



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá

Carlos Aurelio Angulo Romero

Universidad Nacional de Colombia
Facultad de Ciencias Económicas
Programa de Maestría en Administración
Bogotá, Colombia

2021

Evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá

Carlos Aurelio Angulo Romero

Trabajo final presentado como requisito parcial para optar al título de:

Magíster en Administración

Directora:

Jenny Marcela Sánchez Torres

Línea de Investigación:

Estrategia y Organizaciones

Grupo de Investigación en Gestión y Organizaciones (Griego)

Universidad Nacional de Colombia
Facultad de Ciencias Económicas
Programa de Maestría en Administración
Bogotá, Colombia

2021

Dedicatoria

*A mi familia, por todo el apoyo que he recibido
para culminar este proyecto.*

Agradecimientos

A mi esposa que siempre me ha apoyado con las metas que me he impuesto.

A mi tutora Jenny Marcela Sánchez Torres por la guía brindada para sacar adelante este trabajo.

A los profesores, de los cuales aprendí bastante y me llevaron a tomar decisiones y a realizar cambios en mi empresa; lo anterior me ha permitido orientarla a un constante crecimiento.

A mis compañeros de estudio que me brindaron su apoyo en cada curso culminado y que a través de sus experiencias en las organizaciones en las cuales ejercen su actividad actual, me aportaron información de valor para mi desarrollo personal y profesional.

A las empresas que me compartieron un poco de su tiempo y experticia, para poder llevar a cabo este trabajo, mi eterna gratitud.

Resumen

Evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá

Las nuevas Tecnología de la Información y Comunicación han motivado cambios significativos en la dinámica comercial y organizacional de las empresas en el mundo, sobre todo en lo referente a la forma como se analizan los datos para la toma de decisiones. Debido a lo anterior, el presente trabajo de maestría aborda la evaluación de la capacidad de análisis de datos de las pymes que se desenvuelven en el sector de desarrollo de software de la ciudad de Bogotá, buscando caracterizar los modelos de diagnóstico empleados en el entorno, para determinar los niveles de madurez del parámetro observado en las entidades estudiadas, en función de las dimensiones de Organización, Infraestructura, Gestión de recursos, Análisis y Gobernanzas.

Así las cosas, se plantean recomendaciones para fortalecer dicha capacidad de análisis. Para tal fin, se empleó una metodología de tipo descriptivo, de enfoque mixto, que combina técnicas cuantitativas y cualitativas, con un diseño de investigación no experimental. Dentro de los resultados obtenidos fue posible identificar el nivel de madurez de la capacidad de análisis de datos de las empresas que participaron en la investigación, con puntos de referencia como la discriminación por ingresos recibidos por las mismas. De igual manera se realizó un estudio minucioso de los hallazgos en relación con las dimensiones organizacionales consideradas. Finalmente se sugieren acciones que podrían ser de utilidad para el cierre de las brechas de desarrollo identificadas.

Palabras clave: Capacidad de análisis, Analytics Maturity (TDWI), pymes.

Abstract

Evaluation of the data analysis capacity of software development SMEs in the city of Bogota

The new Information and Communication Technologies have led to significant changes in the commercial and organizational dynamics of companies in the world, especially in terms of how data are analyzed for decision making. Due to the above, this master's thesis addresses the evaluation of the data analysis capacity of SMEs that operate in the software development sector in the city of Bogota, seeking to characterize the diagnostic models used in the environment, to determine the levels of maturity of the parameter observed in the entities studied, according to the dimensions of Organization, Infrastructure, Resource Management, Analysis and Governance.

Thus, recommendations are made to strengthen this analytical capacity. To this end, a descriptive methodology was used, with a mixed approach, combining quantitative and qualitative techniques, with a non-experimental research design. Among the results obtained, it was possible to identify the level of maturity of the data analysis capacity of the companies that participated in the research, with reference points such as discrimination by income received by them. A detailed study of the findings was also carried out in relation to the organizational dimensions considered. Finally, actions are suggested that could be useful for closing the development gaps identified.

Keywords: Analysis capacity, Analytics Maturity (TDWI), SMEs.

Contenido

	Pág.
Introducción	15
1. Los modelos de capacidad de análisis de datos: una revisión de la literatura....	21
1.1 Método.....	21
1.1.1 Planificación	22
1.1.2 Búsqueda de información y análisis preliminar	23
1.2 Analítica de datos.....	25
1.3 Capacidad de análisis de datos.....	30
1.4 Modelos para medir la capacidad de análisis de datos.....	35
1.5 Resumen del capítulo.....	46
2. Ajuste del modelo de evaluación de madurez TDWI para su aplicación	48
2.1 El Modelo de TDWI (Transforming Data With Intelligence™)	48
2.1.1 Niveles de madurez.....	49
2.1.2 Dimensiones de madurez	51
2.2 Resumen del capítulo.....	54
3. Aplicación del modelo TDWI.....	55
3.1 Método.....	55
3.2 Resultados generales de la aplicación del modelo TDWI	57
3.2.1 Organización	58
3.2.2 Infraestructura	59
3.2.3 Gestión de recursos.....	60
3.2.4 Análisis	61
3.2.5 Gobernanza.....	61
3.3 Resultados discriminados de la aplicación del modelo TDWI	62
3.4 Evaluación e interpretación de las brechas de la capacidad de análisis de datos en las Pymes	72
3.5 Resumen del capítulo.....	76
4. Recomendaciones para el fortalecimiento de la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá	79
4.1 Recomendaciones generales para fortalecer la capacidad de análisis de datos.....	79
4.1.1 Organización	79
4.1.2 Infraestructura	80
4.1.3 Gestión de recursos.....	80
4.1.4 Análisis	80
4.1.5 Gobernanza.....	81

4.2	Resumen del capítulo	81
5.	Conclusiones y recomendaciones	83
5.1	Conclusiones	83
5.2	Recomendaciones	86
6.	Bibliografía.....	89
A.	Anexo: Cuestionario aplicado a las empresas.....	103
B.	Anexo: Informe final enviado a las empresas que participaron en la investigación.....	128

Lista de figuras

	Pág.
Figura 1-1: Estrategias para la búsqueda para la definición de capacidad de análisis de datos	23
Figura 1-2: Número de documentos encontrados en Scopus según ecuaciones de búsqueda	24
Figura 1-3: Tipo de documentos en Scopus según ecuaciones de búsqueda	25
Figura 1-4: evolución del concepto de análisis de datos.....	30
Figura 2-1: Niveles de madurez TDWI en la capacidad de análisis de datos.....	49
Figura 2-2: Dimensiones del modelo TDWI	52
Figura 3-1: Madurez obtenida en pymes desarrolladoras de software de Bogotá DC y alrededores a lo largo de dimensiones del modelo TDWI.....	58
Figura 3-2: Proporción de pymes desarrolladoras de software en Bogotá y alrededores en función de ingresos netos anuales	62
Figura 3-3: Madurez de pymes desarrolladoras de software en Bogotá y alrededores en función de sus ingresos netos anuales.....	64
Figura 3-4: elementos de madurez evaluados en la dimensión <i>Organización</i> en pymes desarrolladoras de software de Bogotá DC y alrededores.....	66
Figura 3-5: elementos de madurez evaluados en la dimensión <i>Infraestructura</i> en pymes desarrolladoras de software de Bogotá DC y alrededores.....	68
Figura 3-6: elementos de madurez evaluados en la dimensión <i>Gestión de recursos</i> en pymes desarrolladoras de software de Bogotá DC y alrededores	69
Figura 3-7: elementos de madurez evaluados en la dimensión <i>Análisis</i> en pymes desarrolladoras de software de Bogotá DC y alrededores.....	69
Figura 3-8: elementos de madurez evaluados en la dimensión <i>Gobernanza</i> en pymes desarrolladoras de software de Bogotá DC y alrededores.....	71

Lista de tablas

	Pág.
Tabla i-1: Cuadro de operacionalización de variables	18
Tabla 1-1: Análisis de datos	26
Tabla 1-2: Capacidad de análisis de datos	31
Tabla 1-3: Comparación de modelos de medición de capacidad de análisis de datos....	39
Tabla 2-1: Puntajes del modelo TDWI para medir la capacidad de análisis de datos	51

Introducción

El impacto de los procesos de análisis de datos en la organización comienza a cobrar más relevancia, dado que el mundo avanza hacia una dinámica en la cual las organizaciones deben hacer uso de la información que se produce en el entorno, tanto interno como externo, con el fin de comprender de mejor manera el contexto en el que operan. No obstante, las características de las pymes dificultan dar este salto tecnológico para gestionar y sacar verdadero provecho de los datos (Coleman, 2016). Por ello, ante toda la dimensionalidad de esta problemática, resulta necesario acompañar a estas organizaciones en su tránsito hacia una transformación digital sostenible.

En consecuencia, la capacidad de análisis de datos para sustentar decisiones de negocio en las pymes parece ser un campo con gran potencial para ser explorado, dado que las posibilidades para extraer información estructurada, mediante procesos analíticos y predictivos propios de la minería de datos, puede contribuir a crear mayor valor en este tipo de organizaciones (Schab et al., 2018). Sin embargo, según Sen et al (2016), aún existen falencias importantes en estas entidades, en el ámbito global, que impiden que la incorporación de estas herramientas sea vista como una oportunidad real para el crecimiento y progreso, basado fundamentalmente en las limitaciones técnicas y de experiencia que se requiere.

En concreto, las causas que se relacionan con el problema identificado están asociadas con las falencias que tienen las pymes para aprovechar los datos de manera estructurada y alineada con los objetivos estratégicos. Además, la incorporación de TI, por sí sola, no garantiza el éxito de estrategias analíticas, puesto que requiere programas de formación y articulación con otros procesos

empresariales. Igualmente, la ausencia de un compromiso sólido de la gerencia para establecer una cultura organizacional basada en información podría estar disminuyendo la capacidad de análisis de datos de las pymes.

Al no atender la situación actual que se presenta, las consecuencias pueden ser varias, entre estas, la desmejora en el proceso de toma de decisiones, las cuales estarían siendo identificadas con base en la intuición o experiencia. En el entorno cambiante al que se enfrenta hoy en día la empresa, mayormente generado por problemas de salud pública, es imprescindible que se cuente con la capacidad para analizar datos de múltiples fuentes y, a partir de estos, decidir las estrategias más efectivas para enfrentar el futuro (Grossman, 2018) (Weber et al., 2017).

De otra parte, la transformación en los patrones de consumo y dinámicas sociales hacen que la industria del software se ajuste a las necesidades del mercado de forma continua, por tal motivo, estar desconectado de esta realidad implicaría posibles pérdidas económicas y financieras, impactando en aspectos clave como la productividad y competitividad. En consecuencia, las principales afectadas serán las pymes desarrolladoras de software de la ciudad de Bogotá y los trabajadores que laboran en ella, complicando aún más su estabilidad y empleabilidad.

En virtud de lo anterior, no solucionar esta problemática pondría a las pymes en contra de las tendencias globales que sitúan a la analítica como un activo crítico para la toma de decisiones, generando un impacto negativo en su competitividad, que amenazaría el crecimiento y sustentabilidad en los próximos años, obligando a la gerencia a redireccionar recursos y disminuir personal en los casos más complejos (Masood & Sonntag, 2020). De este modo, se observa que los efectos pueden tener un alcance mayor al esperado (Villa & Taurino, 2019); por ello, es necesario establecer la siguiente pregunta de investigación que responda a esta inquietud: ¿Cuál es la madurez de la capacidad de análisis de datos en las pymes de desarrollo de software de la ciudad de Bogotá?

Objetivos

Con base en lo descrito anteriormente, se trazan los siguientes objetivos:

General:

- Estimar la madurez de la capacidad de análisis de datos en las pymes del sector de desarrollo de software de la ciudad de Bogotá.

Específicos:

- Caracterizar los modelos para el diagnóstico de la capacidad de análisis de datos en las organizaciones a través de una revisión sistemática de literatura.
- Determinar el modelo de madurez en análisis de datos para su aplicación en las pymes desarrolladoras de software de la ciudad de Bogotá a través de un análisis comparativo.
- Proponer recomendaciones que permitan fortalecer la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá.

Diseño de la investigación

Con el propósito de alcanzar los objetivos señalados, el presente proyecto es de tipo descriptivo, puesto que busca exponer los componentes, elementos y factores característicos del fenómeno analizado (Sampieri, 2018). Por otra parte, se emplea una estrategia de investigación con enfoque mixto que combina técnicas cuantitativas y cualitativas para recabar la información requerida que permita responder a los objetivos del estudio. Al respecto, Creswell y Creswell (2017) indican que este tipo de abordaje permite al investigador utilizar instrumentos como la encuesta, para recolectar datos subjetivos basados en percepciones y transformarlos en mediciones numéricas, apoyándose en procedimientos matemáticos y estadísticos.

A su vez, el marco del diseño adoptado es no experimental, el cual se fundamenta en la observación de los hechos en la realidad sin intervenir o alterar las condiciones (Reio, 2016), haciéndolo el más adecuado en el contexto de estudios que involucran organizaciones en ciencias sociales, dado que el rol del investigador es plasmar lo que identifica tal y como sucede. Seguidamente, se expone de forma detallada aspectos relacionados con el instrumento a aplicar en las pymes desarrolladoras de software de la ciudad de Bogotá (Tabla i-1).

Tabla i-1: Cuadro de operacionalización de variables

Objetivos específicos	Actividades metodológicas	VARIABLES DE TRABAJO	Técnicas de tratamiento de la información	Resultados esperados
Caracterizar los modelos para el diagnóstico de la capacidad de análisis de datos en las organizaciones a través de una revisión sistemática de literatura.	<ul style="list-style-type: none"> ▪ Realizar una revisión literaria de los modelos de evaluación de madurez de análisis de datos. ▪ Detallar los componentes de los modelos identificados, mostrando su funcionalidad y aportes más relevantes. 	<ul style="list-style-type: none"> ▪ Modelos de capacidad de análisis de datos. ▪ Dimensiones y niveles de madurez contemplados. ▪ Existencia de instrumentos de recolección de la información. 	Revisión documental de publicaciones científicas y libros recientes sobre la temática.	Lograr una identificación y comprensión de los modelos de evaluación de madurez de análisis de datos existentes.
Determinar el modelo de madurez en análisis de datos para su aplicación en las pymes desarrolladoras de software de la ciudad de Bogotá a través de un análisis comparativo.	Analizar los modelos de evaluación de madurez identificados en la fase anterior y seleccionar uno de ellos o una combinación entre estos para determinar el modelo a aplicar en el presente estudio.	<ul style="list-style-type: none"> ▪ Integralidad de los modelos de capacidad de análisis de datos. ▪ Replicabilidad de los modelos de análisis de datos. 	<ul style="list-style-type: none"> ▪ Revisión documental de publicaciones científicas y libros recientes sobre la temática. ▪ Análisis comparativo. 	Lograr escoger un modelo de evaluación de madurez de análisis de datos con amplio abaraje (integralidad) y con alta replicabilidad.
Proponer recomendaciones que permitan fortalecer la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá.	<ul style="list-style-type: none"> ▪ Identificar la población y muestra. ▪ Aplicar el instrumento TDWI ajustado a las pymes seleccionadas. ▪ Tabular, procesar, interpretar y exponer los resultados del modelo TDWI. ▪ Definir lineamientos, acciones y actividades que encaminen a las pymes tecnológicas a mejorar y fortalecer la capacidad de análisis de datos. 	<ul style="list-style-type: none"> ▪ Organización ▪ Recursos ▪ Infraestructura ▪ Analítica ▪ Gobernanza ▪ Capacidad de análisis de datos en las pymes (categoría/nivel de madurez). 	<ul style="list-style-type: none"> ▪ Análisis estadístico de los datos recabados del cuestionario TDWI según la metodología planteada por (Halper y Krishnan, 2014) ▪ Resultados de los objetivos 1, 2 y 3. 	<ul style="list-style-type: none"> ▪ Capacidad de análisis de datos de cada una de las pymes desarrolladoras de software seleccionadas en la muestra ▪ Estrategias orientadas a mejorar y fortalecer la capacidad de análisis de datos en la pyme del sector tecnológico.

Fuente: elaboración propia

En consecuencia, los resultados obtenidos permiten obtener un diagnóstico del nivel de madurez del análisis de datos de las pymes desarrolladoras de software y las brechas de desarrollo útiles para contribuir con el aprovechamiento del análisis de datos como una verdadera ventaja estratégica. Asimismo, la investigación propuesta ofrece una visión aplicable del modelo TWDI en un contexto empresarial real, útil como base de estudio para futuras utilidades.

Estructura del documento

El trabajo final de maestría se organizó en cinco capítulos que desarrollan fundamentos que dan respuesta a los objetivos específicos. En este orden de ideas, en el apartado introductorio se expone todo lo relacionado con la problemática identificada, se definen las actividades principales y el diseño metodológico empleado. En el Capítulo I se elabora un marco de referencia que aborda nociones conceptuales, con base en una revisión de literatura relacionada, con el fin de identificar y analizar los modelos de evaluación de madurez creados, así como sus dimensiones, elementos y categorías.

Adicionalmente, en el Capítulo II se realiza una explicación del modelo de evaluación de madurez, profundizado en cada uno de los componentes que lo conforman y la traducción del instrumento de recolección de información del modelo escogido. Finalmente, en el Capítulo III se desarrollan los resultados concretos de la aplicación del modelo determinado en todas las dimensiones definidas; para, con ello, en el Capítulo IV desarrollar una serie de recomendaciones para el fortalecimiento de la capacidad de análisis en las compañías, con el fin de apoyar su transformación hacia escenarios de mayor crecimiento y productividad.

1. Los modelos de capacidad de análisis de datos: una revisión de la literatura

El objetivo de este capítulo es abordar los conceptos y modelos asociados con la capacidad de análisis de datos en las pymes, para ello, se indaga en la literatura reciente, tanto en idioma inglés como español, para ubicar teorías y modelos que permitan establecer una reflexión comparativa de los modelos que se han desarrollado y se practican en este campo, actualmente.

Inicialmente, se presenta el método utilizado para la revisión de la literatura, a fin de proporcionar un esquema general del estado del arte en esta materia. Seguidamente, en la sección 1.2. se conceptualiza el análisis de datos y su evolución en los últimos años, mientras que en la sección 1.3 se realiza una reflexión sobre el concepto de “capacidad de análisis de datos” con base en varios autores. Finalmente, en los apartados 1.4 y 1.5 se discute acerca de los modelos, haciendo especial énfasis en aquel que ha sido seleccionado para la presente investigación.

1.1 Método

Con el propósito de comprender el significado de la capacidad de análisis de datos e identificar los modelos que puedan medir ese parámetro en las pymes, se realizó una revisión sistemática de la literatura tomando como referencia el proceso sugerido por Kitchenham (2004), estableciendo los siguientes pasos:

- Exploración inicial de artículos asociados con la capacidad de análisis de datos en las pymes

- Revisión de modelos para medir la capacidad de análisis de datos en las pymes
- Contraste, estudio y selección del modelo para medir la capacidad de análisis de datos en las pymes

1.1.1 Planificación

Siguiendo el método de la revisión de la literatura, se establecieron las siguientes preguntas orientadoras:

- ¿Cuál ha sido la evolución de conceptos como *análisis de datos* y *capacidad de análisis de datos*?
- ¿Cuáles son los modelos que han venido desarrollándose en los últimos años para medir la capacidad de análisis de datos?
- ¿Cuáles son las características en términos de dimensiones, variables, indicadores, instrumentos plantean cada uno de los modelos?

En este sentido, una vez definidas las orientaciones generales y siguiendo el protocolo sugerido por Kitchenham (2004), fueron utilizados operadores lógicos de búsqueda OR/AND en las mencionadas bases de datos y repositorios, con el fin de auxiliar el proceso y atender criterios de inclusión (OR) y exclusión (AND) en las indagaciones realizadas. Por consiguiente, como criterios de exclusión para la búsqueda de documentos relacionados con la capacidad de análisis de datos se tuvieron en consideración las palabras clave "*data analytics capability*" y "*definition*", de tal forma que cualquier otro término fue excluido de la búsqueda.

Por otra parte, para la búsqueda de modelos de evaluación de madurez relacionados con el análisis de datos, se tuvieron en consideración criterios combinados, es decir, resultados que contuvieran alguna de las palabras clave "*data analytics capability*" o "*data analytics capabilities*" y que, a su vez, asociaran palabras clave que comenzaran por las raíces *method*, *model* o *framework*. De esta

manera, se logró combinar diferentes términos atendiendo los criterios de búsqueda especificados.

A partir de lo planteado, como estrategia de búsqueda para explorar el concepto de análisis de datos, se empleó la instrucción 1; igualmente, para revisar el estado del arte de los modelos de evaluación de madurez de análisis de datos se utilizó la instrucción 2, tal como se detalla en la Figura 1-1.

Figura 1-1: Estrategias para la búsqueda para la definición de capacidad de análisis de datos

Estrategia de búsqueda 1. Fórmula para la definición de capacidad de análisis de datos

TITLE-ABS-KEY (("data analytics") AND (definition))

Estrategia de búsqueda 2. Fórmula para los modelos de análisis de datos

TITLE-ABS-KEY (("data analytics capability" OR "data analytics capabilities") AND

Fuente: elaboración propia a partir de Kitchenham (2004).

