio web tienen JavaScript activado, a menudo por motivos de privacidad u otros. Para estos usuarios, la plataforma de análisis no recopilará ningún dato.

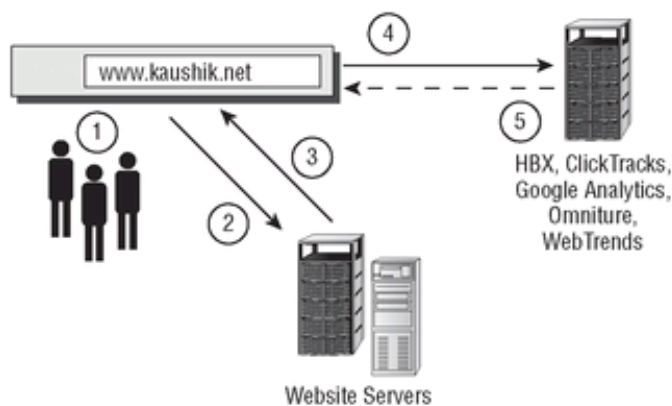


Figura 3. Como se registra los datos con javascripts tags. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.

- **Packet Sniffing:** Esta es una de las formas más sofisticadas de recopilar datos en el sitio web. Un Packet Sniffing (rastreo de paquetes) puede ser una capa de software que se instala en los servidores web. Alternativamente, también puede ser una pieza física de hardware que se conecta al servidor web y centraliza todo el tráfico.

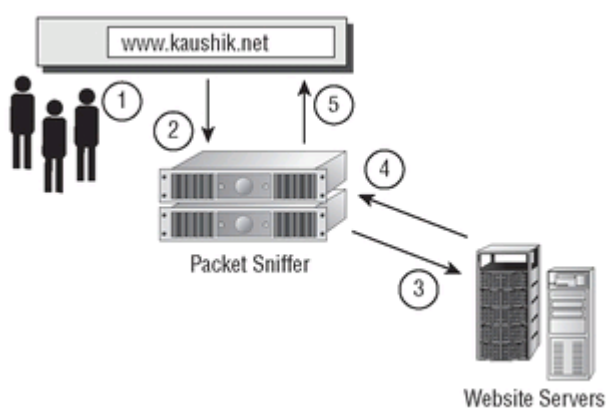


Figura 4. Como trabaja la captura de datos con Packet Sniffing. De “Web Analytics One Hour One Day” por Kaushik, 2007. Copyright 2007 por Wiley Publishing.

Después de explorar la teoría de la analítica web sobre su concepto, historia, métricas y fuentes de datos, podemos concluir que este análisis se inició con datos procedentes de los

registros de los servidores web, y que estos principalmente contienen información técnica y no información comercial. Por lo que, en su ruta evolutiva, hoy en día, las herramientas de analítica web en su carrera por satisfacer los requerimientos de los clientes, se basan en el análisis del flujo de clics.

¿Entonces, podemos oficializar que el análisis de los archivos de registro de los servidores web es ya una fuente obsoleta o en un extremo, es incapaz de colmar las expectativas de interés de algún tipo de usuario en la empresa? La respuesta sería, depende de la definición de la entidad objetivo, de la que deseas estudiar sus indicadores. Como ya hemos explicado desde un inicio, si el objetivo es estudiar el comportamiento más exacto del usuario en las interfaces de tu aplicación web, entonces no hay duda alguna que la herramienta de analítica web a utilizar sería una basada en el flujo de clics. Ya que el análisis resultante, te permitirá evaluar a donde se dirige el usuario, para así poder estudiar su interés en los productos que ofreces o la búsqueda de interés que este se haya definido. ¿Y para qué? Pues, el objetivo es poder tener conocimiento, de qué es lo que más interesa a tus posibles clientes, o de que otros temas averiguaron, aquellos que ya realizaron alguna compra, etc. Tu entidad objetivo es el usuario con el propósito de analizar su comportamiento.

Bien entonces, si hoy en día, planteamos la aplicación de la analítica web basándonos en los archivos de registro de los servidores web para estudiar la entidad “usuario”, con el propósito de analizar su comportamiento, sería como arrojar desde un avión a un tanque sostenido de un paracaídas de tela. Y en consecuencia concluiríamos, que el tanque a pesar de tener proyectiles de grueso calibre, no puede sostenerse en el aire, ni mucho menos combatir contra otros aviones. Pues bien, pero que les parece si planteamos estudiar el comportamiento de la entidad “aplicación web”. Entendamos por aplicación web, a un directorio donde se agrupa una serie de archivos (diferentes formatos) que se ejecutan según la navegación del usuario. ¡Eso ya lo estudia Google Analytics! Es cierto, pero si planteamos el siguiente caso,

por ejemplo, que la gerencia de IT desea hacer un estudio simple del indicador, de número de visitas de las aplicaciones web que descargaron información en formato Excel a usuarios públicos y no a usuarios de la institución; en los últimos tres años y en cada uno de los cuatro servidores con diferentes dominios con los que cuenta la empresa.

¿Cómo hacerlo? Lo primero que se tendría que hacer es recopilar la información de forma individual de cada uno de los dominios registrados en la cuenta de Google Analytics. ¿Y cómo podría subdividir la información por aplicación web? ¿Y si un directorio de publicación cuenta con subdirectorios de aplicaciones web independientes, como podría reconocer la información para cada una? ¿Podemos adicionar una dimensión de estudio en Google Analytics, por ejemplo, Tipo de aplicación web o nombre de aplicación web? Y así podríamos adicionar más preguntas que harían inviable la opción basada en el flujo de clics. Al menos las versiones gratuitas entrarían en esta afirmación.

¿Qué sucede ahora con nuestro tanque? Los registros de los servidores web se convierten ahora si, en una muy buena opción de fuente de información para aplicar analítica web. Pero no por si sola. La data de los registros de los servidores web tiene que ser consolidada en una base de datos y luego integrarla con otras fuentes de información en un datawarehouse. Hablamos de información relacionada a los IPs como País, estado, ciudad etc., Tipo de aplicación, inventario de aplicaciones institucionales, etc. Con el fin de convertirse en la fuente única de información para que a través de una herramienta BI, podamos aplicar la analítica web. Se concluye entonces, que recurrimos a la arquitectura de Inteligencia de negocios para que se pueda analizar los indicadores que propone la analítica web.

Tabla 2

Características a evaluar según la fuente de información

Fuente de Información		
Características	Flujo de clics	Registros de los servidores web

Herramienta a utilizar	Google Analytics	Herramienta de Inteligencia de Negocios
Entidad de estudio	Dominio y su grupo de archivos	Todas las aplicaciones Web De todos los servidores
Propietario de la información	Proveedor de herramienta de Analítica web	Propietario de los servidores web
Proceso de registro de la fuente de datos	A cargo del proveedor de la herramienta de Analítica Web	A cargo de los servidores web
Proceso para recolección de la data	Descarga de archivos en formatos Excel, por cada perfil relacionado a cada dominio configurado para luego consolidarlo	Proceso de extracción de información de los registros de los servidores web.
Relacionar con otras fuentes de información	La herramienta no lo permite. Se requiere un proceso posterior.	Si, se aplica a la data extraída y se integra con otras fuentes en un datawarehouse.

Nogués y Valladares (2017) nos dejan este concepto de inteligencia de negocios, BI pertenece a un mundo analítico. La inteligencia de negocios es un conjunto de herramientas y procesos que lo ayudan a tomar decisiones basadas en datos precisos, ahorrando tiempo y esfuerzo. La idea principal detrás de una herramienta de BI es la posibilidad de analizar fácilmente datos basados en conceptos comerciales sin tener conocimientos técnicos sobre herramientas de bases de datos u otras fuentes que contienen los datos.

Para poder entender la arquitectura de inteligencia de negocios, en la Figura 5, podemos visualizar un gráfico que nos da una idea simple y clara del funcionamiento de la propuesta de BI.

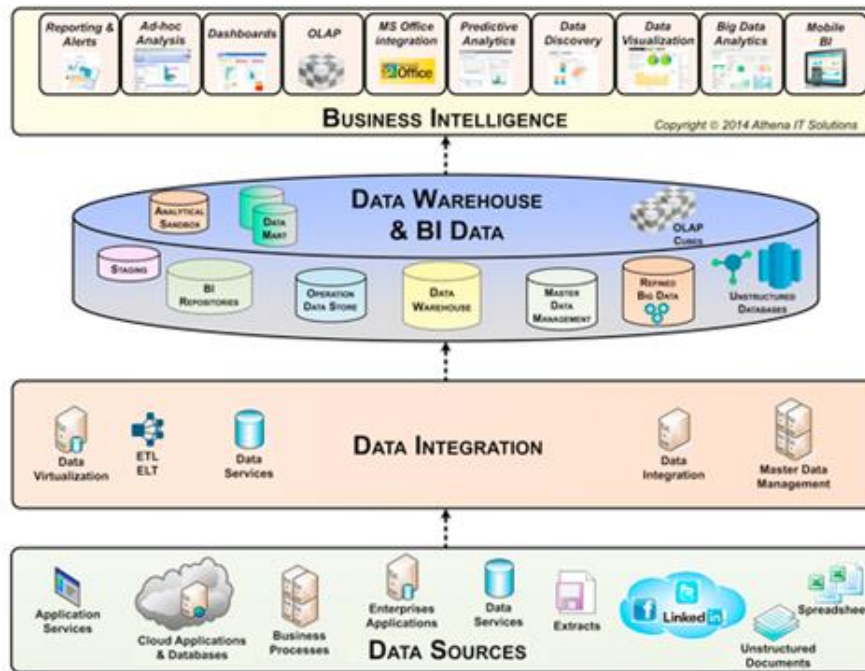


Figura 5. Categorías de la arquitectura técnica de BI. De “Business Intelligence Guidebook” por Rick Sherman, 2015. Copyright 2015 por Elsevier Inc.

De esta figura 5, se puede reconocer que existen cuatro capas de tecnología en la arquitectura BI:

- Fuentes de datos: Cualquier fuente de datos. Esta data puede ser estructurada o no estructurada. Los datos estructurados pueden organizarse en un registro o archivo predefinido y pueden almacenarse en una base de datos u hoja de cálculo. Algunos ejemplos de datos estructurados son datos de ventas, empleados y financieros de una empresa. Los datos no estructurados están libres o no organizados. Archivos de registro, mensajes de correo electrónico, tweets, archivos PowerPoint, documentos de Word y videos son ejemplos de datos no estructurados.
- Integración de datos: En esta fase, la extracción, transformación y carga de datos es el proceso en el cual los datos son tomados del sistema fuente para ser

configurados y almacenados en un datawarehouse. Las herramientas de ETL automatizan las tareas de integración de datos.

- Almacenamiento de datos: Se refiere al datawarehouse cuyo fin es el de almacenar los datos integrados que luego serían consumidos por las herramientas de BI.

Nogués y Valladares (2017) también nos describen acerca del datawarehouse de esta manera: Uno de los conceptos más importantes que viene con una solución de BI es el Datawarehouse. El datawarehouse es la base de los sistemas de BI tradicionales porque es el lugar donde residen los datos que desea analizar, generalmente soportado en un sistema de base de datos. La idea debajo de un datawarehouse es que puede recopilar datos de múltiples fuentes, limpiarlos, garantizar su integridad, garantizar la coherencia y garantizar que se encuentre completa para tener una implementación de BI confiable.

- Inteligencia de negocios y análisis: Esta capa comprende a las herramientas BI que se utilizan para analizar la información. Las mismas que incluyen sus diferentes funcionalidades y servicios dependiendo del proveedor.

La toma de decisiones basada en datos y análisis se ha convertido en un valor incalculable para las organizaciones de todos los tamaños y en todas las categorías de la industria. Los gerentes comprometidos en encontrar formas de hacer un mayor uso de los datos digitales deben comprender la importancia de analizar y usar los datos digitales en la toma de decisiones para mejorar realmente el funcionamiento general y el éxito de una organización. Herrmann (2015) nos describe que existen dos tipos importantes de sistemas de toma de decisiones, los sistemas de control y los sistemas de transformación. Un sistema de control es un sistema de toma de decisiones que monitorea continuamente el estado de otro sistema y toma decisiones

que determinan las instrucciones que deben enviarse al otro sistema para que cumpla sus objetivos. Un sistema de transformación es un sistema de toma de decisiones, que recibe información de fuentes externas y luego toma decisiones que generan nueva información que otras personas usarán.

Para una mayor comprensión del entorno y de los competidores se debe realizar el análisis detallado de la información, lo que permitirá a cualquier tipo de empresa, o de cualquier dimensión, la toma de decisiones con criterios y una alta certidumbre basada en los datos para invertir en estrategias a largo plazo. Entonces, la analítica se respalda en la inteligencia de negocios, conllevando a que esta se convierta en fuerte aporte a los mecanismos de toma de decisiones. Inteligencia de negocios nos provee de información clave como “materia prima” para la toma de decisiones. Luego, son los analistas o usuarios de las herramientas de inteligencia de negocios, los que analizan la información, con el objetivo de encontrar las causas de eventos pasados ocurridos, estudiar el comportamiento actual de la información, y así también, medir la consecuencia de decisiones tomadas anteriormente. De igual manera, el análisis debe incluir la ejecución de predicción de datos para poder fundamentar las decisiones con visión en el tiempo.

Como hemos detallado anteriormente, la analítica web ha contribuido con indicadores para tener una visión del comportamiento de los usuarios en los sitios web, así como como el comportamiento de las aplicaciones, tanto en tráfico como en el estado final expuesto en su publicación. ¿Pero, que más se puede hacer? Lo que conlleva a preguntarnos, ¿Cuál debería ser el paradigma actualizado para aplicar la analítica web y que esta influya en la toma de decisiones de las empresas? Según Kaushik (2007), la respuesta es Trinity, una nueva forma de pensar sobre la toma de decisiones en la analítica web, un concepto que es algo más que un flujo de clics. Ejecutar a Trinity asegurará que pueda construir una plataforma de toma de decisiones de clase mundial que creará una ventaja competitiva sostenible para su empresa.

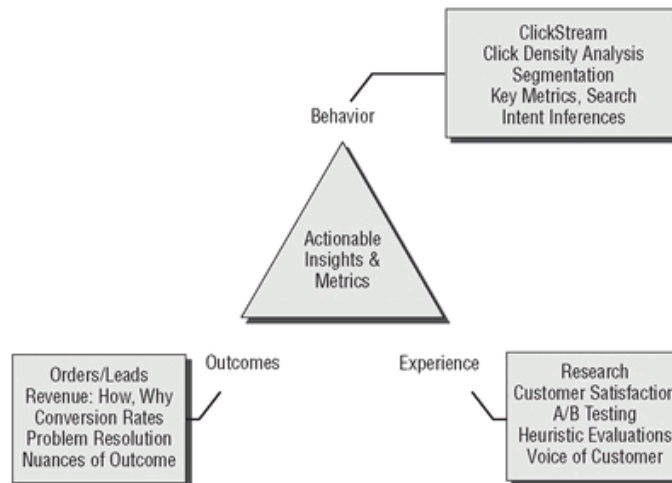


Figura 6. El modelo Trinity. De “Web Analytics One Hour One Day” por Avinash Kaushik, 2007. Copyright 2007 por Wiley Publishing.

Las ideas y métricas accionables son el objetivo que impulsan la diferenciación estratégica y una ventaja competitiva sostenible. Generar estas ideas combinadas con objetivos claros, ayuda a cristalizar los esfuerzos de la organización. Esta estrategia fomenta el comportamiento correcto de la organización y es una excelente manera para que las partes interesadas clave (stakeholders) tomen sus decisiones diarias de asignación de recursos.

El primer componente de la mentalidad de Trinity es el análisis del comportamiento (behavior), lo que tradicionalmente consideramos análisis de datos de flujo de clics. El objetivo del análisis de comportamiento es inferir la intención de nuestros clientes o visitantes del sitio web en función de lo que sabemos de ellos a partir de los datos que tenemos. Después de recopilar datos de flujo de clics, el objetivo es analizarlos desde un plano de referencia más alto. Y para lograr este, debemos incluir un análisis de densidad de clics (overlay site), la segmentación de datos por n niveles para poder descubrir información valiosa, un análisis de los criterios de búsqueda, métricas claves.

El segundo componente de la mentalidad de Trinity es el análisis de resultados. Este elemento de la Trinidad existe para medir si el sitio web está cumpliendo con el objetivo de su existencia. Según la naturaleza de los objetivos con que fueron creados o la misión de la

empresa, podemos mencionar como ejemplo, que, para un sitio web de comercio electrónico, los indicadores a medir estarían dirigidos a los ingresos generados de las ventas en línea. Por otro lado, en un sitio web de soporte, el indicador a medir sería la atención concluida de los problemas reportados por los usuarios. Para cada sitio web los resultados podrían ser diferentes, pero casi siempre serán métricas que afectarán el resultado final y pueden estar vinculadas a las finanzas de la empresa. Las métricas de resultados deben ser analizadas y comparadas con las métricas de comportamiento.

Aunque el elemento de los resultados de la Trinidad es obligatorio, el elemento de la experiencia es quizás el más crítico. El análisis de la experiencia nos permite meternos en la cabeza de nuestros clientes y obtener información del por qué hacen las cosas que hacen. Hay muchas formas de comprender la experiencia y satisfacción de los clientes en el sitio web. Para eso se puede aprovechar las encuestas, o existen metodologías estadísticas cuantitativas y cualitativas muy complejas que se puede aplicar. Otros medios de obtener información son: mediante el uso de la metodología de testing A / B o la prueba multivariante, evaluaciones heurísticas, pruebas de usabilidad de laboratorio, visitas al sitio.

La medición o evaluación de la calidad de la información es el proceso de inspección de la información del negocio para garantizar que satisfaga las necesidades del conocimiento de los trabajadores que dependen de ella. Según Kenett y Shmueli (2016) nos define 8 dimensiones aplicadas a la calidad de la información: Resolución de datos, Estructura de datos, Integración de datos, Relevancia temporal, Cronología de datos y meta, generalización, Operacionalización, Comunicación. Para plasmar una evaluación, debe integrarse diferentes aspectos de un estudio y asignar una puntuación general de Calidad de la información basada en las calificaciones de los expertos. En el presente trabajo se utilizó un enfoque basado en la

calificación que examina un informe de estudio y califica cada una de las ocho dimensiones en una escala de 1 a 5.

A continuación, se realiza una visión general de la gestión de proyectos ágil con Scrum, y compararemos esta metodología con el modelo de proceso tradicional, para así poder contrastar sus especificaciones.

Tabla 3

Comparando modelos de procesos

		Scrum	Tradicional
Gestión del proyecto	Planificación de los productos	Visión del producto, Hoja de ruta, Product Backlog	Hoja de ruta
	Planificación del Proyecto	Product Backlog	Plan de proyecto con hitos
	Planificación de Tareas	Sprint backlog para cada iteración	Fases específicas
	Iteración	Duración de iteración predefinida	De acuerdo con el plan de proyecto
	Duración de Iteración	Típicamente de 1 a 4 semanas	De acuerdo con el plan de proyecto
	Monitoreo de estado	Diariamente con todo el equipo en una pizarra	Por gerente de proyecto, basado en hitos
	Timeboxing	Si	El gerente del proyecto decide
	Entrega	Después de cada iteración (es decir, al final de cada sprint)	Con cada versión o al finalizar el proyecto

	Gestión del cambio	Métricas	Vía la actualización del Backlog Burndown Chart	A través del sistema de gestión de cambios existente Adhesión a plazos, esfuerzo, costos
Gestión de la calidad	Optimización de procesos	Verificación / Testing	Revisión del proyecto, inspección y adaptación (de abajo hacia arriba) Continuamente dentro del sprint	Auditorías internas y externas (por ejemplo, de acuerdo con ISO 9001), programas de optimización de procesos (de arriba hacia abajo) Prueba contra las especificaciones del proyecto durante las fases de prueba
	Validación / Aprobación		Demostración al usuario al final de un sprint (cumplimiento de los criterios de aceptación)	Prueba de aceptación al finalizar el proyecto
Herramientas de desarrollo	Técnicas		Técnicas comunes pero no obligatorias: programación de pares, integración continua, programación de pruebas iniciales, diseño incremental, código limpio, refactorización	Depende del modelo en uso. Algunos incluyen codificación predefinida, guía de herramientas, etc.
Recursos Humanos	Valores y principios		Manifiesto ágil, valores scrum: compromiso,	Gestión de proyectos, modelo de proceso, optimización continua (ISO 9000)

	enfoque, apertura, respeto, coraje	
Organización	Scrum Team, Product owner, scrum Master	Jefe de equipo, gestión de proyectos, organización de proyectos
Entrenamiento	Por iniciativa propia, deseo de mejorar, aprender con el equipo	Plan de formación empresarial, desarrollo de recursos humanos
Asignación de trabajo	Interdisciplinario (funcional cruzado)	Altamente especializado

Nota. Fuente: Adaptado de Linz (2014). Testing in Scrum. Published by Rocky Nook

Hoy en día, las opciones de las herramientas BI se han desarrollado de tal manera que incluyen las funcionalidades de descubrimiento de datos, visualización de datos, análisis en la memoria, análisis en la base de datos, modelado predictivo, ejecución como una aplicación de BI independiente, análisis textuales, análisis de Big Data, entre otros. A continuación, revisaremos dos fuentes de evaluación de herramientas de inteligencia de negocios. La primera es Gartner, publicado en febrero del 2019. La Figura 7, nos muestra en su cuadrante de empresas líderes a cuatro empresas: Microsoft, Tableau, Qlik y ThoughtSpot. Gartner basa sus rankings en encuestas a vendedores y clientes. Su matriz de ubicación gira en la medición por la capacidad de ejecutar estrategias de negocio y la completitud de su visión. Lo que nos lleva afirmar la posición de Qlik, como empresa líder en el mercado global, por lo que se considera que su producto Qlik Sense está ampliamente respaldado por el mismo y nos garantiza un éxito en su implementación para nuestra solución.

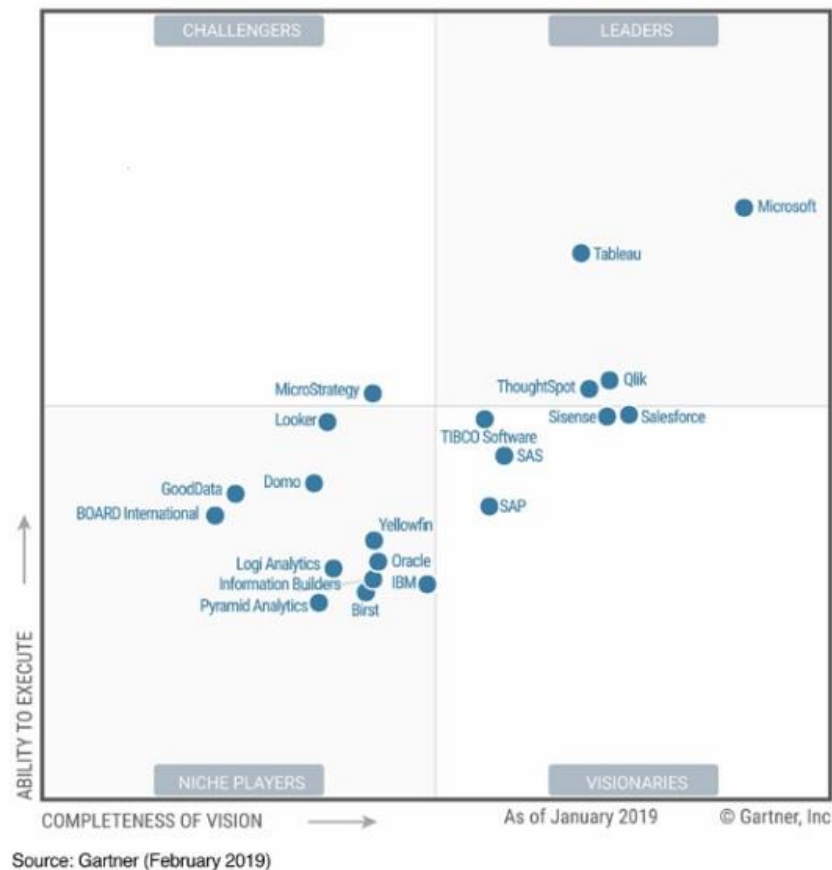


Figura 7. 2019 Gartner Magic Quadrant for Analytics and Business Intelligence Platforms. Recuperado de <https://info.microsoft.com/ww-landing-gartner-mq-bi-analytics-2019.html?ls=Website>.

Otra evaluación importante en el mundo de la inteligencia de negocios, es la basada en los resultados agregados de KPI del BI Survey 18, la más grande encuesta anual a usuarios de software de inteligencia de negocios; casi 2800 informantes de diferentes países de todos los continentes. Todos profesionales relacionados a la inteligencia de negocios. Los resultados del BI Survey, es un producto reconocido por los fabricantes de herramientas BI, al punto que los mismos son publicados en los sitios web de estos últimos. A continuación, en las Figuras 8, 9 y 10 se visualizan la evaluación de algunas funcionalidades de 2 herramientas BI, para nuestro caso hemos escogido la herramienta Qlik Sense y la más conocida, MS Power BI. Con el fin de poder apreciar la comparación entre ambas y también poder analizar el grado de diferencias entre las evaluaciones.



Figura 8. Evaluación de data Discovery. Recuperado de <https://bi-survey.com/business-intelligence-software-comparison>

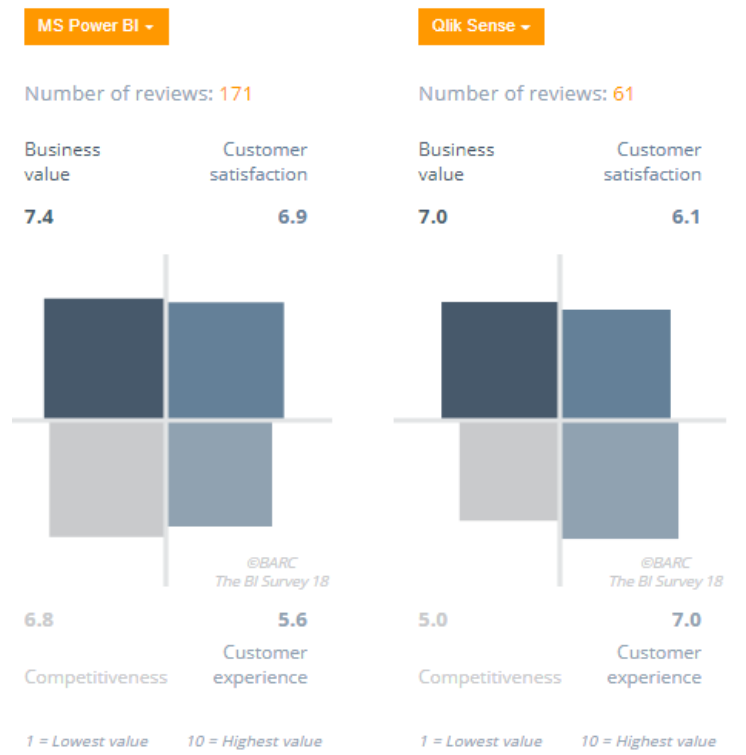


Figura 9. Evaluación de self service reporting. Recuperado de <https://bi-survey.com/business-intelligence-software-comparison>



Figura 10. Evaluación de Embedded analytics. Recuperado de <https://bi-survey.com/business-intelligence-software-comparison>

Finalmente podemos deducir que en varios aspectos de las funcionalidades que evalúa el BI Survey, Qlik sense es muy bien reconocido, si lo comparamos con MS Power BI. Se reconoce la popularidad del MS Power BI en el mercado nacional, pero apreciando esta encuesta internacional, nos abre las puertas a sustentar el uso de la herramienta Qlik Sense, por tratarse definitivamente de una herramienta de vanguardia. Adicionalmente también señalaremos que ya se viene estudiando y aplicando las diferentes funcionalidades de esta herramienta en otros proyectos, dentro de la institución, lo que nos permite con más base, haberla propuesto como parte de la solución de este trabajo.

A continuación, pasamos a describir la institución donde se ha implementado el proyecto. El Instituto Nacional de Estadística e informática – INEI. Este es un organismo técnico especializado, con personería jurídica de derecho público interno, con autonomía técnica y de gestión, dependiente del Presidente del Consejo de Ministros.

Es el organismo central y rector del Sistema Estadístico Nacional, responsable de normar, planear, dirigir, coordinar y supervisar las actividades estadísticas oficiales del país. El INEI tiene rango de Sistema Funcional y su Jefe es la máxima autoridad del Sistema Estadístico Nacional.

