



Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Estudios de Postgrado

Maestría en Artes en Tecnologías de la Información y la
Comunicación

**DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA
ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS, A TRAVÉS
DE APRENDIZAJE AUTOMÁTICO Y REINGENIERÍA DE PROCESOS**

Ing. Juan Fernando Cano Villatoro

Asesorado por la Inga. MBA. Sasha Steffanie Palencia Zetina

Guatemala, noviembre de 2017

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA
ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS A TRAVÉS
DE APRENDIZAJE AUTOMÁTICO Y REINGENIERIA DE PROCESOS**

TRABAJO DE GRADUACIÓN

PRESENTADO A JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

JUANFERNANDO CANO VILLATORO

ASESORADO POR LA INGA. MBA. SASHA STEFFANIE PALENCIA
ZETINA

AL CONFERÍRSELE EL TÍTULO DE

**MAESTRO EN ARTES EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA
COMUNICACIÓN**

GUATEMALA, NOVIEMBRE DE 2017

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Pedro Antonio Aguilar Polanco
VOCAL I	Ing. Ángel Roberto Sic García
VOCAL II	Ing. Pablo Christian De León Rodríguez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Jurgen Andoni Ramírez Ramírez
VOCAL V	Br. Oscar Humberto Galicia Núñez
SECRETARIA	Inga. Lesbia Magalí Herrera López

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

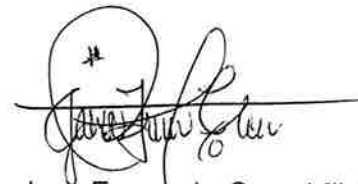
DECANO	Ing. Pedro Antonio Aguilar Polanco
EXAMINADORA	Inga. María Elizabeth Aldana Díaz
EXAMINADOR	Ing. Marlon Antonio Pérez Turk
EXAMINADOR	Ing. Hugo Humberto Rivera Pérez
SECRETARIA	Inga. Lesbia Magalí Herrera López

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS A TRAVÉS DE APRENDIZAJE AUTOMÁTICO Y REINGENIERIA DE PROCESOS

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Posgrado, con fecha agosto de 2016.

A handwritten signature in black ink, appearing to read 'Juan Fernando Cano Villatoro', written over a horizontal line.

Juan Fernando Cano Villatoro



FACULTAD DE
INGENIERÍA - USAC
ESCUELA DE
ESTUDIOS DE POSTGRADO


Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / 24188000 Ext. 86226

APT-2017-024

Como Coordinador de la Maestría en Artes en Tecnologías de la Información y la Comunicación del Trabajo de Graduación titulado **"DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS, A TRAVÉS DE APRENDIZAJE AUTOMÁTICO Y REINGENIERÍA DE PROCESOS"** presentado por el Ingeniero en Sistemas, Informática y Ciencias de la Computación **Juan Fernando Cano Villatoro**, apruebo y recomiendo la autorización del mismo.

Atentamente,

"Id y Enseñad a Todos"


MSc. Marlon Antonio Pérez
Coordinador de Maestría
Escuela de Estudios de Postgrado
Facultad de Ingeniería
Universidad de San Carlos de Guatemala



Guatemala, noviembre de 2017.

Cc: archivo/la

Doctorado: Sostenibilidad y Cambio Climático. **Programas de Maestrías:** Ingeniería Vial, Gestión Industrial, Estructuras, Energía y Ambiente Ingeniería Geotécnica, Ingeniería para el Desarrollo Municipal, Tecnologías de la Información y la Comunicación, Ingeniería de Mantenimiento. **Especializaciones:** Gestión del Talento Humano, Mercados Eléctricos, Investigación Científica, Educación virtual para el nivel superior, Administración y Mantenimiento Hospitalario, Neuropsicología y Neurociencia aplicada a la Industria, Enseñanza de la Matemática en el nivel superior, Estadística, Seguros y ciencias actuariales, Sistemas de información Geográfica, Sistemas de gestión de calidad, Explotación Minera, Catastro.



FACULTAD DE
INGENIERÍA - USAC
ES
ESCUELA DE
ESTUDIOS DE POSTGRADO

Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / 24188000 Ext. 86226

APT-2017-024

El Director de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen y dar el visto bueno del revisor y la aprobación del área de Lingüística del Trabajo de Graduación titulado **"DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS, A TRAVÉS DE APRENDIZAJE AUTOMÁTICO Y REINGENIERÍA DE PROCESOS"** presentado por el Ingeniero en Sistemas, informática y Ciencias de la Computación **Juan Fernando Cano Villatoro**, correspondiente al programa de Maestría en Artes en Tecnologías de la Información y la Comunicación; apruebo y autorizo el mismo.

Atentamente,

"Id y Enseñad a Todos"

MSc. Ing. Murphy Olympo Paiz Recinos

Director

Escuela de Estudios de Postgrado

Facultad de Ingeniería

Universidad de San Carlos de Guatemala



Guatemala, noviembre de 2017.

Cc: archivo/la

Doctorado: Sostenibilidad y Cambio Climático. **Programas de Maestrías:** Ingeniería Vial, Gestión Industrial, Estructuras, Energía y Ambiente Ingeniería Geotécnica, Ingeniería para el Desarrollo Municipal, Tecnologías de la Información y la Comunicación, Ingeniería de Mantenimiento. **Especializaciones:** Gestión del Talento Humano, Mercados Eléctricos, Investigación Científica, Educación virtual para el nivel superior, Administración y Mantenimiento Hospitalario, Neuropsicología y Neurociencia aplicada a la Industria, Enseñanza de la Matemática en el nivel superior, Estadística, Seguros y ciencias actuariales, Sistemas de información Geográfica, Sistemas de gestión de calidad, Explotación Minera, Catastro.



FACULTAD DE
INGENIERÍA - USAC
EP
ESCUELA DE
ESTUDIOS DE POSTGRADO

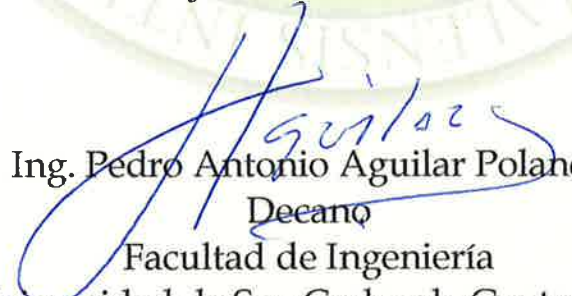
Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / 24188000 Ext. 86226

APT-2017-024

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Postgrado, al Trabajo de Graduación de la Maestría en Artes en Tecnologías de la Información y la Comunicación titulado: **"DISEÑO EXPERIMENTAL CON GROWTH HACKING PARA LA ESTANDARIZACIÓN EN LA INTERPRETACIÓN DE MÉTRICAS, A TRAVÉS DE APRENDIZAJE AUTOMÁTICO Y REINGENIERÍA DE PROCESOS"** presentado por el Ingeniero en Sistemas, Informática y Ciencias de la Computación **Juan Fernando Cano Villatoro**, procede a la autorización para la impresión del mismo.

IMPRÍMASE.

"Id y Enseñad a Todos"


Ing. Pedro Antonio Aguilar Polanco
Decano
Facultad de Ingeniería
Universidad de San Carlos de Guatemala



Guatemala, noviembre de 2017.

Cc: archivo/la

Doctorado: Sostenibilidad y Cambio Climático. **Programas de Maestrías:** Ingeniería Vial, Gestión Industrial, Estructuras, Energía y Ambiente Ingeniería Geotécnica, Ingeniería para el Desarrollo Municipal, Tecnologías de la Información y la Comunicación, Ingeniería de Mantenimiento. **Especializaciones:** Gestión del Talento Humano, Mercados Eléctricos, Investigación Científica, Educación virtual para el nivel superior, Administración y Mantenimiento Hospitalario, Neuropsicología y Neurociencia aplicada a la Industria, Enseñanza de la Matemática en el nivel superior, Estadística, Seguros y ciencias actuariales, Sistemas de información Geográfica, Sistemas de gestión de calidad, Explotación Minera, Catastro.

ACTO QUE DEDICO A:

- Dios** Por demostrar su fidelidad en mi vida, por darme salud y sabiduría para cumplir todas mis metas. Y por ser mi guía todos los días para vencer cualquier obstáculo, por esperar pacientemente en su presencia y ver sus promesas cumplirse.
- Mi madre** Por su gran amor y su apoyo incondicional. Por hacer de mi un ciudadano de bien, con todas sus enseñanzas y cuidados y sobre todo por sus buenos ejemplos y tan gratas vivencias en mi vida.
- Mi padre** Por su lucha incansable para darme lo mejor y su apoyo en el transcurso de mi carrera profesional. Por enseñarme humildad y responsabilidad para concluir todo proyecto que he emprendido en mi vida.
- Mis hermanas** Por levantarme cuando he caído, con palabras de amor y optimismo para animarme a alcanzar mis metas. Por compartir mis alegrías y consolarme en las tristezas.
- Mi familia** Gracias a todos por el apoyo que de alguna manera me brindaron, y en especial a mis tíos y tías, ya que son parte importante en mi vida.

Mis amigos

Parte importante en mi vida, por estar siempre cuando los necesitaba, por escucharme en momentos difíciles y por llenarme de alegrías y gestos de hermandad.

AGRADECIMIENTOS A:

- Dios** Porque estoy consciente que sin sus bendiciones, no estaría por culminar un logro más como profesional y como persona que ama aprender cada día. Por ser el pilar más importante de mis principios y por no soltar mi mano.
- Mi asesora** Por sus conocimientos, tiempo y dedicación para lograr culminar mi trabajo de graduación de manera adecuada, agradecimientos a la ingeniera Sasha Steffanie Palencia Zetina
- Mis catedráticos** Por compartir sus conocimientos a lo largo de la maestría y por su buena voluntad para enseñar con su amplia experiencia profesional, en especial a la ingeniera María Elizabeth Aldana Díaz, por sus consejos para realizar una investigación y experimentación de alto nivel.
- Universidad de San Carlos de Guatemala** Por su excelente formación académica y por el prestigio que vela por mantener, agradecimientos a la Facultad de Ingeniería y a la Escuela de Estudios de Postgrado.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
LISTA DE SÍMBOLOS	IX
GLOSARIO.....	X
RESUMEN	XIII
PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE PREGUNTAS ORIENTADORAS	XV
OBJETIVOS	XVII
MARCO METODOLÓGICO	XIX
INTRODUCCIÓN.....	XXV
1. ANTECEDENTES.....	1
2. JUSTIFICACIÓN.....	7
3. ALCANCES	11
3.1. Investigativos	11
3.2. Técnicos	11
3.3. Resultados	12
4. MARCO TEÓRICO	13
4.1. Growth Hacking	13
4.2. Growth Hacker	14
4.3. Embudo de conversión de Growth Hacking	14
4.4. Canales virales de Growth Hacking.....	16
4.5. Métricas de Growth Hacking	17
4.6. El cambio neto diario	17
4.7. Activos principales diarios	18
4.8. Planificación SOSTAC	18

4.8.1. Análisis de la situación actual	19
4.8.2. Objetivos claros	19
4.8.3. Estrategia de cómo llegar allí.....	19
4.8.4. Las tácticas son métodos específicos	20
4.8.5. Acciones.....	20
4.9. Control.....	20
4.10. Data Mining	21
4.11. Machine learning o aprendizaje automático.....	22
4.12. Clasificación	23
4.13. Regresión	23
4.14. Aplicación de R para machine learning	24
4.15. Clasificación de datos	25
4.16. Clustering	25
4.17. Publicidad en la web	26
5. PRESENTACIÓN DE RESULTADOS	27
5.1. Análisis, diseño de la arquitectura y desarrollo del prototipo experimental	27
5.1.1. Funcionalidad para promocionar publicación en la Fanpage de Facebook.....	28
5.2. Publicación de la Fanpage y sitio web	29
5.3. Exportación de datos a Microsoft Excel	30
5.4. Preparación de datos.....	30
5.4.1. Fuente de datos	31
5.4.1.1. Datos generados en Facebook	31
5.4.2. Comprensión de datos	31
5.4.2.1. Tablas de Datos	31
5.4.2.2. Variables continuas y discretas	32
5.4.2.3. Roles en el análisis	33
5.4.2.4. Distribución de frecuencia.....	34

5.4.3.	Preparación de datos	35
5.4.3.1.	Visión de conjunto	35
5.4.3.2.	Limpieza de los datos	35
5.4.3.3.	Eliminación de variables	35
5.4.3.4.	Mapeo de valores	36
5.4.4.	Tablas y gráficas.....	37
5.5.	Clustering.....	43
6.	DISCUSIÓN DE RESULTADOS.....	49
6.1.	Rendimiento del experimento	50
6.2.	Análisis de la oportunidad de negocio.....	51
6.3.	Herramientas utilizadas.....	52
	CONCLUSIONES	53
	RECOMENDACIONES	55
	REFERENCIAS BIBLIOGRÁFICAS	57
	ANEXOS	61

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Embudo de <i>Growth Hacking</i>	15
2.	Diagrama de arquitectura para la interpretación de métricas de <i>Growth Hacking</i>	28
3.	Imagen de funcionalidad para promocionar publicación.....	29
2.	Histograma de datos crudos.....	34
3.	Histograma de interacción de usuarios.....	36
4.	Gráficas de densidad.....	37
5.	Polígono de frecuencias de “Me gusta”.....	38
6.	Polígono de frecuencias de personas alcanzadas.....	40
7.	Histograma del alcance mensual vrs. la interacción mensual...	41
8.	Matriz de dispersión de métricas de <i>Growth Hacking</i>	42
9.	Matriz de dispersión del promedio de interacciones.....	43
10.	Banner del método simple.....	44
11.	Dendograma del método simple.....	44
12.	Banner del método completo.....	45
13.	Dendograma del método completo.....	45
14.	<i>Banner</i> del método promedio.....	46
15.	Dendograma del método promedio.....	46
16.	Ilustración de <i>Clusters y Outliers</i> del método completo.....	47
17.	Ilustración de <i>Clusters y Outliers</i> del método promedio.....	47
18.	Técnicas de <i>Growth Hacking</i>	49

TABLAS

I.	VARIABLES DE LA ARQUITECTURA.....	XX
II.	CATEGORIZACIÓN DE VARIABLES.....	32
III.	ROLES DE VARIABLES.....	33
IV.	TABLA DE SUMATORIA DE "ME GUSTA" SEMANAL PROMEDIO.....	37
V.	SUMATORIA DE PERSONAS ALCANZADAS MENSUAL PROMEDIO.....	39

LISTA DE SÍMBOLOS

Símbolo	Significado
CSV	<i>Comma-Separated Values</i>
XLSX	Extensión de archivo empleado en <i>Microsoft Excel</i> a partir de su versión 2007.
SAS	Nueva versión del interfaz SCSI compatible con SATA
XML	<i>eXtensible Markup Language</i>

GLOSARIO

<i>Alojamiento</i>	Servicio de internet que provee a los usuarios de almacenamiento en la nube de información como datos, imágenes, videos, entre otros.
<i>Clustering</i>	Significa agrupar los objetos en la información que se encuentra en los datos que describen los objetos o sus relaciones.
<i>Dendograma</i>	Es un tipo de representación gráfica o diagrama de datos en forma de árbol, que organiza los datos en subcategorías que se van dividiendo en otros hasta llegar al nivel de detalle deseado.
<i>Fanpage</i>	Páginas de <i>Fans</i> , son una red social creada por <i>Facebook</i> para todas las empresas o personas emprendedoras que desean crear y diseñar <i>marketing</i> , a través de ellas.
<i>Google Analytics</i>	Es una herramienta de analítica web de la empresa <i>Google</i> . Ofrece información agrupada del tráfico que llega a los sitios <i>web</i> , según la audiencia, la adquisición, el comportamiento y las conversiones que se llevan a cabo en el sitio <i>web</i> .

<i>Interacción</i>	Es un vocablo que describe una acción que se desarrolla de modo recíproco entre dos o más organismos, objetos, agentes, unidades, sistemas, fuerzas o funciones.
<i>Prototipo</i>	Es la primer versión que se lanza para dar a conocer un producto o servicio que reúne gran cantidad de características que poseerá la versión final.
<i>R Studio</i>	Es un conjunto de herramientas integradas diseñadas para el lenguaje de R. Incluye una consola, editor de resaltado de sintaxis que soporta la ejecución directa de código y una variedad de herramientas robustas para trazar.
<i>Red neuronal</i>	Son paradigmas de aprendizaje de procesamiento de información que se asemeja a las neuronas cerebrales, es decir, busca imitar el funcionamiento para un fin específico.
<i>Web</i>	Es un conjunto de páginas relacionadas, que se almacena en Internet y se acceden por medio de un dominio de internet.
<i>XML</i>	Es un lenguaje de etiquetado extensible que se caracteriza por ser legible creado por <i>World Wide Web Consortium</i> y su principal función es compartir datos.

RESUMEN

Aplicar las técnicas de *Growth Hacking* correctamente y realizar un proceso adecuado de interpretación de las métricas con metodologías de *Data Mining* como *Clustering* garantizarán cambios efectivos en los sitios de fans creados en *Facebook*, de acuerdo al análisis de datos y procesamiento con herramientas estadísticas.

La experimentación que se realizó, muestra datos importantes acerca de la interacción de usuarios en las distintas publicaciones, los rangos de edad y las preferencias en una larga colección de datos que al realizar un correcto análisis con *Data Mining*, apoya la toma de decisiones para promover una marca o promocionar una determinada empresa.

El alcance que posee *Facebook* es increíble, ofrece funcionalidades de promoción viral, únicamente debe seleccionarse el país o región y el número de personas a las cuales desea difundirse una determinada publicación, pero lo más efectivo es conocer el mercado y que aceptación tendrán las publicaciones y para esto se necesita básicamente la combinación de contenidos llamativos más las funcionalidades de *Facebook* y un análisis correcto de la interpretación de métricas, para lo cual se presenta la secuencia necesaria en la siguiente investigación experimental.

Las tecnologías utilizadas para la interpretación de métricas son la siguientes: *Fanpage* creada en *Facebook* para la recolección de datos y exportar datos a un archivo CSV o XLSX, procesamiento de los datos obtenidos con *R Studio*, obtener datos estadísticos importantes con QI macros en *Excel*,

así como el análisis de *Clustering* con librerías específicas de R Studio. Otra tecnología son distintas herramientas para el procesamiento de datos como *Mathlab*. Entre las ventajas que ofrece el uso de las herramientas sugeridas, es la facilidad de uso, debido a que durante la etapa de experimentación se realizaron distintas pruebas con diferentes herramientas y en la presentación de resultados se da a conocer el proceso y las herramientas que al final se seleccionaron para cada paso.

Las gráficas resultantes del análisis son tan importantes como los datos que se necesitan para generarlas, por ello, la preparación de los datos es un paso muy importante aunque un poco tedioso, para generar los resultados esperados. Se obtiene un 85 % de efectividad, debido a que se cumplieron las predicciones descritas en el estudio cualitativo previo; Cano, J. (2016). Técnicas de *Growth Hacking* que producen mayor viralidad en los sitios web y redes sociales más comunes de la actualidad. (Estudio cualitativo). Universidad de San Carlos de Guatemala, Guatemala. También debe tomarse en cuenta que la efectividad de la difusión de publicaciones aumentó, debido al proceso de los datos obtenidos del análisis de métricas de *Growth Hacking* y aplicando el método de *Clustering*. Por último, para validar los resultados, se realizaron publicaciones en la *Fanpage*, aplicando las métricas y técnicas resultantes de mayor interés y al hacer el análisis de *Clustering* nuevamente se obtuvo un incremento en las interacciones realizadas por los seguidores en el rango de edades que se había obtenido con anterioridad y con el género que logró mayor interacción.

PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE PREGUNTAS ORIENTADORAS

El *Growth Hacking* es un concepto que ha tomado relevancia por causa de la necesidad de encontrar formas de crecimiento sin necesidad de recurrir en gastos de *marketing* y surge apenas hace un par de años, por lo cual sus métodos no son conocidos y aún no se encuentran estandarizados de tal modo que puedan optimizarse. El *Growth Hacking* no es tan solo una combinación de herramientas de programación y *marketing* sino es reconocer un mercado y atacarlo de manera diferente a la que se hace con los métodos tradicionales, a través de generación de hipótesis, validación y escalabilidad de modelos de negocios, aplicándolo en sitios web y aplicaciones. (Ellis, 2010).

El *Growth Hacking* se basa en tres objetivos del Embudo de Conversión que son: Conseguir visitantes, activar miembros y retener usuarios por medio de métricas, donde cualquier técnica o estrategia que no pueda medirse, simplemente no es útil, pero medir es costoso y las empresas regularmente manejan bajo presupuesto para la medición e interpretación de las mismas. (Patel, 2013).

Lo anterior evidencia que una de las principales debilidades del método de *Growth Hacking* es la falta de estandarización e interpretación de las métricas, porque no es un método de secuencias y reglas y no cuenta con un estándar de las mismas, lo que facilita la mala utilización de las métricas y como consecuencias los sitios web o aplicaciones fracasan, debido a que sin las métricas adecuadas es difícil comprobar la efectividad y el impacto de los

coeficientes críticos para la aceptación de los usuarios (viralidad, facilidad de uso, accesibilidad, etc.)