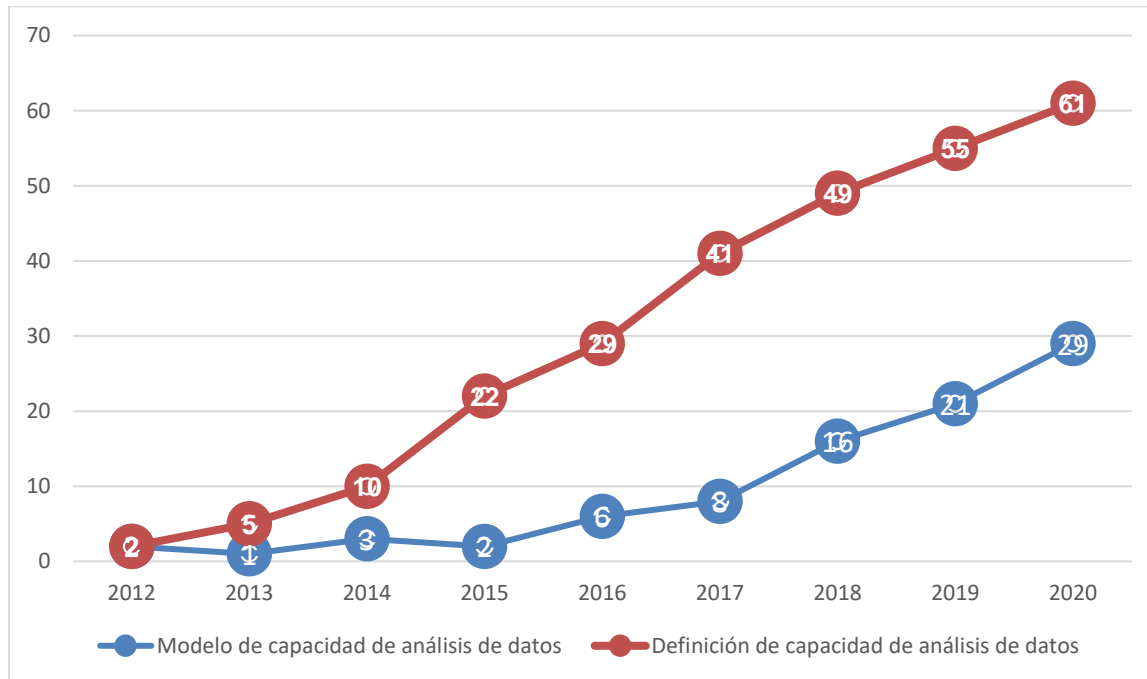
Asimismo, como criterios de calidad para la selección de estudios relacionados con modelos de evaluación de madurez, asociados con el análisis de datos, se definieron: 1) el grado de explicación que el estudio ofrece en cuanto al modelo, 2) el sesgo que pueda contener el estudio en cuanto a la obtención, análisis y exposición de resultados, y 3) los métodos de validación empleados para determinar la robustez del modelo creado. Lo anterior, atendiendo el criterio general de minimización del sesgo y maximización de validez interna y externa expuesto por Alderson et al. (2004) y Khan et al. (2001).

1.1.2 Búsqueda de información y análisis preliminar

Con base en la estructura general de las estrategias señaladas, se obtuvo un total de 237 resultados para la definición de capacidad de análisis de datos (Instrucción 1) y 73 documentos para los modelos (Instrucción 2). La Figura 1-2 resume la

cantidad de documentos encontrados en esta temática, por año, en la base de datos Scopus.

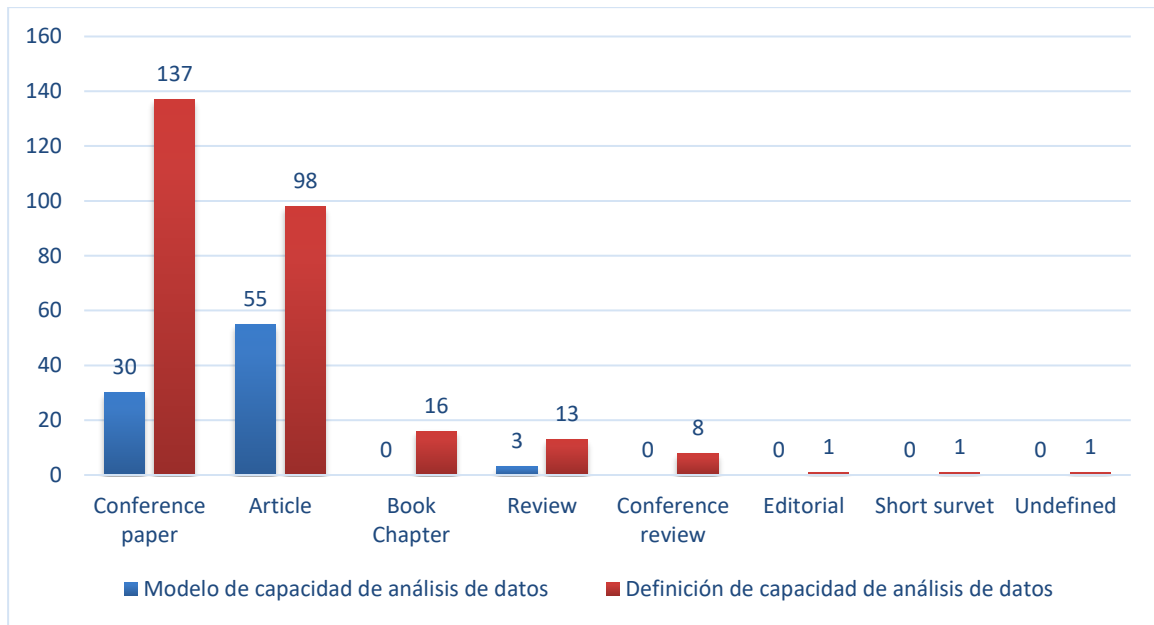
Figura 1-2: Número de documentos encontrados en Scopus según ecuaciones de búsqueda



Fuente: elaboración propia

Se identifica de los resultados que, a partir del año 2015, hubo un incremento importante en las publicaciones relacionadas con la definición de capacidad de análisis, hasta llegar a un pico en 2019; en cuanto a las cifras del 2020, recientemente actualizadas, se observa igualmente un aumento en los documentos, manteniendo la tendencia positiva. Por su parte, las investigaciones asociadas con los modelos comienzan a mostrar un aumento posterior al año 2017; sin embargo, estas son mucho menor con respecto al concepto de capacidad de análisis. Así mismo, se muestra en la Figura 1-3 la clasificación de las publicaciones según el tipo de documento.

Figura 1-3: Tipo de documentos en Scopus según ecuaciones de búsqueda



Fuente: elaboración propia

De acuerdo con lo mostrado, la definición de capacidad de análisis de datos se observa con mayor fuerza en conferencias, seguidos de artículos, lo cual permite inferir que la comunidad científica genera mayor debate en estos espacios, por otro lado, los modelos son desarrollados mayormente en artículos. En este orden de ideas, a continuación, se exponen las posturas teóricas de estas dos variables identificadas en la literatura para medir la madurez de la aplicación de las técnicas de análisis de datos en las organizaciones, dando respuesta a las preguntas orientadoras previamente definidas.

1.2 Analítica de datos

En primera instancia, es importante establecer una definición general de lo que se conoce como dato, entendiendo que este puede ser asumido como el insumo principal para generar información que, posteriormente, se convierta en conocimiento (Dulcé, 2016). Por ello, para que esta relación sea exitosa debe contar con un conjunto de características básicas para que pueda ser aprovechable

en otro tipo de procesos analíticos más avanzados, entre estas se destacan: validez, fiabilidad y accesibilidad (Ye, 2017).

Por otra parte, los datos pueden dividirse en cualitativos y cuantitativos, siendo los primeros aquellos que no están expresados en una escala de medida específica y pueden representar opiniones, percepciones o reflexiones, mientras que los segundos son una expresión numérica de un hecho en concreto, teniendo la capacidad para ser tabulado y procesado (Joyanes, 2013). Realizada esta aclaración, resulta relevante explorar la evolución de los conceptos asociados al análisis de datos durante los últimos años, siguiendo los postulados informados en la Tabla 1-1.

Tabla 1-1: Análisis de datos

Autor	Perspectiva y concepto
Luhn (1958)	Sistema automatizado para difundir información relevante a cualquier punto de la organización, sea industrial, científica o gubernamental fundamentada en procesos que absorben y codifican documentos, a fin de crear perfiles para cada uno de los interesados.
Tukey y Wilk (1966)	Conjunto de procesos estadísticos avanzados para transformar datos en información, apoyado en las ciencias matemáticas y computacionales que viene dado por la experimentación y observación del entorno.
Meyer (1975)	Aplicación de técnicas estadísticas avanzadas para el procesamiento y extracción de información para generar conocimiento científico.
Bryman y Cramer (1992)	Comprensión de comportamientos sociales empleando software especializado como SPSS

Autor	Perspectiva y concepto
	fundamentado en técnicas estadísticas y datos estructurados.
Ramsay (2006)	Modelos funcionales que pueden ser utilizados para mejorar los métodos estadísticos tradicionales con el propósito de fortalecer las estimaciones.
Bihani y Patil (2014)	Proceso que contribuye a la mejora en la toma de decisiones y el éxito en los negocios a través de técnicas de predicción, segmentación y clasificación.
Nagappan et al. (2019)	Proceso de extracción de información útil de amplias bases de datos con fines de explorar y comprender los patrones subyacentes.
Hicks y Peng (2019)	Mecanismo formal para evaluar los datos basados en características objetivas en donde confluyen diversas tendencias de las tecnologías emergentes enmarcadas en el paradigma de ciencia de datos.
Rutkowski, Jaworski y Duda (2020)	Desarrollo de herramientas y métodos para explorar el conocimiento tomando grandes conjuntos de datos, asimismo, se relaciona con la estadística, reconocimiento de patrones, aprendizaje automático, redes neuronales, inteligencia computacional y redes bayesianas.
Ramesh et al. (2020)	Proceso de soporte a los Sistemas de Apoyo a las Decisiones (DSS, por sus siglas en inglés) que permite extraer, transformar, procesar, analizar y visualizar datos que muestren información clave del negocio u organización.

Fuente: elaboración propia con base en Luhn (1958), Tukey y Wilk (1966), Meyer (1975), Bryman y Cramer (1992), Ramsay (2006), Bihani y Patil (2014), Nagappan et al. (2019), Hicks y Peng (2019), Rutkowski et al. (2020), Ramesh et al. (2020)

Con base en lo expuesto en la Tabla 1-1 se puede mencionar que las concepciones de análisis de datos siempre han estado presentes desde los inicios de la ciencia estadística; sin embargo, la capacidad para procesar de forma automática grandes cantidades se potenció con la llegada de los computadores, en las últimas décadas del siglo XX (Eason, 1989). Se observa como Luhn (1958), alto ejecutivo de la naciente compañía International Business Machines (IBM), mostró un proceso analítico que transformaba datos en información coherente según las necesidades organizacionales, denominándolo *sistema de inteligencia de negocios*.

Posteriormente, uno de los estadistas de mayor importancia en Estados Unidos creó técnicas más avanzadas para el análisis de datos, incorporando algoritmos como la Transformada Rápida de Fourier para comprender series temporales o el Análisis Exploratorio de Datos (EDA, por sus siglas en inglés) donde expone el novedoso diagrama de caja y bigotes para mostrar comportamientos, entre otros métodos que hoy en día son ampliamente utilizados (Tukey & Wilk, 1966).

En este sentido, se observa cómo, en los primeros años, el análisis de datos estuvo vinculado directamente a la estadística y su uso se orientó a la resolución de problemas del sector productivo (Meyer, 1975). Con el auge de las ciencias computacionales, la tecnología y la llegada de internet, la cantidad de datos generados en la primera década del siglo XXI comienza a incrementarse notablemente, por ello, el análisis de datos inicia una transformación hacia un uso más extendido, sobre todo en el campo de las ciencias sociales y humanas (Ramsay, 2006).

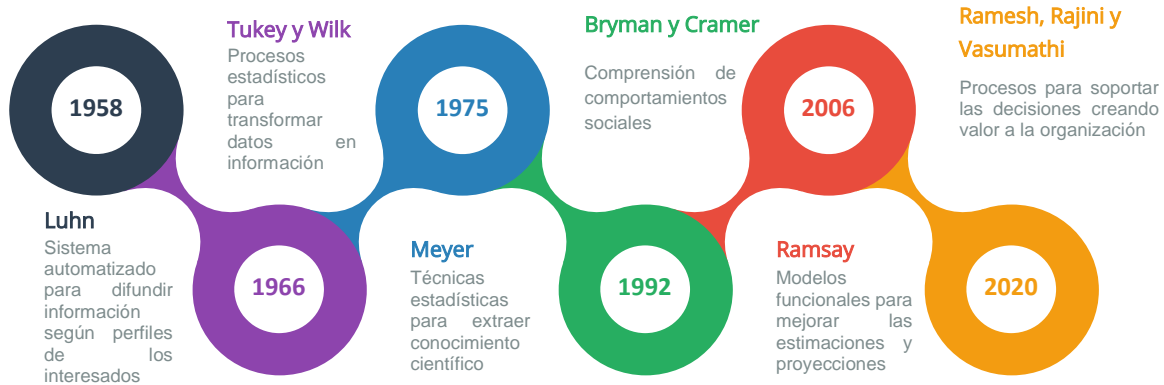
Con este avance, los datos dejan de ser un insumo netamente numérico (cuantitativo) sino que puede ser extraído de forma cualitativa a través de análisis de imágenes, discursos, sensibilidad, etc., todos estos generados desde el ambiente virtual que proporciona el internet, implicando el comienzo de un cambio de paradigma cuyo eje central es dar mayor importancia a las distintas fuentes,

sean estas estructuradas (bases de datos, archivos planos) o no estructuradas (Maxwell & Chmiel, 2014). De acuerdo con Bihani y Patil (2014) las técnicas de análisis de datos están dirigidas a satisfacer cinco grandes áreas: a) exploración; b) retrospectiva; c) predicción; e) toma de decisiones en tiempo real; y, f) transparencia y control.

Así las cosas, se identifica una ampliación del concepto de análisis de datos en donde el objeto no solamente se orienta a resolver problemas del sector productivo o científico utilizando fuentes estructuradas y mediciones precisas, sino que se inicia su aplicación con más fuerza hacia el sector comercial y negocios con el propósito de tomar mejores decisiones (Du et al., 2019). Por ello, hoy en día, la analítica es considerada una herramienta fundamental para comprender el entorno desde sus diferentes dimensiones y *descubrir elementos de valor*, lo que contribuye a decidir la dirección más apropiada de las estrategias, los fenómenos que pueden influenciar en el comportamiento humano y la predicción de eventos fortuitos (Islam, 2020).

En los últimos años, estas técnicas que se derivan de la estadística y han sido potencializadas por las ciencias de la computación, se han constituido en nuevos constructos, tales como la minería de datos, que combinan algoritmos avanzados de aprendizaje automático (*Machine Learning*), inteligencia artificial (IA) y la estadística tradicional, aunque en muchos casos partan de procedimientos matemáticos tradicionales como el *clustering* (segmentación) o análisis de regresión (predicción) (Tan et al., 2006).

En la perspectiva de Hicks y Peng (2019), el análisis de datos ha venido trascendiendo la simple aplicación de técnicas estadísticas, en la medida en que comprende la confluencia de distintas ramas del saber, entre estas la computación y estadística, configurando un área donde aún hay poco consenso sobre lo que realmente significa y abarca. La siguiente infografía muestra un resumen de la evolución del concepto en los últimos años (Figura 1-4).

Figura 1-4: evolución del concepto de análisis de datos

Fuente: elaboración propia con base en Luhn (1958), Tukey y Wilk (1966), Meyer (1975), Bryman y Cramer (1992), Ramsay (2006), Ramesh et al. (2020)

En consecuencia, para el presente trabajo final de maestría, **el análisis de datos se define como el conjunto de operaciones, técnicas y procedimientos que combinan la estadística, computación y matemáticas para extraer información a partir de datos estructurados y no estructurados en una organización** (Martínez, 2014). A continuación, el siguiente apartado expone una revisión de los conceptos asociados con la capacidad de análisis de datos contrastando la postura de diferentes autores en esta materia.

1.3 Capacidad de análisis de datos

En la sección anterior se identificaron algunos conceptos de los autores a lo largo de los últimos años sobre lo que significa el análisis de datos, destacando aspectos clave que han marcado la evolución de su finalidad y propósito; de esta manera, ahora es necesario comprender el concepto de capacidad de análisis, por ello la Tabla 1-2 muestra un breve recorrido histórico.

Tabla 1-2: Capacidad de análisis de datos

Autor y año	Concepto
Denning, et al. (1984)	Se asocia con la capacidad de cómputo de las bases de datos, los operadores utilizados y la determinación de los mejores procesos analíticos.
Irwin (2008)	Combinación de técnicas y métodos para extraer, transformar y procesar datos desde distintos orígenes y formatos.
Marulanda et al. (2013)	Capacidad de las organizaciones para aplicar técnicas de minería de datos para construir procesos de generación de conocimiento en las organizaciones.
Bedeley y Nemati (2014)	Capacidad que está determinada por tres recursos clave: infraestructura de TI, capital humano y arquitectura empresarial.
Bonilla y Rao (2015)	Capacidad de ensamblar, coordinar, movilizar y desplegar recursos basados en el análisis con un propósito estratégico.
Shuradze y Wagner (2016)	Consiste en la coordinación de las dimensiones de infraestructura, experiencia del personal, gestión e integración para el logro de la innovación y agilidad organizacional.
Garmaki y Boughzala (2016)	Capacidad para aprovechar los datos aplicándolo a la toma de decisiones con miras a mejorar el desempeño de la empresa.
Gupta y George (2016)	Competencias de la organización para aprovechar los datos y convertirlos en activos intangibles para mejorar su rendimiento.
Wamba et al. (2017)	Habilidad para adaptarse a los cambios del entorno (capacidades dinámicas) mediante la transformación

Autor y año	Concepto
	de procesos analíticos con el propósito de incrementar el rendimiento empresarial.
Kamioka et al. (2017)	Competencias de los usuarios para utilizar la TI y el potencial para procesar, analizar, así como ofrecer información de valor para la organización.
Ramamurthy (2017)	Integración de procesos organizacionales para extraer información útil que permita optimizar el rendimiento de la empresa, empleando técnicas avanzadas y datalakes (repositorios de datos).
Mikalef et al. (2018)	Capacidad de una empresa para capturar y analizar datos con el fin de generar conocimientos mediante la orquestación y el despliegue de sus datos, tecnología y talento.
Bean y Davenport (2019)	Transformación digital de las organizaciones para tratar los datos como activos intangibles con el fin de tomar decisiones acertadas y coherentes con el entorno global.
Partners New Vantage (2019)	Disposición de las organizaciones para invertir en tecnología disruptiva (Machine Learning, Cloud Computing, Blockchain, FinTech Solutions) a fin de crear soluciones de negocio.
Hao et al. (2019)	Técnicas, métodos y herramientas que se implementan en la organización para identificar oportunidades de mercado, predecir necesidades y analizar decisiones de los clientes por medio de la utilización de grandes cantidades de datos.
Xiao et al. (2020)	Capacidades dinámicas del personal e infraestructura de TI que actúan en conjunto para alcanzar objetivos de innovación en los servicios que se prestan.

Autor y año	Concepto
Yasmin et al. (2020)	Las capacidades de análisis de datos son interdependientes (infraestructura tecnológica, recurso humano, gestión), además, están relacionadas con el desempeño empresarial con más fuerza que con el rendimiento de mercado,
Song, et al. (2020)	Explora la capacidad de análisis en la innovación de servicios; el estudio asume que es un factor clave para incrementar este aspecto en la empresa.
Mikalef et al. (2020)	Conjunto de procedimientos de TI que permiten fortalecer considerablemente la competitividad empresarial, capacidad tecnológica y comercial.

Fuente: elaboración propia con base en Marulanda et al. (2013), Denning et al. (1984), Irwin (2008), Bedeley y Nematí (2014), Bonilla y Rao (2015), Shuradze y Wagner (2016), Garmaki y Boughzala (2016), Gupta y George (2016), Wamba et al. (2017), Kamioka et al. (2017), Ramamurthy (2017), Mikalef et al. (2018), Bean y Davenport (2019), Partners New Vantage (2019), Xiao et al. (2020), Hao et al. (2019), Yasmin et al. (2020), Song et al. (2020), Mikalef et al. (2020)

En este sentido, se observa en primer lugar que, desde las primeras décadas del siglo XXI, la preocupación e interés por contar con una infraestructura tecnológica capaz de capitalizar los datos en información ha sido un reto importante para las empresas, sobre todo en aquellas que aún no ven claridad en los beneficios que esta tendencia puede reportarle.

De acuerdo con Gupta y George (2016), al convertir los datos en un activo que genera valor comienza un cambio de la cultura organizacional donde los procesos se orientan a la explotación de la analítica como un recurso clave para mejorar el desempeño empresarial. Igualmente, Wamba et al. (2017) consideran que estas capacidades son dinámicas, es decir, cambian constantemente en función de las exigencias del entorno; por ello, componentes como la tecnología, el factor humano

y los procesos se convierten los elementos fundamentales para lograr con éxito las políticas de análisis de datos.

En esta misma línea, Kamioka et al. (2017) consideran que la capacidad de análisis viene determinada por la formación y competencias del personal a cargo de la TI en la organización, mientras que Marulanda et al. (2013) asimilan más este concepto a la aplicación de técnicas de minería de datos y la estandarización de procedimientos en este aspecto. Con una visión más holística, los expertos de gran reconocimiento en esta materia, Bean y Davenport (2019), han expuesto reflexiones interesantes sobre este tema, indicando que las corporaciones líderes a nivel mundial presentan dificultades para convertirse en empresas orientadas a los datos, en tanto, específicamente, se basan en los resultados del estudio adelantado por Partners New Vantage (2019) donde se informa que el 72% aun no consolida una cultura de datos y el 53% no considera este insumo como un activo comercial.

Fundamentado en estos hallazgos, resulta más claro afirmar que la capacidad de análisis de datos no sólo significa la implementación de tecnología y software, sino que requiere un cambio en el paradigma de la organización, en donde el recurso humano desempeña un factor clave, pues las competencias analíticas deben estar presentes en la mayoría del personal, independientemente de su actividad dentro de la empresa.

Lo anterior se sustenta en referentes del concepto “capacidad de análisis de datos”, a partir del cual Bean y Davenport (2019) ofrecen una perspectiva moderna y más integral sobre lo que significa este tema para las organizaciones. Los autores señalan que los datos son activos que pueden generar valor agregado al proceso de toma de decisiones, para lo que debe ocurrir un cambio tecnológico profundo en la organización, en aspectos formativos del personal y en el compromiso de los líderes frente a dichos avances.

En definitiva, con base en todas las reflexiones anteriores para este trabajo final de maestría, la capacidad de análisis de datos se entiende como con el conjunto de herramientas, métodos, competencias y habilidades con los que cuenta una organización para extraer, transformar y procesar articuladamente datos provenientes desde diferentes fuentes, sean estas estructuradas o sin estructura, con el propósito de obtener valor, traducida como información útil que apoye la toma de decisiones (Letón & Pedromingo, 2001).

A continuación, se presenta una comparativa de los modelos reseñados por la literatura para medir la capacidad de análisis de datos en las organizaciones.

1.4 Modelos para medir la capacidad de análisis de datos

Los modelos para medir la capacidad de análisis de datos pueden definirse como marcos integrales que describen, explican y evalúan el crecimiento de los ciclos de vida de una determinada área de la organización (Shaaban et al., 2011). En este sentido, el eje central de estos es mostrar la situación actual de una entidad con relación a un aspecto, en este caso, la capacidad de análisis de datos. Al realizar una revisión de la literatura se observó que los modelos cuentan con diversos enfoques a tomar en cuenta para el diagnóstico; algunos poseen dimensiones, variables, indicadores o instrumentos, mientras que otros no; no obstante, existen elementos transversales tales que incluyen la infraestructura de TI, gestión, liderazgo, competencias y habilidades del recurso humano, entre otros (Kumar et al., 2017).