La estructura orgánica del Instituto Nacional de Estadística e Informática, está conformada por: La Alta Dirección, integrada por la Jefatura, Sub Jefatura de Estadística y Secretaría General; Órganos de Asesoramiento, Órganos de Apoyo, Órganos Desconcentrados, conformados por la Escuela Nacional de Estadística e Informática y el Centro de Investigación y Desarrollo; y los órganos de línea estadísticos, constituidos por la Dirección Nacional de Censos y Encuestas, Dirección Nacional de Cuentas Nacionales, Dirección Técnica de Indicadores Económicos y la Dirección Técnica de Demografía e Indicadores Sociales. A

la vez también cuenta con Oficinas en todas las capitales departamentales y en las ciudades de
Chimbote y Tarapoto.

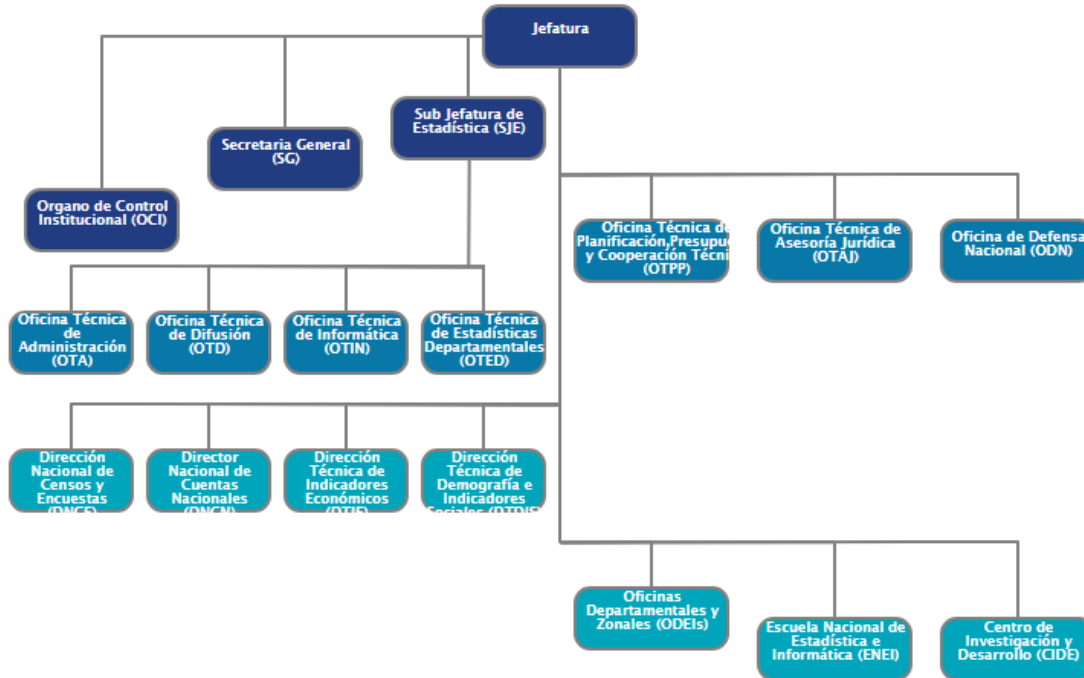


Figura 11. Organigrama Funcional del INEI. Recuperado de <http://www.inei.gob.pe>

La Oficina Técnica de Informática, es la Dirección encargada del soporte técnico institucional y del desarrollo y mantenimiento de los sistemas con fines administrativos, gestión de encuestas y difusión de resultados de los distintos proyectos que lleva a cabo la institución. Con el correr del tiempo, está claro que se han implementado una serie de aplicaciones web, de los diferentes tipos de uso por parte del usuario final; los mismos que han consumido recursos y han tenido una cobertura de usuarios dentro de la institución y fuera de la misma, de origen nacional y extranjero. Este creciente número de aplicaciones ha conllevado a su vez a una serie de problemas que hoy en día requiere que se les atienda. A continuación, para poder describir el detalle del proceso actual referido a la analítica web en el INEI, utilizaremos un diagrama de Ishikawa.

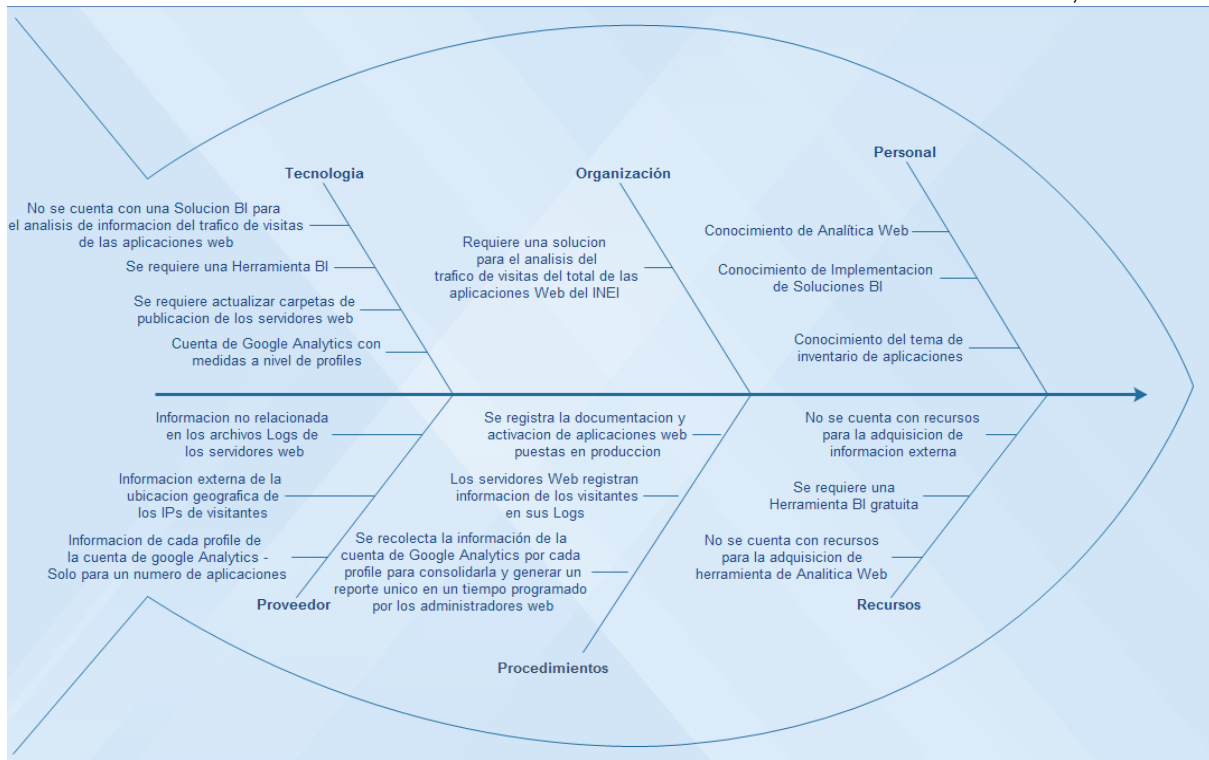


Figura 12. Diagrama de Ishikawa de la problemática actual

En la Figura 12, En el diagrama de Ishikawa describimos la realidad detallada de la problemática que trataremos en esta investigación.

- El personal de la unidad de Producción, a cargo de la tarea; cuenta con la iniciativa de dar solución al problema institucional de la presente investigación. Además de contar con conocimientos de analítica web, implementación de soluciones BI y del tema del inventario de aplicaciones.
- La alta dirección de OTIN, requiere implementar analítica web para realizar el análisis del tráfico de visitas de las aplicaciones web, con el fin solucionar los casos, del estado actual de los directorios de publicación de los servidores web, optimizar el tiempo de generar los reportes de analítica web de las aplicaciones web y también llegar a cubrir el total de las aplicaciones web.
- En los procesos actuales, no contamos con una solución BI para el análisis de la información del tráfico de visitas de las aplicaciones web. Para tal fin, se requiere

también una herramienta BI en su versión gratuita. Los servidores web están configurados para registrar sus archivos log, pero en las diferentes gestiones de la administración web no se ha realizado una actualización de las carpetas de publicación de las aplicaciones web. Por lo que hoy en día se pueden encontrar en línea, por ejemplo, aplicaciones de proyectos que han finalizado o incluso de aplicaciones de testing. También se tiene que señalar, que se dispone de una cuenta de Google Analytics que está configurado a una cantidad de directorios de publicación de aplicaciones.

- Para la implementación de la nueva solución, se cuenta con la fuente no relacionada de los archivos logs de los servidores. A esto se suma las tablas maestras disponibles de modo público que refieren a la ubicación geográfica de los IPS a nivel mundial. Así también de la información generada por la cuenta de google analytics a una cantidad de aplicaciones web.
- Los administradores web tienen la función de registrar en el inventario de aplicaciones; información de cada aplicación web que se publica en los servidores web de producción. Así mismo si alguna configuración de las mismas ha cambiado (bases de datos, url, usuarios de conexión). Configuran los servidores web para la generación de sus archivos de registro y en la actualidad, para la generación de reportes consolidados por aplicación, tienen que recolectar la información de la cuenta de Google Analytics que lo emite por cada ruta configurada, siendo en total 60 aproximadamente; este proceso lo realizan cada tres meses, durante 6 días laborales. La recolección es un proceso manual, es decir, trasladan el valor emitido en las visualizaciones de Google Analytics hacia las

pantallas del Sistema de estadísticas web que registra el número de visitas por aplicaciones. El proceso en mención se visualiza en la figura 13.

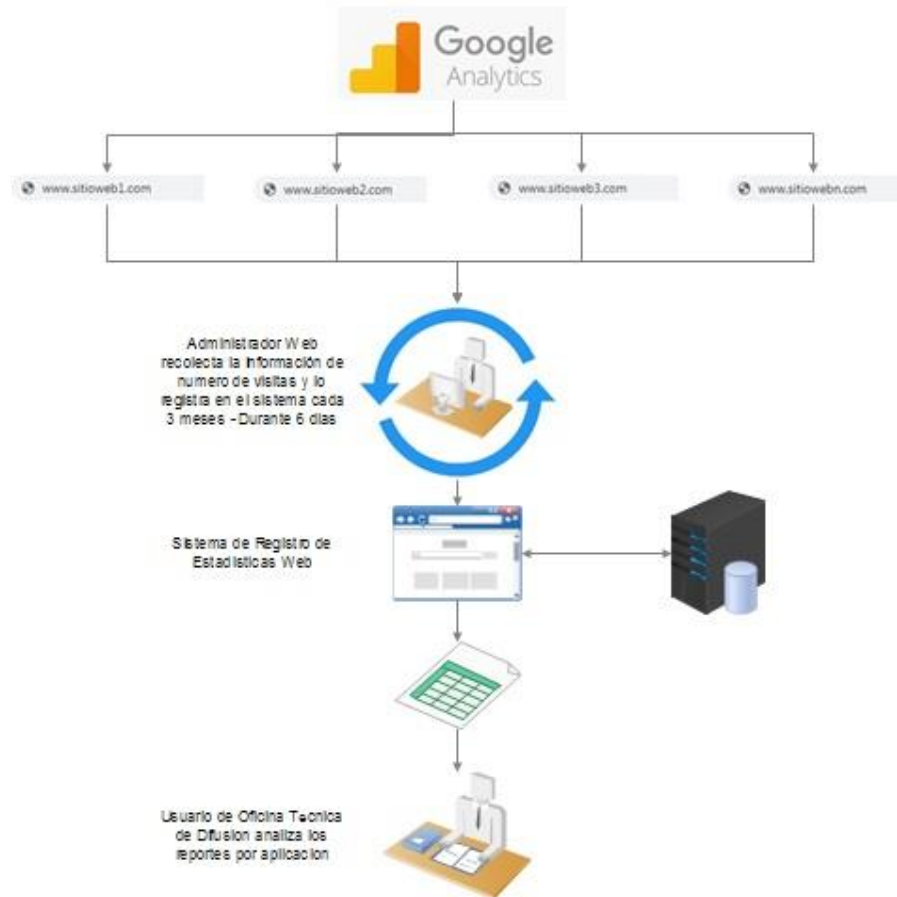


Figura 13. Proceso actual de recolección de información de la cuenta de Google Analytic

Luego el usuario de la Oficina Técnica de Difusión dispone del listado de aplicaciones con la información de número de visitas, esta métrica es la única que en la actualidad se recupera de la cuenta de Google Analytic. En el capítulo de Anexos, incluimos pantallas de la cuenta de Google Analytics como del Sistema de estadísticas web.

- No se cuenta con recursos financieros para la adquisición de información de fuentes externas. Tampoco para la adquisición de una herramienta de analítica

web ni una herramienta BI, por lo que se buscará una herramienta BI en su versión gratuita.

Ahora, revisaremos los resultados del pre test aplicado a las variables definidas en este trabajo:

Tabla 4

Análisis Pre test – Criterios de evaluación de las 3 dimensiones para la toma de decisiones

Dimensión	Medición	Escala
	<u>Número de visitas por Aplicación</u>	
Análisis del comportamiento	El Sistema de Registro de Estadísticas Web visualiza el total de visitas por aplicación, permitiendo a su vez la exportación en formato Excel, para que el usuario final pueda trabajar la información.	Escala [1 - 5]
	<u>Difusión de información estadística oficial</u>	
Análisis de resultados	Solo se mide el número de visitas por las aplicaciones de difusión	Escala [1 - 5]
	<u>Encuesta de satisfacción</u>	
Análisis de Experiencia	Aún está en fase de aprobación por parte de las direcciones responsables	Escala [1 - 5]

En la tabla 4 se visualiza los criterios evaluados a las 3 dimensiones de la toma de decisiones referente a la analítica web, cuyo indicador es visualizado y analizado a través del sistema de registro de estadísticas web. La valoración de las dimensiones se realizó a través de una ficha de evaluación, y que se aprecia en el anexo 8. Para la valoración de cada una de las

dimensiones, se utilizó la escala de Likert del 1 al 5, siendo el 1, la expresión de la mínima importancia y el 5 la máxima.

Tabla 5

Pre Test - Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web

Evaluación del personal	Análisis del comportamiento	Análisis de resultados	Análisis de la Experiencia
1. No es importante	0	0	3
2. Poco Importante	0	1	0
3. Neutral	0	2	0
4. Importante	3	0	0
5. Muy Importante	0	0	0

Fuente: Información recolectada con la ficha adjunta en el anexo 8

Tabla 6

Pre Test – Análisis de la Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web

Evaluación del personal	Pre Test Frecuencia	Pre Test Porcentaje
1. No es importante	3	33.33
2. Poco Importante	1	11.11
3. Neutral	2	22.22
4. Importante	3	33.33
5. Muy Importante	0	0
Total	9	100

Fuente: Información recolectada con la ficha adjunta en el anexo 8

Según la tabla 6, nos indica que la calificación de importancia aprobatoria de las dimensiones de toma de decisiones, con la solución anterior; es de 33.33%. Se considera aprobatoria los valores 4 y 5 de la escala Likert.

Tabla 7

Datos del Pre test de Tiempo de proceso de generación de información para la toma de decisiones

Número de Proceso	Medición	Valor Pre test
1	Días laborales	6
2	Días laborales	5
3	Días laborales	7
4	Días laborales	6
5	Días laborales	5
6	Días laborales	6
7	Días laborales	7
8	Días laborales	6
9	Días laborales	6

Fuente: Información recolectada con la ficha adjunta en el anexo 9

La tabla 7, es el resultado de la observación realizada a los procesos de generación de información con la solución anterior.

Tabla 8

Análisis Pre test – Criterios de evaluación de las 8 dimensiones de la calidad de información

Dimensiones	Proceso	Medición
Resolución de datos	La información recolectada, basada en el único indicador procesado, nos define el número de visitas que se registra en una cantidad numérica, y se encuentra dimensionada por aplicación y por número de semana del mes	Escala [1 - 5]
Estructura de datos	El indicador está definido como un tipo de dato numérico. De igual manera la dimensión, número de semana del mes.	Escala [1 - 5]

	La información no presenta casos de datos dañados o faltantes	
Integración de datos	No se realiza integración de diferentes fuentes de información, solo se recolecta la proveniente de la cuenta de Google Analytic y se adiciona en el modelo de datos del Sistema de Registro de Estadísticas Web	Escala [1 - 5]
Relevancia temporal	La recolección de información se realiza cada tres meses, lo que conlleva a que esta brecha de tiempo afecte a la relevancia de la información.	Escala [1 - 5]
Cronología de datos y meta	La información está definida en una serie de tiempo (semana de mes) y se definió el indicador número de visitas como el objetivo a medir.	Escala [1 - 5]
Generalización	La información recolectada procede de un grupo de aplicaciones de la población total de aplicaciones. Se asume que son las más representativas y que cuyo análisis se puede generalizar en el estudio total. Se desconoce el total de aplicaciones que puedan estar produciendo información de número de visitas.	Escala [1 - 5]
Operacionalización	El valor del dato se mantiene desde su origen hasta el repositorio de base de datos donde residirá para su análisis. No cambia a través de los procesos, se recopila el mismo dato generado por Google Analytics.	Escala [1 - 5]
Comunicación	La información recolectada está disponible a usuarios responsables de la supervisión del sitio web de la institución	Escala [1 - 5]

(Oficina técnica de difusión), de manera visual y en listados de archivos Excel para el estudio respectivo. El resultado del análisis no es publicado, ni internamente al personal del INEI, ni es incluido en algún boletín hacia el público externo. Esto por tratarse de información con carácter de privado.

Fuente: Información recolectada con la ficha adjunta en el anexo 6

En la tabla 8, se señala los criterios evaluados para las 8 dimensiones de la calidad de la información generada con la solución anterior. Y a las que se les aplicara una evaluación con la escala Likert.

Tabla 9

Datos del Pre test de la evaluación de la calidad de información

Evaluación del personal	Medición	Valor Pre test
Personal 1	Evaluación de las 8 dimensiones	3.19
Personal 2	Evaluación de las 8 dimensiones	2.95
Personal 3	Evaluación de las 8 dimensiones	3.15
Personal 4	Evaluación de las 8 dimensiones	3.02

Fuente: Información recolectada con la ficha adjunta en el anexo 6

En la tabla 9, se obtiene el valor obtenido de la evaluación de las 8 dimensiones por cada persona que registro sus valores en la ficha del anexo 6. El valor final de cada uno fue obtenido con la fórmula de Kenett y Shmueli.

Tabla 10

Pre Test - Preguntas de satisfacción del usuario con las interfaces de análisis

Código de pregunta	Pregunta
P1	¿El sistema es de fácil uso?
P2	¿Está usted satisfecho con la variedad de tableros que ofrece el sistema?
P3	¿La interfaz del sistema es amigable para usted?
P4	¿Los tableros emitidos son fáciles de analizar?
P5	¿Cómo considera la presentación de los tableros analíticos?
P6	¿Cómo considera el tiempo de carga de los tableros analíticos?
P7	¿La información que le brinda los tableros le ayuda a realizar una mejor analítica web?

Fuente: Información recolectada con la ficha adjunta en el anexo 10

En la tabla 10, se detallan las preguntas realizadas en la ficha de evaluación sobre la satisfacción del usuario, la misma que se adjunta en el anexo 10.

Tabla 11

Pre Test - Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis

Evaluación del personal	P1	P2	P3	P4	P5	P6	P7
1. Desaprobación Plena	0	0	0	0	0	0	2
2. Desaprobación Simple	0	3	0	0	2	0	1
3. Indecisión o Indiferencia	0	0	3	3	1	2	0
4. Aprobación Simple	3	0	0	0	0	1	0
5. Aprobación Plena	0	0	0	0	0	0	0

Fuente: Información recolectada con la ficha adjunta en el anexo 10

Tabla 12

Pre Test – Análisis de la Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis

Evaluación del personal	Pre Test	Pre Test
	Frecuencia	Porcentaje
1. Desaprobación Plena	2	9.52
2. Desaprobación Simple	6	28.57
3. Indecisión o Indiferencia	9	42.86
4. Aprobación Simple	4	19.05
5. Aprobación Plena	0	0
Total	21	100

Fuente: Información recolectada con la ficha adjunta en el anexo 10

La tabla 12, nos indica que la calificación de aprobación de satisfacción, con la solución anterior; es de 19.05%. Se considera aprobatoria para este caso, los valores 4 y 5 de la escala Likert.

Analizando las problemáticas encontradas en la realidad al inicio del proyecto, entonces, debemos definir que la presente investigación se focaliza en la implementación de una solución de inteligencia de negocios para la Oficina Técnica de Informática, con alcance en el análisis de las métricas del tráfico de visitas de las aplicaciones web del INEI, según lo establecido en la analítica web; que le permita apoyar en la toma de decisiones de manera positiva, y poder enriquecer el análisis que ya se realiza con Google Analytics. Además, también de poder contribuir con la actualización del inventario de aplicaciones, así como optimizar los tiempos de generación de reportes, contemplando el total de aplicaciones web institucionales.

Se pretende entonces, mediante la implementación de esta solución, brindar al usuario de OTIN, una plataforma que tenga acceso a un solo repositorio de datos centralizado y contar con tableros de control con el fin de optimizar, analizar, compartir y monitorear la información

a nivel de detalle, identificando en el tráfico de las visitas a las aplicaciones web, el comportamiento y segmentación de los usuarios con respecto a las mismas. Entonces, la herramienta de inteligencia de negocios debe visualizar, los resultados globales, comparativos entre valores de dimensiones, series históricas (año y mes) de la información agrupada en dimensiones como: tipo de usuario según origen (público o institucional), servidores, tipo de usuario INEI según la locación de la sede donde labora, ubicación geográfica (País, Estado y Ciudad), y a su vez, calculando indicadores de la analítica web como : número de visitantes, número de visitas, tiempo de carga, Megabytes descargados y cargados por parte del usuario, etc.

A continuación, se definen los detalles del alcance:

- Implementación basada en la metodología Ágil Scrum.
- Identificar las métricas a implementar según la analítica web.
- La extracción y procesamiento de datos se realizará de los registros web de cinco servidores web IIS de Microsoft, contemplando el total de sus aplicaciones. Para el presente trabajo, no se han tomado en cuenta los servidores que no son plataforma Microsoft.
- El script de carga de los archivos de registros web, será codificado en Store Procedures de MS SQL.
- El datawarehouse residirá en una base de datos MS SQL.
- Se tiene que integrar otras fuentes de información, como la base de datos de la relación de IPs con su ubicación geográfica y la información del inventario de aplicaciones.
- El desarrollo de los tableros se realizará con la herramienta BI, Qlik Sense Desktop, en su versión gratuita.

- La solución BI será utilizada por las unidades de Control de Calidad, Infraestructura, Desarrollo y Producción para los fines correspondientes de cada fin que persiguen.
- Para aplicar predicciones a la información obtenida, se utilizará la herramienta Studio R.

Como Limitaciones se encontraron los siguientes:

- La cuenta de Google Analytics que registra la información de las aplicaciones web del INEI, lo realiza por cada aplicación configurada en un profile. Entonces se tienen tantos profiles como aplicaciones web. Lo que dificulta la tarea de recolección de datos.
- La herramienta BI Qlik Sense en su versión gratuita no contempla la operabilidad directa con las librerías de Studio R, para poder implementar predicciones automáticas desde el entorno de la herramienta. Por lo que tenemos que invertir un tiempo para aplicar el proceso de predicción, convirtiendo la información de la base de datos SQL en el formato propio de la herramienta, en este caso, Studio R. Luego aplicar las funciones de predicción de serie de tiempo, y el resultado obtenido volver a exportarla hacia la data en el datawarehouse.

A continuación describiremos las tesis recopiladas y consideradas como Antecedentes para el presente trabajo de investigación. Miguel Angello Alva Quiliche y Milagros Roxana Callan Carbajal para su investigación “SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DIRECCIÓN EJECUTIVA DEL HOSPITAL LA CALETA” encontró como problema, que la dirección ejecutiva del Hospital La Caleta necesita tomar decisiones oportunas y acertadas, ya que es vital para una mejora significativa en la atención a las personas que acuden al centro de salud, que en su mayoría son

personas provenientes de provincia o de lugares alejados al distrito donde se encuentra el hospital. Las decisiones tomadas por el directorio se ven reflejadas en: la atención que recibe el paciente, el tiempo que demora en ser atendido un paciente, adquisición de tecnología de última generación, políticas de crecimiento y expansión de servicios. Entonces se plantearon el siguiente objetivo general: Mejorar la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta a través de una solución de inteligencia de negocios. Sus objetivos específicos fueron los siguientes: realizar el levantamiento de los requerimientos, analizar la fuente de datos del sistema operacional, diseñar la solución de inteligencia de negocios, implementar la solución de inteligencia de negocios, realizar Pruebas mediante solución de inteligencia de negocios.

En este trabajo se define las siguientes conclusiones: la metodología Hefesto es eficiente para investigaciones relacionadas a la construcción de un Datawarehouse y flexible ya que es independientemente de las herramientas que se usan para su implementación. El esquema en estrella es el más simple de interpretar y optimiza tiempo en las acciones propias para diseñar las dimensiones, hechos y las uniones pertinentes entre estas tablas. Pentaho es una herramienta muy intuitiva, manejable, multiplataforma, sencilla al momento de integrar los datos y lo más importante que es Open Source. La integración de información se realizó con la herramienta Pentaho Data Integration y los reportes de los usuarios se muestran en un entorno web con Pentaho BI Server.

José Luis Carhuallanqui Bastidas, en su investigación “Diseño de una solución de inteligencia de negocios como herramienta de apoyo a la toma de decisiones en el área de ventas de la empresa farmacéutica Dispefarma”, se plantea como objetivo general, diseñar una solución de inteligencia de negocios que brinde información oportuna para la toma de decisiones en el área ventas de la empresa Dispefarma. Por otro lado, definió como objetivos

específicos los siguientes ítems: reducir el tiempo de generación de reportes con una solución de Inteligencia de Negocios, incrementar el nivel de Calidad de la información con una solución de inteligencia de negocios, mejorar el nivel de satisfacción sobre la información solicitada.

En este trabajo se obtuvo como conclusiones lo siguiente, La solución de Inteligencia de Negocios mejora significativamente la toma de decisiones en el área de Ventas de la empresa Dispefarma. Se redujo el tiempo promedio en elaboración de informes solicitados por la Gerencia. Se realizaron Dashboards e indicadores de Gestión para el proceso de ventas de la organización bajo diversas perspectivas que permite un mejor análisis de la información. Se identificó los mejores clientes por producto y los clientes fidelizados en el tiempo lo cual facilitará el desarrollo de estos clientes a fin de ofrecerles nuevas promociones, descuentos o campañas especiales por su lealtad a la empresa. La solución de Inteligencia de Negocios mejoró significativamente la calidad de la información haciendo más confiables los reportes generados a partir de ella brindando un soporte para la toma de decisiones en la empresa.

Gustavo Fernando Zegarra Fuentes, en la investigación “SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS ORIENTADA A MEJORAR LA TOMA DE DECISIONES EN LAS OPERACIONES MINERAS DE EXTRACCIÓN Y METALURGIA DE HOCHSCHILD MINING”; plantea como objetivo general lo siguiente: Brindar soporte a la toma de decisiones empleando una solución de inteligencia de negocios de las operaciones de extracción y metalurgia diseñada e Implementada a la medida de Hochschild Mining. Y como objetivos específicos lo siguiente: Aumentar la eficacia en la toma de decisiones de la Gerencia de Operaciones de la Minera, acelerar el proceso de obtención de reportes del estado de operaciones mineras a 12 horas, realizar las mediciones de los procesos de extracción y

metalurgia haciendo un uso eficiente de los recursos humanos y de TI de la minera, disponer en todo momento de información de las operaciones mineras de extracción y metalurgia.

Finalmente define como conclusiones lo siguiente: Se evidenció que en el proceso de toma de decisiones a nivel de la gerencia de operaciones mineras, el tener una alta disponibilidad de la información, así como una síntesis fiable de la misma sobre los procesos de extracción y metalurgia contribuyó en una mejora en la toma de decisiones, lo que propició la consecución de metas. Se comprobó una disminución del tiempo que se tomaba para la elaboración de reportes de estado de operaciones mineras de 2 a 3 días hábiles a un intervalo entre 6 y 18 horas con un tiempo medio de tan solo 12 horas. Se concluye que se logró un mejor aprovechamiento de recursos tanto humanos, con menor carga de horas-hombre por día, así como tecnológicos al optimizar el uso de infraestructura para implementar la solución de inteligencia de negocios. Se evidenció que mediante esta solución de inteligencia de negocio se logró una alta disponibilidad de la información sin perjudicar la performance del servidor principal de Base de datos de la minera.