También es importante mencionar que muchas veces se tienen las métricas correctas, pero no se sabe qué hacer con éstas, existen varias aplicaciones y sitios web que simplifican la captura de métricas, pero por ser un método sin estandarización y ciertamente confuso además de aplicarse en contenido diverso de sitios *web* y aplicaciones; todos estos factores mencionados anteriormente dificultan la interpretación y la estandarización del mismo y al momento de fracasar un sitio web o aplicación incurre en un gasto más para la empresa, sin aportar ningún beneficio.

Por lo expuesto anteriormente, se formula la siguiente pregunta principal:
¿Las técnicas de *Data Mining* y análisis de procesos pueden mejorar la estandarización e interpretación de métricas de *Growth Hacking* utilizadas en sitios web o aplicaciones?

Además para apoyar esta pregunta pueden proponerse las siguientes preguntas auxiliares:

¿Qué componente tecnológico se debe implementar en la arquitectura de *Growth Hacking*, para optimizar su procesamiento y efectividad?

¿Cómo se puede estandarizar la secuencia de los procesos del método de *Growth Hacking*?

¿Cómo puede determinarse el rendimiento y mejoras en la interpretación de métricas de *Growth Hacking*?

OBJETIVOS

General

Implementar las técnicas de *Data Mining* y análisis de procesos en un diseño experimental, para mejorar la interpretación y estandarización de métricas de *Growth Hacking*, utilizadas en sitios web o aplicaciones.

Específicos

1. Identificar qué componente tecnológico se debe implementar en la arquitectura de *Growth Hacking*, para optimizar su procesamiento y efectividad.
2. Estandarizar la secuencia de los procesos de interpretación de métricas del método de *Growth Hacking*, utilizando métodos de reingeniería.
3. Determinar el rendimiento y mejoras que pueden implementarse en la interpretación de métricas de *Growth Hacking*.

MARCO METODOLÓGICO

Tipo de investigación

El tipo de investigación tiene un enfoque cualitativo. La investigación definió el componente tecnológico para implementarlo en la metodología de *Growth Hacking*.

Luego de contar con la implementación del componente tecnológico, se realizó el análisis de la información con técnicas específicas para el procesamiento de datos, con el fin de continuar el estándar de secuencia de procesos de interpretación de métricas de las metodología de *Growth Hacking*, es relevante mencionar que *Growth Hacking* no es una metodología de procesos ordenados cuando se trata de implementar e interpretar las métricas, por tanto, se debe seguir la metodología propuesta.

Por último, se determinó el rendimiento y mejoras que pueden implementarse en la interpretación de métricas de *Growth Hacking*, para ello se utilizó la investigación cualitativa, con el fin de establecer cuáles de las métricas son las más eficientes y que técnicas lograban mejores resultados.

Diseño de investigación

La investigación propuesta para el trabajo de graduación, se realizó mediante el diseño experimental, la investigación, la experimentación y la utilización de distintos instrumentos de investigación, se tiene como fin principal:

- Descubrir mediante el proceso de análisis de métricas, las respuestas a las variables del estudio para la creación de la secuencia de pasos de interpretación de métricas de *Growth Hacking* con diferentes componentes tecnológicos, por medio de la publicación de un sitio de fans en *Facebook*.

Alcance de la investigación

La investigación tiene un alcance descriptivo, por medio del análisis de *Data Mining* y todo el proceso en el análisis de datos que conlleva se pueden observar patrones en los usuarios y contenidos que ayudan al crecimiento de la comunidad de usuarios, esto contribuye a que una marca se dé a conocer más rápido y llegue a más personas.

Variables consideradas en la investigación

En la tabla I, se detallan las variables, sub-variables e indicadores que se consideraron para desarrollar el proyecto.

Tabla I **Variables de la arquitectura**

Variable	Descripción	Sub_variable	Indicador
Arquitectura de Growth Hacking	Variable que considera la interpretación y estandarización de métricas de la arquitectura Growth Hacking.	<ul style="list-style-type: none"> ✓ Componente tecnológico • Métrica 	<ul style="list-style-type: none"> ✓ Cantidad de procesamiento ✓ Nivel de efectividad • Porcentaje de rendimiento

Fuente: elaboración propia.

Fases de la investigación

A continuación se detallan las fases realizadas para llegar concretar el desarrollo de la investigación del trabajo de graduación, estas fases describen en qué consisten y qué instrumentos de recolección de información se utilizaron para la investigación.

- **Fase I Revisión documental**

Se realizó la revisión documental de la arquitectura de *Growth Hacking* para determinar el componente tecnológico que se implementó, para optimizar su procesamiento y efectividad. También se revisaron bibliografías, con el fin de aplicar técnicas de análisis de la información y reingeniería de procesos para la estandarización en la secuencia de interpretación de métricas de *Growth Hacking*. Además, se realizó la revisión bibliográfica acerca de las métricas y las formas de interpretación más efectivas, para mejorar el rendimiento de las mismas.

- **Fase II Diseño y desarrollo del prototipo**

Se diseñó y desarrolló un prototipo de *Fanpage* implementando la viralización automática de *Facebook*, la cual ayudó a optimizar la arquitectura de *Growth Hacking*. Se demostró la secuencia de los procesos de interpretación de métricas del método de *Growth Hacking* de forma práctica debido a que se determinó que el estándar es funcional comparado con otros sitios *web*, además se implementaron mejoras en la definición de las métricas de *Growth Hacking*.

- **Fase III Experimentación**

- Después de implementar *Clustering* en la arquitectura de *Growth Hacking* se realizaron mediciones acerca del procesamiento y efectividad comparada con otros sitios web que han implementado esta metodología sin aplicar *Clustering* y los resultados no son claros, ni concluyentes.
- Para estandarizar la secuencia de los procesos de interpretación de métricas de *Growth Hacking*, se utilizó la metodología para el análisis y evaluación de los procesos luego se realizaron las etapas del rediseño o reingeniería de procesos para alcanzar un estándar.
- Se realizó la selección e implementación de los sistemas de medición de acuerdo a las métricas de *Growth Hacking* implementadas en el prototipo, con el fin de realizar mejoras en la interpretación de las mismas.

- **Fase IV Análisis de la información**

Se realizó el análisis de la información con el fin de optimizar el procesamiento y efectividad de la arquitectura de *Growth Hacking* con *Clustering* ya implementado, además de analizar la secuencia de procesos estándar que debería realizarse para la implementación e interpretación de métricas de *Growth Hacking* y con el mismo análisis pudo determinarse las mejoras en las interpretaciones de las mismas.

- **Fase V Creación del estándar de interpretación de métricas**

La creación del estándar de interpretación de métricas del método de *Growth Hacking* contiene una serie de requerimientos que deben cumplirse, entre ellos el procesamiento correcto de los datos obtenidos del sitio de fans. Luego se elaboró diagrama general de la arquitectura (Ver Figura 2) utilizando el método de *Clustering* aplicado al estándar en la secuencia de los procesos de interpretación de métricas del método de *Growth Hacking* y al mismo tiempo aplicando las mejoras que puedan realizarse para un mejor rendimiento.

- **Fase VI Redacción del informe final**

La redacción del informe final de la investigación incluye todo el proceso que tiene que realizarse para llegar a los resultados esperados, luego de optimizar la arquitectura de *Growth Hacking*, el estándar de la secuencia de los procesos de interpretación de métricas de *Growth Hacking* y las mejoras que se implementaron para el mismo proceso.

INTRODUCCIÓN

Las técnicas de *Growth Hacking* están basadas en *marketing* tecnológico, existen varios sitios web que se han vuelto virales y de gran popularidad, debido a la implementación de dichas técnicas. Entre los sitios y redes sociales más populares que han implementado *Growth Hacking* están: *LinkedIn*, *Facebook*, *Twitter*, entre otros.

En el siguiente estudio experimental se investigaron las técnicas y métricas más utilizadas y las que aportan mayores beneficios a determinados sitios web o aplicaciones. En la investigación experimental se realizó un prototipo que ayudó a entender el comportamiento de los usuarios en sitios de fans creados en *Facebook*, a través de métricas e interpretación de las mismas.

Growth Hacking es un método de reciente aparición y definición, por lo que el presente estudio da a conocer conceptos básicos, métricas que pueden aplicarse y cómo deben interpretarse, siendo este el punto clave para lograr la optimización y estandarización del método, debido a que en el método de *Growth Hacking* todo debe ser medible, de lo contrario no es útil y no brindará ningún resultado que pueda utilizarse. El estudio experimental determinó también una secuencia de pasos para saber cómo aplicar dichas técnicas y cómo realizar el análisis e interpretación de métricas de *Growth Hacking*. Para la estandarización en la interpretación y análisis de métricas se hizo uso de métodos auxiliares como *Clustering* y reingeniería de procesos, tratando de crear un estándar funcional y que sea aplicable dependiendo del tipo de sitio *web* al que se aplique.

A continuación se presentan los siguientes capítulos para cumplir con todos los requerimientos del trabajo de investigación y a la vez con la experimentación:

Antecedentes: como se puede observar el término de *Growth Hacking* es relativamente reciente, por lo tanto es poca la bibliografía y estudios científicos que se ha realizado acerca de este tema aplicado a la tecnología, por lo tanto, se dan a conocer casos y estudios puntuales donde se han aplicado técnicas de minería de datos y reingeniería de procesos en el campo del *marketing* tecnológico además de métricas que deben tomarse en cuenta para la realización de proyectos en general y diseño de sitios *web*.

Justificación: El presente trabajo tiene como propósito incidir sobre la línea de investigación acerca del componente tecnológico y la forma correcta de analizar los procesos de *Growth Hacking*, debido a que no existían estándares para la interpretación de métricas que fueran válidos para todos los modelos de negocios y fases de una empresa en proceso de crecimiento y se trata de brindar un aporte que contribuya con algunas empresas que desean implementar *Growth Hacking* en sus sitios *web* o *Fanpage*, pero no tienen los conocimientos para interpretar las métricas y tomar decisiones con los datos recopilados.

Alcances: Los alcances investigativos y técnicos están dirigidos a cumplir con los objetivos específicos del presente trabajo de investigación. Además los resultados de los alcances son la implementación y experimentación de lo investigado en las primeras fases para implementar las técnicas de *Data Mining* y análisis de procesos en un diseño experimental que mejoran la interpretación y estandarización de métricas de *Growth Hacking*.

Marco teórico: Se dan a conocer conceptos importantes que complementan el tema central de investigación que es *Growth Hacking*, y también las técnicas y métodos que se utilizaron en la implementación del prototipo, estos conceptos son importantes para entender cómo experimentar y entender el comportamiento del análisis de *Clustering*, para la estandarización en la interpretación de métricas además de un estándar en los procesos de interpretación. Se detallaron los métodos estadísticos y la forma en la que se presentan los resultados apoyándose en gráficas estadísticas para una mejor comprensión de los mismos.

Metodología: En este capítulo se brindan las variables que son necesarias para medir el cumplimiento de los objetivos; también se detalla el tipo de estudio que en este caso en particular es una investigación experimental y las fases de investigación como la revisión documental, para recolectar información sobre el tema y aplicar las técnicas ya conocidas para el análisis de la información. Finalmente, la metodología que se utilizó para presentar los resultados y la redacción del informe final.

Presentación de resultados: Se representan los resultados de forma clara y concisa utilizando métodos de fácil entendimiento para el lector y tratando de demostrar los puntos medulares de dicha investigación.

Análisis y discusión de los resultados: Es la fase final de la investigación puesto que no basta únicamente con la exposición de los resultados en sí, sino que se analizan y presentan una discusión de los mismos para conocer el aporte que brinda dicha investigación. Para la discusión de resultados se hace uso de la estadística descriptiva, debido a que es la que más se acopla al tipo de estudio que a continuación se presenta a detalle.

1. ANTECEDENTES

Zandi indica en su estudio titulado: Marco para la selección de técnicas inteligentes del negocio basado en arquitecturas inteligentes, indica que las empresas construyen almacenes de datos y utilizan técnicas inteligentes del negocio para identificar y desarrollar nuevas oportunidades. Las nuevas oportunidades pueden proporcionar una ventaja de mercado competitiva y la estabilidad a largo plazo. Por lo tanto, estas técnicas son reconocidas como una estrategia importante y utiliza recursos para impulsar nuevas oportunidades de negocio. Como se sabe una arquitectura inteligente de negocios se desarrolla mediante la definición de la arquitectura y la selección de una técnica inteligente de negocio apropiada, para el almacenamiento de datos; la arquitectura es una importante decisión así lograr nuevas estrategias para la organización. La importancia de una selección de la arquitectura apropiada desde la perspectiva de las aplicaciones o inteligencia del negocio es cuádruple. (Zandi, 2014).

- Permite que varias empresas se conviertan en una parte de la empresa virtual con el menor coste y mayor productividad.
- Es esencial para alinear los requisitos derivados de las aplicaciones de inteligencia del negocio con las necesidades empresariales en la arquitectura o toma de decisiones.
- Mejora la interoperabilidad entre múltiples arquitecturas inteligentes y heterogéneas en una multiempresa o empresa virtual.
- Se trata de un soporte pivotante, facilitador para crear un ambiente donde sea necesario los datos para fluir libremente y con seguridad entre todas las empresas de la empresa virtual o multiempresa.

El estudio anterior indica la importancia de la estandarización de métodos y arquitecturas utilizando técnicas inteligentes de negocios, para el crecimiento de una empresa y para lograr el alineamiento de los requisitos necesarios para la toma de decisiones, contribuye con el proyecto de estandarización de métricas e interpretación de la mismas utilizando *Growth Hacking*, debido a que demuestra varios puntos clave al tratar de definir una arquitectura inteligente para lograr nuevas estrategias, lo cual es aplicable a cualquier *startup* o marca que desea crecer de forma escalable (Zandi, 2014).

Du, Jiao y Tsen en su estudio titulado: Estimación del tiempo de terminación del proyecto y los factores de análisis para la gestión de proyectos de ingeniería concurrente, indican aspectos muy importantes acerca de las tareas que suelen ser interdependientes entre sí y que requiere mucha iteración antes de su culminación, en el que el crítico método vía de evaluación / programa y la técnica de la crítica (CPM / PERT) pueden no ser aplicables para ayudar a estimar la duración del proyecto. Además, la realización de un proyecto a gran escala en un entorno dinámico tiene que lidiar con varios factores al mismo tiempo. Al estimar el tiempo de finalización del proyecto, la investigación previa a menudo se centró en un tema de interés y asumió los otros factores que causan poco efecto sobre la duración total del proyecto. El objetivo del estudio de estimación del tiempo de terminación del proyecto y los factores de análisis para la gestión de proyectos de ingeniería concurrente, es desarrollar un marco de investigación, para ayudar a estimar el tiempo de terminación del proyecto y analizar los principales factores que afectan a la estimación de proyectos de ingeniería concurrentes y complejos. En este estudio se denota la importancia de la recopilación de datos, donde se preparan los datos necesarios, incluyendo la estructura de tarea del proyecto, las relaciones de trabajo, y las características de los miembros del equipo cuantificados, donde las tareas se asignan dinámicamente a los miembros apropiados, según el nivel de

conocimientos de cada miembro de la tarea, la capacidad de trabajo en equipo, la disponibilidad de horario de trabajo, y mejora la curva de aprendizaje. El análisis de datos, donde los factores importantes al momento de concluir el proyecto son estudiados basándose en el análisis de los resultados. (Du, Jiao, Tsen, 2001). El estudio anterior menciona puntos importantes como la recopilación de datos, la preparación de los datos, la estructuración de datos, mejora de la curva de aprendizaje basándose en datos obtenidos, luego del análisis de resultados. Estos puntos son importantes para cualquier proyecto de ingeniería, porque no solo indica los puntos importantes para estimar el tiempo de un proyecto sino a delimitar el estudio de los datos para cada proyecto.

En el libro de Análisis Predictivo y Data Mining de Vijay Kotu publicado en el año 2015. Habla acerca de descubrir con éxito patrones utilizando la minería de datos es un proceso iterativo. Proporciona un marco para resolver el problema basándose en la minería de datos. El proceso de cinco etapas que se describe y proporciona directrices sobre recolección experiencia en el tema; la exploración de los datos con estadísticas y visualización; la construcción de un modelo de uso de algoritmos de minería de datos; probar el modelo y su despliegue en un entorno de producción; y finalmente reflexionar sobre los nuevos conocimientos adquiridos en el ciclo.

Durante los años de evolución de las prácticas de minería de datos, diferentes marcos para el proceso de minería de datos se han presentado por diversos organismos académicos y comerciales, como el proceso de la cruz estándar de la industria para la minería de datos, descubrimiento de conocimiento en bases de datos, etc. En cuanto a la clasificación o predicción de la clase Vijay Kotu utiliza la información de los predictores o variables independientes para clasificar las muestras de datos en dos o más clases distintas o cubos. La clasificación es la tarea de minería de datos más utilizada

en los negocios. Hay varias maneras de construir modelos de clasificación, en diferentes capítulos se definen y discuten la ejecución de seis de los algoritmos de clasificación más utilizados: árboles de decisión, regla de inducción, vecinos más cercanos, ingenuo bayesianos, redes neuronales artificiales, y máquinas de vectores soporte (Kotu, 2015).

En el estudio de Sinergia en la reingeniería de procesos con sistema de planificación de recursos empresariales desarrollado por Santosh Pattanayak en el año 2015, que trata acerca de procesos de reingeniería en negocios siendo este un nuevo concepto de estrategia de negocio, hizo su impacto en los sectores industriales con el advenimiento de la liberación, la privatización y la globalización de la economía india. Pero no puede obtener los beneficios deseados en términos de rentabilidad y sostenibilidad de una organización sin abordar adecuadamente los temas y problemas de los interesados externos en toda la cadena de suministros.

Las tecnologías de la información y comunicaciones (TIC), han impulsado conceptos funcionales cruzados como Enterprise Resource Planning (ERP) integrando a la perfección todas las áreas funcionales y provocando una mejora significativa en los procesos de negocio, así como la productividad de las empresas. Con la operación del servicio de rápido crecimiento, ERP ha ampliado su alcance de sinergización volviendo automáticas las funciones de oficina con funciones de front office.

El cliente y el entorno empresarial centrado deben verse como bienes del capital en el sector de fabricación, siempre es de enorme importancia analizar críticamente el flujo de los procesos de negocio con el cambio de escenario de mercado. Funcionalidad como la planificación de un sistema ERP proporciona a las empresas de negocios la tan buscada medios para gestionar la capacidad

de producción, la disponibilidad de materiales y horarios de envío. Sinergizar es el reto de los grupos de interés en general; como múltiples problemas, tanto técnicos como no técnicos, deben abordarse simultáneamente, mientras que se pone un modelo estratégico para trabajar, curiosamente, no mucho trabajo se ha hecho en esta área. (Pattanayak, 2015).

En el estudio de McMullan y Kervin acerca de la venta de juegos de azar en Internet, publicado en octubre de 2012; examina el diseño web y la ingeniería, la publicidad y el *marketing*, y las características pedagógicas presentes en una muestra aleatoria de 71 sitios de póker internacionales obtenida desde el directorio de Casino City en el verano de 2009. Se codificó para 22 variables relacionadas con el acceso, la apelación, jugador protección, atención al cliente, en el lugar de seguridad, el uso de imágenes, texto y lenguaje, materiales interactivos y envoltentes, productos promocionales y programas, patrocinios, las celebridades, los recursos de tutela, programas de juego responsable, y el contenido temático. Se exploran varias preguntas: ¿Cómo están los sitios de póker en Internet diseñados y estilizados para crear confianza y familiaridad y atraer y retener a los consumidores? ¿Qué tan común son las nuevas formas de medios sociales como "*marketing* viral", "comercialización de la red" y "palabra de ratón" de publicidad para llegar y que constituyen los jugadores de póker como consumidores? ¿Qué mecanismos hacen desplegar sitios web de juegos de azar de Internet para socializar a los clientes a comprar sus productos y entrar en una cultura de póker? ¿Cuáles son los mensajes publicitarios en sitios web master y cuáles son sus implicaciones para el juego responsable y problemas con el juego? Este estudio demuestra muchos puntos importantes que apoyan las técnicas de Growth Hacking aplicado a sitios web y la importancia de crear confianza en los clientes y la ingeniería que debe utilizarse en el diseño de sitios web para que este apoye al crecimiento escalable de la empresa. (McMullan y Kervin, 2012).

2. JUSTIFICACIÓN

El presente trabajo tiene como propósito incidir sobre la línea de investigación de “Tecnologías de la Información y la Comunicación para desarrollo de nuevos mercados e industrias”, definida por la Universidad de San Carlos de Guatemala, mediante un estudio acerca del uso de *Data Mining* y análisis de procesos para estandarizar e interpretar métricas del método de *Growth Hacking*.

Se debe tener claro que en un método como *Growth Hacking*, las métricas son la base fundamental del mismo y que estas métricas son de suma importancia para la toma de decisiones en una *startup* o marca no conocida. También otro de los aspectos que no se debe pasar por alto es el período sobre el cual se van a trabajar las métricas, esto va ligado intrínsecamente al modelo de negocio, aunque estas pueden variar según resultados. Existen muchas herramientas para métricas de sitios *web* y aplicaciones una de las más conocidas, *Google Analytics* que a pesar de ser muy poderosa como herramienta de medición y muy popular no es la más indicada para métricas granulares y específicas, debido a su complejidad y su visión general del producto a un alto nivel. Por los aspectos mencionados y sabiendo que existen tanto métricas como sistemas de medición poco eficientes, es necesario crear un método innovador y que sea funcional para medir e interpretar las métricas generadas por *Growth Hacking*.