En términos de la evolución, los modelos para medir la capacidad de análisis de datos se impulsaron, inicialmente, desde el campo del desarrollo de software (Humphrey, 1989) con el fin de controlar las fases ya completadas, así como la gestión de cambios necesarios para culminarlas, en función de un objetivo a largo plazo. Por lo tanto, los modelos de capacidad de análisis de datos se fundamentan

en observar los cambios que ocurren a través del tiempo, según una serie de etapas que contienen características específicas y ciclos de vida propios, ayudando así a delimitar el momento en el que se está preparado para avanzar a otros niveles de mayor complejidad (Prieto et al., 2015).

Dicho lo anterior, los autores Cates et al. (2005), en su modelo denominado *Ladder of Business Intelligence* (LOBI), se enfocan en la efectividad y eficiencia del proceso de toma de decisiones; además, se fundamenta en la construcción de un plan sólido para la arquitectura de TI, la definición de un cuadro de mando integral, la asignación de roles y procesos de inteligencia de negocio clave para la organización, todos estos aspectos articulados para alcanzar un escenario de mayor efectividad.

Seguidamente, el modelo de Davenport y Harris (2017), denominado *Delta Model*, indica que las organizaciones deben medir la capacidad analítica en tres ejes fundamentales, a saber: a) estructura organizativa; b) gestión humana y c) tecnología; por consiguiente, definen cinco etapas de madurez que pueden alcanzarse, indicando que aquellas que no cuentan con este tipo de procesos tienen una visión limitada de los clientes, mercado y competidores utilizando el instinto para sus decisiones.

En segundo lugar, se encuentran las que tienen procesos incipientes y localizados, pero que aún están desconectados del resto de la entidad; en tercera instancia, las que tienen aspiraciones analíticas con la construcción de métricas e indicadores; en cuarto puesto, las compañías analíticas tienen el respaldo de la directiva y en todos los niveles; mientras que en quinto lugar se posicionan las que aplican el liderazgo basado en datos con una amplia arquitectura de TI.

De forma similar, el *Business Information Maturity Model* (Ladley, 2010) se concentra en la importancia del BI definiendo tres factores clave: a) alineación y

gobierno de datos; b) aprovechamiento y c) entrega. En cuanto a la primera, se define como la estructura de políticas que sustenta la analítica y está en concordancia con los objetivos organizacionales, mientras que la segunda (b) hace referencia extraer información útil de los procesos de datos; así, el tercer elemento (c) aclara que se debe contar con una distribución que garantice el acceso al conocimiento.

En función de lo anterior, el modelo cuenta con un primer nivel en donde la información no se encuentra estructurada, en consecuencia, los almacenes de datos (*Data Warehouse*) vienen a solventar estas falencias ofreciendo acceso y rapidez de cálculo. Por su parte, en el segundo nivel se define más formalmente las necesidades de información del negocio y se evalúa la inversión en analítica, avanzando hacia preguntas más profundas como por qué se requiere dicho conocimiento, a quién va dirigida, entre otras. Por último, en el tercer nivel, se establece el modelo de BI donde todas las partes de la compañía se encuentran articuladas y se explota como activo para la toma de decisiones de todos los miembros.

De otro lado, el Modelo de Madurez *Service Oriented Business Intelligence* (SOBI) (Shaaban et al., 2011) se sustenta en cinco niveles de madurez que declaran la evolución de la organización sobre dos ejes clave: la orientación al servicio y la infraestructura de BI. En cuanto al primer punto, se destaca que la integración de los procesos analíticos permite una rápida respuesta a las necesidades de los clientes, fortaleciendo la capacidad operativa; asimismo, la estructura de BI facilita el procesamiento de grandes volúmenes de datos, proporciona un marco fiable para la toma de decisiones y genera buenas prácticas para la gestión de los datos para responder aspectos críticos del negocio.

Con relación al *Business Intelligence Maturity Model* (BiMM) (Dinter, 2012) cabe señalar que se esfuerza por garantizar la implementación del BI en todos los niveles jerárquicos de la compañía, observando cada una de las etapas de desarrollo

describiendo en detalle las acciones, estrategias y procesos que se deben ejecutar para alcanzar la madurez, a diferencia de otros modelos que no realizan este tipo de acompañamiento, sino que se basan en el diagnóstico. En este orden de ideas, dicho modelo no determina qué nivel de madurez es el óptimo para cada organización, dado que éste se define de acuerdo con los objetivos que persigue la organización.

Así las cosas, el *Gartner's Maturity Model for Business Intelligence and Performance Management* (Olszak, 2013) evalúa tres dimensiones críticas relacionadas con el personal, procesos, así como las métricas y tecnología. Adicionalmente, establece cinco niveles de madurez, en el cual el más bajo corresponde al desconocimiento, caracterizado por la existencia de datos inconexos con gran cantidad de errores que son difíciles de aprovechar. En segundo lugar, el nivel táctico implica pequeñas inversiones en analítica, sin embargo, los indicadores y reportes se utilizan solo en los departamentos con fines operativos y el personal no cuenta con la suficiente capacitación para extraer conocimiento de valor.

En el tercer nivel de enfoque la organización logra transformar la información en decisiones que hacen crecer el negocio, por lo tanto, se comienza a percibir los beneficios de la BI, mientras que en la cuarta etapa (estratégica) se inicia una articulación con los objetivos a largo plazo de la organización, extendiéndose la analítica de datos a otras áreas externas como la gestión de clientes y proveedores. Posteriormente, en el nivel cinco, el BI desempeña un rol crucial en todos los segmentos de operación, se garantiza un acceso a todos los usuarios y se considera como un activo de la entidad.

Tomando en consideración los anteriores análisis, se presenta la Tabla 1-3, la cual resume algunos de los modelos aplicados en el campo del análisis de datos, inteligencia de negocios y *Big Data*.

Tabla 1-3: Comparación de modelos de medición de capacidad de análisis de datos

Nombre	Autor y año	Descripción	Dimensiones	Instrumento de recolección	Niveles de madurez
<i>Analytics Maturity Model (TDWI)</i>	(Halper, 2020)	<p>Generalidades:</p> <p>Proporciona un marco para que las organizaciones entiendan dónde están, dónde han estado y a dónde tienen que ir todavía en su despliegue analítico.</p> <p>Contexto de aplicación:</p> <p>Este modelo ha sido utilizado en organizaciones diversas especialmente las desarrolladoras de programas de soporte para la transformación digital.</p> <p>Debilidades de este modelo:</p> <p>Algunas mejoras que podrían proponerse para este modelo es la inclusión de dimensiones asociadas con los cambios o imprevistos que quedan por fuera de los 5 elementos de base definidos.</p>	<ul style="list-style-type: none"> ▪ Organización ▪ Infraestructura ▪ Gestión de recursos ▪ Analítica ▪ Gobernanza 	Sí (cuestionario)	<ul style="list-style-type: none"> ▪ Naciente ▪ Pre-adopción ▪ Establecido. ▪ Madura. ▪ Visionaria
<i>Ladder of Business Intelligence (LOBI)</i>	(Cates et al., 2005)	<p>Generalidades:</p> <p>Describe los niveles de madurez en cuanto a la</p>	<ul style="list-style-type: none"> • Hechos • Datos • Información • Comprensión 	No se especifica	No posee

Evaluación de la capacidad de análisis de datos
de las pymes desarrolladoras de software de la ciudad de Bogotá

Nombre	Autor y año	Descripción	Dimensiones	Instrumento de recolección	Niveles de madurez
		<p>efectividad y eficiencia en la toma de decisiones.</p> <p>Contexto de aplicación: Empresas diversas que soportan su operación en tecnologías que se renuevan de forma permanente. Ej.: consumo masivo, software, servicios de salud, entre otros.</p> <p>Debilidades: No facilitó una guía asociadas a la comprensión del modelo lo que pudo dificultar la interacción con el mismo y las dimensiones propuestas.</p>	<ul style="list-style-type: none"> • Disponibilidad • Intuición 		
Modelo Delta	(Davenport & Harris, 2017)	<p>Generalidades: Se sustenta en cinco factores clave de éxito que denotan la capacidad y activos en las necesidades de la organización con el fin de lograr las iniciativas de analítica.</p> <p>Contexto de aplicación: Empresas con enfoque en datos y vida digital.</p> <p>Debilidades: Se consideraron criterios solo para iniciativas de</p>	<ul style="list-style-type: none"> • Datos • Empresa • Liderazgo • Objetivos • Analítica 	Sí (cuestionario)	<ul style="list-style-type: none"> • Analíticamente impedido • Análisis localizados • Aspiraciones analíticas • Empresas analíticas • Analítico Competidor

Nombre	Autor y año	Descripción	Dimensiones	Instrumento de recolección	Niveles de madurez
		analítica, dejando por fuera otros componentes de interés.			
<i>Business Information Maturity Model</i>	(Ladley, 2010)	<p>Generalidades: Este modelo se centra en aumentar la importancia del BI. Se centra en la visión y los cambios culturales en el uso de la información.</p> <p>Contexto: Empresas con procesos establecidos en tecnologías de la información.</p> <p>Debilidades: Solo contempló tres niveles de análisis para la implementación del modelo.</p>	<ul style="list-style-type: none"> • Alineación y gobernanza • Aprovechamiento • Entrega 	Sí (cuestionario)	<ul style="list-style-type: none"> • Nivel 1 • Nivel 2 • Nivel 3
<i>Modelo de Madurez Service Oriented Business Intelligence (SOBI)</i>	(Shaaban et al., 2011)	<p>Generalidades: Combina dos paradigmas de arquitectura que se articulan entre sí, una estructura de BI y servicios.</p> <p>Contexto: Empresas diversas con soporte tecnológico en todos los procesos ya establecidos.</p> <p>Debilidades: Solo contempló tres grandes dimensiones, dejando de</p>	<ul style="list-style-type: none"> • Tecnología • Organización • Experticia en los negocios 	Sí (checklist)	<ul style="list-style-type: none"> • Inicial • Inmaduro • Controlado • Gestionado • Maduro

Nombre	Autor y año	Descripción	Dimensiones	Instrumento de recolección	Niveles de madurez
		lado otros aspectos relacionados con las tipologías de empresas que podrían poner en ejercicio el modelo.			
<i>Business Intelligence Maturity Model (BiMM)</i>	(Dinter, 2012)	<p>Generalidades: Consiste en un sistema multidimensional y estructura jerárquica.</p> <p>Contexto: Empresas con base TIC.</p> <p>Debilidades: Solo abarcó tres dimensiones dejando por fuera algunos componentes que según otros autores serían de alta relevancia.</p>	<ul style="list-style-type: none"> • Funcionalidad • Tecnología • Organización 	Sí (cuestionario)	<ul style="list-style-type: none"> • Individual información • Grupos de la Información • Integración • Inteligencia de información • Empresa y gestión
<i>Gartner's Maturity Model for Business Intelligence and Performance Management (BI & PM)</i>	(Olszak, 2013)	<p>Generalidades: El modelo es un medio para evaluar la madurez de los esfuerzos de una organización en materia de BI y PM y cuán maduros deben llegar a ser para alcanzar los objetivos empresariales.</p> <p>Contexto: Se enfoca a organizaciones BI y PM.</p>	<ul style="list-style-type: none"> • Personal • Procesos • Métricas y tecnología 	Sí (cuestionario)	<ul style="list-style-type: none"> • Sin conocimiento • Táctico • Enfocado • Estratégico • Penetrante

Nombre	Autor y año	Descripción	Dimensiones	Instrumento de recolección	Niveles de madurez
		<p>Debilidades: El nivel de especialidad del modelo podría requerir más capacitación frente a su implementación.</p>			
<p>Metodología Integrada Decisiones Multicriterio (MCDM)</p>	<p>(Yasmin et al., 2020)</p>	<p>Generalidades: Se emplea un análisis multicriterio difuso (FUZZY) combinado con el método para evaluar causa y efecto denominado IF-DEMATEL.</p> <p>Contexto: Organizaciones con base TIC.</p> <p>Debilidades: No se desarrolló un check list o cuestionario guía, así también no propone niveles de madurez.</p>	<ul style="list-style-type: none"> • Infraestructura tecnológica • Capacidades de Gestión • Recursos humanos 	<p>No (Datos previamente registrados)</p>	<p>No posee</p>

Fuente: elaboración propia con base Eckerson (2007), Cates et al. (2005), Davenport y Harris (2017), Ladley (2010), Shaaban et al. (2011), Dinter (2012), Olszak (2013) y Yasmin et al. (2020).

En función de los modelos expuestos en la Tabla 1-3, es preciso mencionar que la información y la capacidad para gestionarla adecuadamente es un factor crítico que se encuentra, transversalmente, en casi todas estas herramientas de diagnóstico. En el caso concreto del modelo TDWI, diseñado por Halper y Stodder (2014) y actualizado en 2020 por Halper (2020), cabe destacar que este permite a las organizaciones saber en dónde se encuentran con respecto a los niveles de madurez, qué han hecho y qué necesitan para seguir avanzando.

Posterior a la revisión de los modelos de madurez señalados en la Tabla 1-3, resulta relevante argumentar que se escoge el modelo TDWI dado que ofrece una perspectiva integral sobre el estado en el que se encuentran las organizaciones, además, la metodología de aplicación e interpretación de resultados ha sido reseñada en investigaciones similares como confiable y válida (Eckerson, 2007). Adicionalmente, las dimensiones que establece contemplan los tres ejes fundamentales identificados en otros marcos, tales como la infraestructura tecnológica, personal y gestión de datos, asumiéndose que este proporciona una mirada completa de los avances en la capacidad de análisis de datos en las empresas desarrolladoras de software.

Asimismo, Hatta et al. (2015) señalan que el modelo TDWI, además de ser ampliamente utilizado y fácil de aplicar, posee menos limitaciones y complicaciones con respecto a otros métodos, dado que se basa en la autovaloración de las dimensiones que posteriormente son estandarizadas con base en un proceso de transformación de puntajes para definir el nivel de madurez.

Por lo tanto, se concluye que el modelo TDWI ofrece una metodología fiable, comprobada y mejor ajustada a los fines de la presente investigación, puesto que permite evaluar la capacidad de análisis de datos tomando en consideración cinco (5) dimensiones: organización, infraestructura, gestión de recursos, análisis y gobernanza. Este modelo provee una metodología coherente para medir y monitorear el estado de la analítica en la organización, indicando así los posibles

mecanismos, acciones y estrategias que se deben tomar para desarrollar una cultura basada en los datos. El modelo TDWI es explicado con mayor detalle en el Capítulo II.

1.5 Resumen del capítulo

En el Capítulo I se elaboró un marco de referencia que abordó nociones conceptuales sobre el análisis de datos y la capacidad de las organizaciones para aprovechar estos datos. Asimismo, se realizó una revisión de la literatura relacionada, con el fin de identificar y analizar los modelos de evaluación de madurez creados, así como sus dimensiones, elementos y niveles (categorías/etapas) de madurez.

Posterior a lo enunciado, se determinó que el modelo de evaluación de madurez a utilizar en el presente estudio corresponde al modelo TDWI; esto, producto del análisis comparativo realizado con los modelos relevantes identificados en la revisión literaria. Lo anterior, permitió dar cumplimiento a los objetivos específicos 1 y 2 de la presente investigación, los cuales consistieron en caracterizar los modelos que más contribuyen al desarrollo de la temática bajo estudio y determinar el modelo de madurez en análisis de datos a aplicar en pymes desarrolladoras de software de la ciudad de Bogotá, respectivamente.

Cabe resaltar que la revisión de literatura mostró una amplia variedad de modelos para evaluar la madurez de la aplicación del *Big Data* y BI en las organizaciones. De esta manera, como aspectos relevantes a destacar, se resalta que cada uno de estos modelos, independientemente de sus particularidades y características, apuntan a un solo objetivo: diagnosticar el estado actual de las organizaciones e identificar qué acciones concretas pueden ser adelantadas para avanzar en esta materia. Asimismo, producto de analizar a profundidad las distintas conceptualizaciones de análisis de datos y lo que significa la capacidad para procesarlos, se concluye que las variables tecnológicas (infraestructura), de

personal (formación, competencias y habilidades) y de liderazgo (cambios en la cultura organizacional) son elementos cruciales para el éxito de estas iniciativas.

2. Ajuste del modelo de evaluación de madurez TDWI para su aplicación

El presente capítulo tiene como objetivo describir en forma detallada el modelo de evaluación de madurez de análisis de datos a aplicar en pymes desarrolladoras de software de la ciudad de Bogotá: el modelo TDWI. Lo anterior se realiza tomando como referencia la documentación oficial de dicho modelo. Además, se elabora un ajuste del instrumento de recolección de la información con el fin de adecuarlo al idioma español Latinoamérica y, así poderlo aplicar a la población bajo estudio sin el temor de experimentar posibles limitaciones con relación al idioma original del instrumento (inglés) al momento de su aplicación.

2.1 El Modelo de TDWI (Transforming Data With Intelligence™)¹

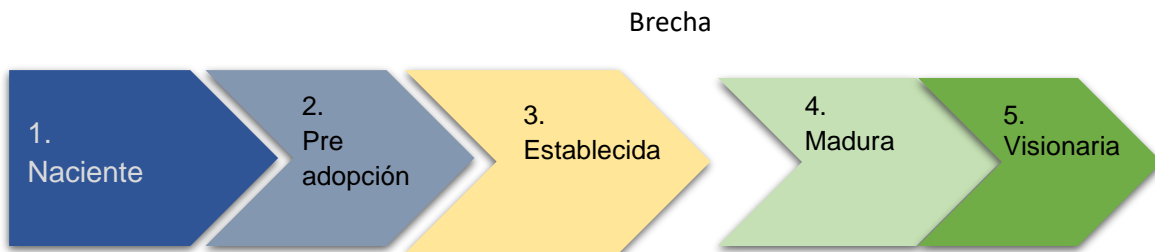
Con la apertura del *Big Data* y el avance de sus técnicas potenciadas por las ciencias de la computación, muchas compañías han venido interesándose cada vez más por este tipo de soluciones de *Big Data*, IA y BI para apoyar su proceso de toma de decisiones; no obstante, existen brechas importantes que aún las mantienen fuera de una transformación digital sostenible.

En virtud de lo citado, el modelo TDWI busca ofrecer una solución para medir el nivel de madurez de las organizaciones, en términos de análisis de datos, a través

¹ Apartado elaborado con base en Halper & Stodder (Halper & Stodder, 2014) y Halper (Halper, 2020)

de una escala compuesta por cinco niveles o categorías de madurez, de menor a mayor: naciente, pre-adopción, establecido, madura y visionaria; de tal manera que los anteriores niveles de madurez se calculan con base en cinco dimensiones: organización, infraestructura, gestión de recursos, análisis y gobernanza. La Figura 2-1 ilustra cada uno de los detalles previamente mencionados.

Figura 2-1: Niveles de madurez TDWI en la capacidad de análisis de datos



Fuente: Halper (2020)

Con relación al instrumento de recolección de información provisto por la compañía TDWI, este consta de 53 preguntas de selección múltiple con única respuesta, el cual busca recabar información de calificación a las dimensiones de madurez para, posteriormente, emitir una calificación global que permita ubicar a las organizaciones en un determinado nivel; de esta manera, el ajuste-traducción realizado a dicho instrumento se consigna en el Anexo B. A continuación, se describen cada uno de los niveles y dimensiones de madurez empleados en el modelo TDWI.

2.1.1 Niveles de madurez

Naciente: en el estado naciente, la organización no tiene la conciencia sobre la utilidad del análisis de datos y el valor que puede reportar al negocio, por lo tanto, apenas se manejan hojas de cálculo que sirven como soporte a las transacciones, mas no son herramientas para la toma de decisiones. En cuanto a la gestión de datos, la estrategia de TI aún no se encuentra definida, debido a esto la calidad y consistencia de los datos son bajos, así como el volumen que se maneja. Por su

parte, en la analítica se observa un esfuerzo por aprovechar los datos con modelos, sin embargo, solo se emplean hojas de cálculo como principal método de análisis y solo con fines de reporte a los interesados puntuales de la organización.

Pre-adopción: en el segundo nivel de madurez, la gerencia comienza a dar los primeros pasos sobre los elementos que debe cubrir para implementar la analítica. En este sentido, se estimula al personal a investigar, formarse y leer en torno al tema para comenzar el proceso de familiarización. En este estado, la inversión es incipiente y puede realizarse en sistemas relativamente accesibles que utilicen lenguajes como Python o R, abiertos y ampliamente utilizados en el mercado. En el ámbito organizacional, se inicia la conformación de grupos de experimentación con el propósito de descubrir los problemas de negocio que pueden ser resueltos con plataformas de *Big Data*.

Establecido: a este tercer nivel de madurez se llega cuando el análisis de datos ha sido establecido en la organización. En éste se llevan a cabo pruebas de concepto para comprobar la viabilidad de proyectos. Por ello, en dicha transición se cubre un amplio espacio de tiempo, tratando de observar las limitaciones y dificultades del proceso a fin de corregir los elementos que puedan estar fallando y que puedan atentar contra la implementación de la analítica. En función de esto, otras áreas de la organización se involucran con mayor compromiso y se elaboran estrategias conjuntas para todos los departamentos.

Madura: al cruzar la brecha entre las pruebas y la puesta en marcha real, se consigue el nivel de madurez. En este nivel los usuarios finales interactúan formalmente con los sistemas de información, por lo tanto, obtienen conocimientos que trasladan a los procesos de negocio de los cuales son responsables, logrando así un impacto positivo en la toma de decisiones. Con relación a la organización, los líderes comienzan a comprender que la analítica es un factor clave para lograr la competitividad y el crecimiento; de este modo, colocan la innovación en el

análisis como una estrategia empresarial y están más dispuestos a invertir en la tecnología necesaria para asegurar este elemento.

Visionaria: en el quinto y mayor nivel de madurez se ubican aquellas organizaciones visionarias y maduras, en términos de análisis de datos. Dichas organizaciones se caracterizan por la adopción de una cultura organizacional basada en los datos y una política de datos claramente definida e implementada, en donde los datos son accesibles por las diferentes partes que integran la organización y no solamente por el departamento de TI. No obstante, llegar a este grado resulta complejo puesto que requiere inversiones significativas y procesos de analítica más profundos.

De otra parte, y con base en las definiciones expuestas, los puntajes definidos para cada nivel de madurez son mostrados en la Tabla 2-1.