Ahora procederemos a describir la justificación del presente trabajo de investigación, entonces diremos que nos basamos en lo sostenido por Kaushik (2007), donde manifiesta que, cada vez más, una piedra angular en la implementación de una estrategia de analítica web óptima, es tener un entorno de datawarehouse que le permite tener datos más complejos en un entorno fácil de producir información. Los datawarehouse suelen ser muy flexibles cuando se trata de importar datos de distintas fuentes externas, en consecuencia, le permiten al usuario realizar el análisis requerido.

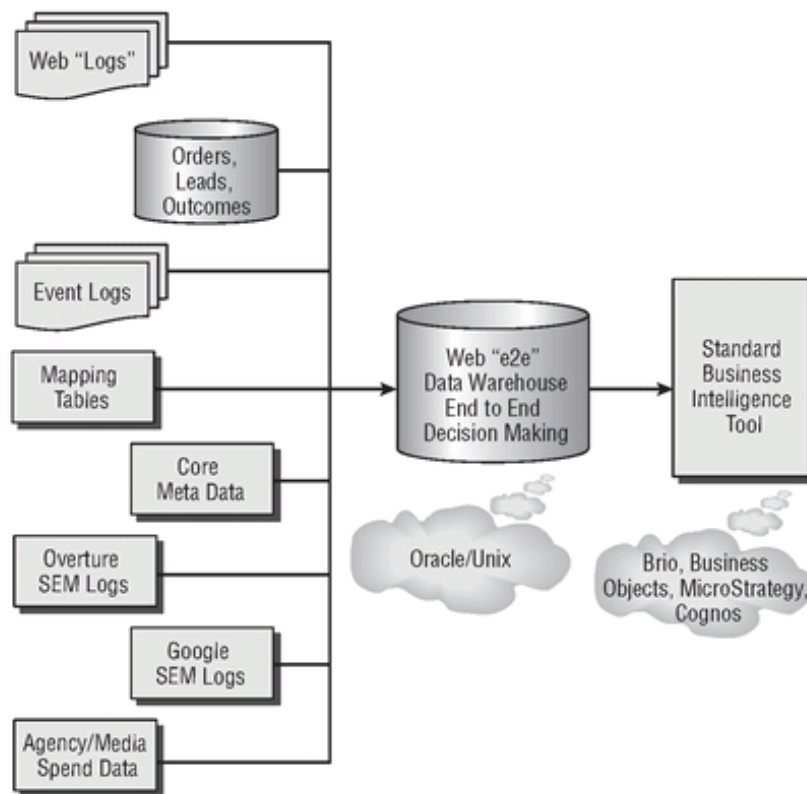


Figura 14. Como se visualiza un datawarehouse Web e2e. De "Web Analytics One Hour One Day" por Kaushik, 2007. Copyright 2007 por Wiley Publishing.

Tener implementado una solución basado en un datawarehouse, nos brinda una gran flexibilidad en términos de trabajar con muchas más fuentes de datos (por ejemplo, registros de los servidores, otras fuentes de Internet, por ejemplo, la base de datos de IPs mundial, datos de búsqueda de Google, datos de otras oficinas de la empresa y otros canales de datos como redes sociales). Esto permite crear realmente una vista de end to end (e2e) del comportamiento y los resultados del tráfico de visitas de manera efectiva a través del tiempo.

1.2. Formulación del problema

¿En qué medida la aplicación de Analítica Web permite la Toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?

Problemas específicos:

¿De qué manera la aplicación de Analítica Web influye en el tiempo de generación de información para la toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?

¿De qué manera la aplicación de Analítica Web influye en la calidad de información para la toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?

¿De qué manera la aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?

1.3. Objetivos

1.3.1. Objetivo general

Determinar en qué medida la aplicación de Analítica Web permite la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

1.3.2. Objetivos específicos

Acelerar el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

Determinar en qué medida la aplicación de Analítica Web influye en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

Determinar en qué medida la aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

1.4. Hipótesis

A continuación, se describe la hipótesis general y las específicas.

1.4.1. Hipótesis general

La aplicación de Analítica Web, influye significativamente en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

1.4.2. Hipótesis específicas

La aplicación de Analítica Web, acelera el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

La aplicación de Analítica Web influye en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

La aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.

CAPÍTULO II. METODOLOGÍA

2.1. Tipo de investigación

Según Zorrilla (1993), la investigación se clasifica en cuatro tipos: básica, aplicada, documental, de campo o mixta. Entonces, según el uso que se pretende dar al conocimiento estudiado en el presente trabajo, enmarcamos este, en el tipo de Investigación Aplicada. La investigación aplicada, guarda íntima relación con la básica, pues depende de los descubrimientos y avances de la investigación básica y se enriquece con ellos, pero se caracteriza por su interés en la aplicación, utilización y consecuencias prácticas de los conocimientos. La investigación aplicada busca el conocer para hacer, para actuar, para construir, para modificar. Por otro lado, si evaluamos el criterio del lugar y los recursos de donde se obtiene la información, definimos el tipo de investigación como Mixta. La investigación documental es aquella que se realiza a través de la consulta de documentos (libros, revistas, periódicos, memorias, anuarios, registros, códigos, constituciones, etc.). La de campo o investigación directa es la que se efectúa en el lugar y tiempo en que ocurren los fenómenos objeto de estudio. Entonces finalmente, la investigación mixta es aquella que participa de la naturaleza de la investigación documental y de la investigación de campo.

Por otro lado, Carlos Sabino (1996) nos presenta la siguiente clasificación: Por su nivel de aplicación y por el tratamiento que se da al objeto de estudio. Analizando los tipos investigación que se encuentran en el primer criterio definimos este trabajo en tipo de

Investigación Aplicada. Las investigaciones aplicadas son la respuesta efectiva y fundamentada a un problema detectado, analizado y descrito. La investigación aplicada concentra su atención en las posibilidades fácticas de llevar a la práctica las teorías generales, y destina sus esfuerzos a resolver los problemas y necesidades que se plantean los hombres en sociedad en un corto, mediano o largo plazo. Y bajo el criterio del tratamiento que se da al objeto de estudio, tenemos tres clasificaciones: exploratorias, descriptivas y explicativas. Por consiguiente, este trabajo es definido del tipo explicativo.

Los estudios explicativos no solo coberturan la descripción de conceptos o las relaciones entre ellas. Realizan la verificación de una hipótesis partiendo de material teórico, identifican las relaciones de causalidad y proponen nuevas hipótesis para futuros estudios. Este tipo de investigación exige una alta concentración y capacidad de análisis y síntesis por parte del investigador, ya que las variables en estudio deben ser meticulosamente estudiadas.

2.2. Población y muestra (Materiales, instrumentos y métodos)

La población a estudiar en el presente trabajo de investigación, son los 13 tableros que visualizan la información del datawarehouse, que se genera e integra de diferentes fuentes de información. Y que luego de ser dimensionada y calculado sus indicadores, se obtiene información de 6 indicadores para su análisis:

- Número de visitas: total y agrupado por diferentes dimensiones.
- Número de visitantes: total y agrupado por diferentes dimensiones.
- Tiempo de carga a las aplicaciones: total y agrupado por diferentes dimensiones.
- Tamaño de información consumida desde las aplicaciones: total y agrupado por diferentes dimensiones.
- Tamaño de información cargada hacia las aplicaciones: total y agrupado por diferentes dimensiones.

- Casos de Error en las aplicaciones: total y agrupado por diferentes dimensiones.

En la siguiente tabla (ver Tabla 13) detallaremos los materiales que emplearemos para la investigación.

Tabla 13

Materiales utilizados

TIPO DE MATERIALES	DESCRIPCIÓN
Equipos computacionales	<ul style="list-style-type: none"> • PC Dell Intel Core i7-6700 CPU 3.40GHz / 16 GB RAM • Impresora Canon MP280
Software	<ul style="list-style-type: none"> • MS-SQL SERVER 2016 • Qlik Sense Desktop • R Studio
Servidor	<ul style="list-style-type: none"> • Sistema Operativo: Windows Server 2008 R2 Enterprise SP1 • RAM: 16GB • HDD: 1.4TB (dividido en 3 unidades) • Version MS -SQL: Microsoft SQL Server 2008 R2 (SP2) - (X64)
Materiales de escritorio	<ul style="list-style-type: none"> • Papel bond A4 80 g • Lapiceros • Cuaderno • Folder plastificado • Resaltadores • Clips • Tajadores
Servicios	<ul style="list-style-type: none"> • Agua. • Luz. • Internet. • Impresiones.

En la siguiente tabla (ver Tabla 14) detallaremos los instrumentos que se emplearon para la recolección de información.

Tabla 14

Instrumentos de recolección de información

INSTRUMENTOS
Anexo 6 – Ficha de Evaluación de las dimensiones de la calidad de la información.
Anexo 8 – Ficha de Evaluación de las dimensiones de toma de decisión en la analítica web.
Anexo 9 – Ficha de Observación para el registro del indicador: Tiempo de generación de información para la toma de decisiones.
Anexo 10 – Ficha de Evaluación de las preguntas sobre la satisfacción del usuario con las interfaces de análisis.

2.3. Procedimiento

En la actualidad, la Oficina Técnica de Informática viene impulsando en el personal de los equipos de analistas de sistemas y desarrolladores, el poder implementar soluciones informáticas con herramientas de última tecnología. Para nuestro caso, el conocimiento sobre inteligencia de negocios y analítica web, fueron los impulsores para ir en búsqueda de casos en donde poder implementar. En realidad, en el INEI, podemos encontrar un mar de casos que esperan ser atendidos. EL análisis de información es un segmento importante de este nicho de casos. Por conocimiento propio de interactuar con los administradores web, se observó el caso de la información recopilada y visualizada de Google Analytics. Por lo que se encontró fácilmente donde poder aplicar los conocimientos de analítica e inteligencia de negocios. Además de procesamiento de información y de diseño de datawarehouse. Es decir, todos los insumos para poder construir una solución de inteligencia de negocios donde aplicar la

Analítica Web con el objetivo de dar solución a problemas en los procedimientos actuales y también poder contribuir a una óptima toma de decisiones.

Para el desarrollo del producto se utilizó la metodología de desarrollo ágil SCRUM. La misma que como metodología sustentamos su elección anteriormente. Pero a esto debemos sumar, la premisa de la dirección de OTIN que impulsa la aplicación de nuevas tecnologías en todos los campos de la construcción de sistemas.

2.3.1 Cronograma

Tabla 15

Cronograma

Ítem	Tarea	Duración (días)	Inicio	Fin
1	Cronograma del proyecto	88	02/09/2019	25/09/2019
2	Inicio	1	02/09/2019	02/09/2019
3	Planificación	3	03/09/2019	05/09/2019
4	Sprints	81	06/09/2019	17/01/2020
5	Sprint 1	18	06/09/2019	01/08/2019
6	Análisis y Diseño	4	06/09/2019	11/09/2019
7	Construcción	11	12/09/2019	26/09/2019
8	Despliegue	3	27/09/2019	01/10/2019
10	Sprint 2	16	02/10/2019	06/11/2019
11	Análisis y Diseño	4	02/10/2019	07/10/2019
12	Construcción	22	08/10/2019	31/10/2019
13	Despliegue	4	01/11/2019	06/11/2019
14	Sprint 3	8	07/11/2019	18/11/2019
15	Análisis y Diseño	2	07/11/2019	08/11/2019
16	Construcción	4	11/11/2019	14/11/2019
17	Despliegue	2	15/11/2019	18/11/2019
18	Sprint 4	12	19/11/2019	04/12/2019
19	Análisis y Diseño	2	19/11/2019	20/11/2019
20	Construcción	8	21/11/2019	02/12/2019
21	Despliegue	2	03/12/2019	04/12/2019
22	Sprint 5	27	05/12/2019	17/01/2020

23	Construcción	22	05/12/2019	03/01/2020
24	Despliegue	10	06/01/2020	17/01/2020
25	Seguimiento y Control	89	06/09/2019	17/01/2020
26	Cierre	5	20/01/2020	24/01/2020

2.3.2 Roles y Responsabilidades

Tabla 16

Roles

ROLES	RESPONSABILIDADES
Jefe de Proyecto	<ul style="list-style-type: none"> • Crear el plan de proyecto. • Establece los plazos, fases y entregables del proyecto. • Controla el alcance del proyecto. • Organizar el entorno de trabajo para los miembros del equipo central. • Monitorea el avance del proyecto, el desempeño y las necesidades del equipo en general. • Asigna los roles, responsabilidades y tareas a cada miembro del equipo. • Reporta el avance del proyecto al comité de seguimiento. • Gestionar los riesgos del proyecto y las expectativas de los usuarios. • Informa con anticipación y propone alternativas de solución al comité de seguimiento sobre cualquier problema que genere atraso o inconveniente en la ejecución del proyecto. • Ayuda a garantizar que se cumplan los hitos. • Promover la comunicación y coordinación en todo el equipo.
Analista de Calidad	<ul style="list-style-type: none"> • Encargado del levantamiento de información. • Elaboración de los prototipos (Storyboarding). • Elaboración del documento de especificaciones. • Elaboración de Plan e informe de Pruebas. • Elaboración de Manual de usuario. • Validación del producto.
Especialista BI	<ul style="list-style-type: none"> • Definición del Modelo de datos • Creación de los Procesos ETL.

- Traducir los requisitos de información empresarial en aplicaciones flexibles.
- Comprender el uso, los matices y la arquitectura de la herramienta de presentación que se utiliza. En nuestro caso Qlik Sense.
- Diseño y construcción de los tableros de análisis de información.
- Diseño y construcción de las interfaces especiales requeridas.

2.3.3 Sprint Backlog

En el Sprint Backlog se identifica los requerimientos a desarrollar en la implementación. Cada requerimiento es catalogado en prioridad según la necesidad de la organización y llevada a cabo por un usuario líder, el Product Owner. Para lograr un Sprint Backlog óptimo, se requiere de constante y fina interacción con el usuario de negocio, para poder identificar conjuntamente los requerimientos necesarios en la implementación de la solución BI.

Tabla 17

Sprint Backlogs

SPRINT	ACTIVIDADES
Sprint I: Especificar los requerimientos del negocio.	Creación de historias de usuario.
	Creación de Requerimientos Funcionales.
	Creación de Requerimientos No Funcionales.
	Creación de Arquitectura de BI.
Sprint II: Diseño del modelo dimensional.	Creación de Store Procedure
	Creación de Base de Datos Modelo Dimensional.
	Creación de Diccionario de Datos

Sprint III. Definición de Indicadores a desarrollar	Identificación de Indicadores Claves de Gestión.
Sprint IV. Diseño de Prototipos	Creación de Prototipos de tablero.
Sprint V: Explotación de los indicadores, mediante los componentes gráficos de la herramienta BI	Desarrollo de tableros de análisis Desarrollo de Pruebas
Sprint VI: Elaboración de Manuales	Guía de Usuario Manual del Sistema

2.3.4 Desarrollo Sprint 1

El Sprint I, comprenderá el levantamiento preliminar de información, desarrollo de las historias de usuarios, definición de los requerimientos funcionales y no funcionales y el diseño de la arquitectura de la solución BI a implementar.

2.3.4.1 Entregables

Tabla 18

Entregables Sprint 1

ENTREGABLES
Creación de Historias de usuario.
Creación de Requerimientos Funcionales.
Creación de Requerimientos No Funcionales.
Creación de Arquitectura de BI.

2.3.4.2 Historia de Usuarios

La historia de usuario, es la representación de un requerimiento de parte del actor usuario. En donde el mismo explica en un lenguaje común lo que necesita que el sistema atienda para dar solución a su requerimiento.

A continuación, se muestra un extracto de las historias de usuarios:

Tabla 19

Control de Acceso a usuarios

Historia de usuario	
Número: 1	Usuario: Todos
<u>Quiero:</u> Controlar acceso a Usuarios	
<u>Para:</u> Mantener la seguridad del aplicativo y acceso a los datos a analizar.	
<u>Descripción:</u>	
Antes de iniciar la aplicación se solicita el nombre de usuario y su clave para que tenga acceso a las pantallas de análisis que corresponda.	
En un inicio existe un solo usuario: usu_analista	
Se puede crear otros usuarios con accesos a diferentes tableros de análisis.	

Tabla 20

Visualización de total de información cargada

Historia de usuario	
Número: 2	Usuario: usu_analista
<u>Quiero:</u> Visualizar las características del universo de información cargada en la fact_table. Con totales de registros por servidor, año, tipo de usuario, estado de navegación, registros que si/no se le asignó la ubicación geográfica, etc.	
<u>Para:</u> Obtener resultados generales de las características del universo de información cargada.	

Descripción:

El dashboard deberá permitir la visualización del total de registros cargados por servidor, año, tipo de usuario, estado de navegación, registros que si/no se le asignó la ubicación geográfica.

Tabla 21

Análisis de información de aplicaciones

Historia de usuario

Número: 3 Usuario: usu_analista

Quiero: Visualizar las características de la información catalogada por la dimensión Aplicación Web

Para: Analizar la información según la dimensión Aplicación Web

Descripción:

El dashboard deberá permitir el análisis de las aplicaciones web por año, servidor, tipo de aplicación, estado de línea.

Tabla 22

Visualización de características de información de Visitantes y Visitas

Historia de usuario

Número: 4 Usuario: usu_analista

Quiero: Visualizar la información catalogada por las características del número de visitantes y visitas.

Para: Analizar la información según del número de visitantes y visitas.

Descripción:

El dashboard deberá permitir el análisis del número de visitantes y visitas según las dimensiones de servidor, año, estado de navegación, tipo de aplicación, Sede de local el INEI, Origen General (Nacional o extranjero).

Tabla 23

Análisis de información según estado de navegación

Historia de usuario	
Número: 5	Usuario: usu_analista
<u>Quiero:</u> Visualizar las características de la información catalogada por la dimensión Estado de navegación.	
<u>Para:</u> Analizar la información según la dimensión Estado de navegación.	
<u>Descripción:</u> El dashboard deberá permitir el análisis de la información según la dimensión Estado de navegación, adicionando al mismo, las dimensiones: servidor, año, estado de navegación general (Con Error, Sin Error), Código de estado de navegación, tipo de aplicación.	

Tabla 24

Análisis de información geográfica por País según número de visitantes

Historia de usuario	
Número: 6	Usuario: usu_analista
<u>Quiero:</u> Visualizar las características de la información catalogada por la dimensión País según el número de visitantes.	
<u>Para:</u> Analizar la información de la dimensión País según el número de visitantes	
<u>Descripción:</u> El dashboard deberá permitir el análisis geográfico por País según el número de visitantes. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.	

Tabla 25

Análisis de información geográfica por País según número de visitas

Historia de usuario

Número: 7 Usuario: usu_analista

Quiero: Visualizar las características de la información catalogada por la dimensión País según el número de visitas.

Para: Analizar la información de la dimensión País según el número de visitas

Descripción:

El dashboard deberá permitir el análisis geográfico por País según el número de visitas. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

Tabla 26

Análisis de información geográfica por Ciudad según número de visitantes

Historia de usuario

Número: 8 Usuario: usu_analista

Quiero: Visualizar las características de la información catalogada por la dimensión Ciudad según el número de visitantes.

Para: Analizar la información de la dimensión Ciudad según el número de visitantes.

Descripción:

El dashboard deberá permitir el análisis geográfico por Ciudad según el número de visitantes. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

Tabla 27

Análisis de información geográfica por Ciudad según número de visitas

Historia de usuario

Número: 9

Usuario: usu_analista

Quiero: Visualizar las características de la información catalogada por la dimensión Ciudad según el número de visitas.

Para: Analizar la información de la dimensión Ciudad según el número de visitas.

Descripción:

El dashboard deberá permitir el análisis geográfico por Ciudad según el número de visitas. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

Tabla 28

Análisis de listado de aplicaciones

Historia de usuario

Número: 10

Usuario: usu_analista

Quiero: Visualizar el listado de aplicaciones, resultado del proceso de carga.

Para: Analizar la información por cada aplicación

Descripción:

El dashboard deberá permitir el análisis de cada aplicación (número de visitantes, número de visitas, tiempo de carga) según las dimensiones servidor, año, mes, tipo de aplicación, aplicación web, estado de navegación, estado de línea.

Tabla 29

Análisis por Mes de diferentes indicadores

Historia de usuario

Número: 11

Usuario: usu_analista

Quiero: Visualizar la información por mes de diferentes indicadores.

Para: Analizar la información por mes de diferentes indicadores.

Descripción:

El dashboard deberá permitir el análisis del comportamiento de los indicadores de las aplicaciones por mes de cada año. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de navegación, tipo de usuario y estado de línea.

Tabla 30

Análisis indicadores por Departamentos del Perú

Historia de usuario

Número: 12

Usuario: usu_analista

Quiero: Visualizar la información de indicadores por departamento del Perú.

Para: Analizar la información de indicadores por departamento del Perú.

Descripción:

El dashboard deberá permitir el análisis de indicadores por departamento del Perú. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación y aplicación web.

Tabla 31

Predicción por Número de Visitantes

Historia de usuario

Número: 13

Usuario: usu_analista

Quiero: Visualizar el número de visitantes por mes de los años cargados y la predicción del número de visitantes para los años 2020, 2021 y 2022.

Para: Analizar la información por mes del número de visitantes de los años cargados y la predicción del número de visitantes para los años 2020, 2021 y 2022.

Descripción:

El dashboard deberá permitir el análisis del número de visitantes de los años cargados y la predicción del número de visitantes para los años 2020, 2021 y 2022 por Servidor.

Tabla 32

Predicción por Número de Visitas

Historia de usuario

Número: 14

Usuario: usu_analista

Quiero: Visualizar el número de visitas por mes de los años cargados y la predicción del número de visitas para los años 2020, 2021 y 2022.

Para: Analizar la información por mes del número de visitas de los años cargados y la predicción del número de visitas para los años 2020, 2021 y 2022.

Descripción:

El dashboard deberá permitir el análisis del número de visitas de los años cargados y la predicción del número de visitas para los años 2020, 2021 y 2022 por Servidor.

2.3.4.3 Diagrama UML de Requerimientos Funcionales

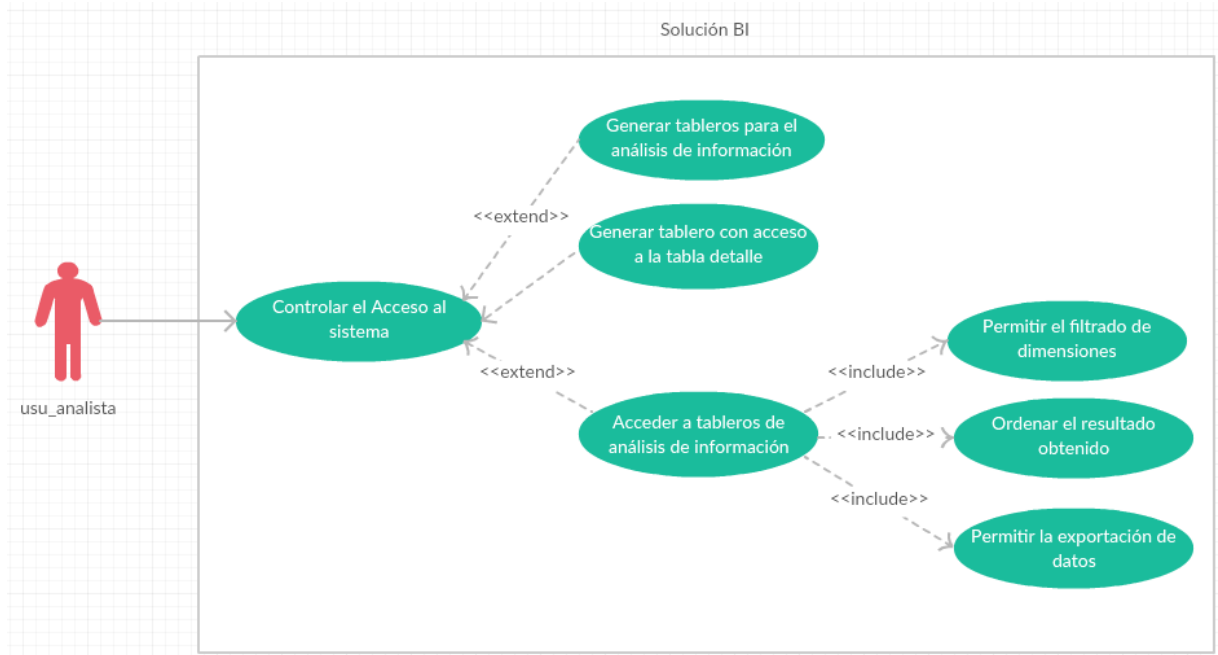


Figura 15. Diagrama UML de Requerimientos Funcionales

En la figura 15 podemos apreciar los requerimientos funcionales definidos en nuestro análisis, y también se incluye las funcionalidades propias de cada interface de nuestra solución implementada.

2.3.4.4 Requerimientos Funcionales

Tabla 33

Requerimientos Funcionales

Requerimiento	Prioridad	Descripción
REQ_CAS Controlar el Acceso al sistema	Alta	Deben existir el perfil de usuario: usu_analista, con accesos a los distintos tableros.
REQ_TAN Generar tableros para el análisis de información	Alta	Crear tableros que permitan la generación de Análisis de los indicadores en base a las variadas dimensiones.

REQ_TDE Generar tablero con acceso a la tabla detalle (Fact Table)	Alta	Crear tableros que permitan visualizar la data granular cargada en la Fact Table, con el fin de obtener información por cada solicitud.
REQ_FIL Permitir el filtrado de dimensiones	Alta	Realizar filtrados para obtener diferentes valores en las operaciones que se apliquen a los indicadores. Estos filtrados deben permitir al usuario mantenerlos activos en todas las pantallas.
REQ_ORD Ordenar el resultado obtenido	Alta	Al visualizar un resultado en tabla, el sistema debe permitir ordenar el total por alguno de los campos.
REQ_EXP Permitir la exportación de datos	Alta	Las interfaces de los tableros deben permitir la exportación en formato Excel y PDF, de la data o los gráficos, según sea el caso
REQ_TAB Acceder a tableros de análisis de información	Alta	Estos tableros deben permitir la generación de Análisis de los indicadores en base a las variadas dimensiones.

2.3.4.5 Requerimientos No Funcionales

Tabla 34

Requerimientos No Funcionales

Código	Tipo	Descripción
--------	------	-------------

SOF_001	Software	Se requiere instalar la Herramienta BI en la PC de los usuarios que analizarán la información.
SOF_002	Software	Disponibilidad de conectarse a diferentes tipos de bases de datos.
USA_001	Usabilidad	Los usuarios serán capaces de encontrar el material de manual de usuario rápidamente.
USA_002	Usabilidad	Los usuarios estarán en condiciones de utilizar correctamente cualquier funcionalidad principal del sistema tras la lectura del material de manual de usuario.
POR_001	Portabilidad	En caso de que el sistema evolucione a una nueva versión, las funcionalidades actuales no se verán afectadas.