Data Mining parece ser el método ideal para el estudio de métricas, debido a que son una gran cantidad de datos generados de diferentes fuentes, porque *Data Mining* en esencia es un proceso analítico diseñado para explorar

los datos, normalmente de negocios o relacionados con el mercado. Se planea utilizar el método para buscar patrones consistentes y/o relaciones sistemáticas entre variables y luego validar mediante la aplicación de patrones detectados a nuevos subconjuntos de datos. Además de buscar una forma de estandarizar e interpretar las métricas de *Growth Hacking*, se debe buscar una forma de analizar los procesos para que este método se vuelva más eficiente cuando sea aplicado en un sitio web o en una aplicación.

No existen estándares de métricas que sean válidos para todos los modelos de negocios y fases de una empresa en proceso de crecimiento. Muchas veces se cae en el error de querer medir sitios o aplicaciones de diferente tipo o función utilizando las mismas métricas y como consecuencia, los resultados no son los esperados o no indican ningún parámetro que se pueda tomar en cuenta para tomar decisiones que apoyen el desarrollo de la empresa. Si no se tiene bien definido el modelo de negocio y no se sabe en qué etapa de crecimiento se encuentra la empresa es muy probable que no se tenga el conocimiento de las métricas que aportarán mejores resultados para el análisis de resultados obtenidos de las métricas. Otro aspecto muy importante son las herramientas, debido a que si se cuenta con un estándar de métricas, pero no se tiene un estándar de herramientas tarde o temprano caerá en el mismo proceso desordenado y sin estandarización para la interpretación de métricas.

Por lo tanto, el presente proyecto de investigación es un aporte que contribuye con muchas empresas que emplean *Growth Hacking* en sus sitios web o aplicaciones, pero que no saben cómo interpretar los datos y por consecuencia se les dificulta la toma de decisiones para el crecimiento de su empresa o marca y la inversión en un sitio *web* o aplicación termina siendo un gasto adicional sin ningún beneficio para la empresa, en vez de apoyar el

crecimiento de la misma. Con la implementación del método de *Data Mining* para la estandarización e interpretación de las métricas de *Growth Hacking*, se pretende establecer un estándar en la secuencia de procesos de análisis e interpretación de datos, de tal modo que, se tengan parámetros de dónde iniciar y tener ideas más claras de qué métricas son útiles para tomar en cuenta y qué debe hacerse con las métricas, luego de obtenerlas para su análisis.

3. ALCANCES

3.1. Investigativos

Para el diseño experimental con la *Fanpage* y la implementación de técnicas de *Growth Hacking*, se establecen los siguientes alcances investigativos:

- Se investigó acerca de *Clustering* y determinó cuál herramienta se acopla mejor al tipo de estudio y datos a procesar, para diseñar una arquitectura de *Growth Hacking* que optimiza la interpretación de métricas, su procesamiento y efectividad.
- Se investigó el método de reingeniería para la estandarización de procesos en la interpretación de métricas de *Growth Hacking*.
- Se investigaron las técnicas para mejorar el rendimiento en la interpretación de métricas de *Growth Hacking*, después de implementar el diseño experimental y haciendo un comparativo con los sitios *web* que han aplicado técnicas e interpretación de métricas de *Growth Hacking*.

3.2. Técnicos

Growth Hacking es un concepto que surgió hace un par de años y se define como una combinación de análisis, curiosidad y creatividad para hacer crecer el número de visitantes o usuarios en una página *web*, *Growth Hacking* no cuenta con un estándar en la interpretación de métricas, por lo que se pretende combinar técnicas de *Data Mining* específicamente el método de *Clustering* para optimizar su procesamiento y efectividad.

Para la estandarización de procesos de interpretación de métricas de *Growth Hacking*, se combinó con métodos de reingeniería de procesos, realizando una revisión y rediseño de los procesos de *Growth Hacking* con el objetivo de alcanzar mejoras en el rendimiento al momento de obtener las métricas y los resultados de la interpretación de las mismas. Lamentablemente por ser un concepto nuevo existen pocos estudios y antecedentes acerca de las métricas y técnicas de *Growth Hacking*; por lo que se deben estudiar las métricas obtenidas de sitios *web* en los que se han implementado dichas técnicas para comprobar su funcionalidad de forma ideal en distintos sitios *web*. Esto con el fin de determinar tecnológicamente, las mejoras que ha habido conforme se han ido implementando dichas técnicas en el transcurso de los años.

3.3. Resultados

- Diseño experimental e implementación del método de *Clustering*, aplicado a las técnicas e interpretación de métricas de *Growth Hacking*, para optimizar el procesamiento y efectividad de la arquitectura de dicho método.
- Estándar de la secuencia de los procesos de interpretación de métricas del método de *Growth Hacking*, luego de haber aplicado reingeniería de procesos.

Comparativo de tablas estadísticas para determinar el rendimiento y mejoras, luego de haberse aplicado una estandarización en la secuencia de procesos de interpretación de métricas de *Growth Hacking* vrs. procesos de interpretación de métricas del mismo método, aplicados anteriormente.

4. MARCO TEÓRICO

4.1. Growth Hacking

Growth Hacking no es sólo construir viralidad sobre un producto, también hace hincapié en crear un producto tan bueno como sea posible para tanta gente como sea posible. Cuando se habla de producto, hace referencia a un sitio *web*, aplicación o una *Fanpage* y debe ser tan bueno que la gente quiera consumir y se sienta orgullosa para compartirlo. El análisis de información continuo ayuda a definir el comportamiento del consumidor potencial y hace que se puedan identificar las fases por las que pasa para intervenir con trucos, funciones o sistemas para que un usuario pueda influir en otros para hacer crecer un determinado negocio. (Iglesias, 2014).

Andrew Chen define *Growth Hacking* como un híbrido entre técnicas de *marketing* y métodos de codificación y programación, revolucionando el *marketing* tradicional y convirtiéndolo en tecnológico. *Growth Hacking* se auxilia de pruebas A/B, páginas de destino, el factor viral, la capacidad de entrega del correo electrónico; además utiliza la medición cuantitativa, modelado de escenarios a través de hojas de cálculo y un gran número de consultas de bases de datos. *Growth Hacking* es un término nuevo para muchos, pero se ha practicado desde hace tiempo entre los vendedores más destacados del Internet y los gerentes de producto en Silicon Valley. (Chen, 2013)

El *Growth Hacking* no solo trata de *marketing* viral; sino también es la innovación constante que hace mejoras estratégicas en varias fases del plan de *marketing* tecnológico, con el propósito de estimular el crecimiento de las *Startup* que lo implementan. (Brits, 2015)

4.2. Growth Hacker

Surgen con la era moderna de la *Web 2.0* para llegar a un mercado determinado y distribuir una idea, en lugar del *marketing* tradicional un hacker debe atraer y entender el comportamiento del usuario proporcionando un valor inmediato para persuadir. El *Growth Hacker* debe apoyarse en todas las disciplinas, analizando puntos de vista económicos del comportamiento y *Gamification*, con el propósito de hallar el mensaje acertado que atraiga a los usuarios. (Chen, 2013).

Un *Hacker* de crecimiento encuentra una estrategia dentro de los parámetros de un método que sea escalable y pueda repetirse para apoyar el crecimiento, impulsado por producto e inspirado por los datos generados por las diversas tecnologías que se apliquen en el proceso. Un pirata del crecimiento vive en la intersección de los datos, productos y comercialización, vive dentro del equipo de producto y una de sus cualidades debe ser su expresión con lenguaje técnico para poder llevar a la práctica lo que se quiere. Los canales tradicionales de mercadeo a menudo representan un alto costo para adquirirlos y un retorno de la inversión muy bajo, debido a la poca competencia. En esta época actual de usuarios de redes sociales y una buena estrategia para el crecimiento con el debido ajuste del mercado en los productos, dará como resultado un escalonamiento masivo por medio de ciclos virales. (Holiday, 2014).

4.3. Embudo de conversión de Growth Hacking

La idea principal del embudo de conversión de *Growth Hacking* se le atribuye a Dave McClure, quien realizó un estudio acerca de formas diferentes

de hacer *marketing* tecnológico basado en las siglas AARRR (adquirir, activar, retener, referir, retornar). El embudo en *marketing* es una herramienta que se ha utilizado durante mucho tiempo por su funcionalidad, pero se deben de tener los conocimientos básicos del embudo del *marketing* tradicional para entender el subyacente de Dave McClure. La idea principal del embudo es hacer un seguimiento entre las personas que escuchan y hablan del producto (activo) y la compra del producto (pasivo). Las etapas de Aarr son muy similares a las del embudo de *marketing* tradicional, aunque los embudos de *marketing* tradicional se adaptan a cada negocio, en función del producto o servicio que ofrecen. En esencia, la Figura 1 explica el embudo de *Growth Hacking* y las diferencias con el *marketing* tradicional.

Figura 1. Embudo de Growth Hacking



Fuente: <http://www.anabelendominguez.com/servicios/growth-hacking/>

4.4. Canales virales de Growth Hacking

Michael Geer es el pirata del crecimiento responsable de la creciente base de usuarios de *Badoo* a más de 70 millones de personas, impartió una charla en la ciudad de Nueva York en la que se refirió a canales virales que no cuestan un centavo. Si se está buscando convertirse en un pirata del crecimiento, la virilidad es un componente importante todos los días, si se está involucrado con un consumidor o empresa orientada a negocios. Pero la viralidad no es algo instantáneo para un producto o para el bienestar del crecimiento de la base de usuarios. La viralidad es el núcleo sobre el ofrecimiento de valor a los usuarios de manera que van organizándose para hacer referencia a otras personas acerca de un determinado producto.

Michael Geer hace referencia a los siguientes canales virales:

- Libretas de direcciones de correo *web*: Usar un servicio como invitaciones abiertas es un buen método y además asegurarse que se están haciendo APIs (interfaz de programación de aplicaciones) para los cuatro mayores webmail: *Yahoo*, *Hotmail*, *AOL* y *Gmail*.
- APIs de medios sociales: Debe de asegurarse de mantener una presencia activa en las principales redes sociales *Facebook*, *Twitter*, *LinkedIn* y *Foursquare*, y no en los no tan populares y recién llegados como *Google+*, *Yammer*, *Tumblr*, y *Salesforce* para obtener una ventaja sobre la competencia.
- Perfeccionar los contactos: Si se cuenta con un producto móvil debe asegurarse que está construido en la función de la usabilidad

y funcionalidad para la gente y para compartir con sus contactos del teléfono.

- Aplicaciones de Facebook: *Facebook* es denominado el gigante de las redes sociales, cuenta con más de mil millones de usuarios, el equivalente a la séptima parte de la población mundial, por lo que se determina que es la mayor de las redes sociales. (Casanova, 2015)

4.5. Métricas de Growth Hacking

En el *Growth Hacking* o piratería del crecimiento todo gira en torno a los datos y la información que pueda generarse con los mismos a través de la interpretación de las métricas, para tratar de hacer crecer una empresa. Si un *Growth Hacker* no es capaz de medir y analizar la información correcta o requerida simplemente no está haciendo bien su trabajo.

A continuación se presentan los tres indicadores que son absolutamente necesarios para rastrear los datos necesarios si se quiere crecer de forma sostenible y escalable.

4.6. El cambio neto diario

Se debe mantener un ojo en la base de usuarios todos los días, con el fin de saber cuánto ha cambiado, por suerte hay un gráfico simple que se puede utilizar con la variación neta diaria incluyendo cuántos nuevos usuarios están adquiriendo compromiso, y cuántas personas se han caído. El gráfico evalúa el

efecto individual de cada elemento y cuánto tiene de crecimiento, a continuación los elementos y la ecuación para la gráfica:

- Nuevos usuarios ¿Cuántas personas se inscribieron hoy?
- Los usuarios que han vuelto ¿Cuántas personas han utilizado hoy por primera vez en un mes?
- Usuarios batidos ¿Cuántas personas han utilizado por última vez hace exactamente un mes?

4.7. Activos principales diarios

Ahora hay dos problemas con el cambio neto diario, métrica que probablemente ya se identificó y no permite cuantificar o medir y realizar el seguimiento de calidad de los usuarios que esté atrayendo, ni se tendrá una forma de rastrear el avance de la retención. La gráfica deberá mostrar algunos picos durante una gran campaña de adquisición de usuarios, pero no se tiene idea de si esas personas se subscribieron y se olvidarán al día siguiente o en realidad permanecerán con el uso del servicio, aquí es donde la métrica de activos principales diarios entra en juego. (Casanova, 2015).

4.8. Planificación SOSTAC

Por sus siglas (Situación, objetivos, estrategia, tácticas, acciones y control), es justo anticipar que una gran proporción de la innovación y el *Growth Hacking* puede relacionarse con la comercialización en línea, o al menos los sistemas de comunicación en línea. Es apropiado considerar un marco básico que se puede

utilizar para organizar el proceso de comercialización. A continuación se describen los pasos del proceso de planificación SOSTAC:

4.8.1. Análisis de la situación actual

Aquí se consideran las percepciones de los clientes y un análisis FODA que es un análisis del plan de negocios. Debe preguntarse: lo que los competidores e incluso jugadores de oportunidades en otros sectores están aprovechando y que hace falta en la actualidad. Se deben tener en cuenta los recursos y determinar si la externalización será necesaria.

4.8.2. Objetivos claros

¿Qué elementos del embudo de ventas se deberían mejorar?

¿Cómo se puede innovar?

Debe surgir una lluvia de ideas de qué tanto es posible en las comunicaciones de *marketing* lograr y lo que se podría lograr al optimizarlas.

4.8.3. Estrategia de cómo llegar allí

Esta es en realidad los objetivos más amplios y la dirección que se debe seguir para llegar a un determinado punto, para lo cual se debe considerar la forma sistemática actual con la que se opera.

4.8.4. Las tácticas son métodos específicos

Tácticas de hecho son los pasos que se toman para hacer seguro el logro de los objetivos estratégicos. En realidad, la intención es comprometer e influenciar a las personas, incentivar y promocionar, a través del método cliente a cliente el contenido preciso para que pueda ser ejecutada a la perfección. Incluso deben prepararse pruebas A / B opcionales, para esto será requerida la ayuda de un redactor.

4.8.5. Acciones

Ahora deben ejecutarse, ya sea internamente o con los otros sitios, la ejecución es la parte más importante, ya que producirá los resultados que se requieren o se anticipan.

4.9. Control

Se tiene que medir el desempeño, con el fin de hacer ajustes, para aprobar un nuevo presupuesto y para medir la eficacia. Este paso también puede ser una cuestión de análisis, señalando que se hace actualmente y esperar los primeros resultados donde el valor de la vida de los clientes o el negocio sea extendida. (Brits, 2015).

4.10. Data Mining

Aunque algunas de las técnicas de minería de datos son bastante nuevas, la minería de datos en sí no es una nueva tecnología, en el sentido de que las personas han estado analizando los datos de equipos desde los primeros ordenadores que se inventaron y sin las computadoras durante siglos antes de eso. Con los años, la minería de datos ha pasado por muchos nombres diferentes, tales como el descubrimiento de conocimiento, inteligencia de negocio, modelos de predicción, análisis predictivo, y así sucesivamente. La definición de la minería de datos como el usado por los autores es:

La minería de datos es un proceso de negocio para explorar grandes cantidades de datos para descubrir patrones y reglas significativas. La minería de datos es un proceso tecnológico y la consecuencia de la práctica es que las organizaciones que quieren sobresalir en el uso de sus datos para mejorar sus negocios no ven la minería de datos como un espectáculo secundario. En lugar de ello, su estrategia de negocio debe incluir la recopilación de datos, análisis de datos para el beneficio a largo plazo, y actuar sobre los resultados. (Linoff & Berry, 2011).

Data Mining se define como el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. (Witten & Frank, 2011).

Data Mining se refiere al proceso de extraer conocimiento de bases de datos. Su objetivo es descubrir situaciones anómalas y/o interesantes, tendencias, patrones y secuencias en los datos. Su ingreso de datos son los que ya están pre-procesados en las fases anteriores de la metodología, el objetivo es construir un modelo a partir de ellos, el cual pueda producir nuevo conocimiento que sea útil para el usuario. Dentro de *Data Mining* (Minería de

Datos) hay tipos de tareas con sus propios requisitos, las cuales pueden ser resueltas por un algoritmo de *Data Mining* específico. Las distintas tareas pueden ser descriptivas (agrupamiento o *Clustering*, reglas de asociación secuenciales y las correlaciones) o predictivas (clasificación y la regresión). (Molina, 2009).

4.11. Machine learning o aprendizaje automático

El aprendizaje automático se ocupa de enseñar a los ordenadores algo sobre el mundo, de modo que se pueda utilizar ese conocimiento para realizar otras tareas. Por el contrario, la estadística se enfoca en el desarrollo de herramientas para la enseñanza de los seres humanos algo sobre el mundo, de modo que se pueda tomar mejores decisiones. En el aprendizaje automático, el aprendizaje se produce mediante la extracción de la mayor cantidad de información a partir de los datos como sea posible (o razonable) a través de algoritmos que analizan la estructura básica de los datos y distingue la señal del ruido. Después de haber encontrado la señal, o el patrón, los algoritmos simplemente deciden que todo lo demás que sobra es el ruido. Por esa razón, las técnicas de aprendizaje automático también se conocen como algoritmos de reconocimiento de patrones. Se puede "entrenar" a las máquinas para aprender acerca de cómo se generan los datos en un contexto dado, y que permita utilizar estos algoritmos para automatizar muchas tareas útiles. Aquí es donde el término conjunto de entrenamiento proviene, al referirse al conjunto de datos que se utilizan para construir un proceso de aprendizaje automático. La noción de la observación de los datos, aprender de ella, y luego la automatización de algunos procesos de reconocimiento está en el centro del aprendizaje automático. Dos tipos particularmente importantes de patrones constituyen los problemas fundamentales: el problema de la clasificación y el problema de la regresión. (Conway & White, 2012).

4.12. Clasificación

Un modelo de clasificación se construye para asignar observaciones en dos o más categorías distintas. Por ejemplo, un modelo de clasificación puede ser construido para estimar si un cliente comprará o no va a comprar un producto en particular. En otro ejemplo, una modelo de clasificación puede ser construido para predecir si las perforaciones en un área en particular son como resultado del hallazgo de petróleo o no. (Myatt, 2007).

4.13. Regresión

Un modelo de regresión es un modelo matemático que predice una respuesta continua variable. Por ejemplo, un modelo de regresión se podría desarrollar para predecir el volumen de ventas reales o la temperatura resultante de un experimento.

Es importante analizar los residuos sobre la base de una serie de factores, incluyendo los siguientes:

- Variable de respuesta: No debería haber ninguna tendencia en los valores residuales a lo largo la gama de la variable de respuesta, es decir, la distribución debe ser aleatoria.
- Distribución de frecuencias: La distribución de frecuencias de los valores residuales deben seguir una distribución normal.
- Orden de observación: No deberían existir tendencias discernibles en función de cuándo se midieron las observaciones.

4.14. Aplicación de R para machine learning

R es un lenguaje extremadamente poderoso para manipular y analizar datos. Su meteórico ascenso en popularidad dentro de las comunidades científicas de datos y aprendizaje automático lo ha convertido de hecho en un lenguaje de análisis. El éxito de R en la comunidad de análisis de datos se debe a dos factores: R proporciona la mayor parte de la potencia técnica que los estadistas requieren, y R ha sido apoyado por una comunidad de estadistas que también son devotos del código abierto. Hay muchas ventajas técnicas que ofrece un lenguaje diseñado específicamente para la informática estadística, estos datos pueden visualizarse haciendo pasar los resultados a la trama de función, que está diseñado para visualizar los resultados de este análisis en otros idiomas con grandes comunidades, tales como *Python*, *NumPy*, *SciPy* y *matplotlib*.

Además, como en otros entornos de computación científica, el tipo de datos fundamental en R es un vector, los vectores pueden ser agregados y organizados de varias maneras, pero en esencia, todos los datos se representan de esta manera.

El aprendizaje automático se desarrolla en la intersección de las matemáticas y estadísticas tradicionales con la ingeniería de software y la informática. La estadística casi siempre se ha preocupado por aprender algo interpretable a partir de datos, mientras que el aprendizaje de máquina se ha ocupado de convertir los datos en algo práctico y utilizable. (Conway & White, 2012).

4.15. Clasificación de datos

Dada una colección de registros o conjunto de entrenamiento, deben realizarse los siguientes pasos:

- Cada registro contiene un conjunto de atributos, uno de los atributos es la clase.
- Encontrar un modelo de atributo de clase como una función de los valores de otros atributos.
- Objetivo: previamente los registros invisibles deben ser asignados a una clase con la mayor precisión posible.
- Un equipo de prueba se utiliza para determinar la exactitud del modelo.
- Por lo general, el conjunto de datos dado se divide en capacitación y de prueba, con un conjunto de entrenamiento utilizado para construir el modelo y la prueba de conjunto se utiliza para validarlo. (Tan, Steinbach, & Kumar, 2005).

4.16. Clustering

La agrupación es el proceso de examinar una colección de puntos y agrupar los puntos en *clusters*, de acuerdo con alguna medida de distancia. El objetivo es que puntos en el mismo grupo tienen una pequeña distancia uno de otro, mientras que los puntos en diferentes grupos se encuentran en una gran distancia uno de otro.

Sin embargo, la intención era que había tres grupos de alrededor de tres diferentes intersecciones de carreteras, pero dos de los grupos se mezclan unas con otras porque no estaban suficientemente separados. (Leskovec, Rajaraman, & Ullman, 2014).