Tabla 2-1: Puntajes del modelo TDWI para medir la capacidad de análisis de datos

Nivel	Puntaje
Incipiente	<= 5
Pre-adopción	6 – 10
Establecido	11 – 15
Madura	16 – 19
Visionaria	20

Fuente: Elaboración propia con base en Halper (2020)

De esta manera, los parámetros definidos previamente servirán para realizar una identificación objetiva de los hallazgos identificados.

2.1.2 Dimensiones de madurez

Además de los niveles de madurez previamente definidos, el modelo TDWI considera cinco dimensiones para comprender qué sucede con algunas áreas de la empresa: organización, infraestructura, gestión de recursos, analítica y

gobernanza. Estas dimensiones sirven de categorías de agrupamiento que engloban, para el caso de la dimensión *Organización*, aspectos tales como liderazgo, financiamiento, estrategia, cultura y valor. La Figura 2-2 describe cada una de estas dimensiones junto con sus aspectos asociados.

Figura 2-2: Dimensiones del modelo TDWI



Fuente: Halper (2020)

Al respecto de la imagen anterior, a continuación, cada una de las dimensiones se amplía con mayor detalle.

- **Organización:** en esta dimensión se examina el liderazgo, el financiamiento disponible, la estrategia, la cultura y el valor que tienen los datos para la toma de decisiones; por tal motivo, se orienta a resolver preguntas clave como ¿en qué medida la estrategia organizativa, la cultura, el liderazgo y las habilidades soportan un programa de analítica exitoso? además, ¿está la compañía comprometida con lograr este éxito?, ¿se han extendido los análisis a toda la organización y se usan en las decisiones diarias?, entre otros interrogantes.

-
- **Infraestructura:** la dimensión infraestructura hace referencia al software y hardware sobre el cual se sustenta la analítica en la empresa. Por ello, los interrogantes fundamentales de este apartado son los siguientes: ¿hasta qué punto la infraestructura apoya el análisis de todas las partes de la empresa y usuarios potenciales?, ¿qué tecnologías existen para apoyar una iniciativa de análisis? y ¿cómo están integradas estas tecnologías en el entorno existente? Asimismo, dado que tareas como la incorporación de tecnología de soporte, actualización de software o adecuación de las bases de datos (entre otras labores) deben ser realizadas por profesionales calificados, se evalúan aspectos como las habilidades de los empleados y la capacitación que la organización les brinda en términos de TI.
 - **Gestión de recursos:** en esta dimensión se consideran aspectos del recurso humano, financiación, y accesibilidad, volumen y frecuencia de los datos disponibles para la analítica; considerando, además, aspectos críticos en términos de procesamiento, integración y acceso a los datos en todos los niveles de la organización. Asimismo, se evalúan aspectos relativos a los procesos utilizados para extraer los datos, al igual que el fomento de una cultura de toma de decisiones basada en datos dentro de la organización.
 - **Análisis:** en la dimensión análisis se evalúan aspectos relativos al procesamiento de datos y cómo estos son utilizados por la organización, así como el tipo de análisis que son aplicados a dichos elementos. Por ejemplo: modelos analíticos en tiempo real, modelos de *Machine learning*, *Deep learning*, etc. Y preguntas como ¿cuál es el alcance de la analítica dentro de la organización?, ¿cuál es el porcentaje de personas dentro de la organización que ejecuta sus propias tareas de analítica?, ¿están los procesos de analítica incluidos dentro de los procesos de negocio?, entre otras interrogantes.

- **Gobernanza:** finalmente, esta última dimensión está relacionada con las políticas organizacionales que rigen los procesos y estándares relacionados con la extracción, transformación, procesamiento y análisis de datos. Estas políticas deben ser coherentes con las estrategias de negocio y ofrecer la suficiente flexibilidad para que la analítica sea desarrollada en la compañía sin más restricciones que las relacionadas con la seguridad y privacidad en el tratamiento de los datos. De igual modo, se evalúan aspectos de establecimiento de estándares y protocolos para el procesamiento de los datos.

2.2 Resumen del capítulo

En el anterior, se ajustó el modelo TDWI al idioma español y sus componentes principales, tales como: niveles, dimensiones y aspectos de madurez fueron descritos para una correcta comprensión del modelo por parte del lector. Asimismo, el instrumento de recolección de información que se aplicará para recabar los datos insumo del modelo fue consignado como elemento Anexo A. De esta manera se dio cumplimiento al objetivo específico número 3, el cual consistió en ajustar el modelo TDWI para su aplicación en Pymes desarrolladoras de software de la ciudad de Bogotá.

3. Aplicación del modelo TDWI

El presente capítulo tiene como objetivo reseñar la capacidad de análisis de datos de las pymes analizadas según el modelo TDWI. Por ello, se exponen de manera descriptiva cada una de las dimensiones abordadas y el alcance que obtuvieron las organizaciones en materia del procesamiento de datos. Con base en estos hallazgos, se elaboran posteriormente las respectivas conclusiones y recomendaciones.

3.1 Método

Como fue comentado en capítulos previos, se toma como referencia el modelo TDWI; no obstante, cabe aclarar que este se toma en su más reciente versión publicada en el año 2020 (Halper, 2020); de tal forma que se evalúan cinco (5) dimensiones para hallar el nivel de madurez de la organización con respecto al análisis de datos. En cuanto al instrumento de recolección de la información utilizada, se empleó un cuestionario virtual compuesto por 53 preguntas de selección múltiple y única respuesta, descrito en el Anexo A; dichas preguntas fueron traducidas al idioma español para garantizar su correcta comprensión. Así mismo, se incluyeron interrogantes útiles para caracterizar económicamente a las empresas y recabar datos confidenciales de contacto con fines de envío de resultados.

Para asignar el valor de nivel de madurez a cada posible respuesta del cuestionario, fueron tomados en consideración los niveles de madurez provistos por el modelo TDWI en su versión 2020; esto permitió definir objetivamente los parámetros relacionados con cada posible respuesta a escoger en el cuestionario.

Cabe aclarar que, debido a que existen preguntas con menos posibles respuestas que niveles de madurez del modelo TDWI (los cuales son 5), se permiten saltos de niveles entre las respuestas. Lo anterior, con el fin de simplificar la aplicación del cuestionario, teniendo en cuenta sólo las respuestas más representativas.

Por otra parte, se utilizó un muestreo no probabilístico por conveniencia, con base en la accesibilidad y voluntad de las empresas encuestadas. La muestra obtenida fue de 30 empresas pymes, desarrolladoras de software, localizadas en la ciudad de Bogotá DC y alrededores. En consecuencia, a cada empresa encuestada se le brindó una adecuada explicación del objetivo del proyecto, el modelo utilizado y la respectiva política de tratamiento de datos.

Asimismo, en relación con los parámetros detallados en la sección 2.1.1. del Capítulo II, y de acuerdo con la metodología planteada por Halper (2020), cada dimensión tiene una puntuación potencial de 20 puntos, señalando que las organizaciones pueden tener diferentes niveles de madurez en cada una de estas. Dicha puntuación es una variable discreta que, según lo sugerido por Halper (2020), debe ser redondeada al entero más cercano.

En función de lo anterior, la Tabla 2-1 muestra el baremo de interpretación que, a efectos de ejemplificar, si una entidad obtuvo en la dimensión *Organización* un puntaje de 5, esto indicaría que la misma se encuentra en un nivel de madurez *Incipiente*, lo cual señala que la empresa no está utilizando la analítica y que sólo utiliza hojas de cálculo para procesos de registro. En este sentido, no hay un esfuerzo importante en establecer una cultura fundamentada en los datos y, por tanto, las decisiones se toman con base en el instinto.

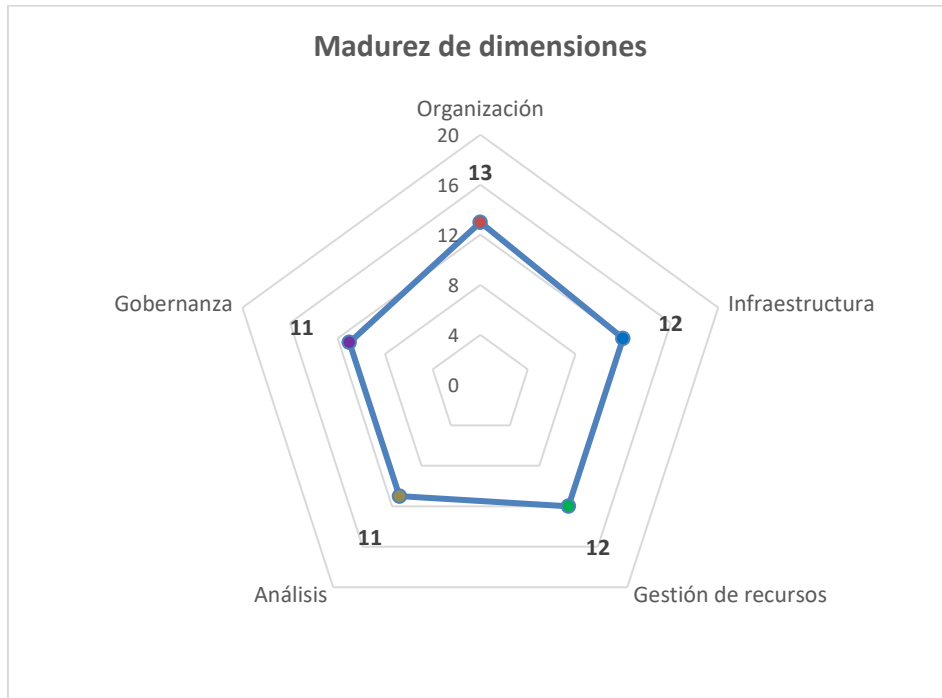
3.2 Resultados generales de la aplicación del modelo TDWI

Una vez se ha preparado el instrumento de recolección de información del modelo TDWI, se procede a aplicarlo al grupo de 30 pymes seleccionadas previamente, bajo los criterios especificados, como son: ser desarrolladoras de software, radicadas en la ciudad de Bogotá y formalizadas en la Cámara de Comercio, de acuerdo con lo especificado en el apartado inmediatamente anterior.

Es relevante así también reportar, que debido a la situación sanitaria global el proceso de aplicación debió sujetarse a la disponibilidad de las empresas y del personal designado para tal fin, lo cual se convirtió en una limitación a destacar, ya que, los cambios y procesos llevados a cabo en a la vida cotidiana de estas organizaciones, pudo interferir de alguna manera con el trabajo de campo, toda vez, que los tiempos y atención del caso fueron reducidos por las implicaciones propias de la virtualización obligada para muchas empresas.

Seguidamente, se presentan en la Figura 3-1 las puntuaciones de madurez obtenidas en cada dimensión contemplada.

Figura 3-1: Madurez obtenida en pymes desarrolladoras de software de Bogotá DC y alrededores a lo largo de dimensiones del modelo TDWI



Fuente: elaboración propia

De la anterior figura es posible observar que todas las puntuaciones para las dimensiones de madurez están contenidas entre el rango de 11 a 15, por lo cual estas se ubican en el nivel de madurez *Establecido*. Resalta, además, que la dimensión que más puntaje obtuvo fue *Organización*, la cual estuvo dos puntos por encima de las dimensiones *Análisis* y *Gobernanza*. En las siguientes subsecciones se explica en detalle la interpretación de los resultados por cada una de las dimensiones evaluadas.

3.2.1 Organización

Las compañías evaluadas en Bogotá DC y alrededores obtuvieron para la dimensión *Organización* un nivel de madurez *Establecido*, esto indica que dichas organizaciones están implementando sólidas estrategias de analítica a lo largo de sus áreas de negocio. En esta etapa, la dirección de la compañía trabaja en conjunto con el departamento de TI en la implementación de proyectos

estratégicos. Usualmente, los cambios positivos que se dan en este nivel de madurez se deben a presiones externas como alteraciones en las preferencias de los clientes o periodos de crisis que originan una necesidad de cambio o reingeniería de procesos que, a su vez, hacen que estas empresas involucren a la analítica como herramienta clave para la toma de decisiones.

En este nivel de madurez los equipos de trabajo se esfuerzan por buscar las herramientas de analítica adecuadas para solucionar los problemas que acontecen en la organización. Así mismo, los pequeños proyectos llevados a cabo por limitados equipos en diferentes departamentos comienzan a dar sus frutos; lo anterior, lleva a estas empresas a ver el análisis de datos como una herramienta imprescindible para el negocio. Es frecuente observar en esta etapa que, a medida que los primeros proyectos van dando resultados favorables, más personas se entusiasman y se unen a estas iniciativas de analítica.

No obstante, la compañía aún no percibe a la analítica como una ventaja competitiva y factor diferenciador para la organización; a la vez que aún no se realizan innovaciones en cuanto a los procesos internos de análisis de datos y una cultura o filosofía organizacional basada en datos aún no ha sido establecida en la empresa. En consecuencia, diversos aspectos deben ser mejorados para quebrar la brecha que separa a las pymes desarrolladoras de software de Bogotá DC del siguiente nivel de madurez.

3.2.2 Infraestructura

Como se expresó anteriormente, el conjunto de compañías encuestadas de Bogotá DC y alrededores obtuvo en todas sus dimensiones un nivel de madurez *Establecido*. Para el caso de la dimensión *Infraestructura*, esto indica que usualmente dichas organizaciones poseen ya algún almacén de datos o *data mart*, por lo general, utilizando solamente componentes de tipo estructurado. En este nivel de madurez muchas de estas compañías comienzan a contemplar la posibilidad de migrar a otras plataformas como la nube y utilizar más tipos de datos,

tales como datos semi estructurados y no estructurados. Así mismo, se comienza a considerar la inclusión de tareas de analítica en la nube con fines de reducir costos y aumentar la capacidad de análisis y procesamiento de datos.

Sin embargo, no se cuenta con un recurso humano calificado en materia de conceptos básicos de infraestructura y análisis de datos en toda la empresa; de modo que las labores de análisis se encuentran centralizadas en el área de TI y gran parte de la responsabilidad de los proyectos de implementación recae sobre este departamento. Así mismo, la infraestructura tecnológica actual no es escalable.

3.2.3 Gestión de recursos

En este caso, el nivel de madurez obtenido por las pymes desarrolladoras de software encuestadas refleja que, muy probablemente, estas empresas ya han formado un equipo responsable por la analítica de datos en la organización; el cual normalmente funciona como un centro de inteligencia de negocio que brinda soporte a diversas áreas de la compañía. En este punto, dichas empresas muy probablemente ya están sopesando la decisión de contratar científicos e ingenieros de datos, propiciando este nivel de madurez la antesala para grandes crecimientos en materia de analítica dentro de la organización.

Como oportunidades de mejora se tiene la inclusión de científicos e ingenieros de datos dentro de los equipos de trabajo. Así mismo, el brindar capacitación al personal ajeno al área de TI en materia de analítica, según lo requieran sus funciones, ayudará a la organización a cubrir la brecha que le impide pasar al siguiente nivel de madurez. De igual manera, la inclusión de profesionales con formación mixta, entre las requeridas por los cargos convencionales y análisis de datos, permitirá potenciar las áreas con carencias en analítica. En términos de financiación, la organización deberá comenzar a emplear métodos que permitan priorizar proyectos de mejora y que estén dentro de sus capacidades económicas.

3.2.4 Análisis

En este caso, el nivel de madurez *Establecido* indica que las pymes desarrolladoras de software de Bogotá DC y sus alrededores ya utilizan reportes, *dashboards*, y algunas herramientas informáticas para soportar el análisis y visualización de datos. En este nivel de madurez las compañías ya comienzan a contemplar la utilización de modelos predictivos de *Machine Learning* para potencializar sus operaciones de producción, marketing, toma de decisiones, entre otras más.

Sin embargo, en este punto, las habilidades de los trabajadores en términos de analítica comienzan a convertirse en un punto débil; esto, debido a que a medida que la complejidad de los modelos aumenta, como en el caso de los modelos de *Deep Learning*, se requieren habilidades y competencias cada vez más específicas. Así mismo, como área de mejora se debe tener en cuenta que cualquier compañía puede tener petabytes de datos, pero sólo aquellas que pueden transformar éstos en valor agregado para los clientes internos y externos son las que consiguen transformar a la analítica en una ventaja competitiva.

3.2.5 Gobernanza

Para la dimensión *Gobernanza*, el nivel de madurez *Establecido* significa que un equipo de gobierno de datos y ciberseguridad ha sido conformado. Usualmente se tienen políticas básicas y determinados protocolos de privacidad establecidos; sin embargo, no se emplea un fuerte y estricto control sobre los mismos; en otras palabras, se comienzan a implementar políticas que perduran en documentos públicos pero que no se implementan en su totalidad.

Las pymes desarrolladoras de software de Bogotá DC y sus alrededores deben comenzar a prever que cuando el volumen de datos almacenado y procesado aumente, interrogantes como ¿a quién le pertenecen los datos?, ¿de dónde vinieron estos datos?, ¿cómo se recolectaron?, ¿son confiables?, entre otras más,

deben ser orientadoras hacia la construcción de sólidas políticas de gobernanza de datos y ciberseguridad.

En conclusión, los anteriores resultados ubican en términos generales a las pymes desarrolladoras de software de Bogotá DC y alrededores en el nivel de madurez *Establecido*.

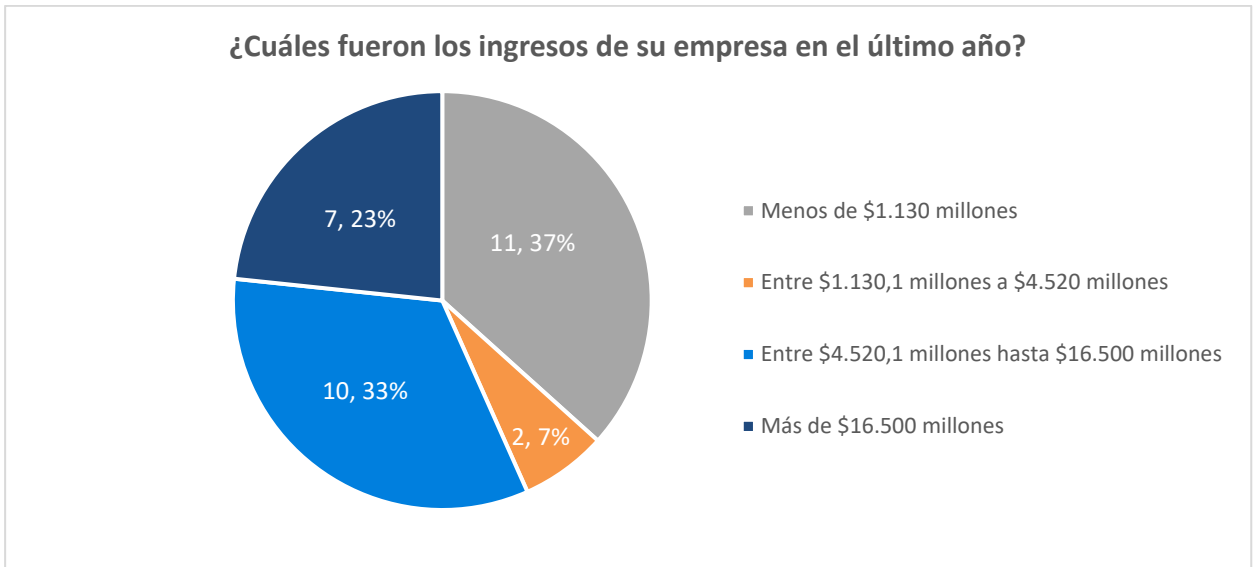
3.3 Resultados discriminados de la aplicación del modelo TDWI

En esta sección se expondrán los resultados alcanzados en cada dimensión de madurez, pero discriminados por categorías alusivas a los ingresos anuales de las compañías bajo estudio; en este sentido, se desea lograr un análisis más profundo de la madurez de las empresas en función de su capacidad económica. Dicho lo anterior, se consideraron las siguientes cuatro categorías de ingresos anuales:

- Menos de \$1.130 millones
- Entre \$1.130,1 millones a \$4.520 millones
- Entre \$4.520,1 millones hasta \$16.500 millones
- Más de \$16.500 millones

De acuerdo con dicha escala de clasificación, las proporciones de empresas en cada escala son mostradas en la Figura 3-2.

Figura 3-2: Proporción de pymes desarrolladoras de software en Bogotá y alrededores en función de ingresos netos anuales

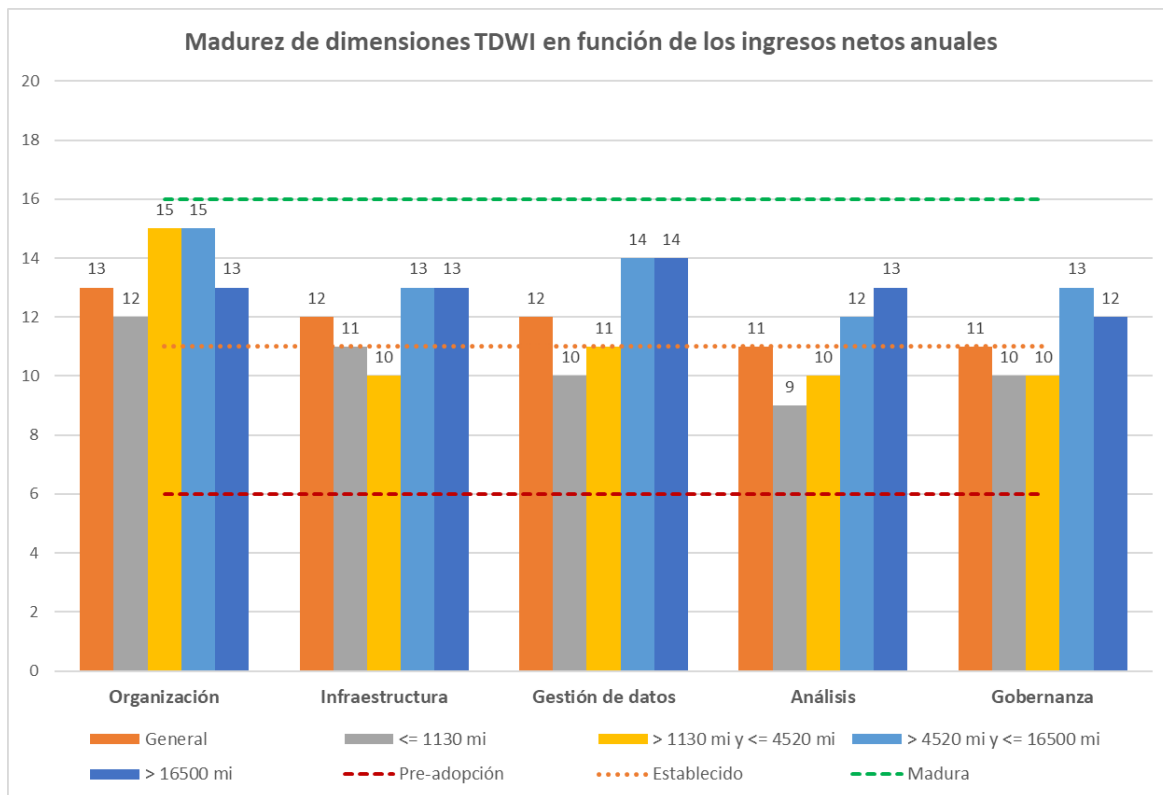


Fuente: elaboración propia

En la Figura 3-2 se aprecian las proporciones de pymes desarrolladoras de software de Bogotá y alrededores, en función de sus ingresos netos anuales, en razón a que dichas categorías fueron tomadas en consideración para realizar un análisis más detallado sobre la madurez de las compañías en relación con su capacidad económica; de esta manera, la siguiente

Figura 3-3 resume los resultados del análisis realizado.