2.3.4.6 Diseño de la Arquitectura

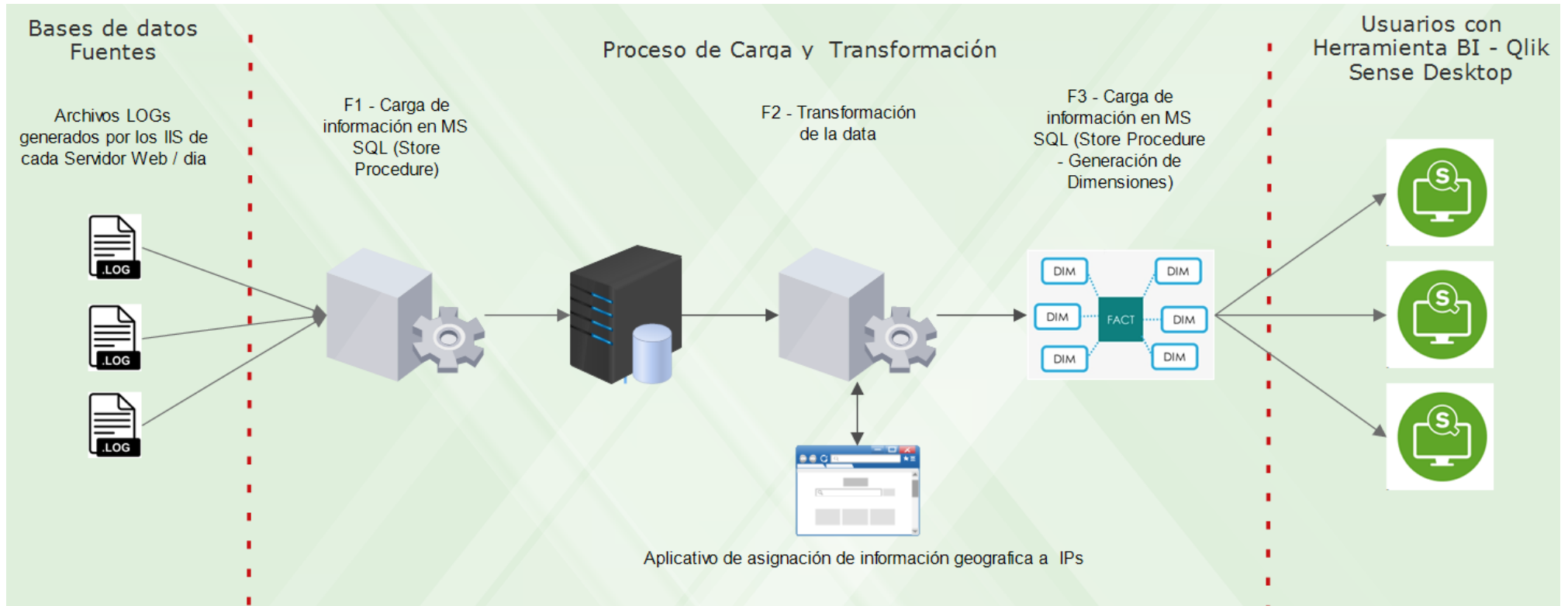


Figura 16. Arquitectura de solución BI

En la figura 16, podemos apreciar los componentes de la solución BI que se propone en este trabajo de investigación.

- Bases de datos fuentes: Los archivos de registro que generan los Servidores Web IIS de Microsoft, con que cuenta la institución
- Proceso de Carga y Transformación: En esta fase se incluye los procedimientos de carga, que capturan la data de los archivos de registro y los procesos de transformación, que asignan los valores de las dimensiones de aplicación web y Ubicación geográfica de los IPs. Por último, se ejecuta el proceso de asignación de las demás dimensiones y el cálculo de los indicadores en el modelo final del datawarehouse.
- Base de datos: la información reside en el motor de Base de datos de MS - SQL. Bajo el modelo estrella de datawarehouse.
- Análisis de la información: Incluye la implementación de los diferentes tableros en la herramienta BI que permite la conexión a la datawarehouse y el análisis de la información a través de indicadores y gráficos de serie de tiempo, filtrados por las diferentes dimensiones definidas.

2.3.5 Desarrollo Sprint II

En el Sprint II, se diseña la base de datos donde se carga la información de los archivos de registro.

2.3.5.1 Entregables

Tabla 35

Entregables Sprint II

ENTREGABLES
Creación de Proceso de carga de información.
Creación de Base de Datos Modelo Dimensional.
Creación de Diccionario de Datos

2.3.5.2 Proceso de carga de información

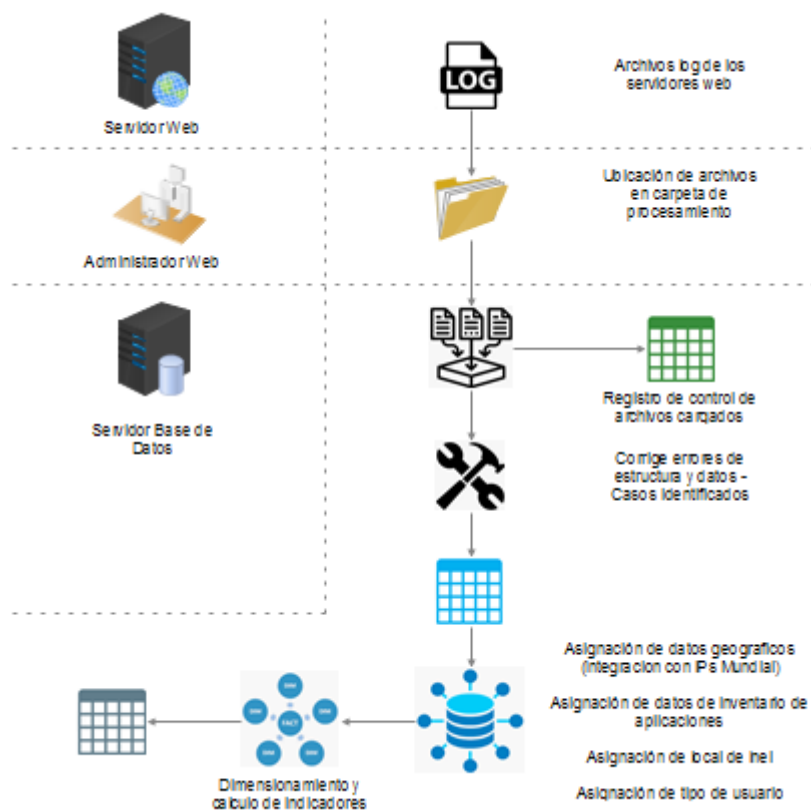


Figura 17. Proceso de carga de información

En la figura 17, se grafica los actores y cada subproceso, del proceso de carga de información en el motor de base de datos MS-SQL. En primera instancia los archivos log de los servidores web son trasladados por el administrador web hacia una carpeta en el servidor MS-SQL, servidor donde se realizará el procesamiento. El Store Procedure inicia su procesamiento generando un registro en una tabla

control para poder registrar el archivo que se ha procesado. Luego corrige algunos casos de estructura y de codificación. Cabe señalar que estos casos ya fueron identificados en las primeras pruebas, por lo que se incluyó su corrección en el algoritmo. Con la data ya cargada, se procede a asignarle la información geográfica relacionada a su IP por cada registro. Así también se asigna la información relacionada al tipo de aplicación, local del INEI y del tipo de usuario. Para luego proceder con el dimensionamiento y cálculo de los indicadores.

2.3.5.3 Base de datos Dimensional

ip2location_db3	Proc_Carga_Log	BI_LOGAPLICACIONES
ip_from	[id-proc]	ANIO
ip_to	[proceso-evento]	MES
country_code	DATE	MES_NOMBRE
country_name	TIME	APLIC_URL
region_name	[s-ip]	NAV_STATUS
city_name	[cs-method]	INEI_LOCAL
	[cs-uri-stem]	USUARIO_TIPO
	[cs-uri-stem-directorio]	NAV_IP_PAIS_COD
	[cs-uri-query]	NAV_IP_PAIS_NOMBRE
	[s-port]	NAV_IP_REGION_NOMBRE
	[cs-username]	NAV_IP_CIUADAD_NOMBRE
	[c-ip]	NRO_VISITANTES
	[cs(User-Agent)]	NRO_VISITAS
	[sc-STATUS]	TIEMPO_CARGA_MAX
	[sc-substatus]	TIEMPO_CARGA_MIN
	[sc-win32-STATUS]	TIEMPO_CARGA_PRO
	[time-taken]	SERVIDOR
	[cs-uri-stem-aplic]	BCS_MAX
	ip_country_code	BCS_MIN
	ip_country_name	BCS_PRO
	ip_region_name	BCS_TOT
	ip_city_name	BSC_MAX
	[c-ipNUmerico]	BSC_MIN
	[c-ipNUmerico2]	BSC_PRO
	INEI_LOCAL	BSC_TOT
	INEI_TIPO_APLICACION	
	USUARIO_TIPO	
	[sc-bytes]	
	[cs-bytes]	

Figura 18. Modelo de Base de datos

En la figura 18 se muestran las tablas de procesamiento y repositorio de la información a analizar. La tabla ip2Location_db3, es la tabla donde reside la relación entre rangos de IPs y la ubicación geográfica correspondiente. En la tabla

Servidor_Aplicaciones se encuentra la información relacionada a las aplicaciones del INEI. La tabla Proc_Carga_Log es donde se almacena la información que proviene de los archivos de registro de los servidores web. Y finalmente la tabla BI_LOGGAPLICACIONES, es el deposito final donde reside la información agregada y dimensionada que se obtiene de la tabla Proc_Carga_Log.

2.3.5.4 Diccionario de datos

Tabla 36

Tabla BI_LogAplicaciones

Ítem	TABLA		BI_LOGAPLICACIONES	
	Campo	Tipo de dato	Tipo Llave	Descripción
1	ANIO	int	PK	Año
2	MES	int	PK	Número de Mes
3	MES_NOMBRE	Varchar (40)	PK	Nombre de Mes
4	APLIC_URL	Varchar (100)	PK	Ruta de aplicación web
5	NAV_STATUS	int	PK	Código de estado de navegación
6	INEI_LOCAL	Varchar (80)	PK	Sede de Local del INEI
7	USUARIO_TIPO	Varchar (80)	PK	Tipo de Usuario
8	NAV_IP_PAIS_COD	Varchar (10)	PK	Código de País
9	NAV_IP_PAIS_NOMBRE	Varchar (90)	PK	Nombre de País
10	NAV_IP_REGION_NOMBRE	Varchar (90)	PK	Nombre de Región o Estado
11	NAV_IP_CIUADAD_NOMBRE	Varchar (90)	PK	Nombre de Ciudad
12	SERVIDOR	Varchar (80)	PK	Nombre de Servidor
13	NRO_VISITANTES	bigint		Número de visitantes
14	NRO_VISITAS	bigint		Número de visitas
15	TIEMPO_CARGA_MAX	bigint		Tiempo de Carga Máxima
16	TIEMPO_CARGA_MIN	bigint		Tiempo de Carga Mínima
17	TIEMPO_CARGA_PRO	bigint		Tiempo de Carga Promedio

18	BCS_MAX	bigint	Bytes transferidos de cliente a servidor - Máximo
19	BCS_MIN	bigint	Bytes transferidos de cliente a servidor - Mínimo
20	BCS_PRO	bigint	Bytes transferidos de cliente a servidor - Promedio
21	BCS_TOT	bigint	Bytes transferidos de cliente a servidor - Total
22	BSC_MAX	bigint	Bytes transferidos de servidor a cliente - Máximo
23	BSC_MIN	bigint	Bytes transferidos de servidor a cliente - Mínimo
24	BSC_PRO	bigint	Bytes transferidos de servidor a cliente - Promedio
25	BSC_TOT	bigint	Bytes transferidos de servidor a cliente - Total

Tabla 37

Tabla Servidor_Aplicaciones

TABLA		Servidor_Aplicaciones		
Ítem	Campo	Tipo de dato	Tipo Llave	Descripción
1	Corr	int	PK	Código correlativo asignado a las aplicaciones
2	Anio	int		Año
3	URLaplic	Varchar (600)		URL de la aplicación, incluyendo servidor
4	Desaplic	Varchar (MAX)		Descripción de aplicación
5	URLAplic2	Varchar (600)		URL de la aplicación, solo directorio
6	Tipo	Varchar (50)		Tipo de Aplicación – Según estado en producción
7	Tipo2	Varchar (50)		Tipo de Aplicación – Según Uso de usuario
8	EnLinea	Varchar (50)		Estado si se encuentra en publicación
9	Servidor	Varchar (50)		Servidor

Tabla 38

Tabla ip2location_db3

TABLA		ip2location_db3		
Ítem	Campo	Tipo de dato	Tipo Llave	Descripción

1	ip_from	float	PK	Rango de IPS desde
2	ip_to	float	PK	Rango de IPS hasta
3	country_code	Varchar (2)	PK	Código de País
4	country_name	Varchar (60)	PK	Nombre de País
5	region_name	Varchar (128)	PK	Nombre de estado región
6	city_name	Varchar (128)	PK	Nombre de Ciudad

2.3.6 Desarrollo Sprint III

El Sprint III, comprenderá la definición de los indicadores a generar o visualizar a raíz del análisis de la información.

2.3.6.1 Entregables

Tabla 39

Entregables Sprint III

ENTREGABLES
Identificación de Indicadores Claves de Gestión.

2.3.6.2 Indicadores

Tabla 40

Indicadores

Indicador	Cobertura
Número de visitas	Total y agrupado por diferentes dimensiones.
Número de visitantes	Total y agrupado por diferentes dimensiones.
Tiempo de carga a las aplicaciones	Total y agrupado por diferentes dimensiones.
Tamaño de información consumida desde las aplicaciones	Total y agrupado por diferentes dimensiones.

Tamaño de información Total y agrupado por diferentes dimensiones.
cargada hacia las aplicaciones

Casos de Error en las Total y agrupado por diferentes dimensiones.
aplicaciones

2.3.7 Desarrollo Sprint IV

El Sprint IV, comprenderá el diseño de los prototipos de los tableros de análisis de información.

2.3.7.1 Entregables

Tabla 41

Entregables Sprint IV

ENTREGABLES
Creación de Prototipos de tableros.

2.3.7.2 Prototipos de tableros

A continuación, ilustraremos los prototipos de los tableros diseñados para los casos definidos en las Historias de usuario.

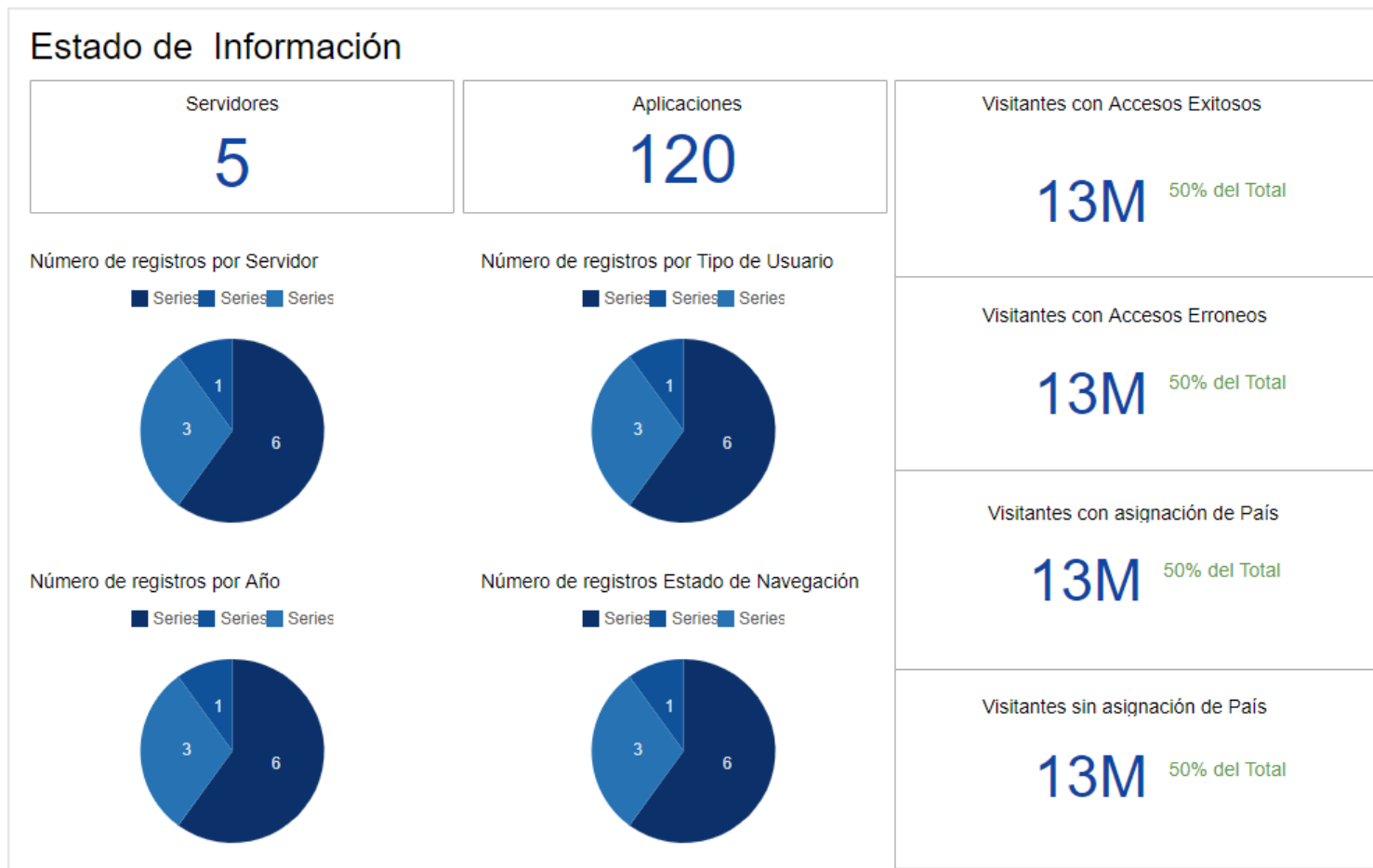


Figura 19. Prototipo de Tablero de datos generales

En la figura 19 se establece el tablero inicial, donde se visualiza indicadores generales de información cargada. Por ejemplo: Numero de servidores, aplicaciones, número de registros por año, por estado de navegación, porcentaje con asignación exitosa de País, etc.

Estado de Información

Servidor

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

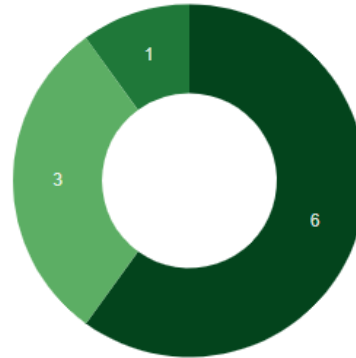
Tipo de Aplicación

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

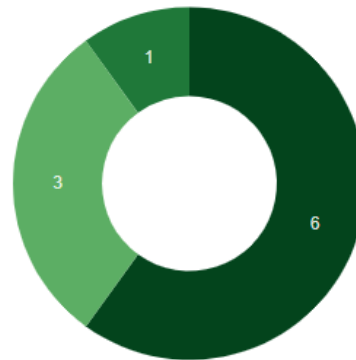
En Línea

Si
No ▶

Aplicaciones



Aplicaciones por Servidor



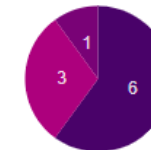
Aplicaciones en Línea - Estado

Series 1 Series 2 Series 3



Aplicaciones en Línea - Tipo

Series 1 Series 2 Series 3



Aplicaciones en Línea que no están en Producción por Servidor

Series 1 Series 2 Series 3



Figura 20. Prototipo de Tablero de Análisis de información de aplicaciones

En la figura 20 se establece visualizar el número de aplicaciones por su tipo definido y por servidor. Y a la vez también filtrado por su estado (Producción / Culminado).

Analisis por Número de Visitas y Visitantes

Servidor

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Año

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Estado de Navegación

Con Error
Sin Error ▶

Tipo de Aplicación

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Número de Visitantes por País



Número de Visitas por País



Número de Visitantes por tipo de Usuario



Número de Visitas por tipo de Usuario



Número de Visitantes por Sedes INEI



Número de Visitas por Sedes INEI



Figura 21. Prototipo de Tablero de Visualización de características de información de Visitantes y Visitas

En la figura 21 se establece visualizar el número de visitantes y visitas por País, por Sedes del INEI y por tipo de usuario. Con la adición de los filtros de: Servidor, año, estado de navegación y tipo de aplicación.

Totales por Estado de Resultado de Navegación

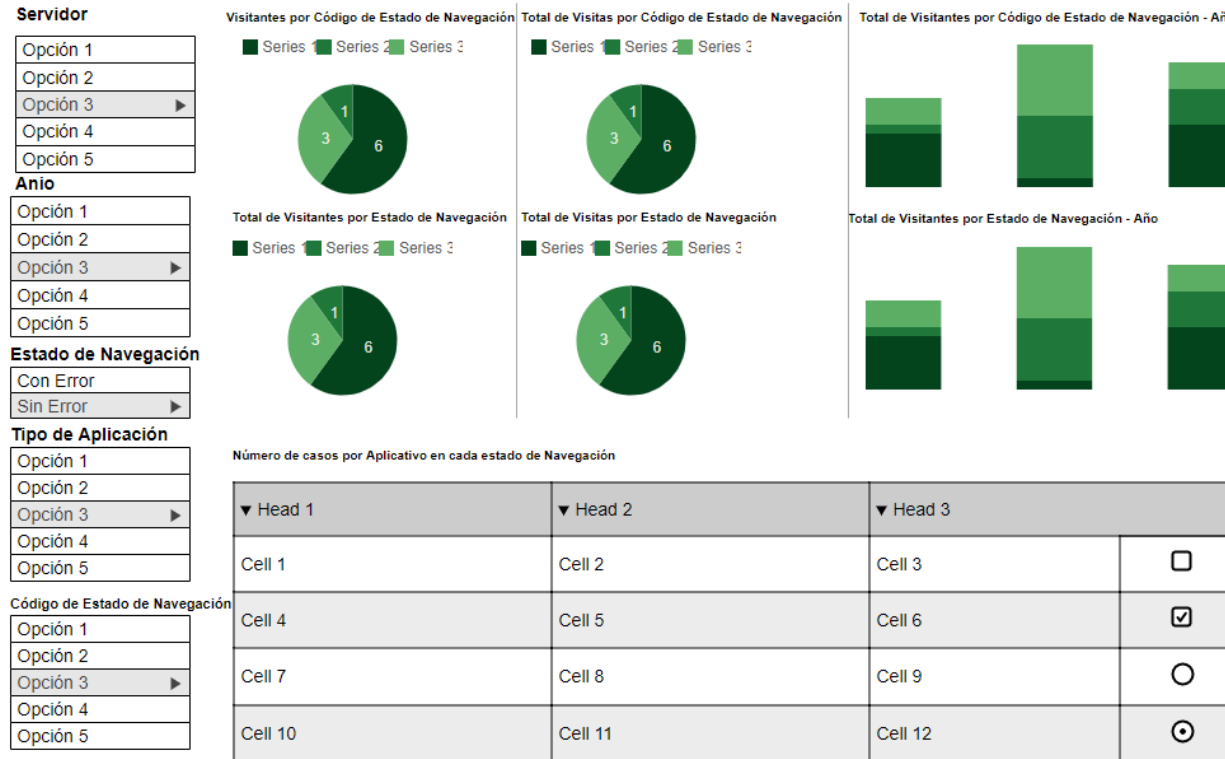


Figura 22. Prototipo de Tablero de estado de Navegación

En la Figura 22 se establece visualizar el número de visitantes y visitas por estado de navegación y por código de estado de navegación. Así también por año y el detalle de casos por estado de navegación. Con la adición de los filtros de: Servidor, año, estado de navegación y tipo de aplicación y código de estado de navegación.

Número de Visitantes por País

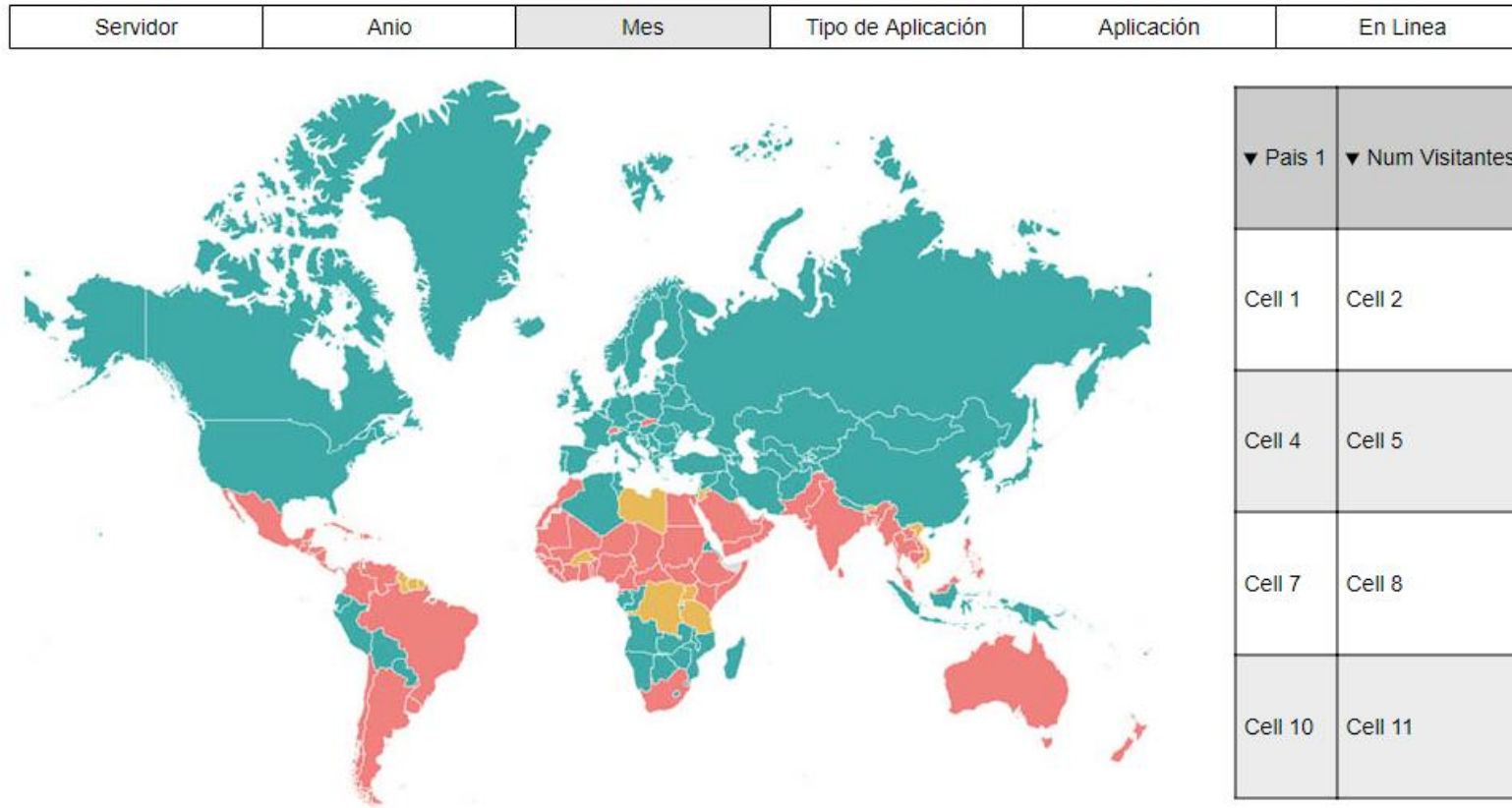


Figura 23. Prototipo de Tablero de Análisis de información geográfica por País según número de visitantes

En la figura 23 se establece visualizar el número de visitantes por País, con los filtros de servidor, año, mes, tipo de aplicación, aplicación y estado de aplicación web.