4.17. Publicidad en la web

Una de las grandes sorpresas del siglo XXI ha sido la capacidad de toda clase de aplicaciones web interesantes para mantenerse así mismos, a través de la publicidad, en lugar de suscripción. Mientras que la radio y la televisión han logrado utilizar la publicidad como su principal fuente de ingresos, la mayoría de medios de comunicación - periódicos y revistas, por ejemplo, han tenido que utilizar un enfoque híbrido, que combina los ingresos de la publicidad y las suscripciones. (Leskovec, Rajaraman, & Ullman, 2014).

5. PRESENTACIÓN DE RESULTADOS

5.1. Análisis, diseño de la arquitectura y desarrollo del prototipo experimental

En la Figura 2 se ilustra la arquitectura del proceso general de la interpretación de métricas de *Growth Hacking*, que es básicamente la generación de datos en el Servidor de *Facebook* y la exportación de los datos diarios, semanales o mensuales a *Microsoft Excel*. Como parte de la propuesta se incluyen los componentes tecnológicos: el procesamiento de datos en *SQL Server* y el método de *Data Mining* seleccionado en este caso *Clustering*, hace uso de la herramienta de *R Studio*, para observar patrones y comportamientos de las publicaciones y personas; con el fin de crear un informe final de interpretación de métricas en un PDF (ver Anexo 4), para realizar mejoras en las futuras publicaciones. En el anexo 3 puede observarse otras gráficas y métodos de *Data Mining* (regresión lineal, otra forma de *Clustering*) implementados en la experimentación con los cuales no se tuvo mayores resultados a diferencia de *Clustering* en sus tres formas.

Figura 2. **Diagrama de la arquitectura para interpretación de métricas de Growth Hacking**



Fuente: elaboración propia, con base a Microsoft Visio.



Con la propuesta de la arquitectura anterior se automatizan varios procesos que al hacerlos manuales serían demasiado complejos y consumirían mucho tiempo, además los resultados aplicando el componente tecnológico propuesto tendría una mayor precisión y son herramientas fáciles de utilizar a diferencia de otros gestores de Bases de Datos y herramientas con mayor complejidad de uso.

5.1.1. **Funcionalidad para promocionar publicación en la Fanpage de Facebook**

Esta funcionalidad que ofrece *Facebook* es realmente eficaz en pocas horas o días la difusión de un determinado contenido puede aumentar y alcanzar a miles de personas las cuales pueden visualizar el contenido promocionado, teniendo la opción de elegir el país, región y número de personas se desea promocionar. Aunque esta funcionalidad es pagada a

continuación puede comprobarse que es funcional a diferencia del *marketing* convencional en sitios *web*, los cuales acarrear altos costos de mantenimiento y alojamiento.

Figura 3. Imagen de la funcionalidad para promocionar publicación

Promociones recientes en Jajaja					+ Crear promoción
	Publicación promocionada Finalizada Sube un meme, video, audio o imagen a http://bi...	6.075 Personas alcanzadas	2.859 Interacciones con publicaciones	Q14,00 Gasto total: Q14,00	Ver promoción
	Publicación promocionada Finalizada Etiqueta a tu ligue, amigo, novio o persona espe...	3.393 Personas alcanzadas	605 Interacciones con publicaciones	Q8,00 Gasto total: Q8,00	Ver promoción
	Publicación promocionada Finalizada Registrarte en nuestro sitio web y gana muchos ...	497 Personas alcanzadas	56 Interacciones con publicaciones	Q8,00 Gasto total: Q8,00	Ver promoción
	Publicación promocionada Finalizada Cuando presentas a tu novia, y tus amigos actúa...	4.602 Personas alcanzadas	2.231 Interacciones con publicaciones	Q10,00 Gasto total: Q10,00	Ver promoción

Fuente: elaboración propia.

En la Figura 3 se muestra el listado de publicaciones promocionadas en la *Fanpage* que se realizó para obtener las métricas de *Growth Hacking*, puede observarse a las personas alcanzadas y el número de interacciones que obtuvo, en un determinado período.

5.2. Publicación de la *Fanpage* y sitio *web*

Con el fin de realizar la experimentación en dos entornos diferentes, además de la *Fanpage* en *Facebook* se creó un sitio *web* con las mismas técnicas de *Growth Hacking* y se intentó viralizar o promocionar de distintas formas, promocionando el enlace en *Facebook*, regalando artículos y prendas

de vestir para ganar seguidores y lamentablemente no se obtuvo la cantidad de datos esperados para poder aplicar *Data Mining* y un proceso de análisis de la datos de manera correcta, por lo cual fueron descartados. Sin dudar el alcance que posee *Facebook* es inmenso y no únicamente para contenidos de entretenimiento y ocio sino también para promocionar un determinado producto o empresa.

5.3. Exportación de datos a Microsoft Excel

Este paso es muy sencillo pero de mucha importancia para el análisis de datos, debe seleccionarse rango de fechas y elegir entre los diferentes tipos de reportes que *Facebook* ofrece, al obtener los datos puede observarse que aunque es un gran volumen de datos, no nos brinda ningún tipo de información para la medición de métricas de *Growth Hacking* y esto es debido a que son datos crudos en diferentes hojas de cálculo y en un mismo archivo, por ello, es necesario seguir todos los pasos desde la preparación de datos hasta obtener datos confiables y aplicar *Clustering*.

5.4. Preparación de datos

La preparación de los datos es una de las partes que más tiempo consume en el análisis de datos o en un proyecto de *Data Mining*.

Se determinó una sola fuente de datos para su respectivo análisis, a continuación se detalla brevemente:

5.4.1 Fuente de datos

La calidad de los datos es el factor más importante que influye en la calidad de los resultados de cualquier análisis. Los datos a menudo son recolectados para responder a específicas preguntas usadas en diferentes estudios en este caso en particular la experimentación.

5.4.1.1. Datos generados en Facebook

Los datos que provienen de esta fuente son los que se obtuvieron de la *Fanpage*, que se publicó con las herramientas que *Facebook* ofrece a cualquier persona que tiene una cuenta en dicha red social, esto con el fin de medir las técnicas de *Growth Hacking* definidas con anterioridad para crear una mayor difusión y viralidad; toma como parámetros métricas de *Growth Hacking*, para medir los resultados y realizar la experimentación.

5.4.2. Comprensión de datos

5.4.2.1. Tablas de datos

Para formar las tablas de datos, se basó en las métricas de *Growth Hacking* y en los datos de mayor relevancia que se generaron de la fuente, con el fin de estandarizar la interpretación de métricas de *Growth Hacking* y proceder con una correcta toma de decisiones.

5.4.2.2. Variables continuas y discretas

Para iniciar la categorización se definió cada variable en términos de tipo de valores que dicha variable puede tomar. Ver Tabla II.

Tabla II. **Categorización de variables**

Variables	Categorización
Fecha: Se refiere a las fechas en que se realizaron determinadas publicaciones.	Continua
Días: Se refiere al número de días que pasaron para determinadas interacciones o resultados a partir de la fecha de publicación.	Discreta
Likes: Se refiere al número de “Me Gusta” que recibió un determinada publicación.	Discreta
Interacción: Define cualquier tipo de actividad que un determinado usuario realizó en la página.	Discreta
Alcance: Trata de la variable que hace referencia al número de usuarios que les apareció una determinada publicación pudiendo ser orgánico o pagado.	Discreta
Género: Se refiere si el contenido tuvo algún tipo de interacción con un hombre o una mujer.	Discreta
Edad: Se definieron 7 rangos de edad	Discreta

para determinar esta variable.	
Idioma: Se refiere al idioma que se habla en determinado país que se visualizó la publicación	Discreta
País: Hace referencia del país donde se interactuó o se visualizó la publicación.	Discreta

Fuente: elaboración propia.

5.4.2.3. Roles en el análisis

La Tabla III despliega los roles de cada variable según su contenido.

Tabla III. **Roles de variables**

Variable	Rol
Fecha	Etiqueta
Días	Etiqueta
Likes	Descriptor
Interacción	Descriptor
Alcance	Descriptor
Género	Etiqueta
Edad	Descriptor
Idioma	Descriptor

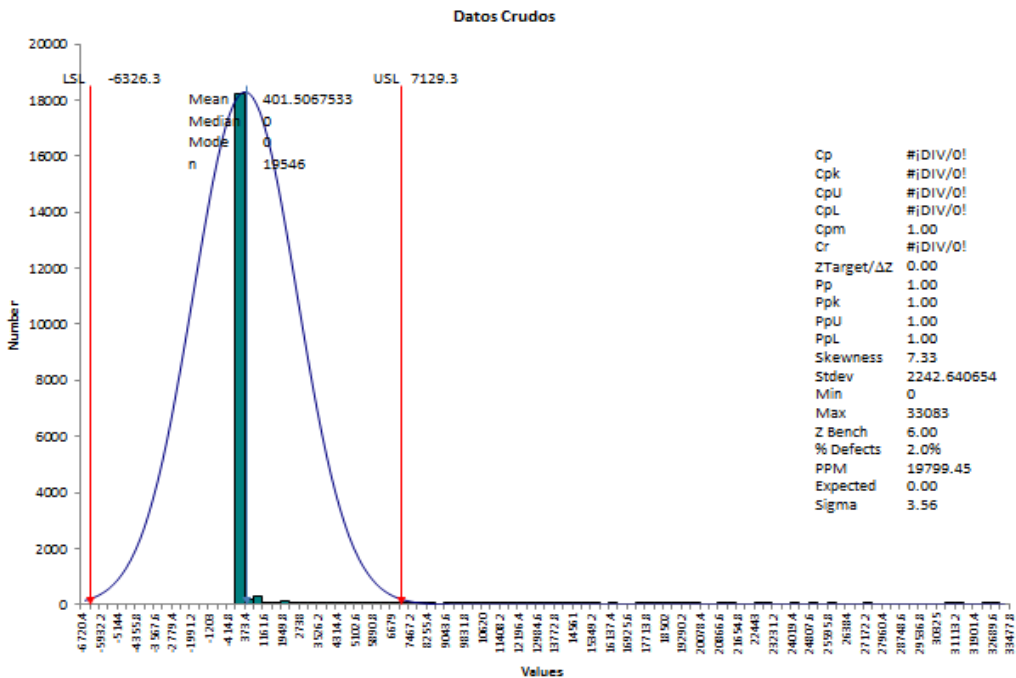
País	Descriptor
------	------------

Fuente: elaboración propia.

5.4.2.4. Distribución de frecuencia

Pueden observarse muchas inconsistencias en los datos, se muestran valores erróneos al momento de calcular la tabla de frecuencias y esto se debe a que el universo de datos crudos, no brindó ningún resultado coherente por falta de preparación de los mismos, por lo que es importante realizar este cálculo para identificar los puntos de error e inconsistencias, y que datos deben ser limpiados, removidos, transformados y segmentados. Ver Figura 4.

Figura 4. Histograma de datos crudos



Fuente: elaboración propia, con base a QI Macros para Excel.

5.4.3. Preparación de datos

5.4.3.1. Visión de conjunto

Luego de crear el análisis y distribución de frecuencias para determinar posibles errores y los cambios necesarios en los datos, se procedió con la visualización de los datos para la revisión de las secciones de los mismos dispuestos para proceder con la limpieza de los datos en general, en caso de ser necesario.

5.4.3.2. Limpieza de los datos

Se corrigieron encabezados de los datos y se eliminaron columnas vacías o errores que se generaron al momento de exportar los datos de los distintos sitios *web*, por cuestiones de compatibilidad entre versiones de Excel o archivos de tipo CVS y herramientas utilizadas como R y *Q/ Macros de Microsoft Excel*.

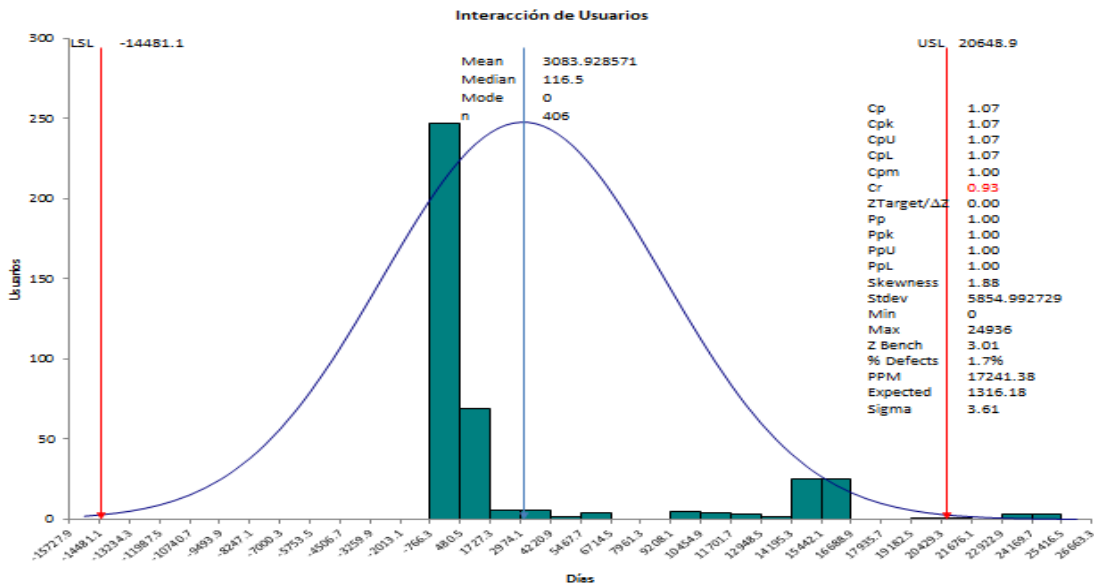
5.4.3.3. Eliminación de variables

Se eliminaron algunas variables (Tiempo de reproducción de videos, personas hablando del contenido, publicaciones vistas en las noticias del muro de *Facebook*, etc.), debido a que no eran necesarias por ser redundantes, estas variables fueron generadas por la misma fuente, además al momento de realizar los cálculos, estas no aportaban ningún dato significativo para el estudio.

5.4.3.4. Mapeo de valores

Se hace referencia al histograma de la Figura 4. Comparado con el de la Figura 5. Puede observarse, que los datos son más consistentes ya tienen sentido y muestra la interacción que se tuvo durante un mes y la barra más alta hace referencia a los usuarios que se alcanzaron pagando la funcionalidad de “Promocionar publicidad” en *Facebook* y las barras más pequeñas, hacen referencia al alcance orgánico que se refiere a los usuarios que se alcanzaron por la difusión manual, compartiendo los contenidos con contactos sin ningún costo.

Figura 5. Histograma de interacción de usuarios

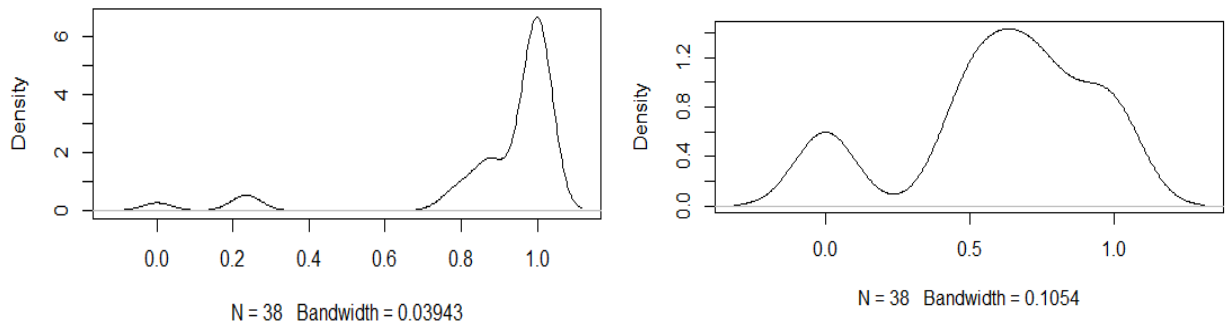


Fuente: elaboración propia, con base a QI Macros para Excel.

Es importante observar el comportamiento de diferentes tipos de gráficas con las variables que se utilizarán. Esto con el fin de determinar errores en los datos, patrones que contribuyen al análisis y agrupaciones de datos que pueden generar resultados interesantes, de forma práctica y eficiente. En la

figura 6 puede observarse la comparación de densidad de los porcentajes de visualización de publicaciones en la semana 1 y los cambios que suceden en la semana 5.

Figura 6. Gráficas de densidad



Fuente: elaboración propia, con base a R Studio.

5.4.4. Tablas y gráficas

En la Tabla IV se agruparon los “Me gusta” de la página de *Facebook* recibidos en una semana promedio, por rangos de edad y género. Se considera que estas variables son de suma importancia en la interpretación de métricas de *Growth Hacking* y obtener información del tipo de contenido de mayor agrado para los usuarios y fans del sitio.

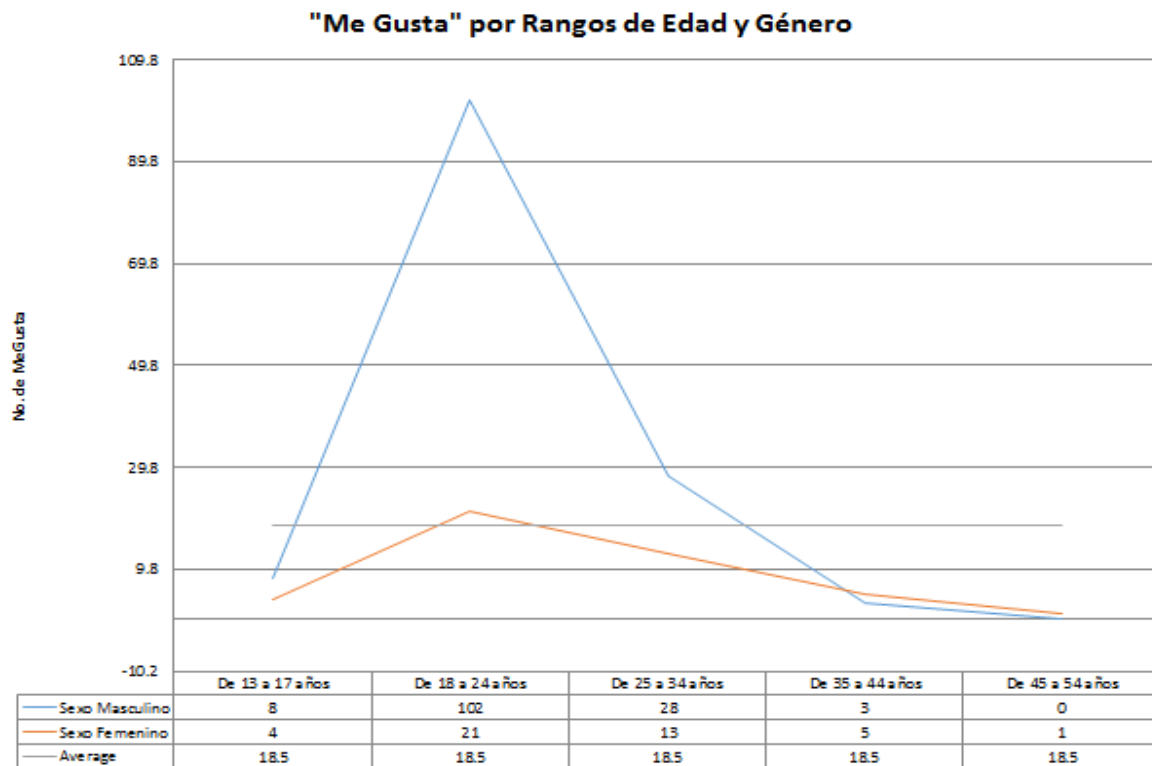
Tabla IV. **Tabla de sumatoria de “Me gusta” semanal promedio**

Rangos de edad	Sexo Masculino	Sexo Femenino	Totales
De 13 a 17 años	8	4	12
De 18 a 24 años	102	21	123
De 25 a 34 años	28	13	41
De 35 a 44 años	3	5	8
De 45 a 54 años	0	1	1
Total	141	44	185

Fuente: elaboración propia.

En la Figura 7, se observa que la mayor incidencia se presenta en hombres en el rango de edad de 18 a 24 años, lo mismo sucede con las mujeres siendo el mayor porcentaje en el mismo rango de edad, pero en menor cantidad que el género masculino. Esto demuestra que existe una mayor interacción en las publicaciones de la *Fanpage* con el género masculino que con el femenino. Los datos observados son el promedio de los “Me gusta” en el lapso de una semana.

Figura 7. **Polígono de frecuencias de “Me gusta”**



Fuente: elaboración propia, con base a QI Macros para Excel.

Se presenta la sumatoria del alcance mensual por rangos de edad y género, estos resultados se logran al aplicar la viralización automática que ofrece *Facebook*, aunque esta herramienta es pagada, dependiendo del

número de alcance de personas es muy eficiente y pueden expresarse los resultados en la Tabla V.