Figura 3-3: Madurez de pymes desarrolladoras de software en Bogotá y alrededores en función de sus ingresos netos anuales



Fuente: elaboración propia

De la anterior

Figura 3-3 se observa cómo las empresas con ingresos netos anuales menores o iguales a 1.130 millones de COP poseen un nivel de madurez *Establecido* sólo para las dimensiones *Organización* e *Infraestructura*. Para el resto de las dimensiones, dichas compañías poseen el nivel de madurez *Pre-adopción*, esto indica que las pymes con ingresos netos anuales menores o iguales a 1.130 millones de COP se encuentran predominantemente en la mencionada categoría.

Con relación a las pymes con ingresos netos anuales mayores a 1.130 millones de COP y menores o iguales a 4.520 millones de COP, se observa un comportamiento similar al de la categoría antes mencionada; esto, debido a que sólo las dimensiones *Organización* y *Gestión de datos* alcanzan la categoría de madurez *Establecido*, mientras que el resto de las dimensiones adquieren el nivel *Pre-adopción*. Cabe resaltar que las pymes en este rango de ingresos obtuvieron una alta calificación en la dimensión *Organización*, restando sólo un punto de madurez para alcanzar la escala de *Madura*.

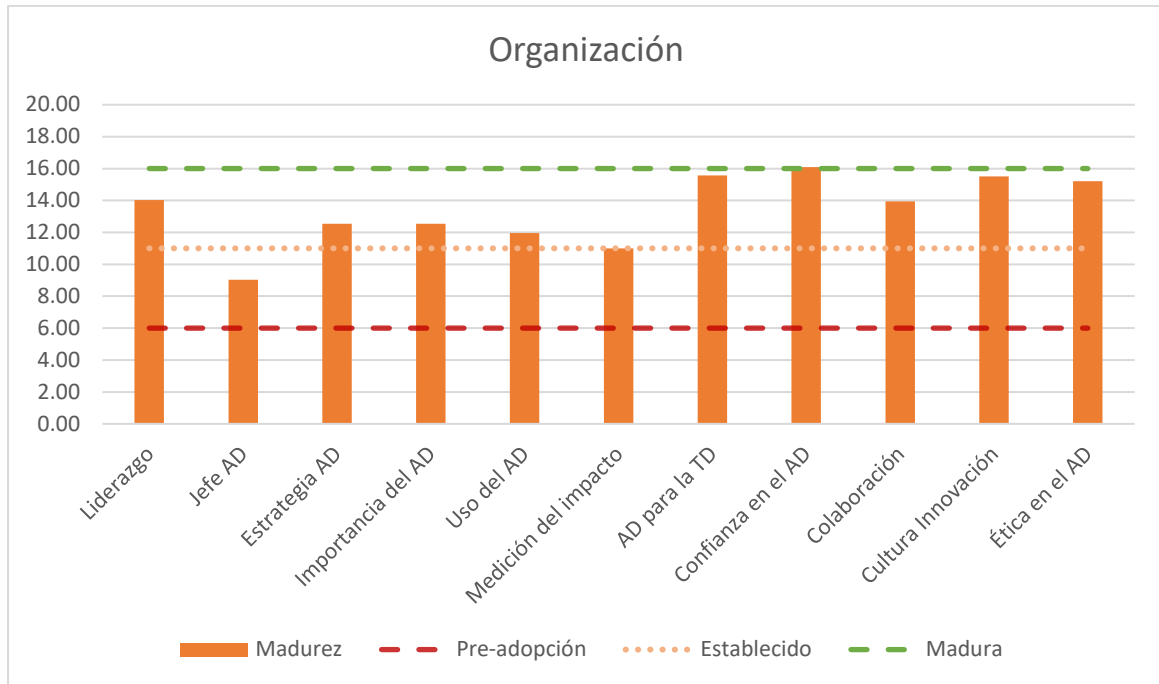
Para el caso de las pymes con ingresos mayores a 4.520 millones de COP y menores o iguales a 16.500 millones de COP, se observa que todas las dimensiones se encuentran dentro del nivel de madurez *Establecido*. Para las

pymes en este rango de ingresos resalta igualmente que la dimensión *Organización* es la que mayor madurez posee, estando ésta a sólo un punto de la categoría o nivel de madurez *Madura*. Del mismo modo, las pymes con ingresos anuales netos superiores a los 16.500 millones de COP mostraron todas sus dimensiones dentro del nivel de madurez *Establecido*; de tal forma que para esta última categoría de pymes la dimensión con mayor madurez fue *Gestión de datos*.

Lo anterior pone en manifiesto que las empresas con mayor capacidad económica mostraron, en términos generales, unos mayores niveles de madurez, siendo notoria la brecha entre las categorías de empresas con ingresos netos anuales menores o iguales a 4.520 millones de COP (las cuales están predominantemente en la categoría *Pre-adopción*) y aquellas con ingresos netos anuales mayores a 4.520 millones de COP (las cuales están predominantemente en la categoría *Establecido*). De igual manera resalta que, independientemente de las categorías económicas contempladas, las pymes muestran un nivel de madurez *Establecido* en la dimensión *Organización*.

Por otra parte, con relación al detalle de los elementos de madurez evaluados por las preguntas del cuestionario aplicado, se realiza a continuación una síntesis de los resultados más relevantes obtenidos. Es así como, con relación a la dimensión *Organización*, se observa en la Figura 3-4 que los elementos mejor evaluados fueron: *Confianza en el análisis de datos (Confianza en el AD)*, *Análisis de datos para la toma de decisiones (AD para la TD)*, *Cultura de innovación (Cultura innovación)* y *Ética en el análisis de datos (Ética en el AD)*.

Figura 3-4: elementos de madurez evaluados en la dimensión *Organización* en pymes desarrolladoras de software de Bogotá DC y alrededores

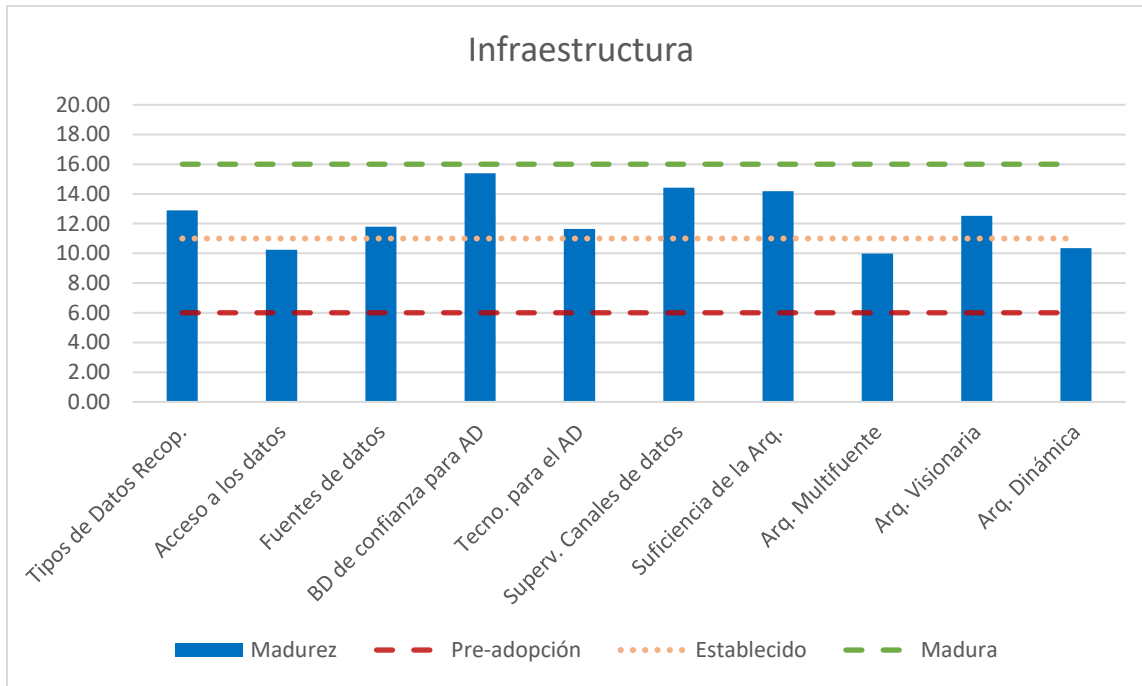


Fuente: elaboración propia.

Se destaca entonces que el elemento *Confianza en el análisis de datos* alcanzó un puntaje de 16,08, lo cual ubica a este aspecto en el nivel de madurez *Madura*, constituyendo así este elemento una gran fortaleza de las pymes desarrolladoras de software de Bogotá DC y alrededores, así como los conceptos antes mencionados, que igualmente poseen una madurez muy cercana al nivel de *Madura*. Por otra parte, se observa que la mayoría de las pymes no poseen un jefe en análisis de datos a cargo de dicha actividad; lo que es evidenciado por el elemento de la Figura 3-4 llamado *Jefe AD*, el cual se ubica en el rango del nivel de madurez *Pre-adopción*.

Una diferenciación igual a la anterior es hecha para la dimensión *Infraestructura*. En esta ocasión, se observa en la Figura 3-5, como fortaleza, el elemento *Bases de datos de confianza para el análisis de datos (BD de confianza para AD)*, el cual se encuentra cerca de llegar al nivel de *Madura*, lo que refleja que los usuarios de las compañías generalmente acceden a fuentes de datos seguros y de alta calidad que permiten generar resultados confiables.

Figura 3-5: elementos de madurez evaluados en la dimensión *Infraestructura* en pymes desarrolladoras de software de Bogotá DC y alrededores

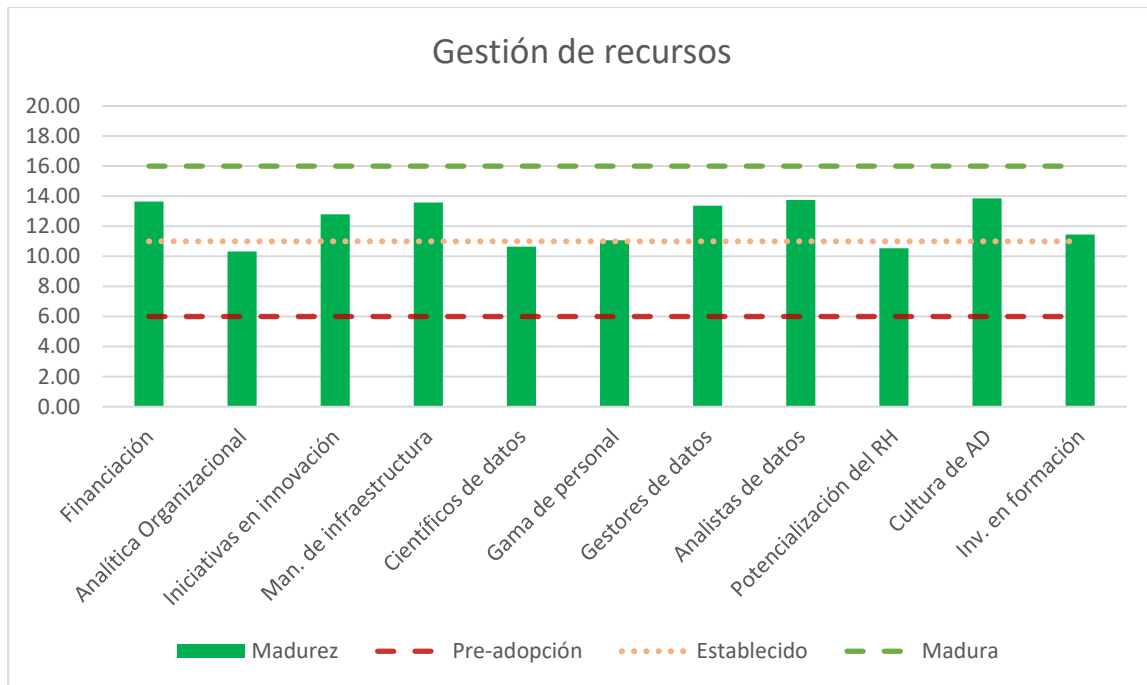


Fuente: elaboración propia

En contraste, resaltan como grandes aspectos de mejora la optimización de la accesibilidad a los datos a lo largo de todos los usuarios y sistemas de la compañía (*Acceso a los datos*), tener una arquitectura multi-fuente que permita a las pymes integrar diversos datos de fuentes dispares para su acceso y análisis (*Arq. Multifuente*) y poseer una arquitectura dinámica o escalable que permita adecuar las capacidades de almacenamiento y procesamiento de las pymes en función de las demandas de estas capacidades a lo largo del tiempo.

Con relación a los elementos de la dimensión *Gestión de recursos*, resalta en la Figura 3-6, como oportunidad de mejora, la inclusión de más aspectos de gestión organizacional en la analítica de datos, con el objetivo de proponer acciones de mejora a nivel organizativo (*Analítica Organizacional*).

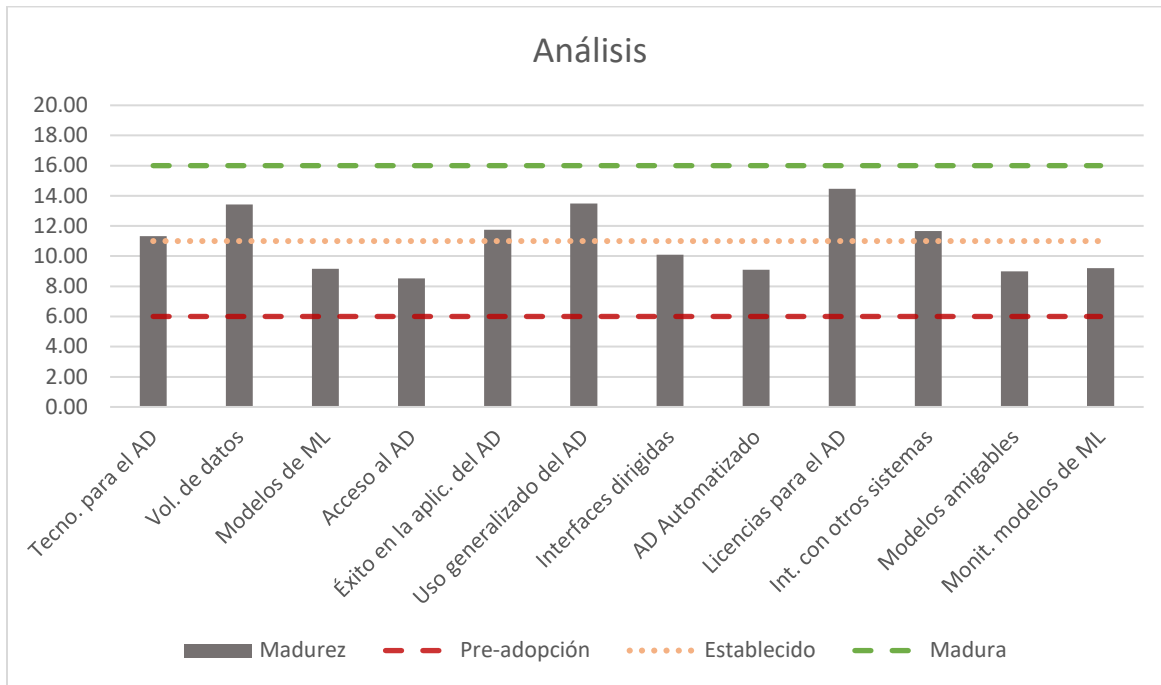
Figura 3-6: elementos de madurez evaluados en la dimensión *Gestión de recursos* en pymes desarrolladoras de software de Bogotá DC y alrededores



Fuente: elaboración propia

Así mismo, destacan, como oportunidades críticas de mejora, la inclusión de científicos de datos (elemento *Científicos de datos*) y la mejora de los conocimientos y habilidades del recurso humano de las compañías en torno al análisis de datos para, por medio de entrenamientos, formar a gran parte de dicho recurso humano en tópicos de analítica (*Potencialización del RH*). Por otra parte, como complemento, en términos de la dimensión *Análisis*, se observa un mayor número de falencias, tomando como referencia mínima de madurez al nivel *Establecido*. Esto puede ser observado en la Figura 3-7.

Figura 3-7: elementos de madurez evaluados en la dimensión *Análisis* en pymes desarrolladoras de software de Bogotá DC y alrededores



Fuente: elaboración propia.

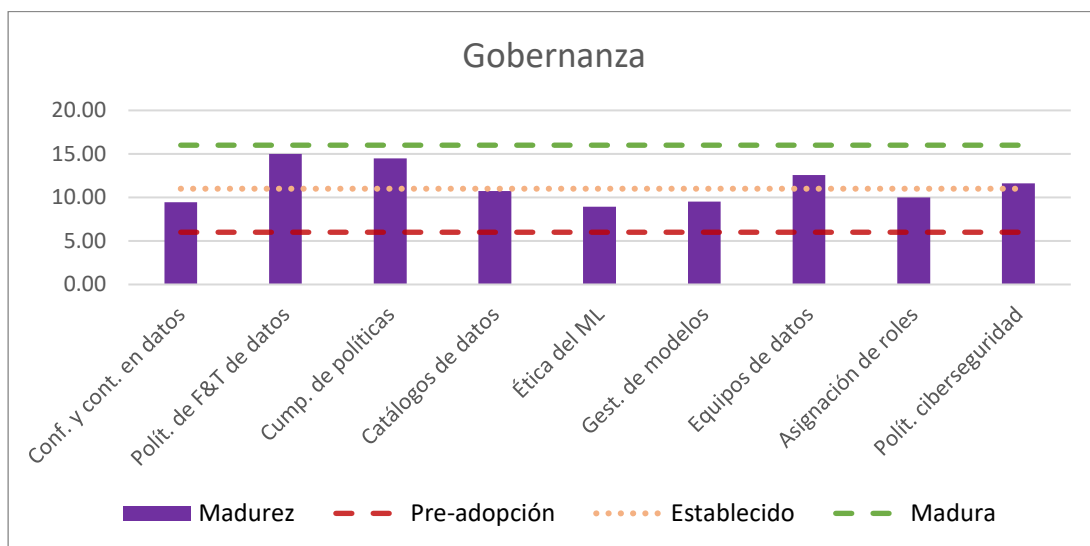
En esta ocasión, es posible destacar, en la Figura 3-7, cómo los elementos *Modelos de Machine Learning (Modelos de ML)*, *Acceso al análisis de datos (Acceso al AD)*, *Implementación de interfaces dirigidas (Interfaces dirigidas)*, *Implementación de análisis de datos automatizados (AD automatizado)*, *Uso de modelos amigables con el usuario (Modelos amigables)* y *Monitorización de modelos de Machine Learning (Monit. modelos de ML)* obtuvieron un puntaje por debajo del umbral del nivel de madurez *Establecido*. Lo anterior da cuenta de una baja inclusión de modelos de *Machine Learning* como estrategia para auxiliar los procesos de toma de decisiones y, por consiguiente, un bajo seguimiento al desempeño de estos modelos.

Así mismo, se observa que los modelos predictivos implementados son poco amigables con el usuario promedio, al igual que las interfaces de los sistemas de analítica utilizados frecuentemente no poseen una GUI amigable y personalizada en función del rol y tipo de usuario que accede al sistema.

Por otra parte, el acceso a las plataformas de análisis de datos no está disponible para todos los usuarios de la organización, lo cual es un punto clave de mejora para avanzar en la capacitación del recurso humano de la compañía, en términos de analítica, tal como se evidenció en el análisis hecho a la dimensión *Gestión de recursos*; esto, debido a que, si los funcionarios de las pymes en cuestión no tienen acceso y contacto con las herramientas de analítica, difícilmente se volverán profesionales competentes en dicha área.

El avance en las áreas de oportunidad antes comentadas permitirá que, naturalmente, con el pasar del tiempo, se puedan manejar datos de tipo *streaming* que actualicen las herramientas de BI y analítica en tiempo real, como es el caso de modelos predictivos que aprenden automáticamente de los nuevos datos que se reciben vía *streaming*. Finalmente, se expone en la Figura 3-8 el anterior análisis aplicado a la dimensión Gobernanza.

Figura 3-8: elementos de madurez evaluados en la dimensión Gobernanza en pymes desarrolladoras de software de Bogotá DC y alrededores



Fuente: elaboración propia

Para la mencionada dimensión se observan como fortalezas la *Definición de políticas de fuentes y tipos de datos (Polít. de F&T de datos)* y la *Aceptación y*

cumplimiento de políticas de gobierno de datos (Cump. de políticas); esto, teniendo en cuenta su proximidad con el nivel de *Madura*. No obstante, cabe aclarar que, desde una perspectiva general, también son áreas de oportunidad ya que no alcanzan a llegar por lo menos al rango de *Madura*. En contraste, como áreas de mejora principales, se precisan implementar acciones para aumentar las políticas de confiabilidad y control de datos (*Conf. y cont. en datos*) y las políticas éticas que rigen los principios de no discriminación y equidad a la hora de construir modelos de *Machine Learning (Ética del ML)*.

De igual forma, se debe trabajar en el control de los modelos de analítica construidos por medio de versionados y la inclusión de metadatos de cada uno de los modelos construidos (*Gest. de modelos*), además de una gestión que permita identificar más claramente los roles de los administradores de datos y los de los equipos de trabajo (*Asignación de roles*).