Número de Visitas por País

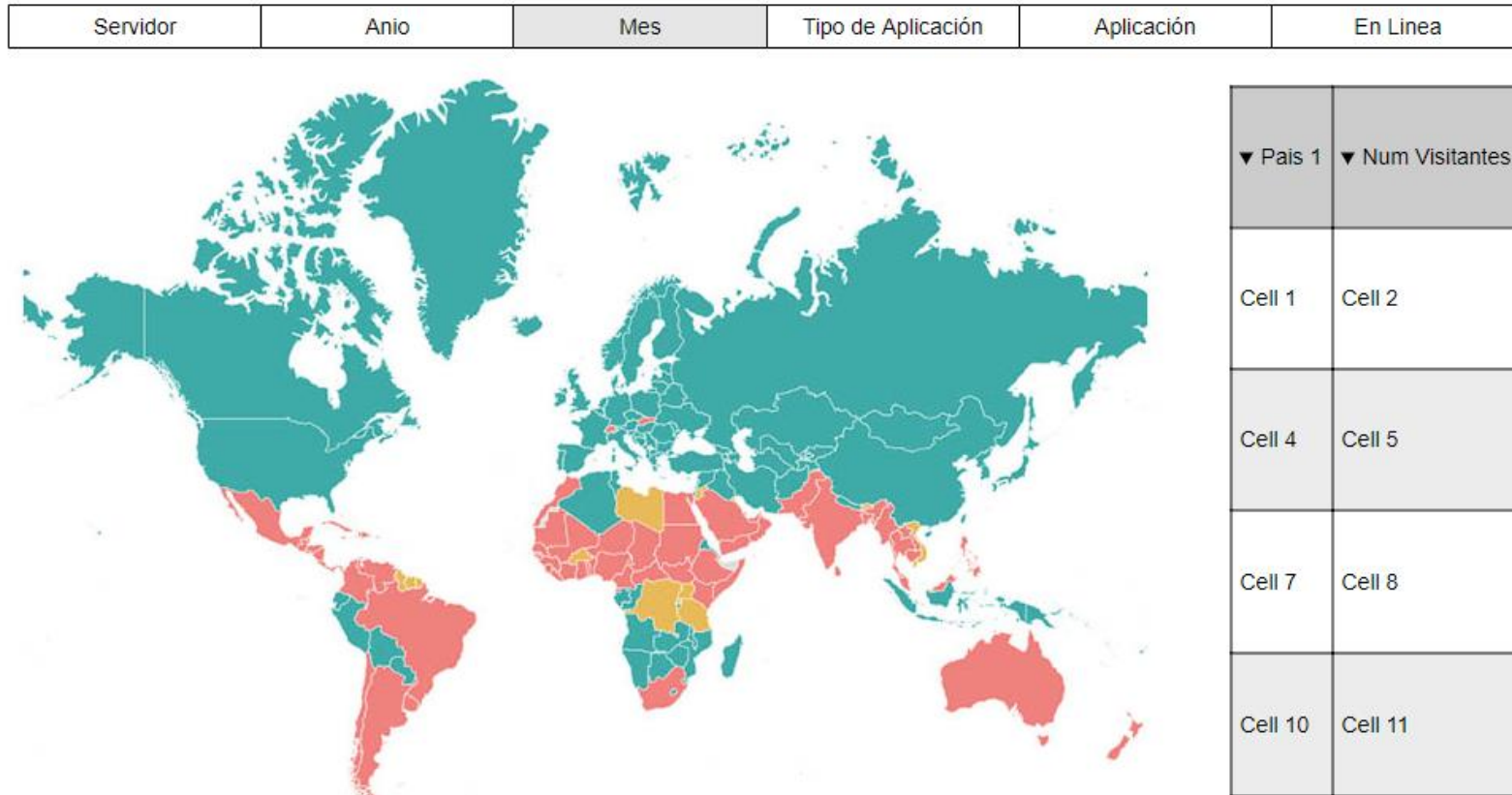


Figura 24. Prototipo de Tablero de Análisis de información geográfica por País según número de visitas

En la figura 24 se establece visualizar el número de visitas por País, con los filtros de servidor, año, mes, tipo de aplicación, aplicación y estado de aplicación web.

Número de Visitantes por Ciudad



Figura 25. Prototipo de Tablero de Análisis de información geográfica por Ciudad según número de visitantes

En la figura 25 se establece visualizar el número de visitantes por Ciudad, con los filtros de servidor, año, mes, tipo de aplicación, aplicación y estado de aplicación web.

Número de Visitas por Ciudad

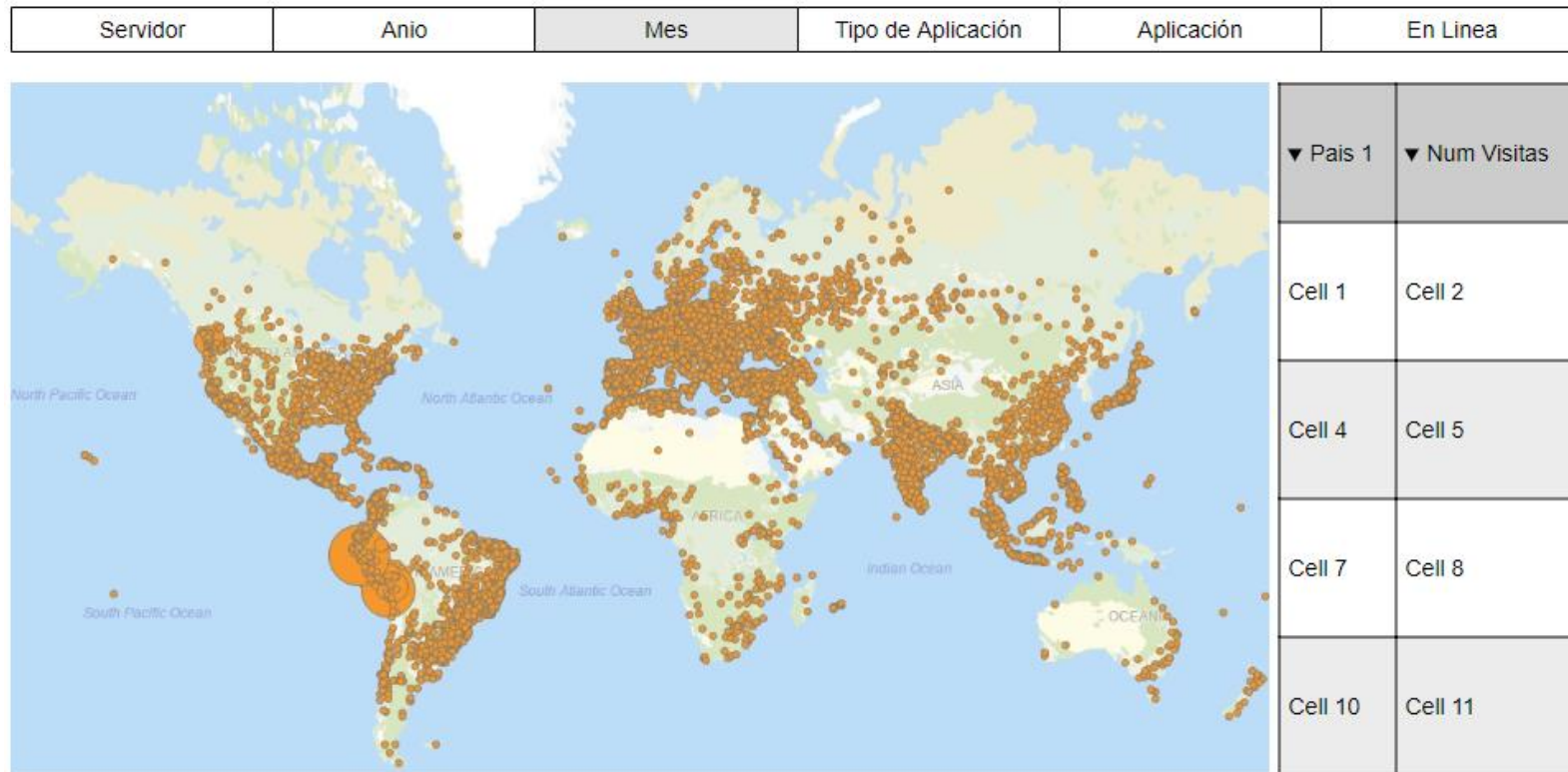


Figura 26. Prototipo de Tablero de Análisis de información geográfica por Ciudad según número de visitas

En la figura 26 se establece visualizar el número de visitas por Ciudad, con los filtros de servidor, año, mes, tipo de aplicación, aplicación y estado de aplicación web.

Listado de Aplicaciones

Servidor	Anio	Mes	Tipo de Aplicación	Aplicación	Estado de Navegación	En Linea
----------	------	-----	--------------------	------------	----------------------	----------

Aplicaciones

▼ Servidor	▼ Tipo de Aplicación	▼ URL	▼ Descripción	▼ Num Visitantes	▼ Num Visitas	▼ Tiempo de Carga Max	▼ Tiempo de Carga Min	▼ Tiempo de Carga Prom

Aplicaciones por Pais y Mes

▼ Servidor	▼ Tipo de Aplicación	▼ Pais	▼ Mes	▼ URL	▼ Descripción	▼ Num Visitantes	▼ Num Visitas	▼ Tiempo de Carga Max	▼ Tiempo de Carga Min	▼ Tiempo de Carga Prom

Figura 27. Prototipo de Tablero de Listado de Aplicaciones

En la figura 27 se establece la visualización del listado de aplicaciones y el detalle de sus indicadores por país y mes.

Análisis de Indicadores por Mes y Año

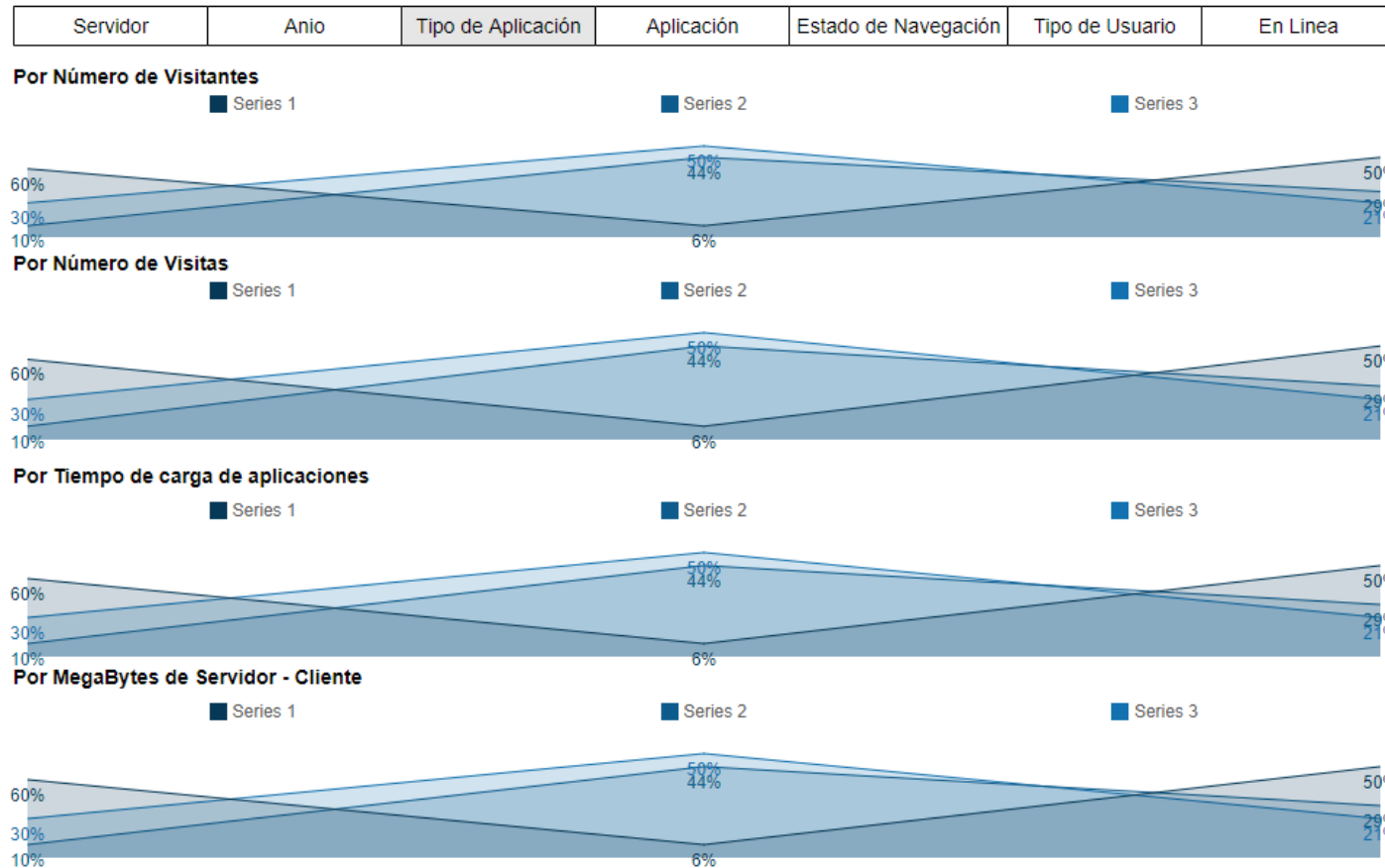


Figura 28. Prototipo de Tablero de Indicadores por Mes y Año

En la figura 28 se establece visualizar por serie de tiempo, los totales de indicadores como el número de visitantes, número de visitas, tiempo de carga de la aplicación, los megabytes de Servidor a Cliente y viceversa.

Número de Visitantes por departamento - PERU

Servidor	Anio	Mes	Pais	Tipo de Aplicación	Aplicación			
▼ Departamento	▼ Num Visitantes	▼ 2019 Num Visitantes	▼ 2018 Num Visitantes	▼ Num Visitas	▼ Accesos sin Error	▼ Accesos con Error	▼ MegaBytes de Cliente - Servidor	▼ MegaBytes de Servidor - Cliente

Figura 29. Figura 29. Prototipo de Tablero de indicadores por Departamento

En la figura 29 se establece el tablero que visualizara una serie de indicadores por departamento.

Predicción Nro Visitantes

Servidor

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Año

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Fuente

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

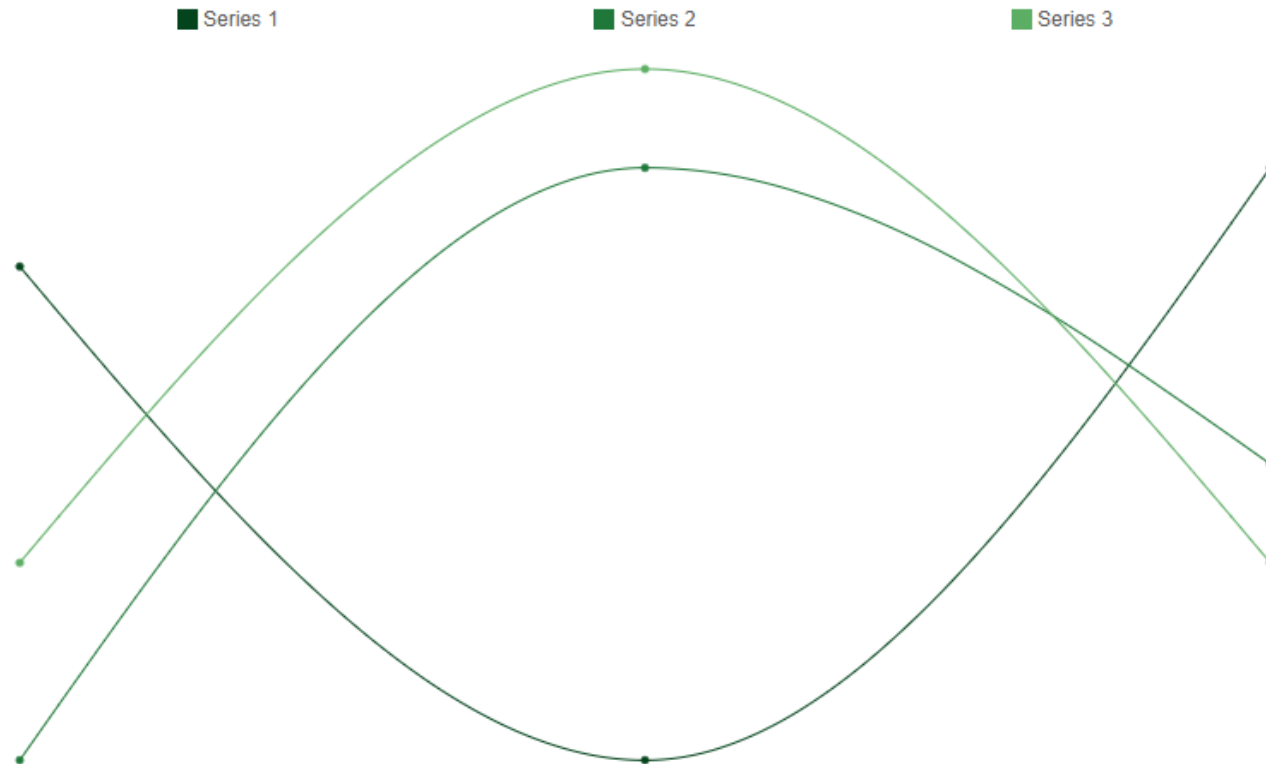


Figura 30. Prototipo de Tablero de Predicción de Número de Visitantes

En la figura 30 se establece la visualización de la predicción del número de visitantes por mes.

Predicción Nro Visitas

Servidor

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Año

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

Fuente

Opción 1
Opción 2
Opción 3 ▶
Opción 4
Opción 5

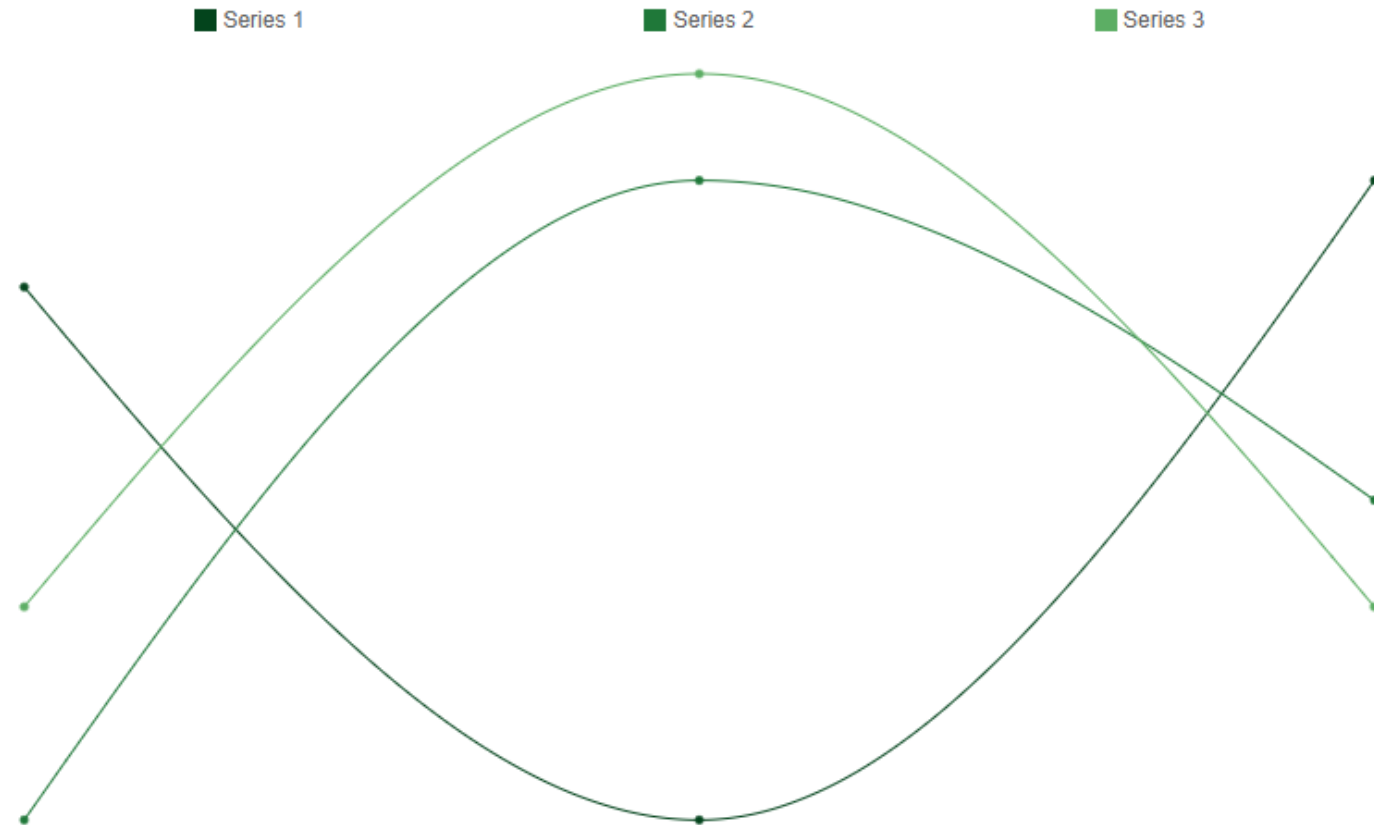


Figura 31. Prototipo de Tablero de Predicción de Número de Visitas

En la figura 31 se establece la visualización de la predicción del número de visitas por mes.

2.3.8 Desarrollo Sprint 5

2.3.8.1 Entregables

Tabla 42

Entregables Sprint 5

ENTREGABLES
Desarrollo de tableros de análisis

2.3.8.2 Implementación de Interfaz de usuario

Los tableros de análisis de información se desarrollaron con la herramienta BI, Qlik Sense Desktop. A continuación, se muestran las pantallas de cada tablero que se definió en los capítulos anteriores.

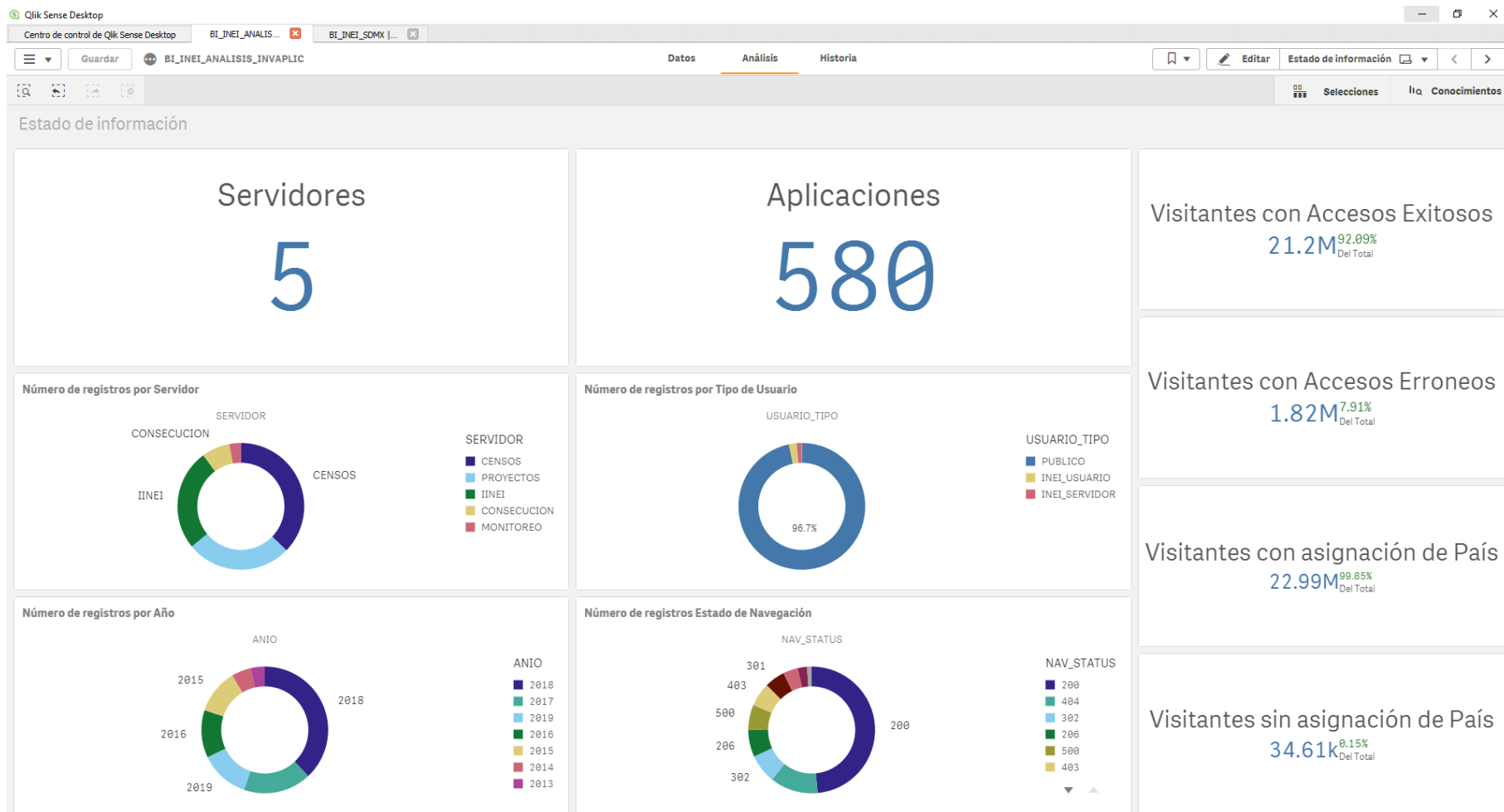


Figura 32. Tablero de estado de la información

En la figura 32 se grafica el tablero inicial, donde se visualiza el total de registros cargados por servidor, año, tipo de usuario, estado de navegación, registros que si/no se le asignó la ubicación geográfica.

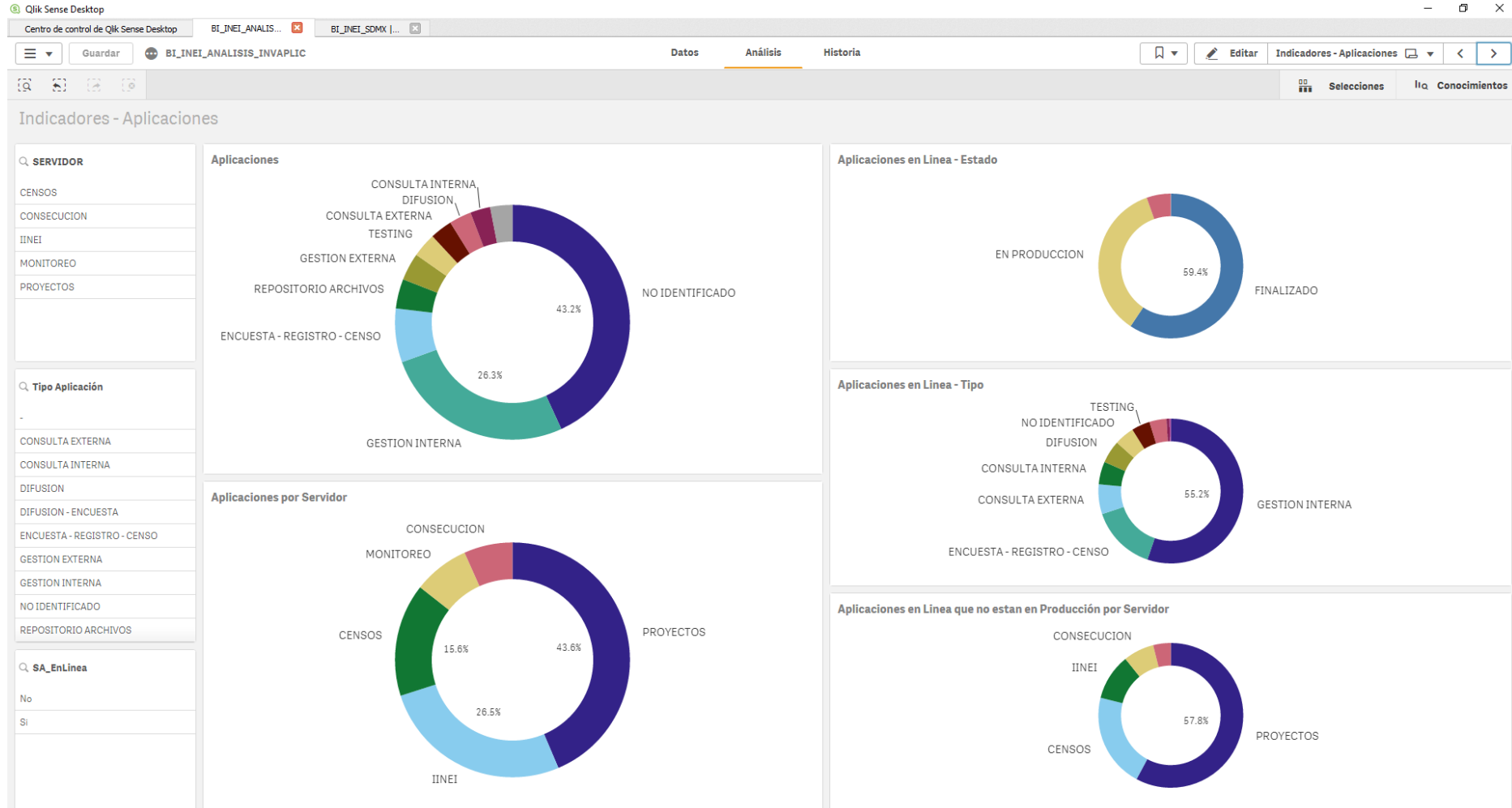


Figura 33. Tablero de Análisis por Aplicación

En la figura 33, el tablero permite el análisis de las aplicaciones web por año, servidor, tipo de aplicación, estado de línea.