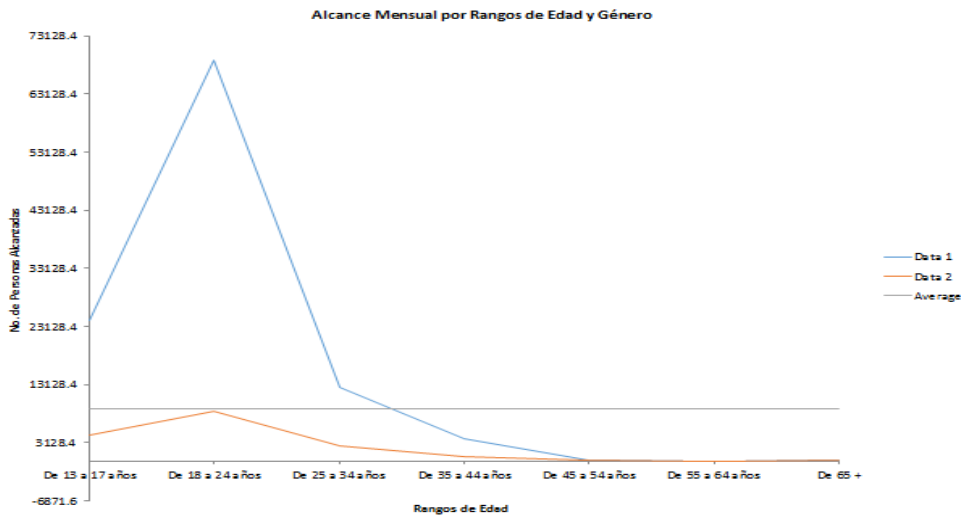
Tabla V. **Sumatoria de personas alcanzadas mensual promedio**

Rangos de edad	Masculino	Femenino	Totales
De 13 a 17 años	24100	4385	28485
De 18 a 24 años	68925	8513	77438
De 25 a 34 años	12727	2501	15228
De 35 a 44 años	3859	750	4609
De 45 a 54 años	224	106	330
De 55 a 64 años	20	19	39
De 65 +	63	42	105
Total	109918	16316	126234

Fuente: elaboración propia.

Es evidente que en el alcance de personas que visualizaron el contenido de la página en *Facebook*, gracias a la herramienta de publicidad pagada, para difundir de forma rápida y efectiva se repite el mismo comportamiento que en la Figura 7. La mayor incidencia en *Facebook* en Guatemala es en el género masculino en el rango de edad de 18 a 24 años. Este dato podría ayudarnos a difundir diferentes contenidos de forma más rápida y efectiva utilizando las técnicas y métricas de *Growth Hacking*. Ver Figura 8.

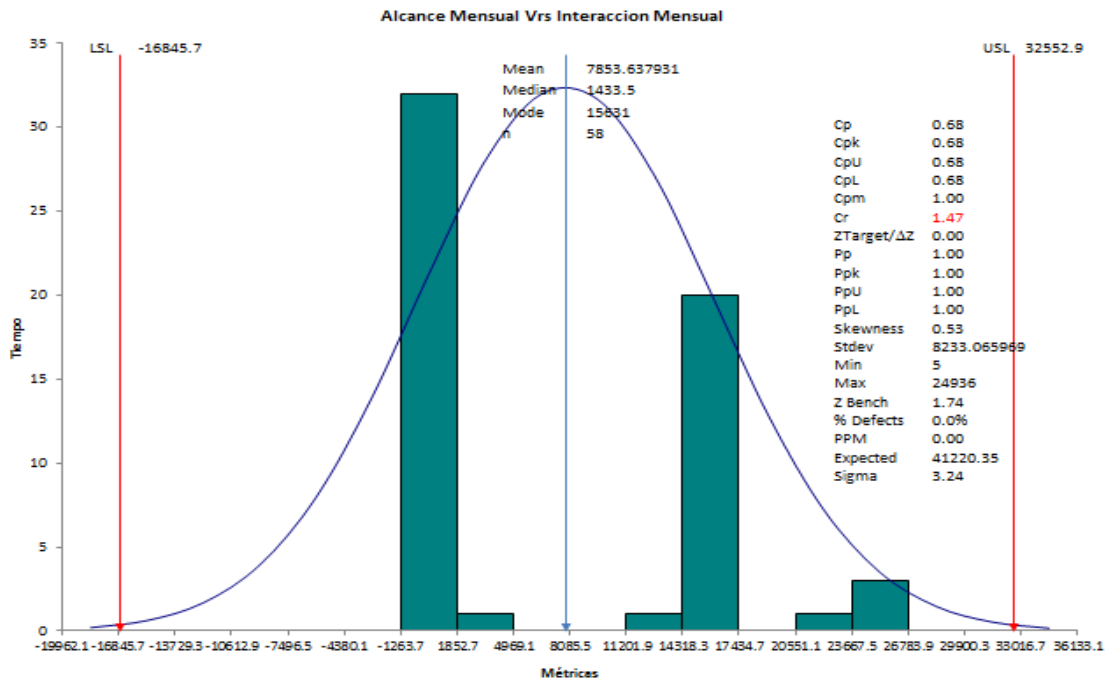
Figura 8. Polígono de frecuencias de personas alcanzadas



Fuente: elaboración propia, con base a QI Macros para Excel.

En la Figura 9 puede observarse que la promoción automática de *Facebook* es efectiva, ya que es correspondida por las interacciones de los usuarios con el contenido publicado en el sitio y el histograma muestra que cuando se difunde de forma manual entendiéndose compartiendo con contactos propios o contactos de amigos, el contenido alcanza a muy pocas personas mientras que con esta funcionalidad aumenta de forma muy notoria el alcance y de forma proporcional aumenta la interacción de los usuarios, lo que significa que puede difundirse cualquier marca o producto orientándolo a un nicho de mercado en específico en este caso a hombres en el rango de 18 a 24 años en las regiones de Guatemala y se tendrían resultados efectivos y rápidos. Además los contenidos visuales son más atractivos para los usuarios lo cual se ha comprobó en las pruebas de experimentación.

Figura 9. **Histograma del alcance mensual vrs. la interacción mensual**



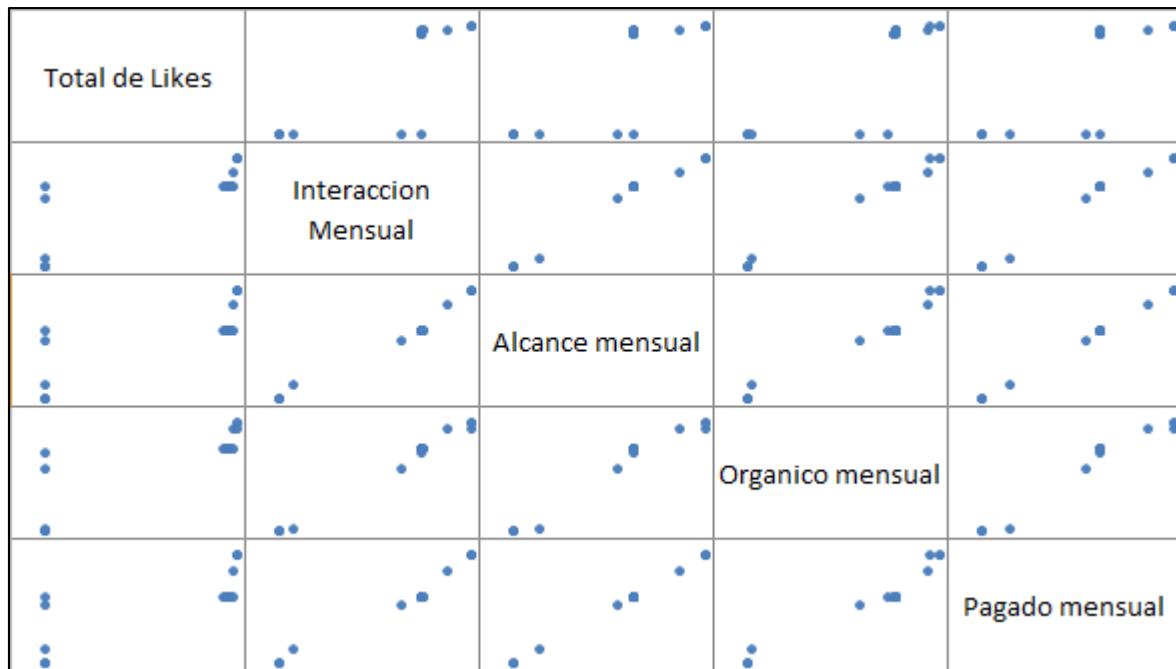
Fuente: elaboración propia, con base a QI Macros para Excel.

En la siguiente matriz de dispersión (Ver Figura 10) puede observarse las relaciones y similitudes entre las métricas de *Growth Hacking* utilizadas en la experimentación, en muchos puntos coinciden los patrones debido a que al momento de pagar la promoción automática de *Facebook* aumentan las otras métricas de manera exponencial, lo cual trae muchos beneficios al contenido publicado y esto contribuiría a aumentar ventas de una determinada marca o ayudaría a darse a conocer una *Startup*; además es una herramienta que está al alcance de cualquier persona que tenga acceso a una cuenta de *Facebook* y aún no se ha explotado todo su potencial.

Otro punto importante y observable es que la cantidad de *Likes* es la métrica que tiene un menor número y se comporta muy diferente a las demás, esto se debe a que los usuarios deben estar realmente convencidos que todo el

contenido publicado en la *Fanpage* es atractivo para volverse fans de la página, de lo contrario solo interactúan con alguna publicación de forma eventual.

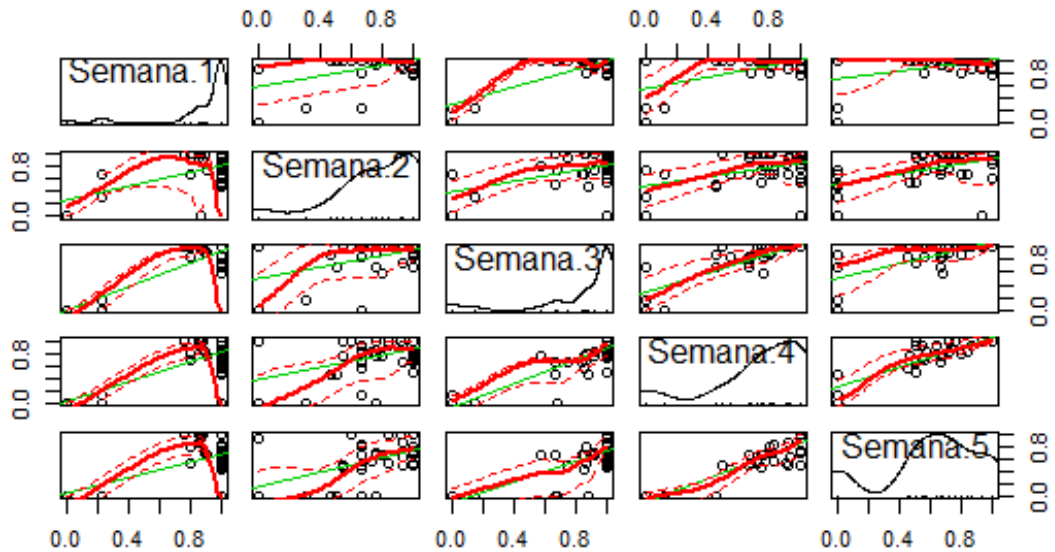
Figura 10. **Matriz de dispersión de métricas de Growth Hacking**



Fuente: elaboración propia, con base a QI Macros para Excel.

En la siguiente matriz de dispersión de la Figura 11, se observan los promedios de interacción con los contenidos gráficos en la *Fanpage*, se tomó como parámetro de tiempo cinco semanas de experimentación, esto con el fin de ver los diferentes comportamientos de las interacciones de los usuarios con las diferentes publicaciones por tipo.

Figura 11. Matriz de dispersión del promedio de interacciones



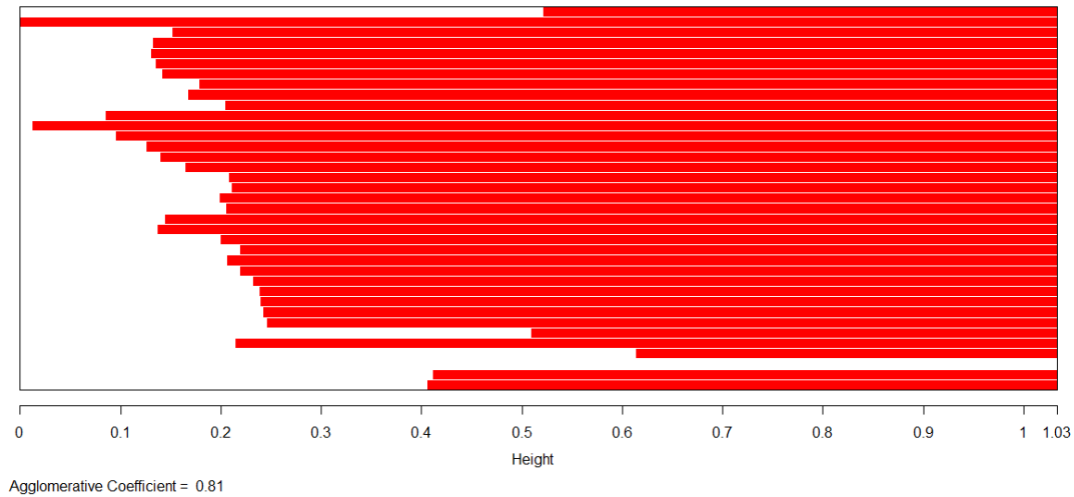
Fuente: elaboración propia, con base a R Studio.

5.5. Clustering

Esta es la fase final de todo el análisis de datos, existen tres tipos de *Clustering* con los cuales se experimentaron, con base a los datos ya procesados y limpios para aplicar *Data Mining*. En la Figura 12, se aplicó el método simple de *Clustering* y puede observarse claramente que los *cluster* se dividen en dos grupos principales y esto demuestra la falta de efectividad para viralizarse de los enlaces, por el contrario los videos e imágenes poseen valores más altos, por su buena aceptación por los usuarios y fans de la página de *Facebook*.

Figura 12. **Banner del método simple**

Banner of agnes(x = a2, method = "single")

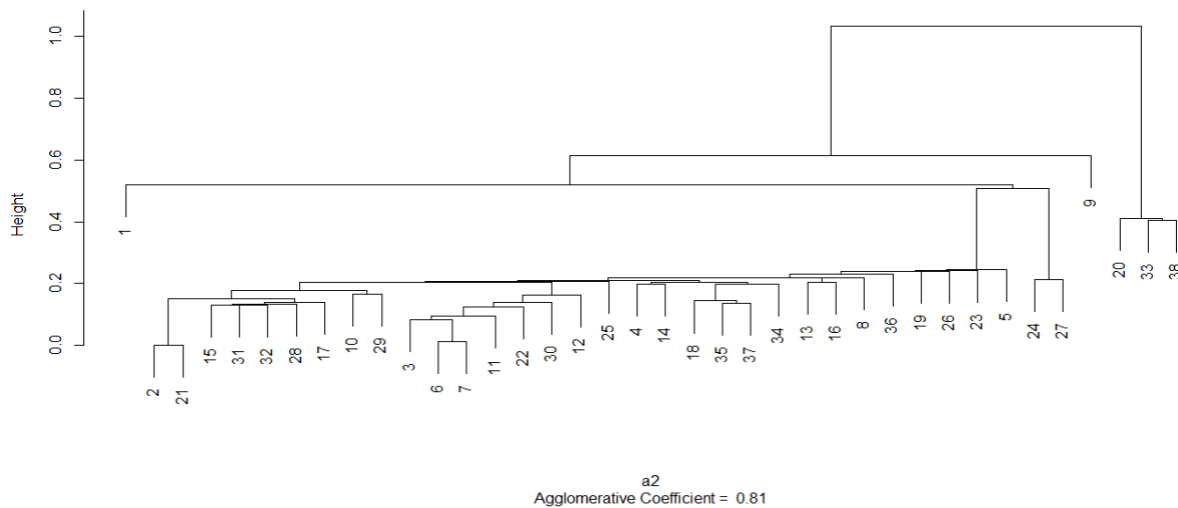


Fuente: elaboración propia, con base a R Studio.

En el dendograma del método simple (Figura 13) se observan tres grandes grupos y se debe a la mayor cantidad de videos publicados y que obtuvieron mayor interacción durante un periodo de cinco semanas.

Figura 13. **Dendograma del método simple**

Dendrogram of agnes(x = a2, method = "single")

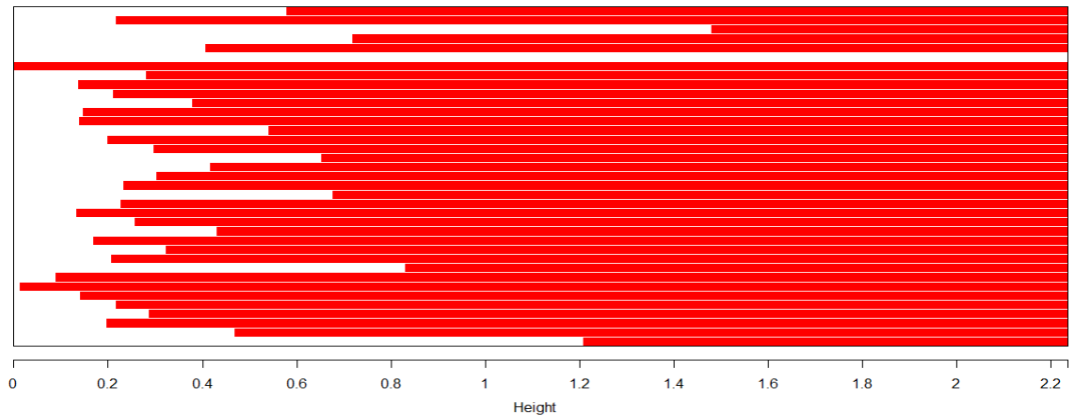


Fuente: elaboración propia, con base a R Studio.

Para la Figura 14, se utilizó el método completo y se notan claras diferencias en los *clusters* además de la variación del coeficiente aglomerativo. El dendrograma en la figura 15 es la representación del *Clustering* efectuado en la figura 14 en ambos se aplicó el método completo y al igual que en el *Banner* puede notarse los mismos dos grupos muy marcados a diferencia de los otros métodos.

Figura 14. **Banner del método completo**

Banner of agnes(x = a2, method = "complete")

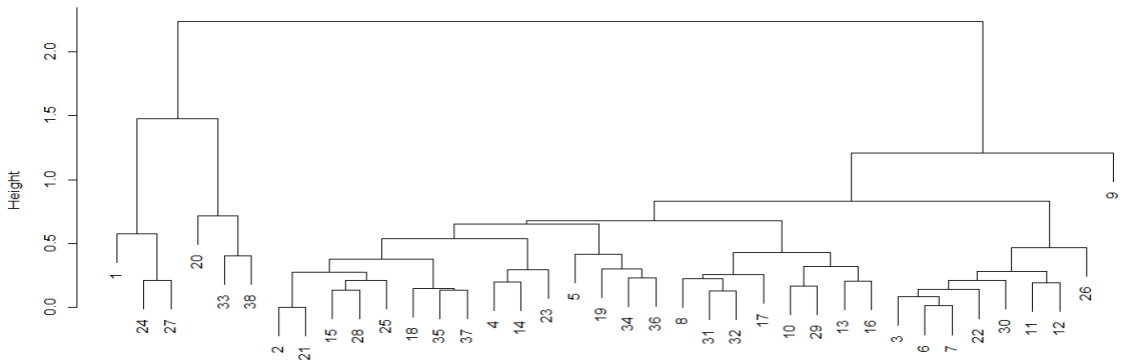


Agglomerative Coefficient = 0.89

Fuente: elaboración propia, con base a R Studio.

Figura 15. **Dendrograma del método completo**

Dendrogram of agnes(x = a2, method = "complete")

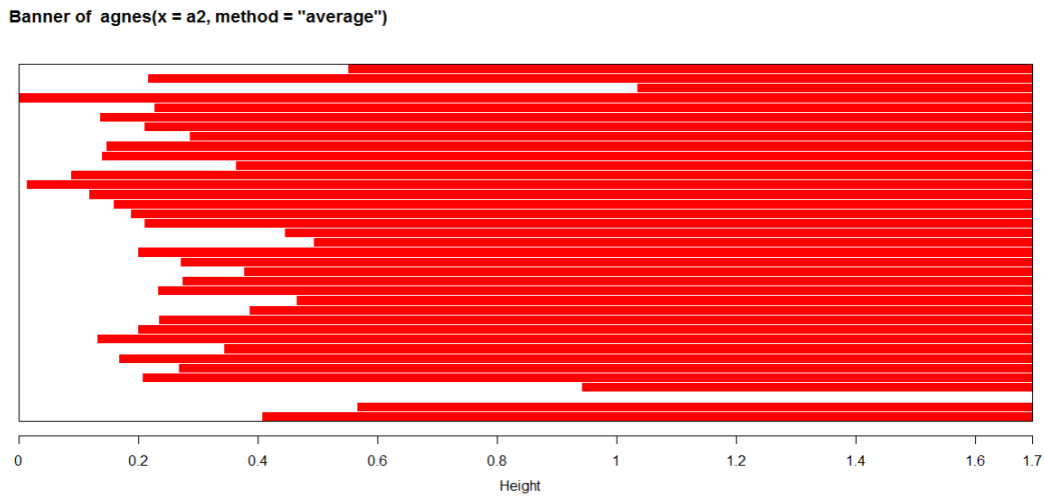


a2
Agglomerative Coefficient = 0.89

Fuente: elaboración propia, con base a R Studio.