3.4 Evaluación e interpretación de las brechas de la capacidad de análisis de datos en las Pymes

A medida que las organizaciones intentan llegar a la categoría *Madura*, se encuentran con una serie de dificultades típicas de la transición entre los niveles de madurez *Establecido a Madura*; estas situaciones hacen que dicho proceso requiera de esfuerzos adicionales si se compara esta transición con los correspondiente al cambio de *Incipiente a Pre-adopción*. Dado que la mayoría de las pymes encuestadas se encuentran enfrentando este desafío, se presenta a continuación una explicación de dichas brechas, en función de cada dimensión de madurez:

- **Organización:** en esta dimensión las brechas son a menudo aspectos culturales y políticos de la organización que impiden que la analítica se vuelva una actividad generalizada en toda la organización. Por ejemplo, para el caso en cuestión puede que una pyme intente incluir la analítica en la

mayoría de sus procesos de negocio (lo cual es una señal de madurez); sin embargo, pueden existir obstáculos como resistencia al cambio, barreras tecnológicas, entre otros aspectos sobre los cuales se debe trabajar para que los empleados tengan confianza en la analítica y se fortalezcan en competencias TIC. Estas actividades pueden llevar mucho tiempo, incluso años; es por ello que una cultura organizacional de toma de decisiones basada en los datos es una herramienta esencial para pasar a nivel de madurez - *Madura*.

- **Infraestructura:** cuanto más madura sea una empresa, mejor uso hará de sus datos; de tal manera que pasará de utilizarlos para generar solamente informes de BI, a emplearlos también para crear modelos predictivos que permitan automatizar procesos y prever situaciones futuras con el objetivo de anticipar riesgos y oportunidades. Para avanzar en el cierre de la brecha a la que actualmente se enfrentan las pymes desarrolladoras de software de Bogotá DC y alrededores, se debe trabajar en la diversificación de los tipos de datos y en el almacenamiento y procesamiento de estos. Así mismo, se recomienda profundizar sobre una arquitectura escalable, coherente y de alta disponibilidad para respaldar el acceso a los datos por parte de múltiples usuarios y satisfacer las demandas internas y externas en cuanto a acceso y capacidad de procesamiento.
- **Gestión de recursos:** en el caso de la mayoría de las pymes desarrolladoras de software de Bogotá DC y alrededores, estas deben mejorar en el fomento y desarrollo de habilidades técnicas para el uso de nuevas tecnologías de análisis de datos, inteligencia de negocios, *Machine Learning* y *Deep Learning*; así como en el uso de datos no estructurados, bases de datos *NoSQL* y técnicas de *Big Data* para realizar análisis más avanzados. Las limitadas habilidades del recurso humano de la compañía son frecuentemente una barrera para avanzar hacia el nivel de *Madura* en las organizaciones; de tal forma que, si las pymes pueden costear la

contratación de personal idóneo para cubrir esta brecha de habilidades, se recomienda hacerlo. No obstante, también es posible utilizar un enfoque mixto entre la contratación externa y la capacitación interna.

- **Análisis:** muchas empresas recopilan grandes cantidades de datos, pero no hacen un uso efectivo de estos. Por tanto, una de las señales de madurez en esta dimensión es el hecho de contar con procesos que permitan transformar grandes volúmenes de datos en información valiosa para la compañía. Para lograr esto, es útil idear sistemas donde los nuevos datos que ingresan a la organización puedan ser analizados e integrados rápidamente a la infraestructura lógica. Para las pymes desarrolladoras de software de Bogotá DC y alrededores, se recomienda evaluar los puntos anteriormente dichos y la inclusión de herramientas de análisis de *Machine Learning* y *Deep Learning*. En este punto, la analítica se puede poner en práctica como parte de un proceso empresarial, de tal manera que esta pueda ser automatizada e integrada en los procesos empresariales.
- **Gobernanza:** una empresa madura entiende que la analítica, mal gestionada, puede ser una bomba de tiempo esperando por explotar, especialmente si se trabaja con datos de fuentes externas. Esto, debido a aspectos de propiedad y privacidad de los datos, legislación aplicable, permisos sobre los datos, origen y confiabilidad de los datos, entre otros aspectos que pueden generar problemas legales y de contaminación de datos. Para las pymes desarrolladoras de software de Bogotá DC y alrededores, lograr avanzar hacia el siguiente nivel de madurez implica implementar y estandarizar políticas de gobernanza de datos y ciberseguridad, contar con herramientas como instructivos de operación y catálogos de datos que permitan poder versionar, validar y monitorear recursos de datos, entre otros aspectos. Así mismo, se debe contar con una

unidad de negocio capaz de supervisar el programa de gobernanza de datos desde una perspectiva holística.

De otra parte, a partir de los hallazgos obtenidos, resulta importante analizar dichos resultados frente a otros estudios relacionados con el tema, realizados previamente y publicados mediante artículos y reseñas académicas. En este sentido, Marulanda, et al. (2013) realizó un estudio con el fin de conocer la manera como las empresas aplican la minería de datos en la gestión de conocimiento, de tal forma que sus resultados identificaron un bajo nivel de uso de estas herramientas, además que las empresas en la ciudad de Bogotá resaltaban por mostrar los indicadores más bajos de aprovechamiento. Lo anterior se relaciona con los hallazgos de la presente investigación, en tanto, las empresas encuestadas muestran un nivel de madurez establecido, lo que indica la necesidad de profundizar aún más en el uso de los datos para la mejora de sus procesos.

Adicionalmente, otro de los aspectos destacables es la cultura necesaria para el adecuado manejo de los datos, por lo cual, Gupta y George (2016) resaltan que, más allá de la recopilación de los datos y el acceso a la tecnología sofisticada, el énfasis debe ser dirigido al fortalecimiento de las habilidades humanas y los recursos intangibles, con el fomento de un aprendizaje organizativo que proporcione el verdadero valor al conocimiento extraído de los datos; de esta manera, los hallazgos de los autores citados se correlacionan con la deficiencia transversal, común a todas las brechas identificadas en esta investigación, en la medida en que se reconoce que la verdadera ventaja es aportada por el buen aprovechamiento que se realiza de la información recolectada en la toma de decisiones en todas las dimensiones de madurez estudiadas.

De otra parte, la dimensión de la infraestructura también obtuvo resultados significativos que identifican un nivel de madurez en el que aún se requieren esfuerzos para mejorar sus arquitecturas que facilite la integración de datos de diversas fuentes. Lo anterior, resulta de gran relevancia, tomando en cuenta lo

expuesto por Kamioka et al. (2017) que logran demostrar las ventajas competitivas que aporta la inversión en el perfeccionamiento de infraestructura y mejora de las capacidades del usuario para el logro de un análisis organizado de la *Big Data*.

En otros estudios autores como Hirvonen & Majuri (2020), han investigado en renglones productivos como la manufactura, que la gran diversidad de sistemas información y los distintos propósitos que cada uno sugiere, puede generar en ocasiones confusión a nivel organizacional, por lo cual establecer una alineación entre estos sistemas y los objetivos de cada entidad es algo vital para el proceso a desarrollar.

Los autores plantean como un lineamiento estratégico, la revisión de los principios que cada sistema trae de forma inherente y contrastarlos con los objetivos de cada entidad en las etapas preliminares, de tal forma, que los aspectos técnicos del ejercicio cotidiano, no se conviertan a futuro en un obstáculo para el desarrollo de las dinámicas que cada sistema de información integra para la generación de resultados (Hirvonen & Majuri, 2020).

3.5 Resumen del capítulo

En el anterior capítulo se expuso la capacidad de análisis de datos de las pymes desarrolladoras de software de Bogotá DC y alrededores, que participaron en la encuesta, conforme a la estructura de valoración del modelo TDWI. Así mismo, se realizó un análisis detallado de madurez por dimensiones y en función de los ingresos netos anuales percibidos por dichas pymes. Se identificó que las pymes con ingresos netos anuales mayores a 4.520 millones de COP mostraron una mayor madurez que aquellas con ingresos netos anuales inferiores. Así mismo, se identificó que la dimensión con mayor madurez, independientemente de su categoría de ingresos, fue la de *Organización*. En contraste, las dimensiones *Análisis* y *Gobernanza* mostraron ser las más débiles para las pymes bajo estudio.

Finalmente, diversos aspectos de diagnóstico y mejora fueron propuestos en el presente capítulo.

4.Recomendaciones para el fortalecimiento de la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá

Fundamentado en las brechas identificadas en la capacidad de análisis de datos utilizando el modelo TDWI, este capítulo pretende dar lineamientos generales y específicos para fortalecer dicho aspecto en las pymes desarrolladoras de software de la ciudad de Bogotá; esto, tomando en consideración la madurez percibida en varios aspectos identificados por medio de la aplicación del instrumento de recolección de información detallado previamente. Dichas recomendaciones son presentadas a continuación.

4.1 Recomendaciones generales para fortalecer la capacidad de análisis de datos

Se exponen a continuación diversas recomendaciones, las cuales se encuentran agrupadas por la respectiva dimensión de madurez a la cual pertenecen.

4.1.1 Organización

- Las pymes deben percibir a la analítica como una oportunidad para generar ventajas competitivas y consolidar un factor diferenciador para la organización
- Trabajar hacia una cultura organizacional de toma de decisiones basada en datos

- Realizar esfuerzos en disminuir al mínimo posible la resistencia al cambio y aumentar las competencias TIC en las organizaciones
- Instaurar cargos de Jefe en Análisis de datos, en donde los profesionales que ocupen este cargo velen por la correcta y estratégica ejecución de los procesos de analítica

4.1.2 Infraestructura

- Capacitar al recurso humano en materia de conceptos básicos de infraestructura y análisis de datos
- Descentralizar del área de TI las labores de analítica y paulatinamente ir dejando que la responsabilidad de proyectos de analítica recaiga en áreas diferentes a la de TI
- Implementar una infraestructura escalable, de alta disponibilidad y basada en la nube
- Diversificar los tipos y fuentes de datos

4.1.3 Gestión de recursos

- Capacitar al recurso humano en materia de analítica e incluir dentro de las nóminas a científicos e ingenieros de datos
- Emplear métodos que permitan priorizar proyectos de mejora en función del beneficio esperado para el cliente y la organización atendiendo aspectos financieros y de tiempo
- Mejorar en el fomento y desarrollo de habilidades técnicas para el uso de nuevas tecnologías de análisis de datos, inteligencia de negocios, *Machine Learning* y *Deep Learning*

4.1.4 Análisis

- Utilizar modelos de *Machine Learning* y *Deep Learning* para potencializar las operaciones de producción, marketing, toma de decisiones, entre otras.

- Monitorear el desempeño de los modelos de *Machine Learning* y *Deep Learning* para realizar una trazabilidad de su desempeño a lo largo del tiempo y poder establecer alertas de reentrenamiento
- Trabajar hacia el aprovechamiento total de los datos
- Contemplar herramientas de *data streaming* en donde los nuevos datos que ingresen a la organización sean analizados e integrados rápidamente a la infraestructura lógica de las pymes
- Integrar a la analítica dentro de los estándares de procesos
- Mejorar la accesibilidad a las herramientas de análisis de datos
- Implementar modelos transparentes que sean fáciles o medianamente-fáciles de usar por un usuario promedio
- Implementar GUIs amigables y personalizadas

4.1.5 Gobernanza

- Realizar un mayor y estricto control al cumplimiento de las políticas de gobernanza de datos
- Consolidar políticas de ciberseguridad y privacidad, propiedad y manejo de datos que se actualicen periódicamente en función de las normativas nacionales e internacionales vigentes, así como de las dinámicas del mercado
- Consolidar políticas éticas que rijan los principios de no discriminación, justicia y equidad a la hora de construir modelos de *Machine Learning* y *Deep Learning*. De igual forma, se debe mejorar la gestión de estos modelos
- Establecer equipos capaces de supervisar el programa de gobernanza de datos desde una perspectiva holística

4.2 Resumen del capítulo

Con base en los resultados obtenidos de la aplicación del modelo de TDWI se proponen una serie de recomendaciones que lleven a las empresas estudiadas a superar las brechas de madurez identificadas para cada una de las categorías

consideradas. De esta manera, para cada una de las categorías de *Organización*, *Infraestructura*, *Gestión de recursos*, *Análisis* y *Gobernanza* se detallan las acciones que se consideran necesarias para fortalecer la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá.

5. Conclusiones y recomendaciones

5.1 Conclusiones

El presente trabajo de investigación tuvo como objetivo estimar la madurez de la capacidad de análisis de datos en las pymes del sector de desarrollo de software de la ciudad de Bogotá. Para ello, se realizó una caracterización de los modelos para el diagnóstico de la capacidad de análisis de datos en las organizaciones, a través de una revisión sistemática de literatura, **observando las generalidades, contextos de aplicación y debilidades de cada propuesta conceptual**, para luego determinar el modelo idóneo a aplicar, definido como TDWI. Dicho modelo de madurez fue traducido al idioma español para facilitar su aplicación; así mismo, fue explicado a lo largo de sus dimensiones y niveles a lo largo del presente documento.

Como factor de impacto en la investigación se resalta que las organizaciones que participaron en el trabajo de campo o aplicación del modelo, se vieron sometidas a las condiciones propias de las restricciones de movilidad, distancia social y virtualización total de actividades (para algunos casos), debido a las medidas de contención para la atención sanitaria derivada a nivel mundial por la irrupción del virus Covid-19.

Luego de aplicar el modelo TDWI en la muestra objetivo, los resultados indican que, en términos generales, las pymes desarrolladoras de software de la ciudad de Bogotá se encuentran en un nivel de madurez *Establecido*, el cual hace referencia a aquellas empresas que ya han comenzado a realizar esfuerzos por implementar

la analítica a lo largo de la organización y han comenzado a obtener pequeños resultados satisfactorios de dichos esfuerzos.

En estas entidades el análisis de datos se encuentra establecido y poco a poco se comienza a volver una actividad cotidiana; sin embargo, aún resta implementar diversos esfuerzos para superar la brecha que separa a las organizaciones del nivel de madurez *Establecido* de las organizaciones en el nivel de madurez - *Madura*. Estos esfuerzos van desde el entrenamiento del recurso humano en tópicos de analítica hasta implementar una cultura organizacional de toma de decisiones basada en datos a lo largo de todos los niveles organizacionales.

Luego de implementar un análisis de madurez, en función de los ingresos netos anuales percibidos por las pymes bajo estudio, se identificó que las empresas con ingresos netos anuales mayores a 4.520 millones de COP mostraron una mayor madurez en todas las dimensiones del modelo TDWI, siendo el nivel de *Establecido* constante a lo largo de las dimensiones evaluadas en las empresas en dicho rango de ingresos. En contraste, aquellas pymes con ingresos netos anuales inferiores a 4.520 millones de COP mostraron estar predominantemente en el nivel de madurez *Pre-adopción*.

Por otra parte, se identificó que la dimensión *Organización* es la que mayor madurez posee, independientemente de los ingresos percibidos por las pymes; en contraste, las dimensiones *Análisis* y *Gobernanza* fueron las más débiles para las empresas bajo estudio, siendo la primera aquella en donde más debilidades se identificaron.

De otro lado, los resultados obtenidos fueron relacionados con estudios previos asociados con el tema, encontrando que el nivel de madurez identificado resulta coherente con otros análisis realizados en pymes colombianas, observando la necesidad de mejorar en los aspectos requeridos para un mejor aprovechamiento.

En relación con esto, igualmente se destacaron resultados de otros actores que resaltan en la ventaja competitiva que proporciona el análisis de datos en la toma de decisiones, la importancia de propiciar el fortalecimiento de las habilidades humanas para el manejo de estos recursos, el valor de invertir en la utilización de una arquitectura que facilite la integración de datos de diversas fuentes y, por último, el valor estratégico de establecer políticas, medidas de seguridad, controles y metodologías para el perfeccionamiento del análisis de los datos.

En esta misma línea, los aportes del análisis de la *Big Data* resultan ser un gran componente estratégico por sí mismo, de ahí la necesidad detectada en profundizar en las políticas, sistemas de seguridad, controles y metodologías para evitar desviaciones y perfeccionar el análisis de los datos. Aun así, Ramamurthy (2017) también resalta que, para el adecuado aprovechamiento de los beneficios asociados, es necesario tener en cuenta la capacidad dinámica, el número de empleados y la edad de la empresa, en el diseño de las políticas requeridas.

En términos generales, los hallazgos demuestran que las pymes colombianas desarrolladoras de software aún tienen mucho trabajo por realizar para el aprovechamiento del análisis de los datos, lo cual no es ajeno al comportamiento global identificado por Bean y Davenport (2019), quienes identificaron que, si bien los empresarios reconocen un gran valor en la *Big Data*, aun fallan en incluir su desarrollo y aprovechamiento en los planes estratégicos de sus organizaciones; por lo cual aquellas empresas que orienten sus esfuerzos al mejoramiento de su toma de decisiones, seguramente verán impactos positivos en el logro de sus objetivos.

Adicionalmente, diversos aspectos de diagnóstico y mejora fueron propuestos en él, con el objetivo de coadyuvar a fortalecer la capacidad de análisis de datos en las pymes desarrolladoras de software de la ciudad de Bogotá, detallando en las acciones necesarias para el cierre de brechas en cada una de las dimensiones incluidas en este estudio.

Con base en los resultados obtenidos, vale la pena destacar los beneficios de la metodología escogida para el análisis, en tanto el modelo TWID brinda una visión comparativa del desarrollo del análisis de datos en las empresas estudiadas, permitiendo enmarcar las prácticas reales de dichas organizaciones en parámetros de medición y valoración destacables. Al respecto, un análisis de los comportamientos estratégicos de las empresas aporta una identificación oportuna y útil para la determinación de las brechas de desarrollo que permita el establecimiento de políticas gubernamentales y gremiales de incentivo para el impacto de las necesidades detectadas; y de esta manera, favorecer el sector de empresas desarrolladoras de softwares.

Finalmente, dentro de las limitaciones se encontró que debido a la extensión del cuestionario fue necesario realizar las encuestas en tres partes para facilitar su diligenciamiento y aumentar de esta manera la tasa de respuestas; adicionalmente las limitaciones propias de las preguntas incluidas en el modelo utilizado, obligaron a considerar permisible los saltos en los niveles de madurez entre las respuestas; no obstante, a pesar de ello, fue posible establecer que el modelo utilizado arrojó resultados confiables objetivos sobre las dimensiones de madurez estudiadas. Otra limitación está relacionada con el hecho de que el modelo utilizado para medir las capacidades no contempla análisis estadísticos más profundos como la correlación entre variables, lo cual enriquecería el análisis alcanzado.

5.2 Recomendaciones

Una alta madurez en procesos de analítica permite a las pymes tomar mejores decisiones, las cuales les permitirán crecer a un ritmo acelerado; no obstante, no todas las pymes necesitan estar en la máxima categoría, en razón a que la naturaleza de sus operaciones, sus condiciones económicas y hasta el aporte de valor a las pymes y sus clientes internos y externos son aspectos claves a evaluar

a la hora de decidir qué tipo de nivel de madurez implementar la analítica en las organizaciones.

En este sentido, se recomienda a futuras investigaciones crear modelos que indaguen, además del estado actual de la compañía, su visión futura; de esta manera, se podrá tener un panorama más general de dónde están las compañías y a dónde desean llegar. Lo anterior, facilitará la creación de *roadmaps* u hojas de ruta que permitan a las pymes crear sus propios planes de acción para llegar desde dicho punto actual hacia una meta futura acorde con las características y necesidades de cada organización.

Así mismo, se sugiere la realización de estudios empíricos que permitan estimar cuánto tiempo le llevará a las pymes desarrolladoras de software pasar de un nivel de madurez a otro en función de los ingresos anuales percibidos. Esto, analizando a su vez las inversiones que se deban realizar en diferentes áreas y la repercusión que éstas tengan en las utilidades de las compañías y la satisfacción del cliente a través del tiempo. Lo anterior, coadyuvará a seguir construyendo métodos y conocimientos útiles para guiar a las pymes desarrolladoras de software e incluso a las de otros sectores hacia una correcta y satisfactoria implementación de la analítica como factor potenciador de desarrollo.

También se sugiere adaptar y ampliar el modelo utilizado para incorporar los análisis estadísticos detallados detectados como una limitante.

6. Bibliografía

- Alderson, P., Green, S., & Higgins, J. (2004). *Cochrane reviewers' handbook, version 4.2.2*. (The Cochrane Collaboration (ed.)).
https://www.mv.helsinki.fi/home/hemila/karlowski/handbook_4_2_2_Karlowski.pdf
- Bean, R., & Davenport, T. (2019). Companies are failing in their efforts to become data-driven. *Harvard Business Review*.
- Bedeley, R. T., & Nemati, H. (2014). Big Data Analytics: A key capability for competitive advantage. *20th Americas Conference on Information Systems, AMCIS 2014*.
<https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1536&context=amcis2014>
- Bihani, P., & Patil, S. (2014). A comparative study of data analysis techniques. *International Journal of Emerging Trends & Technology in Computer Science*, 3(2), 95–101. www.ijettcs.org
- Bonilla, J., & Rao, B. (2015). Decoding data analytics capabilities from topic modeling on press releases. *Portland International Conference on Management of Engineering and Technology, 2015-Septe*, 1959–1968.
<https://doi.org/10.1109/PICMET.2015.7273249>
- Bryman, A., & Cramer, D. (1992). Quantitative data analysis for social scientists. *Estudios Geográficos*, 53(207), 347.
<http://search.proquest.com/openview/71ba8781f88a269938b860d448e6e8d2/1?pq-origsite=gscholar&cbl=1818083>
- Cates, J. E., Gill, S. S., & Zeituny, N. (2005). The Ladder of Business Intelligence (LOBI): a framework for enterprise IT planning and architecture. *International Journal of Business Information Systems*, 1(1/2), 220.