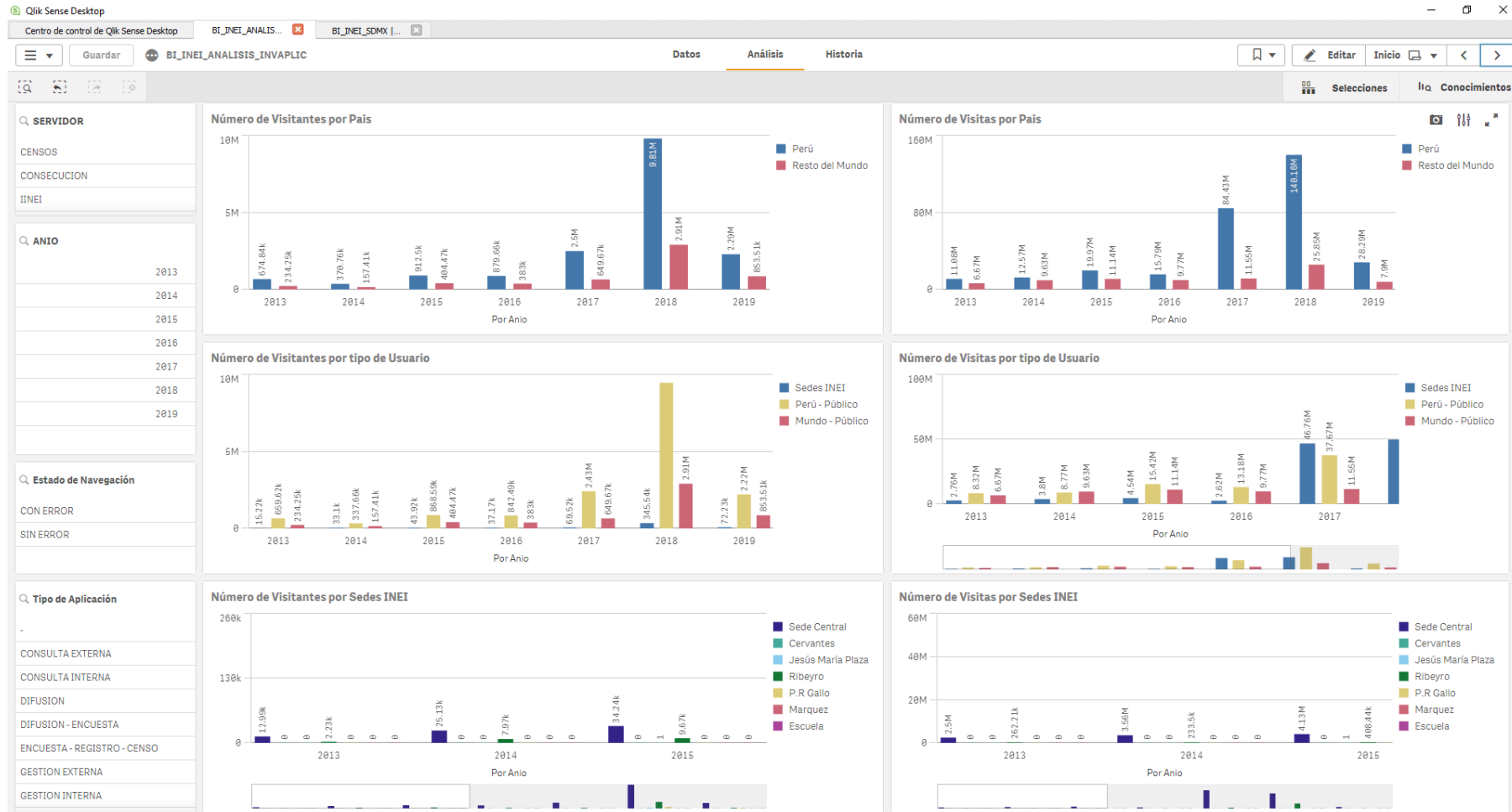


Figura 34. Tablero de Análisis por número de visitantes y visitas

En la figura 34, se grafica el tablero que permite el análisis número de visitantes y visitas según las dimensiones de servidor, año, estado de navegación, tipo de aplicación, visitantes por Sede de local el INEI, Origen General (Nacional o extranjero).

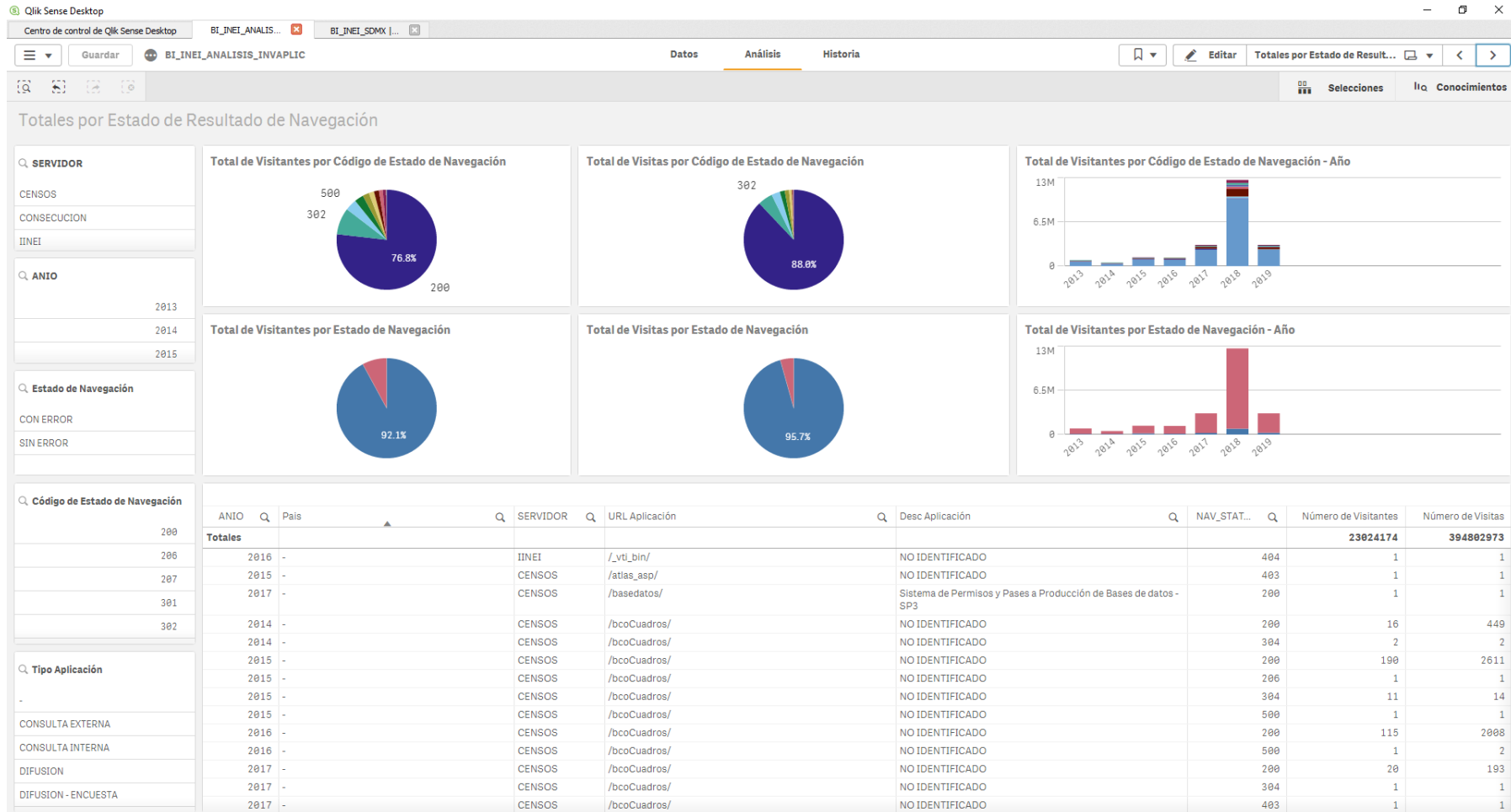


Figura 35. Tablero de Análisis por estado de navegación

En la figura 35, se grafica el tablero que permite el análisis de la información según la dimensión Estado de navegación, adicionando al mismo, las dimensiones: servidor, año, estado de navegación general (Con Error, Sin Error), Código de estado de navegación, tipo de aplicación.

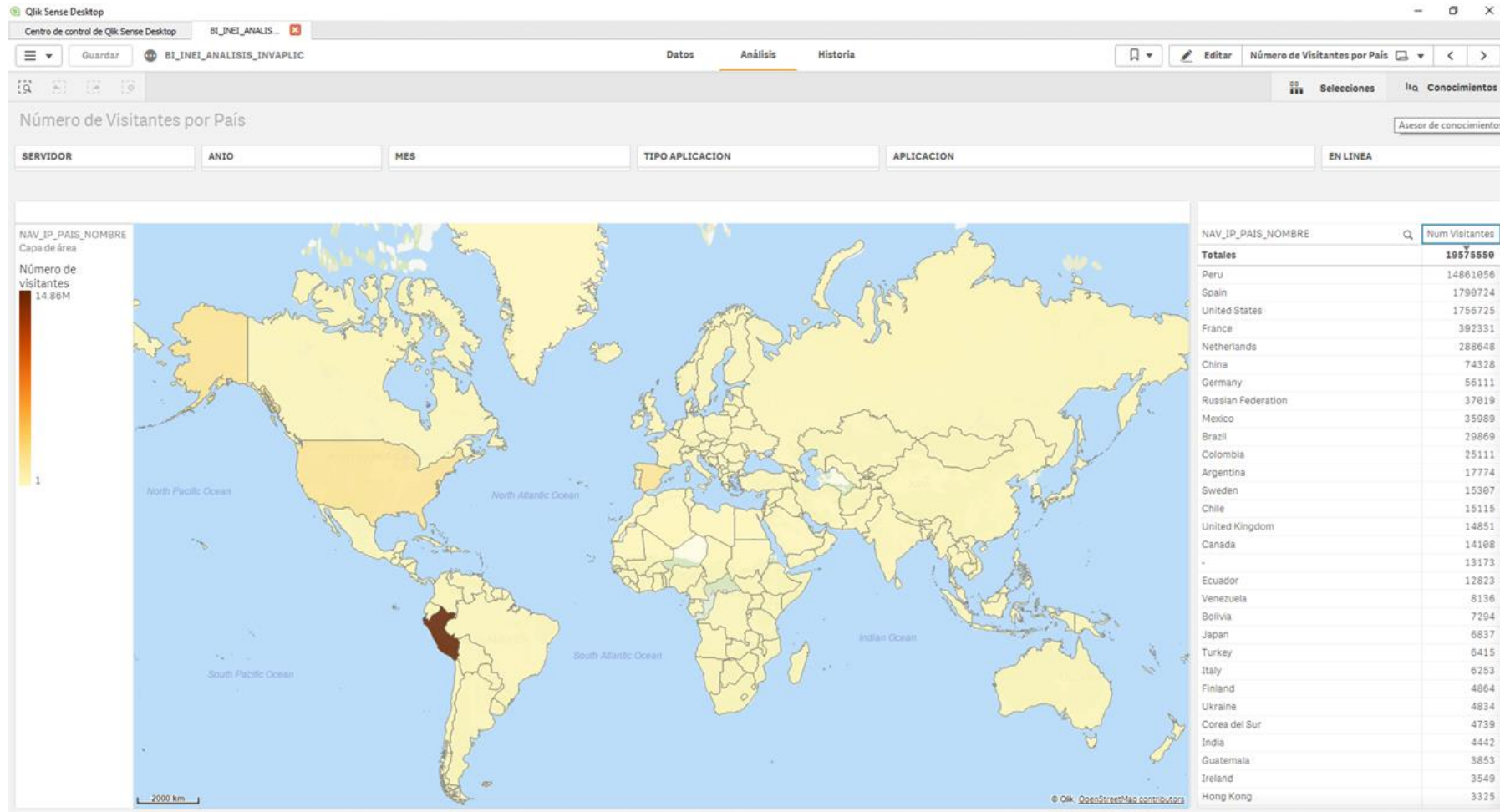


Figura 36. Tablero de Análisis geográfico por País según el número de visitantes

En la figura 36, se grafica el tablero del análisis geográfico por País según el número de visitantes. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

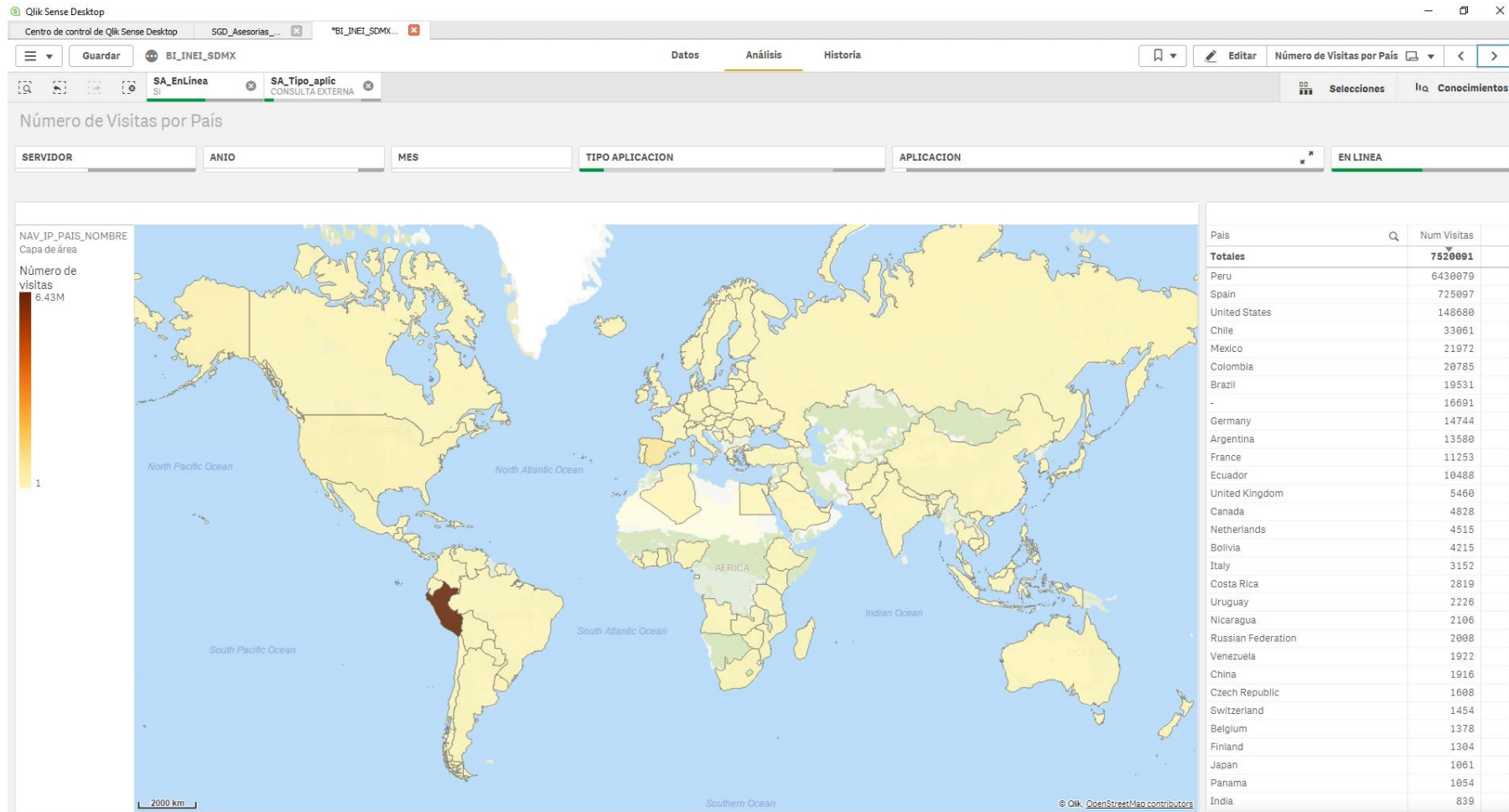


Figura 37. Tablero de Análisis geográfico por País según el número de visitas

En la figura 37, se grafica el tablero del análisis geográfico por País según el número de visitas. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

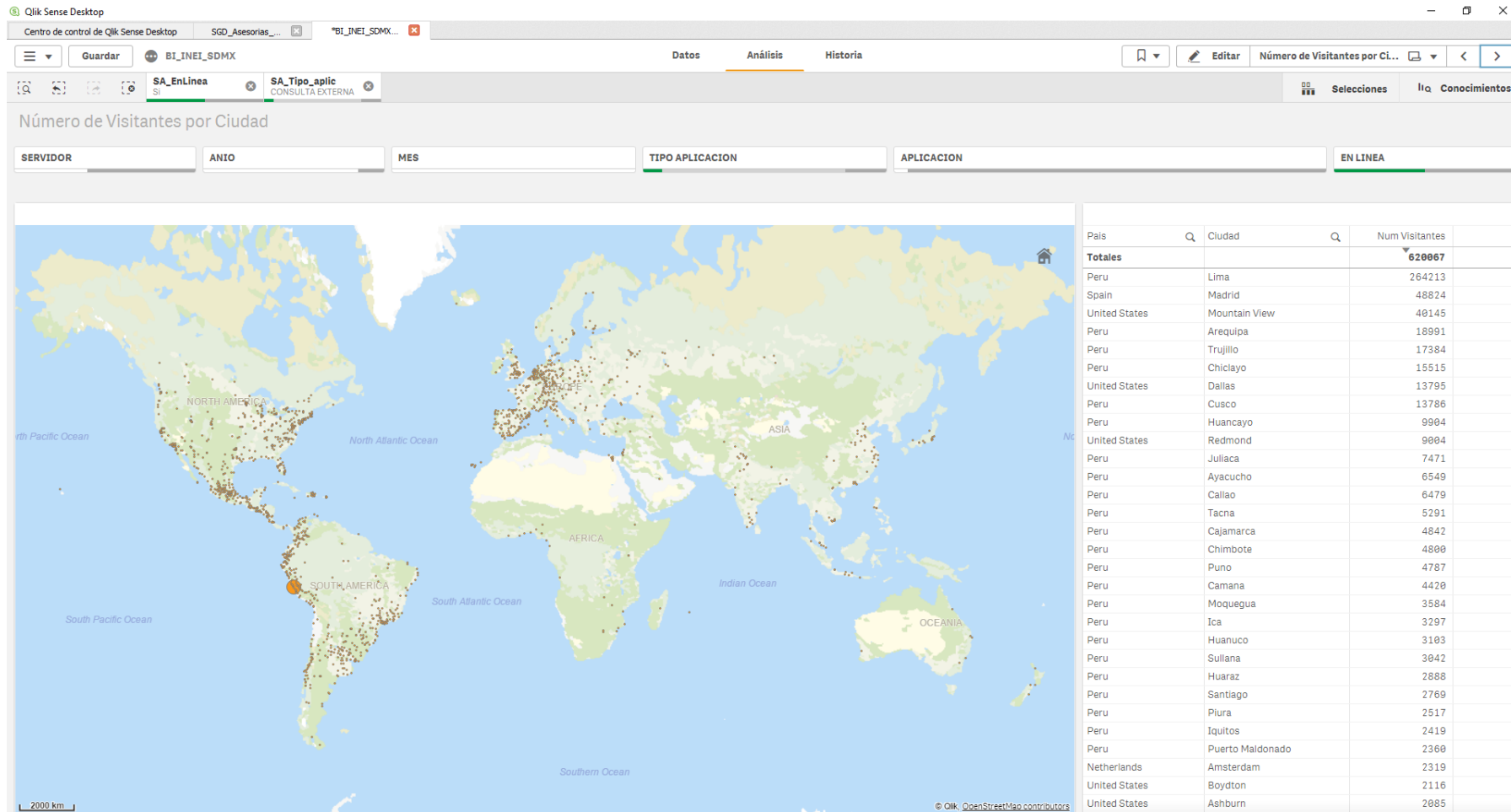


Figura 38. Tablero de Análisis geográfico por Ciudad según el número de visitantes

En la figura 38, se grafica el tablero del análisis geográfico por Ciudad según el número de visitantes. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

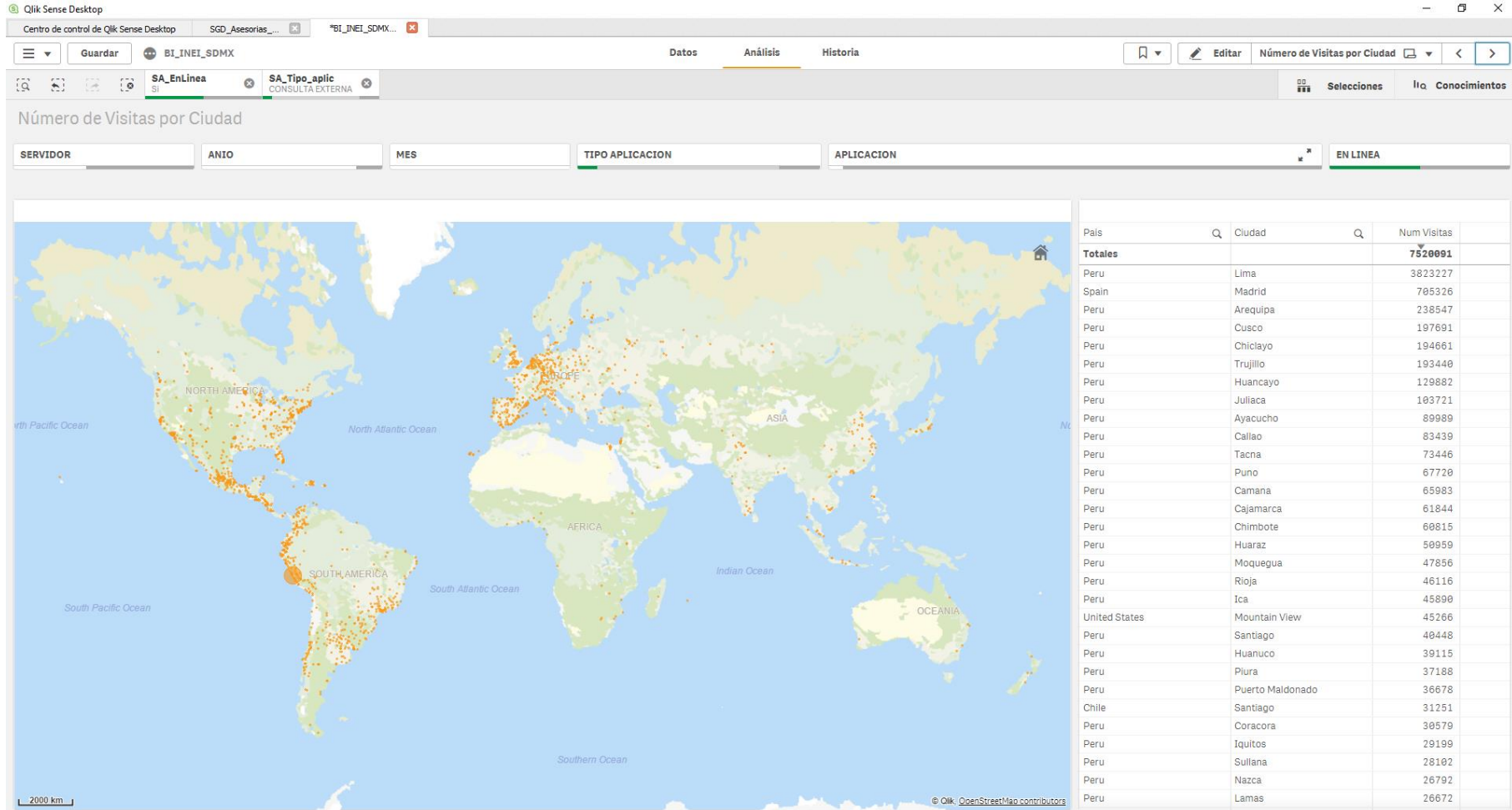


Figura 39. Tablero de Análisis geográfico por Ciudad según el número de visitas

En la figura 39, se grafica el tablero del análisis geográfico por Ciudad según el número de visitas. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de línea.

Qlik Sense Desktop

Centro de control de Qlik Sense Desktop BI_INEI_ANALIS... BI_INEI_SDMX |...

Guardar BI_INEI_ANALISIS_INVAPLIC Datos Análisis Historia

Selecciones Ho Conocimientos

Listado Aplicaciones

SERVIDOR ANIO MES_NOMBRE Tipo Aplic Aplicación Estado Navegación En Línea

SERVIDOR	Tipo de Aplicación	URL	Descripción Aplicación	Num Visitantes	Num Visitas	Tiempo de Carga Máxima (min.)	Tiempo de Carga Mínima (min.)	Tiempo de Carga Promedio (min.)
Totales				23024174	394802973	1514	0.00	0.0339
PROYECTOS	GESTION EXTERNA	/convocatorias/2014/convoca/	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	7500636	47601440	1434	0.00	0.1904
IINEI	-	URL NO VALIDADA	URL NO VALIDADA	2837274	110076148	0	0.00	0.0000
PROYECTOS	REPOSITORIO ARCHIVOS	/web/biblioinetpub/bancopub/	NO IDENTIFICADO	1304438	3640322	837	0.00	0.0844
PROYECTOS	GESTION EXTERNA	/constancias/docentes/	Constancias de Docentes 2013	796724	2876933	5	0.00	0.0008
CENSOS	REPOSITORIO ARCHIVOS	/cpv2007/	NO IDENTIFICADO	745479	9816599	0	0.00	0.0000
CENSOS	DIFUSION	/Censos2007/	Portal Web Censos Nacionales 2007 XI de Población y VI de Vivienda	694882	6313381	0	0.00	0.0000
CONSECUCION	GESTION EXTERNA	/cpv2017/convoca/	Sistema de Consecución de Recursos Humanos - Inscripción de Postulantes	602838	21550036	191	0.00	0.0992
CONSECUCION	GESTION EXTERNA	/convocatorias/2014/convoca/	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	411285	8246457	202	0.00	0.1136
MONITOREO	NO IDENTIFICADO	/cpv2017/convoca/	NO IDENTIFICADO	408098	20299016	191	0.00	0.2189

ANIO	MES	Pais	Tipo de Aplicación	SERVID...	Descripción de Aplicación	URL	Num Visitantes	Num Visitas	Tiempo de Carga Máxima (min.)	Tiempo de Carga Mínima (min.)	Tiempo de Carga Promedio (min.)
Totales							23024174	394802973	1514	0	0.0339
2018	9	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	1409476	9865711	233	0	0.1209
2018	10	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	729413	2987614	30	0	0.4466
2018	6	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	646532	4381626	70	0	0.2422
2018	8	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	625356	4057354	223	0	0.3728
2018	5	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	428618	2650732	288	0	0.3202
2019	5	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	399089	2356773	497	0	0.2361
2018	12	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	362150	2023124	184	0	0.1479
2018	2	Peru	GESTION EXTERNA	PROYECTOS	Sistema de consecución de Recursos Humanos Inscripción de postulantes.	/convocatorias/2014/convoca/	347491	2243975	1434	0	0.6312

Figura 40. Tablero de Análisis por aplicación

En la figura 40, se grafica el tablero del análisis de cada aplicación (número de visitantes, número de visitas, tiempo de carga) según las dimensiones servidor, año, mes, tipo de aplicación, aplicación web, estado de navegación, estado de línea.