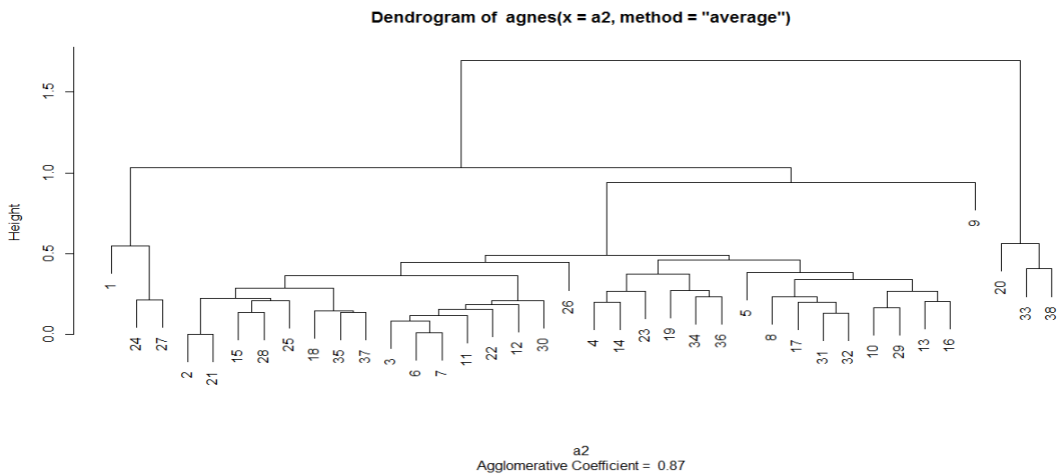
Por último, se experimentó con el método de *Clustering* promedio y su comportamiento es más parecido al método simple que al completo y siguen marcándose los mismos grupos. En la figura 16 pueden observarse las similitudes con su respectivo dendograma en la figura 17, toma en cuenta que son las visualizaciones promedio de cinco semanas, se notan los distintos comportamientos.

Figura 16. **Banner del método promedio**



Fuente: elaboración propia, con base a R Studio.

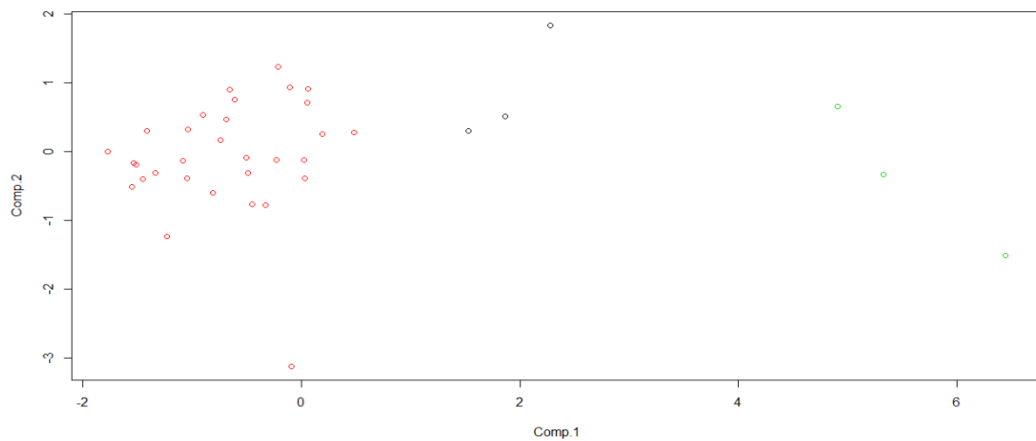
Figura 17. **Dendograma del método promedio**



Fuente: elaboración propia, con base a R Studio.

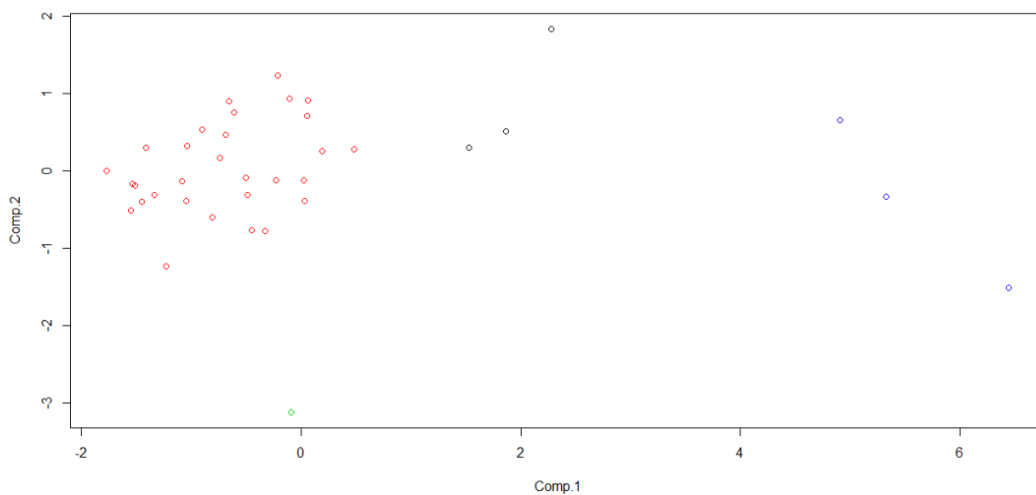
Para observar de mejor manera los resultados de *Clustering*, se aplicó división de grupos de acuerdo al tipo de aglomerado de datos, observándose que son tres grupos y un elemento aislado los que se repiten tanto en el método completo como en el simple. (Ver Figura 18).

Figura 18. **Ilustración de Clusters y Outliers del método completo**



Fuente: elaboración propia, con base a R Studio.

Figura 19. **Ilustración de Clusters y Outliers del método promedio**



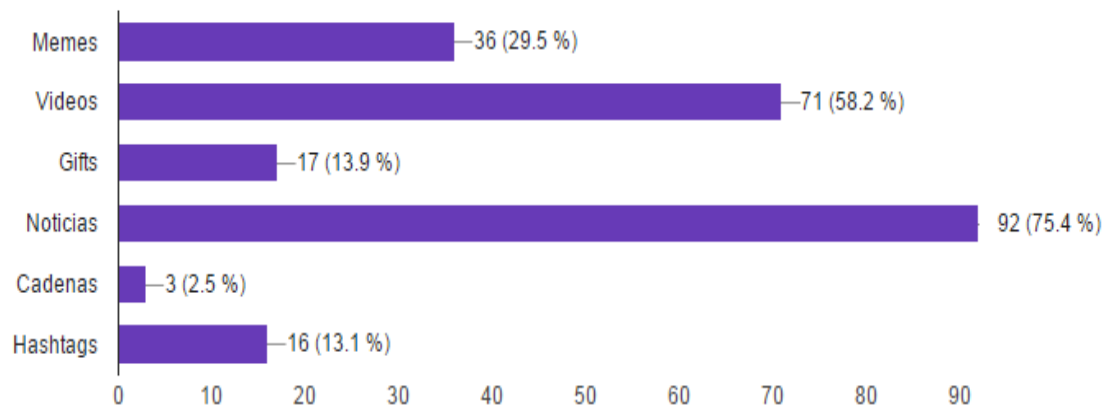
Fuente: elaboración propia, con base a R Studio.

Como puntos importantes en el análisis de *Clustering* anterior pueden observarse los *cluster* muy marcados y la Figura 19 demuestra los puntos aislados, en este caso es uno, provocando una mayor viralidad los videos, que las imágenes y los enlaces publicados.

6. DISCUSIÓN DE RESULTADOS

Para el desarrollo de la *Fanpage* en *Facebook*, se realizó un estudio cualitativo previo de las técnicas de *Growth Hacking* mayormente aceptadas por los usuarios, los resultados a dicho estudio, demostraron que los contenidos gráficos son más populares para la mayoría de personas que utilizan *Facebook* en los menores rangos de edad, lo cual se confirmó en la experimentación y puede observarse el comportamiento en las Figuras 7 y 8, esto se debe a que la población joven tiene mayor familiaridad con las redes sociales y un mayor acceso y uso desde dispositivos móviles a *Facebook*.

Figura 20. **Técnicas de Growth Hacking**



Fuente: elaboración propia, con base al estudio previo.

En cuanto al análisis de las métricas es necesario aplicar la arquitectura sugerida en la Figura 2. Como parte de los resultados del estudio previo acerca de las técnicas de *Growth Hacking*, (Ver Figura 20), se desarrolló una *Fanpage* en *Facebook* y un sitio *web*, en los cuales se aplicaron las mismas

publicaciones y se trataron de aplicar las mismas métricas, lamentablemente los datos de la página *web* no eran resultados de una difusión viral para aplicar un análisis con *Data Mining*, como lo eran los datos de la *Fanpage*, por lo que se tomó la decisión de descartar los datos generados por el sitio *web* y centrarse en la fuente de datos de *Facebook*, entonces, el estándar está orientado específicamente para la interpretación de métricas de sitios creados en *Facebook*, esto es resultado de la funcionalidad automática de “Promocionar publicación”, ya que con esta funcionalidad de *Facebook* se puede difundir a más usuarios que con las técnicas de *marketing* comunes. Por ejemplo con el sitio *web* se hicieron varios experimentos para promocionarlo como: promocionarlo regalando diferentes artículos, promocionando el enlace en la *Fanpage* y haciendo diferentes técnicas de *marketing* común y no se obtuvieron las visitas y registros necesarios para aplicar *Data Mining*.

6.1. Rendimiento del experimento

La efectividad del experimento de interpretación de métricas de *Growth Hacking* y todo su proceso de análisis de datos fue de un 85 %. Toma en cuenta la difusión efectiva, la gran cantidad de datos obtenidos en poco tiempo, la confirmación de la efectividad de las técnicas de *Growth Hacking* y sus métricas y los rangos de edad establecidos desde los resultados del estudio previo.

Para que el proceso de interpretación de métricas de *Growth Hacking* sea efectivo deben considerarse los siguientes puntos:

- Técnicas de *Growth Hacking*: Deben establecerse antes de realizar cualquier tipo de experimentación en una *Fanpage* en *Facebook*.
- Rangos de edad: Antes de realizar cualquier publicación en la página debe conocerse a que nicho de mercado se dirige y en que

rango de edad, con el fin de promocionar o viralizar de forma más rápida y efectiva.

- Género: Además de conocer a qué rangos de edad se les apunta con una determinada publicación, es importante saber qué tipo de contenido tendrá una mayor aceptación por el género femenino o masculino.
- Promocionar publicación: Esta funcionalidad de *Facebook* es un punto importante para llegar a las personas que se proyecta llegar y que la *Fanpage* sea lo más exitosa posible. Se deben conocer los alcances de la funcionalidad y como aplicarla además de elegir la publicación a promocionar y la región donde será difundida.

6.2. Análisis de la oportunidad de negocio

El modelo de negocio para la aplicación de técnicas de *Growth Hacking* en una *Fanpage*, la interpretación de métricas, y el análisis de datos con *Data Mining*, es un modelo ilimitado de oportunidades de negocio debido a que todo las técnicas de *Growth Hacking* fueron diseñadas para hacer crecer las empresas y dar a conocer las distintas marcas, pero no es suficiente con solo aplicar las técnicas de forma correcta, también es necesario medir e interpretar las métricas periódicamente para obtener resultados que colaboren a la toma de decisiones y generen cambios en el modo de aplicar la tecnología a un determinado negocio. El objetivo principal de aplicar técnicas y métricas de *Growth Hacking* a una *Fanpage* es innovar y crecer con bases tecnológicas de *marketing* y no seguir con las mismas técnicas costumbristas sin tener resultados efectivos.

6.3. Herramientas utilizadas

Durante la experimentación se utilizaron varias herramientas para el procesamiento de datos y *Clustering* como *R Studio* y *Mathlab*, obteniendo mejores resultados con *R Studio* por su fácil modo de uso además de otras herramientas que *Microsoft Excel* ofrece para la creación de gráficas estadísticas y cálculos como *QI Macros* siendo una potente herramienta y la cual es recomendable para el proceso de datos.

CONCLUSIONES

1. Las técnicas de *Data Mining* se aplicaron exitosamente, en el análisis de datos generados por la *Fanpage* en *Facebook*, la cual se basó en el diseño experimental aplicando las técnicas y métricas de *Growth Hacking*, surgiendo la estandarización de las métricas utilizadas en el desarrollo de la experimentación. Esta estandarización es aplicable a sitios *web* o sitios creados en una *Fanpage*, teniendo una mayor difusión y resultados en la plataforma de *Facebook*.
2. El componente tecnológico que debe implementarse en la arquitectura de *Growth Hacking*, para optimizar su procesamiento y efectividad son las técnicas de *Data Mining* específicamente *Clustering* que ha demostrado ser efectivo y dar a conocer muchos puntos importantes sobre los datos, y colaborar en la toma de decisiones para realizar cambios en las publicaciones del sitio.
3. Se presentó la estandarización de la secuencia de los procesos de interpretación de métricas de *Growth Hacking* a lo largo de la presentación de resultados, y se observó que todos los procesos son necesarios para llegar a un análisis sin errores aplicando reingeniería de procesos con base a la experimentación anterior.
4. El rendimiento en la interpretación de métricas de *Growth Hacking* demostró ser eficiente en varios aspectos: el correcto procesamiento de los datos, los cambios realizados en publicaciones, el poco tiempo en el que se difundieron las publicaciones y el alcance en número de personas

que visualizaron e interactuaron con las distintas publicaciones. Las mejoras que pueden implementarse son aumentar el contenido visual en las páginas (videos, imágenes, memes, noticias, etc.) en vez de texto, *hashtags* y enlaces a otros sitios, debido a que la mayoría de personas con las que se realizó la experimentación, tuvo preferencia por todos los contenidos visuales y entretenidos.

RECOMENDACIONES

1. Realizar la obtención de datos de *Facebook*, a través de una aplicación que extraiga la información automáticamente por medio del api.
2. Pueden aplicarse otras técnicas de *Data Mining* para la interpretación de métricas de *Growth Hacking*, por lo que se recomienda que en futuros estudios se apliquen diferentes técnicas como: redes neuronales, aprendizaje automático y reglas de asociación para aportar distintos conocimientos.
3. Es aconsejable seguir el estándar propuesto en la secuencia de pasos en el procesamiento de las métricas de *Growth Hacking*; además, se recomienda realizar la experimentación con otras redes sociales que puedan proporcionar una cantidad de datos considerables y provocar viralidad para contribuir con el crecimiento de nuevas marcas y empresas.
4. Desarrollar un componente tecnológico o aplicación para que luego de obtener los resultados sea capaz de postear automáticamente el contenido que provoque mayor viralidad, esto debe convertirse en un proceso de mejora continua y el análisis de métricas reflejará los cambios propuestos.

REFERENCIAS BIBLIOGRÁFICAS

1. Ben Fry. 2007. Visualizing Data. O'Reilly Media, Inc
2. Brits, A. (2015). Growth Hacking 4.0. New York: Kindle Edition.
3. Brown, M. S. (2014). Data minig for dummies. Nueva Jersey, USA: John Wiley and Sons, Inc.
4. Casanova, J. &. (2015). Growth Hacking - Una guía sobre cómo convertirse en un Hacker Crecimiento. Miami FL: White Owl Publishing.
5. Céspedes, H. (2014). Growth Hacking (Piratería del Crecimiento), Start-Up y Emprendimiento. HC Global Group. Santiago de Chile.
6. Chen, A. (2013). The Viral Startup: A Guide to Designing Viral Loops. Corea: Kindle Edition.
7. Conway, D., & White, J. M. (2012). Machine Learning para Hackers. Sebastapol, Rusia: O'Reilly Media, Inc.
8. Edward, C. (2006). Estudio de Procesos Gaussianos para aprendizaje automático. Universidad de Penssylvania. Philadelphia, Estados Unidos.
9. Enterprise Architecture as Strategy: Creating a Foundation for Business. 2006. Jeanne W Ross, Peter. Weill, David C. Robertson.

10. Frick, A. y Frick, S. (2013). Gestión y Desarrollo de Empresas Innovadoras. Diario de Gestión de la Tecnología y la Innovación (vol.8). Santiago de Chile.
11. Holiday, R. (2014). Growth Hacker Marketing. Gran Bretaña: Profile Books LTD.
12. Leskovec, J., Rajaraman, A., & Ullman, J. (2014). Mining of Massive Datasets. Reino Unido: CPI Group Ltd.
13. Linoff, G., & Berry, M. (2011). Las técnicas de minería de datos : para el comercio , ventas y gestion de relaciones. Indiana, USA: Wiley Publishing, Inc.
14. Lombardero, J. (2015). Tesis Doctoral: Problemas y retos de gestión empresarial en la economía digital: estudio comparado y sistémico de competencias directivas. Universidad Camilo José Cela. Villafranca, Madrid.
15. Myatt, G. 2007. Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining. New Jersey. John Wiley & Sons, Inc.
16. REID, R., FRASER-KING, G. y SCHWADERER, D. (2007) Data Lifecycles. E.U.A.: John Wiley & Sons, Inc.
17. Software systems architecture, working with stakeholders using viewpoints and perspectives, 2012. Nick Rozanski, Eoin Woods.

18. Tan, P.-N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining. Michigan: Pearson.
19. Witten, I., & Frank, E. (2011). Data Mining: Practical Machine Learning Tools and Techniques. Auflage: Morgan Kaufmann Publishers Inc.

ANEXOS

ANEXO 1: Publicaciones y porcentajes semanales

La siguiente tabla muestra los porcentajes de las interacciones, durante un periodo de cinco semanas, estos datos se utilizaron para generar las gráficas que se encuentran en el capítulo 5.

Tabla: **Publicaciones y porcentajes semanales**

Tipo de Publicación	Semana 1	Semana 2	Semana 3	Semana 4	Semana 5
Link	1	0.75	0.6784	0	0
Video	1	1	1	1	0.75
Video	0.8889	0.9167	1	1	1
Video	1	0.9333	0.75	0.6678	0.5
Video	1	0.6667	0.8571	0.5487	0.4533
Video	0.875	1	1	1	1
Video	0.8636	1	1	1	1
Video	1	0.5482	1	0.75	0.4567
Video	0.8667	0	1	1	0.9333
Video	0.9167	0.75	0.9167	0.75	0.6667
Video	0.7692	1	1	1	1
Video	0.8571	0.9333	1	0.9286	0.8571
Video	1	0.6667	1	0.8333	0.8333
Video	1	1	0.6667	0.6667	0.6667
Video	1	0.8333	1	0.9091	0.5455
Video	0.8	0.6667	1	0.8	0.8
Video	1	0.5455	1	1	0.5
Video	0.8929	1	0.9143	0.8143	0.7143
Video	0.804	0.99	0.8638	0.6777	0.5482
Link	0	0	0	0	0
Imagen	1	1	1	1	0.75
Imagen	1	1	1	1	1
Imagen	1	0.8	0.5678	0.7598	0.6678
Imagen	1	0.5	0.6667	0.4565	0
Imagen	1	0.7571	1	1	0.7182

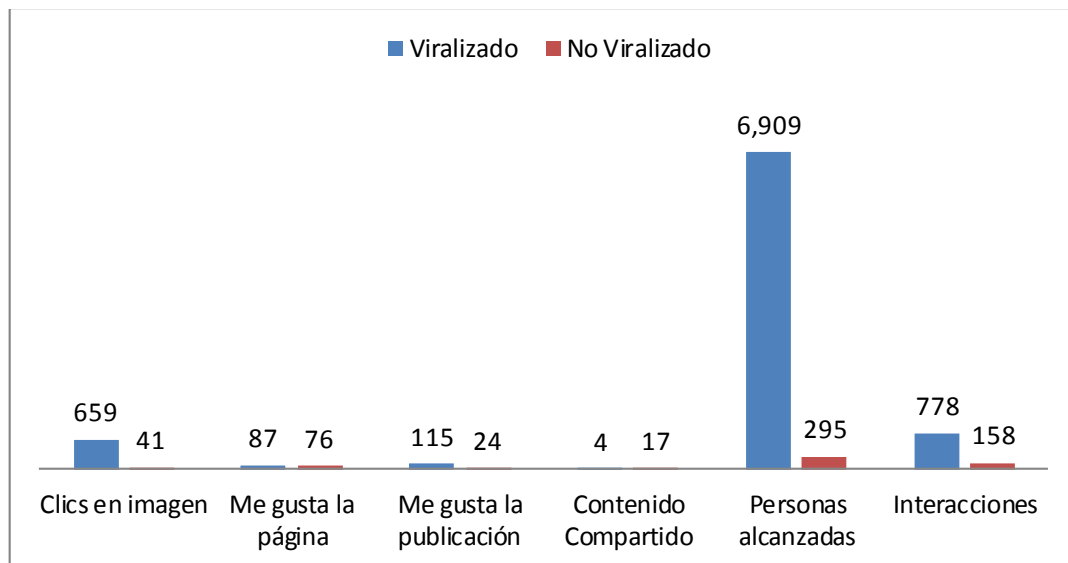
Imagen	1	0.5947	1	1	1
Imagen	1	0.4564	0.8765	0.4564	0
Video	1	0.9333	1	0.9333	0.6323
Video	1	0.6667	0.8333	0.6667	0.6667
Video	1	0.8333	1	1	1
Video	1	0.6667	1	0.9286	0.5
Video	1	0.7273	1	0.8333	0.5647
Link	0	0.6667	0	0.1234	0
Video	1	1	1	0.7273	0.5432
Imagen	1	1	0.8667	0.8	0.8
Imagen	1	1	1	0.5	0.5
Imagen	1	1	0.85	0.7571	0.6714
Link	0	0.307	0.1405	0	0

Fuente: elaboración propia.

ANEXO 2: Funcionalidad promocionar publicación

Al utilizar la función de promocionar publicación de *Facebook* en una *Fanpage* aumenta la viralización de videos, enlaces e imágenes en un tiempo mucho menor que al solo compartir con contactos y amigos agregados en la cuenta para difundirlo.