- <https://doi.org/10.1504/IJBIS.2005.007408>
- Coleman, S. (2016). Data-mining opportunities for small and medium enterprises with official statistics in the UK. *Journal of Official Statistics*.
<https://doi.org/10.1515/JOS-2016-0044>
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Davenport, T., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business School Press.
- Denning, D. E., Nicholson, W., Sande, G., & Shoshani, A. (1984). Research topics in statistical database management. *IEEE Database Eng. Bull.*, 7(1), 4–9.
- Dinter, B. (2012). The maturing of a business intelligence maturity model. *AMCIS 2012 Proceedings*.
<http://elibrary.aisnet.org/Default.aspx?url=https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1083&context=amcis2012>
- Du, X., Liu, B., & Zhang, J. (2019). Application of Business Intelligence Based on Big Data in E-commerce Data Analysis. *Journal of Physics: Conference Series*, 1395(1), 012011. <https://doi.org/10.1088/1742-6596/1395/1/012011>
- Dulcé, H. J. (2016). Datos, información y conocimiento. *Respuestas*, 21(1), 4.
<https://revistas.ufps.edu.co/index.php/respuestas/article/download/642/646>
- Eason, K. D. (1989). *Information technology and organisational change*. CRC Press.
- Eckerson, W. (2007). TDWI benchmark guide: interpreting benchmark scores using TDWI's maturity model. *TDWI Research*, 3–14.
- Garmaki, M., & Boughzala, I. (2016). Conceptualization of Big data analytics Capability based on IT capability: Primary findings. *21st Symposium of the Association Information and Management 2016, AIM 2016*.
- Grossman, R. L. (2018). A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, 38(1), 45–51.
<https://doi.org/10.1016/j.ijinfomgt.2017.08.005>
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics

- capability. *Information & Management*, 53(8), 1049–1064.
- Halper, F. (2020). *TDWI Analytics Maturity Model Guide: Assessment Guide*.
- Halper, F., & Stodder, D. (2014). TDWI Analytics Maturity Model Guide. In *TDWI Research*. https://www3.microstrategy.com/getmedia/9b914607-084f-4869-ae64-e0b3f9e003de/TDWI_Analytics-Maturity-Guide_2014-2015.pdf
- Hao, S., Zhang, H., & Song, M. (2019). Big data, big data analytics capability, and sustainable innovation performance. *Sustainability*, 11(24), 7145.
- Hatta, N. N. M., Miskon, S., Ali, N. M., Abdullah, N. S., Ahmad, N., Hashim, H., Alias, R. A., & Maarof, M. A. (2015). Business intelligence system adoption theories in SMES: A literature review. *ARPN Journal of Engineering and Applied Sciences*.
- Hicks, S. C., & Peng, R. D. (2019). Elements and Principles for Characterizing Variation between Data Analyses. *ArXiv Preprint ArXiv:1903.07639*. <http://arxiv.org/abs/1903.07639>
- Hirvonen, J., & Majuri, M. (2020). Digital capabilities in manufacturing SMEs. *Procedia Manufacturing*, 51, 1283–1289. <https://doi.org/10.1016/j.promfg.2020.10.179>
- Humphrey, W. S. (1989). *Managing the software process*. Addison-Wesley Longman Publishing Co., Inc.
- Irwin, S. (2008). Data analysis and interpretation: emergent issues in linking qualitative and quantitative evidence. In Guilford Publications (Ed.), *Handbook of emergent methods in social research* (pp. 415–435).
- Islam, M. (2020). Data Analysis: Types, Process, Methods, Techniques and Tools. *International Journal on Data Science and Technology*, 6(1), 10. <https://doi.org/10.11648/j.ijdst.20200601.12>
- Joyanes, L. (2013). *Big Data, Análisis de grandes volúmenes de datos en organizaciones* (Alfaomega Grupo Editor (ed.)).
- Kamioka, T., Hosoya, R., & Tapanainen, T. (2017). Effects of User IT Capabilities and Organized Big Data Analytics on Competitive Advantage. *PACIS*, 36.
- Khan, K. S., Ter Riet, G., Glanville, J., Sowden, A. J., & Kleijnen, J. (2001).

Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews.

- Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University, UK and National ICT Australia*. <https://doi.org/10.1.1.122.3308>
- Kumar, V., Goyal, P., & Vandana, R. J. G. (2017). Stakeholder Classification: A Sustainability Marketing Perspective. *EVIDENCE BASED MANAGEMENT*, 111.
- Ladley, J. (2010). *Making enterprise information management (EIM) work for business: A guide to understanding information as an asset*. Morgan Kaufmann.
- Letón, M., & Pedromingo, A. (2001). *Introducción al Análisis de Datos en Meta-Análisis*. Ediciones Díaz de Santos.
- Luhn, H. P. (1958). A Business Intelligence System. *IBM Journal of Research and Development*, 2(4), 314–319. <https://doi.org/10.1147/rd.24.0314>
- Martínez, C. (2014). *Técnicas e Instrumentos de Recogida y Análisis de Datos*. Editorial UNED.
- Marulanda, C. E., López, M., & Mejía, M. H. (2013). Minería de datos en gestión del conocimiento de pymes de Colombia. *Revista Virtual Universidad Católica Del Norte*, 1(38), 158-170–170. <https://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/821/1339>
- Masood, T., & Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, 121, 103261. <https://doi.org/10.1016/j.compind.2020.103261>
- Maxwell, J. A., & Chmiel, M. (2014). Notes toward a theory of qualitative data analysis. In *The SAGE handbook of qualitative data analysis* (pp. 21–34). Sage Thousand Oaks, CA.
- Meyer, S. L. (1975). *Data Analysis For Scientists And Engineers* (Wiley (ed.)). https://doi.org/10.1007/978-3-319-03762-2__1
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2018). Complementarities between information governance and big data analytics capabilities on

- innovation. *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*.
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information and Management*, *57*(2). <https://doi.org/10.1016/j.im.2019.05.004>
- Nagappan, M., Sam, S., Sangeetha, S., Nithya Priya, S., Suguna, N., & Scholar, U. G. (2019). Heart Disease Prediction Using Data Mining Technique. In *Shodhshauryam, International Scientific Refereed Research Journal © 2019 SHISRRJ* (Vol. 2, Issue 10). www.shisrrj.com
- Olszak, C. M. (2013). Assessment of business intelligence maturity in the selected organizations. *2013 Federated Conference on Computer Science and Information Systems*, 951–958.
- Partners New Vantage, L. (2019). *Big data and AI executive survey 2019. Data and Innovation. How Big Data and AI are Accelerating Business Transformation*. <http://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf>
- Prieto, R., Meneses, C., & Vega, V. (2015). Análisis comparativo de modelos de madurez en inteligencia de negocio. *Ingeniare. Revista Chilena de Ingeniería*, *23*(3), 361–371. <https://doi.org/10.4067/S0718-33052015000300005>
- Ramamurthy, A. (2017). Effective information management - A - big- data driven road map for enterprise decision making. *Water Environment Federation Technical Exhibition and Conference 2017, WEFTEC 2017*, *3*, 1854–1865.
- Ramesh, G. S., Rajini Kanth, T. V., & Vasumathi, D. (2020). *A Comparative Study of Data Mining Tools and Techniques for Business Intelligence* (Springer (ed.); pp. 163–173). https://doi.org/10.1007/978-981-13-8253-6_15
- Ramsay, J. O. (2006). Functional Data Analysis. In *Encyclopedia of Statistical Sciences*. John Wiley & Sons, Inc. <https://doi.org/10.1002/0471667196.ess3138>
- Reio, T. G. (2016). Nonexperimental research: strengths, weaknesses and issues

- of precision. *European Journal of Training and Development*, 40(8/9), 676–690. <https://doi.org/10.1108/EJTD-07-2015-0058>
- Rutkowski, L., Jaworski, M., & Duda, P. (2020). *Basic Concepts of Data Stream Mining* (pp. 13–33). https://doi.org/10.1007/978-3-030-13962-9_2
- Sampieri, H. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta* (M. Hill (ed.)).
- Schab, E., Rivera, R., Bracco, L., Coto, F., Cristaldo, P., Ramos, L., Rapesta, N., Pablo Núñez, J., Retamar, S., Casanova, C., De Battista, A., & Herrera, N. E. (2018). Minería de Datos y Visualización de Información. *XX Workshop de Investigadores En Ciencias de La Computación*. <http://hdl.handle.net/20.500.12272/3567>
- Sen, D., Ozturk, M. & Vayvay, O. (2016). An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences*, 235(2016), 159–167. <https://doi.org/10.1016/j.sbspro.2016.11.011>
- Shaaban, E., Helmy, Y., Khedr, A., Nasr, M., & others. (2011). Business intelligence maturity models: Toward new integrated model. *The International Arab Conference on Information Technology (ACIT 2011), Organized by Naif Arab University for Security Science (NAUSS), Riyadh, Saudi Arabia*, 11–14.
- Shuradze, G., & Wagner, H. T. (2016). Governing for agility and innovation in data-rich environments: The role of data analytics capabilities. *24th European Conference on Information Systems, ECIS 2016*.
- Song, M., Zhang, H., & Heng, J. (2020). Creating sustainable innovativeness through big data and big data analytics capability: From the perspective of the information processing theory. *Sustainability (Switzerland)*, 12(5). <https://doi.org/10.3390/su12051984>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Pearson Education India.
- Tukey, J., & Wilk, M. (1966). Data analysis and statistics: an expository overview. *Proceedings of the November 7-10, 1966, Fall Joint Computer Conference*, 695–709. <https://dl.acm.org/doi/abs/10.1145/1464291.1464366>

- Villa, A., & Taurino, T. (2019). SME Innovation and Development in the Context of Industry 4.0. *Procedia Manufacturing*, 39, 1415–1420.
<https://doi.org/10.1016/j.promfg.2020.01.311>
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365.
- Weber, C., Königsberger, J., Kassner, L., & Mitschang, B. (2017). M2DDM – A Maturity Model for Data-Driven Manufacturing. *Procedia CIRP*, 63, 173–178.
<https://doi.org/10.1016/j.procir.2017.03.309>
- Xiao, X., Tian, Q., & Mao, H. (2020). How the Interaction of Big Data Analytics Capabilities and Digital Platform Capabilities Affects Service Innovation: A Dynamic Capabilities View. *IEEE Access*, 8, 18778–18796.
- Yasmin, M., Tatoglu, E., Kilic, H. S., Zaim, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Journal of Business Research*, 114, 1–15.
<https://doi.org/10.1016/j.jbusres.2020.03.028>
- Ye, F. Y. (2017). Measuring knowledge: A quantitative approach to knowledge theory. In *Understanding Complex Systems* (Issue 9789811059353, pp. 155–162). Springer Verlag. https://doi.org/10.1007/978-981-10-5936-0_13
- Alderson, P., Green, S., & Higgins, J. (2004). *Cochrane reviewers' handbook, version 4.2.2*. (The Cochrane Collaboration (ed.)).
https://www.mv.helsinki.fi/home/hemila/karlowski/handbook_4_2_2_Karlowski.pdf
- Bean, R., & Davenport, T. (2019). Companies are failing in their efforts to become data-driven. *Harvard Business Review*.
- Bedeley, R. T., & Nemati, H. (2014). Big Data Analytics: A key capability for competitive advantage. *20th Americas Conference on Information Systems, AMCIS 2014*.
<https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1536&context=amcis2014>
- Bihani, P., & Patil, S. (2014). A comparative study of data analysis techniques.

- International Journal of Emerging Trends & Technology in Computer Science*, 3(2), 95–101. www.ijettcs.org
- Bonilla, J., & Rao, B. (2015). Decoding data analytics capabilities from topic modeling on press releases. *Portland International Conference on Management of Engineering and Technology, 2015-Septe*, 1959–1968. <https://doi.org/10.1109/PICMET.2015.7273249>
- Bryman, A., & Cramer, D. (1992). Quantitative data analysis for social scientists. *Estudios Geográficos*, 53(207), 347. <http://search.proquest.com/openview/71ba8781f88a269938b860d448e6e8d2/1?pq-origsite=gscholar&cbl=1818083>
- Cates, J. E., Gill, S. S., & Zeituny, N. (2005). The Ladder of Business Intelligence (LOBI): a framework for enterprise IT planning and architecture. *International Journal of Business Information Systems*, 1(1/2), 220. <https://doi.org/10.1504/IJBIS.2005.007408>
- Coleman, S. (2016). Data-mining opportunities for small and medium enterprises with official statistics in the UK. *Journal of Official Statistics*. <https://doi.org/10.1515/JOS-2016-0044>
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Davenport, T., & Harris, J. G. (2017). *Competing on Analytics: The New Science of Winning*. Harvard Business School Press.
- Denning, D. E., Nicholson, W., Sande, G., & Shoshani, A. (1984). Research topics in statistical database management. *IEEE Database Eng. Bull.*, 7(1), 4–9.
- Dinter, B. (2012). The maturing of a business intelligence maturity model. *AMCIS 2012 Proceedings*. <http://elibrary.aisnet.org/Default.aspx?url=https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1083&context=amcis2012>
- Du, X., Liu, B., & Zhang, J. (2019). Application of Business Intelligence Based on Big Data in E-commerce Data Analysis. *Journal of Physics: Conference Series*, 1395(1), 012011. <https://doi.org/10.1088/1742-6596/1395/1/012011>

- Dulcé, H. J. (2016). Datos, información y conocimiento. *Respuestas*, 21(1), 4.
<https://revistas.ufps.edu.co/index.php/respuestas/article/download/642/646>
- Eason, K. D. (1989). *Information technology and organisational change*. CRC Press.
- Eckerson, W. (2007). TDWI benchmark guide: interpreting benchmark scores using TDWI's maturity model. *TDWI Research*, 3–14.
- Garmaki, M., & Boughzala, I. (2016). Conceptualization of Big data analytics Capability based on IT capability: Primary findings. *21st Symposium of the Association Information and Management 2016, AIM 2016*.
- Grossman, R. L. (2018). A framework for evaluating the analytic maturity of an organization. *International Journal of Information Management*, 38(1), 45–51.
<https://doi.org/10.1016/j.ijinfomgt.2017.08.005>
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049–1064.
- Halper, F. (2020). *TDWI Analytics Maturity Model Guide: Assessment Guide*.
- Halper, F., & Stodder, D. (2014). TDWI Analytics Maturity Model Guide. In *TDWI Research*. https://www3.microstrategy.com/getmedia/9b914607-084f-4869-ae64-e0b3f9e003de/TDWI_Analytics-Maturity-Guide_2014-2015.pdf
- Hao, S., Zhang, H., & Song, M. (2019). Big data, big data analytics capability, and sustainable innovation performance. *Sustainability*, 11(24), 7145.
- Hatta, N. N. M., Miskon, S., Ali, N. M., Abdullah, N. S., Ahmad, N., Hashim, H., Alias, R. A., & Maarof, M. A. (2015). Business intelligence system adoption theories in SMES: A literature review. *ARPN Journal of Engineering and Applied Sciences*.
- Hicks, S. C., & Peng, R. D. (2019). Elements and Principles for Characterizing Variation between Data Analyses. *ArXiv Preprint ArXiv:1903.07639*.
<http://arxiv.org/abs/1903.07639>
- Hirvonen, J., & Majuri, M. (2020). Digital capabilities in manufacturing SMEs. *Procedia Manufacturing*, 51, 1283–1289.
<https://doi.org/10.1016/j.promfg.2020.10.179>

- Humphrey, W. S. (1989). *Managing the software process*. Addison-Wesley Longman Publishing Co., Inc.
- Irwin, S. (2008). Data analysis and interpretation: emergent issues in linking qualitative and quantitative evidence. In Guilford Publications (Ed.), *Handbook of emergent methods in social research* (pp. 415–435).
- Islam, M. (2020). Data Analysis: Types, Process, Methods, Techniques and Tools. *International Journal on Data Science and Technology*, 6(1), 10.
<https://doi.org/10.11648/j.ijdst.20200601.12>
- Joyanes, L. (2013). *Big Data, Análisis de grandes volúmenes de datos en organizaciones* (Alfaomega Grupo Editor (ed.)).
- Kamioka, T., Hosoya, R., & Tapanainen, T. (2017). Effects of User IT Capabilities and Organized Big Data Analytics on Competitive Advantage. *PACIS*, 36.
- Khan, K. S., Ter Riet, G., Glanville, J., Sowden, A. J., & Kleijnen, J. (2001). *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University, UK and National ICT Australia*. <https://doi.org/10.1.1.122.3308>
- Kumar, V., Goyal, P., & Vandana, R. J. G. (2017). Stakeholder Classification: A Sustainability Marketing Perspective. *EVIDENCE BASED MANAGEMENT*, 111.
- Ladley, J. (2010). *Making enterprise information management (EIM) work for business: A guide to understanding information as an asset*. Morgan Kaufmann.
- Letón, M., & Pedromingo, A. (2001). *Introducción al Análisis de Datos en Meta-Análisis*. Ediciones Díaz de Santos.
- Luhn, H. P. (1958). A Business Intelligence System. *IBM Journal of Research and Development*, 2(4), 314–319. <https://doi.org/10.1147/rd.24.0314>
- Martínez, C. (2014). *Técnicas e Instrumentos de Recogida y Análisis de Datos*. Editorial UNED.
- Marulanda, C. E., López, M., & Mejía, M. H. (2013). Minería de datos en gestión

- del conocimiento de pymes de Colombia. *Revista Virtual Universidad Católica Del Norte*, 1(38), 158-170–170.
<https://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/821/1339>
- Masood, T., & Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, 121, 103261.
<https://doi.org/10.1016/j.compind.2020.103261>
- Maxwell, J. A., & Chmiel, M. (2014). Notes toward a theory of qualitative data analysis. In *The SAGE handbook of qualitative data analysis* (pp. 21–34). Sage Thousand Oaks, CA.
- Meyer, S. L. (1975). *Data Analysis For Scientists And Engineers* (Wiley (ed.)).
https://doi.org/10.1007/978-3-319-03762-2__1
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2018). Complementarities between information governance and big data analytics capabilities on innovation. *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*.
- Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information and Management*, 57(2). <https://doi.org/10.1016/j.im.2019.05.004>
- Nagappan, M., Sam, S., Sangeetha, S., Nithya Priya, S., Suguna, N., & Scholar, U. G. (2019). Heart Disease Prediction Using Data Mining Technique. In *Shodhshauryam, International Scientific Refereed Research Journal © 2019 SHISRRJ* (Vol. 2, Issue 10). www.shisrrj.com
- Olszak, C. M. (2013). Assessment of business intelligence maturity in the selected organizations. *2013 Federated Conference on Computer Science and Information Systems*, 951–958.
- Partners New Vantage, L. (2019). *Big data and AI executive survey 2019. Data and Innovation. How Big Data and AI are Accelerating Business Transformation*. <http://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf>

- Prieto, R., Meneses, C., & Vega, V. (2015). Análisis comparativo de modelos de madurez en inteligencia de negocio. *Ingeniare. Revista Chilena de Ingeniería*, 23(3), 361–371. <https://doi.org/10.4067/S0718-33052015000300005>
- Ramamurthy, A. (2017). Effective information management - A - big- data driven road map for enterprise decision making. *Water Environment Federation Technical Exhibition and Conference 2017, WEFTEC 2017*, 3, 1854–1865.
- Ramesh, G. S., Rajini Kanth, T. V., & Vasumathi, D. (2020). *A Comparative Study of Data Mining Tools and Techniques for Business Intelligence* (Springer (ed.); pp. 163–173). https://doi.org/10.1007/978-981-13-8253-6_15
- Ramsay, J. O. (2006). Functional Data Analysis. In *Encyclopedia of Statistical Sciences*. John Wiley & Sons, Inc. <https://doi.org/10.1002/0471667196.ess3138>
- Reio, T. G. (2016). Nonexperimental research: strengths, weaknesses and issues of precision. *European Journal of Training and Development*, 40(8/9), 676–690. <https://doi.org/10.1108/EJTD-07-2015-0058>
- Rutkowski, L., Jaworski, M., & Duda, P. (2020). *Basic Concepts of Data Stream Mining* (pp. 13–33). https://doi.org/10.1007/978-3-030-13962-9_2
- Sampieri, H. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta* (M. Hill (ed.)).
- Schab, E., Rivera, R., Bracco, L., Coto, F., Cristaldo, P., Ramos, L., Rapesta, N., Pablo Núñez, J., Retamar, S., Casanova, C., De Battista, A., & Herrera, N. E. (2018). Minería de Datos y Visualización de Información. *XX Workshop de Investigadores En Ciencias de La Computación*. <http://hdl.handle.net/20.500.12272/3567>
- Sen, D., Ozturk, M. & Vayvay, O. (2016). An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences*, 235(2016), 159–167. <https://doi.org/10.1016/j.sbspro.2016.11.011>
- Shaaban, E., Helmy, Y., Khedr, A., Nasr, M., & others. (2011). Business intelligence maturity models: Toward new integrated model. *The International Arab Conference on Information Technology (ACIT 2011)*, Organized by Naif

- Arab University for Security Science (NAUSS), Riyadh, Saudi Arabia*, 11–14.
- Shuradze, G., & Wagner, H. T. (2016). Governing for agility and innovation in data-rich environments: The role of data analytics capabilities. *24th European Conference on Information Systems, ECIS 2016*.
- Song, M., Zhang, H., & Heng, J. (2020). Creating sustainable innovativeness through big data and big data analytics capability: From the perspective of the information processing theory. *Sustainability (Switzerland)*, *12*(5).
<https://doi.org/10.3390/su12051984>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Pearson Education India.
- Tukey, J., & Wilk, M. (1966). Data analysis and statistics: an expository overview. *Proceedings of the November 7-10, 1966, Fall Joint Computer Conference*, 695–709. <https://dl.acm.org/doi/abs/10.1145/1464291.1464366>
- Villa, A., & Taurino, T. (2019). SME Innovation and Development in the Context of Industry 4.0. *Procedia Manufacturing*, *39*, 1415–1420.
<https://doi.org/10.1016/j.promfg.2020.01.311>
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365.
- Weber, C., Königsberger, J., Kassner, L., & Mitschang, B. (2017). M2DDM – A Maturity Model for Data-Driven Manufacturing. *Procedia CIRP*, *63*, 173–178.
<https://doi.org/10.1016/j.procir.2017.03.309>
- Xiao, X., Tian, Q., & Mao, H. (2020). How the Interaction of Big Data Analytics Capabilities and Digital Platform Capabilities Affects Service Innovation: A Dynamic Capabilities View. *IEEE Access*, *8*, 18778–18796.
- Yasmin, M., Tatoglu, E., Kilic, H. S., Zaim, S., & Delen, D. (2020). Big data analytics capabilities and firm performance: An integrated MCDM approach. *Journal of Business Research*, *114*, 1–15.
<https://doi.org/10.1016/j.jbusres.2020.03.028>
- Ye, F. Y. (2017). Measuring knowledge: A quantitative approach to knowledge

theory. In *Understanding Complex Systems* (Issue 9789811059353, pp. 155–162). Springer Verlag. https://doi.org/10.1007/978-981-10-5936-0_13

A. Anexo: Cuestionario aplicado a las empresas

Aplicación del modelo TDWI para la evaluación de la capacidad de análisis de datos para las pymes desarrolladoras de software



La Universidad Nacional de Colombia a través del grupo de investigación GRIEGO ha venido aplicando el "Modelo TDWI para la evaluar la capacidad de análisis de datos en las pymes desarrolladoras de software", con el objetivo de identificar fortalezas y debilidades y proponer recomendaciones para fomentar la capacidad de análisis de datos.

Para cumplir con el objetivo propuesto hemos acondicionado el siguiente cuestionario con las siguientes secciones:

1. Caracterización.
2. Organización.
3. Recursos.
4. Infraestructura de datos
5. Capacidad de análisis
6. Gobierno

El resultado del grado de madurez de la capacidad de análisis de datos le será enviado vía email, dentro de los cinco días hábiles posteriores al diligenciamiento por su parte de esta encuesta. El resultado del sector de software en el ámbito nacional, será enviado al finalizar el estudio en aproximadamente tres meses.