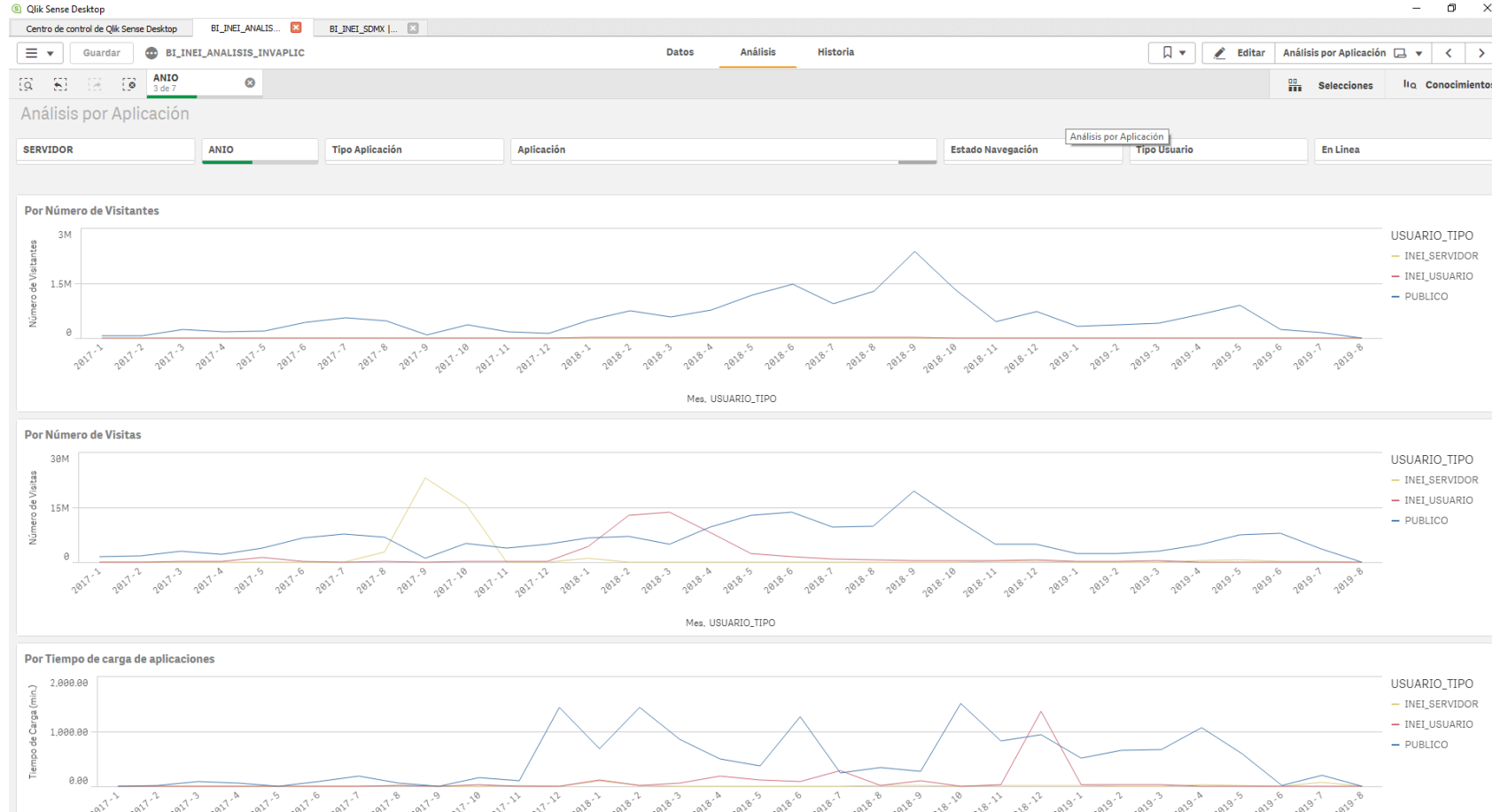


Figura 41. Tablero de Análisis de aplicaciones por mes

En la figura 41, se grafica el tablero del análisis del comportamiento de los indicadores de las aplicaciones por mes de cada año. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación, aplicación web, estado de navegación, tipo de usuario y estado de línea.

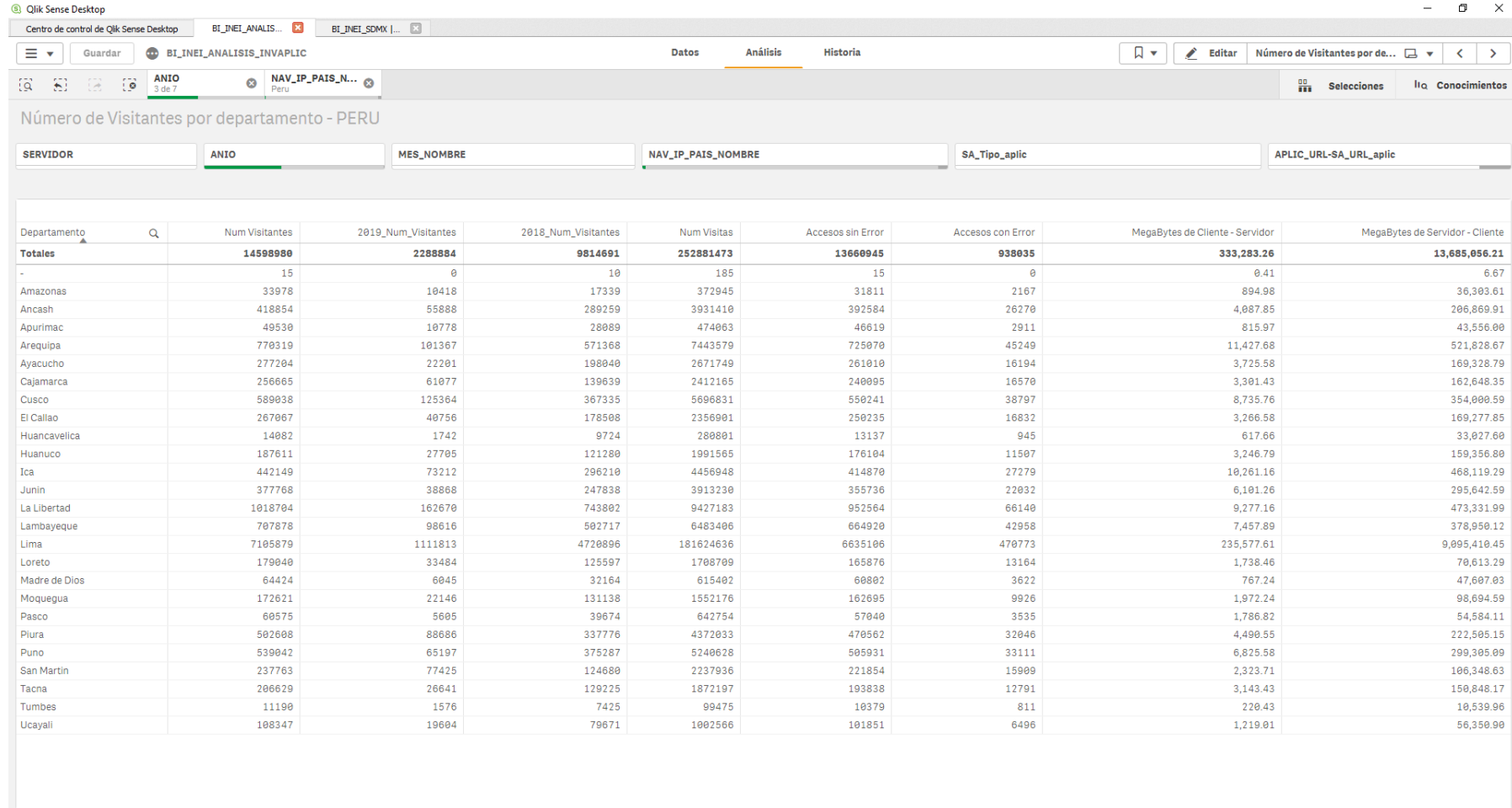


Figura 42. Tablero de Análisis de indicadores por departamento

En la figura 42, se grafica el tablero del análisis de indicadores por departamento del Perú. Adicionando al mismo, filtros como servidor, año, mes, tipo de aplicación y aplicación web.

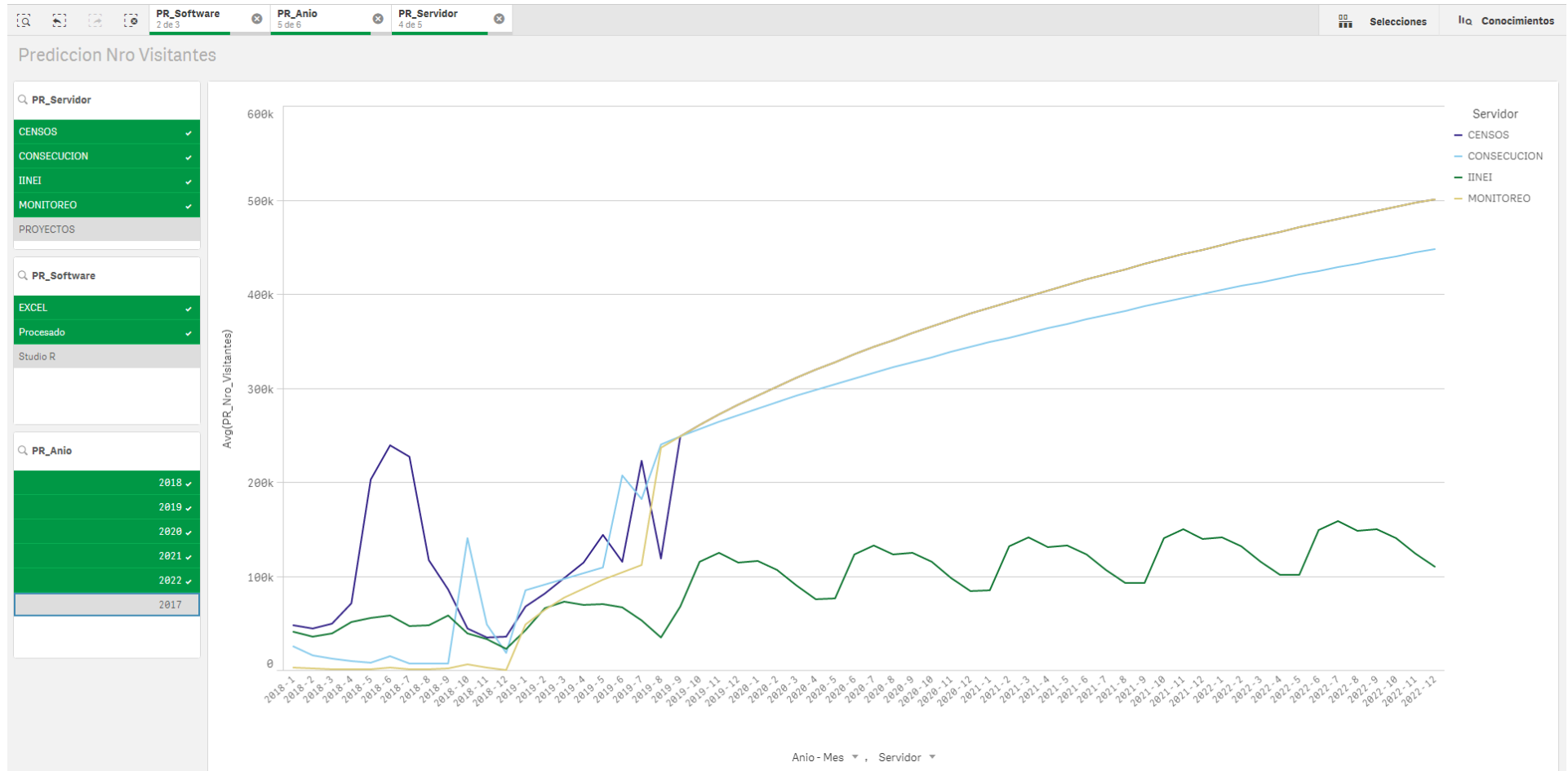


Figura 43. Tablero de Predicción de Número de Visitantes

En la figura 43, se grafica el tablero de la predicción del indicador número de visitantes.

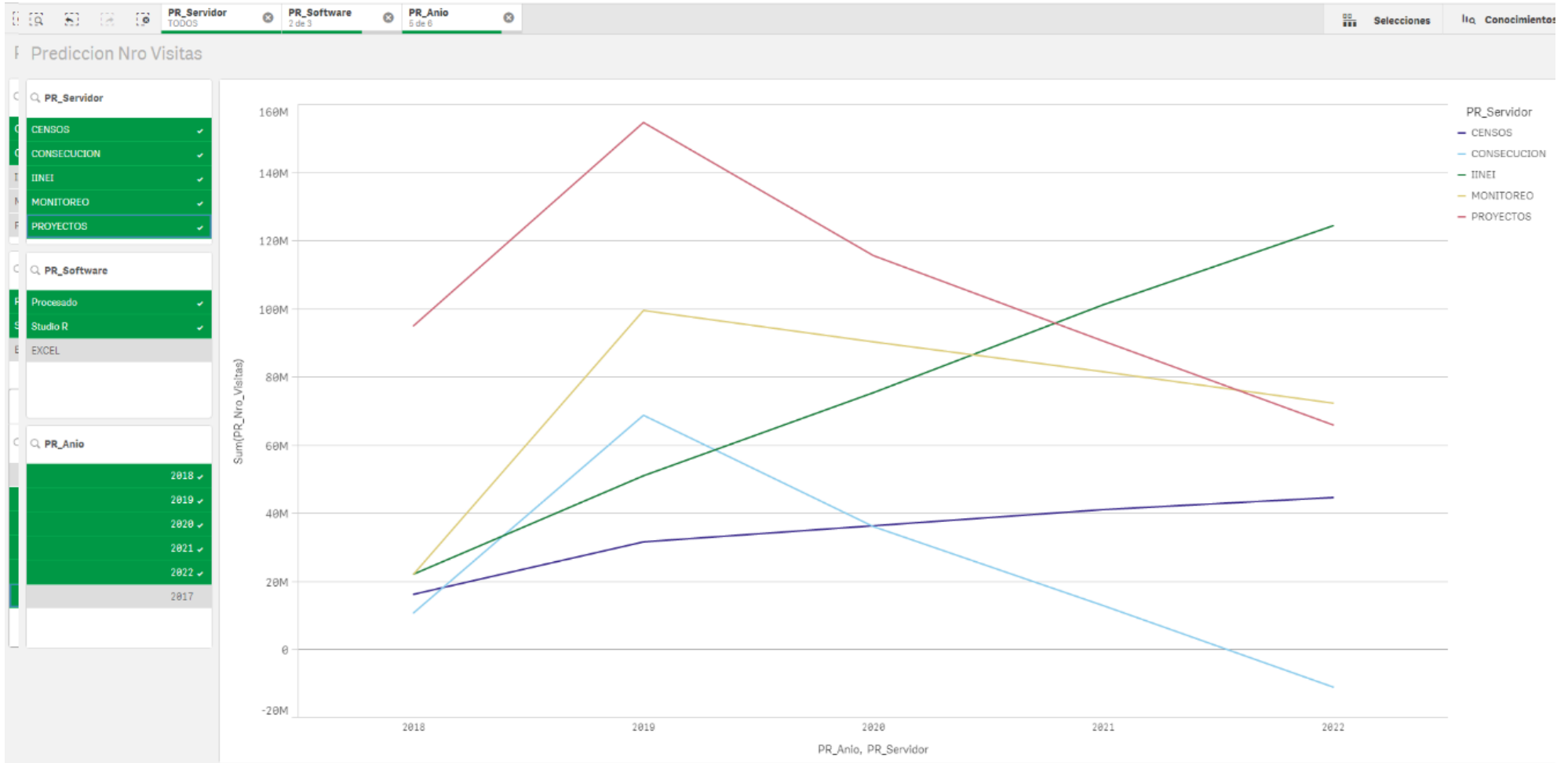


Figura 44. Tablero de Predicción de Número de Visitas

En la figura 44, se grafica el tablero de la predicción del indicador número de visitas.

CAPÍTULO III. RESULTADOS

El estudio del presente informe de tesis es pre-experimental, ya que se realiza un pre test y un post test y cuyos datos recopilados serán comparados a través de la prueba estadística T-Student, con el objetivo de comprobar si las hipótesis planteadas se cumplen. Se utiliza la prueba de Shapiro Will con la aplicación de la herramienta IBM SPSS y las condiciones siguientes:

- Sig. < 0.05 adopta una distribución no normal.
- Sig. > 0.05 adopta una distribución normal.

A continuación, detallaremos la contrastación de cada hipótesis.

3.1 Contrastación de Hipótesis

3.1.1 Hipótesis General

Para la evaluación de la hipótesis planteada, contrastamos los resultados obtenidos en el pre test y en el pos test.

Tabla 43

Análisis Pos test – Criterios de evaluación de las 3 dimensiones para la toma de decisiones

Dimensión	Medición	Escala
Análisis del comportamiento	El análisis del comportamiento de las aplicaciones según los indicadores, se aprecia en los tableros	Escala [1 - 5]
	construidos	
Análisis de resultados	Los indicadores trabajados, todos son muy importantes para el análisis del comportamiento de las aplicaciones de difusión de información a	Escala [1 - 5]
	través de las distintas dimensiones.	

Análisis de	Aún está en fase de aprobación por parte de las	Escala
Experiencia	direcciones responsables	[1 - 5]

En la tabla 43 se visualiza los criterios evaluados a las 3 dimensiones de la toma de decisiones referente a la analítica web, cuyos indicadores son visualizados y analizados a través de los tableros de mando. La valoración de las dimensiones se realizó a través de una ficha de evaluación, y que se aprecia en el capítulo de anexos. Para la valoración de cada una de las dimensiones, se utilizó la escala de Likert del 1 al 5, siendo el 1 la expresión de la mínima importancia y el 5 la máxima.

Tabla 44

Post Test - Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web

Evaluación del personal	Análisis del comportamiento	Análisis de resultados	Análisis de la Experiencia
1. No es importante	0	0	3
2. Poco Importante	0	0	0
3. Neutral	0	0	0
4. Importante	2	1	0
5. Muy Importante	1	2	0

Fuente: Información recolectada con la ficha adjunta en el anexo 8.

Tabla 45

Post Test – Análisis de la Frecuencia de la evaluación de las dimensiones de la toma de decisiones en analítica web

Evaluación del personal	Pre Test	Pre Test	Post Test	Post Test
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
1. No es importante	3	33.33	3	33.33
2. Poco Importante	1	11.11	0	0

3. Neutral	2	22.22	0	0
4. Importante	3	33.33	3	33.33
5. Muy Importante	0	0	3	33.33
Total	9	100	9	100

Fuente: Información recolectada con la ficha adjunta en el anexo 8.

Para el presente trabajo se considera una calificación de importante, para considerar a la Analítica Web como influyente en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI, a los valores 4 y 5 en la escala de Likert. En consecuencia, según la tabla 45 nos indica que la calificación de importancia aprobatoria de las dimensiones de toma de decisiones, con la nueva solución alcanza el 66.66%. Se considera entonces que la Hipótesis planteada: La aplicación de Analítica Web, influye significativamente en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web; es válida.

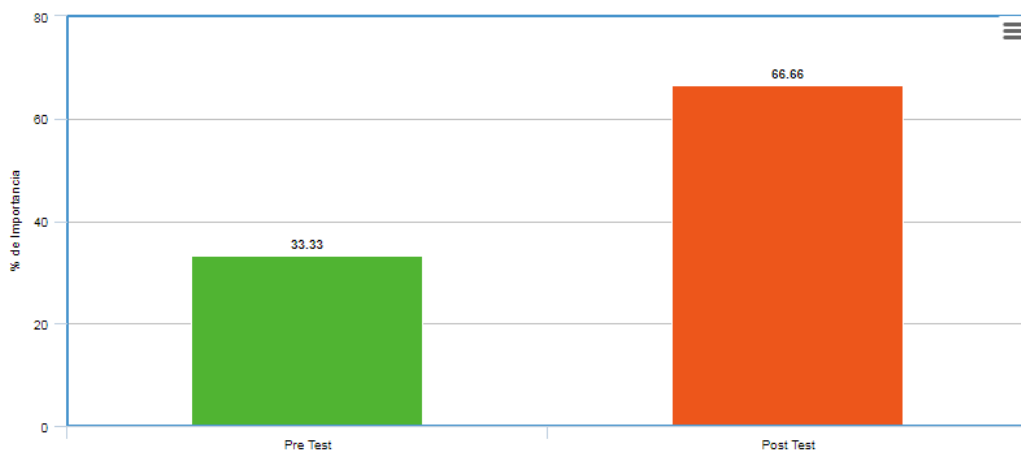


Figura 45. Porcentaje de valoración de las dimensiones de toma de decisiones

3.1.2 Hipótesis Específicas

3.1.2.1 Tiempo de proceso de generación de información para la toma de decisiones.

Para la valoración del Post Test, se tomó como referencia un día laborable por cada servidor procesado. Por única vez, ya que se procesó información de años anteriores (2 como mínimo). El procesamiento de archivos de un intervalo de tiempo menor puede agrupar varios servidores en un solo día. Pero se hicieron las pruebas con la primera métrica para corregir y optimizar, el script de carga y transformación.

Tabla 46

Datos del Post test de Tiempo de proceso de generación de información para la toma de decisiones

Número de Proceso	Medición	Valor	Valor
		Pre test	Post test
1	Días laborales	6	6
2	Días laborales	5	6
3	Días laborales	7	5
4	Días laborales	6	5
5	Días laborales	5	5
6	Días laborales	6	5
7	Días laborales	7	4
8	Días laborales	6	4
9	Días laborales	6	4

Fuente: Información recolectada con la ficha adjunta en el anexo 9.

Para el indicador en evaluación se realiza la prueba de normalidad Shapiro Wilk, en un universo de 9 procesos de prueba y el tipo de medición es cuantitativo.

Tabla 47

Pruebas de normalidad de Tiempo de proceso de generación de información para la toma de decisiones

Shapiro-Wilk

	Estadístico	Casos	Nivel de Significancia
Pre Test	0.83	9	0.055
Post Test	0.84	9	0.060

Fuente: Información recolectada con la ficha adjunta en el anexo 9.

Se observan los resultados de la prueba de normalidad en la Tabla 47, tanto para los datos recolectados de Pre Test y Post Test; el nivel de significancia de Pre Test es de 0,055 y el del Post Test es de 0.060, siendo ambos mayores al margen de error establecido 0,05; por lo tanto, los datos obtenidos poseen una distribución normal y son paramétricas.

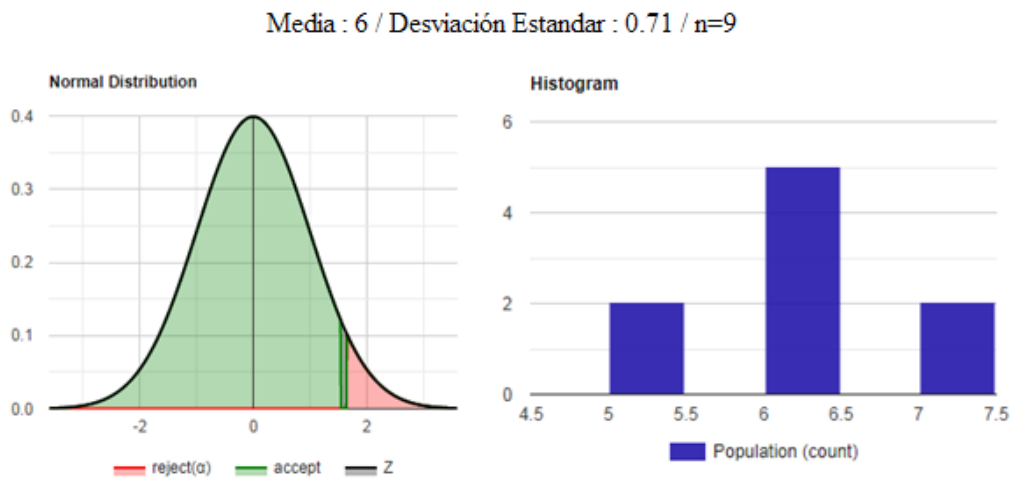


Figura 46. Tiempo de proceso de generación de información para la toma de decisiones

Media : 4.89 / Desviación Estandar : 0.78 / n=9

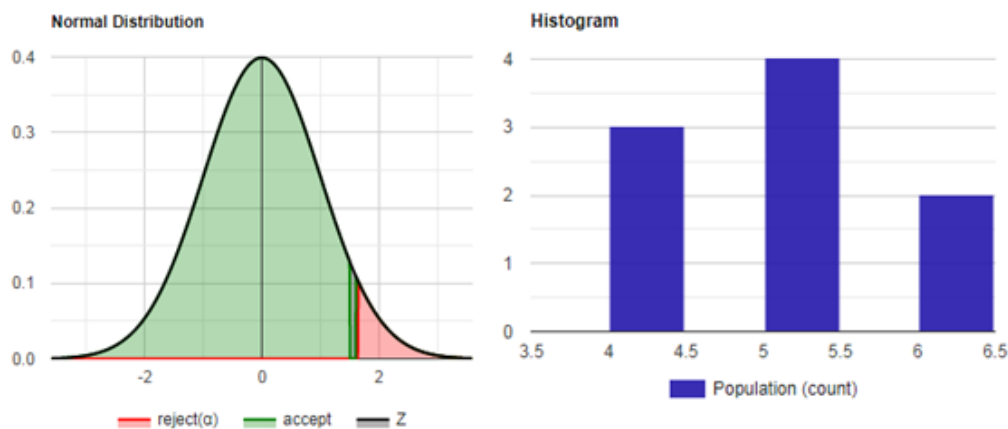


Figura 47. Tiempo de proceso de generación de información para la toma de decisiones

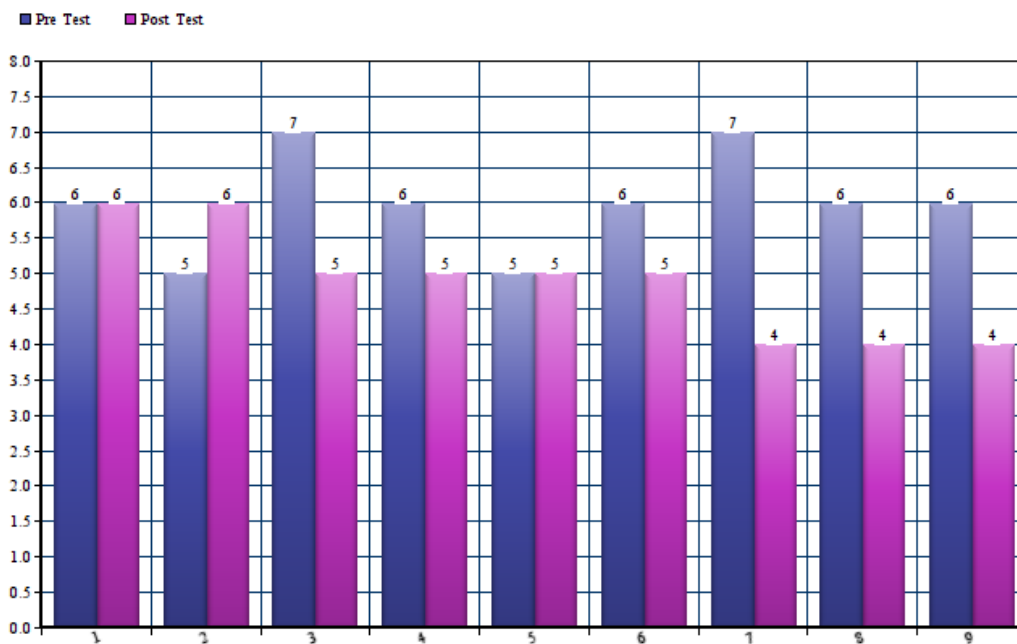


Figura 48. Comparativo de los tiempos obtenidos en el pre test y el post test

Se deduce que el tiempo de proceso de generación de información en el Post Test ha disminuido respecto al del Pre Test; asimismo se comprueba la comparación de medias, teniendo que la media en el pre test es de 6 y la media en el post test es de 4,89. Para evaluar las medias del tiempo de proceso de generación de información, se utiliza la prueba relacional ya que el diseño de la población es observado antes y después (Pre y Post Test).

	Diferencias relacionadas					t	gl	Sig (bilate ral)
	Media	Desviación Tip.	Error tip. de la media	95% Intervalo de confianza para la diferencia				
				Interior	Superior			
Pre Test – Post Test	1.11	0.71	0.423	0.14	2.09	2.626	8	0.030

Figura 49. Tiempo de proceso de generación de información para la toma de decisiones

De acuerdo al resultado generado por medio de las muestras relacionadas, el valor de t es de 2.626; el cual es mayor a 0.060; en tal motivo, el valor de t obtenido a través de la prueba de hipótesis se encuentra en la región de rechazo de la hipótesis nula. Por lo tanto, con un nivel de significancia del 0.030 y un margen de error del 5%, se rechaza la hipótesis nula y se concluye que: El sistema informático reduce el tiempo de generación de información para la toma de decisiones. Se considera entonces que la Hipótesis planteada: La aplicación de Analítica Web, acelera el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web; es válida.

3.1.2.2 Calidad de la información.

Para la evaluación de la hipótesis planteada, contrastamos los resultados obtenidos en el pre test y en el pos test.