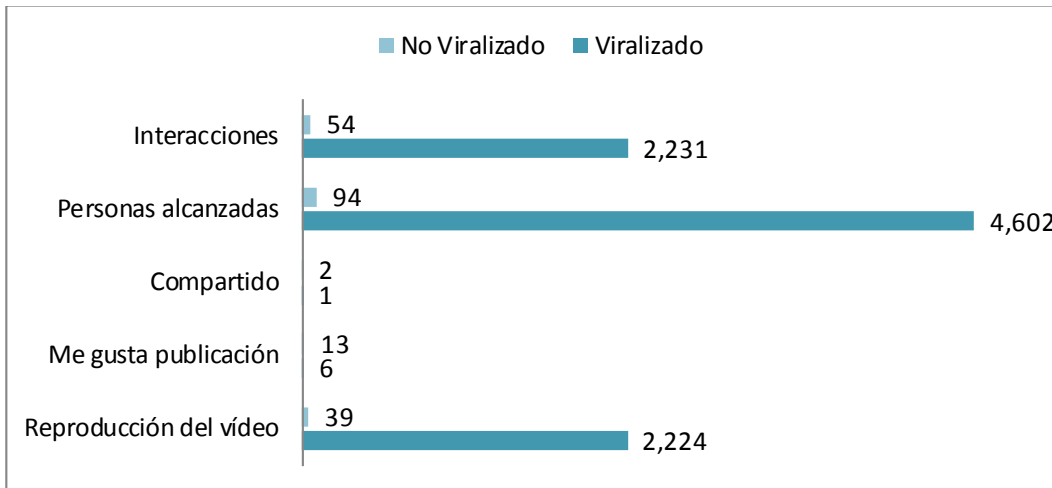
Figura: Barras de funcionalidad “Promocionar publicación” de videos



Fuente: elaboración propia.

Cuando se utiliza la funcionalidad “Promocionar publicación” con los videos puede observarse un comportamiento similar que en las imágenes y las interacciones y personas alcanzadas crecen rápidamente con esta funcionalidad de *Facebook* (Ver Figura de Barras de Interacciones), puede compartir contenidos de cualquier tipo de empresa o producto que se quiera promocionar y obtener un impacto considerable; toma en cuenta edades, tiempo de viralización y la región o país donde desea difundirse.

Figura: **Barras de interacciones con vídeos viralizados**



Fuente: elaboración propia.

Al aplicar la funcionalidad de *Facebook* “Promocionar publicación” a los enlaces es menos eficiente que las imágenes y videos debido a que el alcance mayor fue de 2,526 personas y se recibieron muy pocas interacciones (Ver Figura de Barras de “Promocionar publicación de enlaces”). Realizando una comparativa visual entre las tres graficas puede notarse que publicaciones son las de mayor aceptación y las que alcanzaron una viralización más eficiente. Además puede observarse claramente el alcance que posee la funcionalidad de *Facebook* comparado con el método manual de compartir con los contactos de una determinada cuenta.

Figura: Barras de “Promocionar publicación” de enlaces



Fuente: elaboración propia.

ANEXO 3: Descripción de R Studio

La estadística se puede hacer con muchos softwares estadísticos, incluso con hojas de cálculo, pero para el procesamiento de una gran cantidad de datos y además para aplicar Data Mining el más recomendado es R Studio.

Características técnicas de R Studio

- Es un software libre, se distribuye bajo licencia GNU, lo cual significa que se puede utilizar y mejorar.
- Es multiplataforma, hay versiones para Linux, Windows, Mac, etc.
- Se puede analizar en R cualquier tipo de datos.
- Es una herramienta muy potente.
- Su capacidad gráfica difícilmente es superada por otro paquete estadístico.
- Es compatible con la mayoría de los formatos de archivos de datos (.csv, .xls, .sav, .sa, etc.)
- Existe una gran gama de técnicas estadísticas implementadas y por ser libre cada vez existen más funcionalidades.

Librerías utilizadas

```
> library(xlsx) ##Esta librería es específica para leer y manipular datos de Excel
> library(car) ##La librería car se utiliza para dibujar matrices con diferentes datos
> scatterplot.matrix(a1) ##Dibuja la matriz con los datos incluidos en la variable a1
> a2= a1[,c(2,3,4,5,6)] ##Con esta instrucción pueden quitarse columnas de una tabla
> scatterplot.matrix(a2) ##Dibuja la nueva matriz sin la columna 1
> View(a2) ##Muestra la matriz a2
```

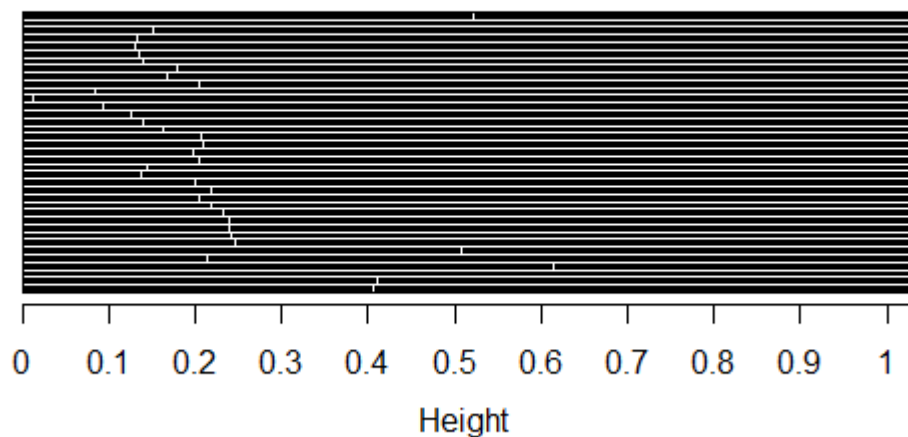
```

> library(cluster) ##Esta libreria es específica para aplicar Clustering en R
> pc1 = princomp(a2, cor = T) ##Esta función ordena los datos con correlación T y los
introduce en una matriz
> a.cp = pc1$scores[ , 1:5] ##Esta función se utiliza dibujar la matriz correlacionada en una
diferente escala
> div.grupo.ag1 = cutree(ag1,3) ## Divide en grupos los datos aglomerados incluidos en
las distintas variables
> div.grupo.ag2 = cutree(ag2,3)
> plot(a.cp, col = div.grupo.ag1)
> plot(a.cp, col = div.grupo.ag2) ##Dibuja la matriz en diferentes grupos y colores
contenidos en la variable a.cp
> div.grupo.ag3 = cutree(ag3,3)
> div.grupo.ag3 = cutree(ag3,4)
> plot(a.cp, col = div.grupo.ag3)

```

La siguiente gráfica de banner muestra la combinación de *clusters* con el método single, aunque es recomendable separar los datos por grupos y dibujarlos como en la figura 18. Para entender mejor los resultados y observar patrones que puedan colaborar con el análisis.

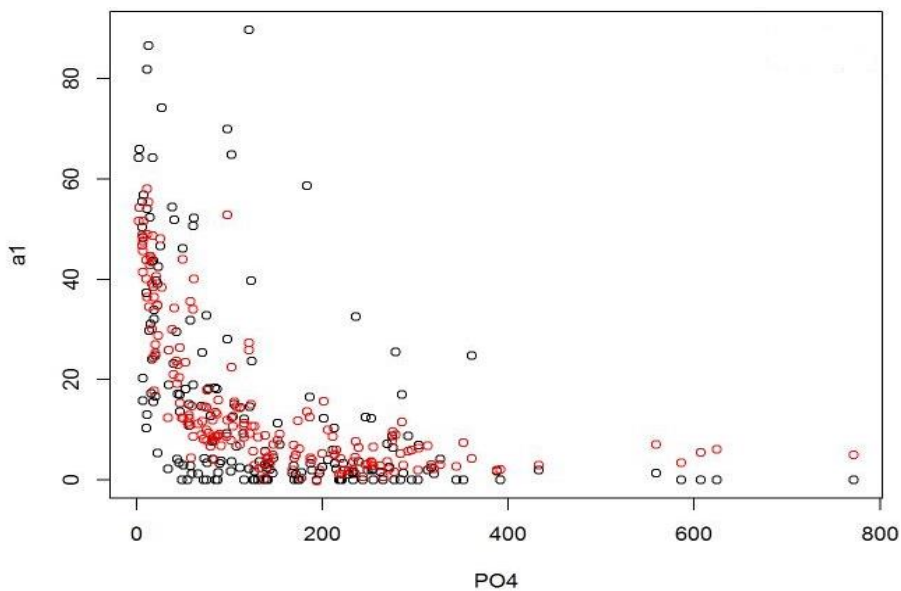
Banner of agnes(x = a2, method = "single")



Agglomerative Coefficient = 0.81

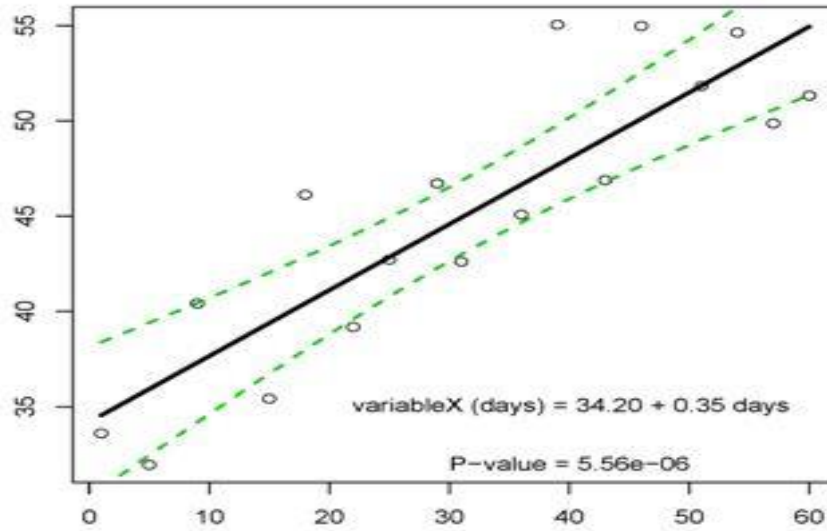
Fuente: elaboración propia, con base a R.

En la siguiente gráfica pueden observarse los usuarios que interactuaron en determinadas publicaciones, pero no proporciona una idea clara de que decisiones tomar en futuras publicaciones en la Fanpage.



Fuente: elaboración propia, con base a R.

En la siguiente gráfica de regresión lineal se compararon publicaciones de videos e imágenes, pero tampoco brinda al usuario ideas claras de desiciones a tomar.



Fuente: elaboración propia, con base a R.

ANEXO 4: Reporte de interpretación de métricas de Growth Hacking en Fanpage de Facebook

a) Top de publicaciones medidas

1) Publicación ID: 35972561

Tipo de publicación: video

Público objetivo: hombres y mujeres

Edades: 13 a 30 años

6.593 Personas alcanzadas		
2.741 Reproducciones de video		
31 Reacciones, comentarios y veces que se compartió		
19  Me gusta	14 En publicación	5 En contenido compartido
2  Me encanta	1 En publicación	1 En contenido compartido
11  Me divierte	11 En publicación	0 En contenido compartido
0 Comentarios	0 En publicación	0 En contenido compartido
0 Veces que se compartió	0 En publicación	0 En contenido compartido
327 Clics en publicaciones		
290 Clics para reproducir 	24 Clics en el enlace	13 Otros clics 

2) Publicación ID: 35972568

Tipo de publicación: video

Público objetivo: hombres y mujeres

Edades: 13 a 45 años

4.560 Personas alcanzadas		
2.269 Reproducciones de video		
8 Reacciones, comentarios y veces que se compartió		
5  Me gusta	4 En publicación	1 En contenido compartido
2  Me divierte	2 En publicación	0 En contenido compartido
0 Comentarios	0 En publicación	0 En contenido compartido
1 Veces que se compartió	0 En publicación	1 En contenido compartido
147 Clics en publicaciones		
141 Clics para reproducir 	0 Clics en el enlace	6 Otros clics 

3) Publicación ID: 35972576

Tipo de publicación: video

Público objetivo: hombres y mujeres

Edades: 13 a 45 años

Rendimiento de tu publicación		
3.745 Personas alcanzadas		
616 Reproducciones de video		
3 Me gusta, comentarios y veces que se compartió		
1 Me gusta	1 En publicación	0 En contenido compartido
2 Comentarios	2 En publicación	0 En contenido compartido
0 Veces que se compartió	0 En publicación	0 En contenido compartido
11 Clics en publicaciones		
7 Clics para reproducir <i>i</i>	0 Clics en el enlace	4 Otros clics <i>i</i>
COMENTARIOS NEGATIVOS		
0 Ocultar publicación	0 Ocultar todas las publicaciones	
0 Reportar como spam	0 Ya no me gusta esta página	

4) Publicación ID: 35972571

Tipo de publicación: gift

Público objetivo: hombres y mujeres

Edades: 13 a 45 años

2.598 Personas alcanzadas		
23 Me gusta, comentarios y veces que se compartió		
22 Me gusta	22 En publicación	0 En contenido compartido
0 Comentarios	0 En publicación	0 En contenido compartido
1 Veces que se compartió	0 En publicación	1 En contenido compartido
25 Clics en publicaciones		
0 Visualizaciones de fotos	16 Clics en el enlace	9 Otros clics <i>i</i>
COMENTARIOS NEGATIVOS		
0 Ocultar publicación	0 Ocultar todas las publicaciones	
0 Reportar como spam	0 Ya no me gusta esta página	

5) Publicación ID: 35972550

Tipo de publicación: gift

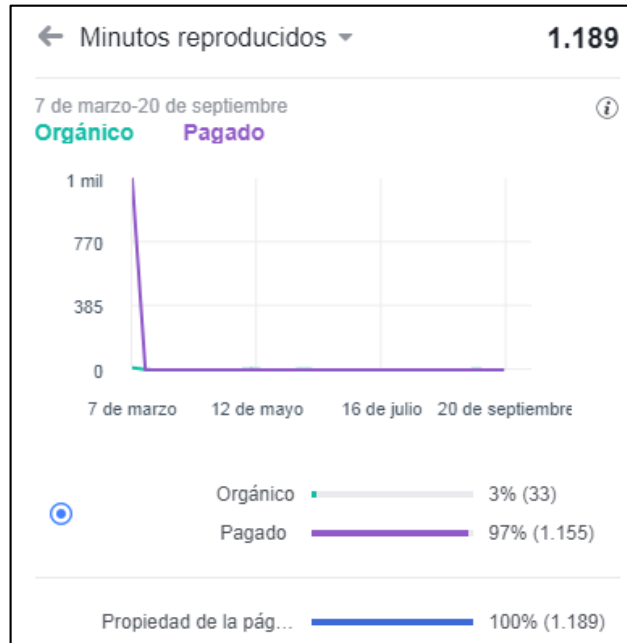
Público objetivo: hombres y mujeres

Edades: 13 a 45 años

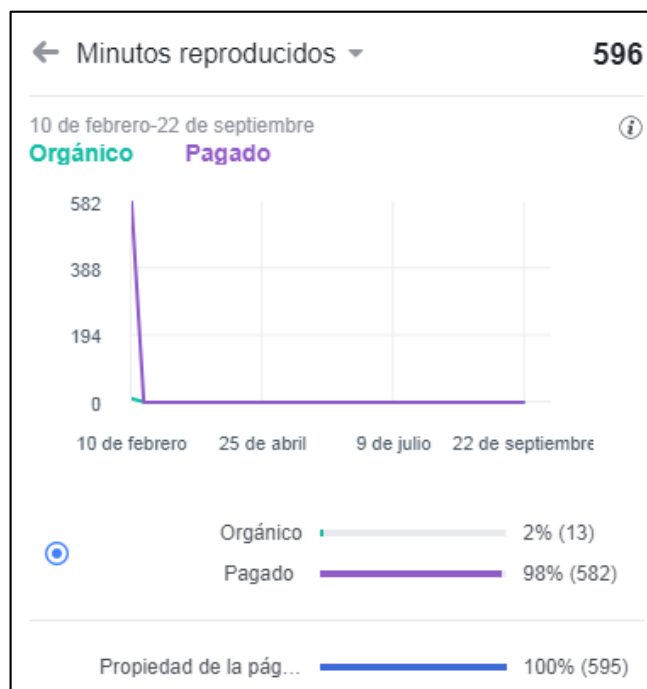
643 Personas alcanzadas		
70 Reacciones, comentarios y veces que se compartió		
68  Me gusta	68 En publicación	0 En contenido compartido
1  Me encanta	1 En publicación	0 En contenido compartido
1 Comentarios	1 En publicación	0 En contenido compartido
0 Veces que se compartió	0 En publicación	0 En contenido compartido
7 Clics en publicaciones		
0 Visualizaciones de fotos	2 Clics en el enlace	5 Otros clics 

b) Rendimiento de Publicaciones

1) Publicación ID: 35972561



2) Publicación ID: 35972568



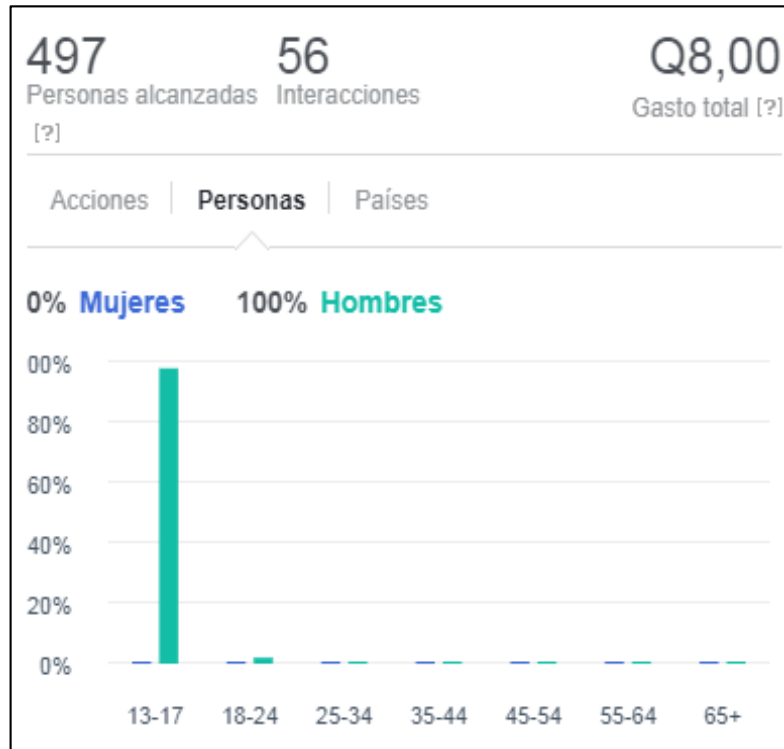
3) Publicación ID: 35972576



4) Publicación ID: 35972571

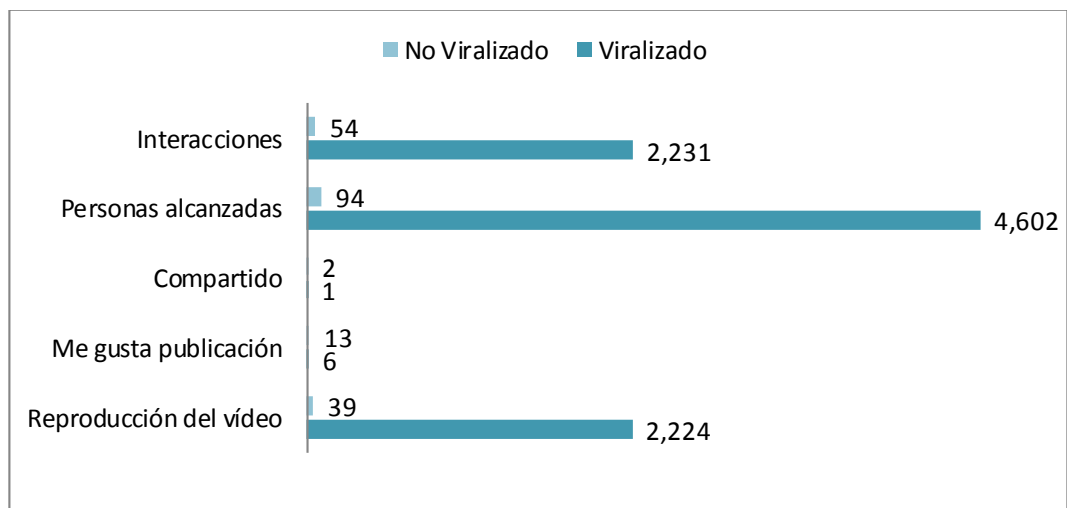


5) Publicación ID: 35972550

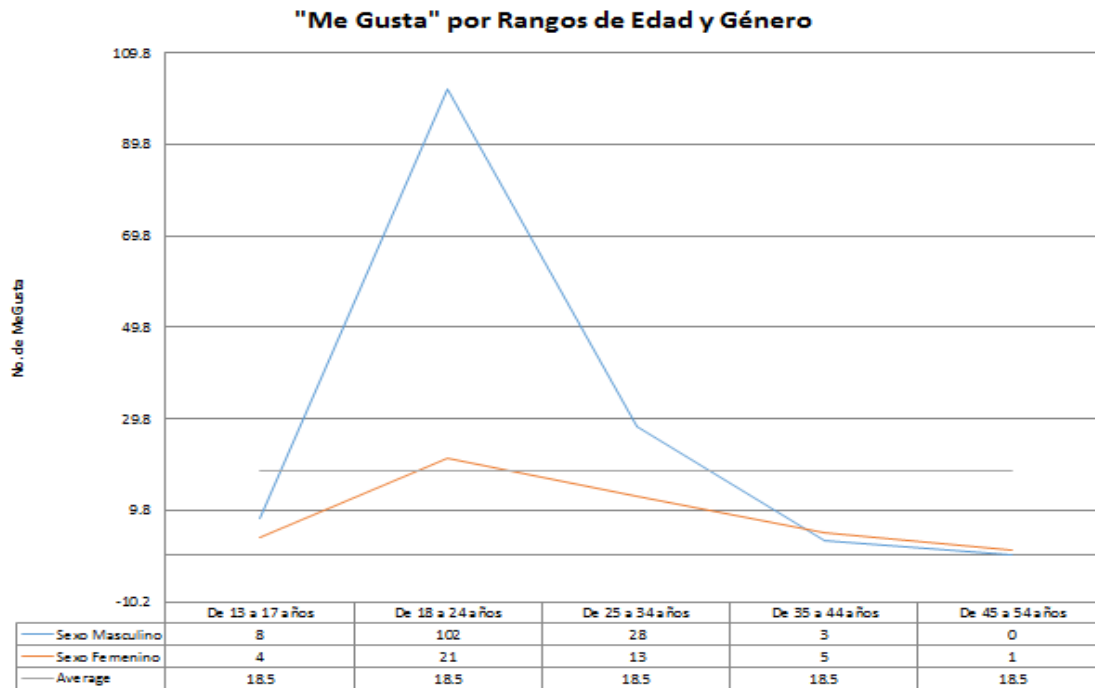


c) Estadísticas de publicaciones

Barras de Interacciones con vídeos viralizados



Puede observarse claramente el alcance que posee la funcionalidad de Facebook comparado con el método manual de compartir con los contactos de una determinada cuenta.