El tiempo de diligenciamiento del documento es de aproximadamente quince (17) minutos.

El grupo de investigación agradece su disposición para participar, la información que se proporcione se utilizará exclusivamente para fines de investigación y será estrictamente confidencial.

Agradecemos su interés y participación.

Aviso de confidencialidad

La información recolectada servirá únicamente para los fines de este estudio. El origen de los datos y el análisis de las respuestas atenderá a procesos de caracterización de las prácticas en análisis de datos de la industria, y se mantendrá confidencialidad en las publicaciones y productos académicos asociados, sobre cualquier información societaria, técnica, jurídica, financiera, comercial, de mercado, estratégica, de productos, nuevas tecnologías, patentes, modelos de utilidad, diseños industriales, modelos de negocios y/o cualquier otra relacionada de las organizaciones encuestadas.

Dirección de correo electrónico *

Texto de respuesta corta
.....

Nombre de la persona que responde la encuesta *

Texto de respuesta corta
.....

Posición o cargo que ocupa en la empresa *

Texto de respuesta corta
.....

Nombre de la empresa *

Texto de respuesta corta
.....

¿Cuáles fueron los ingresos de su empresa en el último año? *

1. Menos de \$1.130 millones
2. Entre \$1.130,1 millones a \$4.520 millones
3. Entre \$4.520,1 millones hasta \$16.500 millones
4. Más de \$16.500 millones



Región donde está localizada su empresa *

- Caribe: Atlántico, Bolívar, Cesar, Córdoba, La Guajira, Magdalena, Sucre y el Archipiélago de San Andrés, Pr...
- Centro-Oriente Y Bogotá D.C: Boyacá, Cundinamarca, Norte de Santander, Santander y Bogotá D. C
- Centro-Sur-Amazonía: Tolima, Huila, Caquetá, Putumayo y Amazonas
- Eje cafetero y Antioquia: Caldas, Quindío, Risaralda y Antioquia.
- Llanos: Arauca, Casanare, Guainía, Guaviare, Meta, Vichada y Vaupés
- Pacífico: Cauca, Chocó, Nariño y Valle del Cauca.

Ciudad *

Texto de respuesta corta

Sección 2 de 6

Organización



Descripción (opcional)



1. ¿Su liderazgo apoya y promueve el análisis de datos en toda la empresa? *

- De ningún modo
- La empresa tiene múltiples conceptos de análisis de datos, pero no lo apoyan realmente.
- Apoyan los esfuerzos del análisis de datos y están empezando a promoverlo. Eso incluye el uso de la analí...
- Apoyan firmemente los esfuerzos del análisis de datos, lo utilizan para tomar decisiones y lo promueven e...

2. Su compañía tiene un Jefe en Análisis de datos que está a cargo de los esfuerzos de *

- No tenemos a nadie a cargo del análisis de datos en mi compañía.
- El analisis de datos está cargo del área de TI de la empresa
- Tenemos uno o más vicepresidentes o directores de análisis en mi empresa, que están a cargo de los anál...
- Tenemos un jefe de análisis de datos

3. ¿Su empresa cuenta con una sólida estrategia para respaldar sus esfuerzos de análisis de datos? *

- No, y no tenemos planes de hacerlo
- No, pero planeamos hacerlo en el próximo año
- Sí, estamos en el proceso de armar una estrategia
- Sí, tenemos una sólida estrategia en marcha para el análisis de datos

4. ¿El análisis de datos es una parte importante de la estrategia de transformación digital de su empresa? *

- No, no tenemos una estrategia de transformación digital
- Sí, estamos en el proceso de unir nuestra estrategia de transformación digital con nuestra estrategias en ...
- Sí, el análisis de datos es una parte importante de la estrategia de transformación digital de mi empresa

5. ¿Cuál es el porcentaje de unidades/áreas de negocio en su empresa que utiliza el análisis de datos para la toma de decisiones del día a día? *

- Menos del 25%
- 26-40%
- 41-55%
- 56-70%
- Más del 70%

⋮

6. ¿Su organización ha medido el impacto del análisis de datos en general? *

- No, no hemos medido ningún impacto
- No, no hemos medido ningún impacto, pero creemos que hemos ganado valor
- Sí, hemos medido el impacto

7. ¿Su organización utiliza el análisis de datos para tomar decisiones? *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

⋮

8. ¿Hay una cultura de confianza en el análisis de datos en toda su empresa? *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

9. Hay una fuerte colaboración en materia de análisis de datos en su organización *

- No, no hacemos análisis de datos
- No, no colaboramos en el análisis de datos. El departamento de informática está a cargo
- Todavía no, pero nos estamos moviendo en esa dirección
- Sí, el área comercial y de informática trabajan regularmente juntos como deben
- Sí, el área comercial, de informática y otros trabajan juntos porque quieren y ven su colaboración como úti...

...

10. Hay una cultura de innovación en su empresa que se extiende hasta el análisis de datos *

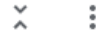
- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

11. Hay una fuerte base ética en su organización que se extiende hasta el análisis de datos *

- Completamente en desacuerdo
- En Desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

Sección 3 de 6

Infraestructura de datos



Descripción (opcional)

12. ¿Qué tipos de datos recopila su organización actualmente como parte de sus esfuerzos de análisis de datos? *

- Ninguno
- Datos estructurados (por ejemplo, registros de tablas) de fuentes internas solamente
- Datos estructurados más al menos uno de los siguientes: datos semiestructurados (por ejemplo, JSON, X...
- Datos estructurados más al menos tres de los siguientes: datos semiestructurados (por ejemplo, JSON, X...

13. Los empleados pueden acceder a los datos según sea necesario, incluidos los datos estructurados y no estructurados, mediante una plataforma de acceso unificada bien definida *

- No, en absoluto
- Sólo si pasan por el departamento TI
- Sí, la mayoría de los analistas comerciales y científicos de datos pueden acceder y hacer uso de los datos,...
- Sí, utilizamos tecnología como las opciones de intercambio de datos para ayudar a organizar y acceder a l...

14. Su organización a menudo hace uso de múltiples fuentes de datos para el análisis de datos *

- No
- Sí, con datos estructurados
- Sí, con diferentes tipos de datos, incluidos los datos semiestructurados y otras fuentes de datos no tradici...
- Sí, con diferentes tipos de datos y hacemos un buen trabajo integrándolos
- Sí, con datos estructurados, semiestructurados y no estructurados de fuentes internas y externas...son sól...

15. Su organización tiene una base de datos de confianza para el análisis de datos *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

⋮

16. Su organización está utilizando los siguientes tipos de tecnología para la gestión de datos: *

- Usamos sólo archivos planos u hojas de cálculo
- Tenemos un almacén de datos (data warehouse) o un data mart
- Usamos nuestro almacén de datos (data warehouse) junto con un lago de datos (data lake), pero están ai...
- Utilizamos una gama de tecnologías, incluyendo nuestro almacén de datos, lago de datos, nube, u otro tip...
- Utilizamos una gama de enfoques que forman una plataforma de datos bien diseñada para el acceso a los...



17. Su organización es capaz de orquestar y supervisar múltiples conductos o canales de datos *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo



18. Su organización cuenta con una arquitectura de datos de toda la empresa para el análisis que pueda soportar el crecimiento de los usuarios *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo



19. Su organización ha diseñado su arquitectura para integrar diversos datos de fuentes dispares * para su acceso y análisis

- Aun no
- Utilizamos múltiples fuentes, pero todos los análisis integrados utilizan datos estructurados
- Sí, e integramos tanto datos estructurados como semiestructurados porque son todos los datos necesari...



20. Su arquitectura de datos está diseñada para escalar *

- En realidad no
- Todavía no, pero estamos avanzando hacia una arquitectura más flexible y escalable
- Sí, confiamos en la capacidad de la compañía para escalar por caso de uso, incluyendo las fluctuaciones ...

21. Su arquitectura está diseñada a escala bajo demanda, para adaptarse a las necesidades del usuario *

- No
- Nos estamos moviendo en esa dirección
- Si

Sección 4 de 6

Recursos



Descripción (opcional)

22. Su organización tiene un proceso de financiación bien establecido tanto para la tecnología como para el análisis de datos. Está impulsado tanto por los negocios como por la tecnología de la información *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Estoy muy de acuerdo

23. La estrategia de analítica de su organización incluye un componente organizativo que permite a su organización ejecutar sus análisis de datos. Esto podría incluir la financiación de un centro de excelencia, equipos de innovación y similares *

- No, y no estoy seguro de que sepamos qué es eso
- Todavía no, pero sabemos que esto es importante y algunas personas quieren esto
- Sí, estamos en el proceso de organizarnos para ejecutar
- Sí, tenemos grupos/equipos como este y estamos trabajando para ampliarlo
- Sí, tenemos una inversión significativa en este tipo de cosas que incluye la formación y el apoyo a las inici...

...

24. Su empresa invierte en iniciativas de gestión del cambio *

- No, y no tenemos planes de hacerlo
- No, pero planeamos hacerlo en el próximo año
- Sí, estamos en el proceso de hacer esto ahora
- Sí, tenemos eso en marcha para proporcionar formación en gestión del cambio, pero es sólo para los ejec...
- Sí, hemos puesto eso en marcha en toda la organización

25. Su equipo está esforzándose por mantener su infraestructura de datos para apoyar el análisis de datos *

- Sí, es una dificultad; hay muy pocos recursos
- Estamos tratando de trabajar de forma más inteligente y utilizar la tecnología para ayudar a impulsar la pr...
- Estamos trabajando de forma más inteligente y tenemos los recursos que necesitamos

26. Su compañía ha contratado científicos de datos como parte de sus esfuerzos de análisis para mejorar sus capacidades de analítica o de análisis de datos *

- No, y no tenemos planes de hacerlo
- No, pero estamos planeando hacer esto dentro del próximo año
- Sí, hemos contratado a algunos científicos de datos
- Sí, nuestros científicos de datos son parte del equipo de análisis de datos
- Sí, nuestros científicos de datos son parte del equipo de análisis de datos y colaboran con la empresa

⋮

27. Su organización emplea a una gama de personal para tratar los diferentes aspectos del ciclo de vida de los análisis. Esto incluye ingenieros de datos y equipos de operaciones, por ejemplo, para ocuparse del análisis en producción *

- No, y no estoy seguro de que estemos pensando en ingenieros de datos o DevOps
- No, pero nos damos cuenta de que esto es importante y podemos tratar de trabajar en ello de manera ad h...
- Estamos poniendo un grupo/personal dedicado a esto
- Sí, tenemos miembros del equipo dedicados con un mandato y habilidades específicas para esto.

28. Su organización cuenta con un equipo para ejecutar la gestión de datos para su análisis *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

⋮

29. Su organización cuenta con un equipo con la capacidad para ejecutar análisis de datos *

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

30. Su empresa cree que puede mejorar las habilidades de sus analistas de negocios o comerciales para convertirse en científicos de datos *

- No, no estamos en el punto en el que necesitamos científicos de datos
- Sí, pero necesitarán la ayuda de otros
- Sí, pueden construir modelos, especialmente con las herramientas fáciles de usar del mercado
- Tenemos todos los datos científicos que necesitamos

⋮

31. Su organización es experta en el manejo de datos. Los usuarios comerciales, así como los analistas comerciales o de negocios, pueden utilizar los datos para obtener información. *

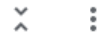
- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

32. Su organización invierte en formación para análisis de datos *

- No, y no tenemos planes de hacerlo
- No, pero sugerimos que el personal lea y aprendan ellos mismos
- Sí, financiamos únicamente la formación interna
- Sí, programamos regularmente capacitaciones financiadas y alentamos, a los empleados a asistir para as...

Sección 5 de 6

Capacidad de análisis



Descripción (opcional)



33. ¿Cuál de las siguientes tecnologías utiliza su organización para analizar sus datos? *

- No realizamos análisis de datos
- Usamos hojas de cálculo
- Usamos hojas de cálculo junto con informes, paneles y descubrimiento visual
- Usamos lo anterior junto con el descubrimiento de datos de autoservicio; también estamos empezando a ...
- Usamos lo anterior junto con análisis predictivo / aprendizaje automático (machine learning) contra múlti...
- Usamos lo anterior junto con otras técnicas como NLP (Procesamiento de lenguaje natural), aprendizaje p...

34. Su organización maneja grandes volúmenes de datos para el análisis de los mismos (por ejemplo, más de 5 TB) *

- No
- Todavía no, pero vamos en esa dirección
- Sí, utilizamos la analítica contra grandes volúmenes de datos

⋮

35. ¿Cuántos modelos de análisis predictivo / aprendizaje automático tiene su empresa en producción? *

- Ninguno
- 1
- 2 a 10
- Docenas
- Cientos

36. ¿Qué porcentaje de personas en su empresa tiene acceso a análisis de datos? *

- Menos del 25%
- 26-40%
- 41-55%
- 56-70%
- Mayor al 70%

⋮

37. Su organización tiene éxito en la articulación de problemas comerciales que requieren análisis de datos. Sabe cuándo utilizar técnicas específicas para resolver diferentes problemas *

- No tenemos ningún éxito en hacer esto
- No tenemos mucho éxito en hacer esto
- No tenemos ni éxito ni fracaso en hacer esto
- Tenemos éxito en hacer esto
- Tenemos mucho éxito en hacer esto

38. El análisis de datos es utilizado por los equipos de toda la organización, cuando es necesario *

- No, se utiliza principalmente en finanzas, operaciones o TI
- Actualmente no, pero más organizaciones se están interesando en él
- Sí, utilizamos análisis en toda la organización

⋮

39. Sus soluciones de análisis de datos están impulsadas por la persona para proporcionar la mejor interfaz de usuario a la persona adecuada (por ejemplo, analistas comerciales, usuarios comerciales, científicos de datos, ingenieros de datos, etc.). Por ejemplo, los científicos de datos pueden usar cuadernos de ciencia de datos, mientras que los analistas de negocios pueden preferir las GUI de arrastrar, soltar y navegar, y los usuarios en general pueden querer consumir paneles e informes o utilizar acciones automatizadas sugeridas por sus modelos *

- No, solo tenemos una herramienta y todos se quejan de ella, ya que no es adecuada para nadie
- Estamos pensando o planeando instalar herramientas que puedan ser utilizadas por diferentes personas
- Sí, utilizamos múltiples herramientas para que cada una de nuestras personas tenga el entorno adecuado
- Sí, utilizamos una plataforma de datos para que cada persona tenga un entorno que maximice la eficienci...
- Igual que la anterior y también nuestra plataforma hace que el intercambio de datos entre entornos sea lo ...

40. Su organización utiliza análisis de datos automatizados (por ejemplo, sistemas que construyen modelos) *

- No
- No, pero estamos explorando estas soluciones
- Los usamos para democratizar la analítica y aumentar la productividad
- Sí, utilizamos una plataforma de datos para ayudar a impulsar la productividad y ponemos controles en to...

41. Su organización utiliza tecnologías de código abierto para el análisis de datos *

- Sí, solo usamos código abierto
- Sí, usamos el código abierto junto con productos comerciales; apoyamos cualquier cosa que el científico/...
- No, solo usamos productos comerciales

⋮

42. Los análisis se ponen en funcionamiento/despliegan en un sistema o sistemas empresariales *
o en una o varias aplicaciones de su organización

- No, y no tenemos planes de hacerlo
- No, pero estamos planeando hacer esto el próximo año
- Sí, estamos intentando hacer esto ahora en tableros (dashboards) o cuadros de mando
- Sí, hacemos esto de forma rutinaria en aplicaciones de tablero o cuadro de mando (dashboards)
- Sí, rutinariamente ponemos en marcha nuestros análisis, incluyendo los cuadros de mando (dashboards) ...

⋮

43. Los resultados de los modelos construidos utilizando inteligencia aumentada tienen
elementos de explicación para aumentar la transparencia tanto para los expertos como para los
no expertos *

- No aplica, no estamos usando estas herramientas
- No, no que hayamos visto
- Sí, solo usamos paquetes que tienen estas características



44. Su organización monitorea los modelos de análisis predictivo / aprendizaje de máquinas (machine learning) para detectar su deterioro. *

- No aplica, no tenemos modelos en producción
- No, y no tenemos planes de hacerlo
- Todavía no, pero lo estamos pensando
- Sí, estamos intentando hacer esto ahora
- Sí, hacemos esto con regularidad
- Sí, lo hacemos habitualmente con controles automáticos y asignamos el tiempo necesario para que nuest...

Sección 6 de 6

Gobierno



Descripción (opcional)



45. Los datos son confiables y se controlan a través de las plataformas de su organización *

- No, tenemos muchos silos de datos que no están controlados
- Confiamos en los datos que usamos para informar que vienen de nuestra bodega de datos (DW), pero no ...
- Estamos empezando a establecer procesos para el gobierno de los datos más allá del DW u otras fuentes ...
- Tenemos un sólido plan de gestión de datos que esboza las políticas y procesos clave; estos se siguen en ...

46. Su organización comprende la (s) fuente (s) de sus datos y cuenta con las políticas adecuadas para manejar diferentes tipos de datos *

- Totalmente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

⋮

47. Los usuarios aceptan y cumplen las políticas de gobierno de datos *

- Totalmente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

48. Su organización utiliza herramientas como catálogos de datos para ayudar a los usuarios a acceder a datos confiables *

- No, y no tenemos planes de instalar un catálogo de datos
- No, pero lo estamos pensando
- Estamos en el proceso de seleccionar un proveedor de catálogo ahora
- Sí, tenemos un catálogo de datos y la gente ha aceptado usarlo
- Sí, tenemos un catálogo de datos, pero no todo el mundo lo usa

49. Los procesos de despliegue del modelo de análisis predictivo / machine learning están en marcha en su organización. Por ejemplo, los modelos deben ser comprobados para que no sean incorrectos o poco éticos (por ejemplo, que no tengan sesgos raciales) antes de ser puestos en producción. *

- No aplica, no tenemos modelos en producción en mi empresa
- Tenemos modelos implementados, pero no verificamos si son correctos. Confiamos en nuestros científic...
- Estamos en proceso de implementar controles sobre modelos en producción
- Contamos con un sólido proceso de control de modelos

50. En su organización existen políticas modelo de gestión. Los modelos deben ser controlados por versión y los metadatos capturados para cada modelo puesto en producción *

- No aplicable - no tenemos modelos en producción
- Solo tenemos unos pocos modelos, por lo que nuestros científicos de datos se ocupan de ellos
- Tenemos un enfoque de sistema de archivos para realizar un seguimiento de los modelos y ocuparnos del...
- Estamos empezando a utilizar registros de modelos y otras tecnologías para capturar metadatos sobre m...
- Utilizamos herramientas y tecnologías para gestionar nuestros modelos. Los modelos están versionados ...

51. Su empresa tiene un equipo de gestión de datos y análisis con representantes de toda la empresa, incluidos los principales interesados. Las funciones y responsabilidades están claramente definidas *

- Totalmente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

⋮

52. El rol de administrador (es) de datos está establecido y los roles y responsabilidades de esa persona (o equipo) están claramente identificados *

- No
- Estamos en proceso de identificar a los administradores de datos
- Hay administradores de datos en mi empresa

53. Las políticas de seguridad se establecen y se cumplen para todas las formas de datos en su empresa *

- No
- Los almacenes de datos (datawarehouse) están protegidos y controlados, pero no necesariamente en fue...
- Sí, existen políticas de seguridad para todo tipo de datos sensibles
- Sí, hemos pensado detenidamente cómo tratamos los diferentes tipos de datos en nuestro equipo de gobi...
- Sí, hemos pensado detenidamente y hemos puesto en práctica cómo tratamos los diferentes tipos de dat...

B. Anexo: Informe final enviado a las empresas que participaron en la investigación

Resultados individual de evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá >



Carlos Aurelio Angulo Romero <caangulor@unal.edu.co>
para piatorresv ▾

📧 11:53 (hace 2 horas)

Reciban un cordial saludo en nombre del Grupo de Investigación en Gestión y Organizaciones - GRIEGO de la Universidad Nacional de Colombia.

Compartimos el resultado individual de la evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá.

Cualquier inquietud o duda con gusto será atendida.

Cordialmente,

--

Carlos Angulo
Estudiante de maestría Universidad Nacional de Colombia
Universidad Nacional de Colombia
caangulor@unal.edu.co



Resultados individual de evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá >



Carlos Aurelio Angulo Romero <caangulor@unal.edu.co>
para jose.palacios ▾

📧 11:47 (hace 2 horas)

Reciban un cordial saludo en nombre del Grupo de Investigación en Gestión y Organizaciones - GRIEGO de la Universidad Nacional de Colombia.

Compartimos el resultado individual de la evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá.

Cualquier inquietud o duda con gusto será atendida.

Cordialmente,

--

Carlos Angulo
Estudiante de maestría Universidad Nacional de Colombia
Universidad Nacional de Colombia
caangulor@unal.edu.co



Resultados individual de evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá >



Carlos Aurelio Angulo Romero <caangulor@unal.edu.co>
para fairmoon ▾

11:45 (hace 2 horas)

Reciban un cordial saludo en nombre del Grupo de Investigación en Gestión y Organizaciones - GRIEGO de la Universidad Nacional de Colombia.

Compartimos el resultado individual de la evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá.

Cualquier inquietud o duda con gusto será atendida.

Cordialmente,

--

Carlos Angulo
Estudiante de maestría Universidad Nacional de Colombia
Universidad Nacional de Colombia
caangulor@unal.edu.co



Resultados individual de evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá >



Carlos Aurelio Angulo Romero <caangulor@unal.edu.co>
para leonardosanchez24 ▾

11:38 (hace 3 horas)

Reciban un cordial saludo en nombre del Grupo de Investigación en Gestión y Organizaciones - GRIEGO de la Universidad Nacional de Colombia.

Compartimos el resultado individual de la evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá.

Cualquier inquietud o duda con gusto será atendida.

Cordialmente,

--

Carlos Angulo
Estudiante de maestría Universidad Nacional de Colombia
Universidad Nacional de Colombia
caangulor@unal.edu.co



Resultados individual de evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá >



Carlos Aurelio Angulo Romero <caangulor@unal.edu.co>
para gcanas ▾

11:34 (hace 3 horas)

Reciban un cordial saludo en nombre del Grupo de Investigación en Gestión y Organizaciones - GRIEGO de la Universidad Nacional de Colombia.

Compartimos el resultado individual de la evaluación de la capacidad de análisis de datos de las pymes desarrolladoras de software de la ciudad de Bogotá.

Cualquier inquietud o duda con gusto será atendida.

Cordialmente,

--

Carlos Angulo
Estudiante de maestría Universidad Nacional de Colombia
Universidad Nacional de Colombia
caangulor@unal.edu.co