Tabla 48

Análisis Pos test – Criterios de evaluación de las 8 dimensiones de la calidad de información

Dimensiones	Proceso	Medición
-------------	---------	----------

Resolución de datos	La información recolectada de los indicadores procesados, están registrados en cantidades numéricas, y a su vez se encuentra dimensionada por aplicación, estado de navegación, Sede de Local del INEI, Tipo de Usuario, País, Estado, Ciudad y por mes y año	Escala [1 - 5]
Estructura de datos	Los indicadores están definidos como tipos de dato numérico. De igual manera las dimensiones, año, mes y estado de la navegación. El resto de dimensiones son de tipo texto. La información no presenta casos de datos dañados, pero si un porcentaje de 0.015 % de IPs que no se consiguió asignar el país de origen.	Escala [1 - 5]
Integración de datos	Se integró información de las fuentes: archivos de registros de los servidores web, información relacionada a los IPs a nivel mundial, información de las aplicaciones web.	Escala [1 - 5]
Relevancia temporal	Teniendo en cuenta que en un día o dos, se puede procesar la información de todos los servidores, se ha iniciado la operación de carga en un periodo de cada 2 semanas.	Escala [1 - 5]
Cronología de datos y meta	La información está definida en una serie de tiempo (mes y año) y se definió como objetivos a medir, los indicadores ya definidos anteriormente. Todos analizados según las dimensiones a consultar.	Escala [1 - 5]

Generalización	La información recolectada procede del total de aplicaciones en los 5 servidores web de plataforma Microsoft. Se asume que en el análisis podemos generalizar los resultados para el estudio del comportamiento de todas las aplicaciones en los servidores Microsoft.	Escala [1 - 5]
Operacionalización	El valor del dato se obtiene de los archivos de registros de los servidores web. Luego en la fase de transformación es recalculado según las diferentes agrupaciones de las dimensiones definidas.	Escala [1 - 5]
Comunicación	La solución de inteligencia de negocios aún está disponible solo al personal responsable de la administración web en OTIN. El objetivo es poder integrarlo con la información del Sistema de Registro de Estadísticas Web. Pero seguirá el lineamiento de que el resultado del análisis no será publicado, ni internamente al personal del INEI, ni será incluido en algún boletín hacia el público externo. Esto por tratarse de información con carácter de privado. Al menos por ahora así será tratado.	Escala [1 - 5]

Tabla 49

Datos del Post test de la evaluación de la calidad de información

Evaluación del personal	Medición	Valor	Valor
		Pre test	Post test
Personal 1	Evaluación de las 8 dimensiones	3.19	4.02

Personal 2	Evaluación de las 8 dimensiones	2.95	4.08
Personal 3	Evaluación de las 8 dimensiones	3.15	3.94
Personal 4	Evaluación de las 8 dimensiones	3.02	4.20

Fuente: Información recolectada con la ficha adjunta en el anexo 6.

En la tabla 49 se visualiza el resultado de la evaluación aplicada a las 8 dimensiones de la calidad de la información del datawarehouse. El mismo que se realizó a través de una ficha de evaluación, la misma que se aprecia en el anexo 6. Luego de la evaluación indicada por cada entrevistado, se calcula el valor final con la fórmula propuesta por Kenett y Shmueli:

$$Q_t = (D_1 * D_2 * D_3 * D_4 * D_5 * D_6 * D_7 * D_8)^{1/8}$$

Finalmente, en el Pre test se alcanza una media de 3.09 y en el Post test se alcanza una media de 4.06. Se considera entonces que la Hipótesis planteada: La aplicación de Analítica Web influye en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web; es válida.

3.1.2.3 Satisfacción del usuario con las interfaces de análisis

Tabla 50

Post Test - Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis

Evaluación del personal	P1	P2	P3	P4	P5	P6	P7
1. Desaprobación Plena	0	0	0	0	0	0	0
2. Desaprobación Simple	0	0	0	0	0	0	0
3. Indecisión o Indiferencia	0	0	0	0	0	0	0
4. Aprobación Simple	2	2	3	1	2	0	0
5. Aprobación Plena	1	1	0	2	1	3	3

Fuente: Información recolectada con la ficha adjunta en el anexo 10

Tabla 51

Pre Test y Post Test – Análisis de la Frecuencia de la evaluación de las preguntas de satisfacción del usuario con las interfaces de análisis

Evaluación del personal	Pre Test	Pre Test	Post Test	Post Test
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
1. Desaprobación Plena	2	9.52	0	0
2. Desaprobación Simple	6	28.57	0	0
3. Indecisión o Indiferencia	9	42.86	0	0
4. Aprobación Simple	4	19.05	10	47.62
5. Aprobación Plena	0	0	11	52.38
Total	21	100	21	100

Fuente: Información recolectada con la ficha adjunta en el anexo 10

La tabla 51, nos indica que la calificación de aprobación de satisfacción, con la solución anterior; es de 19.05% y con la nueva solución es de 100%. Se considera entonces que la Hipótesis planteada: La aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web; es válida.

CAPÍTULO IV. DISCUSIÓN Y CONCLUSIONES

4.1 Discusión

Según los resultados de la presente investigación, se analizó la influencia de la analítica web, mediante comparativas Pre Test y Post test sobre la calidad de la información, tiempo de generación de información, satisfacción del usuario, así como también, la toma de decisiones. También, las soluciones de Inteligencia de Negocios, demuestran que son herramientas muy

importantes que contribuyen a la creación de conocimiento en base al análisis de datos existentes que provienen de dentro y fuera de la empresa.

La influencia de la analítica web en el tiempo de proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI, en la medición del pre test se obtuvo una media de 6 días y con la aplicación de analítica web resultado 4.89 días. Los resultados obtenidos demuestran que se acelera el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI. Según la investigación realizada por Zegarra (2015), quien en su investigación concluye que los resultados obtenidos fueron que el proceso de obtención de reportes del estado de operaciones mineras (procesos de generación de información), en la medición del Pre Test alcanzo una valoración de 72 horas (3 días) y con la aplicación de la solución de inteligencia de negocios se redujo a 12 horas. Por lo que se puede afirmar que con la implementación de la solución de inteligencia de negocios se logró un aceleramiento en el proceso de obtención de reportes del estado de operaciones mineras en la empresa Hochschild Mining.

La influencia de la analítica web en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI, en la medición del pre test alcanzó una media de 3.09 y en la medición del post test alcanzó el 4.06, los resultados obtenidos indican que existe un aumento de 0.97. En consecuencia, se produce un incremento del 31.4%. Según la investigación realizada por Carhuallanqui (2017), quien en su investigación concluye que los resultados obtenidos fueron que la calidad de la información, en la medición del Pre Test alcanzo una valoración de 1.69 y con la aplicación de la solución de inteligencia de negocios aumentó a un 4.26. Los resultados obtenidos indican que existe un aumento de 2.57%, por lo que se puede afirmar que con la implementación de la solución de inteligencia de negocios se logró un incremento de 152% en la calidad de la información en la empresa farmacéutica DISPEFARMA.

La influencia de la analítica web en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI, en la medición del pre test alcanzó una aprobación de 19.05% y en la medición del post test alcanzó el 100%, los resultados obtenidos indican que existe un aumento de 80.95. En consecuencia, se produce un incremento del 325%. Según la investigación realizada por Carhuallanqui (2017), quien en su investigación concluye que los resultados obtenidos fueron que el nivel de satisfacción sobre la información solicitada, en la medición del Pre Test alcanzo una valoración de 1.35 y con la aplicación de la solución de inteligencia de negocios aumentó a un 4.36. Los resultados obtenidos indican que existe un aumento de 3.01, por lo que se puede afirmar que con la implementación de la solución de inteligencia de negocios se logró un incremento de 223% en la satisfacción del usuario sobre la información solicitada en la empresa farmacéutica DISPEFARMA.

La influencia de la analítica web en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI, en el pre test alcanzo una aprobación del 33.33% y en el post test alcanzo el 66.66%, lo que significa un aumento de 33.33%. En consecuencia, se produce un incremento del 100%. Según la investigación realizada por Alva y Callan (2017), quienes en su investigación concluyen que una solución de inteligencia de negocios mejora la toma de decisiones, en base a la mejora del nivel de satisfacción de los usuarios en 52.60 %, la calidad de la información en 60.00 %, el tiempo de generación de reportes que soportan la toma de decisiones en 98.18%, costo de generación de reportes en 98.16 % y el tiempo de análisis de información y formular alternativas en 84.03 %, en la Dirección Ejecutiva del Hospital La Caleta.

4.2 Conclusiones

Se concluye que el promedio del Tiempo de proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web sin la aplicación de analítica web es de 6 días y con la aplicación de analítica web es de aproximadamente 5 días, motivo por el cual, se afirma que la aplicación de Analítica Web, acelera el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI.

Se concluye que la influencia en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI sin la aplicación de Analítica Web alcanza una media de 3.09 y con la aplicación de Analítica Web alcanza el 4.06, lo que significa un aumento de 0.97. En consecuencia, se produce un incremento del 31.4%. Por lo tanto, la aplicación de Analítica Web, influye significativamente en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI.

Se concluye que la influencia en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI sin la aplicación de Analítica Web alcanza el 19.05% y con la aplicación de Analítica Web alcanza el 100%, lo que significa un aumento de 80.95%. En consecuencia, se produce un incremento del 325%. Por lo tanto, la aplicación de Analítica Web, influye significativamente en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI.

Se concluye que la influencia en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI sin la aplicación de Analítica Web alcanza el 33.33% y con la aplicación de Analítica Web alcanza el 66.66%, lo que significa un aumento de 33.33%. En consecuencia, se produce un incremento del 100%. Por lo tanto, la aplicación de Analítica Web,

influye significativamente en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web en el INEI.

El modelo integrado de fuentes de información, la visualización y el análisis de información a través de una herramienta BI, propuesto por Avinash Kaushik, es el sustento de este proyecto y se afirma que su implementación cumple con el objetivo de aplicar analítica web para la gestión de toma de decisiones.

Los beneficios en la toma de decisiones, se disemina en los usuarios que tienen acceso a la solución. Por ejemplo, propiamente en OTIN, la unidad de producción, con el fin de inventariar el parque de aplicaciones web pueden ahora dar de baja a aplicaciones cuyo periodo de ejecución ha culminado, o analizar la tasa de visitas a aplicaciones de difusión implementadas hace una década atrás y evaluar su mantenimiento o retiro. La unidad de desarrollo puede analizar la información de los errores de conexión a las bases de datos o los casos de no publicación que se presentan en las aplicaciones web. Finalmente, para los administradores web, se convierte en una herramienta que les permita monitorear el funcionamiento de las aplicaciones, el origen de las visitas, casos de error, volumen de información descargado y recibido etc. Para la Dirección Técnica de difusión, primordialmente puede monitorear si las aplicaciones de difusión y publicación de información, están cumpliendo sus objetivos y hacia qué tipo de usuarios.

4.3 Recomendaciones

Es cierto que el tiempo de procesamiento de carga de información de la solución implementada, es más rápida que la realizada por el proceso actual. Pero no la reemplazara. Por el contrario, debe implementarse un algoritmo de procesamiento de carga de la información que viene de la cuenta de Google Analytic, conllevando a que el tiempo actual sea optimizado e incluir en el mismo proceso, la extracción de información de más indicadores.

La fase de integración de información, no solo debe recolectar lo proveído por los archivos de registro de los servidores web, sino también, debe incluir la información que viene de la cuenta de Google Analytic. Así mismo, debe incluir las fuentes de información de otros canales, como las redes sociales y canales de video.

La brecha del tiempo de la ejecución del proceso de carga de información, debe ser más corto, para que el usuario pueda contar con información más actualizada y así poder analizar con más prontitud, conllevando a la mejora de la toma decisiones.

Urge la implementación del sistema de encuesta de satisfacción en el sitio web del INEI. La misma que debe ser accesible desde todas las aplicaciones web y pueda cubrir todos los casos de interrogantes de las diferentes aplicaciones.

Poder contar con la versión licenciada de la herramienta BI: Qlik Sense, para poder explotar más funcionalidades en beneficio de la integración de la información y del análisis de la información.

REFERENCIAS

Albert Nogués y Juan Valladares (2017). Business Intelligence Tools for Small Companies: A Guide to Free and Low-Cost Solutions. Recuperado de <https://learning.oreilly.com/library/view/business-intelligence-tools/9781484225677/>

Avinash Kaushik (2007). Web Analytics One Hour One Day. Recuperado de

<https://learning.oreilly.com/library/view/web-analytics-an/9780470130650/>

Bernard Marr (2016). Key Business Analytics. Recuperado de

<https://learning.oreilly.com/library/view/key-business-analytics/9781292017464/>

Carlos Sabino (1996). El proceso de investigación. Obtenido de

<http://camp.ucss.edu.pe/ingenium/index.php/industrial/126-nuevas-tendencias-para-elaborar-tesis-en-la-carrera-de-ingenieria-industrial>

Jeffrey W. Herrmann (2015). Engineering Decision Making and Risk Management.

Recuperado de <https://learning.oreilly.com/library/view/engineering-decision-making/9781118919378/>

Michael Beasley (2013). Practical Web Analytics for User Experience. Recuperado de

<https://learning.oreilly.com/library/view/practical-web-analytics/9780124046191/>

Rick Sherman (2015). Business Intelligence Guidebook. Recuperado de

<https://learning.oreilly.com/library/view/business-intelligence-guidebook/9780124114616/>

Ron S. Kenett y Galit Shmueli (2016). Information Quality. Recuperado de

<https://learning.oreilly.com/library/view/information-quality/9781118874448/>

Thomas H. Davenport y Jeanne G. Harris (2007). Competing on analytics: The new science

of winning. Recuperado de <https://learning.oreilly.com/library/view/competing-on-analytics/9781422156308/>

Tilo Linz (2014). Testing in Scrum. Recuperado de

<https://learning.oreilly.com/library/view/testing-in-scrum/9781492001508/>

Zorrilla, Arena (1993). Introducción a la metodología de la investigación. Obtenido de

<http://tesisdeinvestig.blogspot.com/2011/05/tipos-de-investigacion.html>

Anexo 1 – Matriz de Consistencia

Problemas	Objetivos	Hipótesis	Variables	Dimensiones	Indicadores	Medición
Problema General	Objetivo General	Hipótesis General	Variable Independiente			
¿En qué medida la aplicación de Analítica Web permite la toma de decisiones acerca del tráfico de visitas de las aplicaciones web?	Determinar en qué medida la aplicación de Analítica Web permite la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	La aplicación de Analítica Web, influye significativamente en la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	Analítica Web	Reducción del tiempo de generación de información para la toma de Decisiones.	Tiempo de generación de información para la toma de Decisiones.	Días laborales
Problema Específico 1	Objetivo Específico 1	Hipótesis Específica 1		Calidad de la información. Las ocho Dimensiones.	Rango de valor del nivel de calidad de información evaluada para todas las dimensiones.	Escala [1 - 5]
¿De qué manera la aplicación de Analítica Web influye en el tiempo de generación de información para la toma de Decisiones acerca del tráfico	Acelerar el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	La aplicación de Analítica Web, acelera el proceso de generación de información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.				

de visitas de las aplicaciones web?						
Problema Específico 2	Objetivo Específico 2	Hipótesis Específica 2		Satisfacción del usuario con las interfaces de análisis	Rango de valor del nivel de Satisfacción del usuario con las interfaces de análisis	Escala [1 - 5]
¿De qué manera la aplicación de Analítica Web influye en la calidad de información para la toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?	Determinar en qué medida la aplicación de Analítica Web influye en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	La aplicación de Analítica Web influye en la calidad de la información para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.				
Problema Específico 3	Objetivo Específico 3	Hipótesis Específica 3	Variable Dependiente	Dimensiones	Indicadores	Medición
¿De qué manera la aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de Decisiones acerca del tráfico de visitas de las aplicaciones web?	Determinar en qué medida la aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	La aplicación de Analítica Web influye en la satisfacción del usuario con las interfaces de análisis para la toma de decisiones acerca del tráfico de visitas de las aplicaciones web.	Toma de decisiones	Aplicación del modelo Trinity	Métricas de las 3 dimensiones del modelo Trinity	Escala [1 - 5]

	de visitas de las aplicaciones web.					
--	-------------------------------------	--	--	--	--	--

Anexo 2 – Administración de cuentas de Google Analytic 1

Solo se muestran las métricas de las vistas destacadas. Las métricas de Web y aplicación no se admiten en este momento.

	Sesiones	Duración media de la sesión	Porcentaje de rebote	Tasa de conversión del objetivo
☆ Bases de Datos				
☆ ANDA (UA-46985108-2)				
☆ Todos los datos de sitios web	-	-	-	-
☆ Desarrollo (UA-46985108-3)				
☆ Todos los datos de sitios web	-	-	-	-
☆ Micro Datos (UA-46985108-1)				
☆ Todos los datos de sitios web	-	-	-	-
☆ Censos				
☆ CPV 2005 - Redatam (UA-47056522-9)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 2007 - Cuadros Estadísticos (UA-47056522-5)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 2007 - Mapa de Pobreza (UA-47056522-7)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 2007 - Principales Indicadores (UA-47056522-6)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 2007 - Redatam (UA-47056522-8)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 81 - Redatam (UA-47056522-14)				
☆ Todos los datos de sitios web	-	-	-	-
☆ CPV 93 - Centros Poblados (UA-47056522-10)				

Figura 50. Visualización de la administración de las cuentas en el servicio Google Analytic del INEI.

Anexo 3 – Administración de cuentas de Google Analytic 2

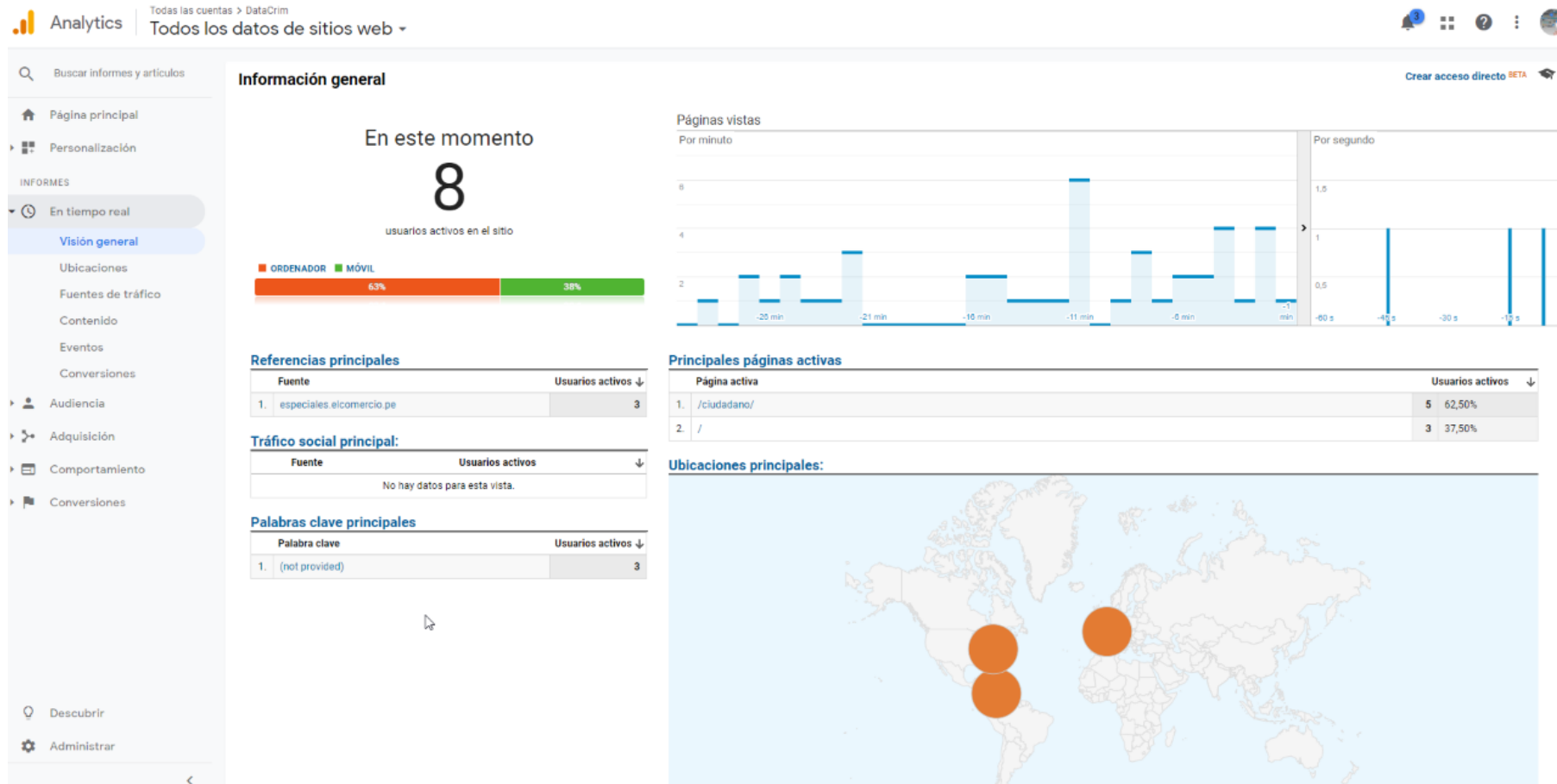


Figura 51. Visualización de la información por cada sitio web configurado en el servicio Google Analytic del INEI.

Anexo 4 – Administración de Sistema de Estadísticas Web 1

Editar Registro - Misión y Visión

Estimado usuario, solo se podrá modificar la cantidad de visitas para este módulo, así mismo tenga cuidado si se va cambiar el archivo de enlaces que contiene este modulo.

Enlace	<input type="text" value="/mision-y-vision/"/>
Año	<input type="text" value="2014"/>
Mes	<input type="text" value="ENERO"/>
Semana	<input type="text" value="semana 1"/>
Visitas	<input type="text" value="84"/>
Agregar Enlaces de Descargas	<input type="button" value="Seleccionar archivo"/> Ningún archivo seleccionado

Figura 52. Pantalla de edición o registro de información en el sistema de Estadísticas Web del INEI

Anexo 5 – Administración de Sistema de Estadísticas Web 2

Estadísticas Portal INEI

Estadísticas por modulo Estadísticas por producto

Búsqueda

Año de Inicio -- Seleccione -- Mes de Inicio -- Seleccione -- [Filtrar](#)

Condición de filtro -- Seleccione --

Año de Fin -- Seleccione -- Mes de Fin -- Seleccione -- [Exportar](#)

MODULO	MODULO	MODULO	2014
Inicio			2343679
Mapa de sitio			16571
Preguntas frecuentes			36231
Contáctenos			16502
Acerca del INEI	Sistema Estadístico Nacional		27476
	Nosotros		13514
	Misión y Visión		11615
	Código de Buenas Prácticas		3142
	Historia		6823
	Jefe del INEI		14377
	Rendición de Cuentas del Titular		0
	Organigrama Funcional		14963
	Seguridad y Salud en el Trabajo		0
Sedes		32350	
Servicios	Catálogo de Publicaciones y Servicios		0
	Centro de Información		77555
	Biblioteca Institucional		16885
Prensa	INEI en los Medios		8727
	Notas de Prensa		263954
Escuela			413
Portal de Transparencia			18
Índice Temático	Economía		265820
	Población y Vivienda		282235
	Empleo		160529
	Sociales		104145
	Medio Ambiente		71115

Figura 53. Visualización de información en el sistema de Estadísticas Web del INEI.

Anexo 6 – Ficha de Evaluación de las dimensiones de la calidad de la información 1

Oficina Técnica de Informática - INEI

FICHA DE EVALUACION DE LAS DIMENSIONES DE LA CALIDAD DE INFORMACION

PRE TEST () / POST TEST ()

El objetivo de la presente ficha, es poder registrar la evaluación de las dimensiones de la calidad de información a la base de datos generada aplicando la Analítica Web. Estas 8 dimensiones están definidas según Kenett y Shmueli, y cuya explicación se difundirá a la entrega de esta ficha al personal a cargo del llenado de la misma. La valoración a designar está basada en una escala de Likert.

Escala de Valoración

1. No es importante
2. Poco importante
3. Neutral
4. importante
5. Muy importante

Datos del Personal

Unidad en la que labora
Nombre
Cargo
Fecha

Dimensión	No es importante (1)	Poco importante (2)	Neutral (3)	Importante (4)	Muy importante (5)	Observaciones
Resolución de datos						
Estructura de datos						
Integración de datos						

Anexo 7 – Ficha de Evaluación de las dimensiones de la calidad de la información 2

Relevancia temporal									
Cronología de datos y meta									
Generalización									
Operacionalización									
Comunicación									



Ing. Michael Malaga Matasich
Director Ejecutivo
Oficina Ejecutiva de Soporte Técnico
Oficina Técnica de Estadística

Michael Malaga Matasich - Jefe de Unidad de Producción



Michael Macavilca Mejía - Analista de Sistemas

Anexo 8 – Ficha de Evaluación de las dimensiones de toma de decisión en la analítica web

Oficina Técnica de Informática - INEI

FICHA DE EVALUACION DE LAS DIMENSIONES DE TOMA DE DECISION EN LA ANALITICA WEB

PRE TEST () / POST TEST ()

El objetivo de la presente ficha, es poder registrar la evaluación de las dimensiones de la toma de decisiones en la Analítica Web, dirigido a las soluciones presentadas. Estas 3 dimensiones están definidas según Avinash Kaushik, y cuya explicación se difundirá a la entrega de esta ficha al personal a cargo del llenado de la misma. La valoración a designar está basada en una escala de Likert.

Escala de Valoración		Datos del Personal				
1. No es importante		Unidad en la que labora				
2. Poco importante		Nombre				
3. Neutral		Cargo				
4. Importante		Fecha				
5. Muy importante						
Dimensión	No es importante (1)	Poco importante (2)	Neutral (3)	Importante (4)	Muy importante (5)	Observaciones
Análisis del comportamiento						
Análisis de resultados						
Análisis de Experiencia						



INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA
MICHAEL MACAVILCA MEJIA

Michael Macavilca Mejia - Analista de Sistemas



Ing. Michael Malaga Matasich
Oficina Ejecutiva de Soporte Técnico
Diseñador Escultivo
Michael Malaga Matasich - Jefe de Unidad de Producción

**Anexo 9 – Ficha de Observación para el registro del indicador: Tiempo de generación
de información para la toma de decisiones**

FICHA DE OBSERVACION PARA EL REGISTRO DEL INDICADOR: TIEMPO DE GENERACION DE INFORMACION PARA LA TOMA DE DECISIONES

PRE TEST () / POST TEST ()

Investigador	Michael John Macavilca Mejia
Institución	Instituto Nacional de Estadística e Informática
Proceso observado	Generación de información para la toma de decisiones
Proyecto	Aplicación de Analítica Web para la toma de decisiones Acerca del tráfico de visitas de las aplicaciones web, En el INEI, Lima 2019

Número de Observación	Fecha de Inicio	Fecha Fin	Instrumento
1			Observación directa del mensaje final de la ejecución del Store Procedure.
2			
3			
4			
5			
6			
7			
8			
9			



Michael Macavilca Mejia - Analista de Sistemas



ing. Michael Malaga Matasich
Director Ejecutivo
Oficina Ejecutiva de Soporte Técnico
Oficina Técnica de Informática
Michael Malaga Matasich - Jefe de Unidad de Producción

Anexo 10 – Ficha de Evaluación de las preguntas sobre la satisfacción del usuario con las interfaces de análisis 1

Oficina Técnica de Informática - INEI

FICHA DE EVALUACION DE LAS PREGUNTAS SOBRE LA SATISFACCION DEL USUARIO CON LAS INTERFACES DE ANALISIS

PRE TEST () / POST TEST ()

El objetivo de la presente ficha, es poder registrar la evaluación de la satisfacción del usuario con las interfaces de análisis de la solución BI aplicando la analítica web. La evaluación consta de 7 preguntas y su valoración está basada en una escala de Likert.

Escala de Valoración
1. Desaprobación Plena
2. Desaprobación Simple
3. Indecisión o Indiferencia
4. Aprobación Simple
5. Aprobación Plena

Datos del Personal	
Unidad en la que labora	
Nombre	
Cargo	
Fecha	

Dimensión	Desaprobación Plena (1)	Desaprobación Simple (2)	Indecisión o Indiferencia (3)	Aprobación Simple (4)	Aprobación Plena (5)	Observaciones
¿El sistema es de fácil uso?						
¿Está usted satisfecho con la variedad de tableros que ofrece el sistema?						
¿La interfaz del sistema es amigable para usted?						

**Anexo 10 – Ficha de Evaluación de las preguntas sobre la satisfacción del usuario con
las interfaces de análisis 2**

¿Los tableros emitidos son fáciles de analizar?			
¿Cómo considera la presentación de los tableros analíticos?			
¿Cómo considera el tiempo de carga de los tableros analíticos?			
¿La información que le brinda los tableros le ayuda a realizar una mejor analítica web?			



Michael Macavilca Mejía - Analista de Sistemas



Ing. Michael Malaga Marasich
Michael Malaga Marasich - Jefe de Unidad de Producción
Oficina Especial de Soporte Técnico
Oficina Técnica de Informática