Fuente: elaboración propia, con base a QI Macros para Excel.

Se presenta la sumatoria del alcance mensual por rangos de edad y género, estos resultados se logran al aplicar la viralización automática que ofrece *Facebook*.

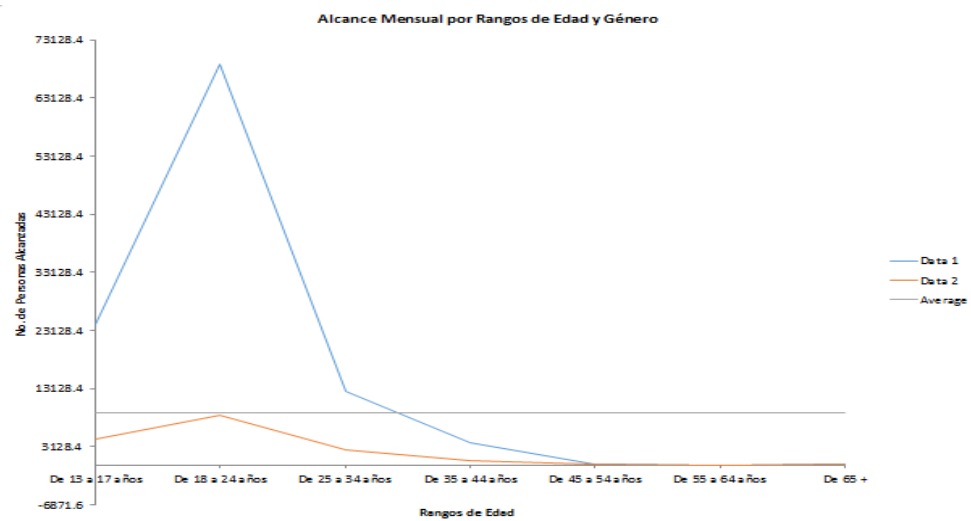
Sumatoria de personas alcanzadas mensual promedio

Rangos de edad	Masculino	Femenino	Totales
De 13 a 17 años	24100	4385	28485
De 18 a 24 años	68925	8513	77438
De 25 a 34 años	12727	2501	15228
De 35 a 44 años	3859	750	4609
De 45 a 54 años	224	106	330
De 55 a 64 años	20	19	39
De 65 +	63	42	105
Total	109918	16316	126234

Fuente: elaboración propia.

La mayor incidencia en *Facebook* en Guatemala es en el género masculino en el rango de edad de 18 a 24 años. Este dato podría ayudarnos a difundir diferentes contenidos de forma más rápida y efectiva utilizando las técnicas y métricas de *Growth Hacking*.

Polígono de frecuencias de personas alcanzadas



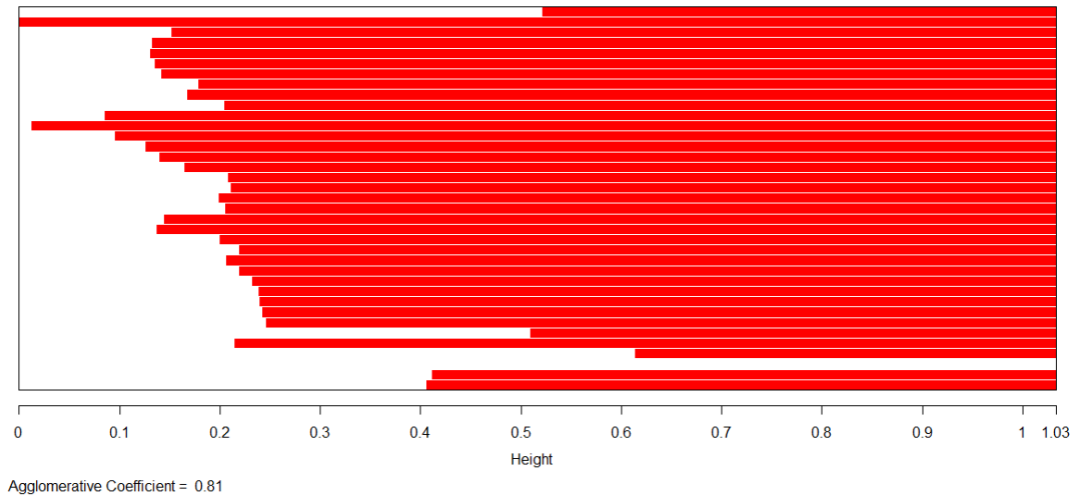
Fuente: elaboración propia, con base a QI Macros para Excel.

Clustering

En la siguiente gráfica se aplicó el método simple de *Clustering* y puede observarse claramente que los *cluster* se dividen en 2 grupos principales y esto demuestra la falta de efectividad para viralizarse de los enlaces, por el contrario los videos e imágenes poseen valores más altos por su buena aceptación por los usuarios y fans de la página de *Facebook*.

Banner del método simple

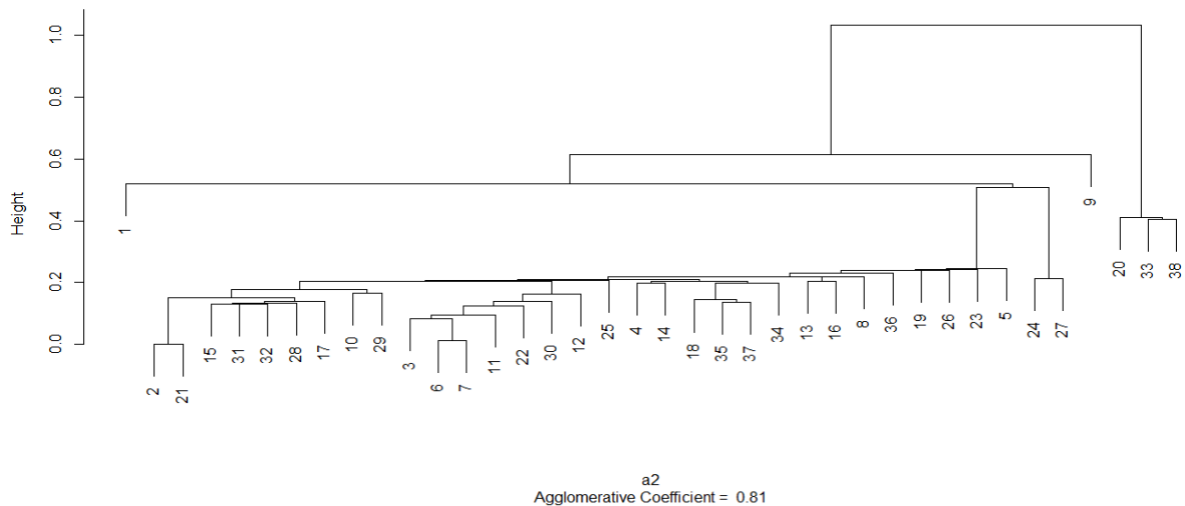
Banner of agnes(x = a2, method = "single")



En el dendograma del método simple se observan tres grandes grupos y se debe a la mayor cantidad de videos publicados y que obtuvieron mayor interacción durante un período de cinco semanas.

Dendograma del método simple

Dendrogram of agnes(x = a2, method = "single")

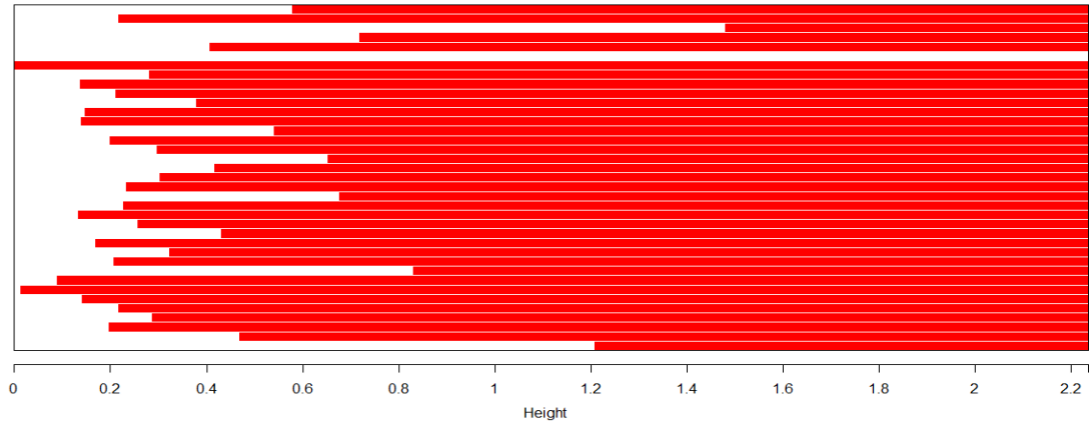


Se utilizó el método completo y se notan claras diferencias en los *clusters* además de la variación del coeficiente aglomerativo. El dendograma es la representación del *Clustering* efectuado, en ambos se aplicó el método

completo y al igual que en el *Banner* puede notarse los mismos dos grupos muy marcados a diferencia de los otros métodos.

Banner del método completo

Banner of agnes(x = a2, method = "complete")

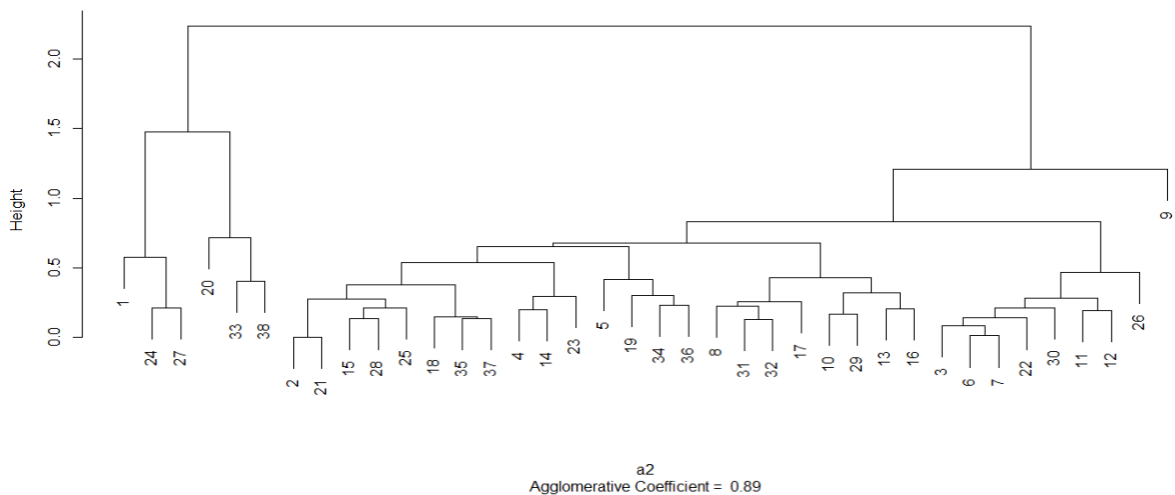


Agglomerative Coefficient = 0.89

Fuente: elaboración propia, con base a R Studio.

Dendograma del método completo

Dendrogram of agnes(x = a2, method = "complete")



a2
Agglomerative Coefficient = 0.89

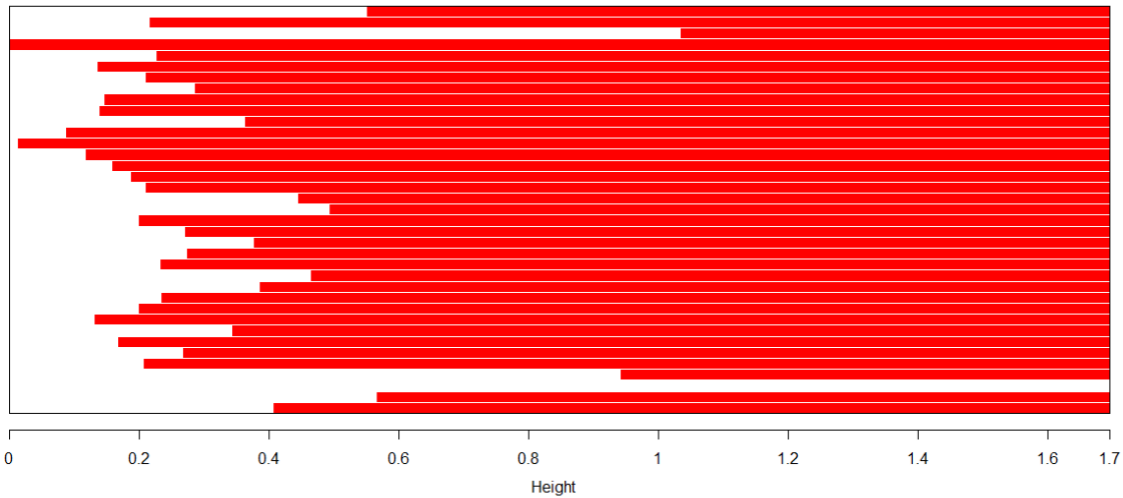
Fuente: elaboración propia, con base a R Studio.

Por último, se experimentó con el método de *Clustering* promedio y su comportamiento es más parecido al método simple que al completo y siguen

marcándose los mismos grupos. En la gráfica pueden observarse las similitudes con su respectivo dendograma, toma en cuenta que son las visualizaciones promedio de cinco semanas, se notan los distintos comportamientos.

Banner del método promedio

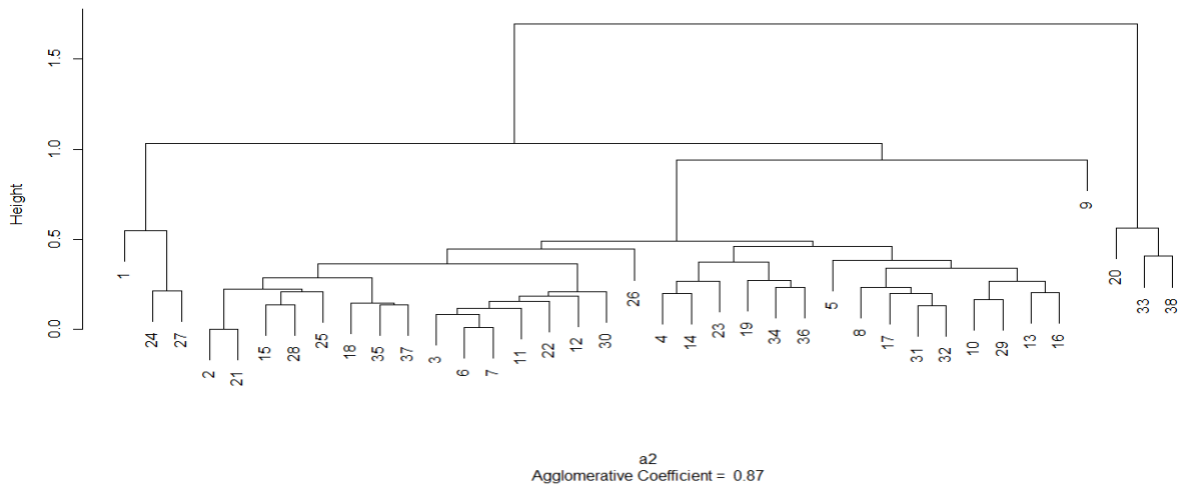
Banner of agnes(x = a2, method = "average")



Agglomerative Coefficient = 0.87

Dendograma del método promedio

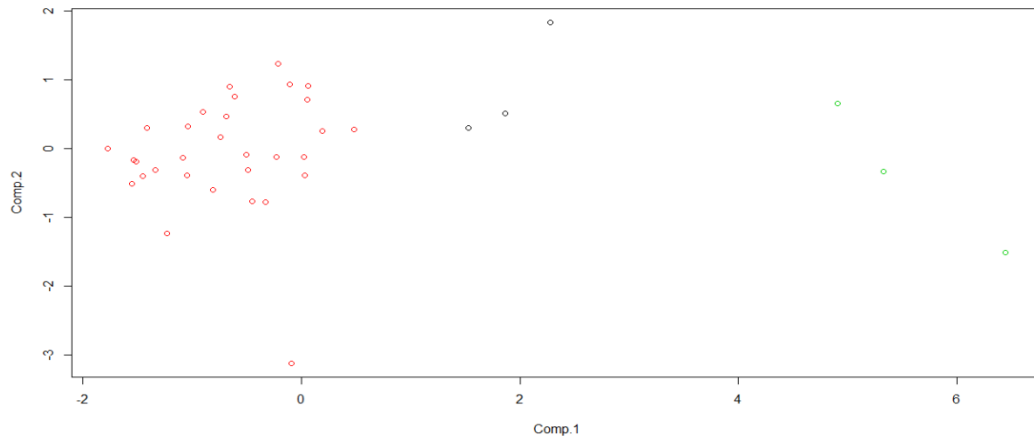
Dendrogram of agnes(x = a2, method = "average")



Para observar de mejor manera los resultados de *Clustering* se aplicó división de grupos de acuerdo al tipo de aglomerado de datos, observándose

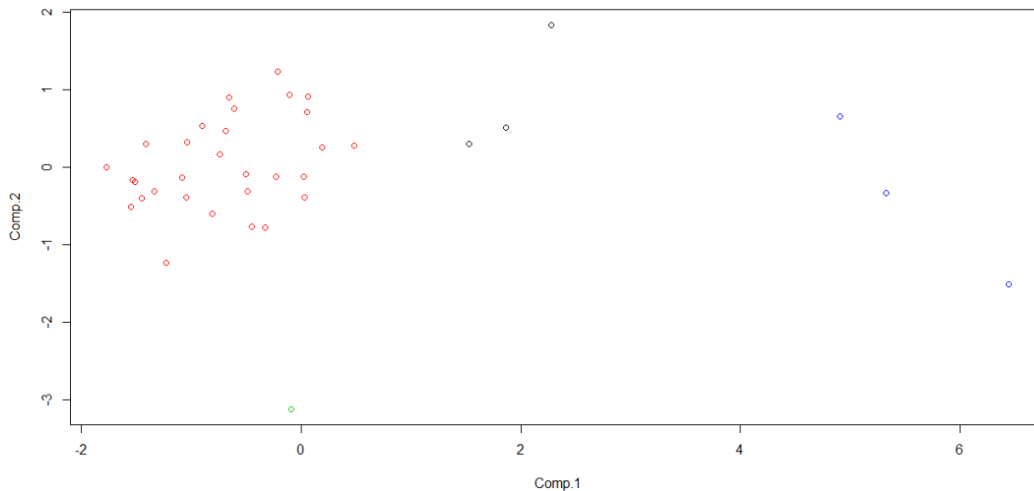
que son tres grupos y un elemento aislado los que se repiten tanto en el método completo como en el simple.

Ilustración de Clusters y Outliers del método completo



Fuente: elaboración propia, con base a R Studio.

Ilustración de Clusters y Outliers del método promedio



Fuente: elaboración propia, con base a R Studio.

Como puntos importantes en el análisis de *Clustering* anterior pueden observarse los cluster muy marcados y demuestra los puntos aislados, en este caso es uno, provocando una mayor viralidad los videos, que las imágenes y los enlaces publicados.

d) Puntos estratégicos para la toma de decisiones:

El rendimiento en la interpretación de métricas de *Growth Hacking* demostró ser eficiente en varios aspectos: el correcto procesamiento de los datos, los cambios realizados en publicaciones, el poco tiempo en el que se difundieron las publicaciones y el alcance en número de personas que visualizaron e interactuaron con las distintas publicaciones. Las mejoras que pueden implementarse son aumentar el contenido visual en las páginas (videos, imágenes, memes, noticias, etc.) en vez de texto, hashtags y enlaces a otros sitios, debido a que la mayoría de personas con las que se realizó la experimentación, tuvo preferencia por todos los contenidos visuales y entretenidos.